# Internet Bottlenecks:

## *the Case for Edge Delivery Services*

Though the interface to the World Wide Web is fairly user friendly and straightforward, the inner workings of the Internet are quite complicated. This complexity exists on two levels: 1) the Internet as an entity unto itself, and 2) the Internet as an interacting ecosystem of networks.

This paper explores the complexities of the Internet, providing insight into how and why the Internet works as well as it does, while also identifying the traffic bottlenecks inherent in its design – bottlenecks that can impede good performance and scalability in the face of exponential growth in Internet traffic, thus reducing Web site effectiveness. Internet content delivery services break these bottlenecks, enabling content providers to derive real value from doing business on the Web.

The Internet is a communication infrastructure that interconnects a global community of end users and content providers. The key to its success is the ease with which end users and content providers can use it. The Internet appeals to end users because it operates 24x7 and provides access to an almost unlimited amount of useful data and services, all via a single, simple user interface – a browser. It is also easy for content providers to connect their services to the Internet. By simply purchasing a single network connection from an ISP, a content provider can instantly reach an audience of 377 million users — albeit with varying degrees of success.[1]
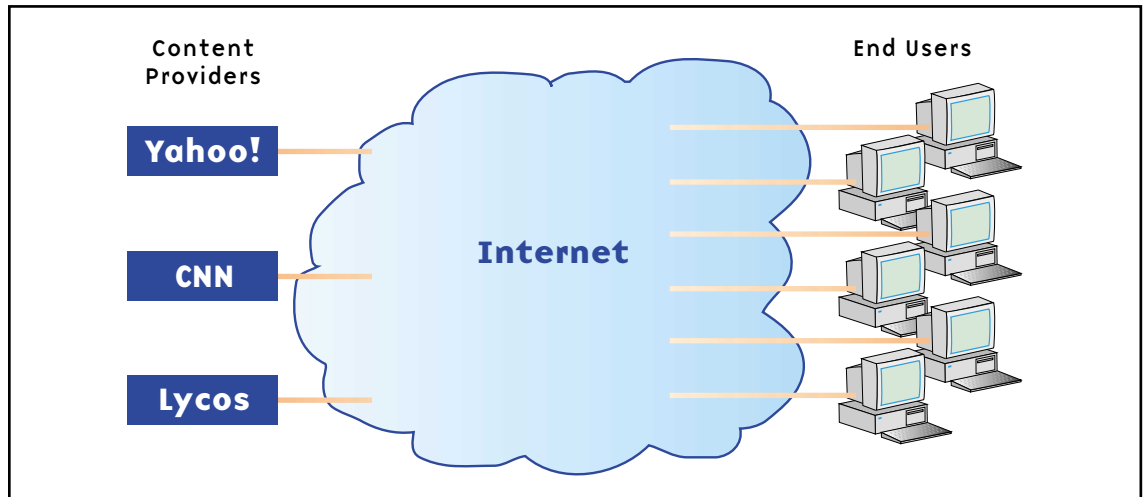


**Figure 1 - Connecting to end users**

While leveraging the Internet offers unprecedented opportunities for content providers to reach new audiences and to take advantage of a new distribution channel, these same businesses cannot reliably conduct e-commerce, serve ads and digital goods, build their brands, and achieve a return on investment if the underlying infrastructure of the Internet is unstable.  The instability and unpredictability of the Internet's performance can have a negative impact on the relationship between businesses and their customers. According to Jupiter Research, nearly half of Web users responding to a Jupiter Consumer Survey said that they have stopped using a preferred site for some amount of time because of performance difficulties; five percent said they abandoned the site permanently.[2] By the same token, a high-performing Web site translates into customer satisfaction and increased market share, which can be correlated to increased purchase conversion and customer loyalty.

Businesses must gain an understanding of the nature of the Internet in order to address the challenges posed by its structure. Once the structural issues of the Internet are understood, businesses can implement the solution to these obstacles, and subsequently reap the rewards of conducting business via the Internet.

