BEAR STEARNS

Equity Research
Technology

Internet Infrastructure & Services

Introducing Internet 3.0
Distribution and Decentralization Change Everything

- **THE RIPPLES CHANGE THEIR SIZE.** This report examines the impact of distribution (of assets and resources across geographies) and decentralization (of a network's assets and resources outward toward "edge" devices) on the Internet. It highlights the companies that will benefit or be hurt by this trend. Finally, it profiles emerging companies that are taking advantage of these trends and pushing the Internet toward its next phase, which we call Internet 3.0.

- **WHAT HAPPENS IN INTERNET 3.0?** Thick client devices dominate the network. Edge devices communicate directly with each other. Increased traffic among devices places strain on networks, forcing network carriers to increase capacity. All of this is driven and guided by the laws of network dynamics.

- **EXTENDING TRADITIONAL MARKETS, INTRODUCING NEW ONES.** We have identified four emerging areas in the field of distribution and decentralization: distributed processing, distributed storage services, distributed network services, and decentralized collaboration.

- **INTEL, MICROSOFT, SUN: INCUMBENTS TAKE CHARGE.** Intel, Microsoft, and Sun Microsystems have emerged as thought leaders, each waving a flag and jockeying for position in Internet 3.0. We also profile a number of private companies that are providing thought leadership on the next phase of the Internet. We highlight their impact on incumbents and existing markets, and quantify new market opportunities.

Chris Kwak
(212) 272-7792
ckwak@bear.com

Robert Fagin
(212) 272-4321
rfagin@bear.com

www.bearstearns.com

MAY 2001

Table of Contents — Page

Private Company Profiles **Page**

Exhibits	Page

Exhibits	Page

Exhibits	Page

Executive Summary

"Don't let the rush obscure the gold." – Gene Kan, Founder of Infrasearch[1]

This report examines the impact of distribution (of assets and resources across geographies) and decentralization (of a network's assets and resources outward toward "edge" devices) on the Internet. It analyzes the evolution of technology and uncovers patterns that are fundamental to network systems development. It examines the role that distribution and decentralization have played in the development of the Internet and the World Wide Web. It highlights the companies that have benefited and will benefit from them and identifies which public companies may be hurt by them. Finally, it profiles emerging companies that are taking advantage of these trends and pushing the Internet toward its next phase, which we call Internet 3.0. We believe that in five years, every company will either embrace distribution and decentralization, or face significant disadvantages.

This report is not for everyone. It is intended for those with a stake in understanding how the evolving Internet will change computing, networking, and communications. It focuses on what we believe is happening to the Internet, the potential impact of this evolution on computer, network, and communications systems, and the effects of these changes on the systems that rely on the Internet. We believe Internet 3.0 will be extremely disruptive.

It is a story with a long history and lots of parts that move in fits and starts, weaving cycles with beginnings and endings. We believe we are exiting one such cycle (Internet 2.0, the Web era) and entering a new one (Internet 3.0, the two-way Web). Understanding the rules, who's playing, and where we are offers us a panoramic view of what may be lost and gained.

INTRODUCING INTERNET 3.0: THE RIPPLES CHANGE THEIR SIZE

What in retrospect may be called Internet 1.0 began in the 1960s with a group of academics and engineers. TCP/IP (Transmission Control Protocol/Internet Protocol), the two fundamental network and transport protocols on the Internet, formalized the Internet; the birth of the modern network was the result. Internet 2.0, the Web era, commenced in the early 1990s. The browser and the modern router pushed Internet 2.0 to its current state; the result was the proliferation of devices connecting to the network and the growth of content to drive the use of these devices. Internet 3.0 has just begun. XML, the relatively new Extensible Markup Language specification, is grounding it, and the development of new software is guiding it. While Internet 3.0 will feature many of the same participants as Internet 2.0, it will, we believe, also highlight several emerging innovations and technologies.

Internet 1.0 was about the network. Internet 2.0 has been about devices "living" on top of that network. Internet 3.0 will be about the software weaving together the devices living on top of the network, and how users with these devices will be able to communicate seamlessly with each other and use systems resources more effectively across geographies and operating platforms. We are captivated by the idea that the

[1] Infrasearch – a peer-to-peer search technology provider – was recently acquired by Sun Microsystems to become part of Bill Joy's Project JXTA, a platform for distributed computing.

intelligence of the network will continue to migrate to “edge” devices, (i.e., devices as close as possible to an end-user rather than a central resource). Networks will continue to expand and much of their functionality will be distributed and decentralized. We believe this is already happening in part with initiatives like Microsoft .NET. While the architecture of the current Internet and Web will continue to be relevant, we believe Internet 3.0 will involve significant architectural changes.

KEY ASSUMPTIONS OF INTERNET 3.0

- **The Edge: The Undiscovered Country.** In Internet 2.0, the Web era, everything has collapsed into the network center. Servers dominate the network with content, processing, and storage; PCs are largely presentation-level devices whose participation is consumption. In Internet 3.0, the center begins to unravel and diffuse. The edge becomes the Internet, and devices do more with what have to date been dormant native resources. PCs become dominant, and all clients are thick. The domain name system is no longer the only addressing system utilized by networks. The browser is no longer the only gateway to the Internet.

- **Faster, Stronger, More.** We are ruled by laws that produce faster, stronger, and more sophisticated technologies. While innovation drives these technologies, we are also guided by prevailing laws that are seemingly deterministic in their predictability and accuracy. The interplay and consequence of these laws suggest fantastic developments to come.

- **XML: Objectifying Everything.** Since 1990, the language of the Web has been HTML, the HyperText Markup Language. We are entering a new era where XML is quickly becoming the lingua franca not just for the Web, but the entire Internet, across all platforms and applications. XML enables anything on the Internet to be treated as an object. This has important ramifications for resource utilization and end-to-end communications.

- **Virtualization.** As a corollary, much of what is happening is predicated on the virtualization of systems resources. We define virtualization as the method by which computing and network systems resources, which may be dispersed and fragmented, are harnessed and utilized in aggregate, to create a virtual pool of resources, and managed seamlessly. While this kind of centrally managed system that creates a virtual cluster is not new (mainframes have done this for years), we believe the scope and reach of the next generation of virtualization is even broader. StorageNetworks, for example, has been virtualizing storage, as have private storage services providers; Zambeel, a privately-held provider of distributed storage systems, is focused on the distributed storage opportunity. LoudCloud has been attempting to virtualize the datacenter, as have private companies including Mimecom and Logictier. Privately held Ejasent and Parabon have recognized the value in virtualizing processing power. There is no shortage of interest from public companies in virtualizing assets, and certainly no shortage of emerging private companies continuing to challenge public companies. Most of these companies use meshed network topologies and XML to overcome geographic limitations.

- **Public Companies Who Understood This.** More than a handful of companies recognized the trend toward distributed and decentralized architectures. Early on, companies like DoubleClick, Exodus, and Inktomi took advantage of the need to distribute systems resources. More recently, Akamai and others have begun to capitalize on distribution and decentralization. We believe even newer entrants (e.g., StorageNetworks) are taking advantage of distributed architectures to better utilize systems resources.

- **A Stake in the Ground.** Microsoft, Intel, and Sun Microsystems have each claimed a plot of land on the distribution and decentralization frontier — Microsoft with .NET, Intel with the Peer-to-Peer Working Group, and Sun Microsystems with Project Juxtapose (JXTA, pronounced jux•ta), the third leg of its Java-Jini-JXTA triumvirate which we describe below. Each has recognized and begun to articulate the significance of distributed and decentralized resources to the next generation of the Internet. Each of these companies has dedicated major resources to executing its vision. Microsoft is aiming to become the dominant platform for the next generation of Web services; Intel is attempting to create a set of standards for peer-to-peer computing; and Bill Joy of Sun is drafting JXTA, a platform for distributed computing.

THE THREE MARKET PHASES OF INTERNET 3.0

We believe Internet 3.0 will be marked by three distinct phases:

Phase I: Embrace

In the first phase of the Internet 3.0 cycle, public companies embrace distributed and decentralized architectures. Already, Microsoft has laid out a visionary approach to distributed Web services by introducing .NET and more recently, HailStorm — its attempt to connect all users on the Internet with rich XML-based feature sets. McAfee ASaP (formerly myCIO), a subsidiary of Network Associates, has been offering Rumor, a peer-to-peer virus signature file-updating service, since October 2000. Exodus has already partnered with privately held distributed processing developer United Devices to offer its customers enhanced performance measurement and load-testing services.

- **Investments in the Class of Internet 3.0.** We believe the entrants to watch for (notwithstanding Microsoft, Sun Microsystems, and Intel) are EMC, PC OEMs (Dell and IBM in particular — Dell because of its dominant position and IBM because of its patents on distributed processing), enterprise software vendors (e.g., Oracle, Veritas, SAP, Interwoven), device manufacturers (e.g., Palm, Handspring, and RIMM), companies with complex database applications like eBay, Amazon, and Yahoo!, xSPs (e.g., Exodus, Digex, StorageNetworks, and Akamai), and even ISPs like AOL Time Warner and EarthLink.

- **More Partnership Announcements.** We believe Scale Eight's partnership with Akamai, United Devices' partnership with Exodus, and Groove's inclusion in Microsoft HailStorm are the tip of the iceberg.

- **Client Devices Continue to Become Thicker.** The laws of network dynamics indicate that the client devices will be thicker in the future because the cost of decentralizing (i.e., localizing) content, processing, and data on the client device will be cheaper than centralizing assets and resources. Client devices will be able to do more as a consequence.

- **Look for Messaging to Lead, Applications to Follow.** Keep an eye out for instant messaging to lead the adoption of Internet 3.0 within enterprises. In particular, look for Yahoo!, Microsoft and AOL Time Warner to lead the charge, positioning themselves with Jabber and other open source XML-based messaging platforms. Applications, like collaboration and value-added network services will follow the adoption of instant messaging platforms.

Phase II: Compete

We believe the second phase of the Internet 3.0 cycle will emerge when private companies threaten to take customers away from strong public incumbents. We believe private companies like Scale Eight, for example, could potentially take advantage of distribution and decentralization and steal customers away from incumbent storage vendors. Similarly, Entropia, a distributed processing concern, has already partnered with SolidSpeed, a content acceleration service provider, to offer performance measurement and load-testing services across SolidSpeed's customer base.

- **Attack Large Markets First.** The Class of Internet 3.0 are expected to attack the largest markets first: large enterprises and service providers. The tell-tale sign that the competitive landscape is in full bloom is when one in the Class of Internet 3.0 wins a large contract at an incumbent's expense. Such an event, in each sector, will mark the beginning of Phase II.

- **Expect Incumbents to Downplay Significance of Internet 3.0 Businesses.** Recognizing the potential impact of a new business, and unable to maneuver quickly enough and reposition business models, incumbents will begin to downplay the significance of Internet 3.0 businesses. This is precisely what happened with mainframe and minicomputer manufacturers when the PC emerged; this is what off-line retailers said when the Web was launched; this is what many carriers had been saying about Voice-over-IP; and this is what many have been saying about outsourced storage.

- **Handful of Companies Come Public.** On the heel of large design wins, a handful of Internet 3.0 members will begin to IPO. While still early, we believe the most likely candidates are Groove Networks, Scale Eight, and United Devices.

Phase III: Win

The final phase of the Internet 3.0 cycle should materialize when a few of these private companies generate enough interest and critical mass to unseat incumbents. Moreover, in Phase 3, private companies will successfully carve out entirely new markets.

- **Incumbents Fall.** As EMC is to IBM, Dell is to Compaq, and Microsoft is to Novell, Internet 3.0 is to incumbents. In every cycle, incumbents fall to newcomers. Internet 3.0 will be no different. Some incumbents will lose because their models are incompatible with the distributed and decentralized direction of computing and networking. Others will fail because they failed to recognize the impact of distribution and decentralization on computing and networking.

- **Acquisition — Last Ditch Effort.** Unable to recover, and too late to partner, some incumbents will try to acquire competitors to regain edge.

We believe we are in the midst of Phase 1. Large public companies are slowly recognizing the trend toward distribution and decentralization of networks and systems resources. It is our belief that framing the landscape and recognizing the early participants in this disruptive phase will prove invaluable in understanding the impact these emerging technologies will have on the Internet.

We believe companies that embrace distribution and decentralization will have a significant advantage going forward due to what we believe are the inherent benefits of distributing and decentralizing data assets and system resources. Overcoming the technical obstacles in the migration to distributed and decentralized architectures will require engineering sophistication. We believe only a handful of companies will have the required technical competence.

FORGING NEW MARKETS

Processing, storage, and bandwidth are the fundamental resources of computing systems. Applications ride on top of these resources. In Internet 3.0, each of these resources is optimized and maximized.

We have identified four emerging areas in the field of distribution and decentralization: distributed processing, distributed storage services, distributed network services, and decentralized collaboration.

- **Distributed Processing.** Distributed processing is defined as the technique by which large processing tasks are segmented into component blocks, delivered to a number of machines where these blocks are processed, and reassembled by a central machine. The fundamental idea behind distributed processing is that more machines can process more data faster. Distributed processing technologies allow several processing devices (such as servers or PCs) to be linked together to share processing resources. Such technologies make particularly efficient use of processor "downtime" (i.e., resources not in active use). Within a particular organization for instance, the likelihood is good that not all desktop computers are concurrently being utilized to their fullest capacity. Distributed processing harnesses these relatively cheap and underused resources to perform otherwise costly tasks. The distributed processing market is currently dominated by a handful of private companies, including United Devices, Entropia, DataSynapse, and Parabon. While there are limitations to what a distributed processing network can do, we believe the opportunities available to these companies are sizable enough to warrant some attention. We estimate the addressable market opportunity for distributed processing software and services will be close to $8 billion in 2004. Because the market for distributed processing software and services is new, we believe the adoption of the technology may take some time. However, in the bio-informatics industry (which requires massive processing power) and in load-testing, there are ample near-term opportunities.

- **Distributed Storage Services.** Distributed storage is defined as a network topology in which storage assets are componentized and distributed across geographies in order to create a clustered, wide-area storage network. Freeing data from geographic isolation and expensive equipment and maintenance costs is a significant value proposition. While storage equipment costs can be astronomical, the most significant portion of a storage infrastructure's costs is software and maintenance/services; we believe the "80/20" rule applies to storage as well — the cost of software and maintenance of storage is about 4x the cost of the actual hardware assets. We believe the overwhelming need to manage data more efficiently and cheaply is giving birth to what we refer to as distributed storage companies. These companies are attempting to allow corporations and service providers to emancipate data from geographically isolated storage environments and enable customers to leverage existing geographically distributed resources. Because distributed storage is a completely new industry, we find measuring the market opportunity extremely difficult. Nevertheless, we regard the distributed storage market as at least the size of the outsourced storage services opportunity, since, in fact, distributed storage companies like Zambeel will likely initially target the Storage Service Provider (SSP) market. As such, we believe distributed storage services could be a $8.8 billion addressable opportunity in 2004.

- **Distributed Network Services.** Distributed network services are defined as intelligent network services that accelerate the delivery of data, propagate data more efficiently, or minimize the cost of operating existing network services using distributed resources across geographies. Companies like Inktomi and Akamai have been offering technologies and services for content distribution networks for the past few years. While the initial market focus of these companies has been on the wide area network (WAN), in Internet 3.0, we believe distributed network services will find traction across the WAN, MAN (metropolitan area network), and LAN (local area network). As such, we believe the distributed network services market is a larger opportunity than that of today's content delivery network market. We estimate the addressable market for distributed network services will be at least $7.4 billion in 2004.

- **Decentralized Collaboration.** Decentralized collaboration is defined as a peer-to-peer system for real-time communications that facilitates the creation of workgroups in which the user's device and its assets become central. Decentralized collaboration generally includes such features as instant messaging, file sharing/transfer, co-browsing, and personal information sharing. We anticipate decentralized collaboration efforts will continue to evolve to facilitate transactions and enable tight integration into ERP (enterprise resource planning), and supply and demand chains. Decentralized collaboration is a brand new market brought to the fore by companies like Groove Networks, the brainchild of Lotus Notes creator Ray Ozzie. The premise of decentralized collaboration is that real-time, fluid communications among individuals drives the velocity of transactions and increases productivity. Being able to communicate with partners in a supply chain or demand chain in real-time without the hassle of a browser-based system can increase the level of "touch" between parties in a network. Collaboration in a decentralized, peer-driven model permits a network of individuals to communicate in a manner akin to everyday life. By integrating into and developing applications on top of existing platforms, decentralized collaboration software vendors have the opportunity to tap into the installed base of tens of thousands of corporations. We believe corporations with a geographically dispersed workforce or ones with a large number of workgroups will be the first to adopt these technologies. For example, Groove Networks recently announced GlaxoSmithKline, Raytheon, and the U.S. Department of Defense as customers. The market for decentralized collaboration is new and untested, and we believe conservative assumptions are more prudent at the outset. We estimate that decentralized collaboration software represents a $1 billion market opportunity in 2004.

We profile a number of companies leading the charge toward the next phase of the Internet. While we believe most of these companies will either be acquired or go out of business, a handful of them should survive and become dominant players in existing or new markets within five years. Certainly, their pioneering efforts will form the foundation of the new Internet.

Napster: The Year of Living Dangerously

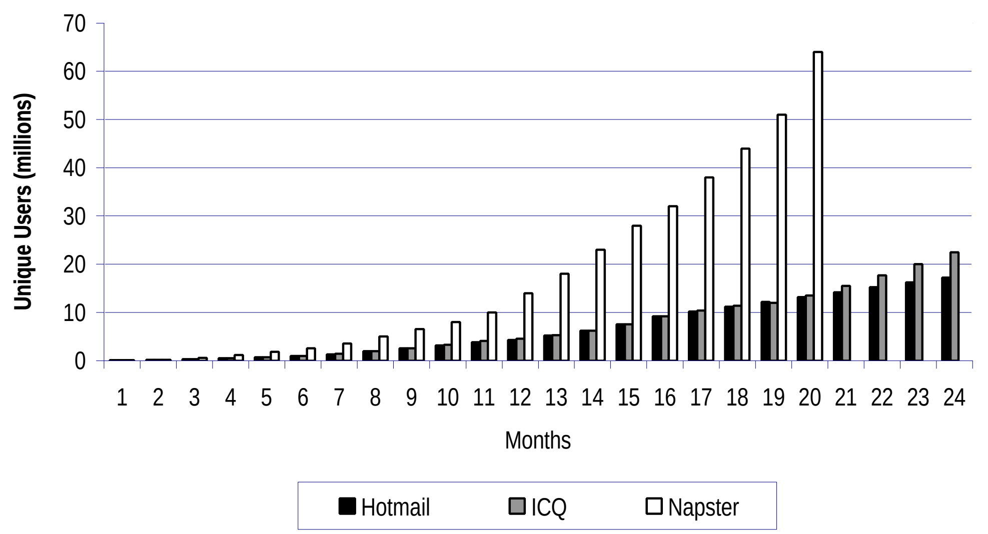
"There's a freedom about the Internet: As long as we accept the rules of sending packets around, we can send packets containing anything to anyone."

– Tim Berners-Lee, creator of the World Wide Web

65 MILLION SERVED

Napster, the popular file-sharing network used to swap MP3 music files, registered over 65 million users in just over 20 months. Compare this to Hotmail, the e-mail service begun in 1996 and acquired by Microsoft in 1998, which registered 30 million users in three years, and ICQ from Mirabilis, the instant messaging (IM) service acquired by America Online, which, coupled with AOL Instant Messenger (AIM), registered 50 million users in two years. Even the Netscape Navigator browser had an installed base of only 38 million after 20 months of its release.[2]

Exhibit 1. Hyper Adoption



Source: Harris Interactive.

Given that all three of these services are free, why has Napster's adoption rate far exceeded those of Hotmail and ICQ? And what could this tell us about the impact of particular network topologies and architectures?

THE VALUE OF DISTRIBUTION

E-mail uses the public DNS (domain name system) whereby an identifier like *ckwak* is attached to a domain name like *bear.com*. An e-mail sent to *ckwak@bear.com* is delivered to the recipient via the public DNS. The logical domain name address, *bear.com*, is mapped to a physical IP address (147.107.43.3), and a mail server delivers the e-mail to *ckwak*, an individual user on the *bear.com* network.

Instant messaging works a bit differently. Instant messaging operates within its own name space — that is, AOL's ICQ has its own directory that maps a user ID to the appropriate IP address. Instant messaging does not use the public DNS. When one

[2] No doubt, there are more users connected to the Internet now than when ICQ launched, which also had a larger installed base to tap into than when Hotmail launched. However, Napster's qualities explain, in large part, ceteris paribus, why Napster would have the steepest adoption curve across the three services.

sends an instant message to a friend using ICQ, their instant messenger software determines their friend’s IP address (AOL maintains a mapping of users’ IDs to their IP addresses during their sessions) and thus delivers messages immediately.

The architecture of Napster is very similar to ICQ. Napster users must download and install a software client that connects the user to Napster’s central resolution server (maintained by Napster) that maps user IDs to IP addresses, maintains a metadatabase of locations of songs, powers searches across the database, and connects users directly to one another to complete the download.

For instant messaging to work effectively, it requires a persistent *presence* (not just a persistent *connection*). Presence detection is one of the core elements of instant messaging — the system knows when a user is “on” the network. However, like e-mail, instant messaging users communicate in a linear, *two-way* fashion. Without the recipient’s awareness of the message, no receipt can be generated. Therefore, like e-mail, instant messaging requires what we refer to as *recognition* in order for the communication to have meaning. In addition, e-mail and instant messaging require users to constantly publish. E-mail and instant messaging both follow the following system: publish, consume, publish, consume in linear fashion — that is, one writes, another responds, one responds, and so forth.

Exhibit 2. Linear Transactions/Communications in E-Mail and Instant Messaging

x-axis (time)

Source: Bear, Stearns & Co. Inc.

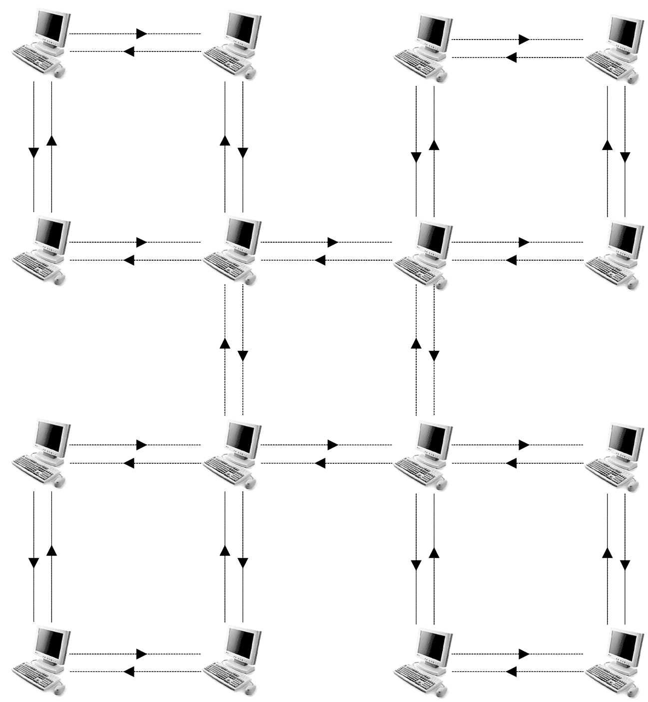
Functionally, e-mail and instant messaging require the two parties (sender-recipient) to “know” each other. A mutual “handshake” of sorts must take place in order for the message to have meaning.

Napster is very different. Participants on Napster are virtually anonymous, in the sense that no party needs to personally know another on the network in order to join, transact, and benefit from the network. In the Napster network, participation is a one-way street. Once a user sets access control rules (setting which files on one’s computer will be shared with others on the Napster network), transactions are generally active in one direction, the file requestor’s. The sender is laissez-faire.

In addition, Napster is unlike e-mail and instant messaging in how it values a user on the network. In e-mail and instant messaging, the value of a new user is primarily realized by the user and those who know the user. With e-mail and instant messaging, one receives little to no value from an individual whom one does not know and will never likely know or with whose circle of friends one does not intersect or interact. This is not the case with Napster.

The value of a user on Napster is captured and spread across the entire network. That is, while one benefits from joining the network (as the network as a whole is a resource to an individual), the entire network also benefits from the individual’s participation. What’s more, once another user has replicated an individual’s data, he or she can now act as a potential distributor as well, so the individual’s contribution is multiplied by use.

Exhibit 3. Non-Linear Simultaneous Transactions/Communications on Napster



Source: Bear, Stearns & Co. Inc.

Napster requires little to no personal communication between parties. The vast majority of Napster users know nothing about the people from whose computers they download music. The reason for this is that access control limits, once set, enable participants of a network to move fluidly around the network and to adapt naturally to changing environmental factors. This presumes participants are allowed to relinquish some assets (e.g., control) in order to gain others (music). It is, we believe, very telling about what certain network topologies enable the network to accomplish.

Exhibit 4. Characteristics of Network Services

	Hotmail	ICQ	Napster
Content Type	Text	Text	File
Architecture	Client-Server	Client-Client (1)	Client-Client (1)
Address Scheme	chris_kwak@hotmail.com	groovieguy	kwak
Protocol	SMTP	IM	FTP/HTTP
Interaction	Singular	Singular	Multi-threaded
Content Origin	Created	Created	Duplicated
Duration	Delayed	Instant	Instant
Data Resides On	Server	Client Device	Client Device

(1) While ICQ and Napster both require a central registry server to oversee the mapping of logical names (groovieguy, kwak) to physical addresses (IP address), once the mapping is done, the transaction between users is unmediated and therefore direct.

Source: Bear, Stearns & Co. Inc.

Legal concerns aside, we believe there are significant network topology lessons to be learned from Napster:

- **Exponential Growth.** Networks, given a coincidence of exogenous and endogenous factors, can propagate exponentially. In Napster's case, these factors include software, persistent broadband connections, abundant disk, standard formats, protocols, and peripherals.

- **Content Utility.** The fact that music emerged as the first ubiquitously exchanged content type on a decentralized basis is telling. Music files possess characteristics of storability, exchangeability, popularity, and ubiquity, along with easy searchableness.

- **Fragmentation and Distribution Yield Redundancy.** Distributed autonomous devices connected on a network create massive redundancy even on less-than-reliable PC hard drives. Redundancy on a massive scale yields near-perfect reliability. Redundancy of this scope and reach necessarily utilizes resources that lead to a network topology of implicitly "untrusted" nodes. In an implicitly untrusted network, one assumes that a single node is most likely unreliable, but that sheer scale of the redundancy forms a virtuous fail-over network. Enough "backups" create a near-perfect storage network.

- **The Consumer and the Provider — a Two-Way Street.** Consumers provide and providers consume. Publishing becomes painless. Discounting even the laundry list of problems associated with HTML editing/publishing, publishing content for a unique domain name (chriskwak.com) demands a plethora of resources: up-front costs, interactions with an ISP, and some technical acrobatics. Publishing Web sites, while much easier than before, can still pose problems. Even with free homepage sites (like Geocities), limitations such as storage capacity, advertising nuisances, and lost branding render publishing difficult and the network underutilized. For various reasons, publishing a product even on eBay, for example, can often prove difficult.

- **Utility, Ease, Cost, Necessity.** Usage is a function of four variables: utility of use, ease of use, cost of use, and necessity of use. The Web is ubiquitous in large part because the browser allows universal navigation, and the browser is ubiquitous because the graphical user interface and the rules associated with its use (hyperlink, toolbars, HTML, HTTP, URL addressing) are simple. It's free; it is necessary in order to surf on the Web; surfing on the Web has high utility.

- **Fluidity and Viscosity of Data.** How easily digital data flow — the *fluidity of data* — is crucial to their exchange and propagated use; if effort must be exerted for data to flow from sender to recipient (through firewalls, sessions, passwords, intermediate storage stops, and so forth), the high *viscosity of data* will render it less mobile. Digitization of data decoupled data from physical media and made data potentially fluid. Networks can make data actually fluid. To achieve fluidity and to overcome viscosity, data must realize geographic, platform, and hardware independence.

- **What's in a Name Space?** Before instant messaging, there was but one primary name space — the domain name space (DNS). Instant messaging requires a separate name space because unlike the DNS, it needs to be able to map a user ID specifically to an IP address — that is, generally, domain names are associated with servers or server clusters in a datacenter (at a content provider, service provider, or corporation), while instant messaging maps an ID directly to the IP address of a client device like a PC. Alternative name spaces enable applications like Napster to operate free of the DNS and its limited ability to map a domain name directly to a client device.

- **Metadata.** Metadata — data that describes data — make data searchable. As such, metadata lie at the core of data discovery and search. Simple metadata enable *searchability*.

WHAT DOES ALL THIS MEAN?

We believe we are in the midst of a transformation in network topologies and architectures (of which Napster and Microsoft .NET are but two examples) that is slowly changing the landscape of computing, networking, and communications. We refer to this as Internet 3.0,[3] or what has come to be known as the two-way Web, the radically evolved offspring of Internet 1.0 (1970s-1980s) and Internet 2.0 (1993-2000), with adaptability characteristics that dramatically improve the way we communicate, transfer data, use network and system resources, increase the velocity of data and transactions, streamline workflow, maximize productivity, and optimize existing and potential computing and network systems. A fundamental wave of change is making its way through computing, networking, and communications, and we believe distribution and decentralization, with XML as its foundation, are at the core of a transformation toward universal connectivity. While Internet 1.0 and 2.0 have shaped computing, networking, and communications in astounding ways, we believe Internet 3.0 will be the focus of the computing industry over the next five to ten years.

The technologies and networks in the coming generation will favor some incumbents over others. We believe those who lead in the personal computing industry will be

[3] Dave Winer, one of the authors of the Simple Object Access Protocol (SOAP), which we discuss below, introduced us to the phrase "Internet 3.0" in March 2001.

beneficiaries from a system resources perspective. Networking equipment manufacturers, we believe, will find a second breath in the coming period, as XML increases the load across networks and as greater fluidity of data results in increased traffic on networks. As an example, according to Webnoize, an authority on the digital entertainment economy, Napster users downloaded roughly 3 billion songs in January 2001; assuming a typical three-minute song is 3 Megabytes, that translates into over 9 Petabytes of data transmitted across the Internet in January 2001 by Napster users alone.

Some who have prevailed in Internet 2.0 — the Web era — may see their roles minimized in the coming years. Those companies who have thrived during Internet 2.0 by riding the centralization wave and who have built business models solely on that trend will have to adapt or risk being obliterated by nimbler parties. This is not to say that companies with business models based on centralized architectures like Web hosters will not have a place in Internet 3.0, but merely that they will need to recognize the importance of weaving geographically distributed assets.

Internet 3.0 is not a business model or fancy technology gimmick. It is a way of networking, organizing data, and collaborating. Understanding the explicit and implicit advantages of a distributed and decentralized system leads to an overwhelming conclusion: There are too many efficiencies in distributed and decentralized systems to ignore. Internet 3.0 has been incubating for the last few years, and the developmental phase has matured enough to make some of these technologies viable.

To understand the influence of distribution and decentralization in shaping Internet 3.0, we need to view the evolutionary computing timeline holistically to see how we got to where we are today, and where we are headed. Most importantly, understanding the history of the development of the Internet and Web will help us understand the notions of synchronized and cyclical network development, which in turn will help us to understand the current state of synchronized technology development and where we are in the computing and networking cycle.

GENEALOGY OF COMPUTING

The large, inflexible, and expensive mainframe systems that proliferated in the 1960s were highly centralized because they had to be. As a function of scarce computing (processing and disk) and physical (floor space) resources, client devices that linked to mainframes were extremely thin and boasted very little power. Thin clients were essentially dummy terminals used for inputting and outputting text-based data. Few processing tasks were undertaken or data stored by the thin client. Centralizing processing and storage functions at the mainframe maximized the economic return of these networks. The processing power, disk space, and application intelligence of clients (because of economic feasibility and limited software development) rendered them effectively dumb and powerless. Mainframe networks persist today because of many performance and reliability advantages, like handling large volumes of data. Nevertheless, the general network topology of a mainframe system remains the same after 40 years. Today, thin clients continue to be nothing more than dummy terminals (compared to bulkier clients, like PCs) connected via communications lines to a central mainframe system.

In 1970, DEC released the first minicomputer system. The impact was the gradual loosening of centralized computing’s stranglehold. The minicomputer helped to stretch the fabric of the local computing environment.

Around the same time, Robert Metcalfe, then a Harvard Ph.D. candidate, was drafting a protocol for reliable connectivity across a local environment. In 1973, Ethernet was born.

The arrival of the personal computer provided the ultimate kickstart to fragmentation and segmentation of networks, as networks pushed into corporate departments and the power of the computing network began to shift toward the individual PC.

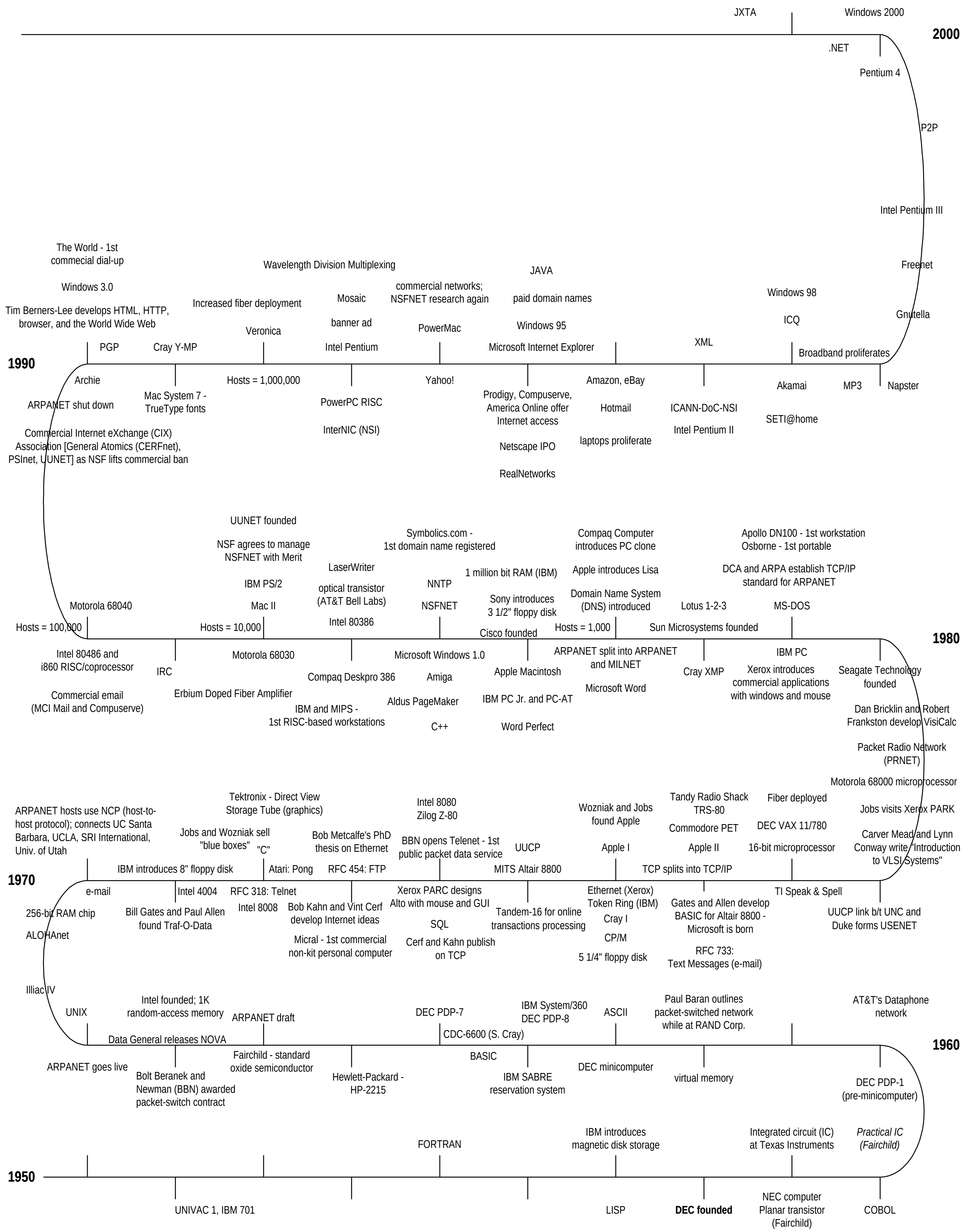
When the Internet began to be commercialized in the late 1980s, more and more devices began to connect to networks. This network development blossomed in 1993 with the introduction of the Mosaic browser, the first commercialized graphical browser for the World Wide Web. Devices connected to the network could now do more than ever.

The consequence of the browser was the introduction of repositories of information in the “Internet cloud.” Large server farms surfaced to house and serve content, and Web hosting concerns arose to manage this equipment in a centralized environment.

What is important to note is that these developments did not occur in isolation. The very nature of computing, processing, and networking is predicated on the push and pull of each on the other. Computing is a function of improved microprocessors; the utility of improved processing is limited by software development; software development has significantly less value without collaboration; collaboration is impossible without network development; network development is difficult without protocols in place.

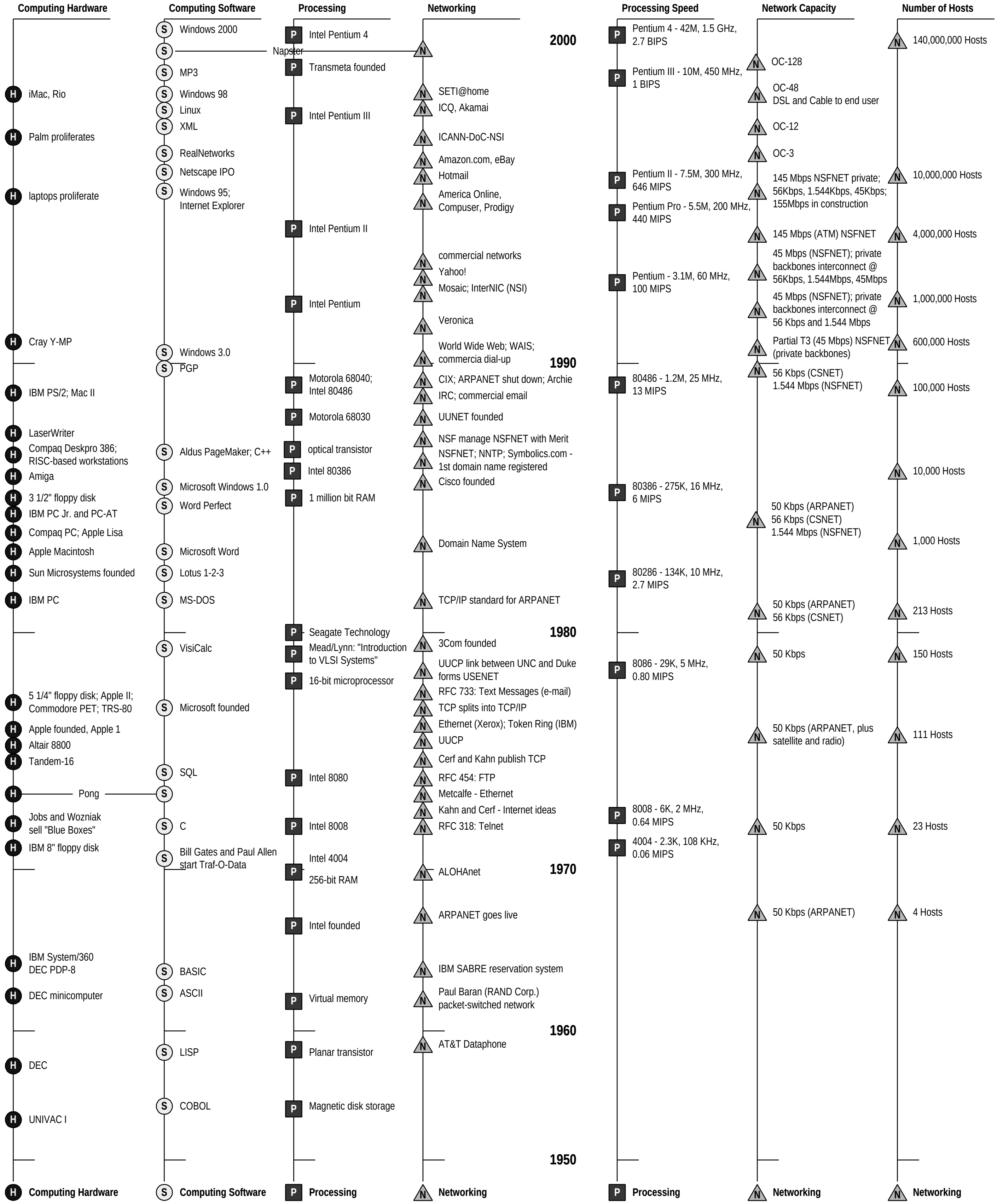
We believe the synchronism of developments in computing, processing and networking is astonishing. We highlight below what we believe are particularly fascinating parallels in time across these industries.

Exhibit 5. Computer and Internet Development Timeline



Source: Bear, Stearns & Co. Inc.

Exhibit 6. Synchronized Development — Processing, Computing, and Networking



Source: Bear, Stearns & Co. Inc.

The synchronized nature of these developments and discoveries is an interesting sidebar to computing history. Allowing that developments in processing and networking are a function of laws, such as Moore's Law and Gilder's Law, we believe the synergies between these laws are less clearly understood. While the laws of computing and networking underlie technologic evolution, an explanation of the interplay of these laws is often missing. Below, we outline the laws relevant to this report, and highlight some of the synergies between these laws.

The Laws of Network Dynamics

The Twelve Networking Truths

1. *It has to work.*
2. *No matter how hard you push and no matter what the priority, you can't increase the speed of light.*

   (2a) *(corollary) No matter how hard you try, you can't make a baby in much less than 9 months. Trying to speed this up *might* make it slower, but it won't make it happen any quicker.*

3. *With sufficient thrust, pigs fly just fine. However, this is not necessarily a good idea. It is hard to be sure where they are going to land, and it could be dangerous sitting under them as they fly overhead.*
4. *Some things in life can never be fully appreciated nor understood unless experienced firsthand. Some things in networking can never be fully understood by someone who neither builds commercial networking equipment nor runs an operational network.*
5. *It is always possible to agglutinate multiple separate problems into a single complex interdependent solution. In most cases this is a bad idea.*
6. *It is easier to move a problem around (for example, by moving the problem to a different part of the overall network architecture) than it is to solve it.*

   (6a) *(corollary) It is always possible to add another level of indirection.*

7. *It is always something.*

   (7a) *(corollary) Good, Fast, Cheap: Pick any two (you can't have all three).*

8. *It is more complicated than you think.*
9. *For all resources, whatever it is, you need more.*

   (9a) *(corollary) Every networking problem always takes longer to solve than it seems like it should.*

10. *One size never fits all.*
11. *Every old idea will be proposed again with a different name and a different presentation, regardless of whether it works.*

    (11a) *(corollary) See rule 6a.*

12. *In protocol design, perfection has been reached not when there is nothing left to add, but when there is nothing left to take away.*

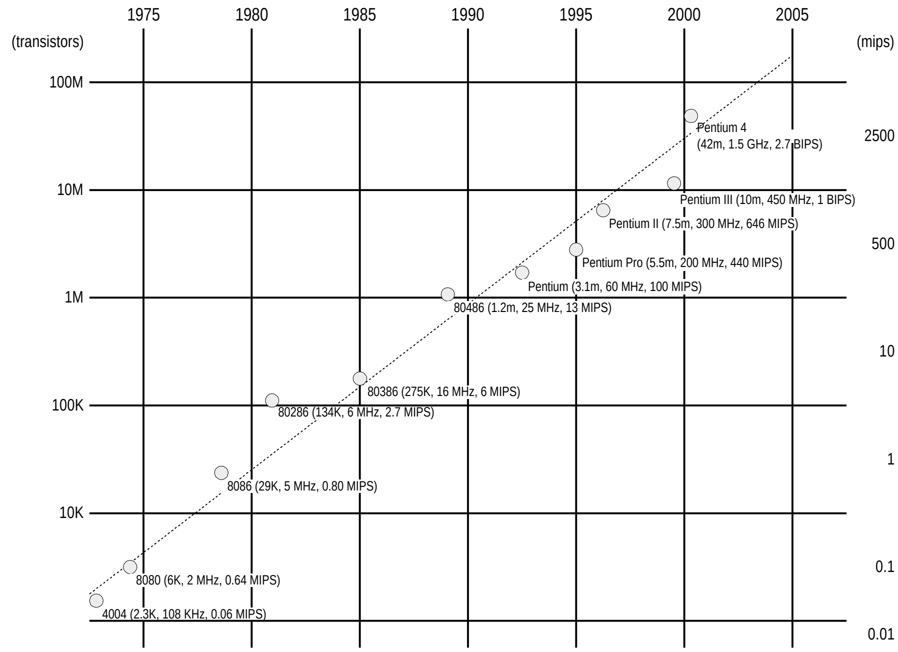
– Ross Callon
Bay Networks, 1996

MOORE'S LAW

In 1965, *Electronics* magazine published "Cramming More Components onto Integrated Circuits." Intel founder Gordon Moore (then Director of Research and Development Laboratories at Fairchild Semiconductor) predicted that processor densities would double every 12-18 months into the foreseeable future.[4] In particular, Moore posited that by 1975, it would be possible to squeeze up to 65,000 components on a silicon chip while maintaining cost constancy. This observation has been incredibly accurate over the last 35 years. We have come to refer to it as Moore's Law.

Though Moore's observation was in reference to circuit densities of semiconductors, it has more recently come to describe processing power.

Exhibit 7. Moore's Law



Note: (transistors, clock speed, MIPS — Millions of Instructions per Second)

Source: Intel; Bear, Stearns & Co. Inc.

Moore's Law has been the single most pervasive law of the computing world. Until recently, the primary force of computing has been chip density, as this drives processor and storage performance. Until the laws of physics on the quantum level are breached, it is likely that Moore's Law will continue to be on the mark.[5]

[4] Moore viewed the first planar transistor in 1959 as the genesis of this exponential phenomenon. In commercial production since 1961, the earliest planar transistor showed a chip containing 2^5, or 32 components. Moore also knew that the next integrated circuit still in the lab but to be unveiled in 1965 would contain about $2^{(5+1)} = 2^6 = 64$ components. Looking ten years out meant a chip would contain $2^{(6+10)} = 2^{16} = 65{,}536$ components. The log graph of this phenomenon is a straight line.

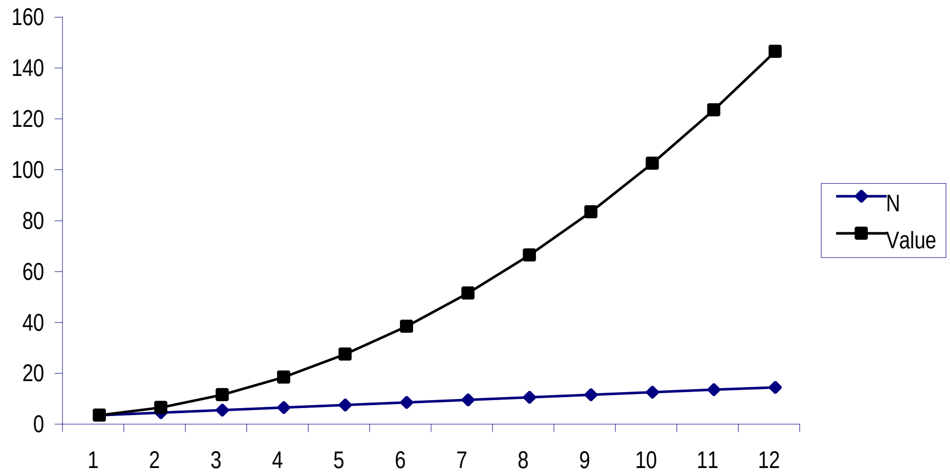
[5] Unless Moore's second law (Economics) takes over. In 1995, Moore argued in an *Economist* article: "What has come to worry me most recently is the increasing cost. This is another exponential. In today's dollars, the cost of a new fab has risen from $14 million in 1966 to $1.5 billion in 1995. By 1998, work will begin on the first $3 billion fabrication plant. Between 1984 and 1990, the cost of a fab doubled, but chip makers were able to triple the performance of a chip. In contrast, the next generation of fabs will see cost double again by 1998, but this is likely to produce only a 50% improvement in performance. The economic law of diminishing marginal returns appears to be setting in. If this exponential trend continues, by 2005, the cost of a single fab will pass the $10 billion mark (in 1995 dollars) or 80% of Intel's current net worth."

