INTERNET RESEARCH GROUP

THE NETWORK PROVIDERS BUSINESS CASE

FOR

INTERNET CONTENT DELIVERY

Peter Christy
Vice President
Internet Research Group
399 Main St., Los Altos, CA 94022
(650) 559-2103
pchristy@irgintl.com
www.irgintl.com

© Internet Research Group, 1999

About the Internet Research Group

The Internet Research Group is the leading consulting, market research, and information services firm focused exclusively on Internet infrastructure technologies for e-business and e-commerce. IRG's mission is to help major Internet technology vendors and key startups make faster and better decisions about their product, market entry, and market development strategies. IRG provides technical product and market analyses, develops authoritative reports on specific markets (such as Internet traffic management and Web caching products), produces focused informational Web sites around specific technologies, and creates white papers and other information products. IRG's consulting division helps companies develop successful market entry and development strategies for new products in early adopter markets. For more information, visit IRG's Web site at www.irgintl.com.

Caching and Internet content distribution are two key focus areas for which IRG offers Web resource centers. For more information about caching, please visit www.caching.com. For more information about Internet content distribution, please visit www.cddcenter.com.

For more information about this paper, please contact:

Peter Christy
Vice President
Internet Research Group
(650) 559-2103
pchristy@irgintl.com

© Internet Research Group, 1999

Contents

© Internet Research Group, 1999

Executive Summary

The rise of the Internet as a primary communications channel in our society has revealed significant performance problems that affect Network Provider revenues and customer satisfaction. Content delivery systems are a solution to these performance problems. Using Akamai Technologies' network as an example, this paper explains how content delivery systems reduce Network Provider and subscriber satisfaction problems. This paper will address the business case for all Network Providers, including ISP operations centers, universities, corporate campuses, broadband providers, and all other locations with large numbers of Web visitors.

Internet Performance Issues

A Web user's experience is directly related to the speed at which his browser retrieves and displays pages from Web servers. This performance can be affected by Web server processing delays (due to excess demand on the server providing the content); Internet delays (overloads at network peering points or routers between the user and the original content server); or "last mile" delays (low-speed, dial-up connections between the user and the Network Provider).

Content Delivery Systems

Content delivery systems resolve performance problems related to Web server processing delays and Internet delays. Fundamentally, content delivery systems create and maintain up-to-date copies of popular or high-bandwidth Web content in cache servers at multiple locations at the edges of the Internet. As a result, users requesting popular Web content may well have those requests served from a location much closer to them (a local Network Provider's data center), rather than from much farther away at the original Web server. By serving content requests from much closer to the user, content delivery systems reduce the potential for Web server overloads and Internet delays.

The Akamai Network System

The Akamai Network System is a content delivery service that consists of thousands of servers located at Network Provider data centers around the world. Akamai contracts with high-volume Web content providers such as Yahoo to distribute their frequently accessed content. Akamai uses the Internet to distribute copies of this content to its FreeFlow servers, and then, using proprietary logic from its network operations center, Akamai uses the Domain Name Services (DNS) system to reroute standard browser contents for its customers' content to the optimal FreeFlow server at the location nearest the user. These servers (nominally configured with 1 GB of DRAM and two 18 GB disk drives) are provided at no cost to Network Providers willing to host them. They integrate seamlessly with a provider's network, storing popular, high-bandwidth content and serving it to the Network Provider's subscriber base. The servers are typically configured in groups of five with two Ethernet switches, although smaller and larger configurations are also available.

Benefits of the Akamai Network

The primary benefits of the Akamai Network for Network Providers are lower infrastructure costs, higher subscriber satisfaction and quality of service, and reduced subscriber churn. By serving popular content from local servers at the Network Provider's data center, the FreeFlow system saves Network Providers the cost of retrieving this content from origin servers that may be thousands of miles away across expensive communications links. And, by serving subscribers from a local server, Network Providers are able to offer a significantly higher level of service, presenting popular Web content quickly and reliably, and thereby also reducing churn rate and its associated costs.

© Internet Research Group, 1999

Internet Research Group

Overview

The Web and the Internet have revealed the power of universal information access, and the world hasn't been the same since. What started 30 years ago as an academic and research tool has become the foundation for modern industrial and commercial activity. Successful Web sites literally experience a doubling of load every 90-100 days. Internet access has driven PCs into more than 50 percent of U.S. homes faster than anyone expected. Billions of dollars are being invested in additional data transmission capacity, which for now seems at best likely to just keep up with demand. And as the Internet evolves, Web access performance – the absence of delays or transmission problems – becomes more important, often relating directly to a Network Provider's brand value and subscriber satisfaction.

Using Akamai's FreeFlow Internet content delivery system as an example, this paper explains emerging Internet content delivery services and how they can make a big difference in achievable Web performance, thereby reducing Network Provider bandwidth costs and boosting subscriber satisfaction for providers that participate in content delivery systems.

Internet Performance Issues

The Internet miracle is based on a growing set of standardized, interconnected networks, and a standard for information publishing and viewing (the Web and browsers). We access information on the Web by giving our browser the name of a Web site, after which that content is fetched and displayed using the Internet to access and transport the data. If everything goes well, and all the links are high-speed, fetching Web data seems instantaneous. But as we all know, the Web is rarely instantaneous. Internet performance problems often arise in any of three general areas:

- **Web server processing delays:** Traffic loads that are ten or more times greater than the typical load on the site as news events or fads drive high access. The site's servers are unable to handle high loads and thrash, halting delivery of data.
- **Internet delays:** Delays somewhere in the Internet between the information requester and the Web server (often at the peering points between networks).
- **Last mile delays:** Delays in the connection between the subscriber and the Internet (e.g. due to a 14.4 Kbps dial-up modem connection).