## *Structure of the Internet*

By definition, an "internet" is a network of networks. The Internet is a well-known example, being made up of thousands of different networks (also called Autonomous Systems or AS's) that communicate using the IP protocol (see Figure 2). These networks range from large backbone providers such as UUNet and PSINet to small local ISPs such as Visinet in Toronto and Access Toledo in Ohio. Each of these networks is a complex entity in itself, being made up of routers, switches, fiber, microwave links, and ATM technology. All of these components work together to move data packets through the network toward their specified destinations.
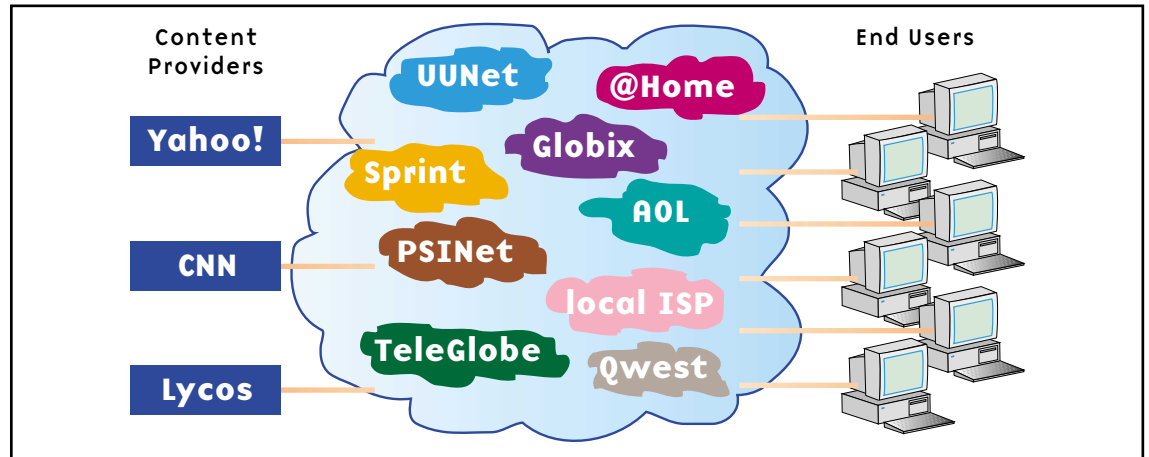
**Figure 2 - The Internet is a network of networks**

In order for the Internet to function as a single global network interconnecting everyone, all of these individual networks must connect to each other and exchange traffic. This happens through a process called **peering**. When two networks decide to connect and exchange traffic, a connection called a peering session is established between a pair of routers, each located at the border of one of the two networks. These two routers periodically exchange routing information, thereby informing each other of the destination users and servers that are reachable through their respective networks.

There exist thousands of peering points on the Internet, each falling into one of two categories: public or private. Public peering occurs at major Internet interconnection points such as MAE-East, MAE-West, and the Ameritech NAP, while private peering arrangements bypass these points. Peering can either be free, or one network may purchase a connection to another, usually bigger, network.

Once the networks are interconnected at peering points, the software running on every Internet router moves packets in such a way as to transport each request and data packet to its correct destination. For scalability purposes, there are two types of routing protocols directing traffic on the Internet today. **Interior gateway protocols** such as OSPF and RIP create routing paths *within* individual networks or AS's, while the **exterior gateway protocol** BGP is used to send traffic *between* different networks. Interior gateway protocols use detailed information on network topology, bandwidth and link delays to compute routes through a network for the packets that enter it. Since this approach does not scale to handle a large-scale network composed of separate administrative domains, BGP is used to link individual networks together to form the Internet. BGP creates routing paths by simply minimizing the number of individual networks (AS's) a data packet must traverse. While this approach does not guarantee that the routes are even close to optimal, it supports a global Internet by scaling to handle thousands of AS's and allowing each of them to implement their own independent routing policies.[3]

Peering points and routing protocols thus connect the disparate networks of the Internet into one cooperating infrastructure. Connecting to one of these networks automatically provides access to all Internet users and servers.

This structure of the Internet as an interconnection of individual networks is the key to its scalability. It enables distributed administration and control of all aspects of the system: routing, addressing, and network buildout. However, inherent in

this architecture are four types of bottlenecks that, left unaddressed, can slow down performance and impede the ability of the Internet to handle a quickly growing number of users, services, and traffic. As illustrated in Figure 3, these bottlenecks occur at the following points in the Internet infrastructure:

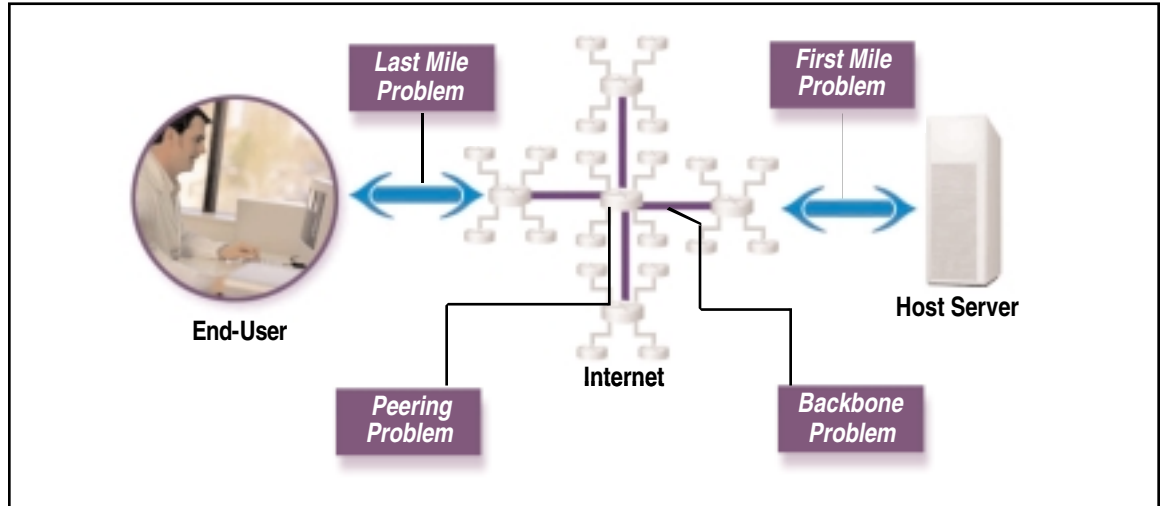1. First Mile    2. Peering Points    3. Backbone    4. Last Mile



**Figure 3 - Location of Internet Bottlenecks**

## *Bogged Down at the Beginning*

The first type of bottleneck is a function of the centralized model used to serve content on the Internet today. Each content provider sets up a Web site in a single physical location and disseminates data, services, and information to all Internet users around the world from this central location (see Figure 4). This means that the speed with which users can access the site is necessarily limited by its **First Mile** connectivity – the bandwidth capacity of the Web site's connection to the Internet. In order to accommodate a growing number of users in this model, not only must each content provider keep buying larger and larger connections to his or her ISP, but so must the ISP continually expand its internal network capacity, and the same goes for neighboring networks. Since it is impossible in this type of approach to keep up with exponential growth in Internet traffic, the centralized model of content serving is inherently unscalable.
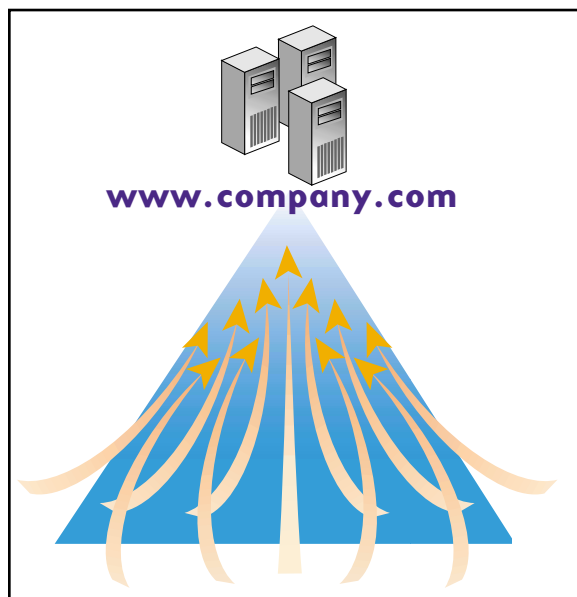
There are numerous examples of first mile bottlenecks impeding Web site performance and availability. These include Web sites covering the Clinton Impeachment hearings,[4] serving the Star Wars' movie trailers,[5] and reporting on the Mars Lander.[6] In all of these situations, the demand for the content exceeded the first mile capacity of the Web site.



**Figure 4 - The first mile bottleneck**

## *Peering: Points of Congestion*

The second type of Internet bottleneck occurs at peering points – the interconnection points between independent networks. The reasons why peering points are bottlenecks are mostly economic.

First of all, networks have little incentive to set up free peering arrangements, since there is no revenue generation opportunity in that type of arrangement, but there are considerable setup costs. At the same time, none of the large networks is going to agree to pay another large network for peering, because from a traffic perspective, they would both benefit equally from such an arrangement. As a result, large networks end up not peering with each other very much and so the limited number of peering points between them end up as bottlenecks.

One of the most common types of peering occurs when a small network buys connectivity from a large network. The issue that comes up in this situation is the long time it takes to provision the necessary circuits. Although there is plenty of dark, or unused, fiber already in the ground, the telephone companies who own it prefer to install new fiber for each requested circuit in order to increase their revenues. As a result, it takes 3-6 months to provision new circuits for peering. By the time a requested circuit is installed, traffic may have grown beyond expectations, resulting in a full link – a bottleneck.

It is also the case that it doesn't make economic sense for a network to pay for unused capacity. Therefore, a network generally purchases just enough capacity to handle current traffic levels, so every peering link is full. Unfortunately, this practice of running links at full capacity creates severe traffic bottlenecks. Links run at very high utilization exhibit both high packet loss rates as well as very high latency (because of router queuing delays) even for the packets that can get through. As a result, web performance slows to a crawl.