METCALFE'S LAW

Metcalfe's Law — articulated by Bob Metcalfe, founder of Ethernet in 1973 while then at Xerox PARC — stipulates that the value of a network is proportional to the square of the number of devices on the network.[6] In technical form, the number of possible connections in a network equals the square of the number of devices on the network.

Exhibit 8. Metcalfe's Law

network value	~	(# of users) x (value per user)
value per user	~	(# of users)
network value	~	$(\text{\# of users})^2$



N = number of devices
V = value of network

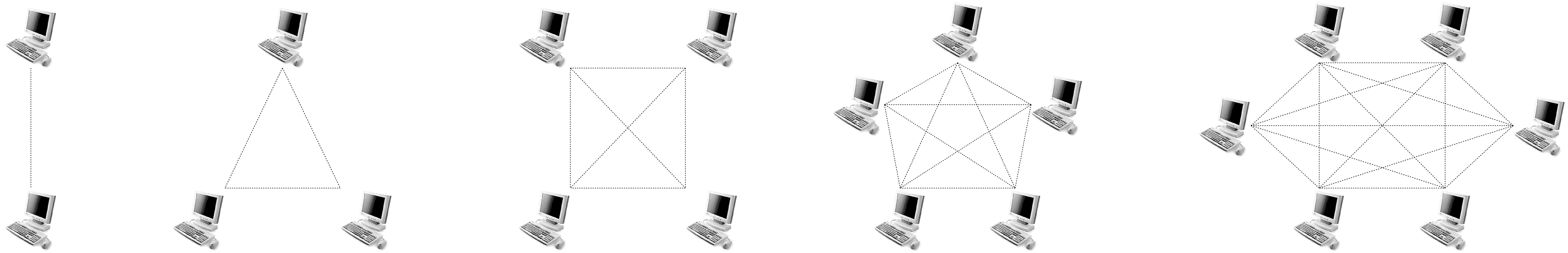
Source: Bear, Stearns & Co. Inc.

The original statement from Bob Metcalfe was apparently (according to one source): "The power of the network increases exponentially by the number of computers connected to it. Therefore, every computer added to the network both uses it as a resource while adding resources in a spiral of increasing value and choice."

One criticism of Metcalfe's Law has been that it has no predictive power. We will show otherwise.

6 For Napster, however, we point out that the value of the network is proportional to the number of files at the node, where the number of files is a function of the storage capacity at the node: $v = (n \propto f \propto s)^2$, where v = value, n = number of nodes, f = number of files, s = storage capacity of node.

Exhibit 9. Metcalfe's Law Illustrated



Number of devices (N)	2	3	4	5	6
Number of possible connections (C)	1	3	6	10	15
Square of number of devices (N^2)	4	9	16	25	36

Value of network (V)

$C \sim (N^2 - N)/2$

$V \sim N^2$ according to Metcalfe's Law. (As we discuss in this report, in a perfect networked environment, V should also equal C.)

$V \sim C \sim (N^2 - N)/2$

As N grows

$V \sim N^2/2$

However, when we assume the number of potential connections is bi-directional (between two devices), then $V \sim (N^2/2)*2 \sim N^2$

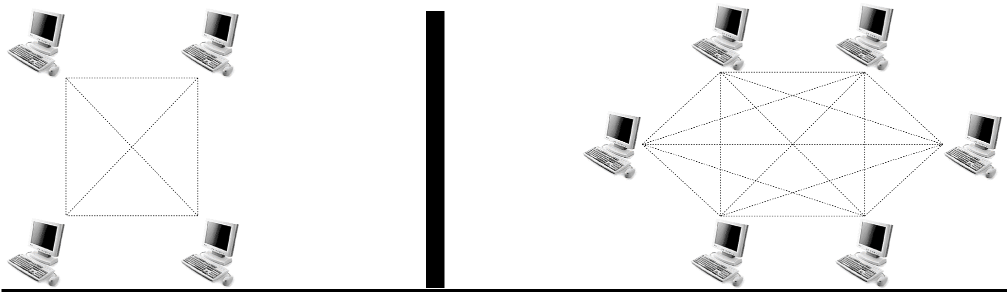
Note: We believe the value of a network is also a function of both the number of connections among devices and the bandwidth capacity of the network. In addition, we believe the value of a connection is a function of storage capacity and shareable information stored.

Source: Bear, Stearns & Co. Inc.

ALLOWING RESOURCES TO BE FREELY SHARED ACROSS NETWORKS MAXIMIZES METCALFE'S LAW

The architecture and topology of a network clearly has an impact on the value of Metcalfe's utility function. For example, should a network be compartmentalized, as in Exhibit 10, there would actually be two networks, each with a different utility function.

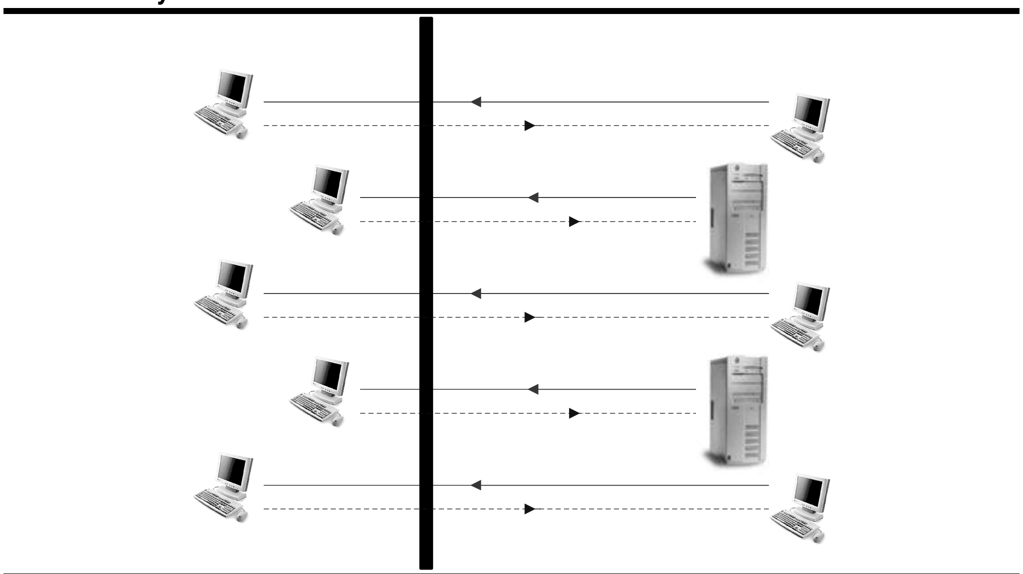
Exhibit 10. Isolated Networks — "Magnets"



Source: Bear, Stearns & Co. Inc.

The Internet is a collection of these fragmented, isolated islands, or what we refer to as *magnets*. Magnets are networks or network segments that isolate devices or device clusters from other devices or device clusters. A good example is a corporate intranet. The rise of the intranet with firewalls and proxy servers and other security devices and measures (like Dynamic Host Configuration Protocol and Network Address Translation: see Appendix B) has enabled sub-networks essentially to hide behind walls, disconnecting and fragmenting networks. *Devices on an intranet benefit from the public Internet, however, they do not typically contribute to the public Internet at large.* Data, in these instances, flows *uni-directionally*, or inward.

Exhibit 11. Asymmetric Dataflow



Source: Bear, Stearns & Co. Inc.

Even in an open network, the asymmetry in dataflow hampers fluid communication and leads to the underutilization of resources. Most information is drawn from servers by edge devices (like PCs), and in such instances, dataflow cannot be symmetric.

We believe there has to be symmetry in dataflow (connectivity) to maximize network value, since without symmetry, some portion of the devices connected to a network is underutilized.

If we look at a physical connection as a two-way link, then the value of a network is directly proportional to the square of the number of devices.

Exhibit 12. Bi-Directionality Is a Centerpiece of Maximizing Network Value

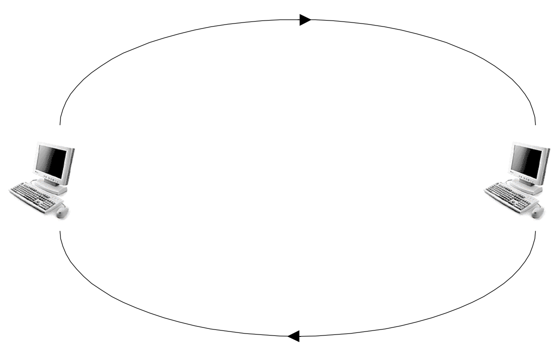
$$V \sim (N^2/2) * (\text{2 connections per single physical connection}) \sim N^2$$

Source: Bear, Stearns & Co. Inc.

From this, we catch a glimpse of the significance of *bi-directionality* in calculating and augmenting the value of a network. This is an explicit problem with centralized client-server (cC-S) systems: connections are mostly *uni-directional*. Servers are "gravity wells" and therefore push and pull data centrally.

We believe every connection in Metcalfe's paradise should be implicitly two-way symmetric. That is, a connection is a link between two devices; it assumes that dataflow is symmetric and that the devices can mutually communicate with one another.

Exhibit 13. Symmetric Dataflow

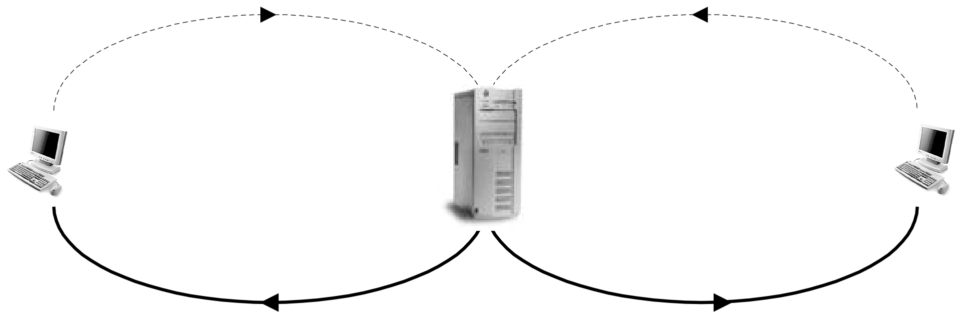


Note: In this scenario, the PC on the right (PC_R) and PC on the left (PC_L) share a symmetric connection. One example that illustrates this symmetry is Napster, where users can download and upload almost equally.

Source: Bear, Stearns & Co. Inc.

However, there are scenarios where the connection between two devices is asymmetric. While the first device may be able to utilize resources of the second, the second may not be equally able to access or use the first.

Exhibit 14. Asymmetric Dataflow



Note: In this scenario, the PCs demand more of the server than the server of the PCs. This is a typical relationship between a server and a client device. The result is asymmetry. An example of asymmetry is between a Web server hosting a streaming media broadcast and a PC. Once the PC demands the data, the flow of data is uni-directional.

Source: Bear, Stearns & Co. Inc.

This is clearly what happens between *unequal* devices — say, between a PC and server.

Asymmetries abound in networks and are often artificially constructed. Our belief is that asymmetries, while sometimes necessary, introduce network inefficiencies. If devices on a network can operate more symmetrically, systems resources are better utilized. In fact, asymmetric connections introduce uneven dataflow and resource strains. A good example is intranet-based streaming, where streaming content flows in one direction — from the streaming device to the user device.

CRITICAL MASS OF A NETWORK

What is critical mass and when do we achieve it? Stated differently, when does a network reach a point when the next entrant/participant in the network contributes zero marginal value? In an auction exchange of five participants, for example, the entrance of a sixth participant increases the value of the network meaningfully. In an exchange with ten million participants, would the next entrant add close to zero marginal value? We do not think so. We believe a network has reached critical mass when the pool of data (the virtual database) contains enough objects such that each participant's search or query can be satisfied by that installed base of data. Therefore, if the $10^6 + 1$ participant brings with him a basket of items not present in the network before his entrance, his value is significant.

With Napster, over 65 million users being on the network does not necessarily indicate that the 65 million + 1 user adds near-zero marginal value to the network. If there are no Beatles songs on the network before the 65 million +1's entrance, then that individual adds significant value to the network, since searches for Beatles songs can now be satisfied. If Napster's network can satisfy all searches, the network has achieved critical mass.

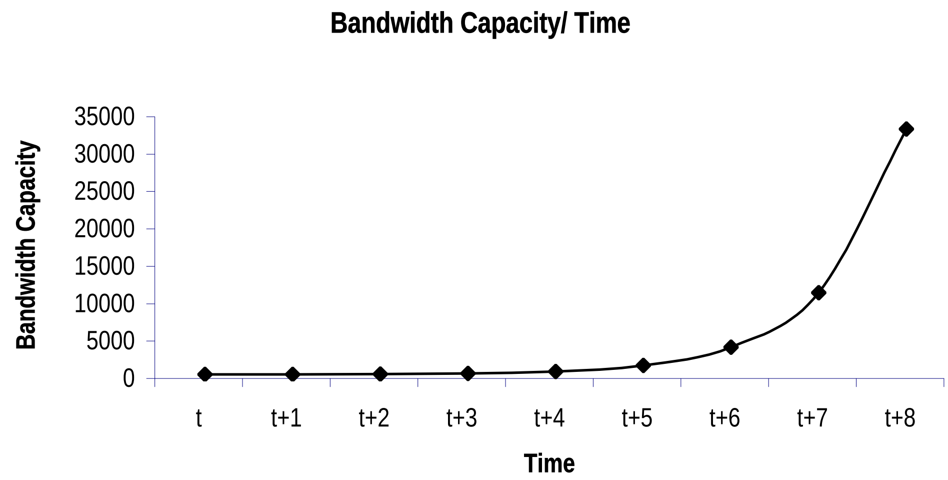
Even when a network achieves critical mass, additional entrants can still add significant value. Generally, these additional entrants add a layer of redundancy.

GILDER'S LAW

George Gilder has predicted that network capacity (measured in bandwidth) triples every 12 months. Like Moore's Law, advances in technology and continued R&D

make Gilder's Law possible. Because bandwidth capacity, according to Gilder, triples every year, and chip density doubles every 12-18 months, Gilder's Law is the dominant law of the networked world.

Exhibit 15. Gilder's Law (in Kbps)

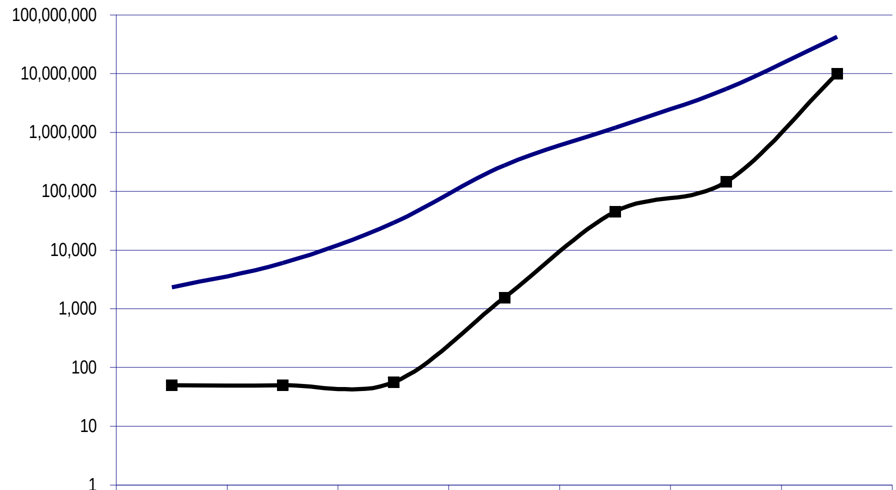


Source: Bear, Stearns & Co. Inc.

Gilder's Law posits that bandwidth capacity growth outstrips Moore's Law. While some believe comparing the number of transistors on a chip to bandwidth capacity in a "pipe" may not be an apples-to-apples comparison, we believe otherwise. Comparing the throughput of a single line of fiber to the number of transistors on a chip is analogous because in both instances we are comparing density — the density of a chip (processing) to the density of fiber (bandwidth). When we plot transistors (number) with bandwidth capacity (in Kbps) logarithmically, we see that the delta between Moore's Law and Gilder's Law shrinks over time. In fact, based on Exhibit 16, bandwidth capacity (measured in Kbps) will most likely surpass transistor density (measured in units) within the next five to ten years.

Note that we plot bandwidth growth against Moore's Law because we are trying to benchmark bandwidth growth against a "standard." Moore's Law is sufficiently linear (logarithmically) to serve as an accurate benchmark.

Exhibit 16. Moore and Gilder (transistors in units, bandwidth in Kbps)



	1970	1975	1980	1985	1990	1995	2000
Transistors	2,300	6,000	29,000	275,000	1,200,000	5,500,000	42,000,000
Bandwidth (Kbps)	50	50	56	1,544	45,000	145,000	10,000,000

Ratio (bps per transistor)							
	22	8	2	6	38	26	240

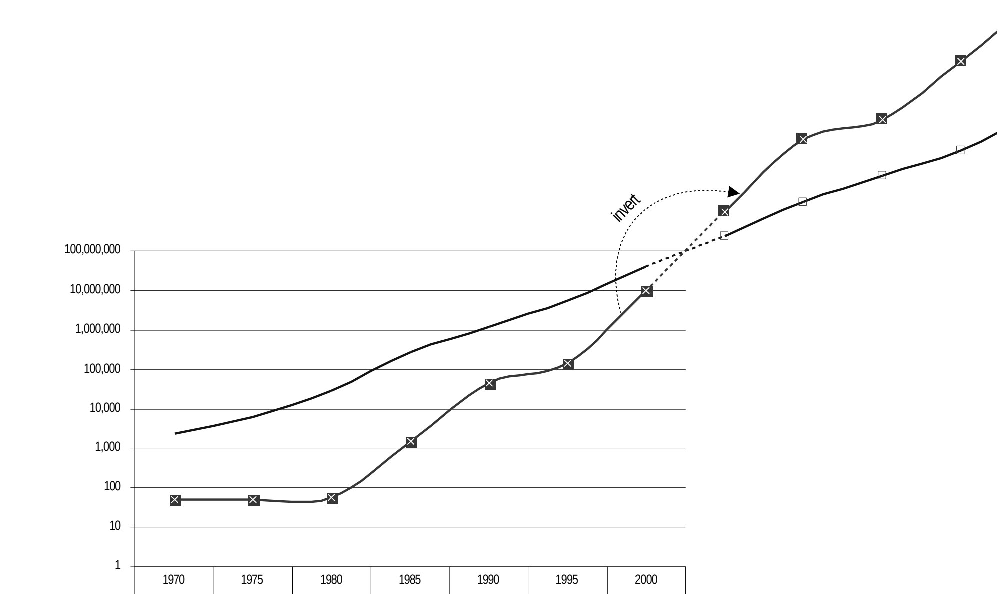
Note: The graph of bandwidth capacity tells us that the bandwidth density over time looks more like a step function than the graph of transistor density over time. The punctuated growth (when the graph of bandwidth capacity curves upward after having flatlined) in bandwidth capacity indicates that optical technologies are particularly susceptible to bursts of innovation, with periods of "dry spells," unlike semiconductor development, which has been and continues to be extremely linear.

Source: Bear, Stearns & Co. Inc.

We believe the overarching message from this graph is that the density of bandwidth continues to grow faster than the density of chips. That is, even if the semiconductor industry could no longer increase the density of chips, the networking industry would continue to increase bandwidth density. Looked at another way, the bandwidth per transistor ratio increases over time.

Our decision to use Moore's Law lies in that law's ability to act as a linear benchmark to compare Gilder's Law. We believe one way of looking at this Moore-Gilder mapping is to invert the graph of Gilder's Law and project it into the future.

Exhibit 17. Inverting Gilder (transistors in units, bandwidth in Kbps)



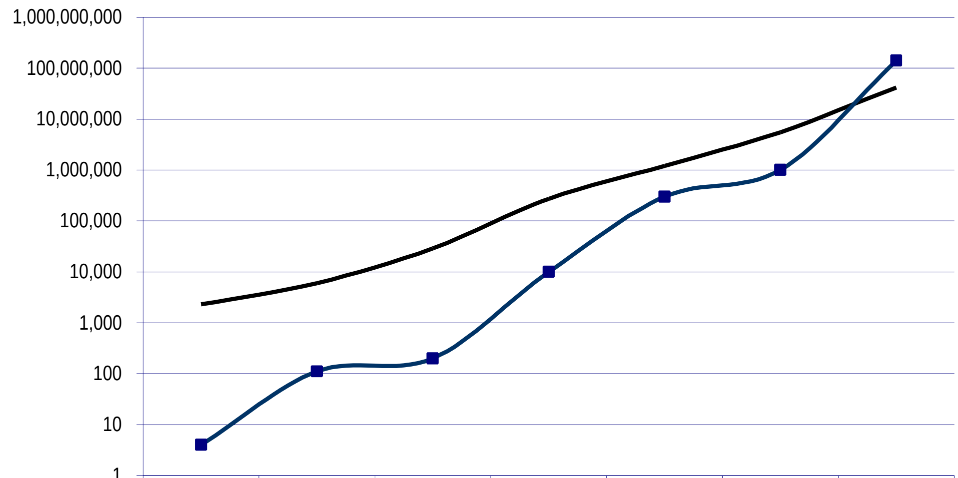
	1970	1975	1980	1985	1990	1995	2000
Transistors	2,300	6,000	29,000	275,000	1,200,000	5,500,000	42,000,000
Bandwidth (Kbps)	50	50	56	1,544	45,000	145,000	10,000,000

Source: Bear, Stearns & Co. Inc.

If both Gilder's Law and Moore's Law are indeed laws, the graph of their development into the future should resemble their historical growth.

More interesting is what Moore's Law tells us about the number of devices on a network.

Exhibit 18. Moore and Metcalfe (in units)



	1970	1975	1980	1985	1990	1995	2000
Transistors	2,300	6,000	29,000	275,000	1,200,000	5,500,000	42,000,000
Networks	4	111	200	10,000	300,000	1,000,000	140,000,000
Ratio (hosts per transistor)	0.002	0.019	0.007	0.036	0.25	0.18	3.3

Source: Bear, Stearns & Co. Inc.

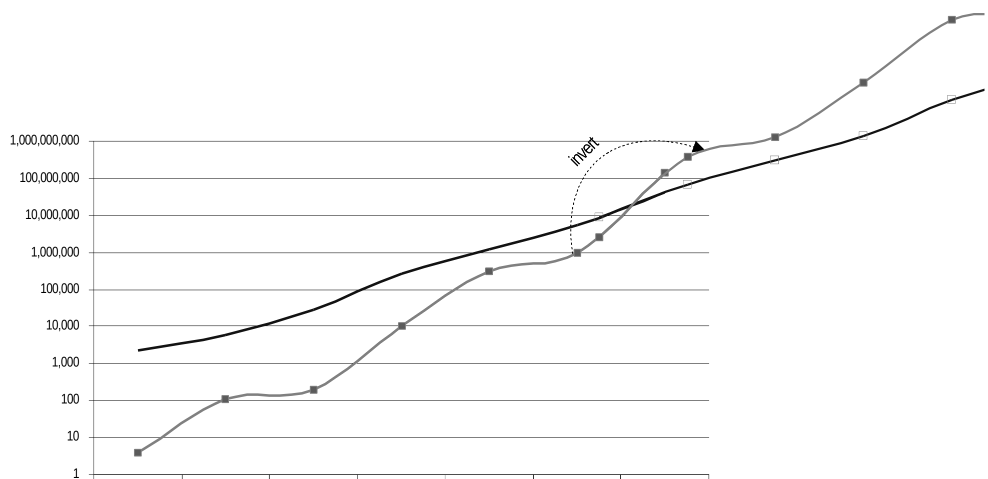
Like bandwidth capacity growth, the growth in the number of network hosts on the Internet is greater than the growth in the number of transistors on a chip. Like Gilder's Law, Metcalfe's Law[7] grows much faster than Moore's Law. In fact, the number of hosts on the Internet surpassed the number of transistors on a chip very recently.

We observe that the ratio of the number of hosts per transistor has been rising steadily over time. Therefore, the graphs of both Metcalfe-to-Moore and Gilder-to-Moore are extremely similar.

Again, the significance of this graph lies in the densities of chips and network connections and the density of devices per transistor. The inverted graph of Metcalfe's Law plotted along Moore's Law offers us a view of the potential number of network nodes going forward.

[7] We are here using Metcalfe's Law as a proxy for the density of nodes on the Internet, since the value of the Internet is directly proportional to the number of devices on it.

Exhibit 19. Inverting Metcalfe (in units)



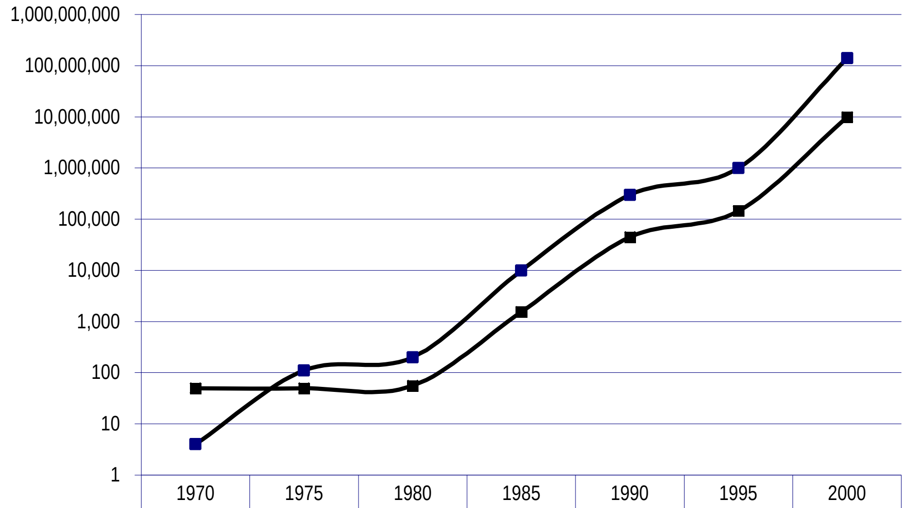
	1970	1975	1980	1985	1990	1995	2000
Transistors	2,300	6,000	29,000	275,000	1,200,000	5,500,000	42,000,000
Networks	4	111	200	10,000	300,000	1,000,000	140,000,000

Source: Bear, Stearns & Co. Inc.

Plotting either Metcalfe's Law or Gilder's Law against Moore's Law yields very little insight except directional growth.

If we plot bandwidth capacity against the number of hosts over time on a logarithmic scale, we see an immediate relationship between network bandwidth and the number of hosts connected to the network.

Exhibit 20. Metcalfe and Gilder: Plotting the Laws of Network Dynamics
(bandwidth in Kbps, hosts in units)



	1970	1975	1980	1985	1990	1995	2000
Bandwidth (Kbps)	50	50	56	1,544	45,000	145,000	10,000,000
Hosts	4	111	200	10,000	300,000	1,000,000	140,000,000
Ratio (hosts to bandwidth, Δ)	1:13	2:1	4:1	7:1	7:1	7:1	14:1

Source: Bear, Stearns & Co. Inc.

The parallel motion of bandwidth capacity and the number of hosts over time suggests that the development of the two is tightly coupled; and though we cannot say with certainty which is the leading indicator, since Gilder’s Law is a predictive law, so too now is Metcalfe’s Law.

The number of hosts always leads the amount of bandwidth available to them. That is, in 1995, we see that network capacity was about 100,000 Kbps (100 Mbps) and that the number of hosts on the network was about 1,000,000. Over time, the ratio of hosts to bandwidth increases. Put in another way, each host is consuming less bandwidth as a percentage of available bandwidth. This is attributable in large part to the increasing number of networks weaving the Internet. The number of networks able to host devices proliferated in the 1990’s with the growth of new networks like Qwest, Level 3, and Global Crossing. However, because there is only one Internet (while there are innumerable fiber lines capable of carrying the highest capacity) the comparison between bandwidth capacity and the number of nodes on the Internet is indeed an apples-to-apples comparison.

The delta between hosts and bandwidth increases over time. In the beginning (around 1975), roughly 100 hosts were on the 50Kbps NSFNET network. That’s a 2:1 ratio, measured in number of hosts to bandwidth capacity. By 1985, that ratio had become 7:1, which remained through 1995, at which point the ratio turned abruptly to 14:1.

The introduction of the Web in 1993 certainly increased the number of hosts on the network. The proliferation of wireless devices also added to the growing number of network hosts.

We believe the trend will be an increasing delta between the number of hosts and bandwidth capacity, in line with the historical trend. Given an amount of bandwidth capacity, over time the network will host more devices on that bandwidth. Metcalfe's Law will be more relevant in the next phase of the Internet.

As a vacuum wants to be filled, the network naturally seeks to maximize Metcalfe's Law. If we extrapolate, we estimate by 2005 there could be three billion devices on the Internet; Bell Labs estimates that the number of devices connected to the Internet will equal the number of people on the globe by 2008, or roughly 6.7 billion. We estimate bandwidth capacity could reach 160 Gbps per fiber by 2005. Currently, OC-768 (40 Gbps) is the fastest optical standard in practice. By 2005, should our extrapolation prove accurate, we could see optical fiber rates reaching OC-3072 (160 Gbps). (See Exhibit 21).

Exhibit 21. Optical Fiber Data Rates

Fiber Deployment	Speed
OC-1	51.84 Mbps
OC-3	155.52 Mbps
OC-12	622.08 Mbps
OC-48	2.488 Gbps
OC-96	4.976 Gbps
OC-192	9.953 Gbps
OC-768	39.813 Gbps
OC-3072	159.252 Gbps

Source: Bear, Stearns & Co. Inc.

Obviously, this has powerful implications for providers of fiber-based bandwidth capacity and equipment providers who sell into the fiber-based bandwidth markets. Continued capital expenditures by the major fiber bandwidth providers suggest there will be a need for all of the available capacity. The success of optical network equipment companies may depend on their ability to develop technologies to handle increasingly fast fiber and focus on equipping fewer, but higher capacity, fibers. One of the reasons we believe in the ultimate success of distributed and decentralized computing is because decentralization and distribution place greater burdens on networks. If our theories on the synchronized increase of network resources and data are correct, then the network burden of decentralization and distribution will drive the need for more bandwidth capacity.

Decentralization? What's Wrong with Centralization?

Centralized client-server (cC-S) is by definition directional and hierarchical. Directionally, a client device pulls data from a server and generally stores data onto servers. Hierarchically, the client connects to a server and is subject to the control and administration of server resources. The cC-S architecture is ideal for many applications and tasks. For example, central management of resources has been a hallmark of secure networks. Network file management and collaborative applications have benefited tremendously from stable and secure central management. However, the topology of the cC-S network inevitably yields inefficiencies, bottlenecks, and wasted resources.

Most enterprise applications and hosted services on the Internet require high capital expenditures and maintenance costs. The introduction of a streaming media offering within an enterprise requires investment in streaming server hardware and software, caches, and often load balancers.

cC-S systems are also high-maintenance. Servicing cC-S systems can be prohibitively expensive, much more costly than equipment and software. For example, Oracle derives over 70% of its revenues from services, not its database software and applications. Microsoft derives over 30% of revenues from services; this would likely be much higher if Microsoft did not sell Windows to consumers. According to IDC, the cost of managing storage is four times the cost of storage hardware.

CENTRALIZED CLIENT-SERVER SYSTEMS HAVE PROBLEMS

cC-S systems do many things extremely well. Access rights and security settings are more easily managed on a cC-S network. Centralizing storage resources enables network administrators to perform quicker backups and more efficient failover management. However, cC-S systems also have their disadvantages.

- **Your Network Administrator Should Be Your Best Friend.** In cC-S, a central administrator (usually a departmental network head) handles network permissions. The administrator must meet the needs of all clients on the network and therefore establishes broad network settings, like maximum storage limits and file access rights. Because it would require constant maintenance, granular attention is generally unavailable.

- **"The San Francisco Mail Server Is Down . . ."** cC-S systems create single points of failure. Below is an e-mail failure notice employees of one *Fortune* 500 company recently received:

> On behalf of the San Francisco office: The SF server is down and possibly will be for most of the day. We are unable to access e-mail if you need to reach us.

Even if the network has a backup system in place, the failure of a central server or storage system can cause at least a major delay and at worst the death of the network. We have all experienced e-mail when a Microsoft Exchange server has gone down. Often, this occurs when traffic load exceeds server capacity — as can occur when a series of "blast" e-mails with large attachments or an e-mail virus overwhelms the messaging system.

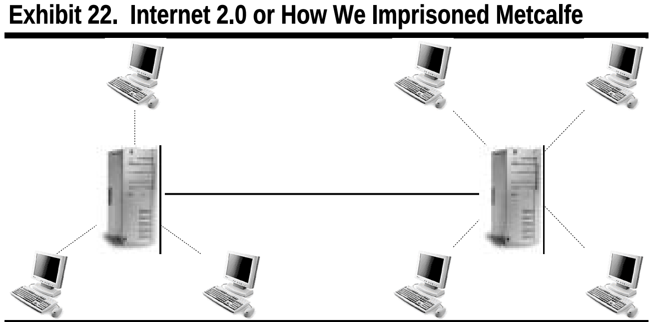
- **Geographic Locality Matters.** cC-S systems are geographically isolated. Before the modern Internet, enterprises designed their NOCs (Network Operations Centers) to serve a geographically isolated enterprise environment since architecting a distributed network was prohibitively expensive and technologically unavailable. As enterprises (networks) have expanded over the last ten years, this centralized network scheme has not kept pace with the demands of geographically dispersed access points and users. So, in a cC-S system, accessing data from a satellite office, the road, or home can be difficult. cC-S systems inhibit easy access to, and the fluid movement of, data.

- **cC-S Hardware and Software Ain't Cheap.** A cC-S system is expensive, as it requires client hardware and software, server hardware and software, a plethora of storage devices and software, and maintenance. Virtually none of these components has been outsourceable. Only recently have outsourced (and distributed) storage and software services become available to corporations so they can lower their cost-of-ownership of data.

IN CC-S, WORKFLOW IS INTERMEDIATED

In a cC-S system, interaction between users is mediated by a server; this creates fragmented "magnets" — isolated sub-networks whose devices cluster around a central machine and whose communications go through the central machine. Instead of extending Metcalfe's Law of ever-expanding connectivity, the cC-S topology often destroys potential connections. This is precisely what has happened in Internet 2.0.

Exhibit 22. Internet 2.0 or How We Imprisoned Metcalfe



Source: Bear, Stearns & Co. Inc.

Exhibit 22 is a classic hub-and-spoke network design model. Hub-and-spoke designs, like cC-S systems, offer an important advantage: The network can filter anything through the server, since all data must move through the server. Because routers[8] are generally isolated to Layers 2-3 of the OSI model (Open Systems Interconnection), they cannot filter data according to type (voice, streaming video, Word documents, etc.). This is one of the primary reasons Layer 4-7 equipment manufacturers like ArrowPoint (acquired by Cisco), Alteon (acquired by Nortel), F5, and Packeteer have

[8] We would also like to emphasize the oft-overlooked role of the modern router in transforming the networking landscape and in shaping Internet 2.0. Routers by definition operate below what are regarded as the "intelligent layers" of the Open Systems Interconnection (OSI) model.

proliferated. Layer 4-7 devices are capable of comprehending the type of data flowing through the network.

The most noteworthy bottleneck of the hub-and-spoke model, we believe, is this: While servers can process requests in parallel (simultaneously), every server (software) has a characteristic threshold after which it slows to a crawl or "crashes." The way hub-and-spoke networks have been designed, servers are expected to go down. The problem is serviced. Problems arise; problems are serviced. This is the reason services fees (administrative and support personnel) are the primary component of expenses for large corporate IT budgets and the largest portion of revenues for established enterprise software vendors.

In the ideal world, internetworking between devices and networks would be truly "Metcalfeian," where individual devices like PCs could communicate directly with others.

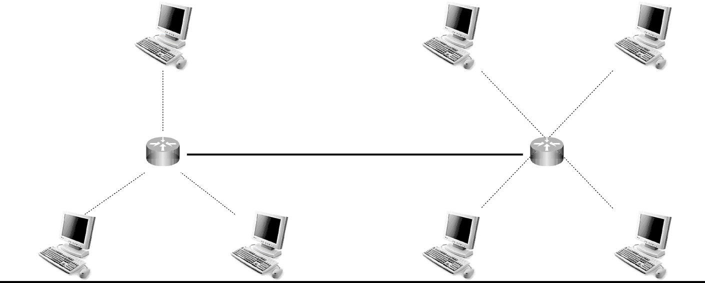
Exhibit 23. OSI Model

	Layers of the OSI Model		Representative Examples	Representative Vendors	Network Topology
7	**Application** (e-mail, file transfer, client-server)	message →	SMTP (e-mail)	Microsoft	Microsoft .NET
6	**Presentation** (encryption, data conversion)	message →	MPEG (video)	RealNetworks	RealPlayer
5	**Session** (starts, stops)	message →	AppleTalk (name binding)	ArrowPoint (L-5 URL switch)	Akamai
4	**Transport** (ensures file/message delivery)	message →	TCP (Transmission Control Protocol)	Alteon (L-4 load balancer)	Exodus
3	**Network** (routes data across networks)	packet →	IP (Internet Protocol)	Cisco (router)	Level 3, Qwest
2	**Data Link** (transmits packets from node to node)	frame →	LAN (Local Area Network)	3Com (hub)	UUNET
1	**Physical** (medium, devices, topology)	bit →	Cable, Microwave	Corning (fiber)	Cablevision

Source: Bear, Stearns & Co. Inc.

Because routers operate in Layers 2-3 of the OSI model, their contribution to network intelligence is limited. The intelligence of a network must reside on devices — e.g., servers and PCs. The processing, direction, and organization of data must occur at the inner (server) or outer (PC) nodes. Since routers continue to be the primary intermediating network points, we expect the intelligence of the network to come from devices.

Exhibit 24. Routers Disintermediate Servers



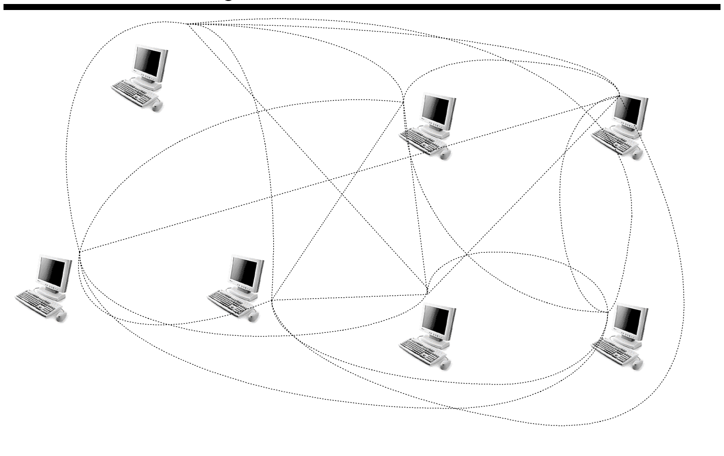
Source: Bear, Stearns & Co. Inc.

In this scenario, we remove the server from the transaction. When data flows from user to user, no server intermediates the transmission. The router merely directs the data; the user device is central. Data moves from user to user, user device to user device. This is what happens in a cC-S system anyway: We download files temporarily onto our PC from a server, work on it, and then resave it to the server; files are hosted by the PC as long as we are "in" the file. In a cC-S system, however, files are always transmitted from the server to the client device and back to the server.

Exhibit 24 also has device-to-device conversation implications. While the router is the central director in a network, we do not "talk" to a router as we do to a server. The router does not store or process content; it merely directs packets toward a destination. Replacing the server with a router creates a more fluid network. (Ironically, a router's latency is measured by how long a packet is stored by the router.)

dC-S systems are about making the router the central point of the network, and forcing processing, storage, and intelligence to edge devices. If the server is squeezed out of the picture, then processing, storage, and intelligence must reside on client devices.

Exhibit 25. Liberating Metcalfe



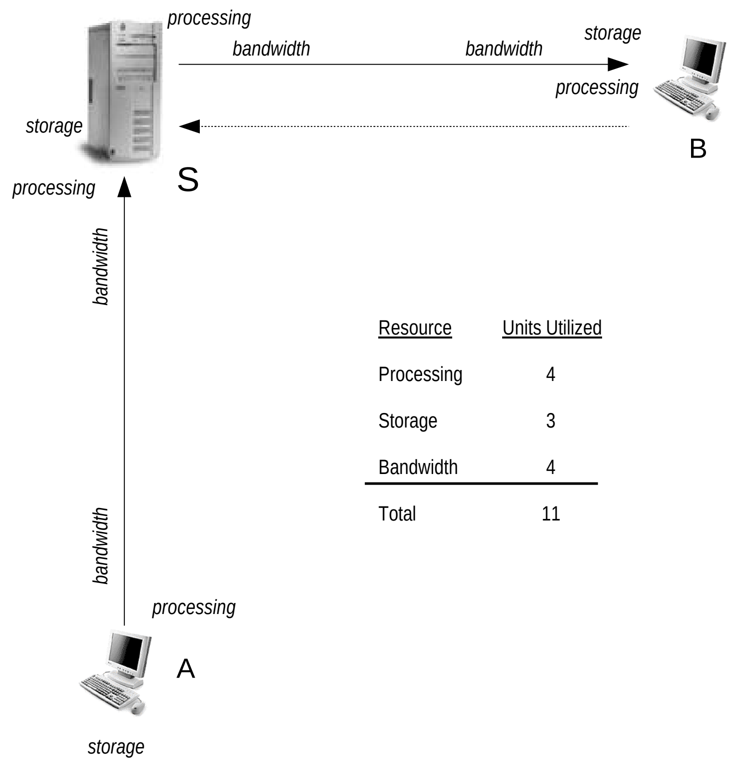
Source: Bear, Stearns & Co. Inc.

In this iteration, we see the number of connections surging. This is the result of devices communicating directly with other devices, not through servers acting as intermediaries for devices.

IN A CENTRALIZED SYSTEM, RESOURCES ARE WASTED

Another way to evaluate a system is to examine the system's resource utilization. Within a computing or network system, the three resources are processing, storage, and bandwidth. More efficient utilization of these resources (without introducing negative externalities) improves the system's efficiency. In addition, there is an elegance in creating a system that boasts no extraneous variables (see Networking Truth 12)[9]. With this in mind, we offer a simple illustration of a transaction in a cC-S environment detailing the terms of connectivity and latency, and exposing the limited velocity of data and inefficient resource utilization.

Exhibit 26. What's Wrong with Centralization



Resource	Units Utilized
Processing	4
Storage	3
Bandwidth	4
Total	11

Source: Bear, Stearns & Co. Inc.

In this transaction, user A has a file that user B wants. In order for A to deliver a file to B, A must upload the file to a centralized server or storage resource (S). In the process, A utilizes storage, processing, and bandwidth resources. S utilizes bandwidth, processing, and storage in order to capture the file and store it, and must expend processing and bandwidth resources to send the file to B. B uses bandwidth, processing and storage resources to receive the data.

DECENTRALIZATION IS LIBERATING

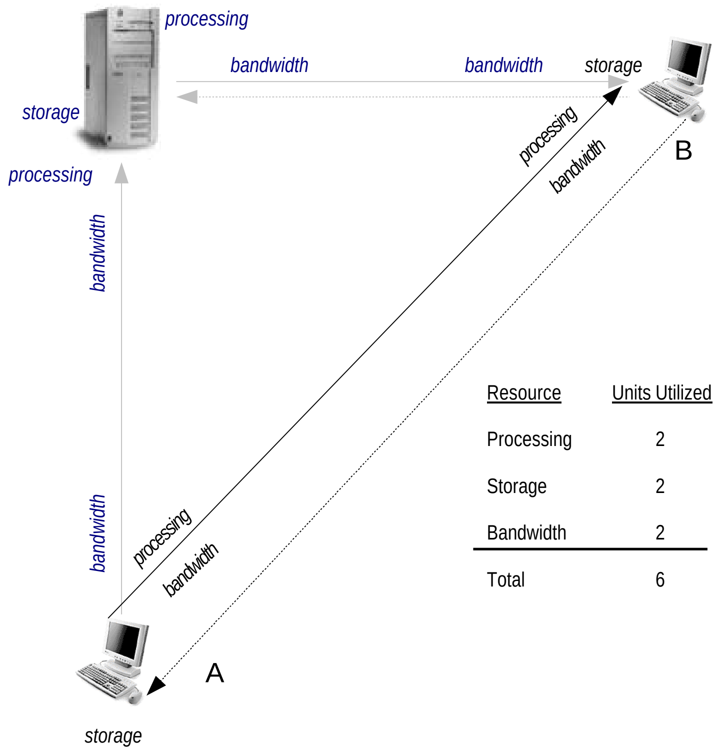
In a dC-S model, since the server is removed from the client-server-client (C-S-C) transaction scheme, client devices can communicate directly with other client devices (C-C).

In a decentralized model, A utilizes bandwidth, storage, and processing to send a file. To capture the file, B simply mirrors this resource utilization.

---

9 In protocol design, perfection has been reached not when there is nothing left to add, but when there is nothing left to take away.

Exhibit 27. What's Right About Decentralization



Resource	Units Utilized
Processing	2
Storage	2
Bandwidth	2
Total	6

Source: Bear, Stearns & Co. Inc.

In a decentralized system, the router becomes the sole intermediary since we push intelligence and resources (of the server) outward to edge devices.

Of course, not all applications can efficiently use a decentralized model. But in this particular file-transfer example, using a decentralized peer-to-peer model saves five system-resource units without compromising upload/download speed.

The limitation in A's upload speed is present in whatever system the device operates; and B's download speed is the same in whatever system it operates. Both are limiting factors. But, in a decentralized system, no additional steps are introduced. In a cC-S system, even SSL (Secure Socket Layer) is done C-S-C, when it could be done C-C. By more effectively utilizing device assets, dC-S systems enjoy not only performance improvements but cost advantages as well.

The Napster network demonstrates the advantages of a dC-S architecture and the cost disadvantages of a cC-S system. We estimate that if Napster were built on a cC-S architecture, for the number of songs "on" its network at its peak, Napster would have had to purchase over 5,000 NetApp F840 Enterprise Filers to host all the songs shared among its users (before the Ninth Circuit's injunction in February 2001).

Exhibit 28. If Napster Were a Centralized Network

Number of Users (millions)		65,000,000	(a)
Number of Songs per User		171	(b)
Number of Bytes per Song		3,000,000	(c)
Number of TB Total		33,345	
Number of TB in NetApp F840		6	
Number of F840s required		5,558	
Price per F840	$	120,000	(d)
Total Cost of Hosting Content	**$**	**666,900,000**	

(a) At its peak, roughly 65 million users were on Napster.

(b) Studies indicated about 171 songs per person before the injunction in February.

(c) We assume the average 3-minute song ripped at 128 Kbps is roughly 3MB.

(d) A 6TB NetApp F840 filer was originally priced at $117,200 when it was launched last year.

Note: We acknowledge the redundancies on Napster that would not exist in a cC-S architecture. We merely show the absolute costs associated with offering service in such a network.

Source: Bear, Stearns & Co. Inc.

Webnoize estimates that Napster users downloaded close to three billion songs in January 2001. How much bandwidth cost would that have entailed for Napster in a cC-S service offering?

