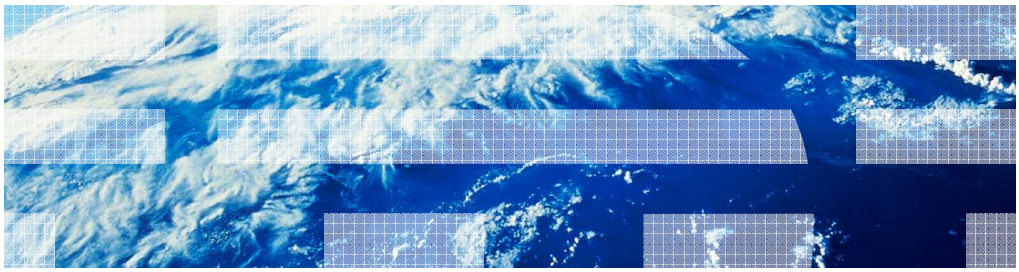
WebSphere DataPower Release 1.0.0

XE82 – WebSphere Application Accelerator for Public Networks



This presentation discusses the WebSphere® Application Accelerator for Public Networks (WAXPN).

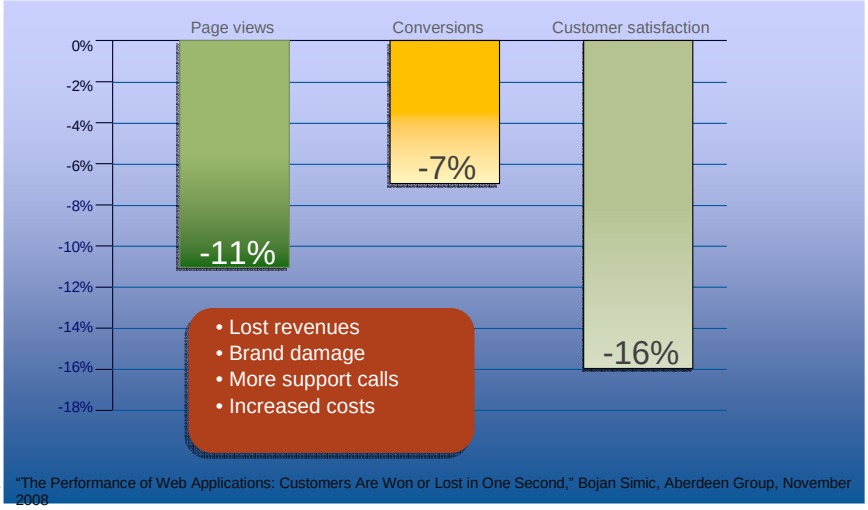
WebSphere Application Accelerator for Public Networks Combining DataPower with the Akamai EdgePlatform

- WebSphere Application Accelerator for Public Networks (WAXPN) is a managed service sold by IBM and subscribed to by your customers
  - WAXPN leverages the Akamai Internet fabric to accelerate and secure web applications to users
  - WAXPN requires deployment of a DataPower® XI50/XS40 with application optimization (AO) or an XE82 (which comes with AO as a standard option)
  - IBM also provides a WAXPN toolkit to enhance how DataPower and Akamai work together
    - The toolkit is a set of stylesheet templates that can be customized to provide:
    - Enhanced caching capabilities
    - Adaptive Load Optimization
    - Security Handshake providing reduced DDoS Risk
    - SureRoute services to identify optimal Internet routes

WebSphere Application Accelerator For Public Networks (WAXPN) is a managed service sold by IBM and subscribed to by your customers to accelerate and secure web applications to users. WAXPN uses the Akamai Internet fabric to accelerate requests across the Internet. WAXPN requires deployment of a DataPower XI50/XS40 with application optimization (AO) or an XE82 (which comes with AO as a standard option). IBM also provides a WAXPN toolkit to enhance how DataPower and Akamai work together. The toolkit is a set of stylesheet templates that can be customized to provide:

- Enhanced caching capabilities
- Adaptive Load Optimization
- Security Handshake providing reduced DDoS Risk
- SureRoute services to identify optimal Internet routes

Background: Internet response time challenges negatively impact revenue and user satisfaction



1. "The Performance of Web Applications: Customers Are Won or Lost in One Second," Bojan Simic, Aberdeen Group, November 2008
2. Source: Internet World Stats, Usage and Population Statistics, http://www.internetworldstats.com/stats.htm, December 22, 2010

Internet response time challenges negatively impact revenue and user satisfaction.