© Internet Research Group, 1999

Performance problems are important to content providers and consumers, but they are more important to Network Providers, who rely on performance as a key differentiator of their service. Performance problems drive subscribers to sample competitive services. Typical Network Provider churn rates are over 4 percent per month, which compounds to an astounding 50 percent per year loss rate, and with new subscriber acquisition costs equal to 2-10 months of revenue, a lower churn rate can directly and significantly impact business profitability.

How Performance Affects Internet Players

In an ideal world, clicking on a page link would be instantaneous – the new page would be in place faster than the mind could absorb its content. At best, today's world is a far cry from that ideal. Human response time is less than a second, but the typical page response time for leading sites is more like 10 seconds, according to data measured by Keynote Systems. And 10-second page responses are hardly the worst case: a site can be down or inaccessible, or there can be instability somewhere in the Net between user and server (packet loss at a peering point or "route flap"). A flash crowd could also hit the site, driving the computer systems into severe, thrashing overload.

Performance problems affect everyone in the Internet value chain:

- Network providers have lower customer satisfaction, higher subscriber turnover, and higher connectivity costs as they attempt to improve content transmissions or subscriber access by increasing Internet capacity and peering with additional networks.
- Web users seeking information must wait for it to be delivered, which wastes time and decreases their satisfaction.
- Content providers seeking to deliver their information see their content's impact decreased through access delays or slow delivery. As a result, site designs are "dumbed down," and lack rich content.

Internet Content Delivery: Changing the Transmission Model

The traditional Internet model of browsers connecting to content servers over the Internet is the root of many of these problems. Using a distant browser to request files from an origin content server was the elegant idea that created the Web and accelerated the Internet revolution (*see Figure 1*). But as the Internet increases in commercial value, as content becomes richer, and

© Internet Research Group, 1999

where performance improvements translate directly into business value, we start to see the limitations in this simple model:

- Servers can't keep up with the peak loads presented, unless the site is built with gross overcapacity.
- Beyond the borders of the U.S., network capacity diminishes rapidly. What works well here is intolerably slow for overseas users.
- The data exchange points (peering points) between the various networks that constitute the "Internet" become overloaded, so data is lost and must be retransmitted.
- The bulk of a Network Provider's traffic is "backhaul" data to retrieve rich content from distant servers.

As content becomes richer and more complex, the chances increase that these problems will impinge on performance, bandwidth costs, and the satisfaction of valuable subscribers.

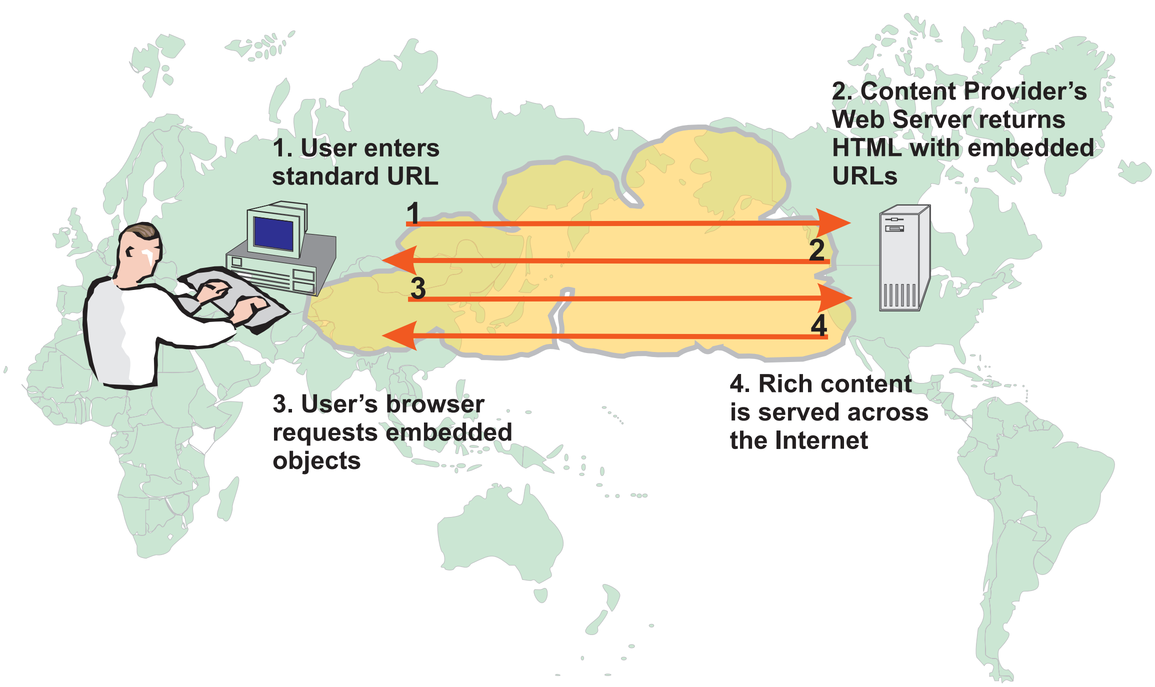

Figure 1. Traditional Internet content delivery.

How an Internet Content Delivery service improves performance

A content delivery service alleviates these problems by moving demanding content (popular and/or bandwidth intensive) closer to the people requesting it. Rather than serving content from the origin Web site, the content distribution model makes copies of key content on a multiple content delivery servers sites distributed through the Internet, close to the users requesting that content (*see Figure 2*). This approach addresses all of the problems previously described:

- With the hottest content off-loaded to multiple distributed content servers, the load on the origin server diminishes.
- The connection from a local content delivery server to the user is shorter than the connection to the origin server, is less subject to delays, and has higher bandwidth, producing increased subscriber satisfaction.
- The path from the user to the content server transits fewer peering points (often none if the content server is within the same network as the user), greatly reducing the sources of packet loss and data retransmission.
- As a system shared by many content providers, a content delivery service dramatically increases the "hit ratio" associated with caching, thereby reducing Network Provider bandwidth costs.