Exhibit 29. How Much It Would Cost Napster to Serve Content in a Centralized Environment

Number of Songs Downloaded		3,000,000,000	
Number of MB per Song		3	(a)
Total MB Downloaded		9,000,000,000	
x 8 bits per byte		8	
Total Mbits Downloaded		72,000,000,000	(b)
per Day		2,400,000,000	(c)
per Hour		100,000,000	(c)
per Minute		1,666,667	(c)
per Second (Mbps)		27,778	(d)
OC-12 (Mbps)		622	(e)
Number of OC-12 Lines Required		45	(f)
Price per OC-12 Line per Month	$	150,000	(g)
Total Cost of Bandwidth per Month	**$**	**6,698,821**	

(a) We assume a typical 3-minute song ripped at 128 Kbps is roughly 3MB.

(b) We convert MB (Megabytes) to Mbits (Megabits) in order to convert data storage to data throughput.

(c) We divide the total number of Mbits by (30 days X 24 hours X 60 minutes X 60 seconds = the total number of seconds in a month) to arrive at a Megabits per second (throughput) value.

(d) The minimum throughput rate required to download all 3 billion files over a month is 16,204 Mbps.

(e) An OC-12 offers 622 Mbps throughput capacity.

(f) In order to download 16,204 Mbps, Napster would require the equivalent of 26 OC-12 lines.

(g) In January 2001, an OC-12 would have cost roughly $150,000 per month.

Source: Bear, Stearns & Co. Inc.

The above example illustrates the power of a decentralized file-transfer system: Napster users' client devices act as servers, saving the company those internal capital investments.

Exhibit 30. How Much Is Sunk into the Napster Installed Base

Number of Users (millions)		65,000,000	(a)
Number of PCs		65,000,000	
Cost per PC	$	1,500	(b)
Total Hardware Cost of Network	**$**	**97,500,000,000**	(c)
Cost of ISP per Month	$	23	(d)
Narrowband (90%)	$	20	
Broadband (10%)	$	50	
Total Network Cost	**$**	**1,495,000,000**	(e)

(a) At its peak, Napster had an installed base of 65 million. We assume each user had a PC.

(b) Assume the cost of a standard PC is $1,500.

(c) Cost of hardware shared among Napster users.

(d) Cost of ISP service per user. While we believe a great number of users are in enterprise and academic institutions, we believe the cost of the network is ultimately absorbed somewhere along the network, and because access is cheapest at the edge, we use end-user access costs as a proxy.

(e) Cost of bandwidth shared among Napster users per month.

Source: Bear, Stearns & Co. Inc.

System resource utilization and management are optimized and maximized in a dC-S network. This is possible in a dC-S environment because unlike a cC-S network, a dC-S network allows the resources on a client to be fully utilized. Because devices can disintermediate servers in dC-S networks, these networks are really “client-client” networks.

Having provided this background on some of the disadvantages of a cC-S architecture and some of the advantages of a dC-S architecture, we now expand on decentralization and Internet 3.0.

Decentralized Computing Defined in Internet 3.0

"The enduring magic of ethernets stems from the law of the microcosm, favoring distributed terminals over centralized hierarchies, peer networks of PCs over mainframe pyramids. The microcosm's relentless price/performance gains on chips have endowed Metcalfe's peer-to-peer scheme with ever more powerful peers, at ever lower prices. Medium-independent from the outset, the Metcalfe systems do not require central switching. In an ethernet system the intelligence is entirely in the terminals, not in the network itself, and most of the bandwidth is local (where some 80 percent of traffic resides) . . . Intelligence in terminals is a substitute for intelligence in networks; switching and routing functions migrate from the center of the Web to the increasingly powerful computers on its fringe."

– George Gilder

Defining Decentralized Computing

What is decentralized computing? What are the hallmarks of decentralized computing? And how do we know it's not a variant of what's been around since the browser? What else do we need for decentralized computing to become pervasive?

All of these questions need to be answered for us to understand decentralization's impact on the systems around us. Just as digitization enabled us to decouple information from physical media, decentralization frees resources from the physical media on which they reside and the geographic isolation that imprisons them.

The overbearing insistence of the browser-to-information marriage on the Internet has crippled Internet application development in many respects. Until now, the infrastructure buildout over the last five years has been driven primarily by the Web, and not by the Internet at large.

We define *decentralized computing in Internet 3.0* as computing where network architecture shifts or extends the focus to the nodes along the network edge (where content origination, consumption and control, and resource allocation and management reside) and where resources are virtualized in a distributed manner.

The Rules of Decentralization in Internet 3.0

We believe there are five fundamental rules of decentralization in Internet 3.0.

Rule 1. The Browser Is Not Alone

Corollary: DNS is no longer the only addressing system

The Web browser is the one piece of software most associated with the Web. It was created by Tim Berners-Lee, developed by Mosaic, marketed by Netscape, and ultimately dominated by Microsoft. Amidst all of the commotion, the browser has changed relatively little. The browser has always been presentation-layer software with little functionality other than to display HTML and associated graphics. It is, in a number of ways, extremely elegant. It is universally accessible and easy to use. It is critical for "surfing" on the Web, and it is free. Therefore, the browser meets all four of our criteria for the adoption of a technology: utility, ease of use, affordability, and necessity.

The browser, however, is a limiting factor in fluid communications and end-to-end connectivity. With the browser, there are too many sessions. The browser is DNS-based and as such, connectivity is staggered. Knowing a domain name tells us nothing about the location of a particular device, since a domain name, as the phrase suggests, points to a domain (of machines or devices) and not necessarily to a specific machine. Being able to locate an exact device on the Internet offers certain advantages which we discuss below.

DNS has been good to us and continues to serve us well, but it will not be the only name system for the Web. Just to illustrate the power of alternative name systems, ICQ-AIM, which has been around for less than four years, has over 100 million registered identifiers. Napster, within 20 months, had 65 million registry entries. The DNS, now entering its fifteenth birthday, has about 35 million entries. It's not even close.

Rule 2. Client or Server or Both — Does It Really Matter?

The PC is our primary communications device for the Web. It is not the central unit of the Web; the Web server is. Why shouldn't the PC — with as much processing power and disk as servers three years ago — be utilized more effectively? In Internet 3.0, the PC (along with other devices) is both a client and a server because it wants to be and it can. In addition, we believe client devices will grow thicker, driven by the three lows of network dynamics.

Rule 3. XML Rules

Standards like XML are dominating the Internet right now. With XML, and associated protocols (like SOAP — Simple Object Access Protocol: see Appendix B), the Web is starting to look more like an organized space than a mess. This was the vision of Berners-Lee when he articulated the concept of a "Semantic Web," a Web in which the underlying language codifies every object on the network, where data is defined, logical, and operates within a grammar that empowers us and machines to organize data on the Internet more coherently. The vision of the Semantic Web is to have data objects describe themselves and to enable devices and their users to do more with this descriptive data.

Rule 4. Virtualization

We defined "virtualization" earlier as the method by which computing and network systems resources, which may be dispersed and fragmented, are harnessed and aggregated to create a virtual pool of resources that can be used on demand and managed seamlessly. Virtualization of systems resources is being researched by public and private companies alike. Virtualization has had and will have a strong impact on storage, processing, and databases.

Rule 5. Metcalfe Maximization Principle

We refer to the way devices have inertia to connect with other devices as the Metcalfe Maximization Principle. In Internet 3.0, we decentralize, we explicitly encourage devices to communicate directly with peers. Direct connections that may otherwise have been unnecessary or impossible are permitted and encouraged in

Internet 3.0. This device-to-device connectivity allows the increased velocity of data and transactions to pervade the entire network.

Metcalfe's Law will cease to be relevant when all devices on networks have direct access to each other, resources are utilizable from anywhere, and all data exhibits perfect fluidity. This is impossible right now, given the isolation of databases, devices, and resources. However, we can come close. Metcalfe's Law takes one step further towards its own obsolescence, as distribution and decentralization of data and resources compel the actualization of the distributed, connected network that Tim Berners-Lee envisioned when he developed the World Wide Web.

ADVANTAGES OF DECENTRALIZATION IN INTERNET 3.0

One of the more powerful ideas of decentralization in Internet 3.0 is the emphasis on the primacy of users' ownership and control of data and resources. The only two assets that matter in a system, we posit, are the data and the system resources associated with the user. Decentralization liberates the user from the grip of central management, especially when it is advantageous to both central management and the user.

Decentralization has a significant impact on the resources available to us, how we utilize these resources, and what we can accomplish with these freed resources. As we saw in the cC-S network, centralization introduces a host of bottlenecks to efficient communications. Decentralization frees data to move as users deem most appropriate, and emancipates system resources for maximum utilization.

- **Creators Should Control Creations.** One of the problems of cC-S environments is the central management and control of network resources. In cC-S, a central administrator handles network permissions. Because the network administrator is constrained by her resources (namely time), she cannot tend to the specific needs of each user on the network. Therefore, network settings and permissions are broad and rarely tailored specifically to each user's desires or needs. In a decentralized environment, broad permissions can be set, at which point, the end-user can set more granular permissions, such as document management and access-control permissions. Within limits, this way of administering control rights can free a sizable portion of end-users' data to move about more freely.

- **Nodes Go Down, But the Network Does Not.** One of the great things about the Web is that should a site like CNN.com go down from, say, a denial-of-service attack, one can access news from an unlimited number of sites. That's because the Web is a decentralized virtual database, a supercluster of independent networks. As such, should any sub-network fail, we can route around it to find other networks. This was the original vision of the packet-switched network. The most scalable and successful network to date — the Internet — is constructed this way. The Internet is a dC-S network: Individual nodes can fail, but the whole remains supported by the distributed fabric of the network. And like other dC-S systems, the Web has a registry server (DNS) and broad permissions and rules (TCP/IP, HTTP, HTML, IPv4), which allow devices, users, and systems to evolve freely. In a cC-S network, recall, the failure of a single node can bring down the entire network. Networks do not have to be constructed this way, and clearly, the most scalable one is not.

- **Locality Is Born.** cC-S systems are geographically isolated. Enterprises design NOCs (Network Operations Centers) to serve a geographically isolated enterprise environment, often because distributing the network is extremely difficult technologically. As enterprises expand, the network does not always keep pace with the distributed nature of the enterprise. Therefore, accessing data from a satellite office, the road, or home is almost always painfully slow and often impossible. Geography is often the single greatest limiting factor in accessing data.

  Networks have already begun to use mirroring and replication techniques to make data assets available across geographies. We believe we are on the cusp of an era where data is as fluid and data assets as connected as the networks that link them.

- **Devices Are Already Installed.** If we can develop software to connect devices to one another (which is ultimately the point of a dC-S architecture), we can use the latent resources on the devices already deployed across enterprises and networks.

- **Hardware and Software Ain't Cheap.** A cC-S system is expensive. A cC-S system requires client and server hardware and software — that is, a cC-S system requires client hardware, client software, server hardware, and server software, as well as a multitude of storage devices, software, and maintenance.

  In a dC-S architecture, the network makes generous use of the resources on edge devices. Each device has processing, memory, and storage assets. In dC-S, each of these resources is utilized more effectively.

  Again, the problem with a cC-S system is the inability of devices to communicate directly with one another. The intermediation by servers for virtually every function creates fragmented magnets and increases costs. Instead of subscribing to Metcalfe's Law, the cC-S topology destroys potential connections. This is precisely what has happened in Internet 2.0 and what Internet 3.0 will undo.

Categories of Decentralization in Internet 3.0

We believe there are three fundamental computing resources underlying networks: processing, storage, and bandwidth. Applications run on top of these resources. Below, we categorize the three resource segments and highlight decentralized collaboration as an application running on top of them. We believe decentralization technologies will be pervasive over time; the first business models to have emerged are distributed processing, distributed storage services, distributed network services, and decentralized collaboration.

DISTRIBUTED PROCESSING

Decades old, distributed processing was only recently popularized by SETI@home, the Search for Extra-Terrestrial Intelligence project headed by Dr. David Andersen of UC Berkeley. We believe the near-term market opportunity for distributed processing alone is small relative to other distributed applications, but as a platform distributed processing could become a centerpiece to a variety of applications. In fact, we think the software provided by many distributed processing vendors could become an integral component of enterprise middleware for time-critical applications. Of all of the categories of decentralization, we are particularly fond of distributed processing because it treats processing as a resource to be virtualized and distributed globally on the Internet. It implies that processing is no longer isolated to a chip or large supercomputer clusters or server farms. Distributed processing, harnessable by good technology, could potentially commoditize processing even further.

We highlight the advantages and disadvantages of distributed processing in Internet 3.0. We also highlight the market opportunity, the players most likely to benefit, and the impact on the computing landscape.

DISTRIBUTED STORAGE SERVICES

Distributing storage has been a mantra for some time. IBM has offered distributed storage management through its ADSTAR Distributed Storage Manager product family, now offered within Tivoli Storage Management Solutions. In fact, the dominant architecture for distributed storage, Fibre Channel Storage Area Networks, suffers from two major limitations: distance (up to 10 kilometers) and clustering (up to 127 nodes). Despite ever larger, nimbler, and more sophisticated storage technologies, storage services have been less about sophisticated distributed architectures and more about leaking centralized systems.

We believe next-generation storage technologies that have been incubating will abolish distance limitations and resource-utilization constraints. What began as a slow stretching of the storage fabric is now quickly becoming a race to obliterate geographic restrictions. The next generation of distributed storage services is being marshaled by companies with unique technology assets that make distributing and decentralizing data assets as simple as turning on a switch. This will involve unique caching, global load balancing, and rich metadata technologies.

We believe companies like Zambeel and Scale Eight are the next breed of distributed storage companies, offering unique and extremely complex technologies to virtualize, distribute, and decentralize storage.

DISTRIBUTED NETWORK SERVICES

Network services accelerate the delivery of data, propagate data more efficiently, or lower the cost of operating existing network services. Moving data across geographies, offering geographic perspective, adding network intelligence that comes from diverse perspectives, and offering speed, efficiency, and cost savings — these are the advantages of distributed network services. Already, Akamai and its customers have benefited tremendously by taking advantage of distributed network architectures.

We highlight some of the interesting technologies and services on the market and those being developed. We believe the explosion of network technologies and the plenitude of bandwidth will enable network services to become superdistributed, ultimately helping to minimize the dependency on server-centric solutions by employing native assets on the network to displace centralized machines.

Public companies Inktomi and Akamai have been offering higher-level network technologies and services for several years on the premise that companies are willing to spend on technologies and services that solve network bottlenecks while offering high ROI. Private companies are now following in these public companies' footsteps, only this time the consideration includes resource utilization of existing assets. Whereas companies like F5 and CacheFlow offer purchasable technologies that improve network performance and introduce cost savings, emerging companies are introducing technologies that take advantage of deployed assets (like servers and PCs) by clustering them into a virtual pool of resources on the LAN, through the MAN, and across the WAN.

DECENTRALIZED COLLABORATION

Collaboration software has not seen much innovation since Microsoft introduced Outlook in 1997. That was 14 years after the introduction of Lotus Notes.

Of the four categories in Internet 3.0, decentralized collaboration is going to take the longest to generate momentum, because it requires us to change our behavior and learn new rules and interfaces. While distributed processing, distributed storage, and distributed network services work under the hood, decentralized collaboration operates on the dash. Nevertheless, decentralized collaboration has the potential to impact every aspect of workflow and communications. We believe once a network has a critical mass of users with decentralized collaboration software, the productivity gains will be too large to ignore.

We believe the next wave of collaboration software has already begun to gain some traction. Multinationals and workgroup-oriented enterprises are clearly the first adopters of decentralized collaboration technologies and services.

We detail the major players and highlight early wins and the potential market opportunity for decentralized collaboration systems.

Distributed Processing: Processing as a Resource

In distributed processing (DP), a large data aggregate is broken into blocks, delivered to individual processors, processed, returned, and reassembled.

Why isn't every processing task distributable in this fashion? Why has there been a need for supercomputers or high-end servers?

The answer is that certain data must be processed serially — that is, each data-block (subunit) must be processed before proceeding onto the next subunit. Other types of data can be processed in parallel and treated like jigsaw puzzles — broken, processed piecemeal, and reassembled. DP is currently focused on the latter.

DP has been around for over three decades. In fact, Xerox PARC scientists John Shock and Jon Hupp created a program called a "worm" that crawled into machines, replicating itself and performing simple computational tasks. In another case, Apple scientist Richard Crandall used idle processing cycles on machines to do large factoring and prime number search.

In both cases, DP was done on a LAN. That's because on a LAN, protocols, languages, and machines are generally uniform enough to minimize variables.

The rise of the Internet as we know it has changed everything. Accessing processing cycles across large geographies has become a simple matter of sending data blocks over IP. Initially, efforts focused on encryption-breaking tasks. Distributed.net, one of the earliest projects focused on utilizing DP to attack encryption, has prevailed in many RSA Security challenges. SETI@home, made famous by the movie *Contact*, has successfully registered over three million participants in its computing grid. Using users' PCs, SETI@home has been able to search for signs of extra-terrestrial life by processing data gathered from radio satellites. We note that the single spark for distributed computing on the Internet was the introduction of a mass-deployed software client by SETI@home.

BENEFITS

The major benefit of DP is virtually unlimited processing resources. Most databases cannot be fragmented for DP; those that can enjoy substantial potential cost savings. For example, ASCI White, IBM's supercomputer, offers 10 TFLOPS[10] at a cost exceeding $100 million. Comparatively, a virtual cluster of 100,000 installed PCs outfitted with Intel Pentium IIIs offers similar performance. The ASCI White from IBM is extremely expensive, floorspace intensive, and electricity consumptive. What's more, with Moore's Law still guiding processor development, fast machines become slow quickly. In the spirit of Moore's Law, we compare a single Intel P4 introduced in 2000 to an SGI Cray from 1989, and IBM's ASCI White to a virtual cluster of 100,000 PCs.

[10] Floating Operations per Second. A FLOP is a measure of the number of operations executed per second. A TFLOP (TeraFLOP) is a trillion FLOPs.

Exhibit 31. Comparing Processing Metrics and Costs

	1989	2001		2001	2001
Brand	SGI Cray Y-MP8/4128	Intel Pentium 4	Brand	IBM ASCI White	Intel Pentium 4
CPUs	4x 166 MHz	1.5 GHz	CPU Capacity	SP Power3-II MHz	100,000 1.5 GHz
Maximum CPUs	4x	1x	Maximum CPUs	8,192x	100,000x
Maximum RAM	1 GB	1 GB +	Total RAM	6 TB	12.8 TB (a)
Disk		50 GBs	Maximum Disk	160 TB	5 PB (b)
GFLOPS	~ 1.3	~ 1.8	TFLOPS	~ 12.3	~ 600 (c)
Power Usage		100 watts	Power Usage	1.2 Mwatts ~ 1,000 homes $1,584 per day	240 Mwatts (d) across 100K homes $36,000 total per day $0.36 per PC
Weight	> 5,000 lbs	< 1 oz	Weight	> 106 tons	NM
Cost	~ $14,000,000	< $1,200	Cost to Access	~ $110,000,000	Minimal

Note: We compare a single Intel Pentium 4 to an SGI, and 100,000 Pentium 4s to IBM ASCI White.
(a) Assume an average PC has 128 MB of RAM.
(b) High-end PCs now come with a 50 GB hard drive. 50 GB x 100,000 = 5 PB.
(c) Intel P4 has 4 FLOPs per cycle; 1.5 GHz x 4 = 6 GFLOPs x 100,000 PCs = 600 TFLOPs.
(d) Assume an active PC operates at 100 watts of electricity. Left on all day, 100 x 24 = 2,400 watts = 2.4 Kw hours x $0.15/Kwhour = $0.36 per day.

Source: *Electronic News*; Bear, Stearns & Co. Inc.

The lesson of a DP grid like the virtual cluster of 100,000 PCs highlighted in Exhibit 31 is: Assets are distributed and so are liabilities.

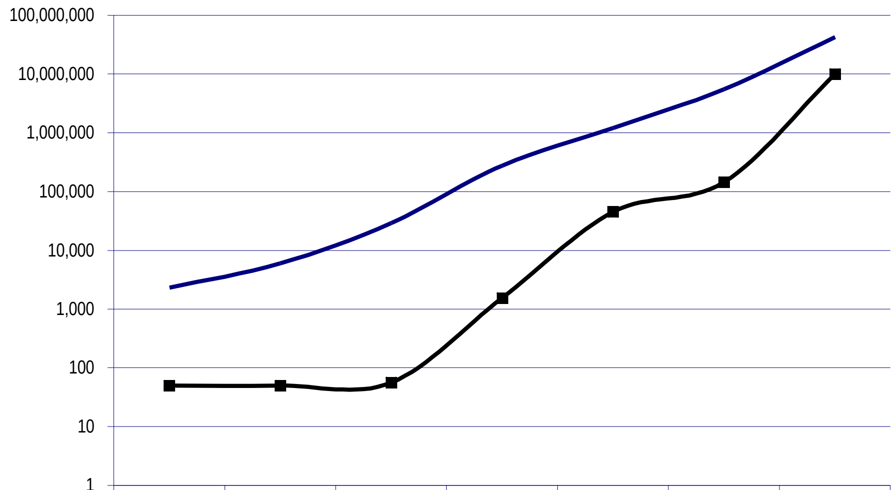
In a DP grid, the cost of the processing assets is absorbed by users (i.e., the PC owners). That is the advantage. The cost of 100,000 Intel P4-based PCs is roughly $200,000,000. Adding electricity consumption, the cost of owning and running the 100,000 PC cluster for a year would be $213 million. And that is just the point. Distributing assets also distributes liabilities (the cost of owning and running the cluster). While the cost, space, and electricity to power a large processing grid is spread across the entire network, distribution also introduces exogenous concerns, like security, reliability, and scalability. *The rewards of distributing processing tasks must outweigh the risks associated with distributing processing assets, and therefore, making a system of distributed assets and liabilities profitable is the business of distributed processing vendors.*

As we have indicated, DP networks are able to target particular markets. While supercomputers can process any kind of data, for now, DP networks are limited to simple parallel processing tasks, tasks that can be fragmented, processed, and then re-assembled.

Ideally, the processing task is one where the compute-to-data ratio is high. When the compute-to-data ratio is high, the amount of processing required of a block is meaningful compared to the amount of bandwidth required to deliver that block. That is, DP (at least across the Internet) would be inefficient if a 3MB block could be processed on a typical PC in five seconds. The compute-to-data ratio would be extremely low, and the cost of bandwidth to deliver the data block would be too high to make the entire project worth distributing.

Historically, the compute-to-data ratio has been too low for DP on the Internet to be possible. Bandwidth capacity has risen, and prices have fallen; processing power has continued to increase. If we revisit our Moore-Gilder graph, we see that only recently have the curves of Moore's Law and Gilder's Law truly come close to converging. In the LAN, the bandwidth capacity has been available. Recall, Intel has employed intranet-based DP for years.

Exhibit 32. Moore and Gilder



	1970	1975	1980	1985	1990	1995	2000
Transistors	2,300	6,000	29,000	275,000	1,200,000	5,500,000	42,000,000
Bandwidth (Kbps)	50	50	56	1,544	45,000	145,000	10,000,000

Source: Bear, Stearns & Co. Inc.

As bandwidth capacity continues to grow, the compute-to-data ratio will continue to grow. That is, when bandwidth becomes so abundant that the cost of transmitting large blocks of data is marginal, processing will be the more valuable resource, and processing resources will demand an even greater premium. In such an environment, it makes even more sense for a system to take advantage of installed processing resources distributed across networks.

One of the big attractions of distributed computing is leveraging untapped processing resources on a network for enterprise-wide tasks. Intel, for example, has been using DP (NetBatch) in-house for over 10 years. To paraphrase Intel CTO Patrick Gelsinger, the company has increased the utilization of its aggregate computing capacity from 35% to over 80%. On a recent chip project, Intel was able to accelerate its validation process by eight weeks. Gelsinger concluded that NetBatch has saved Intel more than $500 million.

Most PCs sit completely idle at least 66% of the time.[11] Even when the PC is in use, the processor is not at peak utilization unless the user is doing processor-intensive tasks — like rendering a photo, calculating a dense spreadsheet, or participating in a multi-player game.

[11] We calculate this by using eight hours as an average workday. eight hours/24 hours = 33%. (We assume enterprise desktops are left on 24 hours per day.) If we include lunch, bathroom, and eBay breaks, we believe the capacity utilization could be as low as 10%.

Exhibit 33. Desktop Processor Utilization

	Processor Capacity (MHz)	Used	Capacity/Used
Desktop	500	25%	50%

	$/Processor	$/Used	$/Used Processor	$/Unused
Processor/Desktop	$ 350	$ 88	$ 44	$ 306
x1000	$ 350,000	$ 87,500	$ 43,750	$ 306,250

Source: Bear, Stearns & Co. Inc.

We believe up to 90% of the cost of a microprocessor goes unused on a PC. If we aggregate the number of PCs across an enterprise, we quickly realize that the single most expensive component of a computer is at 10% capacity utilization.

Once a DP network has a captive user base, other value-added services can be offered using the network. For example, load-testing and performance measurement are two obvious opportunities. We believe there is real demand for Web site performance measurement from end-users' perspectives. We believe DP networks could offer interesting storage services, particularly in the enterprise. We think this kind of service is much further away and would most likely require the participation of PC OEMs.

Exhibit 34. Market Opportunity for Distributed Processing

Bioinformatics and Life Sciences	Genomic research and pharmaceuticals industry could benefit significantly from utilizing distributed processing solutions, as greater processing could quicken the pace of research.
Environmental/Space Sciences	Oil discovery, large image translation and processing.
Financial	Financial services concerns require processor-intensive systems in order to run Monte-Carlo simulations and real-time risk analysis. The market opportunity could be meaningful. This would be a software sale because the system would be in-house.
Entertainment	Graphics rendering is processor and storage intensive. By parceling tasks across existing resources, it is possible to cut costs and speed up projects.
Storage	Once the device is captive, the disk and processor are potentially resources. Distributing storage is a potential application. We believe the enterprise could be the first to adopt distributed storage on PCs; non-critical data (such as archiving) will likely be the first target.
File Sharing	If a distributed processing client becomes ubiquitous on a network, a natural glue is spread across the network. Sharing files and collaborating with others on that network can be a natural extension.

Source: United Devices; Bear, Stearns & Co. Inc.

RISKS

Risks associated with DP, in real terms, are few. Data is always encrypted and the entire process of delivering and collecting blocks of data is automated. In the intranet, risks are mitigated even further primarily because of the uniformity of devices, software, and platforms. On the Internet, matters become a bit trickier, since there are multiple protocols, devices, platforms, and uncontrollable exogenous user variables. Early projects like SETI@home experienced falsified and damaged data.

Nevertheless, much of the concern regarding DP is groundless, in our view. Napster is significantly more invasive than DP projects, and over 65 million users voluntarily participated in Napster. Because all of the information in DP is encrypted, most of the security risk is measured. In addition, most of the DP vendors have a boxing mechanism in place so the processing application does not interfere with or hamper the performance of other applications running on the client device. We see the most significant issue facing DP not as adoption, but as the size of the opportunity.

Exhibit 35. Distributed Processing Companies

Company	Funding		Investors	Customers
DataSynapse	$5.25 million	Series A	Rare Ventures, Neocarta Ventures, The NYC Investment Fund, Silicon Alley Venture Partners, Wand Partners	First Union
Entropia	$23 million	Series B	Mission Ventures, RRE Ventures	Envive, SolidSpeed
Parabon	$6.5 million	Seed	Undisclosed	
Porivo	$2 million	Seed	Undisclosed	
United Devices	$13 million	Series A	SOFTBANK Capital, Oak Investment Partners	Exodus, iArchives, NFCR (sponsored by Intel)

Source: Bear, Stearns & Co. Inc.

Below, we discuss in more detail the companies we believe have the best chance of succeeding in the space.

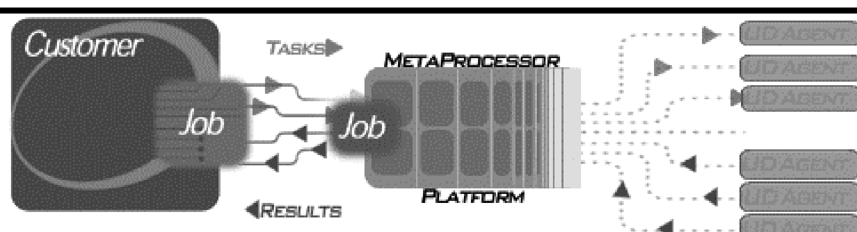
UNITED DEVICES

The Operating System of Distributed Processing

United Devices (UD) is the brainchild of several ex-Microsoft employees. Founded in 1999, UD established itself as a front-runner in the DP market by announcing its first customer, Exodus, in October 2000. In addition, UD recently launched a search-for-cancer project with The National Foundation for Cancer Research (NFCR) Centre for Drug Discovery in the Department of Chemistry at the University of Oxford, sponsored by Intel. We believe these two milestones give UD an advantage.

UD's network is architected like other DP networks. Large data objects are broken into smaller data blocks, parsed to processing agents where the smaller blocks are processed, and returned to a central server for reassembly.

Exhibit 36. United Devices Network Flow



Source: Company data.

We believe UD's advantages lie in its already significant captive userbase. As of April 27, 2001, UD's network boasted 318,000-plus users with 443,000-plus devices that have contributed 39,834,919 hours of CPU time, up from 286,658 users with 368,714 devices that had contributed 24,062,118 hours of CPU time the week prior.

UD has been focused on establishing itself as the largest and most pervasive network on the Internet. We believe this vision is what clearly separates UD from others: While many competitors are focusing primarily on signing enterprise customers, UD has been focused on growing its processing network to be able to offer the vast processing resources on it to potential customers, to tie into their enterprise networks or to be used on a standalone basis. However, we also regard this approach as UD's Achilles heel: Supply of processing resources requires demand to be useful, and thus far demand for Internet-based DP has not materialized.

Products and Services

UD's products and services center around the company's MetaProcessor Platform. Like most DP companies, UD offers both services and software.

- **The MetaProcessor Platform.** The MetaProcessor platform is UD's enterprise software offering. The MetaProcessor Platform enables enterprises to weave internal resources into a processing platform.

- **Global MetaProcessor Service.** UD also offers customers like Exodus the ability to utilize the globally distributed installed base of users running UD's client.

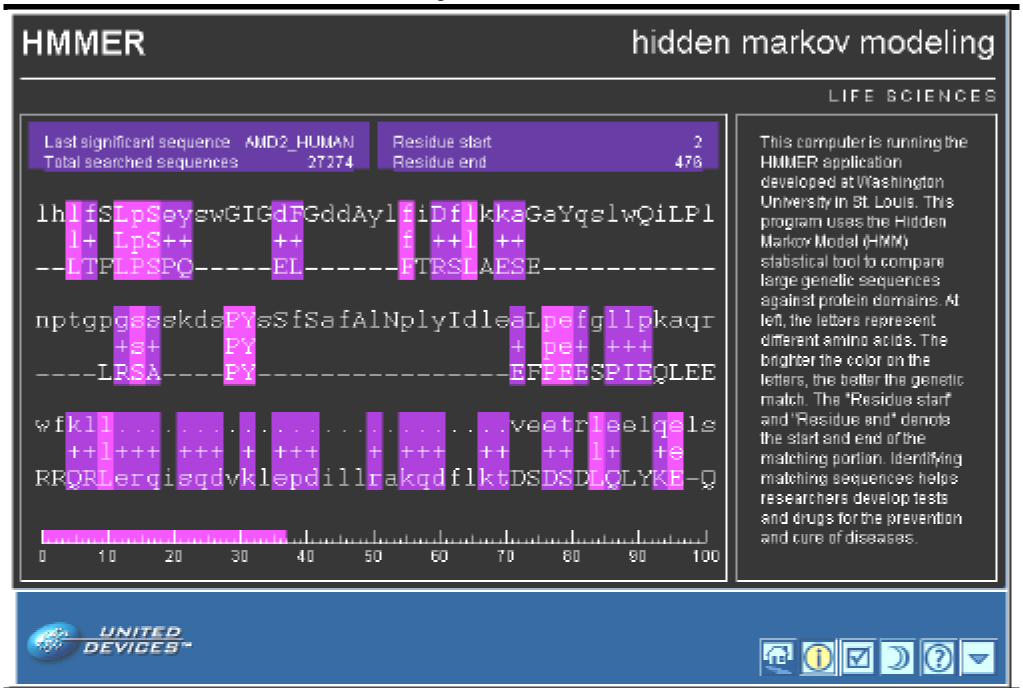
The company has made the MetaProcessor SDK available. UD has several bioinformatics applications that are already ported to the MetaProcessor platform, including:

- **HMMERsearch and HMMERpfam.** Performs profile hidden Markov model based matching of protein and genetic sequences;

- **BLAST.** Performs protein and genetic sequence matching via the Basic Local Alignment Search Tools (family includes: BLASTX, BLASTN, TBLASTX, PBLAST, TBLASTN);

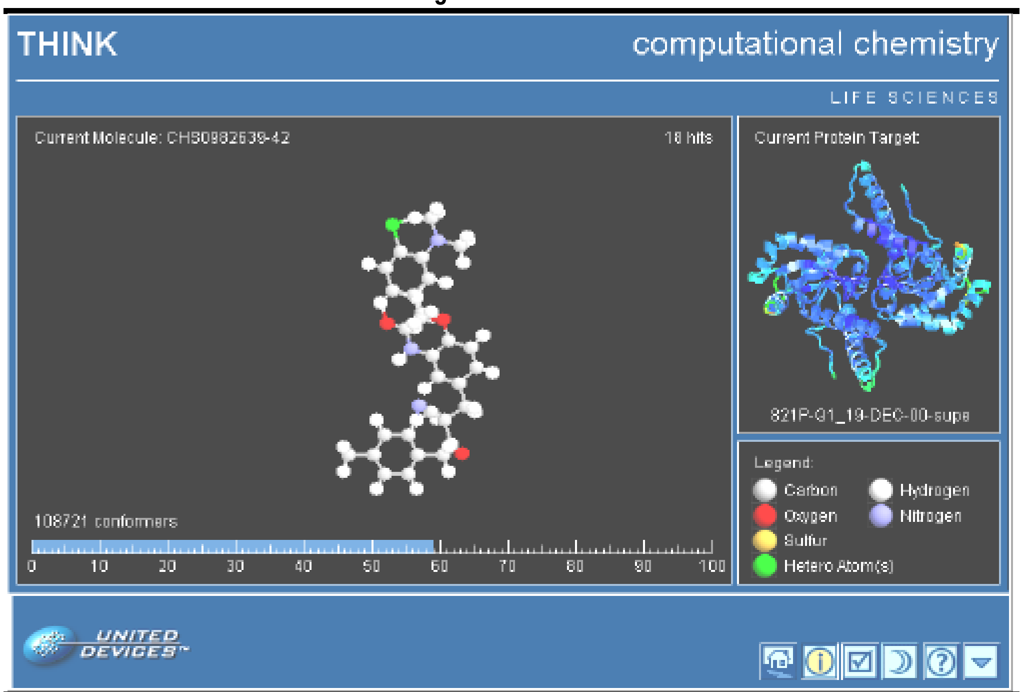
- **SIM4.** Performs DNA sequence matching.

Exhibit 37. United Devices Client Running HMMER



Source: United Devices; Bear, Stearns & Co. Inc.

Exhibit 38. United Devices Client Running THINK



Source: United Devices; Bear, Stearns & Co. Inc.

Customers

- **Exodus, Inc.** On November 6, 2000, United Devices announced an agreement with Exodus Communications whereby Exodus Performance Labs would use UD's Global MetaProcessor platform for distributed Web site testing.

Exodus acquired Web site monitoring company Service Metrics in 1999. However, like Keynote, Service Metrics only offers testing and measurement from POPs, which cannot measure performance of Web sites from an end-user's perspective. While testing Web sites and measuring performance from distributed POPs offers advantages, it does not offer a view from the actual users' perspective. By partnering with UD, Exodus Performance Labs can now offer its customers load testing of Web sites from actual users' computers.

The reality of simulated load/stress testing is that it is expensive and artificial. Load testing requires server farms and returns results from a simulation environment, and does not capture the real experience of an end-user.

UD's deal with Exodus is exclusive and revenue generating.

- **National Foundation for Cancer Research (NFCR) Sponsored by Intel.** UD and the NFCR Centre for Drug Discovery in the Department of Chemistry at the University of Oxford, England, have teamed up in the search for new drugs. As a first step to find a cure, the NFCR Centre must evaluate the cancer fighting potential of hundreds of millions of individual molecules. It is anticipated that about 24 million hours of computer time will be required to accomplish this task — considered to be the largest computational chemistry project ever undertaken.

  We believe Intel's sponsorship of UD on this project speaks volumes about UD's technology and clout, and ties with Intel.

- **iArchives.** iArchives is a government-sponsored service that is indexing everything from archived microfilm and newspapers to government records and genealogical annals. Each of these documents, once digitized, is searchable and accessible via a browser. UD is offering processing services to iArchives to digitize all of the files.

Target Market

While focused primarily on bio-informatics/life sciences and network services (load/stress testing and performance measurements), UD is targeting the gamut of heavy computing and distributed computing opportunities.

We believe the low-hanging fruit is in bio-informatics, financial services, and load-testing and performance measurement. As the technology matures, we believe all relevant and serviceable scientific, military, and academic research requiring heavy computing could be outsourced to networks like UD. The pot of gold lies in enterprises that require heavy processing for data mining, processing-intensive applications, and the like.

Management

UD management (Ed Hubbard, CEO; Dr. David Andersen, CTO and Founder of SETI@home; Jikku Venkat, VP of Engineering; and Lance Hay, CFO) has a long history with Dell, Intel, Microsoft, and IBM. We believe this is one of the company's most valuable assets. Industry father Dr. Dave Andersen brings experience and credibility.

ENTROPIA

When Gridlock Speeds Up Everything

Entropia has become famous within the DP community for its sponsorship of The Great Internet Mersenne Prime Search (GIMPS), a project that searches for a particular type of prime numbers.[12]

Like UD, Entropia has won the backing of several leading venture capital funds. Entropia is the first DP company to close a Series B round of financing, raising $23 million in its second round, and $30 million to date.

We believe Entropia and UD are the front-runners in the DP race and that both companies offer very similar features.

Exhibit 39. Entropia Network Flow

Source: Company data.

Products & Services

Entropia's products and services offering include:

- **Entropia 2000 Internet Computing Services.** Entropia offers DP services on an outsourced basis with its 2000 Internet Computing Services.

- **Entropia Application Porting & Integration.** Entropia offers implementation and deployment services to port applications across its grid.

- **Entropia 2000 Enterprise Server.** Entropia offers Entropia 2000 Enterprise Server, a distributed computing system for in-house DP implementations.

[12] A prime number is an integer that is divisible only by itself and one; the first primes are 2, 3, 5, 7, and 11. A Mersenne prime is a prime of the form 2^P-1. The first Mersenne primes are 3 = 2^2-1, 7=2^3-1, 31=2^6-1, and 127=2^8-1. To date, only 38 Mersenne primes have been discovered.

What we find particularly compelling about Entropia is the following: While Entropia's software is written in native code (C++), the company has been focused on ensuring the stability of the device on which the client software runs so as not to interfere with other applications running on the device. Entropia calls this "binary sandbox security."[13] By guaranteeing isolation of system resource utilization and applications, Entropia is able to confine Entropia-defined tasks and maintain the integrity of native applications on the client device.

Customers

- **Envive.** In November 2000, Entropia announced Envive, a hosted e-business performance management service provider, as a customer and partner. By working with Entropia, Envive is offering performance testing and monitoring services to its customers.

- **SolidSpeed.** In February 2001, Entropia announced that SolidSpeed, a content delivery network, had selected Entropia to offer Web site measurement services to its customers.

Management

Entropia's management and founders have a deep connection with the distributed computing industry. Dr. Andrew Chien (CTO and co-founder) brings nearly two decades of expertise in supercomputing and large-scale clusters. He is the Science Applications International Corporation (SAIC) chair professor in the Department of Computer Science and Engineering at the University of California, San Diego. He has also held joint appointments with both the National Center for Supercomputing Applications (NCSA) and the National Partnership for Advanced Computational Infrastructure (NPACI), working on large-scale clusters.

Entropia also boasts Dr. Martin Stuart, Vice President of Life Sciences, on its management roster. We believe domain expertise is critical to gaining traction within the bio-informatics and financial services industries and that dedicating resources as Entropia is doing will be highly advantageous in a competitive market.

We regard the company's recent hire of Mike Cox as Vice President of Worldwide Sales and Service as an affirmation of the DP industry in general and Entropia in particular. Cox was vice president and general manager of North American Sales for Hewlett-Packard's Business Customer Organization where he was responsible for enterprise and commercial computing sales, marketing and support services for computer solutions in North America.

DATASYNAPSE

Going Vertical

Privately-held DataSynapse, a New York-based DP vendor, is focused on becoming the middle layer of the application-to-application processing matrix. As such, DataSynapse believes it most closely resembles enterprise middleware companies like TIBCO and New Era of Networks. In DataSynapse's case, enterprise

[13] Several DP vendors use Java precisely because Java offers enhanced security features.

applications are tightly coupled with processing resources distributed across the enterprise.

We believe the zealous focus across select verticals gives DataSynapse domain expertise. DataSynapse's CEO, Peter Lee, a former investment banker, has close ties to the financial services community. Lee's knowledge of the inner workings and the pain-points of financial services organizations should prove advantageous.

We also admire the company's inventiveness in its consumer business. Like its competitors, DataSynapse offers a sweepstakes for participating users. However, unlike its competitors, DataSynapse has already partnered with technology and service providers, like DSL provider Bluestar.net (a Covad company) and consumer firewall provider Zone Labs. We believe DataSynapse's dedication to providing its users with a secure experience and the company's focus on broadband users suggest further opportunities are possible between DP vendors and ISPs, and client device software and hardware manufacturers.

DataSynapse's partnership with Zone Labs and DSL providers should prove to be a leading indicator of the partnerships available to DP vendors. We believe the most likely partnership candidates are companies like Network ICE, software clients like RealNetworks and instant messengers, PC OEMs, OS vendors, and ISPs like America Online, EarthLink, and Excite@Home.

PARABON

Commoditizing Processing

Parabon Computation is a unique facet of DP. Unlike its competitors, Parabon is focused on enabling anyone to be able to tap processing resources on any device across any network. The difference between Parabon and its competitors is that Parabon is enabling anyone with processing cycles to become a potential seller of processing resources.

Frontier, Parabon's distributed processing platform, powers Pioneer, the client-side software download provided by the company. Frontier is being utilized for a variety of applications already, including exhaustive regression, bioinformatics, Monte Carlo simulation for financial modeling, and digital rendering.

Exhibit 40. Graphics Rendering Using Distributed Processing — Before/After

Source: Parabon Computation.

Parabon was the first DP company to publish an SDK.

MARKET OPPORTUNITY FOR DISTRIBUTED PROCESSING

Evaluating the size of the DP market is proving to be a tough task. We believe there are several moving parts to the market that make pinpointing the opportunity difficult.

First, because most of the services that companies like UD, Entropia, DataSynapse, and Parabon offer are not direct substitutes for high-end systems, analogizing the market for high-performance technical systems with DP opportunities is inappropriate; the market size for high-end computing systems is an upper limit proxy. We believe much of the market for DP software is limited for now to financial services and scientific research. The market for performance measurement and load-testing is difficult to measure in light of the volatility in demand across Web-based businesses that require such services. On the flip side, we cannot currently measure the potential new market opportunities that will likely emerge with the introduction of software and applications designed to tap into distributed processing resources. With this in mind, we are publishing what we believe are the realistic opportunities for DP companies.

- **High-Performance Technical Systems.** Many organizations currently purchase supercomputing time or machines. Many more purchase high-end iron to perform lengthy processing tasks.

Exhibit 41. Worldwide High-Performance Technical Systems Revenue and Market Share by Primary Application, 1999

Rank	Application	Revenue ($M)	Market Share of Total (%)	Cumulative Market Share (%)
1	Scientific research and R&D	1,476.2	26.3	26.3
2	Mechanical design and engineering analysis	841.3	15.0	41.3
3	Classified and defense	613.3	10.9	52.2
4	Biological and chemical engineering	467.0	8.3	60.5
5	Electrical design and engineering analysis	446.7	8.0	68.4
6	Geoscience and geoengineering	370.5	6.6	75.0
7	Technical management and support	269.4	4.8	79.8
8	Mechanical design and drafting	205.7	3.7	83.5
9	Economic and financial modeling	188.7	3.4	86.9
10	Simulation	177.0	3.2	90.0
11	DCC and distribution	161.3	2.9	92.9
12	Imaging	108.1	1.9	94.8
13	Other commercial	98.7	1.8	96.6
14	Software engineering	96.3	1.7	98.3
15	Other technical	49.7	0.9	99.2
16	Industrial process analysis	47.0	0.8	100.0
17	Measurement and control	0.0	0.0	100.0
	Total	5,616.9	100.0	

Source: IDC.

Exhibit 42. Worldwide High-Performance Technical Systems Revenue by Primary Application and Type, 1999 ($ in millions)

Overall Rank	Application	Capability Systems	Technical Enterprise Servers	Technical Divisional Servers	Technical Departmental Servers
1	Scientific research and R&D	480.1	172.4	452.8	370.9
2	Mechanical design and engineering analysis	102.9	64.0	310.4	363.9
3	Classified and defense	88.4	90.8	254.6	179.5
4	Biological and chemical engineering	68.1	52.6	136.4	209.9
5	Electrical design and engineering analysis	50.3	15.4	122.1	258.9
6	Geoscience and geoengineering	47.4	25.1	121.3	176.6
7	Technical management and support	9.0	30.4	55.9	174.1
8	Mechanical design and drafting	0.0	0.2	52.6	152.8
9	Economic and financial modeling	15.1	43.0	76.0	54.6
10	Simulation	9.5	7.6	125.5	34.4
11	DCC and distribution	8.6	6.8	70.1	75.8
12	Imaging	10.3	7.2	47.5	43.1
13	Other commercial	24.4	5.8	37.0	31.4
14	Software engineering	19.4	2.6	16.5	57.8
15	Other technical	0.0	2.1	4.7	42.9
16	Industrial process analysis	0.0	0.6	9.4	37.0
17	Measurement and control	0.0	0.0	0.0	0.0
	Total	934	527	1,893	2,264

- **Technical capability computers.** These are systems configured and purchased to solve the largest, most demanding problems.
- **Technical capacity computers.** We have divided the capacity segment into three subsegments based on price band:
  - **Enterprise**, which are systems purchased to support technical applications in throughput environments and sold for $1 million or more.
  - **Divisional**, which are systems purchased to support technical applications in throughput environments and sold for $250,000 to $999,000.
  - **Departmental**, which are systems purchased to support technical applications in throughput environments and sold for less than $250,000.

Source: IDC, 2000.

We believe the effective opportunity for DP in the high-performance technical systems market lies for now in the low-end market for departmental and divisional servers. As such, we believe the current market opportunity for DP in the high-performance technical systems is around $4.2 billion.

Biosciences

We believe biosciences currently represents the largest opportunity for DP companies. We highlight the opportunity in the biosciences market below.

Consider Celera Genomics, the private company that claimed first prize in sequencing the human genome. Celera's infrastructure uses:

Exhibit 43. Celera Genomics

$1 million	Annual Electricity Consumption
70 TB	Storage
1.7 TFLOP	Aggregate Processing Power
900 processors (a)	Server Processors
6 Paracel GeneMatchers	Thousands of Processors Each

(a) Compaq AlphaServer.

Source: IDC.

Compare this to what United Devices can offer.