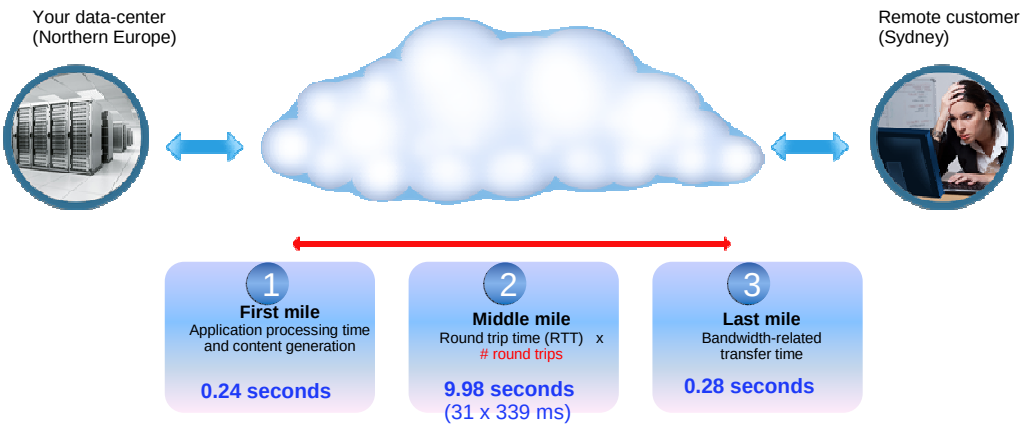
The chart shows the results of a study showing the impact of increased response time with respect to revenue, cost, and so on.

"The Performance of Web Applications: Customers Are Won or Lost in One Second," Bojan Simic, Aberdeen Group, November 2008

Source: Internet World Stats, Usage and Population Statistics, www.internetworldstats.com/stats.htm, December 22, 2010

Response time to a remote user can take 10.5 seconds too slow!

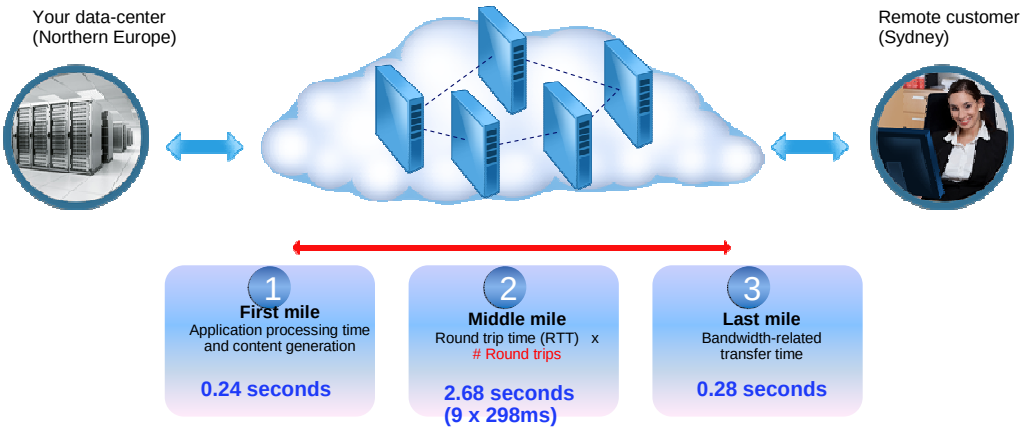
95% of total response time is due to the middle mile



From Stockholm to Sydney, rendering a typical 70k web page with 25 objects requires 31 round trips

Even if the backend application servers respond very quickly, response time can be significantly slowed through the Internet (referred to here as the “middle mile”).