While it has been argued that peering bottlenecks will dissipate due to telecom consolidation, recent research into the structure of the Internet and its traffic composition indicates otherwise. In particular, it shows that the Internet is, and will continue to be, made up of thousands of networks, none of which individually carry a significant portion of user access traffic.

The figure below graphs the percentage of total **access traffic** that is handled by each of the Internet's 7400+ networks.[7] A network's access traffic, measured in bits/sec, is composed of the data packets originating at clients or servers directly connected to the network. In contrast, a network's **transit traffic** is composed of the data packets that originate on other networks. The access traffic for a network represents the amount of traffic its subscriber base of end users and Web sites is sending out onto the Internet, i.e., it is a measure of the amount of traffic the network is contributing the total Internet traffic load. This graph shows that user access traffic is pretty evenly spread out over 7400 networks, with no one network accounting for more than 5%, and most contributing much less than 1%.
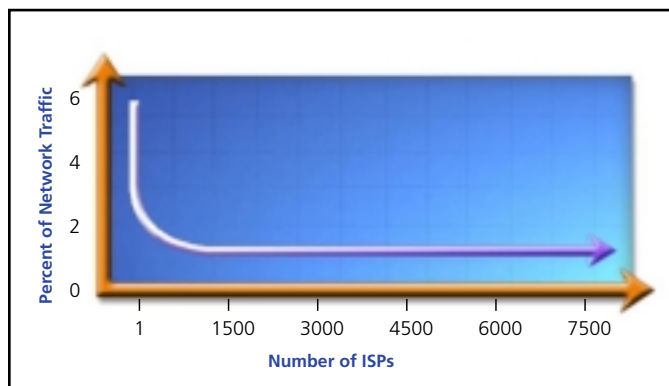


**Figure 5 - Internet user traffic across all networks**

The data in Figure 5 illustrates two key points:
*1. Since there are thousands of networks in the Internet, there are at least thousands of peering points in the Internet.*
*2. Since access traffic is evenly spread out over the Internet's thousands of networks, most traffic must travel through a number of different networks, and, consequently, a number of peering points, to reach its destination. (In contrast, if one network were to account for most of the Internet's access traffic, and, consequently, most of the Internet's subscribers, then most destinations could be reached within that single network and no peering points would have to be traversed).*

Therefore, the peering bottleneck problem is clearly a large-scale problem inherent to the structure of the Internet. Preliminary research into historical data shows that the number of networks making up the Internet is growing, a further sign that peering bottlenecks are likely to persist, rather than disappear. According to Boardwatch Magazine, as of April 2000, there were over 7,400 ISPs in North America alone, up from over just 5,000 in March 1999. [8]

## Breaking the Backbone

The third type of Internet bottleneck is in the capacity of the large long-haul networks that make up the Internet backbone. Because today's centralized model of content serving requires that almost all Internet traffic traverse one or more backbone networks, the capacity of these networks must be able to grow as quickly as Internet traffic.

A network's capacity is determined by the capacity of its cables and routers. Since fiber is cheap, plentiful and able to support high-bandwidth demands, cable capacity is not an issue. Instead, it is the routers at the ends of the fiber cables that limit backbone capacity. At any point in time, the speed of the packet-forwarding hardware and software in routers is limited by current technology. Routers have not historically provided the type of quality of service (QOS) features of carrier switches, and router capacity improvements have not kept pace with the exploding increase in Internet traffic volumes. In fact, many ISPs run IP over switched ATM networks, because IP routers have not been able to keep pace with their traffic demands.[9] However an ATM network is more expensive to deploy and maintain.

And while backbone providers too are spending a great deal of money upgrading their routers to handle more traffic (e.g., UUNet has recently added 10 Gbps OC-192 links to its network), demand will still end up far exceeding capacity, as illustrated in the following example.

Let's compute demand for long-haul Internet capacity. Consider an example application of video-on-demand on the Internet, the personalized Cooking Channel. Instead of watching the broadcast version of the generic Cooking Channel program on TV, each user will be able to put together his own "menu" of recipes to learn, catering to his or her own specific tastes or entertainment plans. How much Internet capacity will the personalized Cooking Channel consume? Viewer monitoring performed by Nielsen Media Research shows that cooking shows rate about 1/10 of a Nielsen rating point, or 100,000 simultaneous viewers. At a conservatively low encoding rate of 300 Kbps[10] and 100,000 unique simultaneous streams, the personalized Cooking Channel will consume 30 Gbps.