Exhibit 44. Celera Genomics Versus United Devices

Celera Genomics		United Devices
$1 million	Annual Electricity Consumption	$4.7 million (b)
70 TB	Storage	5 PB
1.7 TFLOP	Aggregate Processing Power	600 TFLOP
900 processors (a)	Server Processors	100,000
6 Paracel GeneMatchers	Thousands of Processors Each	

Note: We assume UD's network comprises 100,000 Intel P3 PCs, with 50 GB of hard drive space.
(a) Compaq AlphaServer.
(b) Consumed by PCs connected to the network.

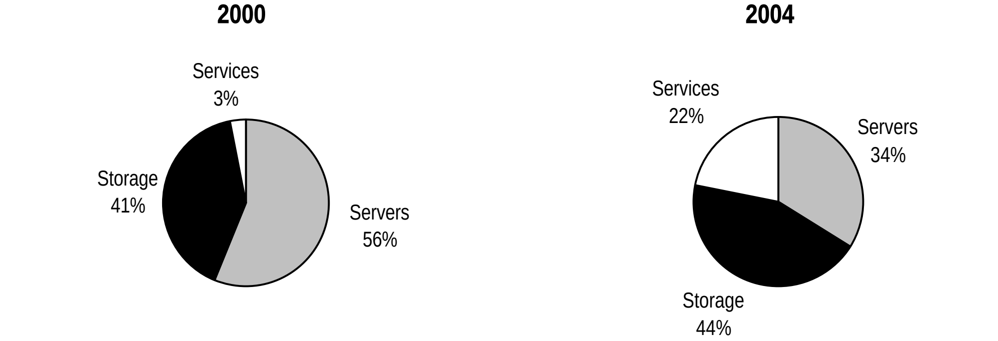
Source: IDC; Bear, Stearns & Co. Inc.

For processing tasks that can be done in parallel, UD can offer over 350x the aggregate processing power of Celera Genomics' infrastructure. In addition, the majority of UD's infrastructure costs are absorbed by consumers. Even the electricity to power a 100,000-PC DP network is consumed by the home PCs.

This is why the biosciences market is so attractive to DP vendors, and why firms like GlaxoSmithKline are interested in DP technologies. Many biosciences firms use vast clusters of Linux servers to create a pool of processing resources. Incyte, for example, has over 3,000 Linux boxes to power its processing needs. Since these processing clusters process data in parallel, they provide an opportunity for DP networks.

Apart from hardware costs, DP vendors can save biosciences firms services and storage costs. Jobs that a DP company performs are automated and require little maintenance. In addition, biosciences companies can farm out storage of the raw data to the DP vendor, since the vendor has centralized storage and distributed storage assets they can tap. DP vendors can enable these firms to tap their own resources (desktops, servers) or supplement this with a large user base on the Internet.

Exhibit 45. Biosciences IT Market Segmentation, 2000 and 2004



Source: IDC.

IDC estimates the biosciences IT market represents a $11.6 billion opportunity in 2004, from $2.2 billion in 2000, representing a CAGR of 50%. We believe DP companies may be able in part to defray some of the hardware, software, and services costs.

Enterprise Software Opportunity

We believe the enterprise software market opportunity for DP in financial services and bio-informatics is accurately measured by the potential total addressable base of customers.

We have created a back-of-the-envelope estimate of the potential market opportunity for enterprise software sales in DP. Based on our knowledge of an enterprise implementation and pricing, we arrive at the following forecast.

Exhibit 46. Distributed Processing Enterprise Software Opportunity

		Firms		
		100	**500**	**1,000**
PCs	**500**	50,000	250,000	500,000
	1,000	100,000	500,000	1,000,000
	5,000	500,000	2,500,000	5,000,000
$/Month/PC	**$ 100**	$ 5,000,000	$ 25,000,000	$ 50,000,000
	100	10,000,000	50,000,000	100,000,000
	100	50,000,000	250,000,000	500,000,000
	200	10,000,000	50,000,000	100,000,000
	200	20,000,000	100,000,000	200,000,000
	200	100,000,000	500,000,000	1,000,000,000
	300	15,000,000	75,000,000	150,000,000
	300	30,000,000	150,000,000	300,000,000
	300	150,000,000	750,000,000	1,500,000,000
Contract Length	**12 months**	$ 60,000,000	$ 300,000,000	$ 600,000,000
	12 months	120,000,000	600,000,000	1,200,000,000
	12 months	600,000,000	3,000,000,000	6,000,000,000
	12 months	120,000,000	600,000,000	1,200,000,000
	12 months	240,000,000	**1,200,000,000**	2,400,000,000
	12 months	1,200,000,000	6,000,000,000	12,000,000,000
	12 months	180,000,000	900,000,000	1,800,000,000
	12 months	360,000,000	1,800,000,000	3,600,000,000
	12 months	1,800,000,000	9,000,000,000	18,000,000,000

Note: $/month/PC takes into account server license fees associated with an enterprise sale.

Source: Bear, Stearns & Co. Inc.

We believe the number of firms likely to use DP software in-house is not large. In fact, outside of financial services, biosciences, and academic and government institutions, the market for DP software could be quite limited. We also conservatively estimate that the number of PCs implemented within those firms (likely *Fortune* 1000) will be in the three-figure area. Many of the DP vendors we have spoken to have indicated cost per seat in the low-three-digit area (this figure incorporates server license fees). Based on these assumptions, we believe the addressable market opportunity for DP software into enterprises is probably around $1.2 billion over the next four years. The actual market will be tiny over the next couple of years.

Performance Measurement/Load Testing

The performance measurement and load-testing market could be a good opportunity for DP vendors. Both UD and Entropia signed their first customers in this space — UD with Exodus Performance Labs, and Entropia with Envive. We reiterate our belief that DP vendors offer their performance measurement/load testing customers a unique perspective — geographic, platform, and systems.

IDC estimates the market for automated software quality (ASQ) tools for distributed environments will reach over $2.2 billion in 2004 from $618 million in 2000, representing a CAGR of 40%. Currently, Mercury Interactive is the leader in this market. As a side note, IDC estimates the market for performance management will reach $5.5 billion in 2004, from $2.7 billion in 1999, representing a 15.3% CAGR.

AGGREGATE MARKET OPPORTUNITY

We believe the following represents a best-efforts forecast of the opportunity for the DP companies in 2004.

Exhibit 47. Distributed Processing — Aggregate Market Opportunity ($ thousands)

	2001	2004	CAGR
High Performance Computing	2,260	4,520	15%
Performance Testing/Measurement	600	2,200	30%
Enterprise Software	120	1,200	59%
Addressable Market Opportunity	2,980	7,920	22%
Assumed Market Penetration	1%	4%	
Implied Market Opportunity	29.8	316.8	81%

Source: Bear, Stearns & Co. Inc.

IMPACT ON OTHER INDUSTRIES

- **Performance Measurement/Load Testing.** We believe the very first industry impacted by DP is performance measurement and load-testing. We count Keynote Systems and Mercury Interactive as the market leaders. We regard Service Metrics (part of Exodus), Micromuse, Computer Associates, and BMC Software as potential competitors. Inversely, we also believe these companies are the most likely partners and acquirers of companies like UD, Entropia, DataSynapse, and Parabon. While Service Metrics and Keynote are service providers, Mercury and the others are primarily software vendors. In the early stages, DP vendors will likely work with the service providers. The software vendors will be impacted later.

- **Microprocessor Manufacturers.** Intel has been working with United Devices closely over the last several quarters. In fact, one of Intel's major goals has been to accelerate the microprocessor uptake cycle. Intel CTO Patrick Gelsinger believes peer-to-peer computing will benefit Intel's efforts to grow its microprocessor market.

- **PC OEMs.** We believe PC manufacturers could have opportunities in DP. Because so much of the benefit of distributed applications rests on processing power and hard-disk space, PC OEMs may benefit from users purchasing next-generation PCs.

- **Software Vendors.** Software vendors like Sungard (financial services), BEA Systems (Web application server), Oracle (database and apps), Adobe (design), and Synopsis (CAD/EDA) could offer value-added software features like DataSynapse's real-time middleware solutions.

DP could also change the way software is designed. If processing becomes an easily-clusterable resource going forward, this could impact the way developers design software to take advantage of processing resources more effectively.

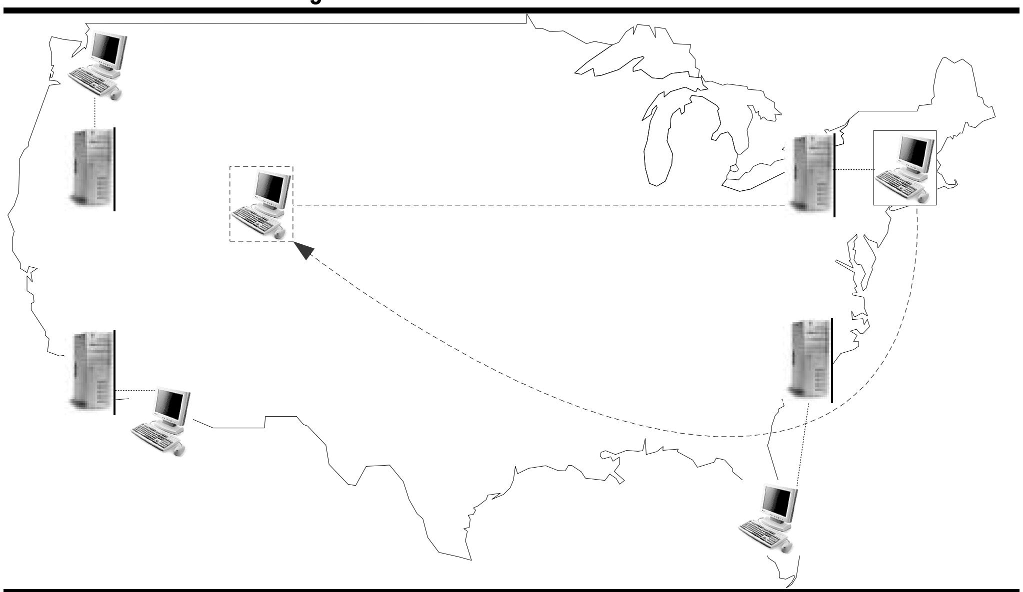
Distributed Storage Services: Collapsing Geographic Isolation

Distributed storage has come to mean Fibre Channel Storage Area Networking (FC SAN) provided by incumbents like EMC, IBM, and Hitachi, and enabled by Brocade, Emulex, and QLogic. Despite the popularity of SANs, two hallmark limitations of FC SAN are geographical reach (FC SANs are usually good to about 10 kilometers over certain grades of fiber) and limited clustering capabilities (up to 127 nodes).

We believe the next generation of distributed storage services will be truly distributed, on a geographic scale covering the globe and clustering thousands of nodes. The call for geographic independence and unlimited clustering across a storage network arises from the nomadic nature of users and the capital intensity associated with centralized proprietary storage systems. Because of the Internet and the Web in particular, users are demanding faster access to data from any point. Service and content providers are searching for cheaper solutions that can serve users faster.

Centralized storage environments have limited architectural flexibility. When an individual requests data from a fixed geographic location close to the storage (store), the delivery of the data is generally fast. However, upon moving to another location, the individual's request traverses the distance between the two geographies.

Exhibit 48. Centralized Storage Environment



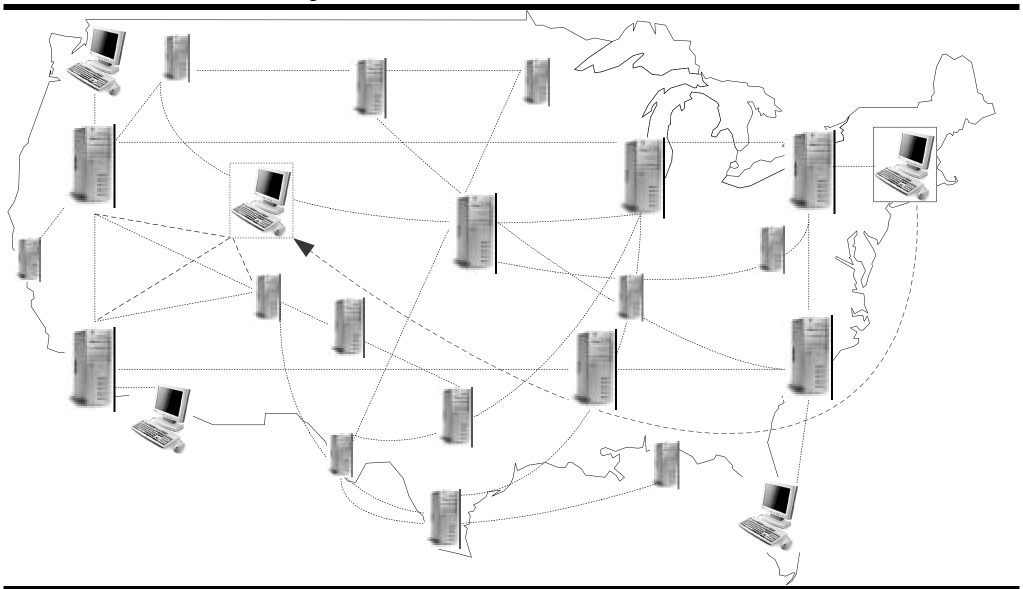
Source: Bear, Stearns & Co. Inc.

The chief network bottleneck in a centralized storage system is therefore latency. In centralized storage, like other centralized systems, no matter where an individual is requesting data from, the distance between the individual and the data is the distance between the individual and the data object located at the core storage facility. The closer the individual is to the core, the faster the response time; the further the individual is from the core, latency grows by a non-linear factor.[14]

This may not be intuitive. Take, for example, a network cache. The network cache within an enterprise stores frequently requested data objects (graphics, pictures, logos) that originally reside on “origin” servers. The distance between the user and the sought-after data in this case is the distance not between the user and the core origin server but between the user and the network cache. By bringing data closer to users, distributed storage networks decrease the time it takes users to access data.

The cost of caching data locally is much lower than expending bandwidth to pull the data from the core origin server. The cost of disk is significantly lower than the cost of bandwidth. Duplicating a data object once on disk across geographically dispersed storage devices for users on the network to access is infinitely cheaper than continually and redundantly transmitting data from a central store. In a distributed storage system, caching or storing a data object across two locations involves a one-time bandwidth expense of duplicating and transmitting the object through the MAN or across the WAN. In a centralized system, being forced to access a centralized storage system from a distance involves both MAN and WAN bandwidth. WAN bandwidth, while cheaper than MAN bandwidth, is still significantly more expensive than disk.[15]

Exhibit 49. Distributed Storage Environment



Source: Bear, Stearns & Co. Inc.

This superdistribution can only succeed with higher-layer technologies: caching, load balancing, and distributed database management. In a distributed storage architecture, data objects are duplicated across geographically-dispersed storage devices using sophisticated caching algorithms. Access to data in such a network must be carefully

[14] Being twice as far away from a storage system does not equate to twice the access time; because of the way networks are often constructed, it could take more than double the time.

[15] Technologies like Gigabit Ethernet are helping to lower prices in the MAN, as GigE is much cheaper than comparable bandwidth on SONET/ATM.

handled so that users are receiving data from the most efficient and cost-effective network node. This requires tight management and monitoring of distributed copies of data objects, a concern absent in centralized systems.

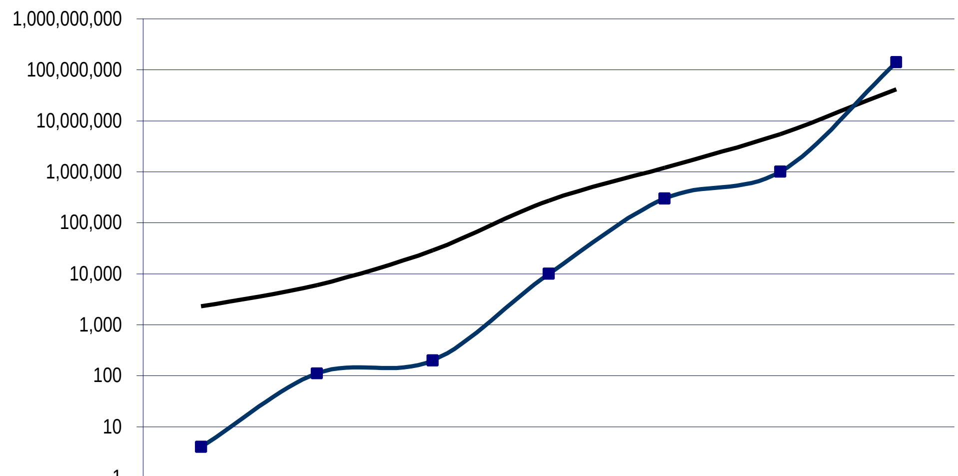
In many ways, this is precisely what Akamai has done with select data objects. Akamai evaluates the user's geographic location and then determines from which machine to serve data to the user. In effect, Akamai decreases the response time of accessing data and the cost of storage by leveraging a large, globally distributed cluster of caches running on commodity hardware.

Akamai's sophistication is not its distribution. Akamai's sophistication lies in its caching, network monitoring, and routing algorithms. This is why we believe the coming generation of distributed storage technologies and services will look a lot like Akamai's network.

Revisiting Moore and Metcalfe

Recall, the graph of Moore's Law and Metcalfe's Law indicates that the growth in the number of network nodes continues to outpace the growth of processor capacity, and hence storage advances.

Exhibit 50. Moore and Metcalfe



	1970	1975	1980	1985	1990	1995	2000
Transistors	2,300	6,000	29,000	275,000	1,200,000	5,500,000	42,000,000
Networks	4	111	200	10,000	300,000	1,000,000	140,000,000

Source: Bear, Stearns & Co. Inc.

We believe the anticipated continued growth in the number of nodes on the network is possible only through international adoption of the Internet and the continued introduction and addition of edge devices. As the number of networked devices and the geographic spread of users continue to grow, having a distributed storage architecture may not only be advantageous, it will likely be necessary. We believe the promise of a fully distributed storage system lies in its capacity to allow the virtual database called the Internet to infinitely scale.

The Storage Matrix

We believe storage can be segmented into the following categories:

Exhibit 51. Storage Matrix

	Technologies	Services
Centralized	EMC	Xdrive
Distributed	Zambeel	Scale Eight

Source: Bear, Stearns & Co. Inc.

We will be focusing on the bottom layer of distributed storage technology and services providers.

Exhibit 52. Distributed Storage Companies

Company	Funding		Investors	Customers
Scale Eight	$26.5 million	Series B	CenterPoint Ventures, Crown Advisors, InterWest Partners, Oak Investment Partners	Akamai, MTVi, Vingage
Zambeel	Unknown		Kleiner Perkins Caufield & Byers, New Enterprise Associates, Integral Capital Partners	stealth

Source: Bear, Stearns & Co. Inc.

As the name indicates, distributed storage services refer to an outsourced storage service where storage assets are distributed across geographies and linked to one another, often on a peer-to-peer basis.

We believe universities are driving much of the development in distributed storage. Two academic projects focused on globally distributed storage are Interet2 and OceanStore.

Internet2

A 180-university consortium established to develop and deploy advanced network applications and technologies, Internet2 has as its goal of alleviating the following:

> *The performance problems of the commodity Internet are well known: Not only can access times be long or unpredictable, availability can be unreliable. Because the performance of the Web is dependent on the real-time performance of wide area networking it cannot be improved without the global application of resources. Finally, and of most importance to the Web caching community, the Web is wasteful in its use of wide area networking due to repetitive transmission of the same data. While caching and mirroring strategies continue to evolve to address these problems, the underlying client/server architecture of the Web does not allow us to*

improve access to certain services by applying resources to those services alone. This structure inhibits the development of a viable economic model for differential investment in new infrastructure.

What Internet2 strives to create is a distributed network of services to maximize the utilization of resources on the Internet. Three principles guide this effort[16]:

1. Collections of content — The kinds of content that can go into a channel are a superset of what is accessible via the Web. It includes the familiar collection of portable files and programs available through Web protocols (e.g., text, graphics, Java applets, multi-media) and also services that utilize non-Web protocols, such as back-end database systems and remote search engines. But an aggregation of such content becomes a channel only when it is more or less deliberately collected and made available to end users. For example, the entire ensemble of digital content used during an academic course — papers, videos, application and applet software, large data sets for analysis, running simulations, online examinations, etc. — could be made into a channel.

2. Policy-based applications of resources — Channels use a combination of network and storage resources to localize collections of content that people want, so that it can be accessed quickly and reliably when they want it. Any given channel will manage its resources according to policies implemented by its creators, the owners of the infrastructure, and the users. So-called "push" channels bring collections of content all the way to the desktop machine, and therefore need no re-engineering of the infrastructure. Server channels, of the kind I2-DSI (Internet2-Distributed Storage Infrastructure) is designed to support, allow the application of resources at the server level, enhancing the delivery of collections of content for a wider population of users.

3. Transparent delivery to end users — The default Internet paradigm is to resolve a domain name to a single host in a globally-consistent manner. The Web builds on this model by supporting a uniform naming scheme (URLs) for objects stored on servers and accessible via Web protocols. Server channels continue this paradigm by redirecting service requests to localized content replicas using a transparent access mechanism that preserves the global consistency of domain-name resolution. For instance, two students taking our hypothetical course in different parts of the world could click on the same hyperlink on their teachers' Web page and be directed to local copies of the same object.

OCEANSTORE

OceanStore comprises a federation of utility providers who cooperate to synthesize a seamless, consistent, highly available, mobile data platform. It is a component of the Endeavour project at Berkeley. Unlike other file systems, OceanStore provides truly nomadic data that is free to migrate and be replicated anywhere in the world.

OceanStore's premise, like Internet2's, is that the Internet as it stands today is highly unreliable, centralized, and geographically localized. OceanStore's vision is to create a globally-distributed cluster of "untrusted" nodes that (in sheer number) creates a

[16] Source: Internet2.

complete fail-over network. This was one of the lessons we learned from Napster in the first part of our report.[17]

Two design goals ground OceanStore: untrusted infrastructure and nomadic data.

- **Untrusted Infrastructure.** Most devices are individually unreliable. By aggregating devices and creating a redundant network, the infrastructure in totality can be made reliable. There has to be a separation between the class of devices storing data and the class of devices delivering the intelligent network and storage management layers. In essence, data stores act as dumb storage buckets with very little intelligence. Separate management devices store intelligent metadata and operate a structured tagging system, caching algorithms, load balancing, and storage management intelligence.

- **Nomadic Data.** OceanStore defines "nomadic data" as data that is allowed to flow freely. Ideally, every data object should be cacheable, anywhere and anytime. (This is what OceanStore refers to as "promiscuous caching.") OceanStore admits that "promiscuous caching" introduces a level of complexity; however, according to the consortium, this kind of caching introduces greater flexibility and functionality. What it requires is "introspective monitoring," a scheme that discovers relationships between objects on a metadata level.

We believe many of the distributed storage services and technology providers are using similar innovations. We profile distributed Internet Storage Infrastructure provider Scale Eight, and distributed storage technology provider Zambeel, as we feel they most clearly exemplify our Internet 3.0 theme.

SCALE EIGHT

Scaling with Demand

When Akamai decided to offer storage services, it partnered with San Francisco-based Scale Eight.

Scale Eight's mission is to become the dominant provider of storage services to the Internet by offering distributed storage services via its proprietary technology. The growing demand for multimedia files and complex data raises the issue of how these objects will be stored and delivered to end-users efficiently and cost-effectively.

Scale Eight has determined that by creating a distributed network infrastructure on commodity storage devices, the company can offer Internet-based storage services at a fraction of the cost of a home-grown Wide Area Storage Network.

[17] Fragmentation and distribution yield redundancy. Distributed autonomous devices connected on a network create massive redundancy even on less-than-reliable PC hard drives. Redundancy on a massive scale yields near-perfect reliability. Redundancy of this scope and reach necessarily utilizes resources that lead to a network topology of implicitly "untrusted" nodes. In an implicitly untrusted network, one assumes that a single node is most likely unreliable, but that sheer scale of the redundancy forms a virtuous fail-over network. Enough "backups" create a near-perfect storage network.

Products and Services

Scale Eight MediaStore is the company's core offering. MediaStore enables customers to access chosen files from any server or browser. Scale Eight enables this by installing a single file server called a MediaPort in a customer's LAN through which the customer can access files. The function of the MediaPort is to cache frequently requested files, and to store and retrieve files from Scale Eight StorageCenters.

The MediaPort is grounded on a proprietary file system called the Scale Eight Global File System (8FS) which can offer customers a holistic image of the geographically distributed file system (aggregate picture of all the LANs). Files that enter Scale Eight's network are mirrored (replicated/cached) across multiple facilities, and load balanced for optimal delivery.

Customers access their files (all file types supported) either through their LAN (through MediaPort) or through a browser (if accessing through the WAN) through a proprietary naming system (8RL, an authenticated URL) that uniquely tags each file.

Scale Eight operates four StorageCenters **—** two in the U.S. (California and Virginia) and one each in London and Tokyo.

Service plans start at 300 GB. Customers can add capacity as needed, in real time. Typically, a managed mirrored TB per month costs $25,000, roughly 80% less than StorageNetworks' cost per TB per month. Scale Eight can offer this kind of pricing because the company uses commodity hardware. The "special sauce" is the software that powers the mirroring, load balancing, routing, and caching.

As an interesting note, Scale Eight boasts David Patterson, Professor of Computer Science at UC Berkeley, as the company's Chief Scientist. David Patterson is the inventor of RAID systems (redundant array of inexpensive disks), which has been the architecture of modern storage systems.

Scale Eight has signed up Akamai as a customer and Exodus as a reseller of its services.

ZAMBEEL

Carrying a Storage Pouch with You

While Zambeel remains in stealth mode, we have pieced together a few details that may provide the reader with a general idea of the company's vision.

Located in Fremont, California, and backed by Kleiner Perkins Caufield & Byers, New Enterprise Associates, and Integral Capital Partners, Zambeel was founded by several industry veterans with expertise in distributed systems.

Zambeel is developing a storage architecture that is aimed at resolving the problems of expensive proprietary storage systems and their highly centralized nature that limit flexibility and geographic coverage.

The company will be introducing a distributed storage architecture using commodity hardware that is able to form clusters of thousands of machines across vast distances.

Unlike Scale Eight, which is offering a service based on its technology, Zambeel is attempting to become the arms merchant for distributed storage systems.

What this fully distributed architecture enables is storage capacity on demand, from anywhere, for all data types. Using complex caching, load balancing, mirroring and data duplication algorithms, and a fresh library of metadata, Zambeel will enable storage service providers and corporations to create a fully redundant, secure failover network spanning vast distances at a fraction of the cost of competitive solutions (SAN and NAS).

OUTSOURCED STORAGE SERVICES PROVIDERS

Storage service providers (SSPs) naturally have distributed architectures in order to mirror data and offer globally-available storage services. Once data enters an SSP's network, it can be shuttled to any node on the network. As such, SSPs are ideally positioned to become fully-distributed storage service providers.

We believe SSPs will have to overcome two technology hurdles in order to take full advantage of distributed storage architectures: virtualization and global Layer 4-7 capabilities. By virtualization, we mean again the ability to utilize resources in an object-oriented manner — where individual servers and storage devices are usable for multiple customers. [Currently, SSPs offer customers their own dedicated storage device(s).] By global Layer 4-7, we are referring to the ability of the devices within the network to be woven into a unified cluster fabric. Each node on the network then becomes a commoditized resource with the distinction of having geographic perspective necessary to serve users within a geographic locale.

Exhibit 53. Outsourced Storage Services Providers

Public Company	Ticker	Service
EDS	EDS	EDS Intelligent Storage Services
Hewlett-Packard	HWP	
IBM	IBM	
StorageNetworks	STOR	STORmanage
Private Companies	**Funding**	**Investors**
Arsenal	$12.6 million Series A	SouthEast Interactive Technology Funds
Creekpath	$23 million Series A	Telesoft Partners, Sequel Venture Partners, Exabyte
Storageway	$36 million Series B	WorldView Technology Partners, Matrix, Montreux, Redpoint
Scale Eight	$26.5 million Series B	Oak Invetment Partners, Centerpoint Ventures, Crown Advisors
Storability	$18 million Series B	Battery Ventures, Madison Dearborn Partners, Lightspeed Venture Partners
Sanrise	$100 million Series B	Crosspoint Venture Partners, Comdisco Ventures, Exodus, Hitachi, Morgan Stanley Dean Witter, Veritas
WorldStor	$33 million Series B	Mid-Atlantic Venture Funds, Keystone Venture Capital, Five Paces Ventures, W.R. Huft Asset Management

Source: Bear, Stearns & Co. Inc.

We believe the SSPs most able to offer virtualization and distribution will likely survive. Otherwise, the capital and margin models may not sustain the businesses.

MARKET OPPORTUNITY

For service providers like Scale Eight, we believe content delivery and rich-media intensive organizations are likely to be their first customers. For companies like Zambeel, we believe it is entirely too early to evaluate market size and opportunity.

Should Zambeel gain traction, the market opportunity would indeed be the entire market opportunity for EMC.

SSPs like StorageNetworks will not be able to exist without distributing storage resources across geographies. In addition, as we have noted, the capital structure of outsourced storage can work only if storage resources are virtualized. Should an SSP have to dedicate a separate storage box for each customer, the capital and margin structures will prove economically unfeasible, and the provider's need for equipment will be a function of the number of customers and not the amount of storage required to serve those customers.

In brief, outsourced storage is trying to decouple data from geographic centralization. In the case of StorageNetworks, once data is on StorageNetworks' network, it can be replicated to each S-POP (Storage Point-of-Presence).

We believe one of the ways to measure the market opportunity of distributed storage providers is by looking at SSPs. While not precise, it gives us a proxy for the distributed storage services market opportunity. We think all outsourced storage services will be distributed in time.

Exhibit 54. Worldwide Storage Utility Spending by Region, 1998–2005E ($ in millions)

	1998	1999	2000	2001	2002	2003	2004	2005	2000–2005 CAGR (%)
United States	0	11	139	559	1,770	3,601	5,608	6,643	*116.7*
Europe	0	0	10	99	432	1,245	2,101	2,697	*206.3*
Canada	0	0	2	11	44	144	252	344	*177.7*
ROW	0	0	2	13	126	508	853	1,021	*248.0*
Worldwide	**0**	**11**	**153**	**682**	**2,372**	**5,498**	**8,814**	**10,705**	***133.9***

Note: Includes all service spending with pure-play SSPs and any storage-on-demand spending with other types of vendors.

Key Assumptions:

- Much of the spending in this space will be done as part of a larger contract; for example, it will be done for Web site hosting, ASP, ISP, or other services.
- As more firms become familiar and comfortable with the SSP model, the customer base for this market will expand to include all types and sizes of organizations.

Messages in the Data:

- Even by 2003, about two-thirds of the spending in this segment will still come from the United States and Canada.
- Year-to-year growth will slow between 2000 and 2005 for the overall segment.

Source: IDC.

Exhibit 55. Worldwide Storage Utility Spending by Vendor Type, 1998–2005E ($ in millions)

	1998	1999	2000	2001	2002	2003	2004	2005	2000–2005 CAGR (%)	2004 Share (%)	2005 Share (%)
Pure-play storage service providers (1)	0	8	143	350	1,292	2,826	3,491	3,708	*91.8*	*39.6*	*34.6*
Storage product suppliers	0	0	1	104	285	415	575	615	*261.2*	*6.5*	*5.7*
Internet data centers/ Web hosting firms (2)	0	1	2	85	395	1,425	3,245	4,425	*366.6*	*36.8*	*41.3*
xSPs	0	1	2	35	60	85	105	117	*125.6*	*1.2*	*1.1*
Resellers/integrators	0	1	2	20	85	115	145	160	*140.2*	*1.6*	*1.5*
Telcos	0	1	3	88	255	632	1,253	1,680	*254.5*	*14.2*	*15.7*
Total	**0**	**11**	**153**	**682**	**2,372**	**5,498**	**8,814**	**10,705**	***133.9***	***100***	***100***

(1) Includes SSP services that are resold or OEMed (i.e., rebranded) by Internet data centers, Web hosting firms, and other organizations.
(2) Represents discrete storage utility offerings that are designed, marketed, and delivered by Internet data centers and Web hosting firms with their own personnel and infrastructure. Does not include sell-through or rebranding of SSP offerings.
Note: Includes all service spending with pure-play SSPs and any storage-on-demand spending with other types of vendors.

Key Assumptions:

- Internet data centers, Web hosting firms, and telecommunications companies will become more serious competitors in this market in coming years.
- Storage product suppliers will find that the utility model conflicts with their product business limited these firms' overall competitiveness in this market.

Messages in the Data:

- Pure-play SSPs will show the slowest growth overall between 2000 and 2005 in this market.
- By 2005, Internet data centers, Web hosters, and other related firms will own the biggest share in this market.

Source: IDC.

Peer-Based Storage

Peer-based storage is least likely to take off any time soon. However, we believe there are distinct advantages to having data coupled with the primary user device (the PC). We therefore highlight some of the advantages of peer-based distributed storage.

We looked at over 275 PCs in an enterprise environment to see how much of a desktop's hard drive was utilized. We discovered that on average, roughly 15% of the total hard drive capacity was actually utilized by enterprise users.

Exhibit 56. Desktop Hard Drive Utilization (HD capacity in Gigabytes)

	HD Capacity	Partition	Used	Unused
Desktop	7.6	3.8	1.1	6.4
x1000	7,591	3,769	1,146	6,445

Note: Partition — Many enterprise desktop hard drives are partitioned so that a portion of the hard drive is kept unusable. In our example, we discovered roughly 50% of the total hard drive capacity was partitioned.

Source: Bear, Stearns & Co. Inc.

50GB hard drives go for roughly $100. Across 1,000 PCs, roughly $85,000 of the hard drive cost would be unutilized, or roughly 6.5 Terabytes.

Since we have yet to hear of such an offering, we believe it would be premature to evaluate market opportunity. However, we can foresee a time when enterprises make use of the latent storage across desktops, workstations, and servers by distributing archived or bulk data like e-mail redundantly across multiple machines for backup.

We believe companies like United Devices, Entropia, DataSynapse, and Applied Meta are well positioned to enable such services in enterprises through their respective software platforms.

We would like to reiterate our view that the value of such an offering is not in the amount of savings associated with such services; memory prices are eroding too fast for that. The value is in virtualizing storage resources in order to weave disparate machines into one large distributed storage network system using complex software.

Exhibit 57. Potential Peer-Based Storage Technology Providers

Company	Funding		Investors	Customers
Applied Metacomputing	$6 million	Series A	Polaris Venture Partners	Boeing, Stanford, U.S. Naval Research Laboratory, NASA's Ames Research Center, Lawrence Livermore National Laboratory
DataSynapse	$5.1 million	Series A	Rare Ventures, Neocarta Ventures, The NYC Investment Fund, Silicon Alley Venture Partners, Wand Partners	First Union
Entropia	$30 million	Series B	Moore Capital, RRE Ventures	Envive, SolidSpeed
Parabon	$6.5 million	Series A	Undisclosed	
United Devices	$13 million	Series A	SOFTBANK Capital, Oak Investment Partners	Exodus, iArchives
XDegrees	$8 million	Series A	Red Point Ventures, Cambrian Ventures	

Source: Bear, Stearns & Co. Inc.

Impact on Other Industries

Distributed storage services and technology providers could have a significant impact on the storage industry. Just as caching vendors like Inktomi have obviated the need to purchase significant numbers of servers in many cases, distributed storage services and technology providers could obviate the need to purchase expensive storage systems. In fact, one of the big attractions of emerging distributed storage vendors is price. Many of the distributed storage services and technology providers are using commodity-type boxes for their hardware needs.

REVIEW OF INTELLIGENT NETWORK TECHNOLOGIES AND SERVICES

Caching

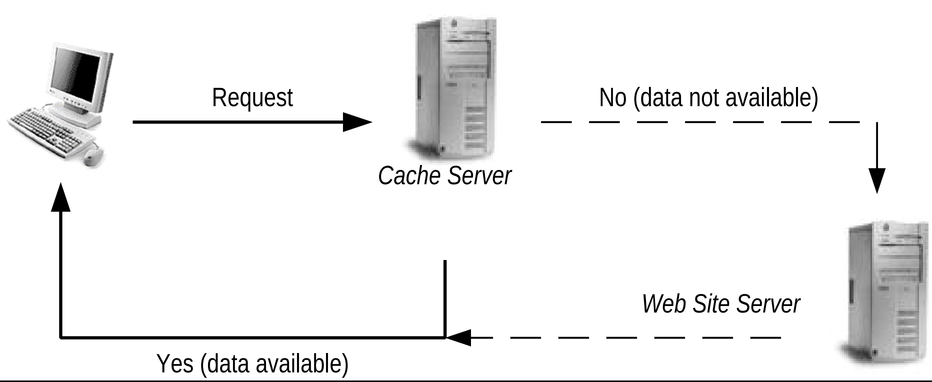
One way to resolve the latency and traffic congestion problems on public networks (such as the Internet) and on private networks (such as a corporation's) is through caching. *Caching essentially stores (caches) the most widely used portions of a Web site, a group of Web sites, or the Internet as a whole on a separate server or other hardware device.* When a user attempts to access information, the Web cache server intercepts the request and checks to see if the information is stored locally. If it is, the information is delivered to the user from the local cache rather than from the content provider's actual Web server, resulting in faster responses to end-users and considerable bandwidth savings for content providers (discussed in more depth below). If the information is not available locally, the request is forwarded to the origin Web server. As the origin server responds, the cache server passes the content to the end-user and simultaneously copies applicable content (since not all content is cacheable) into memory to more effectively serve subsequent requesters. The cache refreshes its database to store only the latest and most requested data.

In summary, caches address the following issues:

- high cost of bandwidth versus storage; and

- slow response times to users due to server and other congestion.

Exhibit 58. How Web Caching Speeds Reponses to Users



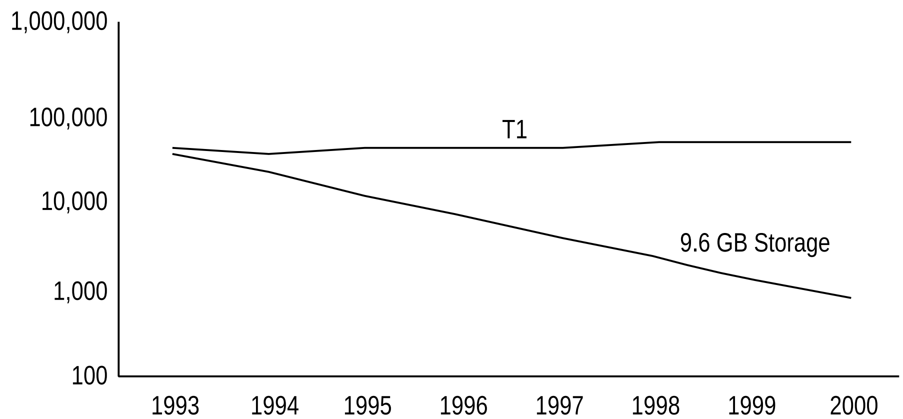
Source: Bear, Stearns & Co. Inc.

As noted, bandwidth savings and improved performance are the two key drivers of caching product sales.

Caching reduces telecommunications costs (cost of deploying additional bandwidth) for Internet service providers and, where high-speed lines are unavailable, provides a viable alternative. Local Internet storage caching is less expensive than network re-transmission and, according to market research firm IDC, becomes more attractive by about 40% per year. Caches are particularly efficient for international traffic and traffic that otherwise moves across large network distances.

Exhibit 59 compares the cost of a disk storage system with comparable network capacity. The basis of comparison is the annual cost of an average WAN (Wide Area Network) service connection rate (the network portion of the connection, not the cost of translation or interface equipment on either end of the network segment) with the cost of storing a full day of data transmitted over the connection. The full day of transmission capacity for the storage system was chosen based on typical cache-to-backup ratios of 1% and derated by about 3.5x to account for less than 100% WAN capacity exploitation and for a 28%-35% level of repeated data. 9.6GB was used as the basis of comparison because that is approximately the amount of data a T1 line (1.544 Mbps, 24 voice channels) can transmit in 24 hours at full rate.

Exhibit 59. Annual Network Line Cost Compared with the Cost of 9.6GB Storage — One Day's Transmission at T1 Line Rates ($)



Source: IDC.

Large ISPs connect to the Internet backbone through one or more T3 lines (44.736 Mbps, 672 voice channels). While bandwidth prices continue to decline, these lines are very expensive (average about $445,000 per year) and are also rather scarce (only available in select cities with installation wait times of one to two months if available). The cost savings benefits of caching versus bandwidth are equally compelling at the higher end of the market.

Exhibit 60. Estimated Bandwidth Cost Savings Derived from Caching

Current Cost Savings		Cost Savings Assuming 30% Decline In Line Costs	
Estimated cost of a T3 line/month	$ 37,083	Estimated cost of a T3 line/month	$ 25,958
x 12 = yearly cost	$ 445,000	x 12 = yearly cost	$ 311,500
Cost per Mbps per month	$ 824	Cost per Mbps per month	$ 577
T3 average data throughput (Mbps)	45.0	T3 average data throughput (Mbps)	45.0
x % of Internet traffic that is cacheable	40%	x % of Internet traffic that is cacheable	40%
= Cacheable traffic (Mbps)	18.0	= Cacheable traffic (Mbps)	18.0
Cache hit rate	40%	Cache hit rate	40%
Cacheable traffic x hit rate = Mbps served from the cache	7.2	Cacheable traffic x hit rate = Mbps served from the cache	7.2
x Cost per Mbps per month = bandwidth savings per month	$ 5,933	x Cost per Mbps per month = bandwidth savings per month	$ 4,153
x 12 = yearly bandwidth savings per T3 line	$ 71,200	x 12 = yearly bandwidth savings per T3 line	$ 49,840

Source: Patricia Seybold Group; Paul Kagan Associates, Inc.; Bear, Stearns & Co. Inc. estimates.

Below, we highlight the more notable caching vendors.

Exhibit 61. Caching Vendor Landscape

Software-Based	Hardware-Based
iMimic Networking	CacheFlow
Inca Technology	Cisco
Inktomi	Eolian
Volera (Novell)	F5
Oracle	Industrial Code & Logic(1)
Squid	InfoLibria
Vixie Enterprises	Lucent
	Network Appliance
	Nortel Networks

(1) Squid-based appliance.

Source: Bear, Stearns & Co. Inc.

Load Balancing/Bandwidth Management

Load balancing/traffic management products are systems that sit in a network and process network traffic streams, switching and otherwise responding to incoming requests from the outside (the Internet) or within a corporate network, by directing these requests to specific Web servers (or server clusters) based on a set of pre-defined rules. Most Web sites are composed of multiple servers that may have different capacity levels. The products we discuss below can "load balance" data traffic on a network, meaning that if a network administrator wishes, the product can direct traffic to any of these multiple servers according to its capacity. These products usually have the ability to test the servers they are connected to for correct operation, and re-route data traffic around a server should that one fail. Most of these devices also have the capacity to recognize the requester and/or the data being requested and prioritize the request and/or the response accordingly. The ability of the load balancer/traffic manager to consider the source of the message or the relative load on each of several replicated servers, and then direct the message to the most appropriate server, increases efficiency and uptime.

Bandwidth management technologies give network administrators at Internet service providers or corporate enterprises the ability to set and enforce policies to control network traffic, ensuring that networks deliver predictable performance for mission-critical applications. Bandwidth management systems can prioritize mission-critical traffic (such as SAP, PeopleSoft, etc.) as well as guarantee minimum bandwidth for the most critical, revenue-generating traffic (e.g., voice, transaction-based applications).

Exhibit 62. Representative Load Balancing/Traffic Management Appliance Vendors

Vendor	Product
Allot Comm.	Allot NetEnforcer
Cisco	LocalDirector
Coyote Point Systems Inc.	Equalizer
F5 Networks	BIG/ip System
HydraWeb	Hydra5000, Hydra2000, Hydra900
Industrial Code & Logic	TrafficCop
Intel	NetStructure (iPivot)
RADWARE	Web Server Director

Source: Company data; Internet Research Group; Bear, Stearns & Co. Inc.

Exhibit 63. Representative Load Balancing/Traffic Management Software Vendors

Vendor	Product
Allot Communications	Allot NetPolicy
IBM	eNetwork Dispatcher
Lightspeed Systems Inc.	IP Magic Suite
Microsoft	WLBS (Windows Load Balancing Service)
Platform	Application Resource Management/ LSF
Resonate	Central Dispatch

Source: Company data; Internet Research Group; Bear, Stearns & Co. Inc.

Exhibit 64. Representative Load Balancing/Traffic Management Intelligent Switch Vendors

Vendor	Product
Nortel Networks (Alteon WebSystems)	Ace Director
Cisco Systems (ArrowPoint Communications)	CSS 11000
Macaroni (FORE Systems/ Berkeley Networks)	ExponeNT switches
Extreme Networks	Summit series
Foundry Networks	ServerIron series
CyberIQ (formerly HolonTech)	HyperFlow series
Nortel Networks	Intelligent Load Balancing Accelar 700, ACE Director

Source: Company data; Internet Research Group; Bear, Stearns & Co. Inc.

Exhibit 65. Representative Traffic Distributor Vendors

Vendor	Product
Cisco	DistributedDirector
F5 Networks	3DNS Controller and 3DNS+ Controller
Intel	Network Application Engine. Product acquired via Intel's acquisition of NetBoost
Legato	Co-Standby Server for NT (distributed operating system). Acquired product via acquisition of Vinca
Microsoft	NT Cluster Distributing Operating System, which includes Convoy Cluster Software, a product acquired via Microsoft's acquisition of Valence Research
Resonate	Global Dispatch
Sun Microsystems	Enterprise Cluster Server cluster operating system

Source: Company data; Internet Research Group; Bear, Stearns & Co. Inc.

Exhibit 66. Representative Load Balancing/Traffic Management Intelligent Switch Vendors

Vendor	Product
CheckPoint	FloodGate-1
Cisco	QoS Device Manager
Packeteer	PacketShaper
Sitara	QoSWorks
TopLayer	AppSwitch

Source: Company data; Bear, Stearns & Co. Inc.

The public Internet's inherent architectural problems have given rise to a variety of data transport solutions designed to mitigate public Internet network delays. In many cases, these solutions are designed to avoid much of the public Internet altogether, instead routing data over privately-monitored networks to ensure satisfactory delivery.

We have identified the following services and technologies designed to avoid the Internet's inherent congestion:

- Content distribution networks and streaming media networks – Akamai, iBEAM, Real Broadcast Network (RBN);
- IP multicast – FastForward Networks (acquired by Inktomi), Bang Networks;
- Optimized routing solutions – InterNAP;
- Collocation facility data exchange and other forms of content peering — Equinix;
- Fiber-based networked operators who are beginning to add "intelligence" to their networks to be able to offer Internet content distribution services — AT&T and Enron; and
- Broadcast solutions, (i.e., vendors attempting to take advantage of the television broadcast infrastructure to deliver rich-media Internet content) — dotcast.

We regard the CDN/SMN (content delivery network/streaming media network) space as the most relevant to the distributed network services sector, as both CDNs and SMNs operate distributed networks.

Exhibit 67. Distributed Network Services

Public Company	Ticker	Technology	Customers
Akamai	AKAM	Proprietary	Yahoo!, 3,400+
Digital Island	ISLD	Inktomi	CNBC, FT, E*TRADE, among others
IBEAM	IBEM	Inktomi	MTVi, Launch.com, among others
Yahoo! Broadcast.com	YHOO	Proprietary	
RealNetworks	RNWK	RealNetworks	
AT&T	T		
British Telecom	BTY		
Deutsche Telecom	DT		
Enron	ENE		
Exodus	EXDS		
Genuity	GENU		
PanAmSat	SPOT		

Private Company	Funding		Investors	Customers
Bang Networks	$18 milion	Series A	Sequoia, Marc Andreesen, Dan Warmenhoven, Nicholas Negroponte, Angel Investors	CBS Sportsline, among others
Cidera	$75 million	Series D	Wheatley Partners, Berger Funds, meVC, Draper Fisher, Munder, GE Equity, among others	Akamai, among others
SolidSpeed	$4.5 million	Series A	Arbor Venture Partners	IBM reseller agreement
Speedera	$20 million	Series B	ABS Ventures, Banc of America, Comdisco, Hewlett-Packard, Oracle, Palo Alto Investors, Trinity Ventures, Continuum Group	Hewlett-Packard, Intel, among others

Source: Company data; Bear, Stearns & Co. Inc.