With an efficient middle mile – Total response time reduced to 3.2 seconds with only nine round trips



1 First mile
Application processing time and content generation
0.24 seconds

2 Middle mile
Round trip time (RTT) x # Round trips
2.68 seconds (9 x 298ms)

3 Last mile
Bandwidth-related transfer time
0.28 seconds

From Stockholm to Sydney, rendering a typical 70k web page with 25 objects requires nine round trips with Akamai

WAXPN with Akamai's network fabric accelerates and secures the middle mile through the Internet.

WAXPN network fabric leveraging Akamai

- Servers deployed into thousands of ISPs close to applications and end users
- 84,000+ Servers
  - 1,500+ Locations
  - 1,000+ Networks
- Akamai “edge” is within one network hop of 90% of Internet users
- Transparent to end-users and applications
- Routing based on DNS resolution

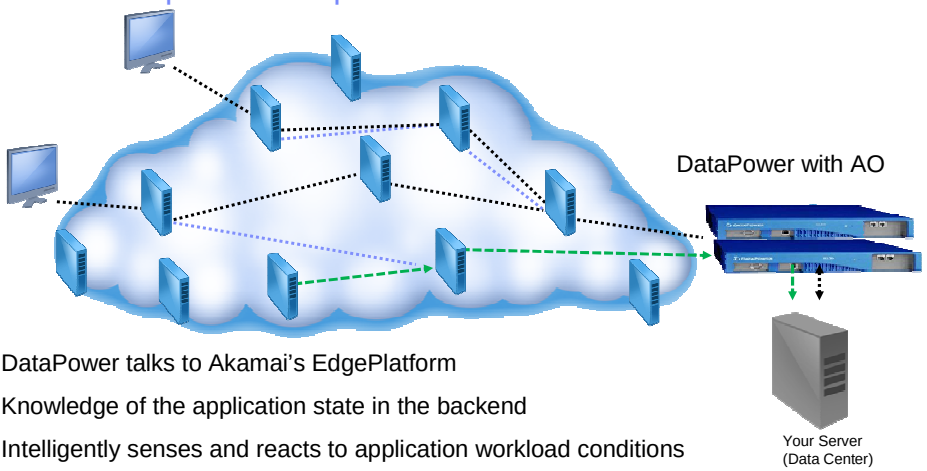
The WAXPN Network fabric consists of more than 80,000 servers in 1500 locations worldwide. A WAXPN edge server is within one network hop of 90% of Internet users. This service uses DNS resolution To route requests to edge servers and is therefore transparent to end users.

DataPower and WAXPN toolkit

WAXPN and DataPower with AO work together to accelerate and secure web applications from the browser through the Internet all the way to the enterprise application servers

WAXPN and DataPower with AO work together to accelerate and secure web applications from the browser through the Internet all the way to the enterprise application servers.

