Akamai Managed Application Acceleration Services for Dynamic Web Applications

Real-World Performance Testing for Dynamic Web Applications

A white paper commissioned by Akamai Technologies, Inc.

© 2007 The Tolly Group
Executive Summary 3

Dynamic Web Acceleration Services 3
Real-World Web Performance Testing 4

Akamai EdgePlatform Delivers Results 7

Route and Connection Optimization 7

Dynamic Mapping
Route Optimization
Connection Optimization

Compression 10

Bringing Content Close to Users 11
Pre-fetching
Advanced Caching

Test Set up & Methodology 12

Detailed Test Results 13
Executive Summary

Enterprise IT professionals face a variety of challenges in delivering and managing centrally located Web-based applications to employees, business partners and customers on a global scale. What’s required is a new dynamic that places significant emphasis on route optimization, transport protocol optimization and data compression to ensure rapid and consistent application response times over the Internet.

Akamai Technologies, Inc.’s Web Application Accelerator (for B2B) and Dynamic Site Accelerator (for B2C) harness such techniques to accelerate and optimize the delivery of business applications on a global scale, yielding performance improvements unmatched by traditional acceleration solutions. These services deliver an end-to-end acceleration solution by offering unique "middle-mile" acceleration techniques like route optimization that allows for "steering" around Internet bottlenecks leading to both a faster and a more consistent/reliable experience.

Akamai’s Web Application Accelerator and its Dynamic Site Accelerator are built on the Akamai EdgePlatform — a massive global network of 20,000 servers strategically deployed across 1,100 networks in 70 countries. This enables businesses to extend their dynamic, highly interactive and secure Web sites to the Akamai EdgePlatform, thus bringing content close to requesting consumers. Akamai’s global reach, along with optimized routing, connection and compression technologies, enables businesses to optimally deliver dynamic Web content and ensures consistently faster Web page response times.

Dynamic Web Acceleration Services

Akamai Technologies commissioned The Tolly Group to benchmark the performance of the company’s Dynamic Web Acceleration Services, which encompasses Web Application Accelerator and Dynamic Site Accelerator — both of which promise to accelerate Web-based enterprise-class content and business processes on a global scale. Since both Dynamic Web Acceleration Services are delivered over the same Akamai global network, the company says users should experience similar performance results.

The Dynamic Site Accelerator service (for B2C) and the Web Application Accelerator (for B2B) enables e-businesses to accelerate dynamic, highly-interactive Web applications securely, resulting in improved performance,
higher availability, and an enhanced worldwide user experience.

Application requests are routed across Akamai’s trusted, globally distributed platform, which accelerates Web application content through a number of techniques, including dynamic mapping of user requests, route optimization technology, connection optimization, and caching technology.

Tolly Group engineers conducted extensive tests pertaining to improving application performance for Web applications. Engineers set up one Web server each in the U.S., United Kingdom and Singapore. The servers fielded requests via HTTPS from a wide variety of clients scattered worldwide. Application response times were measured using the Keynote testing service — both without the benefit of the Akamai service, as well as with the Dynamic Web Acceleration Services enabled. Tests were conducted during April 2006.

Results show that the Akamai Dynamic Web Acceleration Services improved the performance of tested applications via secure Web page download by as much as 5X across geographic scenarios tested for B2B and B2C scenarios. Tests also show that consistency of Web performance improved in addition to the performance gain. These benefits can be attributed to various Web acceleration technologies which are described in later sections of this document.

Real-World Web Performance Testing

For the tests, engineers constructed a test bed of global proportions to represent actual B2B and B2C network conditions that online business users are likely to encounter on a daily basis. The goal was to create a “real-world performance testing environment” to measure Web application acceleration performance of the Akamai Dynamic Web Acceleration Services against live Internet conditions that fluctuated on a time-of-day and day-by-day basis. By utilizing such a real-world scenario, users can gain a deeper appreciation of the end-to-end performance gains delivered by the Akamai service, even as traffic and data requests traverse multiple service provider backbones.

Test results show that Akamai Dynamic Web Acceleration Services can improve download times for Web pages by up to 5X, and often 2X to 3X within an origin server location. (See Figure 1, next page.) These performance improvements become more significant considering that the tests
were conducted against real-world live Internet traffic in which the data requests and responses traversed multiple service provider networks. The performance improvement was acquired by effectively bringing content close to requesting end users.