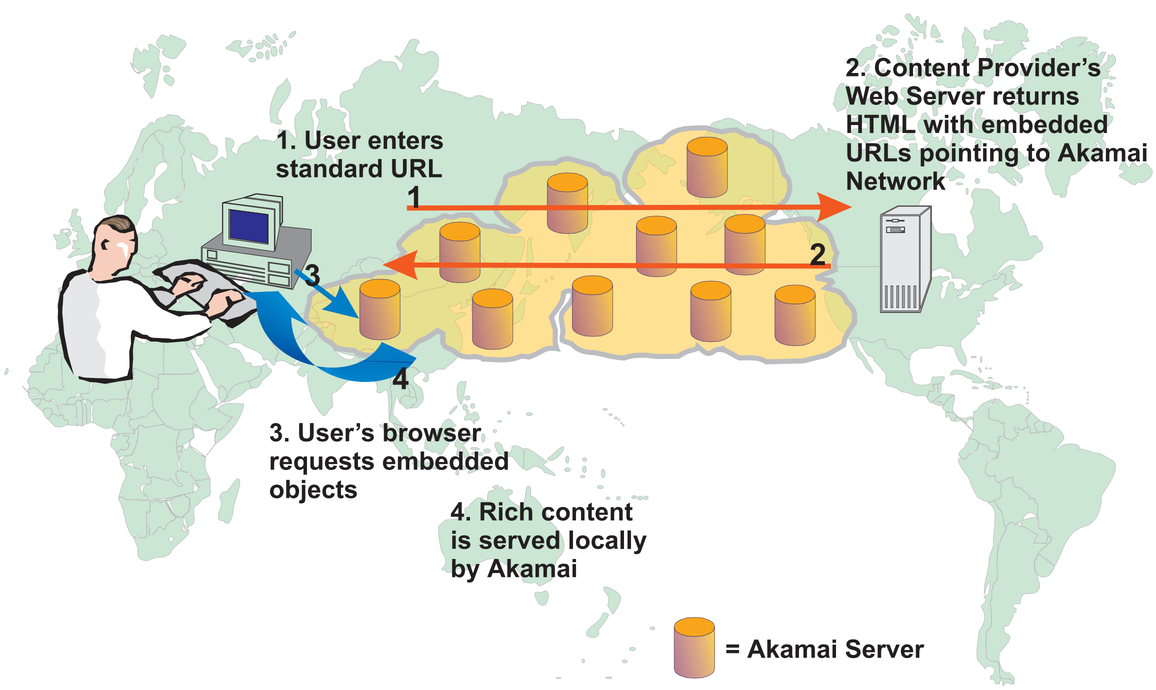

Figure 2. Content delivery with the Akamai Network system.

As the success of systems like Akamai's clearly demonstrates, a shared content delivery service is a much more effective solution to Web access performance issues than is even possible by building bigger servers and adding additional interconnection capacity. By solving a common problem, content delivery services offer common benefits to everyone in the Internet value chain:

- Subscriber Network Providers save on infrastructure costs and improve their customer satisfaction.
- Content consumers get better quality as their access to popular Web site improves in performance and availability.
- Network providers improve the performance and reliability of their service.

ICD systems can directly impact performance because they are practical remedies for these problems, and because they improve performance in the absence of any severe problem. In other words, adding content delivery directly and significantly improves customer satisfaction, decreases subscriber churn, and contributes to the Network Provider's bottom line.

The Akamai FreeFlow System

The Akamai FreeFlow system has three functional components: Akamai servers, a request routing system, and Domain Name Services.

Servers

Akamai's FreeFlow system runs on thousands of servers distributed throughout the Internet at Network Provider operations centers, universities, corporate campuses, and other locations with a large number of Web visitors. Each server in this generation of the service has a minimum of 1GB of RAM and two 18 GB disk drives and performs the functions of a cache, storing a copy of the content that Akamai delivers.

Akamai typically deploys servers in configurations of five with two Ethernet switches (*see Figure 3*), although both smaller and larger configurations of three, nine or eighteen servers are also available. The Ethernet switch is used to provide inter-server communications as well as a 100 Mbps connection to the Internet. The rack-mounted servers are extremely easy to install and are typically functioning within hours with absolutely no changes to a provider's network topology or configuration.

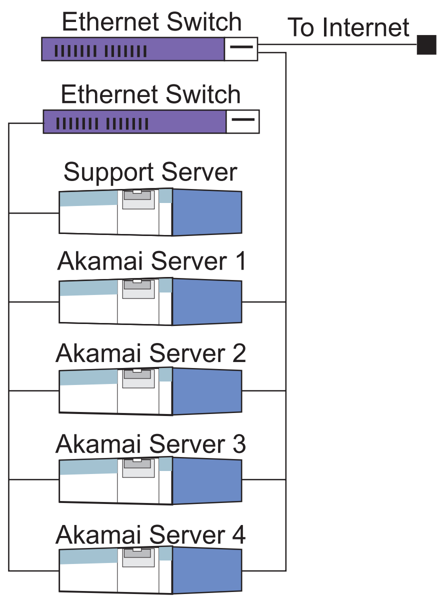

Figure 3. A typical deployment diagram for Akamai servers.

> "It's free and we save on traffic to the U.S. As more content providers sign up with Akamai it will get even better."
> --Kenji Hirota
> KDD Japan
> Akamai Accelerated Network

Through its Akamai Accelerated Network Program, Akamai offers these servers to qualified Network Providers at no cost – the provider's actual investment is rack space, power, and local network connection. In exchange for these modest costs, Network Providers gain the higher bandwidth and improved quality of service for their customers while diminishing their need for upstream Internet capacity. What's more, the Akamai servers are configured to serve only a Network Provider's downstream customers, preventing the possible consumption of upstream bandwidth.