As we have indicated, most of these companies focus on the WAN. We believe the next generation of distributed network services companies will continue to push into the MAN and start penetrating the LAN.

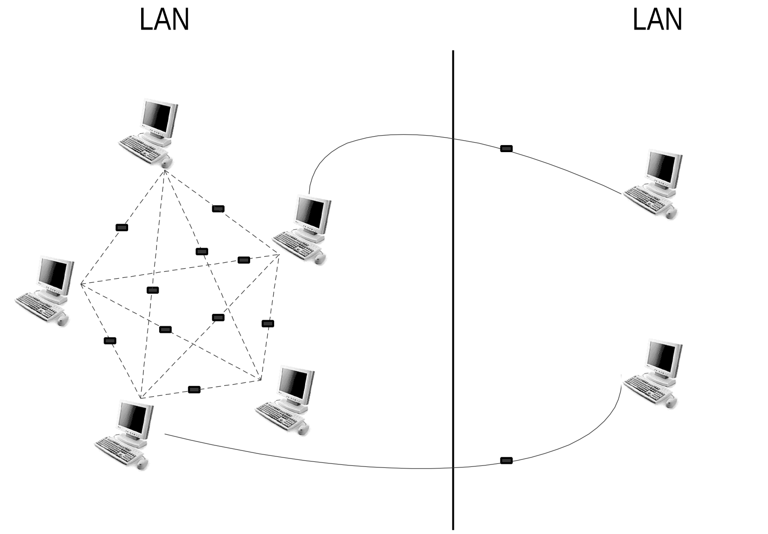
Next-Generation Distributed Network Technologies and Services

The next generation of network services and technologies will continue to offer advantages in caching, load balancing, and bandwidth management. What distinguishes these next-gen services and technologies is their emphasis on further pushing the distribution mantra.

As the first generation of distributed network services companies have discovered, distribution alone is not enough. What makes a distributed network service valuable is the intelligence grounding the network — that is, Akamai is interesting not because its network of servers is distributed globally, but rather because of the "special sauce" in its caching and routing algorithms on the servers and network. Many emerging distributed network services are grounded on intelligent software layers that ride on top of existing distributed assets, in the LAN and across the WAN. Others demand a buildout of network infrastructure.

Until recently, the 80/20 rule applied to LAN versus WAN traffic: the rule specifying that about 80 percent of each user's network traffic is local (remains in the LAN).

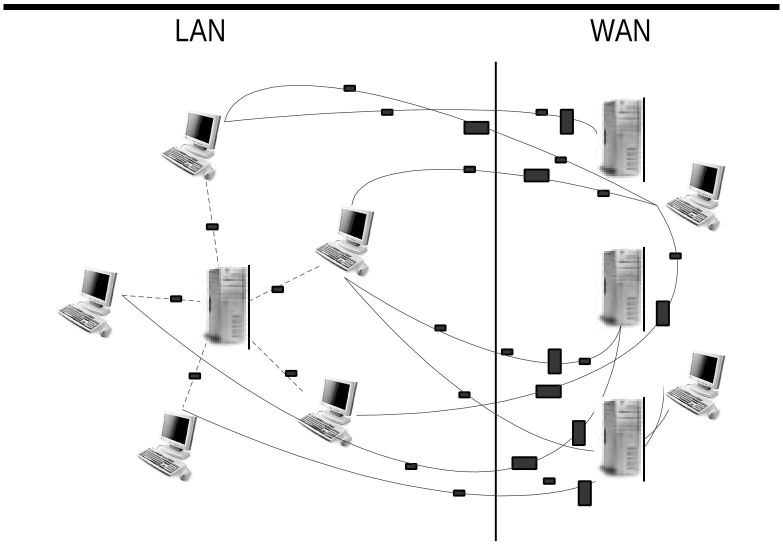
Exhibit 68. Phase I: 80/20 — The LAN Era



Source: Bear, Stearns & Co. Inc.

Over the last five years, network topologies have changed so dramatically that the 80/20 rule no longer applies. In fact, we believe the operative state is actually the reverse of the 80/20 rule, where 80% of a user's traffic is long-distance (leaves the LAN), a state typified by the high Internet traffic growth over the last five years.

Exhibit 69. Phase II: 20/80 — The WAN Era



Source: Bear, Stearns & Co. Inc.

This shift in traffic has been a big reason for the network infrastructure buildout over the past five years. The movement of data from the LAN to the WAN has been the single most important principle behind network buildout: equipment follows data, and over the last decade, data has flowed outward.

We think this reversal (80/20 to 20/80) in traffic flow is now about to change. We believe there will be a resurgence in local traffic over the next five years, driven by internal streaming media, voice-over-IP, and higher data velocity LAN-based network services.

Exhibit 70. Phase III: 80/20 Became 20/80 Becomes 50/50

	Internal (I)	External (E)	Dataflow in I	Cause	Dataflow from I to E	Cause
Phase 1	80	20	Symmetric	Ethernet	Asymmetric	LAN-LAN No Web
Phase 2	20	80	Asymmetric	Server-based	Asymmetric	Web
Phase 3	50	50	Symmetric	Blended Server & Peer	Symmetric	Streaming Media VoIP Collaboration

Source: Bear, Stearns & Co. Inc.

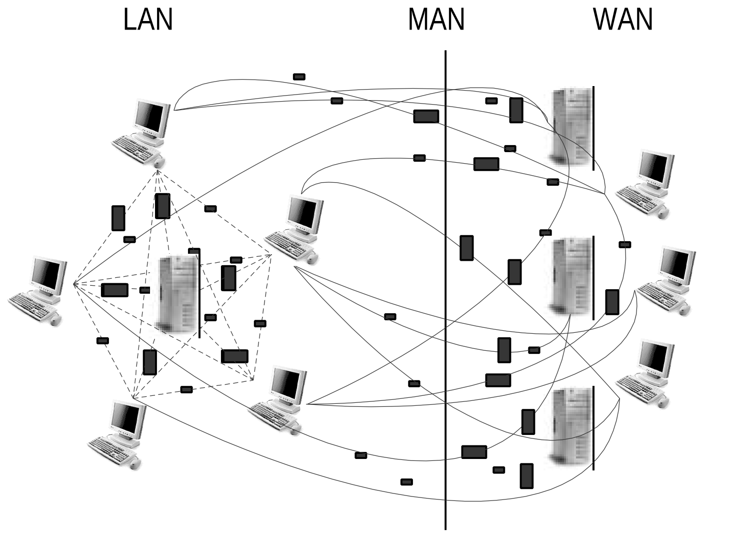
In Phase III (Internet 3.0), we assume the velocity of data continues to increase, not just on the Internet but in the intranet. We believe distributed LAN-based network services will help drive local traffic precisely because data flows more symmetrically in the next generation of LAN-based services.

The attractiveness of distributed network services lies in their ability to localize traffic — lowering bandwidth expenditures and improving response times. By serving and fulfilling as many requests for data from devices on the LAN, enterprises and ISPs can cut costs and improve performance dramatically.

The rise of the CDN and SMN validated the theory that content is better served locally. Within CDN/SMN networks, equipment and content migrate from the WAN to the MAN, from origin servers in datacenters to servers in COs, headends, and POPs. We believe content delivery will continue to migrate to the MAN and start penetrating the LAN. The reason for this is twofold: Smaller distances translate to lower latency, and local data calls are cheaper than long-distance data calls.

We imagine applications like streaming media and VoIP will contribute meaningfully to internal traffic. However, LAN-based network services will be the equalizing variable on a day-to-day basis. While Internet traffic will continue to grow, local traffic will grow faster, we predict, because much of the data served by the Internet is localizable.

Exhibit 71. 50/50 — Weaving the LAN, MAN, and WAN



Source: Bear, Stearns & Co. Inc.

CLASS OF INTERNET 3.0

Below is a list of the companies contributing to the distributed network services theme in Internet 3.0.

Exhibit 72. Distributed Network Services and Technology Providers

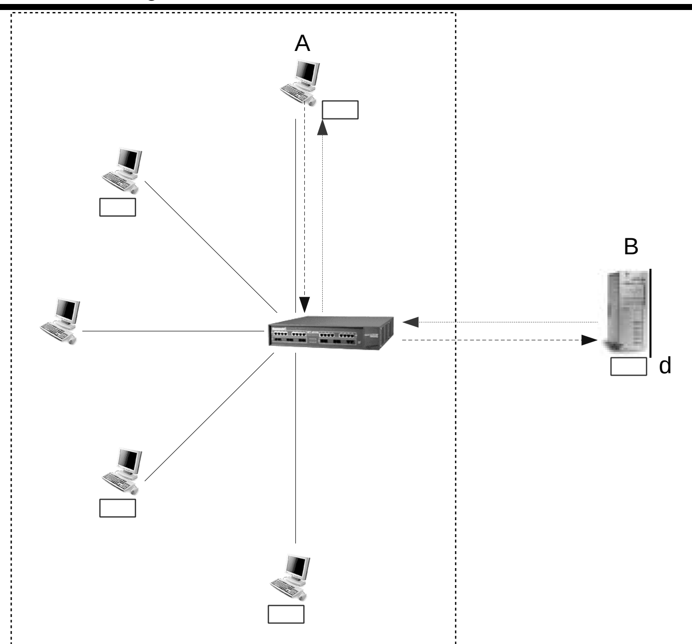
Company	Funding		Investors	Comment
3Path	$10 million	Series B	The BRM Group, Goldman Sachs, Intel Capital, BackWeb Technologies	
Applied MetaComputing	$6 million	Series A	Polaris Venture Partners	Several customers
Centrata	$5 million	Series A	Kleiner Perkins, Common Angels, dot EDU Ventures	Stealth
eMikolo Networks	$4.5 million	Series A	Israel Seed Partners	
Ejasent	$26 million	Series B	Crescendo Ventures, Crystal Internet Ventures, Red Rock Ventures, Technology Crossover Ventures, Bill Joy, BV Jagadeesh	
ExactOne	$4.5 million	Series A	JEGI Capital, SeaCap Ventures, Kaufman Family Partnerships	PartMiner, Gomez.com, Guru.com, among others
Infrasearch/Gonesilent	$5 million	Seed	Marc Andreesen, Angel Investors	Acquired by Sun
Napster	$15 million	Series C	Hummer Winblad Venture Partners, Angel Investors	Bertelsmann stake
OpenCola	$13 million	Series B	Battery Ventures, Mosaic Venture Partners, Torstar Corporation	
Proksim Software	Undisclosed		Société Innovatech du Grand Montréal, T2C2 Capital L.P.	Nortel partnership
Static	$5 million	Series A	Zone Ventures, Tim Draper	Several customers
Uprizer	$4 million	Series A	Intel, Kline Hawkes, Shugart Venture Fund	Ian Clarke, creator of Freenet
XDegrees	$8 million	Series A	Redpoint Ventures, Cambrian Ventures	
Zodiac Networks	Undisclosed	Series A	Kleiner Perkins, Benchmark Capital, The Barksdale Group	Marc Andreesen-led management

Source: Bear, Stearns & Co. Inc.

Publicly traded, Westborough, Massachusetts-based MangoSoft offers a shared browser-cache cluster software product called Cachelink that allows users on a LAN to serve browser-cached content to each other. We outline the advantages of this kind of system below.

Traditionally, if a user on a LAN wanted to retrieve a file on the Internet (despite the fact that other devices on the network may have a copy of the file), the file was delivered via a centralized server outside the network.

Exhibit 73. Getting a File from a Server Outside the LAN



Note: Computer A is looking for "d." Several people on his LAN have "d." Nevertheless, his request is sent beyond the LAN and served by B, the original server.

Source: Bear, Stearns & Co. Inc.

Each time a user requested a file on the Internet, the client device ran to the server housing the file and retrieved it, consuming bandwidth and suffering latency.

Once corporate enterprises began purchasing network caches, a client device on the LAN now had to go only as far as the local cache (in the LAN) to pick up content that would otherwise live on the Internet, since the local cache intelligently stores frequently-requested content.

Until now, this has been one of the only ways to speed up the retrieval of content and lower bandwidth costs at the same time, as network caches are certainly a cheaper solution than bandwidth.

The question has now become: Can we do away with having to purchase network caches for basic caching purposes? We believe the answer is "yes." We believe the next generation of network services will be distributed and will utilize installed devices to service the LAN.

We posit that content providers, service providers, and enterprises could all potentially benefit from distributed network services. Service providers could deploy LAN-based distributed network services to reduce bandwidth expenses. Currently, when an ISP subscriber requests a particular piece of content (say, a picture), the ISP will usually be able to serve the object from its cache farm (if the content is cacheable); this enables the ISP to serve content locally (otherwise, the ISP has to fetch the content from the particular content provider's origin server), which saves the ISP bandwidth expenses. When a content provider posts what turns out to be popular content, unless the content is offloaded onto a cache (in-house or on a CDN), the content provider will often have to purchase additional servers or suffer what has come to be known as the slash-dot effect (/.) or "flash crowding" — a situation that arises when requests for popular content slow the delivery of the content. Instead, content providers could tag certain objects and have them served to end users by peer devices; this would in effect decrease the capex of the content provider and decrease latency for the end user. Enterprises spend a tremendous amount on bandwidth, as users' requests for content are served by either the cache or the content provider's origin servers. Instead of having to purchase caches or spending more on bandwidth, enterprises and service providers could implement LAN-based distributed network services technologies that would enable content to be served by peers.

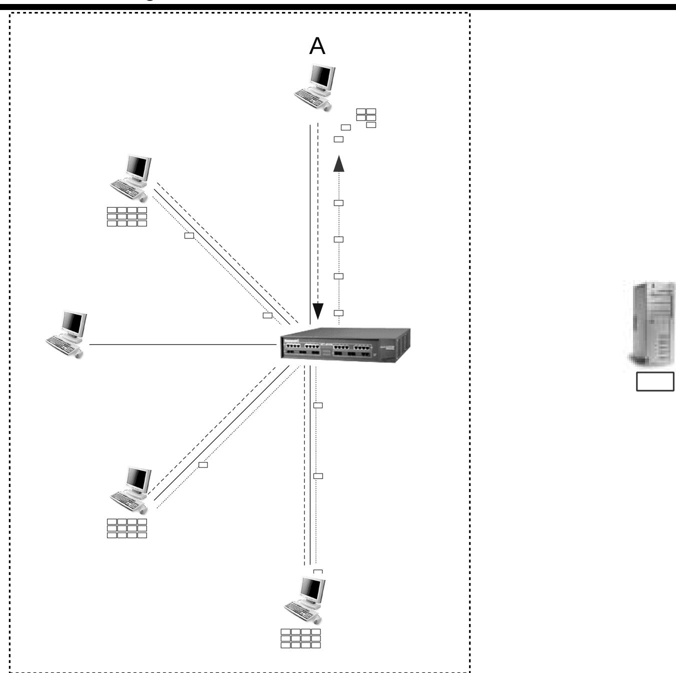
We outline a handful of companies offering distributed network services below.

OPENCOLA

It's in There

The goal of OpenCola, a privately held distributed network services technology provider, is to enable networks to search and share data assets more effectively. Should a user be looking for a particular file on the Internet, OpenCola's Swarmcast solution enables the file to be served by the user's peers on the local network by breaking and then splicing portions of the file from peers. This way, the request does not leave the local network, which results in bandwidth savings (as local bandwidth is cheaper), and the request for the file is answered more immediately (low latency).

Exhibit 74. Getting a File from Peers Within the LAN



Note: This time, when Computer A is looking for a file "d," instead of being forced to download the file from a server outside the LAN, the request is answered by multiple peers, each of whom serves a portion of the file. Computer A then splices the fragments to generate the file.

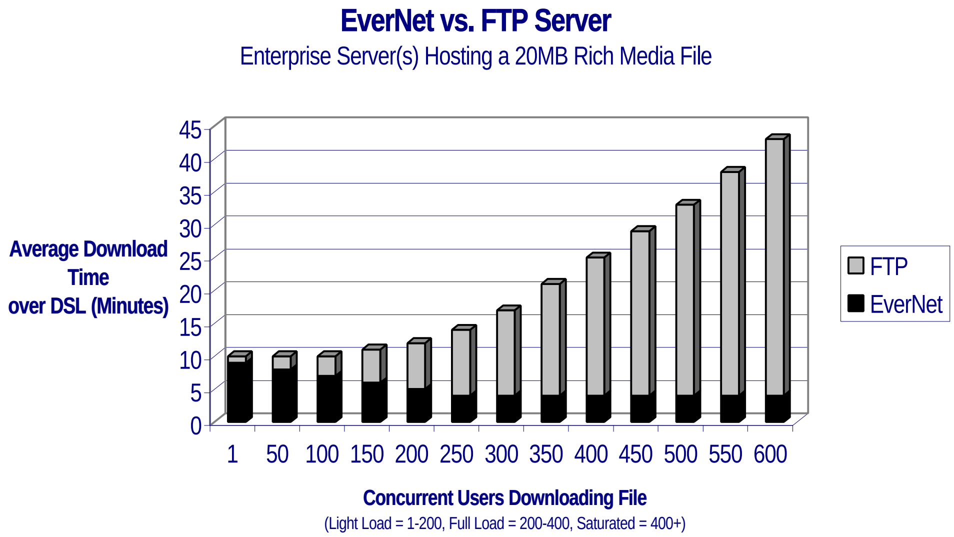
Source: Bear, Stearns & Co. Inc.

In essence, OpenCola acts as a cache cluster. Unlike traditional cache server farms, however, OpenCola enables peer devices to behave like caches and serve portions of an object to another peer.

We believe the value proposition that OpenCola and other distributed network service providers offer becomes more interesting as richer content saturates networks. One ideal use of OpenCola's Swarmcast technology would be to serve the next *Star Wars* trailer. Should the trailer be 20 MB in size, OpenCola's Swarmcast could allow users in a LAN to serve one another, instead of forcing users to go to Apple QuickTime's Web site to download the official trailer (which is delivered by Akamai).

According to EverNet Systems, a privately held distributed network service technology provider, download times for a 20MB file fall dramatically when peer devices serve one another.

Exhibit 75. Accelerated File Transfer



Source: EverNet Systems, Inc.

Serving data from local storage is more cost effective than expending bandwidth to fetch content from origin servers because disk is cheaper than bandwidth.

Exhibit 76. Price per Gigabyte (GB) of Disk and Bandwidth

Disk		Bandwidth	
		1.544	Mbps
		60	seconds/minute
		60	minutes/hour
		24	hours/day
		30	days/month
		4,002,048	Mbits
		8	Mbits/MB
		500,256	MB
		1,000	MB/GB
50	GB	500	GB
$100	per 100 GB	$1,500 - $2,000	cost of T1/month(a)
$2	**per GB**	**$3 - $4**	**per GB**

Note:

Disk — We took an average price of disk from several suppliers.

Bandwidth — We arrive at our cost per GB transferred by calculating the total number of GB transferred on a T1 operating at full capacity for a month. T1's cost around $1,500 per month, on average.

We believe the disparity between disk and bandwidth prices per GB is understated. Using disk in a LAN (as in caching), the same file can be served innumerable times at a fixed cost of disk. With bandwidth, each transmission of the file consumes bandwidth.

(a) Does not include setup fees.

Source: Bear, Stearns & Co. Inc.

Storage prices have been falling quickly, and the steep drop historically in disk prices is enough to make it a cheaper alternative to bandwidth whose prices, while declining, remain much higher per Gigabyte than disk. Storing (duplicating) data is significantly cheaper than delivering data using bandwidth.

STATIC ONLINE

Streaming in Bits

Static Online takes OpenCola a step further. Static Online has been focused on splicing portions of streaming media files (.mpeg, .wav, .mov) live or on-demand. This is particularly difficult to do since, unlike static files, the order of reception of packets of a streaming media file can be critical to its quality.

As one of the only companies with working technology in this area, Static Online has piqued our interest. We believe that should Static Online successfully map distributed multimedia files to a single user, the company could potentially add significant value to enterprises and ISPs, and save customers considerable bandwidth expenses.

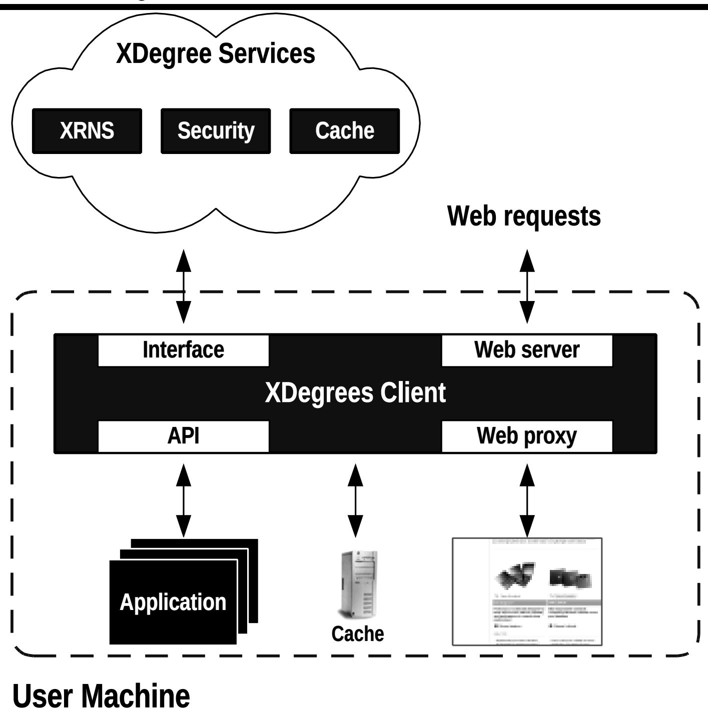
XDEGREES

The X-Factor

Mountain View-based XDegrees is a platform provider for distributed network services. Backed by Redpoint Ventures and Cambrian Ventures, XDegrees provides its core eXtensible Resource Name System (XRNS) technology to allow identification, addressing, security, and access to resources distributed throughout networks.

XRNS is a name system, much like the DNS, that enables a network to tag all items on a network (files, machines, applications, devices, and people) as objects. In so doing, XDegrees powers networks and their resources to be named, stored, directed and searched.

Exhibit 77. XDegrees Architecture

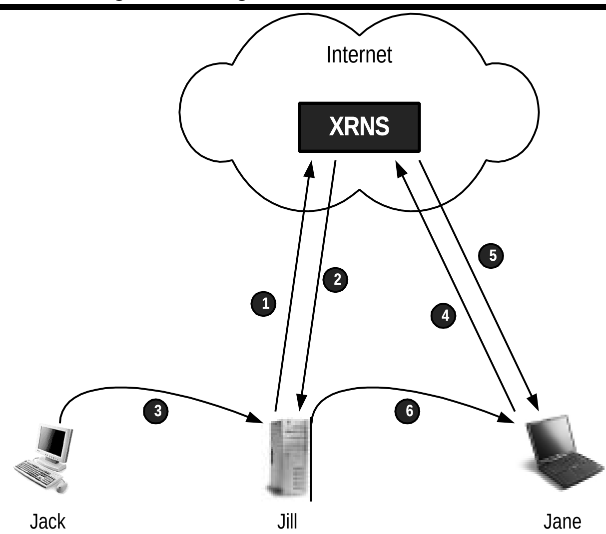


Source: Company data.

Naturally, as a platform vendor, XDegrees offers the foundation and tools to build distributed network applications, with an underlying focus on scalability (hundreds of millions of “nodes”), performance, and reliability.

Below, we articulate how enterprises can use XDegrees to create a file-sharing system, one of many applications possible with XDegrees' XRNS.

Exhibit 78. Sharing Files Using XRNS



Note: This shows how a collaborative Network Application would allow the salesforce of Acme Corporation to easily share documents. Jack shares a price list with Jill and Jane at an XRNS URL (e.g., http://www.acme.com/sales/price_list.doc). Jill's application requests this URL (1) and XRNS returns a pointer to the file on Jack's machine (2). Jill's application then retrieves the file from Jack's machine (3).

Suppose that Jack turns off his machine so it is no longer available. Jane's application requests the same URL from XRNS (4) and XRNS returns that the document has been cached on Jill's machine (5). Jane's application seamlessly retrieves the file from Jill's machine (6).

Source: Company data.

We believe XDegrees is one of the few platform vendors whose focus and technology will likely permit it to gain traction in the distributed network services market.

EJASENT

The Computing Utility — "3 AM Performance, 24 Hours a Day"

Privately held Ejasent, located in Mountain View, California, and backed by Crescendo Ventures, Crystal Internet Ventures, Red Rock Ventures, Technology Crossover Ventures, Bill Joy (Sun Microsystems), and BV Jagadeesh (Co-founder of Exodus Communications), among others, offers what we regard as the next generation of outsourced computing services.

First, Web hosting companies like Exodus offered hosting of complete Web sites (applications, content, databases) on an outsourced basis, across a limited distributed network. While Web hosters are good at hosting applications, they are poor at distributing them. CDNs like Akamai began offering hosting of specific content objects across a highly distributed network. While CDNs are good at distributing content, they are poor at hosting applications.

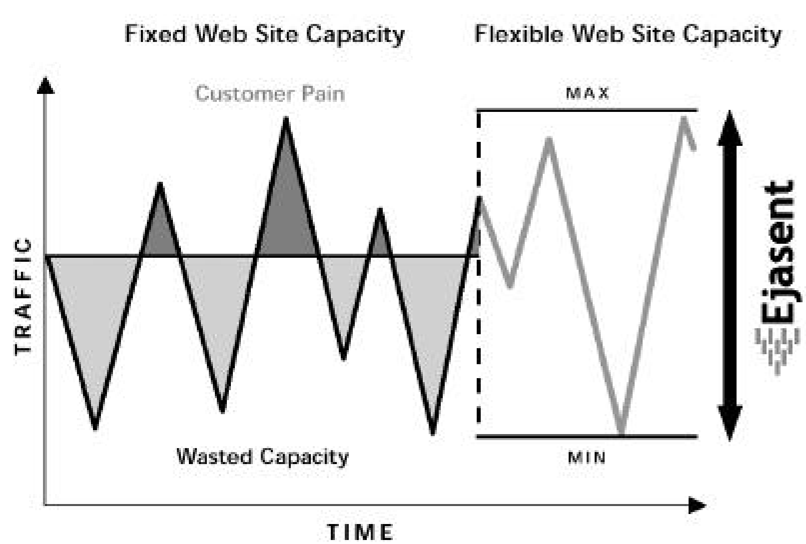
As a computing utility provider, Ejasent is attempting to resolve these limitations by distributing applications processing.

Ejasent offers application processing as a service across its distributed network to offer a “computing utility” that provides “elastic and infinite processing capacity instantaneously.” Clusters of Solaris machines are woven in a fabric across the country. On these machines run applications like BEA Systems Weblogic and BroadVision.

Imagine if Amazon were to announce a 25% discount on books. The incredible traffic spike at Amazon could potentially tax Amazon’s Web servers to a point that they begin to fail. What is the solution to this? Right now, Amazon would have to use load-balancers and caches to manage traffic and purchase additional servers to augment processing capacity. However, server processing power remains a constant — that is, there is a threshold after which servers (software) can no longer process additional requests.

Enter Ejasent. When the next user reaches Amazon’s Web site, the request is redirected to Ejasent’s racks of Solaris machines, running Amazon’s application server systems (Ejasent racks duplicate Amazon’s Web processing systems). Ejasent’s machines process the user’s request on the fly (Amazon’s database is still used, just the processing of the request (i.e., the serving of the pages) is done by Ejasent). This utility offloads a significant amount of processing load from Amazon’s network of Web application servers. Amazon pays Ejasent on a pay-as-you-use basis. (Note: Amazon is not an Ejasent customer.)

Exhibit 79. Ejasent’s Dynamic Computing Utility



Source: Company data.

While incumbents like IBM, Oracle, and HP have all announced computing utility offerings like Ejasent’s, we believe Ejasent (with Sun) is unique in offering a computing utility for applications processing.

The two important technical innovations involved in the service are application switching (whereby a server running one application for Customer A needs to switch to running another application for Customer B) and applications-to-machines separation (where each application runs on a separate machine; in fact, Ejasent has been working closely with Sun in order to posit this kind of tightly-coupled functionality directly into the OS).

Some believe Ejasent's competitors include publicly traded Managed Service Provider (MSP) LoudCloud and privately held MSPs Logictier and Totality (formerly MimEcom). We do not believe Ejasent should be included in this category, particularly because, unlike MSPs, Ejasent's capital structure (virtualized resources shareable by multitudes of customers on the same infrastructure/equipment) is extremely light. By virtualizing the processing power of Solaris machines, Ejasent is able to serve any number of customers on the same machine. (MSPs generally create a separate infrastructure for each customer.) We believe MSPs and storage service providers (e.g., LoudCloud and StorageNetworks) are beginning to offer cap-lite computing utilities.

EJASENT UPSCALE

Products and Services

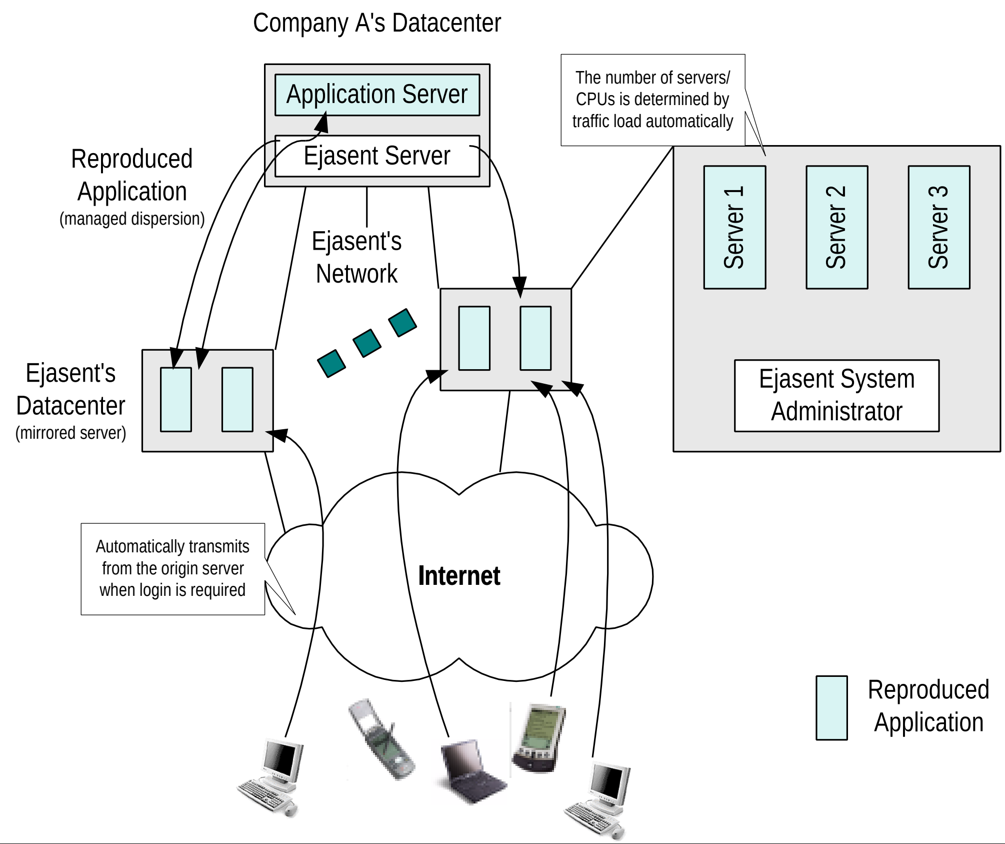
UpScale offers on-demand Web application processing capacity.

- **Ejasent Application Processing Network (APN)** responds in real-time to surge in a customer's traffic load. As the traffic load on a Web site increases and user response times start to degrade, the site can add additional application processing capacity transparently and instantly (within three seconds). This enables the customer to adaptively increase or decrease the processing power necessary to assure site responsiveness regardless of the traffic load or the geographic location of the Web site.

- **No Additional Capital Expenditures.** Web site operators no longer need to invest upfront in hardware to meet their "peak load" requirements. Ejasent UpScale provides processing capacity for all traffic conditions including peak loads and flash crowds, and permits Web-based companies to pay for only the resources it consumes. The result is increased savings in Web site infrastructure including hardware, software, site hosting, and maintenance.

- **Ejasent Instant Application Switching.** Ejasent's patented Instant Application Switching technology enables Web site managers to take a "snap shot" of their Web applications and host these on Ejasent servers placed around the Internet. These AppShots are identical instances of the applications running on the central site. When loads on the central site exceed pre-defined thresholds, Appshots are "activated" in physical locations closest to the points of consumption. This "activation" process takes approximately one second and is executed automatically. The Appshots maintain data synchronization with the central site over a secure, authenticated, and encrypted connection. As traffic loads recede, AppShots are progressively "retired." Applications in the APN are not tied to any particular server. Instead they are scheduled "on-demand" and run on any of the thousands of available processors, thus creating a "virtual single server."

Exhibit 80. Ejasent's Network



Source: Company data; Nikkei Internet Technologies.

Market Opportunity

We believe our forecast for content delivery and streaming media network services is an appropriate proxy for the distributed network services market. The market for Internet content delivery solutions is poised to experience tremendous growth due to secular trends in Internet usage and increasing amounts of rich data travelling over the Web. We estimate the addressable market for Internet content delivery (including streaming media) could reach $7.3 billion in 2004 (Exhibit 81).

Exhibit 81. Worldwide Content Delivery/Streaming Media Network Market Opportunity

	2000E	2001E	2002E	2003E	2004E
Worldwide Internet Users (m)	328	427	521	602	687
Growth		*31%*	*22%*	*16%*	*14%*
x % Active Internet Users	50%	50%	50%	50%	50%
= Worldwide Active Internet Users (m)	164	214	261	301	343
x Duty Rate (Users on at a particular time)	15%	15%	15%	15%	15%
= Internet Users (m)	25	32	39	45	52
x Average Bandwidth per User (Kbps)	159	176	199	234	288
= Peak Bandwidth (Kbps)	3,915,594,334	5,638,443,344	7,769,045,898	10,564,789,679	14,855,159,174
=Peak Bandwidth (Mbps)	3,915,594	5,638,443	7,769,046	10,564,790	14,855,159
x % Traffic is HTTP	50%	50%	50%	50%	50%
= Peak Web Bandwidth Utilization (Gbps)	1,958	2,819	3,885	5,282	7,428
Growth		*44%*	*38%*	*36%*	*41%*
x Average Cost per Mbps	1,500	1,350	1,215	1,094	984
= CDN/SMN Addressable Market Opportunity ($ in millions)	2,936.7	3,805.9	4,719.7	5,776.3	7,309.9
Assumed Market Penetration	5%	6%	7%	8%	9%
Implied CDN/SMN Market Size ($ in millions)	144.5	225.3	326.6	457.4	652.0

Source: Bear, Stearns & Co. Inc. estimates.

IMPACT ON OTHER INDUSTRIES

- **Content Delivery/Distribution Networks.** We believe one of the advantages of using services like Akamai's is in the company's focus on serving "heavy" objects (pictures, multimedia files, etc.). We envision a time when heavy objects can be served from devices in a LAN. Since client-server devices on the LAN (PCs, servers, workstations) are already installed (and hence their cost is sunk), tapping into them to offer network services will likely become commonplace. We believe what will likely happen is that distributed network services providers like Akamai will recognize the larger enterprise opportunity and offer LAN-based services using technologies like Static Online's.

- **Layer 4-7 Vendors.** The imperative for companies to manage network performance and cost is driving the development of higher-layer technologies. We believe many of these young companies will be offering competing software products. By utilizing the installed base of devices on the LAN, corporations may discover that the efficiencies and cost savings associated with layering intelligent software on top of existing devices may outweigh the advantages of purchasing altogether separate Layer 4-7 devices.

- **LAN Network Equipment Vendors.** LAN-based network equipment could experience increased demand if the next turn of the Internet is inward toward the LAN. Moving from the WAN to the LAN traverses the MAN. We therefore believe datanetworking equipment providers focused on the MAN and LAN could see greater demand for their products.

- **LAN Management.** Increased LAN development and traffic should drive adoption of LAN management products and monitoring tools.

"What will on-line interative communities be like? . . . They will consist of geographically separated members . . . communities not of common location, but of common interest. . . .The whole will constitute a labile network of networks — ever changing in both content and configuration. . . . the impact . . . will be very great — both on the individual and on society. . . . First, . . . because the people with whom one interacts will be selected more by commonality . . . than by accidents of proximity."

J.C.R. Licklider and Bob Taylor, "The Computer as a Communication Medium"

Quoted by David Reed in the *Jounal of the Hyperlinked Organization*, January 19, 2001.

In 1984, Ray Ozzie introduced Lotus Notes, a revolutionary collaboration platform. Since then, Lotus Notes has been at the center of collaborative applications development. Since the introduction of Lotus Notes, only Microsoft Outlook (introduced in 1997) has offered anything for collaboration and groupware on a mass enterprise level.

One reason collaboration software has not successfully penetrated enterprises is because collaboration software has not yet met our four criteria for adoption: ease of use, utility of use, cost of use, and necessity of use. New collaboration software, while not necessarily expensive, is an additional expense that is often obsoleted by collaboration programs like Lotus Notes (platform) and Microsoft Outlook (application). In addition, collaboration software requires users to learn new interfaces and rules, and forces users to change everyday workflow behavior. Nevertheless, collaboration software has utility, especially for organizations with distributed workforces or large numbers of workgroups. For these reasons, the adoption of collaboration software across enterprises has been spotty at best.

We believe this is changing. Three factors have reset the collaboration software debate: instant messaging, file sharing, and decentralization. The wide adoption of instant messaging (like ICQ) has validated it as a legitimate communications tool; the popularity of file-sharing networks like Napster has proven that file-sharing across a distributed network of participants is possible and introduces significant advantages. The popularity of instant messaging and file-sharing has also proven that users are willing to learn new user interfaces and rules if the benefits are greater than the pain of learning to operate the new system.

We believe the potential success of collaboration reflects people's preference for the option to communicate in real time — witness the relatively instant popularity of instant messaging.

We see decentralized collaboration — i.e., generally not hosted in a purely client-server architecture — as most likely to gain traction. Our assertion stems primarily from economics: The economics of running a decentralized collaboration system are better than a traditional client-server one. Within a centralized collaboration system, server-side hardware and software are required to integrate client-side software and devices. Within a decentralized collaboration system, there is virtually no server-side hardware or software. The client device requires software, and the system operates from client device to client device, often unmediated by expensive server hardware and software.

The vision of decentralized collaboration is to increase the velocity of communications between and among people in workgroups, across chains. By increasing the velocity of communications, it is believed, transactions velocity will follow.

Messaging Revisited

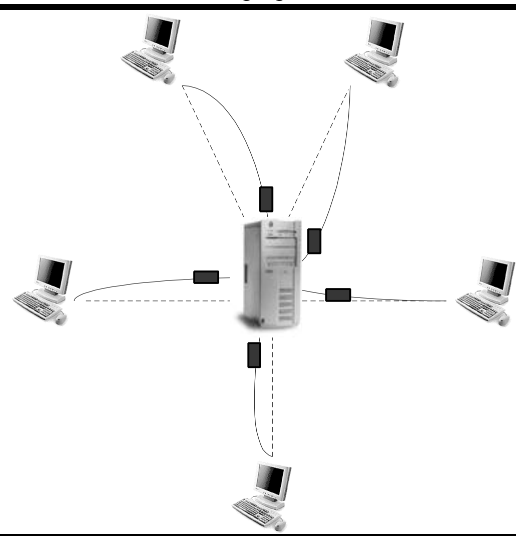
Imagine a network where e-mail is decentralized. What would it look like? E-mail would be created and stored on client devices. There would be no central server to buy to host the e-mail and no central server to crash.

The closest thing we have to decentralized e-mail is instant messaging. To the user, instant messaging, a decentralized messaging system, differs from e-mail primarily in its ability to deliver messages in real-time. Architecturally, instant messaging networks do not require the massive infrastructure costs associated with e-mail networks, because there is no centralized server/storage (unless the network chooses to store messages). Since there is no central intermediating server in instant messaging, it is extremely difficult to bring down the entire network.

The only aspect of a decentralized e-mail system that requires centralized assets is a registry server to resolve user IDs and IP addresses. This means that the network cannot fail unless the central registry server (which is extremely light) should crash.

In a centralized messaging system, the server is the funnel from which messages come and go. The centralization of the messaging infrastructure at the server means the system fails with the server. Once the server fails, the messaging infrastructure collapses. Centralization also means that any edge device could theoretically collapse the central server — that is, should one PC introduce a "crash" variable into the server, the remaining PCs fail automatically with the server.

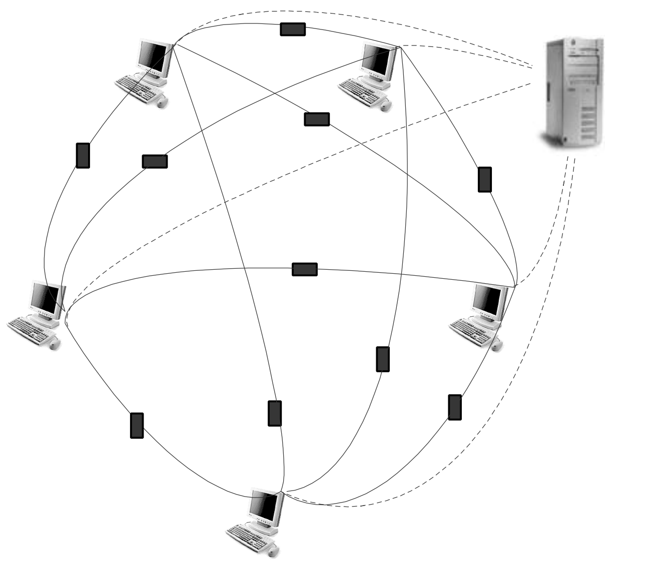
Exhibit 82. Centralized Messaging Infrastructure



Source: Bear, Stearns & Co. Inc.

In a decentralized messaging system, the server functions "outside" the messaging path — that is, the registry server functions only as a mapping machine, mapping user IDs to IP addresses. The messaging occurs directly between devices. Messaging is much more fluid in this scenario, and the only way the entire system can fail is in the event of registry server failure. The chances of the registry server failing are low, and message load will not cause the registry server to crash since the registry server does not handle e-mail data. Again, the registry server only handles mapping of user IDs to IP addresses.

Exhibit 83. Decentralized Messaging Infrastructure



Source: Bear, Stearns & Co. Inc.

Messaging in a decentralized system is much more organic, as messages are allowed to flow and move autonomously from user to user. This dynamic network connectivity represents an evolved messaging system. It maximizes Metcalfe's Law and frees the server (and hence the messaging system) from constantly having to break through its performance bottleneck.

There are a few dozen decentralized collaboration software and services providers.

Exhibit 84. Decentralized Collaboration

Company	Funding		Investors	Customers
Consilient	$2.8 million	Series A	Oak Hill Venture Partners, The Sapling Foundation	
Engenia	$22.7 million	Series C	Cooley Godward, Dominion Ventures, Intel Capital, SpaceVest, Thomson Corp., Topaz Investors, Winfield Capital Corp., St. Paul Venture Capital, Novak Biddle Venture Partners, Vanguard Atlantic, Aurora Funds	PetroVantage, Thomson's, Coherence Ventures
Groove Networks	$41 million	Series B	Accel Partners, Intel Capital	GlaxoSmithKline, Raytheon, U.S. Department of Defense
IKimbo	$6.375 million	Series B	Cross Atlantic Capital Partners, PTEK Ventures, Draper Atlantic, Steve Walker & Associates	PricewaterhouseCoopers
NextPage	$20 million	Series B	Oak Investment Partners, epartners, Dominion Ventures, Amp Capital Partners	ABC Television Network, ABN AMRO, Deloitte & Touche UK, Thomson Learning, West Group
QUIQ	$15 million	Series B	InterWest Partners, BancBoston Ventures, Altos Ventures, Discovery Ventures	National Instruments, Quaero, eChips, AlphaSmart, Network Appliance, Packtion

Source: Bear, Stearns & Co. Inc.

There are various offerings in play in the marketplace. Some, like IsoSpace, offer browser-based collaboration platforms that leverage existing systems (XDegrees introduced a Microsoft Outlook-based file-sharing system back in February 2001). Here, we provide a single case study on decentralized collaboration. We profile Groove Networks, as we feel Groove encapsulates most of the features offered by other decentralized collaboration vendors.

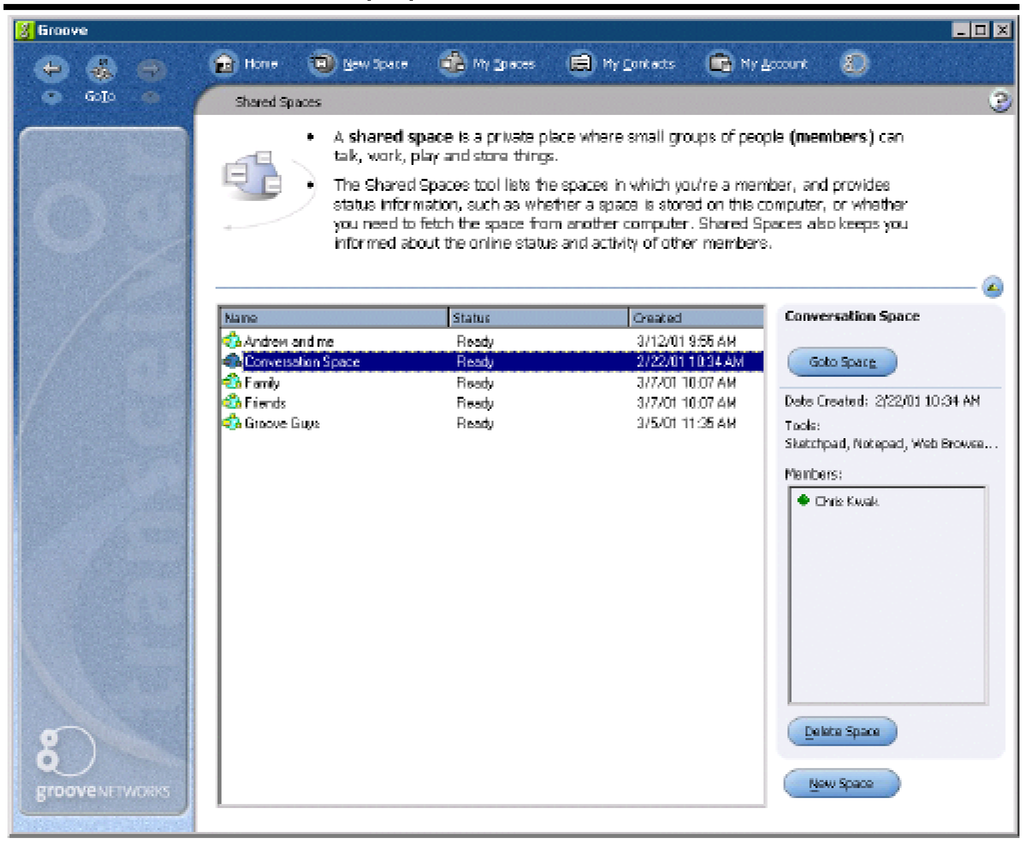
GROOVE

Groove Networks, the brainchild of Lotus Notes visionary Ray Ozzie, demonstrates, in our opinion, some of the best qualities of Internet 3.0. Groove maximizes decentralization, resource utilization, and velocity of data and transactions.

Groove is a platform for the development of peer-to-peer collaboration applications. As such, Groove offers a basic toolset that includes file sharing, instant messaging, calendaring, and co-browsing, among other functions. As a platform, Groove is a base on which developers can build applications.