Interestingly, application acceleration appliances used in enterprise networks offer only a subset of the capabilities featured in Akamai's Dynamic Web Acceleration Services. Akamai's Web-based application acceleration services extend beyond the datacenter to address "middle-mile" performance and reliability bottlenecks, in addition to the traditional "first-mile" bandwidth and protocol optimization techniques. As a result, Akamai provides an end-to-end acceleration solution from the origin all the way to the end-user, regardless of geography.

The consistency improvement with Akamai is apparent in Figures 2-3. Figure 2 shows the application performance results for the Asia Pacific (APAC) region tested, for the U.S.-based origin and APAC users. Note the consistency in response time for the Akamai results and the significant variance in response times without the benefit of the Dynamic Web Acceleration Services. Note how in the chart, the yellow dots are tightly clustered and do not deviate much. These results quantify the benefit of the route optimization applied to wide area traffic. Figure 3 shows the percentage improvement in response time consistency for the three regions...
tested.

So, testing shows that not only does the Akamai Dynamic Web Acceleration Services provide faster response times, but they do so with less deviation than traffic that rides over the general Internet, without the benefit of Akamai’s technology. So, precisely what is at the heart of the Akamai Dynamic

Application Performance Consistency — U.S Origin Server and APAC Users with Application Acceleration
As Reported by Keynote Testing Service

![Graph showing response time consistency for U.S. origin server and APAC users with and without application acceleration.](image)

Elapsed time
- APAC - Internet
- APAC - Akamai

Source: The Tolly Group, November 2006

Figure 2

Percentage Improvement in Response Time Consistency — Origin Servers and Worldwide Users with Application Acceleration
As Reported by Keynote Testing Service

![Bar chart showing percentage improvement in response time consistency.](image)

Origin server location
- APAC
- EU
- US

Note: Origin server totals for percentage of response time improvement is based upon worldwide client access.

Source: The Tolly Group, November 2006

Figure 3
Web Acceleration Services that enables them to deliver faster application response times with minimal variability? To understand that, you need to look closely at the technologies Akamai brings to bear to solve the B2C and B2B Web application performance problem.

**Akamai EdgePlatform Delivers Results**

Akamai’s EdgePlatform is a massive distributed computing network, consisting of more than 20,000 servers worldwide, which acts as an optimized “overlay” to the public Internet.

By overlaying the existing Internet with a network of secure, performance-optimized server hardware, Akamai is able to provide customers with a predictable, scalable, secure platform on which to run and deliver a wide variety of content and e-business applications. The following are key acceleration solutions that Akamai EdgePlatform deploys to improve Web performance and reliability. (See Figure 4.)

**Route and Connection Optimization**

Akamai’s Dynamic Web Acceleration Services provide capabilities for accelerating the delivery of completely dynamic or transactional content which involves an origin server access on every request. These capabilities are delivered through the implementation of route and connection optimization, which includes three key technologies:

- Dynamic Mapping
- Route Optimization
- Connection Optimization

**Dynamic Mapping**

Dynamic mapping ensures that users are dynamically mapped to the optimal entry point on the Akamai network. Dynamic mapping directs user requests for application content to an optimal Akamai edge server using the multi-level Domain Name System. It is completely transparent to networks and end users – end users are already using it today for the
Dynamic mapping finds the closest Akamai edge server for each user based on network proximity. The network proximity is not necessarily the same as geographical proximity. The three main components locate the “closest” edge cache to an end user from a network point-of-view, taking into account packet loss, throughput and latency. Dynamic mapping also takes into consideration the status of the target edge server by checking that the server has the requested object on its hard disk and that the server has resources available to serve the end-user.

**Route Optimization**

Akamai uses “SureRoute for Performance” technology for route optimization. SureRoute takes advantage of Akamai’s extensive network deployment to an HTTP overlay network, which allows users to both improve performance and avoid network failures for communications between edge servers and a customer’s origin site. As a result, fully dynamic and transactional content is delivered more quickly and more reliably to end users, even when it is uncacheable.
SureRoute identifies alternate paths over which an Akamai edge server can communicate with a customer’s origin server and uses these alternatives to either improve the performance of the connection, or to provide for failover in the event that a direct path is congested or otherwise unavailable. (See Figure 5.)

When an Akamai edge server contacts the origin server, the "direct" path typically is a route obtained through BGP (Border Gateway Protocol). When SureRoute is used, alternate paths to the origin server are accessed, for example, by sending a request from the Akamai edge server to other Akamai servers, which then issue their own requests to the origin server. Based on real-time performance measurements of HTTP downloads, latency, and loss frequency, SureRoute may use an alternate path to communicate between an edge server and the origin if it can bypass network congestion.