Now consider WorldCom's UUNET backbone which carries approximately half of all Internet transit traffic. Its network is comprised of multiple hubs of varying capacity, ranging from T1 (1.544 Mbps) to OC48 (2,450 Mbps or 2.4 Gbps). UUNET's Canadian and U.S. networks are linked with more than 3,600 Mbps (or 3.6 Gbps) of aggregate bandwidth between multiple metropolitan areas. However, the equivalent of 32 cooking channels alone could overload the capacity of the world's largest backbone to transfer content from the U.S. to Canada!

This gross mismatch of demand for and capacity of backbone bandwidth makes the backbone an increasingly serious bottleneck.

## The Long Last Mile

Most Internet users understand the last mile bottleneck – it's the limited capacity of their 56 Kbps modem connection to their ISP. A common misconception, however, is that eliminating this bottleneck will solve all Internet performance problems by providing high-speed Internet access for everyone. In fact, by rate-limiting Internet access, 56 Kbps modems are saving the Internet from a meltdown. If all users were able to access the Internet via multi-megabit cable modem or DSL modems, the other three types of bottlenecks would make the Internet unbearably slow. In other words, the Internet is only as fast as its slowest link.

This can easily be deduced from the example of cable access provider @Home. Compared to the total number of Internet users, @Home has a very small subscriber base – only 1 million subscribers compared to 377 million – only 0.3%. However, these high-bandwidth subscribers make up 5% of Internet access traffic. So, a higher bandwidth connection corresponds to higher demand for Internet capacity. If all other 376 million Internet users were to get high-speed connections, the Internet would melt.

Since the Internet is already being run at close to capacity, it's three other core bottlenecks must be alleviated before millions of additional broadband users are unleashed.

"Once the last-mile problem is solved," says David Levy, senior research analyst for Chase H&Q, "the congestion higher up in the Internet backbone becomes more apparent."

## Alleviating the Bottlenecks: the Case for Delivering Content from the Network Edge

The current centralized model of Internet content serving requires that all user requests and data travel through several networks and, therefore, encounter all four types of core Internet bottlenecks, in order to reach their destinations. Due to this model, the entire Internet slows down during peak traffic times such as 7 a.m. – 7 p.m. EST. The consequence of continuing to use this model is that the Internet will not be able to handle the exponential growth in traffic/user demand.

Fortunately, this unacceptable end result can be avoided by replacing centralized content serving with **edge delivery**, a much more scalable model of distributing information and services to end users. In edge delivery, the content of each Web site is available from multiple servers located at the edge of the Internet. In other words, a user/browser would be able to find all requested content on a server within its home network.

The edge represents a symbolic gateway to the unpredictable network of networks that the Internet is – the point where local traffic is added to the flows of national traffic. More importantly, the edge is the closest point in the network to the end user, who is the ultimate judge of service.[11]

Edge delivery solves the first mile bottleneck by eliminating the central point from which all data must be retrieved. By making each Web site's content available at multiple servers, each Web site's capacity is no longer limited by the capacity of a single network link. In contrast, the currently used "solution" to alleviating the first mile bottleneck – outsourced Web site hosting – simply moves the first mile problem to the hosting provider's datacenter. While outsourced hosting is a good idea for many reasons, it does not solve the first mile problem.

Edge delivery solves the peering bottleneck problem by making it unnecessary for web requests and data to traverse multiple networks and thus encounter peering points. Of course, in order to accomplish this goal, the edge delivery servers must be deployed inside ALL networks that have access customers. Since there are thousands of networks and access traffic is spread evenly among them, edge delivery schemes with only 50-100 locations or deployed in only the top 10 networks cannot in any way solve the peering bottleneck problem.

When content is retrieved from the network edge, the demand for backbone capacity decreases, thus alleviating the bottleneck there. Again, this can only be achieved with edge servers deployed at every one of the thousands of Internet access providers.

While edge delivery does not solve the last mile bottleneck issue, it does help to alleviate the growing severity of the other three bottlenecks, thus enabling content delivery closer to end users. This becomes especially important as technological advancements continue to improve last mile bandwidth capacity and thus unleash more and more traffic onto the Internet.

By shortening the path traversed by requests and data, and by enabling complete avoidance of the first mile, peering, and backbone bottlenecks, edge delivery can clearly eliminate the "world wide wait." Not only does this improve the end user's experience, but, according to Jupiter Research, it promises a less expensive way of delivering rich content, in contrast to simple delivery from an origin server.[12]

In fact, even network service providers realize the benefits of edge delivery. In Infonetics Research's August 2000 report: "The Service Provider Opportunity in the US 2000: National ISPs and Tier 2 CLECs", of ISPs surveyed, over half use caching or content delivery technologies now, rising to 69% of respondents in 2001. Among those who use content delivery and caching, 88% will partner with a content distribution service in 2001.