Cache compatibility

In the near future, Network Providers that have invested in third-party caches will be able to join the Akamai network through a simple software upgrade to their cache software. Akamai and Cisco have authored the Open Cache Interface, which will enable third-party caches to communicate with the Akamai network and store content for major Akamai customers such as Apple, Yahoo, and CNN Interactive. To date, CacheFlow, Cisco, InfoLibria, Network Appliance, and Novell have joined this initiative.

INTERNET RESEARCH GROUP

© Internet Research Group, 1999

Request routing

Akamai optimizes access to the content it serves by continuously optimizing how requests for that content are distributed to the thousands of servers that serve content. The first step in that optimization is data collection.

Data Collection

To maximize the performance from its set of servers, Akamai first uses methods of measuring Internet performance, Akamai content request traffic, and server loads.

Data Processing

With all this data collected, FreeFlow uses sophisticated algorithms to compute the optimal mapping of requests to servers, given request and server loads, the available access to the various server centers, and the general condition of various Internet links. Those routings are put in place through the use of the Internet Domain Name Services.

Domain Name Services

Akamai uses the Domain Name Services (DNS) functions within the Internet to steer requests for Akamai-served content to specific Akamai servers. All Akamai content is part of the Akamai name domain. As a DNS server contacts Akamai to resolve an Akamai URL to the IP number equivalent, Akamai considers the location (IP number) of the requesting DNS server and feeds it the optimized routing for that location. Then, the DNS system does the actual work of steering the request to the right Akamai server. As load and Internet conditions change, Akamai re-optimizes the solution in an ongoing fashion in near real time.

Distributed servers magnify bandwidth

It is important to note that Akamai's Network is *not* a facilities-based network. Akamai content is served from individual servers located as close to the consumer as possible, rather than over a dedicated or shared network. Each Akamai server is capable of serving over 100Mbits/second, so the set of 1,700 servers can serve 170 Gbits/second – a huge capacity compared to the requirements of even the largest Web sites – but there is no high-capacity data transmission network needed between those servers or back to the NOC. The only network capacity used is that right at the edges, where bandwidth is the least expensive.

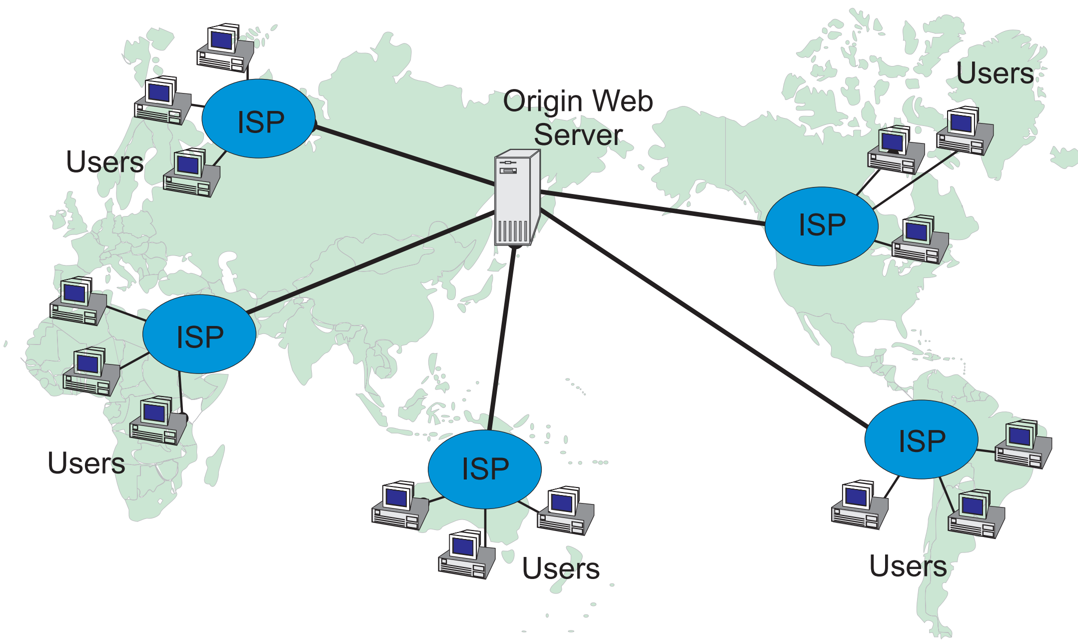

Figure 4. Traditional Internet content access – browsers get content from origin server.