The fundamental idea of Groove is the Shared Space (SS) — a space that collaborators share. It is private and viewable only to those who have agreed to participate in the SS. Participants in an SS can collaborate via a number of predefined tools and applications, and via an unlimited set of applications that developers can create.

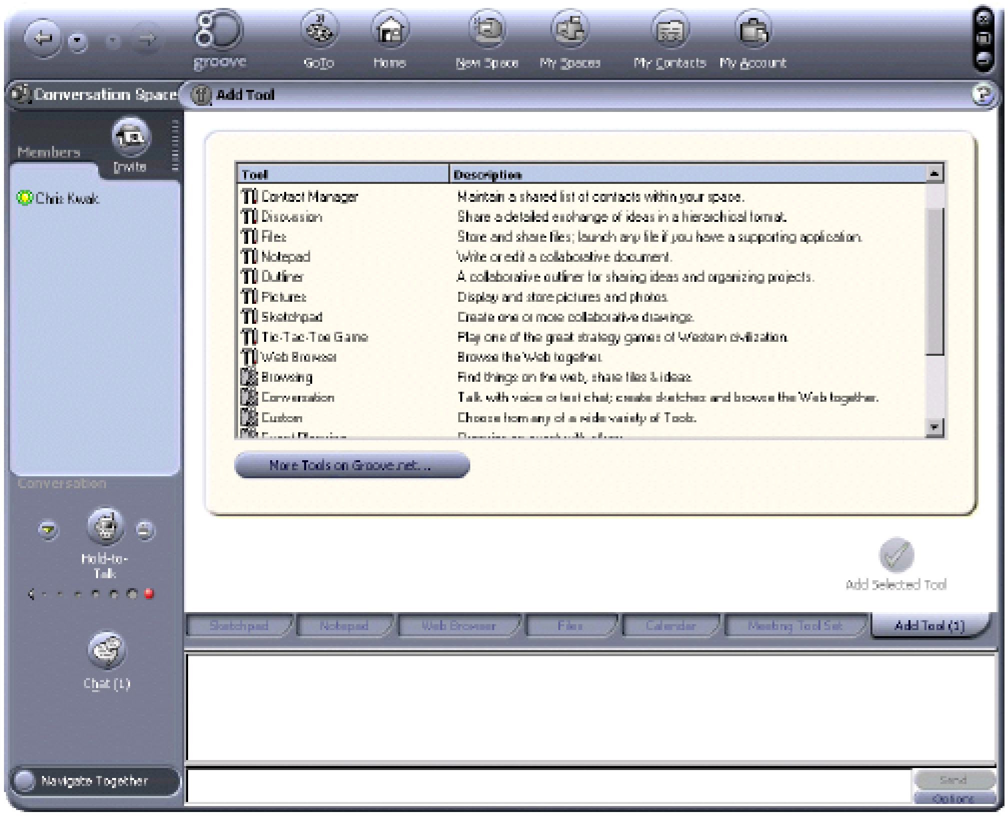
Exhibit 85. Groove User Interface (GUI)



Source: Groove Networks; Bear, Stearns & Co. Inc.

Pre-loaded tools include: contact manager, discussion space creator, file manager, notepad, games, co-browsing, conversation, calendar, etc. Exhibit 86 below highlights some of the other tools and their descriptions.

Exhibit 86. Current Groove Toolset

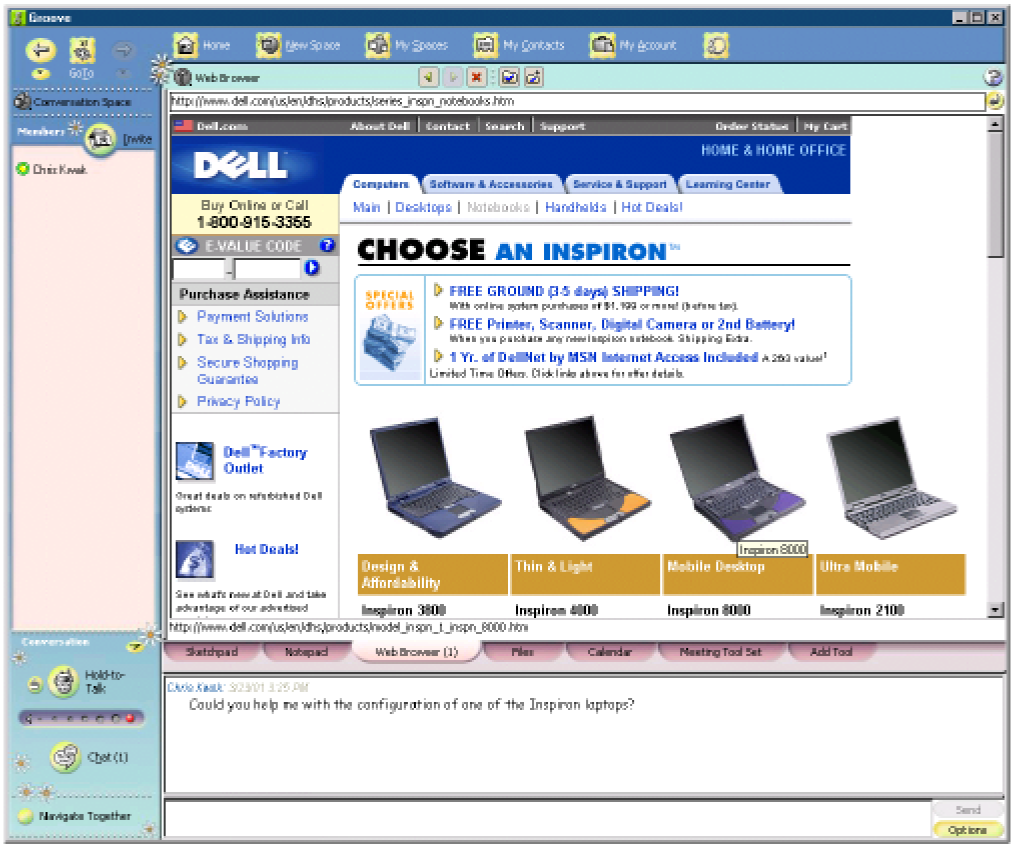


Source: Groove Networks; Bear, Stearns & Co. Inc.

Groove also offers conferencing capabilities, including voice options. Participants in an SS can hold down a “Talk” button to speak over the Internet directly with peers.

Within an SS, users can communicate via real-time chat (instant messaging), connect to e-mail applications, view presentations, and exchange files on a peer-to-peer basis. Users can also navigate the Web together by co-browsing using Groove. We believe this feature, while not new, offers the potential for improved sales and customer support applications.

Exhibit 87. Navigating Together on Groove



Source: Groove Networks; Bear, Stearns & Co. Inc.

Groove is almost always a peer-to-peer application. That is, messages and files are shared directly, unless users are temporarily off-line, files being transferred are too large, users are behind firewalls and require intermediate storage stops, or the number of recipients of a message or file is large. In these cases, Groove utilizes a relay hub[18] to coordinate data flow among devices.

Groove can also be used to create virtual private networks (VPNs) of peers. This is possible with Groove because the identities of users on Groove are authenticated by digital signatures (encrypted keys that detail the identity of users on the network) that reside on the client device and on a central registry server. Each user in a Groove Shared Space connects and is permitted to collaborate with others within the Shared Space by exchanging digital signatures. It is possible with such a strongly encrypted peer system to create an extremely elegant and easy-to-use VPN across the public Internet. With Groove, PKI (Public Key Infrastructure) is obviated by the presence of exchangeable identity markers (digital signatures) on the Groove client.

We believe Groove has distinguished itself from competitors (apart from the deep technology and intuitive user interface) by offering its product as a platform for decentralized collaboration. This means Groove is not simply an application like instant messaging or file transfering. Instead, developers and corporations can

[18] 3Path is an example of a relay service provider.

introduce brand new applications that ride on top of Groove. Groove recently announced that it has already signed partnership agreements with more than 200 companies who intend to develop, deploy and service business solutions based on the company's platform.

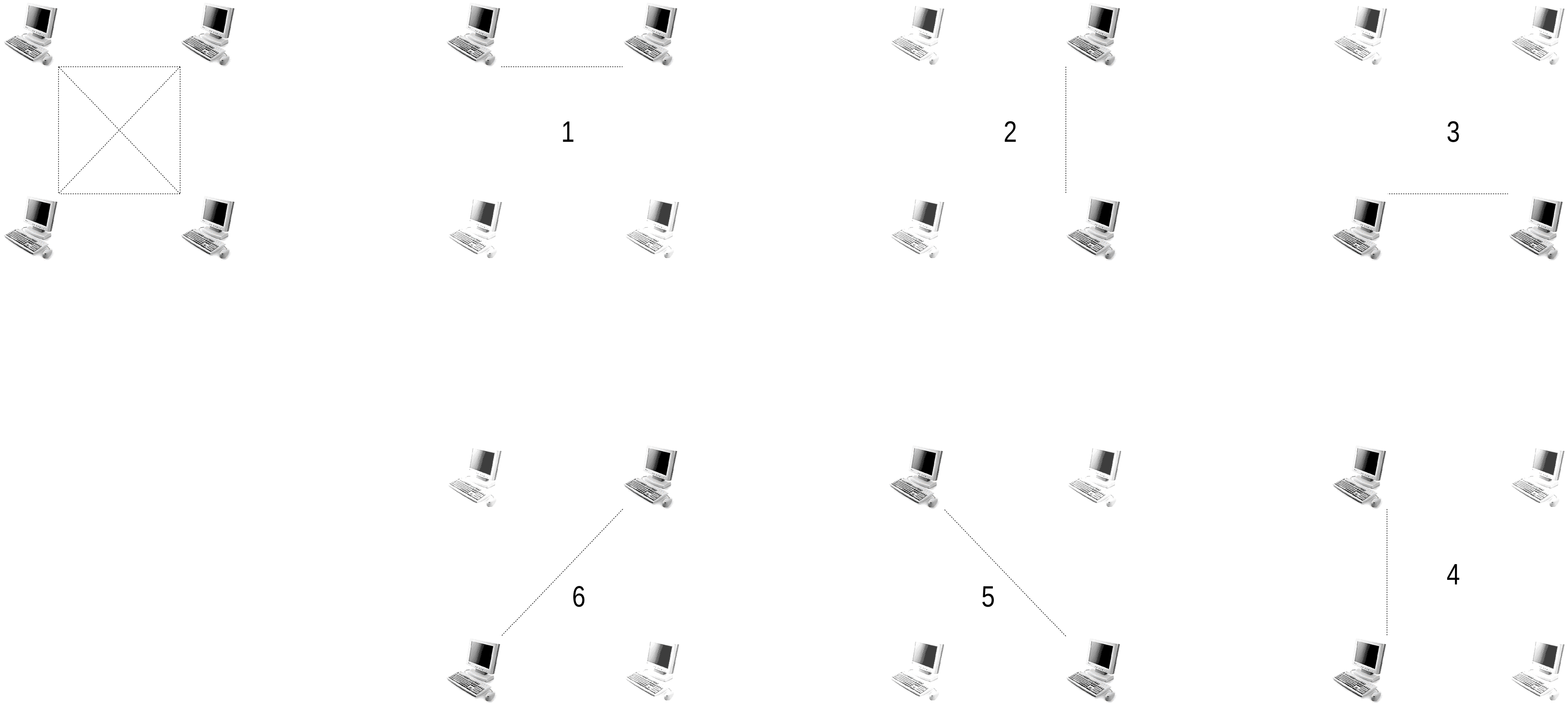
For example, a developer could easily create an API (application programming interface) linking Groove to Microsoft Outlook. Corporations like Ariba could potentially use Groove to create a system linking customers. Dell could create a high-touch customer support front-end that ties into its customer databases using Groove.

REED'S LAW

The fundamental theory behind Groove in particular and decentralized collaboration in general is that unlike Metcalfe's Law, where the number of connections between devices is the square of the number of devices, Reed's Law describes more accurately the way people communicate and collaborate.

According to David Reed, a former research scientist at Lotus, Metcalfe's Law does not capture the power of connections and groups. Metcalfe's Law primarily describes two-device connections and does not capture the potential number of combinations of groups within a network. If there are four devices on a network, the number of potential physical connections is $(N^2 - N)/2$, which equals the number of two-device connections possible in the network. However, permutations of connections *among* objects (not just *between* objects) make the potential number of groups even larger. So, in the case of four devices, we could have each device connecting to each other device (Metcalfe's Law) or grouping with one or other device(s) (Reed's Law).

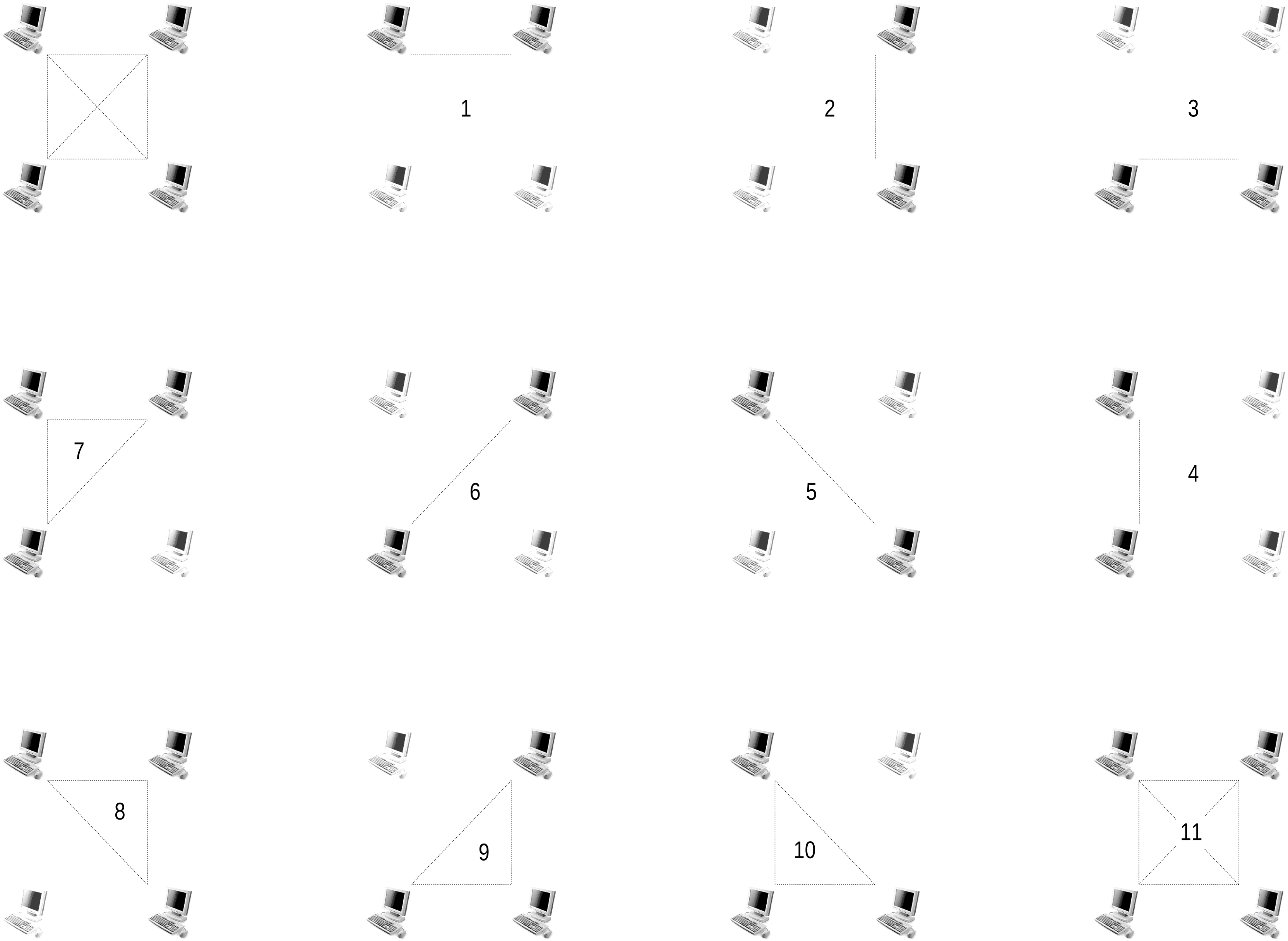
Exhibit 88. Metcalfe's Law with Four Devices



Note: Metcalfe's Law measures the number of potential combinations of coupled pairs. In the case of a network with four devices, $(N^2-N)/2 = 6$. In the event that we are considering two-way connections, the number of potential connections would be [(N2 - N)/2]*2 = 12.

Source: Bear, Stearns & Co. Inc.

Exhibit 89. Reed's Law with Four Devices



Note: Reed's Law helps us measure the number of possible groups in our four-device network. As such, it includes the number of potential connections *between* devices under Metcalfe's Law and the number of potential groups *among* devices. The number of potential groups under Reed's Law equals $2^N - N - 1 = 16 - 4 - 1 = 11$

Source: Bear, Stearns & Co. Inc.

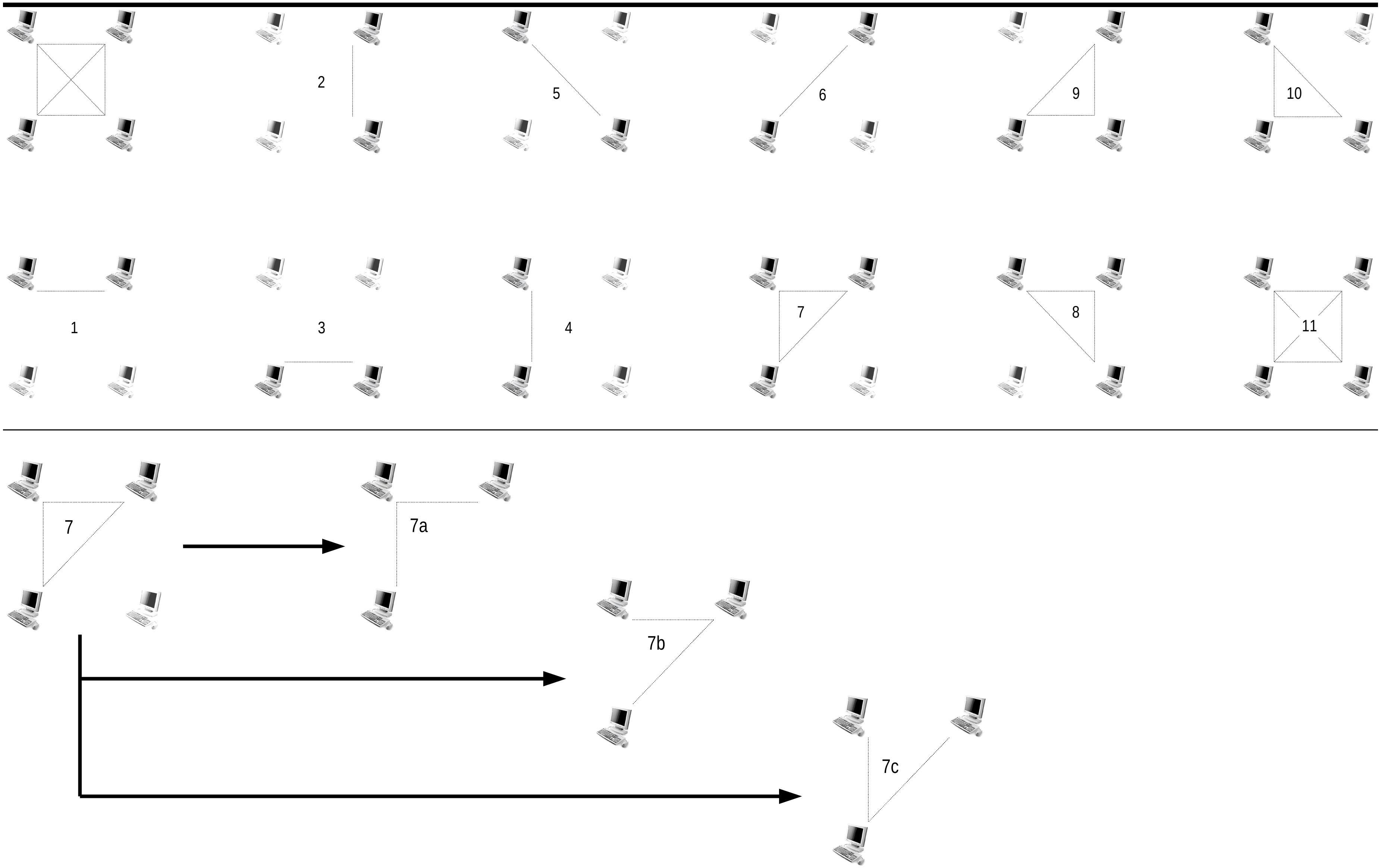
We refer to these formed groups as *clusters*. We believe taking Reed's Law one step further to take into account storage capacity and the number of data objects (the ultimate goal of the network is to share data) increases the number of useful connections, now in reference to connections between data objects and not just devices:

$$2^{N(B)}$$

B equals the number of data objects at a node N (i.e., device). This means that the number of potential combinations with N devices should be expanded to the number of potential combinations of B data objects within N devices across the network.

We believe the number of clusters is actually larger than the number of groups described by Reed's Law. While Reed's Law describes open clusters (i.e., clusters where each person is connected to every other person in the group), we think the actual number of clusters should include those scenarios in which clusters are closed. In a scenario with four people, Reed's Law indicates that a group is defined as a subset of the four individuals (4, 3, 2) in which all the members are directly connected to one another. However, we believe the total number of clusters should include subsets where each member is not necessarily directly connected to one another.

Exhibit 90. Revisiting Reed's Law



Note: We submit that the actual number of clusters is larger than the number of groups under Reed's Law, because clusters can have broken links while groups cannot.

Source: Bear, Stearns & Co. Inc.

Here is an exercise that reveals the power of such clustering: Imagine a conference call with 50 individuals from Company A and 50 from Company B to talk about A's potential acquisition of B. Under Reed's Law, in order for everyone to participate, all 100 individuals should be visible to each other. However, we believe under such an environment, the individuals from A should be able to isolate themselves within the 100 and communicate only with colleagues from A. Likewise, members of B should be able to insulate themselves from A's view.

Another possibility is in a negotiations scenario. Imagine two companies are negotiating a deal. A has 10 individuals on the "call" and B has 20 individuals on the call. As the head negotiators are communicating with one another, members of A should be able to communicate with one another without the individuals from B overseeing their communications.

We believe Groove is a perfect example of how the maximization of a network law (in this case, Reed's Law) is a powerful explanation for why the technology will be adopted.

While most people might regard peer-to-peer collaboration software as the province of commercial users, it will most likely be adopted first by corporations using it across the intranet. One reason is that collaboration software requires a critical mass of users to be especially useful. Corporations that boast a large number of workgroups would require team members to standardize their collaboration on Groove. Outside of the intranet, entities who are part of supply and demand chains may be slower to use collaboration software because the true value of collaborating among partners is achieved when all participants have adopted the software.

Exhibit 91. Processes Supported by Supply-Chain Exchanges — Collaboration Impact

Business Processes	Business Activities	Benefits
Supply-chain collaboration among external business partners	Demand forecasting, inventory checking and replenishment, order management, promotion planning, transportation planning, etc.; self-service activities such as payment or order status checking, among others	Reduction in inventory, faster time to market, more efficient collaboration with suppliers and customers, greater customer satisfaction
Collaborative design and project management	Access to component catalogs by approved suppliers, hosting or renting of design and analysis software, approved product specifications, joint engineering reviews, transmission of engineering change orders, vaults for product life-cycle data, access to project management time lines, among others	Faster time to market because of concurrent design, testing, manufacturing planning, and tool design; greater flexibility in product and manufacturing planning, avoidance of costly redesign, savings in application software, savings in parts acquisitions, better project management

Source: IDC.

MARKET OPPORTUNITY

We believe the critical mass required for mass collaboration software adoption makes the near-term market opportunity limited.

Exhibit 92. Current and Near-Term (12 Months) Use of Instant Messaging by Company Size (% of Respondents)

Q. How likely is it that your organization will deploy instant messaging to your employees in the next 12 months?

	Small (1–99 Employees)	Medium-Sized (100–999 Employees)	Large (1,000+ Employees)
Likely	16.1	20.6	23.6
Somewhat unlikely	4.9	14.7	15.3
Not at all likely	33.3	23.5	20.8
Already deployed	3.7	2.9	6.9
Don't know	3.7	2	1.4

N = 255

Key Assumptions:

- IDC's experience has been that there is a gap of approximately 15% between respondents' stated intentions and actions with regard to technology purchases.
- Numbers may not add to 100% due to rounding.
- Large corporations are more likely to have distributed workforces, use ICE, and be early adopters of IM.

Messages in the Data:

- While corporate adoption is currently low, IT managers will seek to deploy IM applications or hosted services at a much faster rate in the next 12 months.
- As the technology becomes more pervasive and hosted deployments more tested and cost-effective, smaller corporations will roll out IM to their employees at a much faster rate.

Source: IDC, 2000. -

Other key barriers include IT security concerns, user familiarity, and the decision to purchase a brand-new software program for basic functionality that is already offered, though less elegantly and less completely, through Microsoft Outlook, Lotus Notes, and Web-based applications.

Exhibit 93. Market Opportunity for Decentralized Collaboration Software

		2001E	2002E	2003E	2004E	2005E
Employees	**Firms**					
1,000 - 4,999	15,867	47,601,000	47,601,000	47,601,000	47,601,000	47,601,000
5,000 - 9,999	2,033	10,165,000	10,165,000	10,165,000	10,165,000	10,165,000
10,000 - 49,999	1,693	16,930,000	16,930,000	16,930,000	16,930,000	16,930,000
50,000 - 99,999	198	19,800,000	19,800,000	19,800,000	19,800,000	19,800,000
>100,000	116	11,600,000	11,600,000	11,600,000	11,600,000	11,600,000
Total Heads		106,096,000	106,096,000	106,096,000	106,096,000	106,096,000
Penetration		10%	15%	20%	25%	30%
Potential Seats		10,609,600	15,914,400	21,219,200	26,524,000	31,828,800
$ per Seat		$50	$45	$41	$36	$33
Total Opportunity		**$530,480,000**	**$716,148,000**	**$859,377,600**	**$966,799,800**	**$1,044,143,784**

- We are targeting firms with more than 1,000 employees, as we believe those with fewer are much less likely to purchase collaboration software.
- We assume that the number of employees in each bracket is conservatively at the low end of the range.
- Based on early data, we believe early penetration rates are close to 10% of the total employee base at companies.
- We estimate a seat is $50, declining 10% per year.

Source: Dun & Bradstreet; Bear, Stearns & Co. Inc.

Impact on Other Industries

We believe decentralized collaboration software could have a profound impact on the way we communicate. Just as e-mail has increased productivity, decentralized collaboration could introduce significant benefits to users, especially in freeing data and empowering users to communicate in real-time.

We believe decentralized collaboration software could potentially displace current messaging systems. More likely, decentralized collaboration software will work in tandem with existing messaging systems.

More immediately, we think decentralized collaboration software will integrate into supply chains, ERP systems, and customer support front-ends. Since most decentralized collaboration software is grounded on XML, integration into applications and system software of every stripe is inevitable.

Public Companies That Take Advantage of Distributed Networks

Distribution, the first leg of decentralization, has been a powerful motif for more than a handful of public companies. As we have indicated, there are distinct advantages of distributing resources, application logic, and storage.

Below, we highlight several public companies who have taken advantage of distribution and decentralization in building their businesses. We do this in an effort to show the advantages and opportunities in moving toward distribution and decentralization and the foreseeable business models accruing from them.

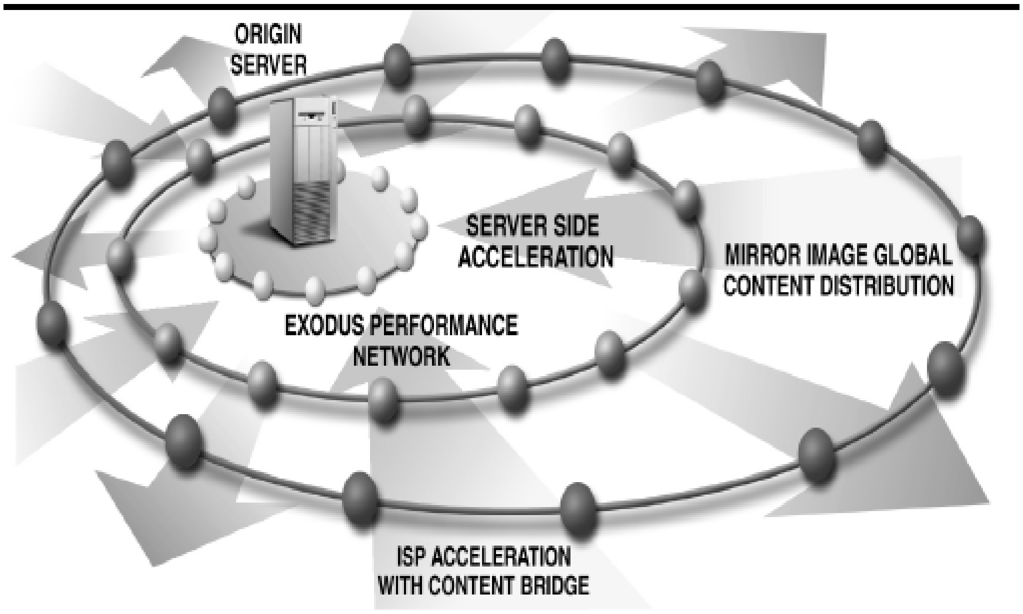
KEYNOTE

Keynote's measurement system relies on over 750 live agents distributed across the Internet. These agents collect real-time data from a growing base of over 50 metropolitan centers worldwide. Keynote has placed agents on all the major networks, which allows the company to show the performance that users experience on each of these backbone networks while connecting to any given Web site. This is different from server clusters, which are expensive and lack geographic perspective.

EXODUS

Exodus has datacenters geographically distributed across the globe not only to serve multiple geographies, but also to mirror content among datacenters. Exodus has created a network of concentric circles. Each layer interconnects with adjacent networks and creates a virtuous circle of distributed networks.

Exhibit 94. Exodus Network Architecture

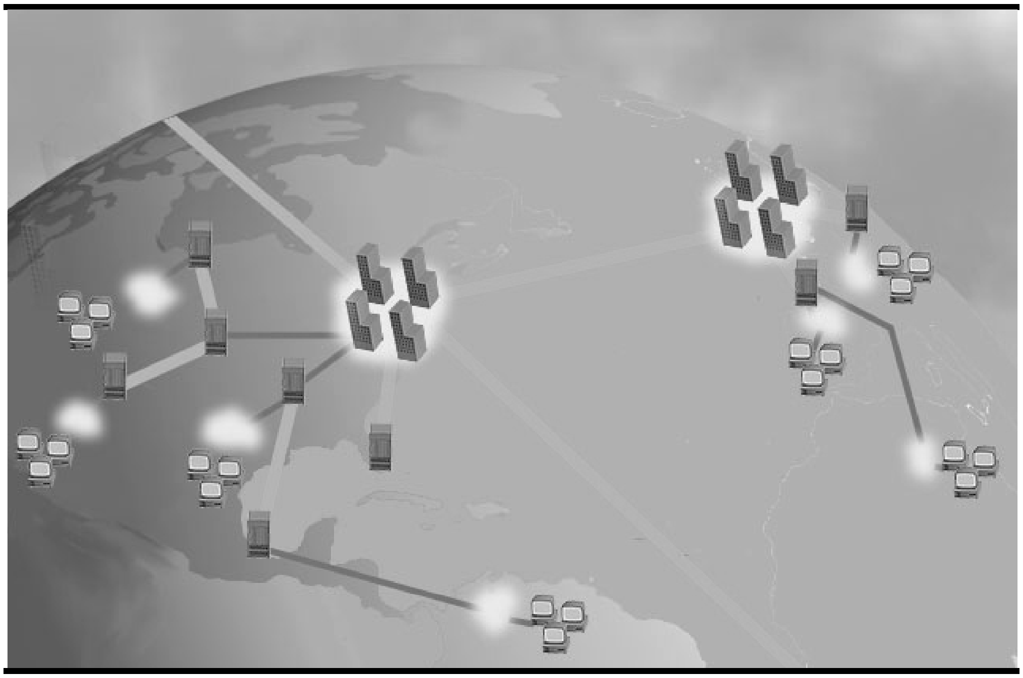


Source: Company data.

CDNS/SMNS

Digital Island and Akamai have distributed servers around the world. Distribution along the edge of the Internet provides advantages impossible in a centralized system.

Exhibit 95. Content Distribution Network

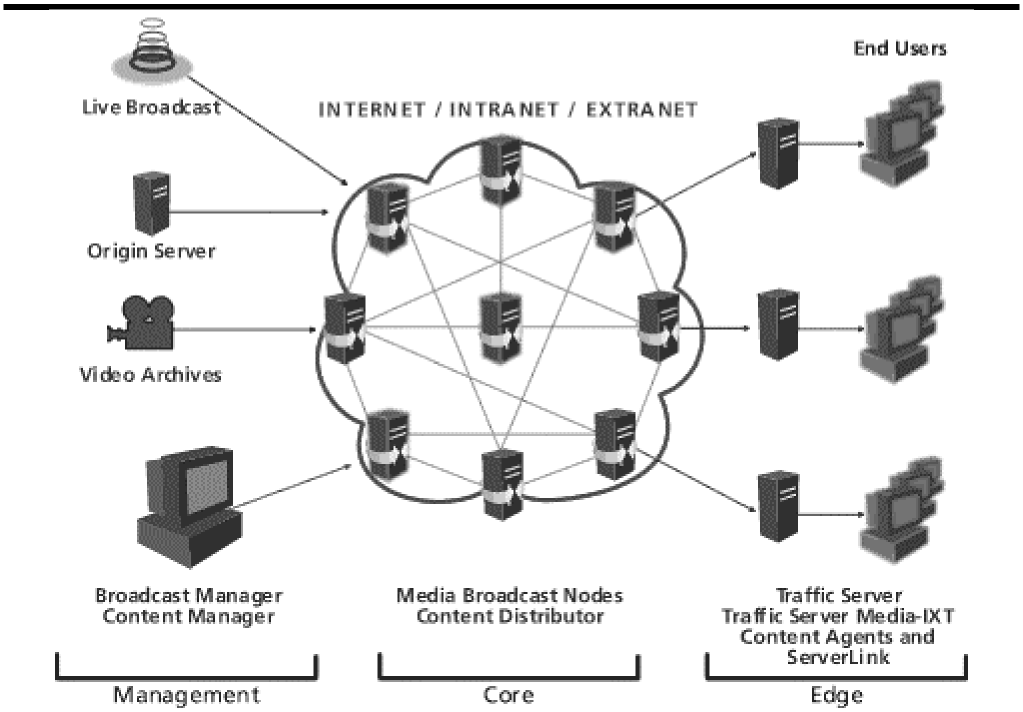


Source: Digital Island.

INKTOMI

Inktomi provides content delivery software that enables corporations and networks to distribute content.

Exhibit 96. Inktomi



Source: Company data.

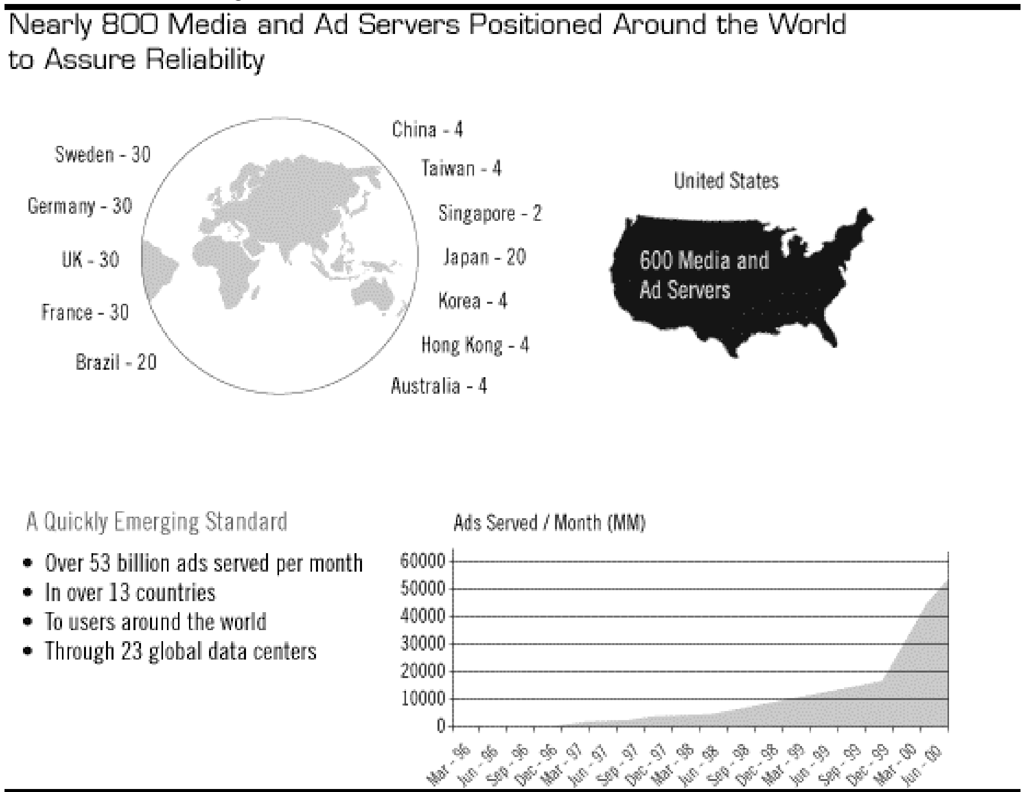
VERISIGN

As the manager of the registry for .com, .net, and .org, VeriSign is responsible for the largest directory on the Internet. The DNS is a distributed and hierarchic directory that is used by corporations and ISPs alike. VeriSign is at the top of this directory system.

DOUBLECLICK

DoubleClick was one of the earliest content distribution networks. Because DoubleClick was responsible for serving ads to globally distributed clients, the company had to build out its network of servers to serve ads. This architecture has enabled DoubleClick to scale with customer demand and to improve delivery times of ads.

Exhibit 97. Ad Serving Network

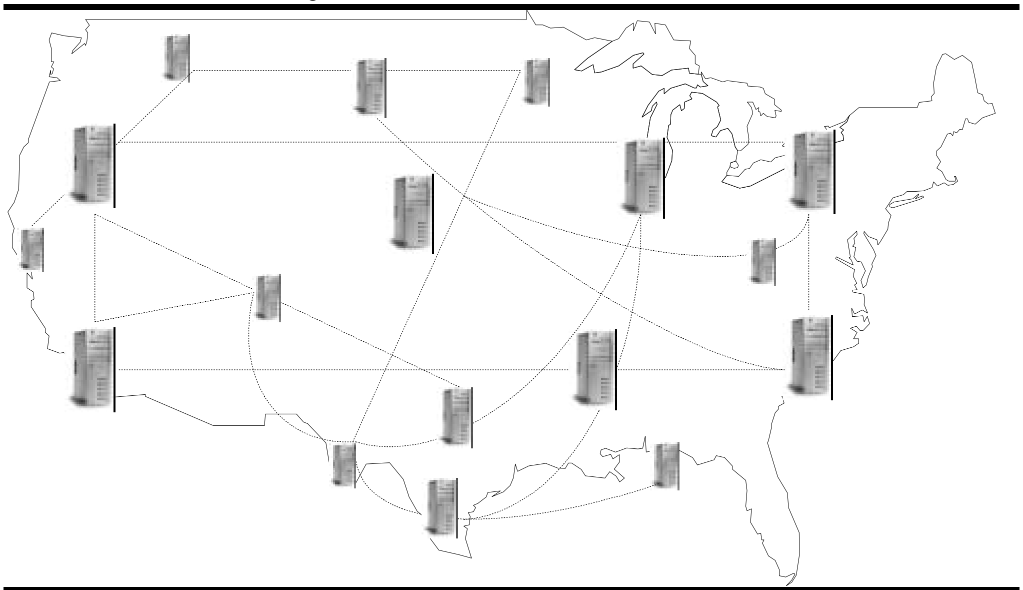


Source: Company data.

STORAGENETWORKS

StorageNetworks has deployed 51 S-POPs (Storage Points of Presence) — clustering around the world. We believe the distributed nature of these storage facilities will ultimately serve StorageNetworks well, as the virtualization of storage assets will enable optimization of resource utilization.

Exhibit 98. Distributed Storage Architecture



Source: Bear, Stearns & Co. Inc.

THE PLANS OF THE BIG THREE: INTEL, SUN, AND MICROSOFT

Having recognized the import of what is happening on the network and the Internet in particular, incumbents are quickly initiating projects and launching public awareness campaigns in order to jockey for position. It is our belief that incumbents in the computing world are realizing that the next wave of growth may have something to do with distribution in general and decentralization in particular. We believe three incumbents — Microsoft, Sun Microsystems, and Intel — are leading the efforts of Internet 3.0. Two of them have much to gain, and one has much to lose.

INTEL PEER-TO-PEER WORKING GROUP

The Peer-to-Peer Working Group is a consortium for advancement of best practices for peer-to-peer computing. Concerns about interoperability, security, performance, management, and privacy generally hamper the adoption and implementation of new technologies. Through the Peer-to-Peer Working Group, Intel is attempting to help set the standards for peer-to-peer computing applications and platforms.

Exhibit 99. Intel Peer-to-Peer Working Group Members

Alliance Consulting	Information Architects
Applied MetaComputing	Intel
BIAP Systems	J.D. Edwards
Bright Station PLC	NetMount
CenterSpan Communications	NextPage
Consilient	OpenCola
Endeavors Technology	OpenDesign
Engenia Software	Outhink
Entropia	Proksim Software
Fujitsu PC Corporation	Science Communications
Global Network Computers	Static Online
GridNode	Symbiant Group
Groove Networks	United Devices
Hewlett Packard	XDegrees

Source: Intel.

We find it particularly telling that Hewlett-Packard and Fujitsu PC are both on the list. We believe PC manufacturers, along with Intel, are approaching distributed processing vendors as possible assets in accelerating the adoption cycle of PCs. We also regard J.D. Edwards' presence as an interesting sidenote. J.D. Edwards provides ERP and SCM (supply-chain management) software. J.D. Edwards is interested in smoothing out the chains to increase transactions velocity and improve workflow across customers and suppliers. Decentralized collaboration software could be a major plus for supply-chain and demand-chain software and services vendors.

Exhibit 100. Intel Peer-to-Peer Working Group Committees

GROUP STEERING COMMITTEE	
Chairperson	
Brian Morrow	Endeavors Technology
INDUSTRY INFLUENCERS	
Andrew Chien	Entropia
Andrew Grimshaw	Applied MetaComputing
Jeffrey Kay	Engenia Software
Tim Mattson	Intel
TECHNICAL ARCHITECTURE COUNCIL	
Chairperson	
Bob Knighten	Intel
COUNSEL MEMBERS	
Greg Bolcer	Endeavors Technology
Steve Bush	OpenDesign
Andrew Grimshaw	Applied MetaComputing
Tom Ngo	NextPage
Damien Stolarz	Static.com
Jikku Venkat	United Devices

Source: Intel.

The Peer-to-Peer Trusted Library (PtPTL) allows software developers to add the element of trust to peer-to-peer applications. It provides support for digital certificates, peer authentication, secure storage, PKI, digital signatures, and symmetric key encryption.

An outgrowth of the success of the company's internal distributed processing effort (NetBatch), Intel is attempting to push the adoption of standards so that Internet 3.0, which is mostly a PC-based architecture, can become pervasive quickly and help accelerate the adoption of more powerful (read: newer) client and next-gen devices. This is in line with our thesis that the PC will become an even more central client device in Internet 3.0. Intel has partnered with Internet 3.0 companies, like UD, and has invested in more than a handful, including Uprizer, Groove, and Engenia.

Java, Jini, JXTA: Sun Microsystems

At a recent O'Reilly P2P conference, Sun Microsystems' chief architect, Bill Joy, unveiled Sun's foray into distributed peer-to-peer systems. Bill Joy has focused Project Juxtapose (JXTA) on the triplet — Searching, Sharing, and Storing information across the Internet. According to Joy, JXTA means "putting things next to each other, which is really what peer-to-peer is about."

The four main concepts of Project JXTA are:

- “pipe” from one peer to another;
- grouping;
- monitor and meter; and
- security.

When juxtaposing this project with Microsoft’s .NET, Joy indicated that Sun is not trying to do something infinitely complicated like .NET. This will be part of the Sun Open Network Environment (Sun ONE) platform at some point. Sun is looking for some simple, more elemental distributed computing technology. To advance its vision, Sun acquired InfraSearch in March 2001, a peer-to-peer search technology provider.

On April 26, Sun unveiled Project JXTA by introducing working documents and components. JXTA started as a project incubated at Sun under Bill Joy and Mike Clary to address peer-to-peer computing. Like Intel, the JXTA founders have as their vision the creation of an open set of standards that will allow interoperability among other peer-to-peer applications and tools.

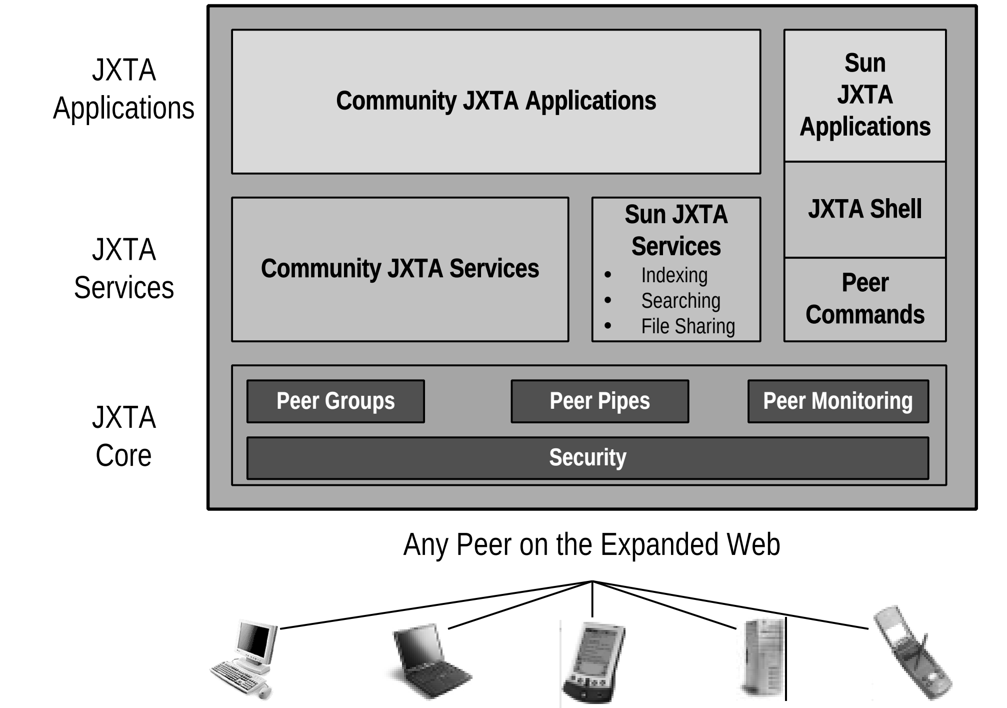
The principles underlying JXTA are:

- promote communication among applications;
- develop administrative commands for peers, peer groups, and groups of peers in the spirit of UNIX pipes and shells;
- keep the core small and elegant;
- support multiple platforms and languages, and micro devices to servers; and
- like Java, address security from the beginning.

The architecture of JXTA works like this: Users connect to the JXTA network by downloading a peer client. By logging onto the network, users connect to peers who have joined the JXTA network. In so doing, users can form groups, communicate, and collaborate.

JXTA is an open source project. JXTA’s tools enable developers to build applications for the JXTA platform that fit seamlessly. As developers build applications, the pool of resources available to the community grows.