WAXPN adaptive load optimization

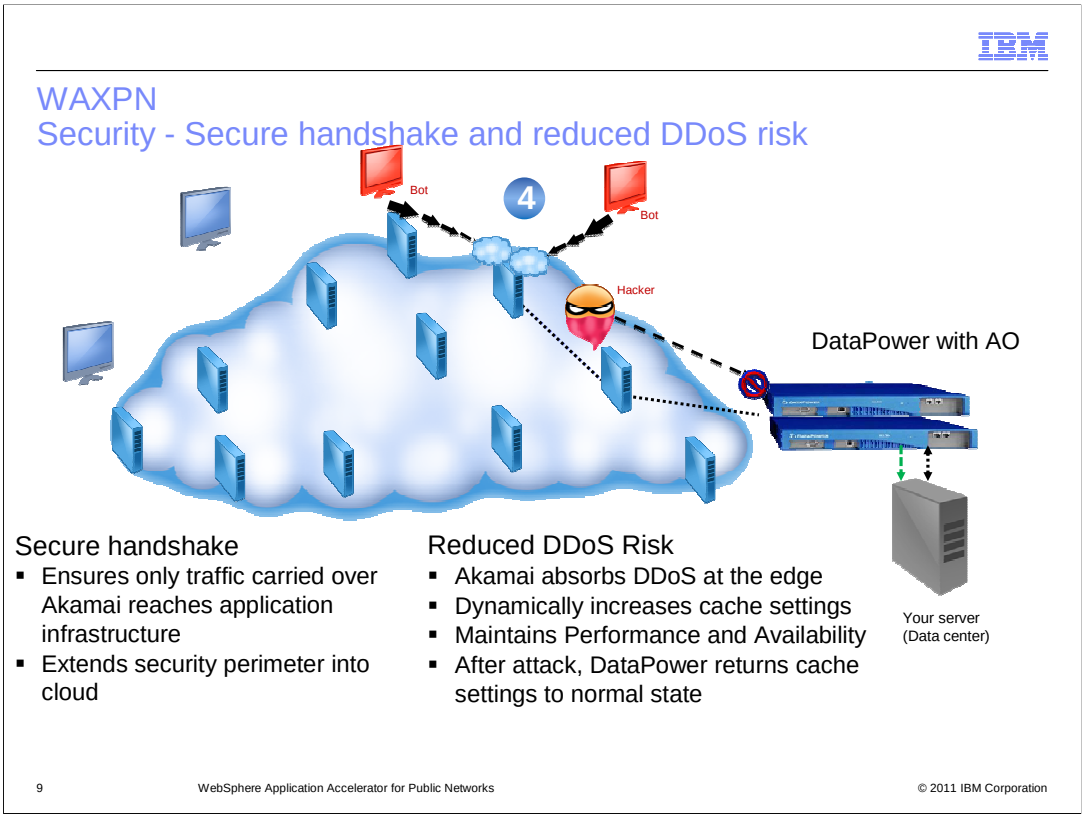


- DataPower talks to Akamai's EdgePlatform
- Knowledge of the application state in the backend
- Intelligently senses and reacts to application workload conditions
- Balances traffic across WebSphere VE Deployment infrastructure
- Dynamically modifies WebSphere Application Accelerator cache settings in the cloud
- Adapts and throttles traffic reaching the data center

During spikes of activity, users will still enjoy snappy response time with information fed from the WAXPN cache.