SureRoute is used to optimize communication between edge servers and the origin server. SureRoute chooses the optimal path to the origin device to ensure that a site is continuously accessible and that dynamic content is delivered to end users via the optimal path.

How Akamai’s SureRoute Works

SureRoute “looks ahead” to find the fastest route to an origin server, ensuring content is delivered quickly.

Source: The Tolly Group, November 2006  
Figure 5
SureRoute enables an Akamai edge server to obtain content from a customer origin server through the Akamai network via an optimal Akamai server near the origin. Inside the Akamai network, proprietary techniques are used to accelerate the content delivery and avoid Internet congestion points and unnecessarily long routes. The optimal path between the origin site and edge server is selected based on real-time data collected by Akamai — based on this data, it may be direct (as recommended by BGP), or indirect through an intermediary Akamai edge server.

**Connection Optimization**

Once the route optimization technology selects an optimal route, Akamai further optimizes this connection using a combination of persistent connections and protocol optimization, leading to improved Web performance and a faster and more reliable end-user experience. The persistent connections avoids TCP slow-start and connection setup for every new user request. The TCP slow-start starts at a low speed and increases the speed during connection, but many Web connections do not last long enough to reach top speed. The persistent connection also reduces the load on the origin server by maintaining connections to a small set of Akamai servers instead of millions of end users. As a result, no new connections need to be negotiated and established. Setting up a new connection is more CPU intensive than maintaining an existing connection.

Additional optimization is provided by tuning TCP connection parameters. The values of these parameters govern the speed of downloads by controlling the rate at which data is transmitted, and the manner in which the server recovers from packet loss. Akamai professionals help the customers to tune these parameters on a per-site basis, resulting in additional performance improvements for the delivery of dynamic application content.

**Compression**

The Web Application Accelerator and Dynamic Site Accelerator services also leverage performance-enhancing compression technology. Compressing content before it is sent to the end-user is especially effective at reducing transfer times for HTML content to users on slower connections. Since application pages are often light on graphics, this can be particularly effective. This and other Akamai enhancements provide improved end-to-end performance, independent of the networks traversed from user to origin.
Despite long-standing “support” for compression in the HTTP protocol, most applications serve content uncompressed, due to a variety of subtle client-side issues. Exceptions need to be made for specific browser versions, operating systems, and content. Akamai’s Web Application Accelerator and Dynamic Site Accelerator take advantage of the benefits of compression by providing the ability to carefully apply compression to only the data that can benefit from such an operation.

**Bringing Content Close to Users**

Delivery of site content via the Akamai EdgePlatform follows a structured path: Each end user request is directed to an Akamai edge server via Akamai’s intelligent DNS. Upon receiving a request for content, an edge server retrieves the appropriate content (HTML page, image, document, etc.) from a local cache or from the origin server if the content is uncacheable. It then applies features selected by the customer, performs any appropriate edge processing, and finally delivers the resulting content to the requesting user.

To ensure a high cache hit rate for cacheable content, such as support documentation or software downloads, requests are directed to the Akamai servers most likely to contain the requested content. Cache coherence features ensure that the content served to end-users is fresh. For content that has been marked as uncacheable, such as personalized account reports or real-time inventory status, it is retrieved from the origin server for each request, as appropriate. Trips to the origin server to refresh and retrieve content are optimized via extensive use of persistent connections (reducing latency through reuse of established connections). Additional connection optimization techniques tune the parameters that govern server-to-server connections, as well as reduce the number of long-haul round trips required to retrieve content embedded in application pages.

**Pre-fetching**

The rendering of a Web page generally requires multiple round-trips between the client and the Web server. The first round trip is the request/response for the Web page HTML, and subsequent round trips are for the requests/responses for the embedded objects (images) within the Web page. By leveraging a prefetching capability designed to reduce the number of long-haul round trips required to retrieve embedded content, Akamai is able to minimize the amount of time it takes a browser/client to load and display a Web page, including the embedded objects.
For dynamic application content, customers often configure Akamai’s edge servers to not cache the HTML pages generated by the origin servers, but to act as a proxy for the request and response. By parsing the HTML returned by the origin server, the edge servers can begin making requests back to the origin host for the embedded objects, while simultaneously sending the HTML back to the client. The client would then parse the HTML page, and send requests to an Akamai edge server for the embedded objects. Because the Akamai edge-server has already retrieved the objects from the origin even prior to receiving a request for these objects from the client, it is able to respond to the client requests immediately, serving the requested objects from its memory cache.