Exhibit 101. JXTA Model



Source: Sun Microsystems.

The JXTA community of partners includes some of the more prominent companies in the class of Internet 3.0. In fact, many of them are included in the company profiles section at the end of our report.

Exhibit 102. JXTA Community

Source: Sun Microsystems.

One of our main themes has been that the PC will gather even more momentum in Internet 3.0. In a peer-to-peer environment, PCs will be the dominant gateways to the

Internet. With JXTA, Sun is trying to penetrate edge devices in general and the PC in particular. We believe Sun has recognized the increasing significance of the PC in the next generation of the Internet and wants to participate in its development. JXTA is an auspicious start, in our opinion.

MICROSOFT'S INTER.NET

If Microsoft could set the clock back ten years and shape the Internet and Web as it saw fit, what would it look like? We think Microsoft's blueprint would resemble .NET, the vision set out by the company last year. One of the pain points for Microsoft has been the company's inability to transfer its Windows dominance to the Web. If Microsoft had its way, it would have established Windows as the operating system for the Web.

Fundamentally, Microsoft .NET is an architecture for designing the Web that weaves all devices, data, and services into a fabric using standards-based programming interfaces and Microsoft software. By offering .NET to the Internet community, Microsoft has decided to make .NET the common platform for applications development and services on the Web. Much of .NET is predicated on the notion of interchangeable parts through XML and SOAP (see Appendix B) — interchangeable in large part because the parts are supplied by Microsoft.

Why has Microsoft waited until now to launch .NET? We believe a few things had to be in place before Microsoft could introduce .NET:

- **First, the Web.** In order to offer Web services, the Web had to reach critical mass, not just among consumers but businesses as well.

- **Content Is King.** The Web also had to grow its distributed database of content and applications. MSN.com and Passport are perfect examples of content and applications not available five years ago.

- **Learn, Trust, Rely.** Users had to grow to learn, trust, and rely on the Web. While this has certainly taken time, services like My Yahoo! have proven that users do trust and rely on the Web for essential applications like calendaring, e-mail, and financial services.

- **XML and SOAP.** A universal glue capable of connecting disparate objects, systems, and devices had to emerge. XML has only recently gained traction, and SOAP has only recently garnered recognition. The Microsoft .NET framework rests on a bed of XML and SOAP.

Products and Services

Microsoft's .NET product portfolio is vast and includes virtually every piece of a distributed Web services framework. We highlight the portfolio below.

Servers

- Exchange 2000 — messaging infrastructure

- SQL Server 2000 — database infrastructure

- BizTalk Server 2000 — document infrastructure
- Commerce Server 2000 — commerce infrastructure
- Internet Security and Acceleration Server 2000 — network infrastructure
- Application Center 2000 — application and datacenter infrastructure
- Host Integration Server 2000 — legacy integration infrastructure

Tools

- .NET Framework
- CLR — runtime program
- Libraries/APIs — messaging schemes
- Active Server Pages — Web application server
- Visual Studio.NET — front-end Web development tool
- C# — programming language of .NET

Not only has Microsoft extended its Web development framework with Visual Studio.NET, it has introduced an alternative to Java with C#.

Services

HailStorm represents Microsoft's first Web service within the .NET framework. HailStorm includes Microsoft Passport (authentication), Hotmail (Web-based messaging, including calendaring), and MSN.com (Web-based content), and will be a subscription-based service.

The initial set of HailStorm services includes:

- myAddress — electronic and geographic address for an identity
- myProfile — name, nickname, special dates, picture
- myContacts — electronic relationships/address book
- myLocation — electronic and geographical location and rendez-vous
- myNotifications — notification subscription, management, and routing
- myInbox — inbox items like e-mail and voice mail, including existing mail systems
- myCalendar — time and task management
- myDocuments — raw document storage
- myApplicationSettings — application settings
- myFavoriteWebSites — favorite URLs and other Web identifiers

- myWallet — receipts, payment instruments, coupons, and other transaction records
- myDevices — device settings, capabilities
- myServices — services provided for an identity
- myUsage — usage report for above services

While the laundry list of HailStorm services might resemble My Yahoo!, because HailStorm uses XML and SOAP, Microsoft is able to unify devices, data, and services into a seamless network. In addition, by leveraging its installed base of messaging products (Exchange, Outlook) and consumer software (Windows), Microsoft should be able to integrate Web- and non-Web-based applications and data fluidly.

Ultimately, we believe Microsoft has the vision of empowering individuals in essence to carry their identities (and all that composes an identity — contacts, messaging, calendaring, content, data) wherever they go. This has been one of the biggest problems with the Web to date — the inability of devices and data repositories to form a mosaic of a user's identity, where updating one facet of one's life requires the manual updating of all of one's properties. We believe .NET in general, and HailStorm in particular, is aimed at helping us aggregate our dispersed data stores into what Microsoft refers to as a constellation that includes our personal features and the services we want while integrating and synchronizing our devices. In essence, .NET helps us to formulate and organize our digital selves.

Our chief concern with this vision arises from the centralized storage of data in .NET, particularly HailStorm. Microsoft owns HailStorm; therefore, Microsoft could potentially control a significant portion of bit-based information about an individual. In fact, Microsoft could in effect own an individual's online identity. We are not bothered so much by this fact; rather, we are troubled more by the dominant presence of the browser in HailStorm.

As such, it is afflicted by all of the problems introduced by the browser that we have articulated, namely, the browser's inability to connect users point-to-point. Transferring files between two devices using browsers is often mediated and indirect. We believe a non-browser-based application (leveraging an instant messenger client, for example) could be more interesting.

Nevertheless, we believe weaving disparate data sets and applications with XML and SOAP enables devices finally to speak with other devices seamlessly. .NET is distributed (though not decentralized) because, in its case, the network is the entire Internet.

Tempering the Tendency to Overstate the Case

"There will be time, there will be time…"

– T. S. Eliot (The Love Song of J. Alfred Prufrock)

Internet 3.0 will happen. But, it will take time.

We must ask ourselves: Is Internet 3.0 going to change the way we interact with family and friends? Will we be more productive? Will we face fewer hassles? Does this change anything? What if this is like push technology, or B2C and advertising-based business models? Could companies, technologies, or industries born of the decentralization theme be relics or sorry remembrances of things past within five years?

We remind readers that decentralization is not a piece of hardware, an operating system, or a brand new communications network. Decentralization is a direction, and as such, its inertia, we believe, will drive the development of the Internet.

Some believe decentralized systems pose insurmountable problems for the enterprise and will therefore never be widely adopted. This was precisely the thinking when the PC was born and when the World Wide Web emerged. Adoption of decentralized systems will naturally take time, particularly within the enterprise. But it will happen, despite resistance, because decentralized systems and applications, as we have shown, offer too many advantages and efficiencies.

In the end, decentralization will prevail, in our view, because it satisfies the four criteria we established at the beginning of this report: utility, ease of use, cost, and necessity. Decentralization enables services that increase productivity; decentralized systems are becoming easier and easier to use; the cost advantages are striking; and decentralized systems can do things that centralized systems cannot.

WHAT TO WATCH FOR

Assuming distribution and decentralization are the driving variables in Internet 3.0, how do we identify where we are in the cycle? For signposts in Internet 3.0, look for the following:

Phase I: Embrace

- **Investments in the Class of Internet 3.0.** We believe the entrants to watch for (notwithstanding Microsoft, Sun Microsystems, and Intel) are EMC, PC OEMs (Dell and IBM in particular — Dell because of its dominant position and IBM because of its patents on distributed processing), enterprise software vendors (e.g., Oracle, Veritas, SAP, Interwoven), device manufacturers (e.g., Palm, Handspring, and RIMM), companies with complex database applications like eBay, Amazon, and Yahoo!, xSPs (e.g., Exodus, Digex, StorageNetworks, and Akamai), and even ISPs like AOL Time Warner and EarthLink.

- **More Partnership Announcements.** We believe Scale Eight's partnership with Akamai, United Devices' partnership with Exodus, and Groove's inclusion in Microsoft HailStorm are the tip of the iceberg.

- **Client Devices Continue to Become Thicker.** The laws of network dynamics indicate that the client devices will be thicker in the future because the cost of decentralizing (i.e., localizing) content, processing, and data on the client device will be cheaper than centralizing assets and resources. Client devices will be able to do more as a consequence.

- **Look for Messaging to Lead, Applications to Follow.** Keep an eye out for instant messaging to lead the adoption of Internet 3.0 within enterprises. In particular, look for Yahoo!, Microsoft and AOL Time Warner to lead the charge, positioning themselves with Jabber and other open source XML-based messaging platforms. Applications, like collaboration and value-added network services will follow the adoption of instant messaging platforms.

- **Emerging New Applications.** Instant messaging and file sharing are two recent applications that have gained instant popularity. We believe on-the-fly VPN creation among an individual's devices where all of an individual's devices (phone, PDA, PC, etc.) will be linked and synchronized securely, and controllable by each other, could help drive Internet 3.0.

- **Further Adoption of XML and the Spread of SOAP.** We believe distributed applications and platforms will have to incorporate XML and SOAP to future-proof themselves.

Phase II: Compete

- **Attack Large Markets First.** The Class of Internet 3.0 are expected to attack the largest markets first: large enterprises and service providers. The tell-tale sign that the competitive landscape is in full bloom is when one in the Class of Internet 3.0 wins a large contract at an incumbent's expense. Such an event, in each sector, will mark the beginning of Phase II.

- **Commodity Hardware, Specialized Software.** Storage vendors like EMC and Network Appliance layer on proprietary software onto commodity hardware (drives, memory). This trend will continue. Look for most of the innovation in Internet 3.0 to come from software. By offering software for commodity devices, many of the Internet 3.0 technology providers will enable enterprises to layer software on top of installed hardware (like PCs and servers) and save tremendous capital expenditures.

- **Expect Incumbents to Downplay Significance of Internet 3.0 Businesses.** Recognizing the potential impact of a new business, and unable to maneuver quickly enough and reposition business models, incumbents will begin to downplay the significance of Internet 3.0 businesses. This is precisely what happened with mainframe and minicomputer manufacturers when the PC emerged; this is what off-line retailers said when the Web was launched; this is what many carriers had been saying about Voice-over-IP; and this is what many have been saying about outsourced storage.

- **Handful of Companies Come Public.** On the heel of large design wins, a handful of Internet 3.0 members will begin to IPO. While still early, we believe the most likely candidates are Groove Networks, Scale Eight, and United Devices.

Phase III: Win

- **Incumbents Fall.** As EMC is to IBM, Dell is to Compaq, and Microsoft is to Novell, Internet 3.0 is to incumbents. In every cycle, incumbents fall to newcomers. Internet 3.0 will be no different. Some incumbents will lose because their models are incompatible with the distributed and decentralized direction of computing and networking. Others will fail because they failed to recognize the impact of distribution and decentralization on computing and networking.

- **Acquisition — Last Ditch Effort.** Unable to recover, and too late to partner, some incumbents will try to acquire competitors to regain edge.

- **New Markets Created.** Brand new markets will emerge, led by the Class of Internet 3.0 or incumbents who positioned themselves in Phase I.

BEAR STEARNS

Private Company Profiles

Funding and Deal Activity; Customer Momentum Already

To date, we believe over $500 million has been invested in brand new companies that are pursuing the development of distribution and decentralization technologies.

Exhibit 103. Funding for Distribution- and Decentralization-Focused Companies

Company	Amount		Date	Investors
Applied MetaComputing	$6 million	Series A	October 2000	Polaris Venture Partners
Centrata	$5 million	Series A	Summer 2000	Kleiner Perkins, Common Angels, dot EDU Ventures
Consilient	$2.8 million	Series A	November 2000	Oak Hill Venture Partners, The Sapling Foundation
DataSynapse	$5.25 million	Series B	September 2000	Rare Medium Group, Neocarta Ventures, The NYC Investment Fund, Silicon Alley Venture Partners, Wand Partners, Henry Kravis, Rick Stowe, David Rockefeller
Ejasent	$25.9 million	Series B	August 2000	Crescendo Ventures, Crystal Internet Ventures, Red Rock Ventures, Technology Crossover Ventures, Bill Joy, BV Jagadeesh, Prabakar Sunderrajan, David Banks
eMikolo	$4.5 million	Series A	October 2000	Israel Seed Partners
Engenia Software	$22.7 million	Series C	July 2000	Cooley Godward, Dominion Ventures, Intel Capital, SpaceVest, Thomson Corp., Topaz Investors, Winfield Capital Corp., St. Paul Venture Capital, Novak Biddle Venture Partners, Vanguard Atlantic, Aurora Funds
Entropia	$23 million	Series B	January 2001	Mission Ventures, RRE Ventures
ExactOne	$4.5 million	Series A	December 2000	JEGI Capital, SeaCap Ventures, Kaufman Family Partnership
Groove Networks	$41 million	Series B	October 2000	Accel Partners, Intel Capital
iKimbo	$6.375 million	Series B	May 2000	Cross Atlantic Capital Partners, PTEK Ventures, Draper Atlantic, Steve Walker & Associates
Infrasearch	$5 million	Seed	November 2000	Marc Andreesen, Angel Investors
Kalepa Networks	$1 million	Seed	June 2000	Angel Investors, Jump Investors, Wilson Sonsini Ventures, Staenberg Private Capital, FirstCorp Leasing
Napster	$15 million	Series C	May 2000	Hummer Winblad Venture Partners, Angel Investors
NextPage	$20 million	Series B	January 2000	Oak Investment Partners, epartners, Dominion Ventures, Amp Capital Partners
Oculus Technologies	Unknown		Unknown	Ford Motor
OpenCola	$13 million	Series B	January 2000	Battery Ventures, Mosaic Venture Partners, Torstar Corporation
OpenDesign	$7 million	Seed	Unknown	Nathan Myhrvold
Parabon Computation	$4.9 million	Seed	Unknown	undisclosed
Proksim Software	$3.9 million	Series A	Unknown	Société Innovatech du Grand Montréal, T2C2 Capital
QUIQ	$15 million	Series B	January 2001	InterWest Partners, BancBoston Ventures, Altos Ventures, Discovery Ventures
Roku Technologies	$6.1 million	Series B	March 2000	PSINet Ventures, Draper Atlantic, Nextel Communications
Scale Eight	$26.5 million	Series B	July 2000	CenterPoint Ventures, Crown Advisors, InterWest Partners, Oak Investment Partners
Static Online	$5 million	Series A	December 2000	Zone Ventures, Tim Draper
United Devices	$13 million	Series A	August 2000	SOFTBANK Capital, Oak Investment Partners
Uprizer	$4 million	Series A	April 2001	Intel, Kline Hawkes, Shugart Venture Fund
WorldStreet	$30 million	Series C	July 2000	Reuters, ING Barings, UBS Warburg, Advanced Technology Ventures IV, North Bridge Venture Partners II, Gilde IT Fund BV, Cambridge Technology Capital Fund, J.P. Morgan Capital, Comdisco Ventures, ABS Ventures, MF Private Capital
XDegrees	$8 million	Series A	November 2000	Redpoint Ventures, Cambrian Ventures
Zambeel	Unknown		2000	Kleiner Perkins, NEA
Zodiac Networks	Unknown		2001	Kleiner Perkins, Benchmark

Source: Bear, Stearns & Co. Inc.

Some of these companies sport very well known investors. Sun Microsystems' Bill Joy and Exodus co-founder BV Jagadeesh has invested in Ejasent. Intel Capital has invested in a handful of private companies. Vinod Khosla of Kleiner Perkins has led investments in Centrata and Zambeel. Nathan Myhrvold has helped found OpenDesign. Ford Motor has invested in Oculus Technologies.

Exhibit 104. M&A Transactions

Company	Date	Type	Amount	Acquirer
Gigabeat	April 2001	Acquisition	NA	Napster
Infrasearch (1)	March 2001	Acquisition	N/A	Sun Microsystems
Scour	December 2000	Acquisition	$9 million	Centerspan Communications
Napster	October 2000	Minority Stake	Up to $50 million loan	Bertelsmann Music Group

(1) aka GoneSilent.

Source: Bear, Stearns & Co. Inc.

Sun acquired Infrasearch to grow its JXTA vision. We believe Intel has been investing heavily in distributed processing and collaboration companies to push PC and faster processor adoption. Microsoft has yet to invest in any company, but they have highlighted Groove in HailStorm and WorldStreet as a poster child of the company's .NET strategy.

We also find it interesting that many of these private companies have already signed customers at public companies' expense.

Scale Eight, a provider of distributed multimedia storage services, has signed Akamai as a customer. Scale Eight has also signed MTVi Group's Web properties (MTV.com, VH1.com, Country.com, and Sonicnet).

NextPage, a distributed collaborative software provider, has signed ABN AMRO, Deloitte & Touche, and ABC Television Network as customers.

ExactOne, a distributed database software provider, has signed PartMiner, Gomez.com, and Guru.com.

Oculus Technologies, a collaboration software provider, has signed Ford Motor Company as a customer.

QUIQ, a collaboration software provider, counts National Instruments and Network Appliance among its customers.

United Devices, a distributed processing concern, has signed a deal with Exodus.

Applied MetaComputing, Inc.

Address

1 Memorial Drive
Cambridge, MA 02142
(617) 374-2500
www.appliedmeta.com

Team

Management

Dr. Andrew Grimshaw	President
David Kopans	CFO
	VP of Business Development

Board of Directors

TBA

Investors

Polaris Venture Partners

$6 million	Series A

Customers

Boeing, Stanford, U.S. Naval Research Laboratory, NASA's Ames Research Center, Lawrence Livermore National Laboratory

Competitors

Parabon
XDegrees

Stats/Misc

25 Employees

Description

Applied Meta is a platform provider for distributed applications.

Authors of Legion, the distributed computing platform, Applied Meta is the brainchild of Andrew Grimshaw and the result of six years of research and development work conducted in the computer science department at the University of Virginia. Originally designed and built for the Department of Defense (DoD) and the Department of Energy (DOE), Legion was conceived as a product that would enable the high-speed LAN and WAN networks of the future to achieve new levels of collaboration and computing performance. Legion is in use at research facilities around the globe, from Stanford University, to the Naval Research Laboratory, to NASA's Ames Research Center.

Legion was first developed at the University of Virginia with NSF, DOE, DOD, and NASA support and in collaboration with NPACI and NAVO MSRC.

Applied Meta has operated a network of more than 4,000 CPUs for the National Science Foundation.

Products and Services

In the coming weeks, Applied Meta will be unveiling its product and service offerings.

Centrata, Inc.

Description

While Centrata remains in stealth, we believe the combination of MIT and Akamai pedigree, Vinod Khosla, and savvy advisors make this company one to keep on the radar screen.

We believe much of the work that Shishir Mehrotra and his team at MIT centered on randomization algorithms and the utilization and management of distributed resources.

Address

130 Shoreline Drive, First Floor
Redwood Shores, CA 94065
info@centrata.com
www.centrata.com

Team

Management

Deepak Bhagat	CEO, Board Member
Vinod Khosla	Chairman of the Board General Partner, Kleiner Perkins

Board of Advisors

Janpieter Scheerder	EVP, Sun
Bill Campbell	Former CEO of Intuit
Rob Bowman	EVP of RealNames
Mike Cassidy	Co-Founder of Direct Hit
Kevin Kinsella	Avalon Ventures
Rajeev Motwani	Prof. of Computer Science, Stanford
Dr. Nancy Lynch	Prof of Electrical Engineering & Computer Science, Head of Theory of Distributed Systems research at MIT

Investors

Kleiner, Perkins, Caufield & Byers
Common Angels
Dot EDU Ventures
Invesco
Presidio Venture Partners

$5 million — Series A

Status

Stealth

Consilient, Inc.

Address
1815 Fourth Street, Suite B
Berkeley, CA 94710
(510) 981-9200
www.consilient.com

Team

Management

Ray Conley — CEO
Erik Freed — President & CTO
Brian Falkenhainer — VP Product Development
Joe Campbell — VP of Engineering
Bob Chamberlain — VP Worldwide Sales & Business Development
Greg Groom — SVP of Strategic Alliances
Sandra Smith — VP Strategy & Marketing

Board of Directors

Jonathan Hare — Chairman
Michael Spence — Vice Chairman, Partner, Oak Hill Venture Partners
Ray Conley — Partner, Oak Hill Venture Partners
Erik Freed — President & CTO
Susan Dawson — Director, The Sapling Foundation

Investors

Oak Hill Venture Partners
The Sapling Foundation
$2.8 million — Series A

Customers

BP Amoco

News

SAIC Signs Strategic Alliance Agreement With Consilient — 3/22/01
BP Amoco Signs Worldwide Agreement With Consilient — 3/20/01

Competitors

iKimbo, Groove Networks, among others

Description

Consilient is an enterprise software company aimed at creating, distributing, and optimizing business processes. The company is focused on solving the following:

- Support the diverse nature of real-world business processes: unpredictable combinations of manual and automated tasks that are distributed across multiple people, infrastructures, and organizations.
- Provide task-focused and information-rich content that individuals need to make informed and effective business decisions.
- Deliver a way to empirically track and record business processes, so that companies can adapt and optimize process flow.
- Extend, enhance, and scale existing IT investments.

The Consilient solution seamlessly addresses each of these challenges with a unique mobile process container and peer-to-peer infrastructure that delivers unprecedented flexibility, scalability, and control over dynamic business processes.

Products and Services

The Consilient Sitelet Technology Platform is a framework for the rapid development of process personalization solutions for business enterprises.

The Sitelet Technology Platform supports the rapid creation, distribution, and execution of portable, interactive process containers called Sitelets. These mobile, self-contained intelligent agents have the ability to dynamically aggregate and organize process content; transport the content between people and systems; and support the interactive discovery, evolution and execution of business processes.

Sitelets work with your existing technology and applications to manage the flow of business processes across the Web, e-mail, or any other network. Sitelets automatically track the process flow as it moves across your organization and among your external partners, allowing you to view the current status of any process and to quickly identify and correct bottlenecks. As a result, companies can distribute process execution without giving up control or losing sight of their broader strategic goals.

DataSynapse, Inc.

Address

408 8th Avenue, Penthouse Suite A
New York, NY 10001
(212) 842-8842
www.datasynapse.com

Team

Management

Peter Y. Lee	Co-Founder & CEO
Jamie Bernardin	Co-Founder & CTO

Board of Directors

Derrick Chen	General Partner, Rare Medium Group
Peg Murphy	Director, NeoCarta Ventures

Board of Advisors

Michael Lobdell	MD, JP Morgan Chase
Steve Brotman	Managing Director, Silicon Alley Venture Partners
Tom Conway	Associate, Wand Partners
William Cruger	MD, JP Morgan Chase
Bernard Goldstein	Director, Broadview International LLC
Richard H. Stowe	Founding Partner, Welsh, Carson, Anderson & Stowe
Volney (Terry) Taylor	Retired Chairman & CEO, The Dun & Bradstreet Corp.

Investors

Rare Medium Group, Rare Ventures
Neocarta Ventures
The NYC Investment Fund
Silicon Alley Venture Partners
Wand Partners
Henry Kravis, Rick Stowe, David Rockefeller

$5.25 million — Series A

Customers

First Union

Competitors

Entropia
United Devices
Parabon

Description

DataSynapse is a provider of business solutions that improve applications turnaround performance in the financial services, energy, and telecom sectors. The company delivers a peer-to-peer distributed computing platform that leverages clients' existing IT infrastructure. DataSynapse's WebProc platform can be integrated with legacy applications in one to two weeks to harness idle, underutilized, and/or dedicated network resources on a client's intranet.

The solution addresses time-sensitive, naturally parallel applications including the vast majority of decision support, mid-office, and product development applications involving Monte Carlo, Value@Risk, stress testing, and other simulations.

Products and Services

DataSynapse's solution enables firms to:

- reduce the turnaround time of compute-intensive processes;
- run more iterations of analytics; and
- leverage existing investment in it infrastructure to achieve cost savings and seize time-to-market opportunities.

Position

There are several key trends driving client attraction to the next generation of distributed computing:

- implementation of Straight-Through Processing (STP), accelerating trade settlement time for all global market participants from T+3 (three days after trade) to T+1;
- strategic objective to move toward intra-day, real-time portfolio pricing and risk management, instead of end-of-day cycles; and
- the huge leap in switched network bandwidth allowing applications to be easily distributed to remote resources anywhere on a client's intranet.

DataSynapse provides a turnkey solution to powering these business-critical applications using the underutilized resources already in place.

Ejasent, Inc.

Address

2490 E. Charleston Road
Mountain View, CA 94043
(650) 230-6300
www.ejasent.com

Team

Management

David Banks	CEO
Rajeev Bharadhwaj	Founder & CTO
Burton Hipp	VP, R&D & Founder
Sandy Abbott	CFO
Ron Anderson	VP, Network Services
Susan Depositar	VP, Human Resources
Robert Freeman	SVP, Marketing
Michael Grant	VP Corporate Development
William Lattin	Chief Security Officer
Bill Ogden	Senior Director, Sales
Robert Rodriguez	VP, Engineering
Rob Rosen	Director Professional Services

Investors

Crescendo Ventures
Crystal Internet Ventures
Red Rock Ventures
Technology Crossover Ventures

Bill Joy	Sun Microsystems
BV Jagadeesh	Exodus
Prabakar Sunderrajan	Exodus
Peter Sevcik	Netforecast
John Katsaros	Jupiter Communications
David Banks	Former CEO, Versant

$26 million	Series B

Description

Ejasent owns and operates a global, distributed, interactive services network that enables eBusinesses to allocate computing resources in real time to meet the needs of variable, unpredictable and opportunistic Internet traffic. Ejasent enables enterprises to exercise granular control over the provisioning, deployment, and scaling of Web infrastructure based on price/user response time goals set by the enterprise. By deploying Ejasent, an enterprise has the ability to transparently and virtually extend its existing central site capabilities without costly, complex, and slow site expansion.

Products and Services

Ejasent UpScale, Ejasent's first interactive service, supplies eBusinesses with on-demand Web application processing capacity instantly and transparently. With UpScale, Ejasent has created a dynamic platform that offers consistent and predictable Web site performance, dramatically improved Web site economics, and infinitely elastic capacity on demand.

Ejasent's Application Processing Network (APN) responds in real-time to rapid increases in an eBusiness's customer load. As the traffic load on a Web site increases and user response times start to degrade, the site can add additional application processing capacity transparently and instantly (within three seconds). This enables the eBusiness to adaptively increase or decrease the processing power necessary to assure site responsiveness regardless of the traffic load or the geographic location of the Web site. The result is that large sites are able to maintain a consistent level of end user quality of experience at a global level.

Ejasent's patented Instant Application Switching technology enables Web site managers to take a "snap shot" of their Web applications and host these on Ejasent servers placed around the Internet. These AppShots are identical instances of the applications running on the central site. When loads on the central site exceed pre-defined thresholds, Appshots are "activated" in physical locations closest to the points of consumption. This "activation" process takes approximately one second and is executed automatically. The Appshots maintain data synchronization with the central site over a secure, authenticated, and encrypted connection.

As traffic loads recede, AppShots are progressively "retired." Applications in the Application Processing Network are not tied to any particular server. Instead they are scheduled "on-demand" and run on any of the thousands of available processors thus creating, in effect, a "virtual single server."

eMikolo Networks, Inc.

Address

300 East 42nd Street, Second Floor
New York, NY 10017
(646) 495-2162
www.emikolo.com

Team

Management

David Butler	CEO
Moshe Raines	President
Ittai Golde	CTO
Oded Regev	Chief Scientist
David Yerushalmi	Director Business Development
Dani Blendis	VP Product Development

Board of Directors

Hanan Gilutz	Founding Investor
Michael Eisenberg	Israel Seed Partners

Investors

Israel Seed Partners

$4.5 million	Series A

Competitors

Inktomi, CacheFlow, Network Appliance, Cisco

Description

eMikolo Networks is developing next-generation Internet infrastructure software built upon its proprietary collaborative technology platform. eMikolo Networks' collaborative technology platform is applicable to a wide array of infrastructure-related areas, including content delivery and Internet storage.

Products and Services

eMikolo has made Demand Driven Access (DDA) technology available to the public.

Demand Driven Access is a stand-alone appliance that plugs into any IP-based network. This plug and play appliance creates an infinitely scalable network by automatically distributing a "thin agent" to all users. The DDA then creates a Common Object List (COL) that contains the largest, most frequently requested Web objects to be delivered over the network. This configurable list, which has a default of 10,000 objects, ensures that all web based content (static, dynamic, and/or streaming) is delivered dynamically to provide highly available services.

The COL serves as the data source for eMikolo's Edge Detection, a patent-pending technology that determines the best route for content to travel at any given moment. Edge Detection fluidly maps current Internet users, and creates transient content caches at the end user level. This method enables fast distribution of updated, local content closer to the end user.

Engenia Software, Inc.

Description

EngeniaUnity is a collaboration platform for instant, dynamic, event-driven collaboration among employees, partners, suppliers, and customers.

Products and Services

Engenia's collaborative services platform includes an applications server, tools, and interfaces with which to build XRM applications. Engenia defines XRM as managing the extended relationships companies have all along their value chain.

Position

Engenia has alliances with GE Global eXchange Services, BEA Systems, and Intel.

Address

1800 Alexander Bell Drive, Suite 100
Reston, VA 20191
(703) 234-1400
www.engenia.com

Team

Management

Jeff Crigler	Co-Founder & CEO
Jeff Kay	CTO
Linda Ciabatoni	President & COO
Joseph Herman	VP, Product Dev..
Ben Steinberg	VP, Professional Services
Becky Beeler	VP, Finance & Operations
Stephen Ford	VP, Sales

Board of Directors

Rick Boswell	General Partner, St. Paul Venture Capital
Jeff Crigler	CEO
Jeff Kay	CTO
Lee Keet	President, Vanguard Atlantic
Roger Novak	Founding Partner, Novak, Biddle
Roger Widing	MD, Spacevest
Randy Werner	General Partner, Dominion Ventures

Investors

Aurora Funds, Inc.
Dominion Ventures
Novak Biddle Venture Partners
Thomson Corporation
Topaz Investors, LLC
Cooley Godward, LLP
Intel Capital
SpaceVest
St Paul Venture Capital
Winfield Capital

$22.7 million	Series C
$30 million to date	

Customers

Ten customers (PetroVantage, Thomson's, Coherence Ventures, among others)

Competitors

MatrixOne, Ariba/Agile,
I2 Technology/Trade Matrix

Entropia, Inc.

Address

4350 Executive Drive, Suite 200
San Diego, CA 92121
(858) 623-2840
www.entropia.com

Team

Management

Robert North	CEO
Dr. Andrew Chien	CTO
Michael Shneyderman	CFO
Scott Kurowski	VP of Bus. Development
Neville Billimoria	VP of Marcom
Dr. Martin Stuart	VP of Life Sciences

Investors

Mission Ventures
RRE Ventures

$23 million	Series B
$30 million to date	

Customers

Envive, SolidSpeed

Competitors

Popular Power
DataSynapse
Porivo
Parabon

Description

Entropia has become famous within the DP community for its sponsorship of The Great Internet Mersenne Prime Search (GIMPS), a project that searches for a particular type of prime numbers.

Entropia has won the backing of several leading venture capital funds. Entropia is the first DP company to close a Series B round of financing, raising $23 million in its second round, and $30 million to date.

Products and Services

- **Entropia 2000 Internet Computing Services.** Entropia offers DP services on an outsourced basis with its 2000 Internet Computing Services.
- **Entropia Application Porting & Integration.** Entropia offers implementation and deployment services to port applications across its grid.
- **Entropia 2000 Enterprise Server.** Entropia offers Entropia 2000 Enterprise Server, a distributed computing system for in-house DP implementations.

EverNet Systems, Inc.

Address

225 Velarde Street
Mountain View, CA 94041
(650) 386-5545
www.evernet.com

Investors

Undisclosed	Seed

Competitors

OpenCola, Static Online, Allibra

Description

EverNet Systems, Inc. is a next-generation content delivery services provider that has developed a software technology for moving large (>1MB) and popular files on the Internet. Two and one-half years in the making and built from the ground up, EverNet's technology transforms Internet content delivery. Based on "distributed download" or multi-peers to peer ("PXP") architectural concepts, EverNet flips the Internet load model — the more in demand a file is, the better the service.

Products and Services

EverNet's proprietary Progressive Active Geometric Mirroring ("PAGM") technology is a centrally managed, three-tier, distributed download, PXP network, consisting of Master Servers, Distribution Servers, and Client software that resides on users' desktops. First a file is "PAGMized" — i.e., scanned for viruses, compressed, encrypted, and a digitally signed "header" is added. It is then accessed from a simple link on a content provider's site and delivered via a network of clients acting as mirroring sites for the requesting client, with the Master Server and Distribution Server authorizing and supervising secure network transactions. The file is broken into small, easily transferable "chunks" and the requesting client intelligently manages the process of receiving and assembling these from the cache or hard drive of multiple provisioning clients, into a complete and error-free download.

ExactOne, Inc.

Address

4465 Brookfield Corporate Drive
Suite 100
Chantilly, VA 20151
(703) 378-0808
www.exactone.com

Team

Management

Bernard Ferre	Chairman & CEO
Raymond Capece	President & COO
Tim Smith	VP Software Development

Board of

Kent Hawryluk	JEGI
Eric Cohen	SeaCap
Joe Heller	Next Level Venture
Richard Houston	
Raymond Capece	President & COO
Bernard Ferret	Chairman & CEO

Investors

JEGI Capital
SeaCap Ventures
Kaufman Family Partnership

$4.5 million Series A

Customers

PartMiner, Gomez.com, Guru.com, 86.com, Spree, University of Tennessee, among others.

Competitors

Cohera
Vignette
AltaVista
IBM Websphere

Description

ExactOne has developed data integration technology to facilitate real-time access and enablement of distributed data, and offers the technology as a hosted service. ExactOne customers have used the technology to deploy real-time comparison sourcing/shopping functionality. The technology enables buyers to simultaneously find and compare products from many suppliers in real-time. For example, PartMiner, has a site called FreeTradeZone in which buyers of electronic components can simultaneously find and compare products from 50 vendors across dynamic attributes such as price, inventory, and availability. Each vendor's electronic product catalog or database is integrated into ExactOne's distributed infrastructure using proprietary software.

To achieve "infinite" scalability, ExactOne has designed a distributed architecture that addresses two distinct problems: supporting an infinite number of queries and supporting an infinite number of target data sources. The resulting distributed architecture includes two major sub-systems: the front-end and the back-end. The front-end server includes the query parameters, the result attributes, and the target databases. The result attributes include part number, manufacturer, availability, quantity, price, and supplier. The target databases are electronic catalogs from 50 suppliers. The back-end servers include ExactOne's proprietary query engine, parsing engine, and post-filtering module. The front-end and back-end are independent software components that share the query workload on a many-to-many basis, without the need for centralized processing. ExactOne's software technology is entirely Java-based and uses accepted industry standards (XML, standard APIs, HTTP, etc.) to deliver maximum performance.

Products and Services

ExactOne's Dynamic Data Access services offers Net markets and PTNs means to deploy virtual and remote catalogs. The company's technology enables buyers to simultaneously find and compare products from many disparate supplier e-catalogs or databases in real-time.

Position

ExactOne has positioned its service as a hosted, cost-effective application. Customers leverage ExactOne's infrastructure which helps them cut the costs of acquiring bandwidth, accessing supplier data, maintaining supplier data up to date, and reporting usage.

Groove Networks, Inc.

Description

Groove is a platform for the development of peer-to-peer collaboration applications. As such, Groove offers a basic toolset that includes file sharing, instant messaging, calendaring, and co-browsing, among other functions. As a platform, Groove is a base on which developers can build applications.

The fundamental idea of Groove is the Shared Space (SS) — a space that collaborators share. It is private and viewable only to those who have agreed to participate in the SS. Participants in an SS can collaborate via a number of predefined tools and applications, and via an unlimited set of applications that developers can create.

Groove is almost always a peer-to-peer application. That is, messages and files are shared directly, unless users are temporarily off-line, files being transferred are too large, users are behind firewalls and require intermediate storage stops, or the number of recipients of a message or file is large. In these cases, Groove utilizes a relay hub to coordinate data flow among devices.

Groove can also be used to create virtual private networks (VPNs) of peers. This is possible with Groove because the identities of users on Groove are authenticated by digital signatures (encrypted keys that detail the identity of users on the network) that reside on the client device and on a central registry server. Each user in a Groove Shared Space connects and is permitted to collaborate with others within the Shared Space by exchanging digital signatures. It is possible with such a strongly encrypted peer system to create an extremely elegant and easy-to-use VPN across the public Internet. With Groove, PKI (Public Key Infrastructure) is obviated by the presence of exchangeable identity markers (digital signatures) on the Groove client.

Address

100 Cummings Center, Suite 535Q
Beverly, MA 01915
(978) 720-2000
www.groovenetworks.com

Team

Management

Ray Ozzie — Founder, Chairman & CEO
Michael Matthews — EVP of Sales & Marketing
Brian Halligan — VP of Sales
David Scult — VP, Business Solutions
Jack Martin — CFO
Ken Moore — Co-Founder, VP of Development, Infrastructure & Online Services
Jack Ozzie — Co-Founder, VP of Development, Platform & Developer Services
Eric Patey — Co-Founder, VP of Development, Tools & Application Components

Board of Directors

Jim Breyer — Managing Partner, Accel Partners
Mitch Kapor — Partner, Accel Partners

Investors

Accel Partners
Intel Capital

$41 million — Series C
$60 million to date

Competitors

Consilient, Engenia, iKimbo, QUIQ

Stats/Misc

200-plus developers

Ikimbo, Inc.

Address

500-A Huntmar Park Drive
Herndon, VA 20170-5100
(877) 945-4626
www.ikimbo.com

Team

Management

Jamey Harvey	EVP & Chief Product Officer
Edwin Miller	President & CEO
Eric Wimer	EVP & COO
Robert Caudill	CTO
Craig McIntire	VP of Engineering

Board of Advisors

Don Telage	Director Network Solutions
Seymour Miles	
Eric Pulier	Chairman and Co-Founder of U.S Interactive
Mark Warner	Managing Director of Columbia Capital

Investors

Cross Atlantic Capital Partners
PTEK Ventures
Draper Atlantic
Steve Walker & Associates

$6.375 million Series B
$7.6 million to date

Customers

PricewaterhouseCoopers

Competitors

Groove Networks, Engenia, QUIQ

Description

Ikimbo is a collaboration applications provider.

Products and Services

- **Omniprise.** Omniprise is Ikimbo's collaborative applications platform. One of the interesting things about Omniprise is its ability to incorporate wireless devices into the fabric of the collaboration network.

Position

By signing PricewaterhouseCoopers as a client, Ikimbo has positioned itself ahead of most collaboration software providers.

NextPage, Inc.

Address

3125 W. Executive Park Way
Lehi, Utah 84043
(801) 768-7500
www.nextpage.com

Team

Management

Brad Pelo	Founder & Chairman
Bill Wesemann	CEO
Nathan Birchall	CFO
Darren Lee	VP of Strategy & Marketing
Bruce Law	VP of Corporate Marketing
John Bara	VP of Business Development
Henry Heilesen	EVP
Tom Ngo	CTO
Patrick McGowan	VP of Engineering
Justin Anderson	VP of Worldwide Sales
Chris Worsley	VP of Services

Board of Directors/Advisors

Brad Pelo	Chairman
Anne Lamont	Oak Investment Partners
John Stevens	Amp Capital Partners
Doug Allred	VP at Cisco
Alan Ashton	Founder of WordPerfect

Investors

Oak Investment Partners
epartners
Dominion Ventures
Amp Capital Partners

$20 million — Series B

Customers

ABC Television Network
ABN AMRO
Deloitte & Touche UK
The Institute of Taxation in Ireland
Thomson Learning
West Group
(150 in total)

Stats/Misc

200 employees
NextPage was founded in July 1999. The company has acquired a professional services firm and another P2P company, netLens.

Description

NextPage launched the NXT 3 e-Content Platform — the first application to enable Peer-to-Peer Content Networking for businesses. The software connects servers together in a network of content, as if that content were in one system or location. For example, a user in Chicago from a multinational corporation can simultaneously access information located in Chicago, London, New York, and other offices as if that information was sitting on his/her desktop. NextPage technology eliminates the need to centralize resources. Content authors store the information on the servers at their location, and the authors can maintain and update the information. No content is centralized, moved, or replicated. The IT department does not have to manage the content, just the application. Companies can connect other external resources to their Content Network. A company can link partners, suppliers, and customers into the network of connected servers. Therefore, users can obtain information inside and outside the company as if it were on their desktops. Also, NextPage technology uses server-based P2P to capitalize on the benefits of distributed technology while ensuring security of business-critical information of enterprises. Companies are using this technology to increase efficiency, to leverage the value of their information assets, and to improve revenues.

Products and Services

- **NXT 3 e-Content Platform,** enabling Peer-to-Peer Content Networking.

NextPage competes against peer-to-peer products that enable information retrieval or file sharing. NextPage technology has some elements of search, portals, and content management, and in some instances may be competitive to those solutions. However, typically these technologies are complementary to NextPage technology, and, in fact, NextPage has formed partnerships with portal and content management companies.

Oculus Technologies Corp.

Address

103 Broad Street, 5th Floor
Boston, MA 02110
(617) 426-4277
www.oculustech.com

Team

Management

Christopher Williams	President & CEO
Matthew Tullis	CFO
Matthew Wall	CTO
Benjamin Linder	VP of Product Development
Robin Waldman	VP of Marketing
Ronald Czik	VP of Engineering
Shaun Meredith	VP of Product Deployment
Marianne Wisheart	Director of Human Resources

Investors

Ford Motor Company

Customers

Ford Motor Company

News

Signs Ford Motor as customer	3/27/01

Description

Oculus Technologies provides P2P collaboration software. Oculus CO provides the "glue" to tie disparate applications, platforms, and organizations together. And CO provides the "grease" to smooth the product development process, speeding up information flow, enabling real-time iterations and improving decision making across the extended enterprise.

Products and Services

Oculus CO takes peer-to-peer (P2P) collaboration to a new level, enabling cross-functional teams to seamlessly exchange data in ways never before possible, developing innovative products faster, smarter, and more economically.

- Works across applications. Extends power of legacy systems & dedicated software.
- Peer-to-Peer environment. Get data directly from source; scaleable.
- Secure, discrete data sharing. Protects intellectual property; maintains model integrity.
- Real-time connectivity. Current and accurate data exchange; timely decision making.

Engineers at Ford Motor, for example, searching for ways to improve the fuel efficiency of its vehicles are using Web collaboration technology to share design changes and other information with engineers and suppliers scattered around several locations. That way, they can instantly analyze how a proposed design change would affect a vehicle's fuel economy. Analysis that might have taken three days can now be completed in less than a minute. That's important as Ford races to make good on a promise to boost the fuel economy of its sport-utility vehicles 25% by 2005. Just as significantly, the technology will shave $5 million to $15 million off a vehicle's development costs. While that's small change on a car that costs $2 billion to develop, the savings would be sizable if it were applied companywide, Ford says.

Position

Oculus Technologies' peer-based collaboration software places it squarely in the middle of the supply-chain, demand-chain virtual exchange space.

OpenCola, Inc.

Description

OpenCola is a provider of distributed network services technologies. OpenCola's technology enables a busienss to centralize what is strategically important (index/reputation) and to distribute costs (processing/bandwidth/storage).

Products and Services

- **OpenCola Folders.** With OpenCola Folders on a desktop, finding a steady stream of timely relevant documents is a matter of dropping "seed" files onto the folder. Thereafter, the folder will fill with a steady load of related files gathered from the Internet. OpenCola Folders use a combination of machine intelligence, human decisions gathered from across the OpenCola network, and direct user feedback to continuously improve results and adjust to users' changing tastes. For content providers, OpenCola Folders provides a "relevance switched" network, where every document is automatically brought to the attention of the users to whom it is relevant.

- **OpenCola Swarmcast.** Swarmcast is a large file distribution system for content and network providers, that serves large files significantly faster, using less Web-server bandwidth. Before Swarmcast, the hottest, most desirable files on-line were also the slowest, least available files — the more concurrent downloads a file had, the worse the service to each of the downloaders became. Content providers' only recourse was to throw money at the problem, adding bandwidth and mirroring servers. Swarmcast converts large, unwieldy files (audiovisual media, software, images, video games, mods, etc.) into nimble chunks that swarm through the network, re-assembling themselves on the user's computer at maximum speed.

Address

81 Langton Street, Suite 13
San Francisco, CA 94103
(415) 437-6161
www.opencola.com

Team

Management

Grad Conn	CEO
Cory Doctorow	Chief Evangelist
John Henson	CTO

Board of Directors

Grad Conn	CEO
John Abraham	Venture Partner, Battery Ventures
Joe Farina	President & COO, Genuity Inc.
Tom Koutoulakis	Partner, Cassels Brock & Blackwell LLP
Vernon Lobo	Managing Director, Mosaic Venture Partners
David Samuel	Managing Director, Mosaic Venture Partners
Mark Sherman	General Partner, Battery Ventures

Investors

Battery Ventures
Mosaic Venture Partners
Torstar Corporation

$13 million Series B
$16 million to date

Competitors

PurpleYogi, EverNet Systems, Allibra, MangoSoft

OpenDesign, Inc.

OPENDESIGN

Address

12721 Bel-Red Road
Bellevue, WA 98005
(425) 467-5500
www.opendesign.com

Team

Management

Edward Jung	President, Acting CEO
Alexander J. Cohen	Chief Evangelist
Gregory Galliford	CFO
Steve Bush	VP of Engineering
James Murray	VP of Technology

Board of Directors/Advisors

Nathan Myhrvold, Ph.D.	Founder
Bill Reid	General Manager Global Networking & Systems Process Automation & Analysis, Microsoft

Investors

Nathan Myhrvold, Ph.D.
Undisclosed amount

Competitors

XDegrees
Applied MetaComputing

Description

OpenDesign's mission is to unleash the power of collaborative commerce with its Smart Application Routers. OpenDesign's proprietary technology combines the best of P2P, client-server, and routing architectures to facilitate business process integration along the supply chain. OpenDesign, Inc. is comprised of a leading team of technologists from CNET, E*Trade, Microsoft, and Netscape and has offices in Bellevue, Washington, and South San Francisco, California.

Products and Services

Smart Application Routers for collaborative commerce.

Position

OpenDesign's router facilitates the integration of disparate data, business logic, applications, and computing environments into flexible applications that can be deployed throughout and across enterprise networks. This router enables dynamic business applications to adapt to real world changes and operate across firewalls and computing environments. Specific benefits of the OpenDesign Smart Application Router include:

- **Efficiency Gains.** Greater automation of business processes, easier integration of partners and applications, lower working capital, and easier outsourcing.
- **Revenue Gains.** Faster response time to opportunities, accelerated time to market, greater ability to meet demand hikes, and support more products and enable better service.
- **Risk Reduction.** More accurate forecasts, assured data integrity, high application reliability with load-balancing, and fault tolerance technology to adjust to traffic peaks.

Parabon Computation, Inc.

Description

Parabon's distributed computing solution harnesses the power of idle computers to power applications on a supercomputing scale. Clients can access the power of a broad network of Internet-connected computers as needed or optimize their own enterprise computing resources by deploying a complete distributed computing solution inside the corporate firewall. A platform-independent development environment empowers developers and ensures applications can run securely, any time, from anywhere.

Products and Services

Parabon offers clients a choice of distributed computing platforms:

- offering unprecedented computing power, on demand, from any Internet-connected computer;
- Frontier Enterprise uses distributed computing technology to harness the wasted power of an organization's internal infrastructure, from PCs to mainframes;
- Prospector brings new efficiency to bioinformatics research with genetic sequence comparison and analysis of massive datasets;
- Exhaustive Regression analyses any number of variables in all combinations; and
- rendering using Monte Carlo algorithm creates photo-realistic animation.