## *Edge Delivery – No Small Challenge*

There are complex challenges to implementing edge delivery, however. Providing a valuable Internet content delivery service requires far more than deploying servers on the edge. A whole set of supporting services and capabilities must be in place to provide a robust infrastructure. At a minimum, the following must be addressed to ensure a comprehensive and valuable solution.

**Content must be deployed on edge servers -** Each content provider must somehow make his or her content available from these edge servers. Despite the technological complexities of edge delivery, it should be simple for Web site owners to implement and use.

**Requires massive deployment of edge servers -** In order to be effective for a significant portion of Internet users, the edge servers must be deployed in thousands of networks – a huge undertaking. This requirement follows from the fact that user access traffic is evenly spread out over the Internet's 7400+ networks.

**Geographic diversity of edge servers is critical -** Optimal performance and reliability depend on the geographic diversity of the edge servers. Servers must be deployed at diverse locations worldwide, close to end users.

**Content management across disparate networks -** Because the ideal edge delivery service is deployed throughout multiple networks, the service must be sophisticated enough to manage content as it traverses these multiple networks. This requires sophisticated algorithms, comprehensive mapping of the Internet, and complex tracking mechanisms to ensure that content was delivered to the end destination.

**Content must be kept fresh and synchronized -** Ensuring content freshness and accuracy requires sophisticated mechanisms that automatically check the host site for changes and retrieve fresh content for delivery to the edge of the network. Some providers use standard caching. Content stored on standard cache servers does not change as the source content changes; thus content freshness is not guaranteed.

**Fault tolerance is crucial -** Truly reliable service providers acknowledge that no single network component is immune to failure. As a result, they design delivery systems to assume failure and automatically adjust to prevent service interruptions – a quality called fault tolerance.

**Monitor Internet traffic conditions -** A quality delivery service must have immediate access to extensive real-time data about Internet traffic so as to map the fastest route from server to end user in any type of traffic, but especially "flash crowds" – sudden heavy demand, expected or not, for a Web site. The service must be able to quickly reroute content in real time in response to congestion and network outages.

**Request routing -** The delivery service must be able to accurately pinpoint the user's geographic location in order to locate the closest and most available edge server from which to deliver content.

**Managing load -** A robust service must keep its servers from getting overloaded and will not only balance load among servers in a region, but also ensure that objects reside in the minimal number of locations necessary to provide good download performance.

**Performance monitoring -** Service providers must be able to determine, instantly and at any time, the status and performance of their edge servers.

**Insightful reporting -** Sophisticated content delivery services make delivery decisions based on detailed data, such as the end user's location and Internet traffic conditions, and should be able to provide meaningful, real-time data to their customers. Additionally, the service should provide each content provider with information on the type and amount of traffic being served by the thousands of edge delivery servers.

**Billing -** A robust edge delivery service must be able to support billions of hits per day, and have the infrastructure in place for real-time logging and billing of those hits, even across multiple networks.

**Comprehensive Serice Level Agreement -** Those services deployed across multiple networks need to provide a single SLA that guarantees their commitment to provide a reliable, single-source solution.  An edge delivery service should offer a SLA that says its performance will be better than what the customer can do on its own - that is the intrinsic value of such a service.

## *Edge Delivery with Akamai*

In response to the growing need for a more scalable model of Internet content delivery, Akamai has built an edge delivery service that accelerates access to Web content by up to 800%. Running on a network of 6000+ servers in 335+ networks and 54 countries[13], Akamai's flagship service, FreeFlow, is quickly approaching the goal of being inside every ISP. Content is spread out to Akamai edge servers using a demand-based pull mechanism, and sophisticated network monitoring and mapping algorithms direct each user request to the optimal edge server. Redundancy and fault-tolerance built into the network architecture ensure 100% availability. The success of edge delivery and Akamai's implementation of it is evidenced by the large amount of traffic served by FreeFlow on a daily basis: over 4 billion hits per day for over 2,800[14] of the most popular sites on the Internet.

The technical benefits of FreeFlow allow Web businesses to focus on core business issues such as marketing, product fulfillment, and content development, instead of on the complex and often intractable technical obstacles to delivering content across the Internet in an efficient and effective way. Through the combination of new technology and an aggressive network deployment strategy, Akamai Technologies enables Web businesses to serve content from a global network of servers, overcoming the limitations of today's restrictive model for delivering Web content to end users. No longer will "flash-crowds" – sudden traffic bursts that can render a Web site or an entire network effectively unreachable – make Web sites victims of their own success.