**Advanced Caching**

Akamai’s Advanced Cache Control provides a comprehensive set of configurable cache settings that enable users to specify, at a granular level, how Akamai edge servers are to cache and serve content. Advanced Cache Control provides users with more than 40 different options for controlling content freshness and expiration, query string result caching, redirects, cookies, and session IDs.

**Test Set up & Methodology**

Tolly Group engineers measured the real-world Web performance for three different origin server locations with or without Akamai application acceleration services. The origin server locations were San Antonio, Texas (U.S.), London (United Kingdom) and Singapore. Each origin server was loaded with 191 KB of a Web page including a 40 KB HTML document, one CSS file, small JavaScript files, flash objects and 40 images (mostly 3 ~ 5 KB, two 10 KB, one 15 KB and one 20 KB file.)

Engineers used the Keynote testing service to measure the real-world Web performance around the world. Engineers picked 53 Keynote testing agents around the world using 25 agents from major cities in the U.S., 15 agents from major cities in the European Union, 11 agents from the major cities in Asia Pacific region, one agent from Mexico City and one agent from Johannesburg, South Africa.
Engineers configured each agent to retrieve a Web page from three different locations with or without Akamai once an hour for two weeks of the test duration and record the response time for every request. Akamai edge servers established secure HTTPS sessions with clients and servers. The testing started on 03 April 2006 and ended on 16 April 2006. The Keynote testing service reported the Web application response time and availability with respect to the agents and the time with or without Akamai Dynamic Web Acceleration services enabled.

**Detailed Test Results**

In the U.S.-based origin server test, 53 clients utilized worldwide took about 5.3 seconds on average to complete downloading a page directly from the server and about 1.2 seconds via the Akamai servers. Viewed from a slightly different lens, a user in an Asia/Pacific country requesting information from a U.S.-based business that utilizes Akamai application acceleration will benefit from a 479% improvement in response time. Likewise, a user based in the United Kingdom would experience a 236% improvement.

---

**Akamai Dynamic Web Acceleration Services Test Bed**

![Diagram of Akamai Dynamic Web Acceleration Services Test Bed](image)

Source: The Tolly Group, November 2006

Figure 6
in response times when Akamai is used, and even U.S.-based users would benefit from a response time improvement of 141%.

In the U.K.-based server test, the same clients took about 5.9 seconds directly from the server (using the standard Internet path) and about 1.6 seconds via the Akamai service. On a percentage basis, that translates into response time improvements ranging from 72% for U.K. users, to 386% for users requesting data from Asia/Pacific countries. (See Figures 1 and 6.)

In the Asia-based test, the worldwide clients consumed about 16.2 seconds to download the requested Web page without Akamai Web application acceleration and about 4.5 seconds with Akamai. In this scenario, tests revealed response time improvements ranging from 228% to 273%, depending upon geographic location of the requesting user. (See Figures 1 and 6.)
Terms of Usage

USE THIS DOCUMENT ONLY IF YOU AGREE TO THE TERMS LISTED HEREIN.

This document is provided, free-of-charge, to help you understand whether a given product, technology or service merits additional investigation for your particular needs. Any decision to purchase must be based on your own assessment of suitability.

This evaluation was focused on illustrating specific features and/or performance of the product(s) and was conducted under controlled, laboratory conditions and certain tests may have been tailored to reflect performance under ideal conditions; performance may vary under real-world conditions. Users should run tests based on their own real-world scenarios to validate performance for their own networks. Commercially reasonable efforts were made to ensure the accuracy of the data contained herein but errors and/or oversights can occur. In no event shall The Tolly Group be liable for damages of any kind including direct, indirect, special, incidental and consequential damages which may result from the use of information contained in this document.

The test/audit documented herein may also rely on various test tools the accuracy of which is beyond our control. Furthermore, the document relies on certain representations by the sponsor that are beyond our control to verify. Among these is that the software/hardware tested is production or production track and is, or will be, available in equivalent or better form to commercial customers.

When foreign translations exist, the English document is considered authoritative. To assure accuracy, only use documents downloaded directly from The Tolly Group’s Web site.

All trademarks are the property of their respective owners.

The Tolly Group is a leading global provider of third-party validation services for vendors of IT products, components and services.

The company is based in Boca Raton, FL and can be reached by phone at (561) 391-5610, or via the Internet at http://www.tolly.com, sales@tolly.com