For organizations with computationally challenging problems that require supercomputing-powered applications, Frontier delivers computing power even to a desktop computer.

Position

Competitors in Internet and Enterprise distributed computing include DataSynapse, Entropia, and United Devices. Parabon is distinguished among distributed computing companies by its security and platform independence, which enable the company to support many and varied independent developers, resulting in a rich set of distributed applications. The company also has the additional offerings of robust packaged applications, professional services, and custom application development.

First distributed processing platform to publish SDK.

Address

3930 Walnut Street, Suite 100
Fairfax, VA 22030-4738
(703) 460-4100
www.parabon.com

Team

Management

Steven L. Armentrout, Ph.D. — Founder, President, CEO
James Gannon — CTO
Mark Weitner — VP of Sales & Marketing
James O'Connor — VP of Engineering
Antony Davies, Ph.D. — Chief Analytics Officer
Peter Wentz — Director of Operations
Paula Armentrout — Director of Human Resources

Board of Directors

Craig Fields, Ph.D. — Chairman of Defense Science Board and former DARPA Director
Steven L. Armentrout, Ph.D. — Founder, President, CEO
Mark Weitner — VP of Sales & Marketing

Professional Advisors

Legg Mason, Hogan and Hartson, Exodus, Ogilvy PR Worldwide, IDC

Investors

$6.5 million — Seed

Customers

25 beta customers

Partners

National Cancer Institute, National Human Genome Research Institute, University of Maryland.

Competitors

Entropia, DataSynapse, United Devices

Stats/Misc

Founded in 1999
Over 50 full-time employees

Proksim Software, Inc.

Description

Proksim Software is a provider of advanced Internet communication infrastructure software. The company's revolutionary networking paradigm provides Internet-based distributed applications with scalability and synchronization in the delivery of time-sensitive information. With Proksim's Telco-grade enabling technology, network service providers and content providers bring value-added services quickly and efficiently to end users. The company targets two high-growth markets with T-SIS (Time Sensitive Information Sharing) for online commercial applications and Net-Z for online infotainment applications.

Products and Services

Net-Z is the duplicated object platform to enable online gaming to be rolled into content service providers' (CSP) infrastructure. Fully tooled to provide simulation, monitoring, and auditing at all levels, it lets CSPs offer rich bandwidth adjusted gaming experience that can be turned into a revenue source.

T-SIS is a platform to create Internet Computing applications to be rolled into CSP infrastructure. Like Net-Z, it is fully tooled to provide monitoring and auditing at all levels. Furthermore, it lets CSPs offer Service Level Agreement type experience at the application level. In fact, not only with guaranteed bandwidth reservation, but on-site and off-site load balancing and/or fault tolerance can be turned into premium revenue.

Position

Focus is on the large multi-player online gaming market.

Address

816 Congress Ave., Suite 1100
Austin, TX, 78701
(512) 493-5764
www.proksim.com

Team

Management

Jean-Guy Lacombe	President & CEO
Harald Lotz	EVP, Worldwide Sales
Laurent Visconti	VP of R&D
Sylvain Beaudry	Founder & CTO
Marie Cloutier	Director of Marcom
Eric Bolduc	Finance Director & Controller
François Vigneault	VP of Bus. Dev.

Board of Directors

Sylvain Beaudry	Proksim Software
André de Fusco	CEO, D-Star Technologies Inc.
Martin Duchaîne	Société Innovatech du Grand Montréal
André Duquenne	Vice President, T2C2 Capital L.P.
Jean-Guy Lacombe	Proksim Software
Guy Lemieux	President and COO, Bell Mobility Radio

Investors

Société Innovatech du Grand Montréal
T2C2 Capital L.P.

$3.9 million — Series A

Stats/Misc

Nortel Networks partnership

QUIQ, Inc.

Description

QUIQ, Inc. offers a customer support solution that reduces costs and enhances loyalty through mass collaboration for Global 2000 companies.

Products and Services

QUIQ Connect

- Core Product Components
- QUIQ Knowledge Network
- QUIQ Index Server
- QUIQ Talk Natural Language Interface

Optional Enhancements

- Expert Escalation and Editorial Workflow
- NNTP Newsgroup Bi-Directional Posting
- Survey Topics
- Search Integration

Position

QUIQ draws customers together with employees, partners, and industry experts to share information and help each other find solutions to their service inquires. QUIQ Connect leverages shared knowledge to dynamically deliver resolutions and to fuel self-service content.

Address

2121 South El Camino Real, 10th Floor
San Mateo, CA 94403
(650) 294-2900
www.quiq.com

Team

Management

Greg Richardson — President & CEO
Dr. Raghu Ramakrishnan — Co-Founder, CTO
Kartik Ramakrishnan — Co-Founder, VP of Business Development
Pranav Mohindroo — VP of Client Services and Products
Brock Alston — VP of Sales
Richard Buchanan — VP of Marketing
Jesus Ortiz — VP of Engineering

Board of Directors

Greg Richardson — President and CEO
Dr. Raghu Ramakrishnan — CTO
Ho Nam — Partner, Altos Ventures
Arnold Silverman — Discovery Ventures
Stephen Bowsher — General Partner, InterWest Partners

Investors

InterWest Partners
BancBoston Ventures
Altos Ventures
Discovery Ventures

$15 million — Series B
$20 million to date

Customers

National Instruments, Quaero, eChips, AlphaSmart, Network Appliance, Packtion

Competitors

Kanisa, AskJeeves, Primus, Service Ware, CoolBoard, Orbital

Scale Eight, Inc.

Address

625 Second Street, Suite 101
San Francisco, CA 94107
(877) 372-0956
www.scale8.com

Team

Management

Dick Watts	President & CEO
Joshua Coates	Founder & CTO
Patrick Rogers	Vice President, Product Mgmt & Business Development
Wendi Norris	VP of Marketing
Dave McDonald	Vice President, Finance & Administration
Chris Bruno	VP of N. American Sales
Jim Dawson	VP of International Sales
Keith Falk	VP of Human Resources

Board of Advisors

David Patterson	Chief Scientist, Professor, U.C. Berkeley
Steve Wallach	VP of Technology, Chiaro Networks
David Culler	Professor, U.C. Berkeley
Donald Lobo	Founding engineer and “Technical Yahoo!,” Yahoo! Inc.
George Symons	VP of Availability Pdts, Legato Systems

Investors

CenterPoint Ventures
Crown Advisors
InterWest Partners
Oak Investment Partners

$26.5 million	Series B

Customers

Akamai, MTVi Group’s Web properties (MTV.com, VH1.com, Country.com, and Sonicnet), Vingage

Description

Scale Eight’s mission is to become the dominant provider of storage services to the Internet by offering distributed storage services via its proprietary technology. The growing demand for multimedia files and complex data raises the issue of how these objects will be stored and delivered to end-users efficiently and cost-effectively.

Scale Eight has determined that by creating a distributed network infrastructure on commodity storage devices, the company can offer Internet-based storage services at a fraction of the cost of a home-grown Wide Area Storage Network.

Products and Services

Scale Eight MediaStore is the company’s core offering. MediaStore enables customers to access chosen files from any server or browser. Scale Eight enables this by installing a single file server called a MediaPort in a customer’s LAN through which the customer can access files. The function of the MediaPort is to cache frequently requested files, and to store and retrieve files from Scale Eight StorageCenters.

The MediaPort is grounded on a proprietary file system called the Scale Eight Global File System (8FS) which can offer customers a holistic image of the geographically distributed file system (aggregate picture of all the LANs). Files that enter Scale Eight’s network are mirrored (replicated/cached) across multiple facilities, and load balanced for optimal delivery.

Customers access their files (all file types supported) either through their LAN (through MediaPort) or through a browser (if accessing through the WAN) through a proprietary naming system (8RL, an authenticated URL) that uniquely tags each file.

Scale Eight operates four StorageCenters — two in the U.S. (California and Virginia) and one each in London and Tokyo.

Service plans start at 300 GB. Customers can add capacity as needed, in real time. Typically, a managed mirrored TB per month costs $25,000, roughly 80% less than StorageNetworks’ cost per TB per month. Scale Eight can offer this kind of pricing because the company uses commodity hardware. The “special sauce” is the software that powers the mirroring, load balancing, routing, and caching.

Static Online, Inc.

Address

The Bradbury Building
304 S. Broadway, 5th Floor
Los Angeles, CA 90013
(213)·617-6900
www.static.com

Team

Management

Jay Haynes	CEO
Damien Stolarz	CTO & Founder
Dave Burchianti	VP of Marketing
Patrick DeMarco	VP of Sales

Investors

Zone Ventures
Tim Draper

$5 million	Series A

Customers

Several paying customers

Competitors

OpenCola, EverNet Systems, Allibra

Description

Static is a provider of a software solution that reduces bandwidth and server costs for streaming media providers. By aggregating and utilizing the unused bandwidth of streaming media viewers, the Static Streamer enables content providers to save significant costs. Because the available bandwidth scales with the actual number of users, content providers can scale efficiently while eliminating costly mistakes in predicting traffic.

Products and Services

- **The Static Streamer** is a software solution that reduces bandwidth and server costs for streaming media providers.

  The Static Streamer works by aggregating the unused bandwidth of users watching streaming media from content providers. Users transparently take advantage of available bandwidth from other users in the network. An initial stream begins at either a content provider or a content delivery network origin server. However, additional requests for the stream are automatically redirected to other unused or under-utilized Static Servers with optimal resources based on a number of different metrics including ping time, quality of connection, and other key criteria.

United Devices, Inc.

Address

12675 Research, Building A
Austin, TX 78759
(512) 331-6016
www.ud.com

Team

Management

Ed Hubbard	CEO & Co-Founder
Dr. David Anderson	CTO
Dr. Jikku Venkat	VP of Engineering
Lance Hay	CFO & Co-Founder
Becky Edgar Melton	VP & General Counsel
David Wilson	VP of Marketing & Bus. Dev.

Investors

SOFTBANK Capital
Oak Investment Partners

$13 million — Series A

Customers

Exodus, iArchives, Sponsor of National Foundation for Cancer Research (NFCR) Centre for Drug Discovery in the Department of Chemistry at the University of Oxford, England

Competitors

Entropia, DataSynapse, Parabon

Stats/Misc

63 full-time employees

Description

United Devices develops and manages the software and infrastructure required to aggregate idle computation, storage, and bandwidth resources via the company's MetaProcessor Platform. The MetaProcessor Platform is a scaleable, secure software package that allows customers to access more computing power at lower overall project costs. It works by tapping into the idle computing resources of thousands of individual PCs connected to a network. The networked resources can be as focused as a private corporate system or as broad as the Internet.

Products and Services

- **MetaProcessor Platform.** Organizations that want to employ distributed computing in their own corporate networks can purchase the MetaProcessor Platform software package.

- **Global MetaProcessor Service.** Commercial and research organizations that need to solve very large-scale computing problems can harness the power of the Internet through this a full-support, outsource service. The service provides access to the combined computing power of the United Devices Member Community.

- **MetaProcessor Applications.** MetaProcessor Applications are ready-to-use with the MetaProcessor Platform on a corporate intranet or with the hosted Internet service.

- **MetaProcessor Software Developers Kit.** The MetaProcessor SDK allows application developers to create custom applications for the MetaProcessor Platform. The SDK provides a programming environment that makes it easy to use existing application code and integrates with the MetaProcessor platform for safe and easy deployment.

Uprizer, Inc.

Description

Uprizer was founded by Ian Clarke, creator of decentralized file-sharing technology, Freenet.

While we do not know the specifics of Uprizer, we believe the company is most likely leveraging anonymous file-transfer technology from Freenet (the least well-know variety of the early three major file-transfer protocols — Napster and Gnutella are the other two) to create LAN- and WAN-based network services.

Address

www.uprizer.com

Team

Management

Steven Starr	Chairman
Rob Kramer	CEO & President
Ian Clarke	CTO
Steve Winshel	SVP Business/Technology
Dave Chura	Senior Director, Engineering

Board of Directors

Steven Starr	Chairman, Uprizer
Rob Kramer	Director, Uprizer
Ian Clarke	Director, Uprizer
Frank Kline	Director, Kline Hawkes & Company
Jon van Bronkhorst	Director, Shugart Venture Fund
Kurt Sehnert	Observer, Intel

Investors

Intel
Marc Andreesen
Kline Hawkes & Company
Shugart Venture Fund

$4 million	Series A

Competitors

OpenCola
Static Online

Stats/Misc

Founded by Ian Clarke, creator of Freenet

WorldStreet Corp.

Description

WorldStreet Corporation is a developer of peer networking solutions. WorldStreet provides the Internet infrastructure and services required to support online, business-to-business communities.

Products and Services

WorldStreet's latest product, WorldStreet Net, is the first peer networking solution initially targeting the financial industry. WorldStreet Net provides a single, integrated communications channel for members of qualified communities of industry professionals. It connects securities market participants in a unified environment in which information and ideas are shared between interested and entitled parties. This collaboration occurs over a distributed network that integrates directly into firms' existing applications, workflows, and business processes.

WorldStreet's customers include leading global investment banks Scudder Kemper, American Century, Janus, Invesco, JP Morgan Investment Management, Independent Investment Advisors, Boston Partners, Glenmede Trust, JP Morgan, Deutsche Bank, ING Barings, UBS Warburg, and WR Hambrecht.

WorldStreet's content partners include the following: Comtex, First Call, Hoover's, Interactive Data, Market Guide, and Technimetrics. WorldStreet is strategically backed by the industry it serves. Its investment partners include J.P. Morgan, American Century, Reuters, Deutsche Bank, Hambrecht & Quist, Cambridge Technology Fund, North Bridge Venture Partners, Advanced Technology Ventures, and Gilde Investment Management.

Position

WorldStreet is initially focused on the financial industry, where the exchange of investment ideas and information can be improved with the company's product offerings.

Address

465 Medford Street
Boston, MA 02129-1454
(617) 918-2500
www.worldstreet.com

Team

Management

Bruce Fador	President & CEO
Rod Hodgman	EVP & COO
Bob Lamoureux	SVP & CTO
Mark J. Fitzpatrick	SVP & CFO
Alexis Kopikis	Founder, SVP of Product Strategy
Shawn Henry	SVP, Product Management & Solutions
Tim Walker	SVP of Sales
Peter Dychkewich	VP of Human Resources
John Riley	VP of Marketing
Tom MacDougall	VP, Information Technology

Board of Advisors

David Marcus — Co-Founder

Board of Directors

Paul Walborsky — Co-Founder

Investors

Advanced Tech. Ventures
American Century
Cambridge Technology
The Deutsche Bank Group
Gilde Investment
Hambrecht & Quist
J.P. Morgan Capital
NorthBridge Venture Partners

$30 million — Series C
$50 million to date

Customers

Deutsche Bank, ING Barings,
JP Morgan

XDegrees Corp.

Address

2307 Leghorn Street
Mountain View, CA 94043
(650) 691-0400
www.xdegrees.com

Team

Management

Michael Tanne	CEO & Co-Founder
Dan Teodosiu	CTO & Co-Founder

Board of Directors

Michael Tanne	CEO & Co-Founder
Dan Teodosiu	CTO & Co-Founder
Jeff Brody	General Partner, Redpoint Ventures
Chris Moore	Associate, Redpoint Ventures
Venky Harinarayan	General Partner, Cambrian Ventures

Board of Advisors

Craig Donato	Former VP, Excite, Current CEO, Grand Central
Joe Pistritto	Former VP Engineering Zaplet, FaceTime, PointCast

Investors

Redpoint Ventures
Cambrian Ventures

$8 million	Series A

Customers

Product has not been launched
Beta customers to be announced in Q2

Competitors

OpenDesign
Applied MetaComputing

Stats/Misc

Employees: 25
CEO Michael Tanne – Founder of AdForce
CTO Dan Teodosiu – Founder of Xift; Sr. Scientist, HP Labs; Ph.D. in Operating Systems, Stanford.

Description

XDegrees provides its core eXtensible Resource Name System (XRNS) technology to allow for identification, addressing, security, and access of resources distributed throughout networks.

Products and Services

XDegrees offers its XRNS server-based resource directory system on both a licensed and hosted basis to enterprises and application developers.

Position

XDegrees provides its core XRNS technology to application developers and enterprises to enable Network Applications that utilize both client-server and peer-to-peer approaches to leverage all resources of internal and Internet-wide networks.

ZAMBEEL

Zambeel, Inc.

Description

Address

www.zambeel.com

Team

Management

Sid Agrawal — Founder & CEO
Dr. Waheed Qureshi — Founder & CTO
Amar Rao — Vice President, Strategy & Business Development
Teresa Malo — Vice President, Finance & Administration

Board of Directors

Sid Agrawal — Founder & CEO
Dr. Waheed Qureshi — Founder & CTO
Dick Kramlich — Co-Founder & General Partner, New Enterprise Associates
Bernie Lacroute — Partner, Kleiner Perkins Caufield & Byers

Board of Advisors

Steve Anderson — Associate Partner, Kleiner Perkins Caufield & Byers
Forest Baskett — Venture Partner, New Enterprise Associates
David Cheriton — Professor, Stanford
Peter Danzig — CTO, Akamai
David Dewitt — Professor, U of Wisconsin
Vinod Khosla — Partner, Kleiner Perkins Caufield & Byers
Anil Nori — VP & CTO, Asera
Scott Sandell — General Partner, New Enterprise Associates

Investors

Kleiner Perkins Caufield & Byers
New Enterprise Associates
Integral Capital Partners

While Zambeel remains in stealth mode, we have pieced together a few details that may provide the reader with a general idea of the company's vision.

Located in Fremont, California, and backed by Kleiner Perkins Caufield & Byers, New Enterprise Associates, and Integral Capital Partners, Zambeel was founded by several industry veterans with expertise in distributed systems.

Zambeel is developing a storage architecture that is aimed at resolving the problems of expensive proprietary storage systems and their highly centralized nature that limit flexibility and geographic coverage.

The company will be introducing a distributed storage architecture using commodity hardware that is able to form clusters of thousands of machines across vast distances. Unlike Scale Eight, which is offering a service based on its technology, Zambeel is attempting to become the arms merchant for distributed storage systems.

What this fully distributed architecture enables is storage capacity on demand, from anywhere, for all data types. Using complex caching, load balancing, mirroring and data duplication algorithms, and a fresh library of metadata, Zambeel will enable storage service providers and corporations to create a fully redundant, secure failover network spanning vast distances at a fraction of the cost of competitive solutions (SAN and NAS).

Zodiac Networks, Inc.

Description

We believe Zodiac is a distributed network services provider and will likely compete with Akamai, Inktomi, OpenCola, Static Online, and EverNet Systems, among others.

The company is currently in stealth.

Address

1350 Villa Street, Suite 200
Mountain View, CA 94041
(650) 623-0980
www.zodiacnetworks.com

Team

Management

Marc Andreesen
Mike Homer
Wade Hennessey

Investors

Kleiner Perkins Caufield & Byers
Benchmark Capital
The Barksdale Group

BEAR STEARNS

Appendices

BEAR STEARNS

On the Internet, information is broken down into packets by the transmission control protocol (TCP) and is later re-assembled in the proper order. The Internet protocol (IP) is responsible for making sure the packets are sent to the right destination.

Each packet is given a header that contains a variety of information, such as the order in which the packets should be assembled with other, related packets.

Routers receive data via input ports. When an input port receives a packet, a software routine called a routing process is run. This process looks inside the header information in the data packet and finds the address where the data is being sent. It then compares this address against an internal database called a routing table. The routing table has detailed information about the ports to which packets with various IP addresses should be sent. Based on what it finds in the routing table, the router sends the packet to a specific output port. The output port then sends the data to the next router — or to the destination itself.

At times, packets are sent to a router's input port faster than the port can process them. When this happens, the packets are sent to a holding area called an input queue — an area of RAM (random access memory) on the router. Each input port processes packets from the queue in the order in which they were received ("FIFO").

If the number of packets exceeds the capacity (length) of the queue, packets may be lost. When this happens, the TCP protocol on the sending and receiving computers can have the packets re-sent.

As TCP creates each packet, it also calculates and adds to the header a checksum, which is a number that TCP uses on the receiving end to determine whether any errors have been introduced into the packet during transmission.

As the packets arrive at their destination, TCP calculates a checksum for each packet. It then compares this checksum with the checksum that has been sent in the packet. If the checksums do not match, TCP knows that the data in the packet has been corrupted during transmission. It then can discard the packet and ask that the original packet be re-transmitted.

TCP/IP is used because the Internet is a packet-switched network meaning there is no single/unbroken connection between sender and receiver. Instead, when information is sent, it is broken into small packets, sent over many routes at the same time, and then re-assembled at the receiving end.

By contrast, the traditional telephone system is a circuit-switched network; once a connection is made, that part of the network is dedicated only to that single connection.

XML

XML is the Extensible Markup Language, extensible because it is not a fixed format like HTML. Unlike HTML, XML is not a single, predefined markup language. It is a metalanguage — a language that describes other languages.

The advantage of XML is that it can be used to define and customize languages, while HTML already defines the way to describe a class of objects or documents.

The primary difference between HTML and XML is this: Whereas HTML defines the rules associated with how things are displayed, XML defines what those things contain. As such, XML is a metalanguage. Put another way, HTML tells us what a thing looks like, while XML tells us what a thing is.

In Exhibit 105, we graphically display an HTML page. It is a typical page on Amazon.com. Below the page, we display portions of the actual HTML code. As we can see, HTML tells us nothing about the content of the code. The HTML merely defines how the documents and objects should appear in a browser.

Exhibit 105. Example of HTML Page

HTML

```
<body>
<b> Harry Potter and the Goblet of Fire (Book 4)</b><br>
by J. K. Rowling, Mary Grandpre (Illustrator)<br>
<b>Our Price: $15.57</b><br>
<b>Availability:</b> Usually ships within 24 hours.<br>
<b>Reading level:</b>
Ages 9-12<br>
<b>Hardcover</b>
 - 734 pages (July 8, 2000)
</body>
```

As distilled, the HTML translates into the following presentation in Excel.

Harry Potter and the Goblet of Fire (Book 4)
by J. K. Rowling, Mary Grandpre (Illustrator)
Our Price: $15.57
Availability: Usually ships within 24 hours.
Reading level: Ages 9-12
Hardcover - 734 pages (July 8, 2000)

Source: Amazon.com; Bear, Stearns & Co. Inc.

The problem with HTML is it has no semantics — that is, there is no indication of what the words and the objects represent. There is no intelligence in the language. HTML is unable to tell us what a word or phrase means.

XML is different. In Exhibit 106, we see a typical Web page. It looks identical to an HTML page. In fact, presentation wise, HTML and XML are indistinguishable. We display the page as it would appear in a browser. Below it, we display the XML code associated with the page.

Exhibit 106. Example of XML Page

Chris Kwak	Harry and the Hindershots	McMillon	200	$4.50
Robert Fagin	Howdy, or Tales of Country Poo	Billibong Publishing	376	$18.95

XML

```
<?XML VERSION = "1.0" ?>
<!DOCTYPE ARTICLE [
<!ENTITY WD "WebDeveloper">
<!ELEMENT ARTICLE (TOPIC)*>
<!ELEMENT TOPIC (TITLE,PUBLISHER,AUTHOR,PRICE,PAGES)>
<!ELEMENT TITLE (#PCDATA)>
<!ELEMENT PUBLISHER (#PCDATA)>
<!ELEMENT AUTHOR (#PCDATA)>
<!ELEMENT PRICE (#PCDATA)>
<!ELEMENT PAGES (#PCDATA)>
]>
<ARTICLE>
 <TOPIC>
              <TITLE>Harry and the Hindershots</TITLE>
              <AUTHOR>Chris Kwak</AUTHOR>
              <PUBLISHER>McMillon</PUBLISHER>
              <PRICE>$4.50</PRICE>
              <PAGES>200</PAGES>
 </TOPIC>
 <TOPIC>
              <TITLE>Howdy, or Tales of Country Poo</TITLE>
              <AUTHOR>Robert Fagin</AUTHOR>
              <PUBLISHER>Billibong Publishing</PUBLISHER>
              <PRICE>$18.95</PRICE>
              <PAGES>376</PAGES>
 </TOPIC>
</ARTICLE>
```

Source: WebDeveloper.com; Bear, Stearns & Co. Inc.

We can see how XML differs from HTML. Unlike HTML, XML can actually tell us what a word or phrase refers to. In the case of Chris Kwak, we know that these words refer to Author (<AUTHOR>Chris Kwak</AUTHOR>). Therefore, in addition to describing how an object should look, XML tells us what the object represents.

This is because XML is object oriented. As an object-oriented language, XML enables us to treat virtually anything on the Internet as an object — words, objects, documents, system resources. By componentizing the language of the Internet, XML enables us to query, share, distill, and call anything on the Internet.

SOAP

SOAP (Simple Object Access Protocol) is a lightweight XML protocol for exchange of information in a distributed and decentralized environment, enabling the access of objects, servers, and services in a platform-independent manner.

A protocol created by Microsoft, DevelopMentor, and Userland Software and backed by companies that include IBM, Lotus, and Compaq, SOAP gives a big push toward the vision of distributed and decentralized networked systems and services. What SOAP enables the network to become is one where data and resources can be treated as objects that can be accessed in a uniform manner. With SOAP, we can do whatever we like with data on the Internet at any point along the origin-destination route.

SOAP has three essential elements:

1. Metadata envelope — defines *what* a message contains and *how* to process it.
2. Encoding rules — set for expressing instances of application-defined datatypes.
3. Convention — for representing remote procedure calls[19] and responses.

To understand the value of SOAP, imagine a dynamically generated Web page. Because the Web is HTML-based, attempting to do anything with data resident in HTML outside of a browser is fruitless. Accessing, exchanging, manipulating data on the Web in a uniform manner is made possible with XML and with protocols like SOAP, which articulates how XML data can be processed.

UDDI

The Universal Description, Discovery and Integration (UDDI) specifications define a way to publish and discover information about Web services. The term "Web service" describes specific business functionality exposed by a company, usually through an Internet connection, for providing a way for another company or software program to use the service.

The industry has begun to recognize the importance of Web services for e-commerce. With UDDI, discovering which partners are offering which services is a matter of linking directly to the partner's Web site. In this approach, a description file on a company's Web site can announce the Web services available to its partners. It automates the entire process of description, discovery and integration. By tying into an XML-based registry, all information regarding a customer or partner can be updated and accessed.

[19] RPC – Remote Procedure Call. RPC is a protocol that one program can use to request a service from another program on another computer in a network. A *procedure call* is a function call. RPC allows users to work with remote procedures as if the procedures were local and, therefore, RPC extends the capabilities of procedure calls across networks and is essential in the development of distributed systems. Each remote procedure call has two sides: a client that makes the call to a server, which sends back a reply.

Exhibit 107. UDDI

Interop Stack	Universal Service Interop Protocols (these layers are not defined yet)
	Universal Description, Discovery Integration (UDDI)
	Simple Object Access Protocol (SOAP)
	Extensible Markup Language (XML)
	Common Internet Protocols (HTTP, TCP/IP)

Source: UDDI.org.

HTTP

HyperText Transport Protocol is critical to what is going on in Internet 3.0. HTTP is the transport specification and details how HTML is moved around the Internet. It is basic and fundamental to content transport on the Web. In Internet 3.0, HTTP remains fundamental because it is able to exchange all kinds of data.

The Goal Is Transparency[20]

One of the goals of a network system is transparency, or what we have referred to as fluidity of data. Transparency allows data to move easily across the network, so that devices can share data and resources. Barriers between nodes hinder fluid communications. We list some of these barriers below.

Firewalls

Firewalls shield companies from unwanted outside users and control internal users' Web habits. There are several ways that firewalls can be implemented:

- A packet filter blocks traffic based on IP address and/or port numbers.
- A proxy server acts as a relay between two networks, breaking the connection between the two.

NAT

A network address translation (NAT) device hides the IP addresses of client devices in an intranet by presenting one IP address to the outside world. The reason for NATs is that when the 32-bit (2^{32}) IPv4 was introduced, a significant portion of the available IP addresses was allocated to a few select groups. A shortage of IP addresses has forced many networks to use NATs. If a network administrator desires to hide individual devices from the network at large, he can use a NAT to make the device invisible by positioning it behind the NAT.

DHCP

Dynamic Host Configuration Protocol is software that dynamically assigns IP addresses to client machines logging onto a TCP/IP network. It eliminates having to manually assign permanent IP addresses.

[20] For more details, please refer to RFC 2775.

The idea behind DHCP, like NAT, is that IP addresses are resources and a dearth of IP addresses has forced networks to try to conserve them or work around the shortage.

In DHCP, a block of IP addresses is shared among devices. The DHCP server will allocate IP addresses dynamically, based on need and time. By allocating IP addresses dynamically, the network administrator is not forced to allocate fixed IP addresses for all devices on the network. When certain devices are not on, the idea is that these devices should no longer need to waste a resource like an IP address.

The problem with DHCP, like NAT, is that devices on the network do not have fixed IP address identities. Therefore, a device may have a different IP address inter- and intra-sessions.

IPv6: The Solution?

The answer to many of these barriers to transparency may be IPv6. IPv6 uses a 128-bit addressing scheme. Address lengths are four times longer than in IPv4.[21] Therefore, the problem with the shortage of IP address all but disappears under IPv6. In addition, IPv6 offers many security and functionality advantages.

IPv6 would indeed help with transparency and solve many problems plaguing the Internet today. Its biggest hurdle has been implementation, as it is a specification that was first introduced in the IETF in 1995.

---

[21] IPv4. The resolution of a domain name query to the corresponding IP address involves several transactions at the heart of the domain name system. As noted, each domain name corresponds to a 32-bit binary IP address in IPv4, the current IP addressing specification. The IP address consists of four binary octets separated by dots. For example, www.bearstearns.com's IP address is 216.178.174.2. Each dot-separated number represents a binary octet. 216 in binary form is an eight-bit number (11111000); 178 (10110010); 174 (10101110); 2 (00000010). Therefore, the binary iteration of 216.178.174.2 is 11111000.10110010.10101110.00000010 (a). In this case, 216 is the .com part of the Web address in a Class C network (b), 178 represents "bear" and 174.2 refers to a specific computer on the Bear Stearns network hosting bearstearns.com. (Administration of IP addresses is carried out by IANA, the Internet Assigned Numbers Authority.)

Once the IP address is known, a user's query can then be resolved to the physical location of the computer hosting the desired information. If a user types in www.bearstearns.com, once 216.178.174.2 is resolved for the domain name, the user can then take that IP address and contact the server hosting www.bearstearns.com. Generally, a query from a client to a name server is performed by a resolver – the library of routines called by applications when they want to translate (resolve) a DNS name. The resolver is able to reach across the network to locate a root name server – the name server that is able to determine which servers on the Internet house first-level domain information – that is, a server that can resolve the .com TLD, for example. The resolver is then able to determine which name server to contact for a second-level domain. In the process, a resolver may need to consult several name servers. A resolver, once it locates a useful name server, will often cache the resource records from these name servers and establish a link to this server for future reference.

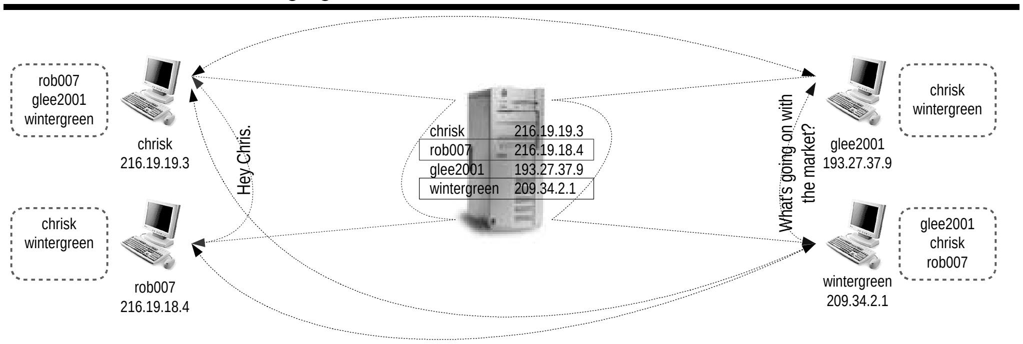
(a) The theoretical number of IP addresses possible in the 32-bit system is 2^32. While this number of potential addresses seemed plentiful at the time, the expansion of the Internet, the wide adoption of domain names, and the proliferation of devices requiring IP addresses (like handhelds) have highlighted the potential shortage of IP addresses.

(b) Several types of classes exist in the IP address structure. Class A, with a decimal value in the first octet between 1-126, is dedicated to certain early networks (ARPA, AT&T); Class B, with a decimal value in the first octet between 128-191, is for secondary networks (Ford Motor, MIT); Class C, 192-223; Classes D and E, 224-247. Each class is dedicated to particular kinds of networks within the domain name space.

INSTANT MESSAGING

Instant messaging (IM) turned mainstream with ICQ from Mirabilis in 1998. Acquired by AOL in 1998, ICQ quickly became the default standard for IM. The architecture of IM depends on a few things: client, name space, and gateways.

Exhibit 108. Instant Messaging Architecture



Source: Bear, Stearns & Co. Inc.

Users download a software client to join an IM network. Once users register with an IM service, their IDs (e.g., chrisk, rob007) are mapped to their IP addresses in a central registry. Linking to other IM users is a matter of querying the central registry to locate users' IDs, and hence their IP addresses. Once the central registry has mapped an individual's ID to an IP address, all messages written to an individual are delivered directly to the individual's device. The central registry identifies and locates the individual's device. Once it has helped identify the individual's device on the network, the central registry gets out of the way.

Below, we list the leading IM clients.

ICQ

With over 100 million registered users, ICQ is by far the largest instant messaging client. Messages are stored by AOL Time Warner.

AOL Instant Messenger

AOL Instant Messenger (AIM) is the sister client to ICQ, with an additional 50 million registered users.

Yahoo! Messenger

Yahoo! has pushed into instant messaging with YIM! by leveraging its 58 million userbase. Yahoo! has begun to offer other functionality like file-sharing on its IM client.

MSN Messenger

Third on the scene, Microsoft has been using its network of Web properties to grab IM share. Microsoft was the first to fully integrate Internet telephony into its IM client. In a recent Jupiter Media Metrix study commissioned by Microsoft, MSN Messenger showed 29.5 million users compared with 29.1 for AIM (does not include ICQ).

Jabber: Letting Us All Gab

Jabber is an XML-based, open-source system and protocol for real-time messaging and presence notification. Jabber provides anywhere access on any device that offers interoperability with other IM and Web-based services, including AOL Instant Messenger, ICQ, MSN Instant Messenger, and Yahoo! Messenger.

Because Jabber is grounded on XML, it can interoperate with any messaging application or platform. The power and promise of Jabber is that it could potentially shatter barriers between proprietary systems. Since XML offers so many advantages (structured data, open formats, and extensibility for content management, directory systems, querying functionality), Jabber may close the loop between disparate messaging networks.

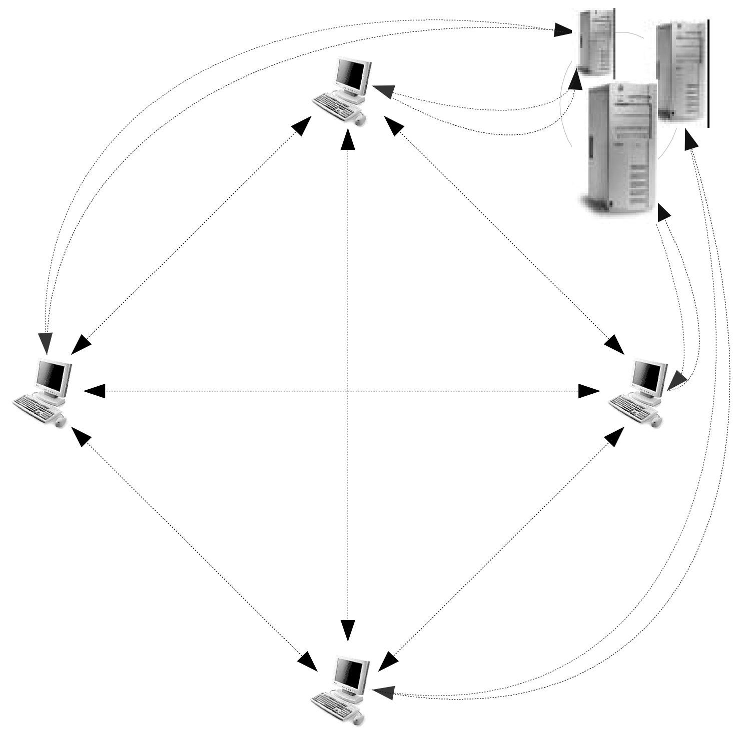
FILE SHARING

Napster popularized file sharing on a peer-to-peer basis. The peer-to-peer file sharing system is extremely elegant. There are several varieties of file-sharing protocols and each has its advantages and disadvantages. We highlight two: Napster and Gnutella.

Napster

Within a Napster-like file-sharing system, the user connects to a central registry server (there are several), much like in IM. This server acts as a gateway to identify users, maintains a metadirectory of user and file information, and connects users to one another.

Exhibit 109. Napster-Style File Sharing



Note: The curved lines represent connections between device and registry server. The registry server answers search queries and returns information (e.g., file name, owner, IP address) to devices. The straight lines represent connections between/among devices.

Source: Bear, Stearns & Co. Inc.

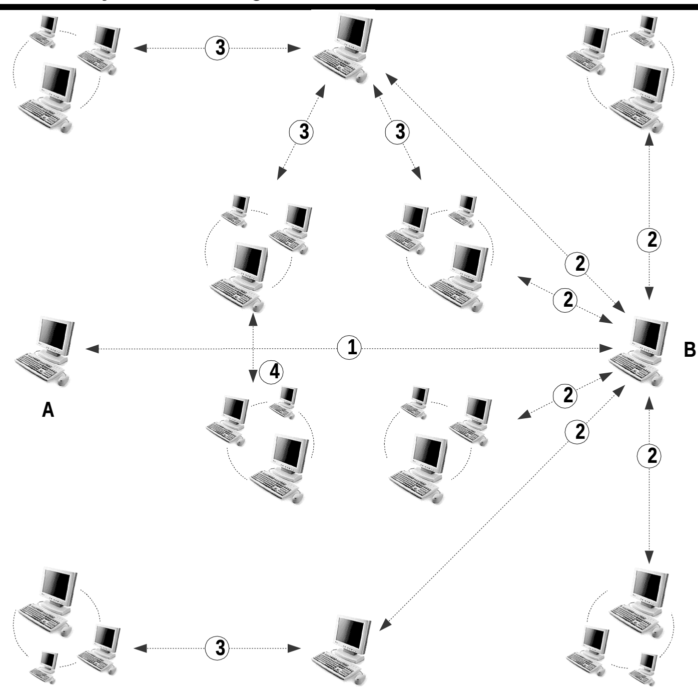
The advantage of the Napster-like system is that the central registry server offers a stable search environment and a metadatabase of songs. While one may be disconnected in the middle of downloading a file from another user's device, one's connection to a Napster server is rarely lost.

The disadvantage with such a system is that searches can reach distant geographies. In adition, because there is ultimately ownership and control of the network (someone maintains the registry servers), the central registry server can be shut down. An example of a Napster-like file-sharing network (apart from Napster) is Napigator.

Gnutella

Gnutella was a genie that a few AOL developers let out of the bottle. Like Napster, Gnutella enables a device to act as a client and a server. However, unlike Napster, a Gnutella network has no central registry server. Connections established by one device to another spiral out in a "six degrees of separation" fashion.

Exhibit 110. Gnutella-Style File Sharing



Note: In a Gnutella system, A establishes a connection with B. A is then connected to the PCs connected to B. Connections spiral out to form a local clustered network of PCs.

Source: Bear, Stearns & Co. Inc.

The advantage of this kind of file sharing system is that there is no central control of a registry server. As such, a Gnutella network is extremely difficult to shut down. Additionally, connecting to a local device ensures that the majority of devices that comprise a Gnutella network are local. TTLs[22] can limit how far a device can be on the network.

The disadvantage of Gnutella is that because the connection to other devices to create the database of users and files is itself prone to random events (e.g., shut down, lost connection), the network can be extremely fragile. Should one lose a connection on the network, one often loses one's connection to everything else. That is, if one loses touch with a friend who's connected to a group of individuals, one will likely lose one's connection to those individuals, unless one establishes a connection with those individuals while one is connected to them. An additional disadvantage is network load. Because discovering other devices on the network willing to participate on the Gnutella network is a matter of PINGing everyone on the LAN, the network can

[22] TTL – Time To Live: In IP, there is a field that indicates the useful life of a data packet. When a data packet is sent, it traverses many routers/devices (hops) before reaching the intended destination. Upon reaching a router/device, the TTL field is decremented by one. If for some reason the TTL field reaches zero, the packet is dropped. Paul Baran, the inventor of packet-switching, included this field so that should a portion of the network be brought down (in a war scenario), the packet wouldn't bounce around endlessly and clog up the network and bog down the router.

The reason why TTL fields are important in Internet 3.0 is that within a distributed network, when a query is sent, there must be a diameter or limit to how far a packet can travel – i.e., how many devices (hops) the packet can traverse. In distributed networks, determining the number of hops a device is from the user and setting the TTL appropriately can have significant ramifications, particularly when evaluating the range (horizon) of a network and the useful resources available to a user. Additionally, because PINGs and queries can consume a significant portion of the local network's bandwidth, setting the TTL field appropriately can be meaningful in tempering bandwidth consumption.

often face congestion. In fact, on a Gnutella network, PINGs can comprise 50% of the message traffic and queries[23] can comprise another 25%. This is why in standard Gnutella, the TTL field is set to seven.

Examples or variants of Gnutella, apart from the original network, include Lime Wire, Bearshare, Toadnode, and iMesh. We are particular fans of Lime Wire.

---

[23] Theodore Hong of Freenet has noted that in relation to Gnutella, even ten queries per second (at 560 bits per query) on a three-connection Gnutella client (if we include the load of other messages on the network) would be enough to saturate a 56 Kbps link [(10 X 560 X 4 X 3) = 67.2 Kbps].

Companies mentioned:

3Com (COMS-5.69)
Adobe Systems (ADBE-39)
Agile Software (AGIL-18)
Akamai Technologies (AKAM-9.16)
Amazon.com (AMZN-14)
AOL Time Warner§ (AOL-51)
Apple Computer, Inc. (AAPL-23)
Ariba Inc. (ARBA-6.34)
Ask Jeeves, Inc. (ASKJ-1.76)
AT&T Corporation (T-22)
Bank of America§ (BAC-55)
BEA Systems (BEAS-34)
BMC Software (BMC-22)
Boeing Company (BA-67)
BP Amoco (BP-53)
British Telecom (BTY-75)
Broadvision (BVSN-5.59)
Brocade Comm. Sys. (BRCD-43)
Cablevision§ (CVC-56)
CacheFlow Inc. (CFLO-7.58)
Celera Genomics§ (CRA-38)
Check Point Software (CHKP-55)
Cisco Systems, Inc. (CSCO-19)
CNET Networks, Inc. (CNET-12)
Comdisco, Inc.§ (CDO-2.37)
Compaq Computer (CPQ-16)
Computer Associates§ (CA-28)
Corning Inc. (GLW-21)
Covad Comm. Group§ (COVDE-1.18)
Dell Computer Corp. (DELL-24)
Deutsche Telekom (DT-22)
Digex Inc.§ (DIGX-15)
Digital Island§ (ISLD-3.36)
DoubleClick, Inc. (DCLK-13)
Dun & Bradstreet (DNB-28)
E*Trade Group (ET-9)
Earthlink Network (ELNK-13)
eBay Inc. (EBAY-56)
Electronic Data Systems§ (EDS-63)
EMC Corp. (EMC-39)
Emulex Corp. (EMLX-38)
Enron Corp. (ENE-57)
Equinix (EQIX-1.27)
Exabyte Corp. (EXBT-1.54)
Excite@Home Corporation (ATHM-4.08)
Exodus Communications, Inc. (EXDS-8.40)
Extreme Networks, Inc. (EXTR-28)
F5 Networks§ (FFIV-11)
First Union (FTU-31)
Ford Motor Company§ (F-28)
Foundry Networks, Inc. (FDRY-17)
Fujitsu Ltd.¤ (FJTSY-68)
General Electric (GE-50)
Genuity Inc. (GENU-2.27)
GlaxoSmithKline (GSK-55)
Global Crossing (GX-13)
Goldman, Sachs§ (GS-95)
Handspring Inc. (HAND-10)
Hewlett-Packard Co.§ (HWP-25)
Hitachi (HIT-105)
iBeam, Inc.§ (IBEM-0.40)
IBM Corp.§ (IBM-114)
Incyte Genomics (INCY-19)
ING Groep (ING-66)
Inktomi Corp. (INKT-7.25)
Intel Corporation (INTC-27)
InterNAP Network Svcs. (INAP-2.95)
Interwoven Inc. (IWOV-15)
J.D. Edwards (JDEC-11)
J.P. Morgan Chase (JPM-47)
Keynote Systems (KEYN-12)
Launch Media (LAUN-0.43)
Legato Systems§ (LGTO-15)
Legg Mason (LM-47)
Level 3 (LVLT-14)
LoudCloud (LDCL-4.40)
Lucent Technologies Inc.§ (LU-9.65)
Mercury Interactive (MERQ-59)
Merrill Lynch (MER-67)
Micromuse (MUSE-35)
Microsoft Corporation (MSFT-68)
Morgan Stanley Dean Witter (MWD-66)
Motorola, Inc. (MOT-15)
National Instruments (NATI-33)
Network Appliance, Inc. (NTAP-23)
Network Associates§ (NETA-12)
New Era of Networks (NEON-6.24)
Nextel Communications (NXTL-17)
Nortel Networks (NT-13)
Novell, Inc.§ (NOVL-4.97)
Oracle Corporation (ORCL-16)
Packeteer§ (PKTR-7.20)
Palm Inc. (PALM-6.86)
PanAmSat§ (SPOT-33)
PSInet, Inc.§ (PSIX-0.19)
QLogic (QLGC-43)
Qwest Communications (Q-38)
RadioShack§ (RSH-28)
RADWARE Ltd. (RDWR-18)
Rare Medium (RRRR-1.09)
Raytheon§ (RTNA-30)
RealNetworks, Inc. (RNWK-8.80)
Research In Motion (RIMM-30)
Resonate (RSNT-3.06)
Reuters Group plc¤ (RTRSY-86)
SAP AG (SAP-36)
SOFTBANK¤ (SFTBF-42)
Sony (SNE-79)
Sportsline.com (SPLN-3.40)
StorageNetworks (STOR-20)
Sun Microsystems, Inc. (SUNW-17)
Sungard Data Systems (SDS-59)
Synopsys (SNPS-57)
Tektronix (TEK-27)
Texas Instruments, Inc. (TXN-37)
Thomson Multimedia (TMS-41)
TIBCO Software§ (TIBX-12)
UBS AG (UBS-148)
VA Linux Systems, Inc. (LNUX-4.50)
VeriSign (VRSN-52)
VERITAS Software§ (VRTS-62)
Versant Object Tech. (VSNT-1.80)
Xerox (XRX-11)
Yahoo! Inc. (YHOO-18)

Above list priced as of the market close on May 15, 2001.

§ Within the past three years, Bear, Stearns & Co. Inc. or one of its affiliates was the manager (co-manager) of a public offering of securities of this company and/or has performed, or is performing, other banking services, including acting as financial advisor, for which it has received a fee.

¤ NOTE TO ACCOUNT EXECUTIVES: Check the CAS system, Chapter 1 Directory, for state Blue Sky restrictions prior to soliciting a client or accepting orders from a client. This report discusses numerous securities, some of which may not be "Blue-sky" qualified for sale in certain states and, therefore, may not be offered to investors in such states. Accordingly, in those states, this report does not constitute a solicitation or offer of such nonqualified securities.

BEAR STEARNS