In addition to FreeFlow, Akamai has continued to add to its portfolio with other innovative edge delivery services, including Akamai Conference, Akamai Forum, Digital Parcel Service, EdgeScape, EdgeSuite, FirstPoint, and FreeFlow Streaming. (See http://www.akamai.com/html/sv/index.html for more information). These services extend Akamai's value proposition from one of site insurance to one of site differentiation. With Akamai's services, today's most visible sites are not only reliable, they are differentiated with new features that can only be delivered from the edge of the Internet.

With the world's largest distributed edge network and an entire suite of Internet content delivery solutions, Akamai is delivering a better Internet.

**"Content delivery partners such as Akamai…put content closer to users, eliminating the need for much of the back-and-forth traffic that bogs down performance. Those companies have inched the bar closer to ideal conditions, and will quickly become standard for site managers. To remain competitive, site managers must use these  services."**
**– Jupiter Research, Performance on the Edge, October 2000**

1    http://www.nua.ie/surveys/how_many_online/index.html

2    Jupiter Research, Performance on the Edge, October 2000

3    For more information on Internet routing and BGP, see the articles posted on http://avi.freedman.net/.

4    http://news.cnet.com/news/0-1005-200-337093.html?tag=st.ne.1002.srchres.ni

5    http://news.cnet.com/news/0-1005-200-339853.html?tag=st.ne.1002.srchres.ni

6    http://www.zdnet.com/zdnn/stories/bursts/0,7407,2403703,00.html

7    This data was obtained from logs of the Akamai network. Since Akamai serves over 4 billion hits/day for more than 2000 Web sites, its traffic

     provides an extremely representative sample of Internet traffic.

8    BoardWatch Magazine's Directory of Internet Service Providers, 12th Edition

9    Merrill Lynch, Internet Infrastructure 2000, July 2000

10   In 5 years, video bit rates will certainly be at least in the 1Mbps range.

11   Merrill Lynch, Internet Infrastructure 2000, July 2000

12   Jupiter Research, Performance on the Edge, October 2000

13  Source: Akamai, as of 11/1/2000

 14  Source: Akamai, as of 11/1/2000

# Glossary

**Access Traffic –** Measured in bits/sec and composed of the data packets originating at clients or servers directly connected to the network.

**ATM –** Asynchronous Transfer Mode; a low-delay, high-bandwidth, multiplexing and packet-switching technique. ATM uses a fixed-length packet, resulting in lower processing requirements and higher speeds.The term "asynchronous" applies, as each cell is presented to the network on a "start-stop" basis – in other words, asynchronously.

**Autonomous System –** A collection of routers under a single administrative authority using a common Interior Gateway Protocol for routing packets.

**Backbone –** The part of the communications network which carries the heaviest traffic.

**Bandwidth –** In telecommunications, bandwidth is the width of a communications channel.

**BGP –** Border Gateway Protocol is an Internet protocol defined by RFC 1163. It is a TCP/IP routing protocol which routers employ in order to exchange appropriate levels of information. When BGP peer routers first establish contact, they exchange full routing tables; subsequent contacts involve the transmission of changes, only. In an intradomain routing environment between Autonomous Systems, IBGP (Internal BGP) is run, allowing the free exchange of information between trusted systems. IBGP is in a class of protocols known as IGPs, or Internal Gateway Protocols. In an interdomain environment, EBGP (External BGP) is run, allowing the routers to exchange only pre-specified information with other pre-specified routers in other domains in order to ensure that their integrity is maintained. EBGP is in a class known as EGPs, or External Gateway Protocols.

**Byte –** A group of eight bits handled as a single logical unit.

**Cable Modem –** A small box that connects your PC to the Internet via your local cable TV provider. Cable modems allow PC users to download information from online services at speeds one hundred times faster than today's fastest telephone modems. However, the actual speed of transmission is dependent on many factors, including, but not limited to, how many other cable modem users are utilizing your local provider's network at that time; how fast the connection is from your cable provider to the Internet; how fast the connections are along the way; how fast the distant server is that you're attached to, etc.

**CLEC –** Competitive Local Exchange Carrier or Certified Local Exchange Carrier. A term coined for the deregulated, competitive telecommunications environment envisioned by the Telecommunications Act of 1996. The CLECs compete on a selective basis for local exchange service, as well as long distance, international, Internet access, and entertainment (e.g. Cable TV and Video on Demand).

**DSL –** Digital Subscriber Line. Digital Subscriber Line is a service that offers a faster Internet connection than a standard dial-up connection. DSL technology uses existing 2-wire copper telephone wiring to deliver high-speed data services to businesses and homes. In its various forms—including ADSL, HDSL, IDSL, R-ADSL, SDSL, and VDSL—DSL offers users a choice of speeds ranging from 32 Kbps to, in laboratory settings, more than 50 Mbps. Over any given link, the maximum DSL speed is determined by the distance between the customer site and the Central Office. At the customer premises, a DSL router or modem connects the DSL line to a local-area network (LAN) or an individual computer.