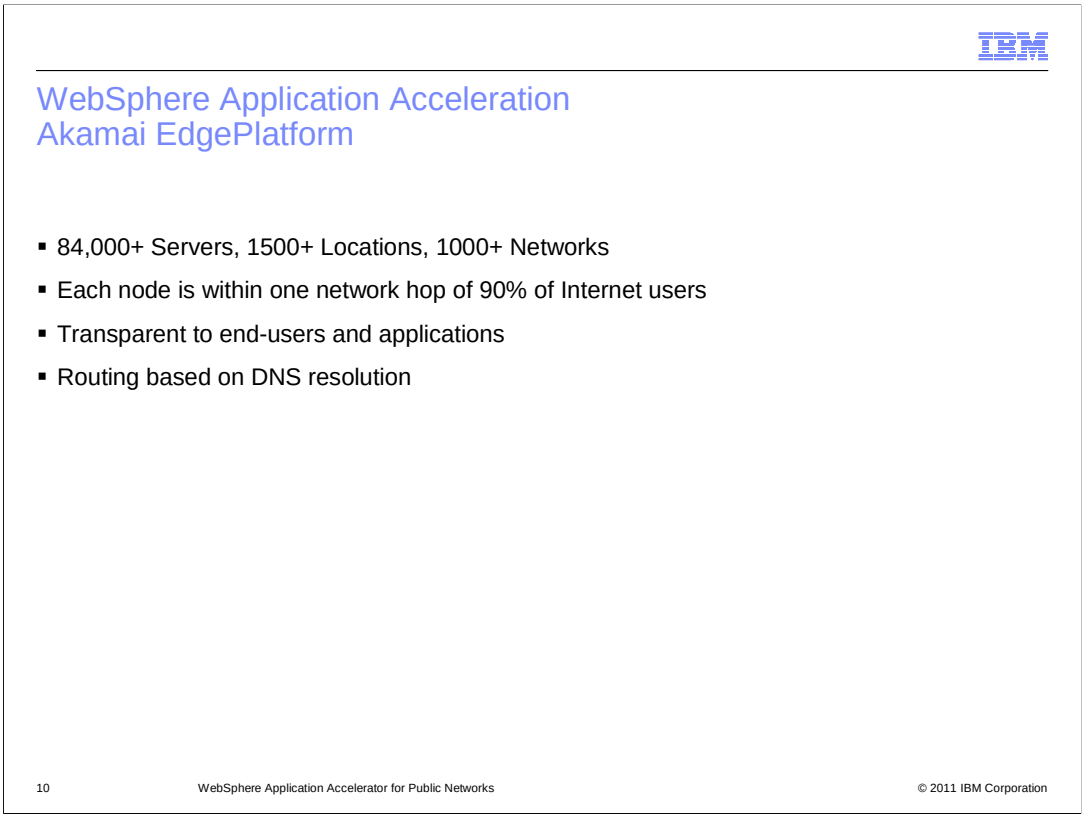
As the backend workload decreases, the caching duration of dynamic data is decreased.

This improves application availability and responsiveness while still providing the freshest data to users.



The WAXPN network fabric is designed to absorb DDoS

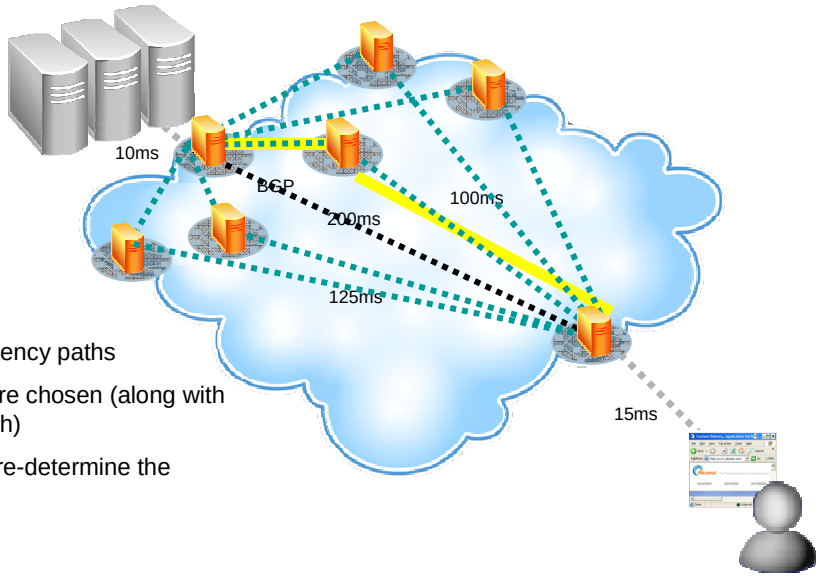
The WebSphere DataPower appliance together with WebSphere Application Accelerator for Public Networks provides a second layer of security protection with a security-enhanced handshake between the Akamai network and DataPower, which helps ensure that only traffic carried over the Akamai network reaches application infrastructure—thereby extending the security perimeter into the cloud and protecting the origin from external threats.



To control the reliability, availability and performance of the Internet IBM has partnered with Akamai Technologies. Akamai owns over 84,000 servers, at 1500 plus locations, geographically distributed across the globe.

Each node or server is within one network hop of 90 percent of all Internet users.