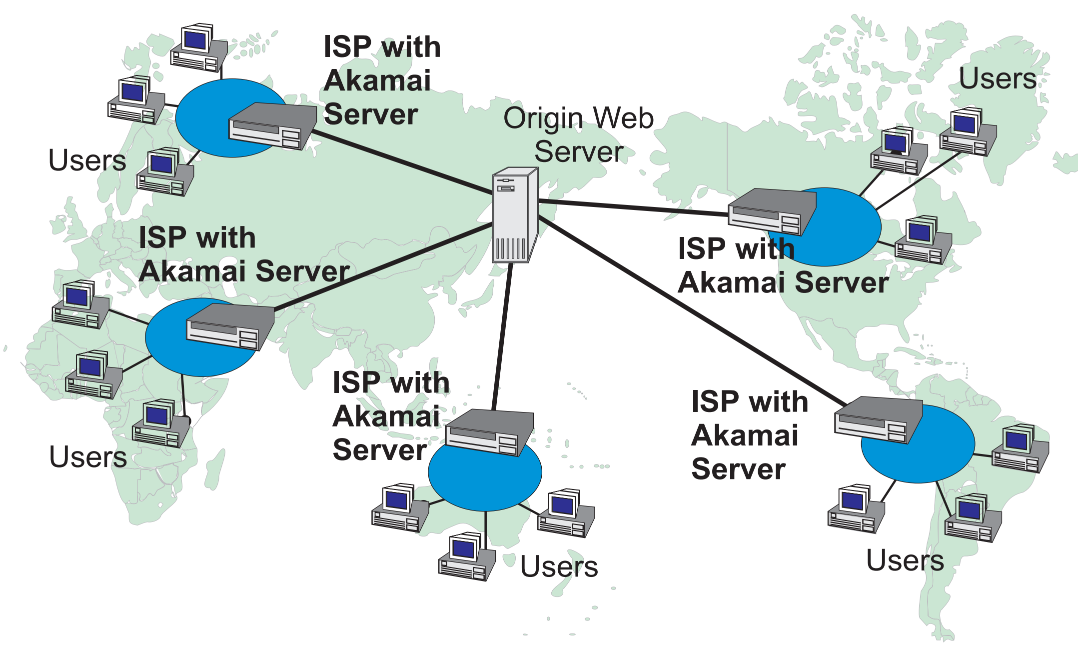

Figure 5. Internet content distribution with Akamai servers – Akamai servers at Network Providers cache hot content and serve it locally.

Akamai contracts with content providers to deliver their content through the Akamai FreeFlow service as if it were a physical network. Content is "Akamaized" by changing the URL references to the Akamaized version. Requests are subsequently served from Akamai servers situated on the edge of the network, near the user. This simple change improves performance under normal conditions and avoids the problems outlined earlier. Akamai logs the rate at which content is served for each customer, aggregating the instantaneous contribution of thousands of servers, and charges the customer as if that content had been served from a centralized server through a single big pipe.

How Subscriber Network Providers Benefit from Content Delivery

The principal benefits of content delivery services from the Network Provider's point of view are savings on infrastructure costs, increased customer satisfaction, and reduced subscriber churn (and the consequent reduction in customer acquisition costs).

Infrastructure savings through reduced bandwidth demand

By adding content delivery servers to its infrastructure, a service provider saves significantly on capital and operating expenses because these servers serve its users exclusively: no requests from outside the provider's network are mapped to the provider's region of servers. The content server delivers content locally that otherwise would have to be fetched from the Internet. The bandwidth needed to connect the server locally is essentially free (abundant) whereas the same bandwidth over the Internet is costly. For most sites, the average bandwidth savings with an offering like Akamai's will be 10-15 percent.

> "Akamai has very clever technology that is saving us money and improving our service. It has actually improved our profile in the marketplace."
>
> --John Lindsay
> Operations Mgr., iiNet Ltd.
> Akamai Accelerated Network

The actual saving is probably considerably larger than the average, because most Network Providers provision for peak load, not for average load. Because the content provided by a content delivery service tends to be exactly the content with the highest probability of flash crowds and surge loads, the content delivery service has a greater impact under these peak conditions. Whereas the impact on average load might be 10 percent, the impact on peak loads could be much greater.

Improved customer satisfaction through better QoS

One of the biggest impacts of Internet content delivery on a Network Provider's business is in the area of customer satisfaction. Viewed as a cost savings, a content delivery system saves some percentage of costs; but from the user's point of view, content delivery can make a black and white difference in performance.

A subscriber Network Provider runs a difficult business. Good access is transparent – fast and reliable. Access performance becomes an issue when it is less than transparent. Even though many performance problems are not related to a provider's network, new Web users attribute *all* of their problems to their Network Provider. There can be many causes for subscriber access problems:

> "It makes the Akamaized parts of the Web run like a scalded cat. In Australia, where people are accustomed to Web pages from the USA arriving in a sluggish or inconsistent manner, the consistent, blazing speed at which Akamai served content arrives is jaw dropping stuff."
>
> --Simon Hackett
> Technical Director
> Internode, Australia
> Akamai Accelerated Network

- Problems in the subscriber access line (dial-up connection, xDSL line, cable)
- Inadequate provisioning of upstream bandwidth – not enough capacity for the number of on-line subscribers either under normal loads or flash crowd conditions
- Problems with the origin server (inadequate capacity, for example)
- Problems in the Internet between the Network Provider and the origin server (packet loss and data retransmission from an overloaded peering point, for example).

It's important to realize that using a content delivery system can make a big difference in all aspects of performance except those due specifically to problems within the Network Provider's network. A content delivery system moves highly popular, highly accessed content from the origin servers to a content delivery system within the provider's network, so it alleviates the other problems as follows:

- Problems in the access line – Because content delivery servers are placed near modem racks, DSLAMs, and cable head-ends, local loops are fully utilized.
- Inadequate provisioning of upstream bandwidth – The local content delivery server reduces upstream bandwidth demand in all conditions and especially in flash crowd circumstances.
- Problems with the origin server – The local content delivery server reduces traffic on the origin server by as much as 90 percent (for popular, served content) and largely eliminates dependence on the server. Again, these improvements are greatest precisely under the most demanding flash crowd circumstances.

© Internet Research Group, 1999

- Problems in the Internet between the Network Provider and the origin server – The content retrieved from the local server isn't subject to Internet delays.

The subscriber satisfaction impact of a content delivery system is amplified because the content delivery system is likely to accelerate access to the most popular, most frequently accessed material. Delays to relatively obscure content can easily be attributed to problems with the server responsible for its delivery, but access to popular sites such as Yahoo! are a natural way to compare the performance of different Network Providers.

> Akamai has a really good shot at solving the global Web content distribution problem, and I think they have it nailed. Because of our partnership with Akamai, I have Web sites I can highlight to my customers, like CNN and Apple, whose enormous download speed becomes a great demonstration of the value offered to them by choosing Internode. It is a strong competitive differentiator in an increasingly challenging market.
>
> --Simon Hackett
> Technical Director
> Internode, Australia
> Akamai Accelerated Network

Lower subscriber churn rates

High churn rates are a general Network Provider problem. Tables 3 and 4 below give some concrete examples of the impact of churn rate on profitability and shows how seemingly small changes in churn rate end up making a big difference in bottom line profitability. Happier customers mean lower churn rates, lower customer acquisition costs (because less is spent to replace departing subscribers), and lower associated administrative costs involved in deleting and adding subscribers.