**Edge Delivery –** An Internet content delivery method by which the content of a Web site is available from multiple servers located at the edge of the Internet.

**Exterior (or External) Gateway Protocol –** The protocol used by a gateway in one autonomous system to advertise the IP addresses of networks in that system to a gateway in another autonomous system.

**First Mile -** bandwidth capacity of a Web site's connection to the Internet.

**Gbps –** Gigabits per second. Gig is one thousand million bits per second.

**Hop –** Each short, individual trip that packets make many times over, from router to router, on their way to their destinations.

**Internet Protocol –** IP. Part of the TCP/IP family of protocols describing software that tracks the internet address of nodes, routes outgoing messages, and recognizes incoming messages. Used in gateways to connect networks at OSI network Level 3 and above.

**ISP –** Internet Service Provider. A vendor who provides access for customers (companies and private individuals) to the Internet and the World Wide Web. The user typically reaches his ISP by either dialing-up with their own computer, modem and phone line, or over a dedicated line installed by the Internet Service Provider, a CLEC, or a local or long distance telephone company.

**Interconnection Points –** See MAE

**Interior (or Internal) Gateway Protocol -** The term applied to any protocol used to propagate network reachability and routing information within an autonomous system. There is no single standard IGP, but RIP is one of the most common.

**Last Mile –** "Last mile" is an imprecise term that typically means the link – usually twisted pair – between an end user and the telephone company central office – local, long distance, or Internet. It doesn't mean a mile, since that "mile" could be less than a mile or several miles. The term has entered the language referring to the problems of your communications making it that last mile. Often that "last mile" runs over old, limited bandwidth copper wire that has been in the ground for years and the quality of whose cable is not optimal.

**Latency –** A fancy term for waiting time or time delay. The time it takes to get information through a network.

**MAE –** A network access point (NAP), or public peering point, where Internet Service Providers (ISPs) interconnect to exchange traffic at the national backbone level. The original MAEs were MERIT Access Exchanges, established by MERIT Corporation, which joined with IBM and MCI to establish the original public Internet backbone that replaced the NSFNET (National Science Foundation NETwork). MAE East, the original MAE, is located in Vienna, VA. MAE West was later established in San Jose, CA. The term MAE was then used to denote either Metropolitan Area Exchange or Metropolitan Ethernet Exchange, both of which are interchangeable.

**Mbps –** Millions bits per second.

**OC-192 –** Optical Carrier Level 192. SONET channel of 9.953 thousand million bits per second (Gbps).

**OC-48 –** Optical Carrier Level 48. SONET channel of 2.488 thousand million bits per second (Gbps).

**OSI –** Open Systems Interconnection. The only internationally accepted framework of standards for communication between different systems made by different vendors. OSI was developed by the International Standards Organization (ISO), whose goal is to create an open systems networking environment where any vendor's computer system, connected to any network, can freely share data with any other computer system on that network or a linked network.

**OSPF –** Open Shortest Path First. A link-state routing algorithm that is used to calculate routes based on the number of routers, transmissions speed, delays and route costs.

**Packet –** Generic term for a bundle of data, usually in binary form, organized in a specific way for transmission. The specific native protocol of the data network (e.g. X.25, Frame Relay, ATM) may term the packet as a packet, block, frame or cell. A packet consists of the data to be transmitted and certain control information.

**Peering –** A relationship established between two or more ISPs (Internet Service Providers) for the purpose of exchanging traffic directly, rather than doing so through a backbone Internet provider. The traditional Internet architecture calls for ISPs and regional carriers to exchange traffic at Network Access Points (NAPs), carrier-class switches and routers, of which there currently are a dozen or so around the world.

**RIP –** Routing Information Protocol. RIP is based on distance-vector algorithms that measure the shortest path between two points on a network, based on the addresses of the originating and destination devices. The shortest path is determined by the number of "hops" between those points. Each router maintains a routing table, or routing database, of known addresses and routes; each router periodically broadcasts the contents of its table to neighboring routers in order that the entire network can maintain a synchronized database.

**Router –** Routers are the central switching offices of the Internet and corporate Intranets and Wide Area Networks (WANs). A router is, in the strictest terms, an interface between two networks.

**T1 –** A speed designation of high speed communication between systems over telephone lines at speeds of up to 1.544M.

**T3 –** A speed designation of higher-speed communication between systems over telephone lines at speeds of up to 44.736M; it contains 28 T1 channels.

**TCP/IP –** Transmission Control Protocol/Internet Protocol; specifications for transport and network layer communication between computers on a network.

**Transit Traffic –** Network traffic composed of the data packets that originate on other networks.

Akamai

www.akamai.com