WebSphere Application Acceleration
SureRoute path Optimization

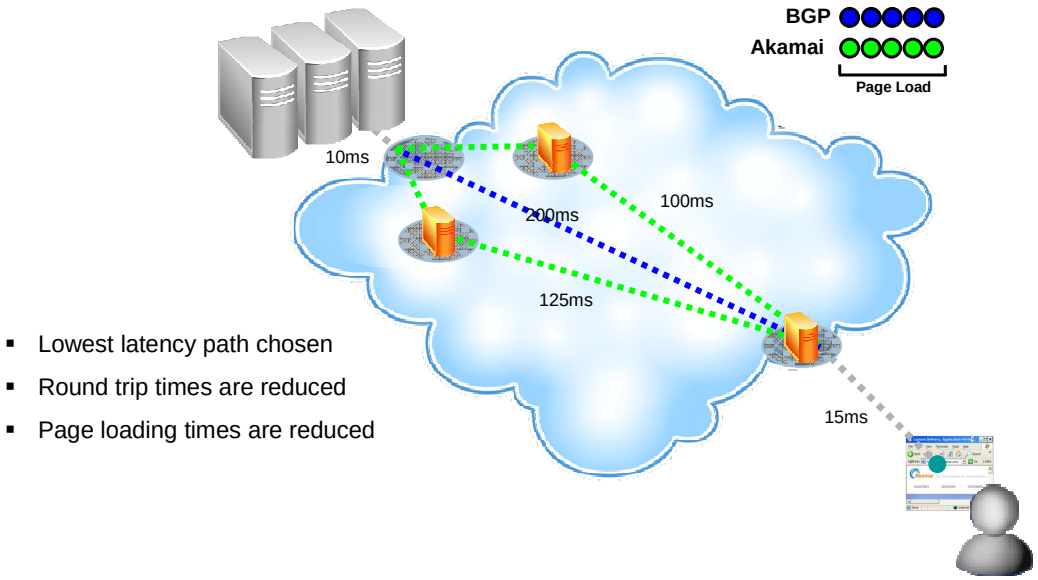


- Discovers low latency paths
- Best two paths are chosen (along with Internet BGP path)
- Request races pre-determine the optimal path

Sure Route path optimization uses parallel requests on disjoint paths to determine and cache the fastest or lowest latency paths across the Internet. The redundant disjoint paths can also be used for fail-over purposes.

For example, in this picture, the best path from the client's web browser to the web servers is the path highlighted in yellow. The B G P path is also identified as the black dotted line.

WebSphere Application Acceleration
SureRoute versus BGP

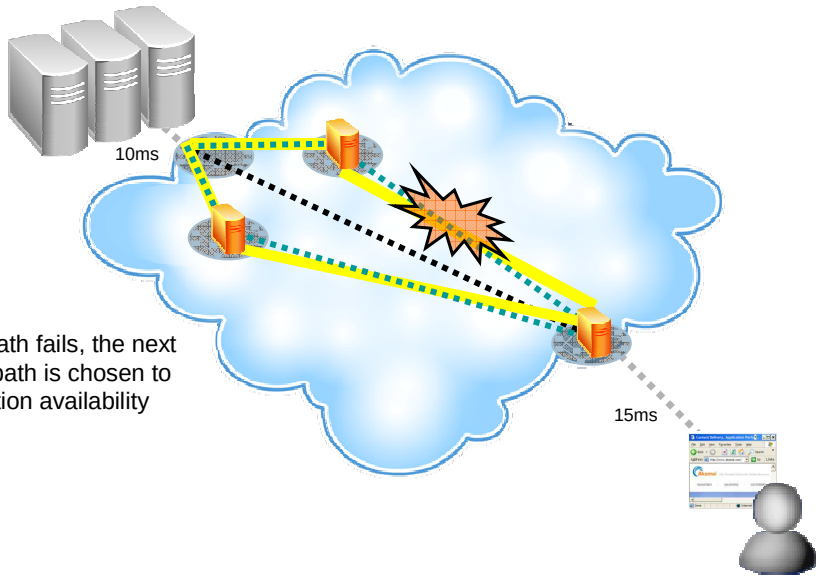


- Lowest latency path chosen
- Round trip times are reduced
- Page loading times are reduced

Sure Route verses B G P. Because the lowest latency path is chosen, the reduction in page rendering times is cumulative. Complex web applications require many round trips between the user agent and the origin server, and even the smallest improvements in latency, add up quickly to a user visible difference.