The Economics of "Churn"

The analysis on the following page shows how reducing the subscriber churn rate (the rate at which subscribers discontinue service) can have a profound the impact on a Network Provider's gross revenue.

Basic Assumptions			
Churn rate	4.4% (Average)	4.0%	3.5%
Average subscriber acquisition cost (in revenue months)	5	5	5
Non-critical Assumptions			
Average subscriber monthly bill	$20	$20	$20
Number of subscribers	100,000	100,000	100,000
Calculations			
Annual revenues	$24,000,000	$24,000,000	$24,000,000
Annualized loss rate	42%	39%	35%
Monthly subscriber loss	4,400	4,000	3,500
Acquisition cost / subscriber	$100	$100	$100
Cost to refresh lost subscribers / mo.	$440,000	$400,000	$350,000
Impact of Reduced Churn Rate			
Savings / month		$40,000	$90,000
Savings / year		$480,000	$1,080,000
Bottom Line Impact			
Savings as a percent of revenue		2%	5%

Table 3. Bottom line impact of specific churn rate reductions
Source: Internet Research Group

In Table 3, we assume that a Network Provider's goal is to maintain or grow the customer base. We assumed an average churn rate of 4.4 percent (the figure given as an average Network Provider churn rate a recent article). We also assume that the average monthly revenue per subscriber is $20. We then calculate the cost to replace lost subscribers at the average rate, and at two reduced churn rates of 4.0 percent and 3.5 percent. The explanation details the calculations in Column B (the 4.4 percent churn rate).

In the **Basic Assumptions** section, we assume the average churn rate, and a subscriber acquisition cost of 5 months (at the $20 monthly average subscriber revenue shown directly below).

Under **Non-critical Assumption**s, we assume 100,000 subscribers and compute annual revenues of $24 million ($20/mo. x 100,000 subscribers).

In the **Calculations** section, we determine the annualized version of the churn rate just to emphasize the severity of the problem: If you lose 4.4 percent of your subscribers a month, that's a 42 percent loss rate on an annualized basis! We multiply the monthly loss rate times the annual

© Internet Research Group, 1999

revenue to get the monthly subscriber loss, which is also the number of new subscribers we need to acquire just to stay even. We calculate the cost of acquiring new subscribers as equal to five months of revenue, or $100. (By saying it is five months of revenue we're just emphasizing the obvious fact that it takes quite a while before the Network Provider gains the benefits of that new subscriber.) This $100 figure includes all costs: advertising and other forms of demand creation, telemarketing, the accounting costs of setting up a new account, and the average cost of technical support for a new user.

When we compute the acquisition cost of each new subscriber (the subscriber's monthly bill times the acquisition cost in months of revenue, or $100 in our case), we arrive at the frightening fact that we have to spend $440,000 month to acquire enough new subscribers just to stay even!

In other columns, we examine scenarios with 4.0 percent and 3.5 percent churn rates, and likewise calculate the costs of keeping the subscriber base constant at these rates. As shown under **Impact of Reduced Churn Rate,** the savings are $480,000 and $1,080,000 per year respectively! We then normalize these amounts under **Bottom Line Impact** by comparing them to our annual revenues and get savings, as a percentage of revenue, of 2 percent and 5 percent respectively (big numbers for a business with thin margins). It's worth noting that the savings as a percentage of revenue is independent of the number of subscribers and the average monthly bill, as long as the subscriber acquisition cost in months remains constant..

Using Content Delivery Systems with Caching

The Akamai content servers contain caches. If you already have a cache, does it make sense to add an Akamai server? If you have an Akamai server, do you still need a cache? In general, these two functions are complementary.

Any cache functions by responding to requests for Web objects. If the object requested is unknown to the cache, it is fetched from the origin server. If the object has been fetched recently, the request is served from the cache. In the case of a conventional Internet cache, the cache sees all requests from the community of users served by the cache. The benefit from such a cache derives from the commonality of use within the community. The cache provides a benefit if the requested object has already been recently requested. In such an application, a cache demonstrates a "hit ratio" (the percentage of requests served from the cache) of about 40 percent.

© Internet Research Group, 1999

An Akamai server operates like a conventional Internet cache but with very different results because of (a) the content served by the cache and (b) the resulting hit ratio. The big difference between the Akamai server and a conventional Internet cache is that the Akamai server *only* gets requests for those objects that Akamai is currently referring to it rather than the much larger and less predictable set of requests fielded by the Internet cache. As a result:

1) The hit rate on those specific Akamai requests is effectively 100 percent.
2) Essentially, all requests are served from the 1Gbyte DRAM buffer on the cache rather than disk, because the Akamai system can minimize the selection of content served by each Akamai server.

The impact of the improved hit ratio is easy to understand and quite profound. Assume for each cache that the service time from the cache (when the object is stored locally) is .1 second and the service time from the Internet (on a miss) is 8 seconds. The average service time from the cache can be computed as follows:

Average_time = hit_time * %_hit + miss_time* %_miss

For both the Akamai server and the traditional Internet cache, we'll assume that the hit_time is .1 second and the miss time is 8 seconds. The difference comes in the hit_ratio: will assume that the Akamai server hits 99 percent of the time where as the Internet cache hits 40 percent of the time.

For the conventional cache:

Average_time = .1*.4 + 8*.6 or about 5 seconds.

For the Akamai use of the cache:

Average_time = .1*.99 + 8*.01 or about 0.2 seconds.

Thus, the average performance of the same cache in the Akamai application is about ***twenty-five*** *times better* than in the Internet caching application. The cache works better in the Akamai application because the stream of requests sent to the cache is very selective, and they all hit. Because they all hit, and because the set of objects served is managed, they all fit in the DRAM buffers, and the performance of the server can be very good.

It's easy for a simple, modern server to deliver in excess of 100Mbits/sec. if the content is in RAM. When a cache serves content from the disk, the performance is limited to about 5 Mbits/second per disk drive (because of the time spent moving the disk heads from one object to

© Internet Research Group, 1999

another). To get 100Mbits/second from a disk-based cache would require 20 or more disk drives – a significantly bigger and more complex system.

In summary, an Akamai server doesn't replace a normal cache. The performance advantage of the Akamai server is due to how it is used, not to the design of the server *per se*. Akamai content servers and Internet caches serve two very distinct purposes: Akamai optimizes content delivery for specific, paying content publishers; while an Internet cache optimizes content access for a community of subscribers. Each solution uses similar server technology but with very different usage characteristics. Having an Akamai server in no way eliminates the value of an Internet caching system. Similarly, Internet caches get only piece of the benefit that the Akamai server does for Akamaized content.

Can a conventional cache be an Akamai server?

Since the fundamental caching function is the same in both Akamai and conventional caches and the demand for floor or rack space is always a consideration for Network Providers, the obvious question is, "Could a conventional cache be substituted for the Akamai server and used to serve both conventional and Akamaized content?" The answer is "Yes, if a few relatively straightforward adaptations were made to the Internet cache." The changes would involve:

1. **Registration with the Akamai network control center.** Obviously, Akamai would have to know that this cache existed so that suitable Akamaized content could be routed to it.

2. **Data measurement and logging.** The cache would need to record information including (a) the load on the cache, (b) the load on all of the Akamai content and (c) the source of Akamai load. This data would be factored into Akamai's distribution optimization and customer billing.

3. **Memory optimization**. You would want the cache to "pin" the Akamai content down into RAM so that it would be available for fast service without requiring any disk bandwidth, and so it wouldn't get displaced by other content.

These adaptations are part of the Open Cache Interface (OCI), which was co-developed by Cisco Systems and Akamai. The result is a single cache that performs both functions well, provided that the system capacity (CPU speed, RAM complement) is adequate for the combined requirements of the two, merged applications. So rather than being competitive technologies, it seems that with a little engineering and design, Internet caches incorporating OCI serve both their traditional function and are fully functional as Akamai system servers.

© Internet Research Group, 1999

Conclusion

The recent introduction of Internet content delivery services represents a fundamental step in improving the performance and reliability of Web access. These services all place content delivery servers at the edges of the Internet near the consumers of that information, and use these servers to cache and serve popular content. The Akamai FreeFlow system uses a sophisticated traffic management scheme to direct requests optimally to a server located close to the requester. When under the direction of such a traffic management and content optimization system, Akamai servers achieve very high performance compared to a typical Internet cache, and they make very significant improvements to the access performance and reliability of the content they serve.

Installing these caches within a provider's network improves the cost-efficiency of the provider's network operations, and more importantly it makes a dramatic difference in the speed and reliability of access to much of the Internet's most popular content. By doing so, Akamai servers significantly improve subscriber satisfaction and diminish subscriber churn, all with a direct bottom-line impact to the Network Provider.

Predicting the Future

The Internet Research Group believes that several market dynamics will enhance the validity of content delivery systems for Network Providers:

1. **Content delivery will become more demanding, not less.** With the increased use of broadband access methods (Cable TV modems and xDSL), there will be increased use of rich (data-intensive) Web objects (e.g. streaming media files). These rich objects impose an increased data transfer burden (more data) and they place a premium on data transmission performance (the better the link to the streaming server, the better the quality of the viewing, for example).
2. **Quality problems on the Internet won't disappear soon.** The strength and the weakness of the Internet is that it consists of a loose federation of interconnected networks. That lack of overall management was essential to the Internet's creation and to its continuing, unabated evolution and growth. But it also means that there isn't any obvious way to make it all work in a coordinated manner, especially considering that one network's priority traffic may be another network's junk mail.

© Internet Research Group, 1999

INTERNET RESEARCH GROUP

Unless there is massive consolidation of network owners and operators, the problems that lead to Internet delays are likely to persist for some time to come.

3. **Data storage will continue to be cheap compared to data transmission**: Content delivery systems not only improve performance, but they ultimately save money because storing data (rather than transmitting it) is the most cost-effective solution. As dramatic as improvements in data communications cost/performance are, they just keep up with the equally dramatic improvement in storage, at best.

The basic reasons for the value of content delivery systems are unlikely to go away anytime soon. Internet content delivery systems will present a very promising solution for Network Providers that want to reduce bandwidth costs while boosting user satisfaction.

© Internet Research Group, 1999

Appendix 1: Glossary of Terms

Akamai Accelerated Network – A service provider that has deployed Akamai servers in its network to improve quality of service for customers while greatly diminishing its need for upstream Internet bandwidth.

Akamaize – The transformation of a URL into an ARL at the origin server, thereby redirecting user browser requests to the Akamai network.

Akamai Resource Locator (ARL) – A Universal Resource Locator that Akamai uses to direct browser requests to content servers in order to optimize the performance of its FreeFlow content distribution system.

Cache – A network device that stores frequently requested Web content, intercepts HTTP requests for that content, and serves those requests itself rather than passing them onto to the origin server.

Churn rate – The rate at which subscriber customers leave a Network Provider, typically expressed as a percentage of the provider's total subscriber base.

Cache Interface Protocol (CIP) – A communications protocol jointly developed by Akamai and Cisco Systems, which allows CIP-compliant caches to seamlessly interact with the Akamai network.

DSLAM (Digital Subscriber Line Access Multiplexer) – A device that connects multiple Digital Subscriber Lines' aggregate traffic to the DSL host's network.

Flash Crowd – A sudden, large surge in traffic to a particular Web site, typically fostered by the availability and broad promotion of new and highly desirable content such as the Victoria's Secret video fashion show or the publication of the Starr Report on the Whitewater investigation.

FreeFlow – A global Internet content delivery service that enables Web site owners to deliver the bulk of their content more quickly and reliably than possible with centralized Web servers.

Content Delivery Server – A server that caches content at a location where the content can be served to local browser users more quickly and at a higher level of quality that it can from the origin server.

Origin Server – A Web server that contains the original copy of content being served to browsers.

© Internet Research Group, 1999

Peering Point – A data sharing point between two of the component networks that make up the Internet. Transmission delays often occur at these points if there is no overall management of them.

Quality of Service (QoS) – The level of Web service that a subscriber perceives, and which is a function of the speed and level of quality at which content is delivered to the browser.

Route Flap – Instability in networks that occurs when large, rapid changes in the availability of bandwidth on different network routes causes the routers involved to send successive and conflicting signals to one another as to which transmission path is best.

Uniform Resource Locator (URL) – A symbolic name (such as www.yahoo.com) that is registered through the Internet Domain Name System (DNS) to a specific IP address, so that browser users specifying the URL will always be directed to the correct system based on its IP number.

© Internet Research Group, 1999

Appendix 2: Measuring Access Performance

It's worth spending a little time talking about how performance can be measured, and why the measurements often yield surprising and informative results. We'll use Keynote Systems as an example, although other approaches exist and also provide valuable data.

Keynote has roughly 100 measurement systems distributed throughout the Internet. Like Akamai, Keynote places these systems at important places that are characteristic of where Internet demand originates. Every 15 minutes or so, each of these systems makes relatively simple measurements of the access performance for a set of Web servers. Some servers are chosen to provide an overall benchmark of Internet performance, while others are measured because the owner (or one of its competitors!) has contracted with Keynote for data.

The Keynote measurements include factors due to server performance and factors due to Internet delays. If the performance of a site degrades at times of the day as seen by all the measurement servers, it's clear that there are server issues. But there are many examples where for a given server at a specific time of day, access is essentially perfect from certain areas of the Internet (suggesting strongly that the server is performing perfectly, probably under light load) whereas the performance is quite poor from other parts of the Internet (*see Figure 6*). In these cases, there is compelling although circumstantial evidence that the problem is in the Internet, and the problem is severe. These are the problems that a content delivery service directly addresses, because all indications are that the problem occurs at the peering points of the federated networks. By placing content servers in all networks, at the edges, Akamai assures that access to the content it serves never has to cross these boundaries.

© Internet Research Group, 1999

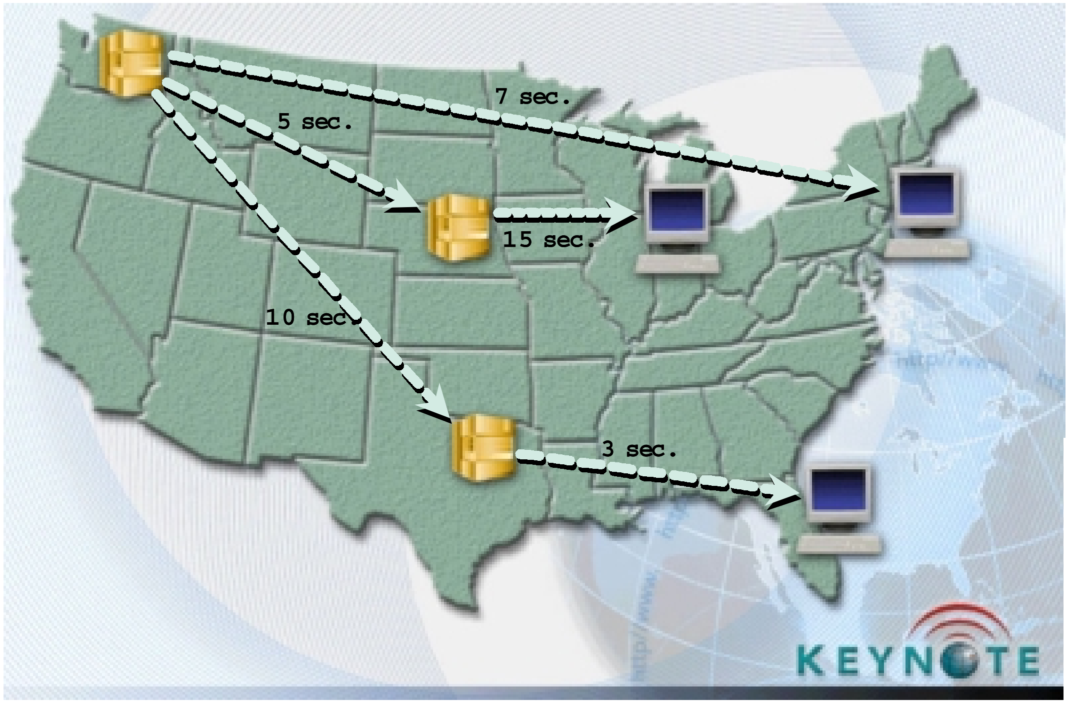

Figure 6. Keynote Systems' measurements reveal wide variances in access speeds to the same origin server from different parts of the Internet, strongly suggesting that Internet delays, rather than server performance, are at fault.

Source: Keynote Systems

Systems like Keynote's are a good way to demonstrate and quantify the kinds of problems that Akamai and other content distribution systems address. Akamai uses Keynote on an ongoing basis to validate the performance of the Akamaized version of the Web site and to assure that the Akamaized performance is an improvement over the unmodified site.