For example, in this picture, the client uses a web browser to access a website. Because the top most path has the lowest latency, this will be used to service the user's request. The request travels from the orange server closest to the user, to the top-most orange server, and eventually reaches the web servers. The web servers return a document which travels the same path back to the user's web browser. Five round trips in this fashion would result in about a 1 second latency improvement over the B G P path, and this improvement gets larger with each new request.

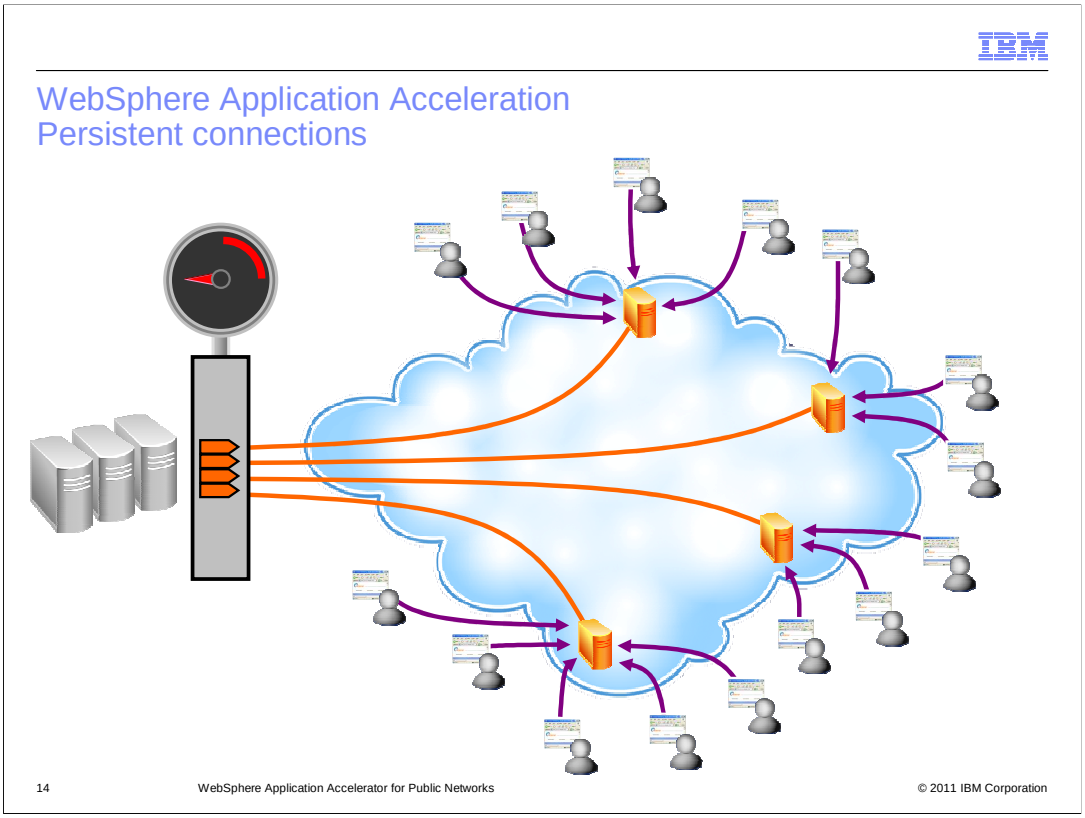
WebSphere Application Acceleration
Automated path failover



- If the optimal path fails, the next best alternate path is chosen to ensure application availability

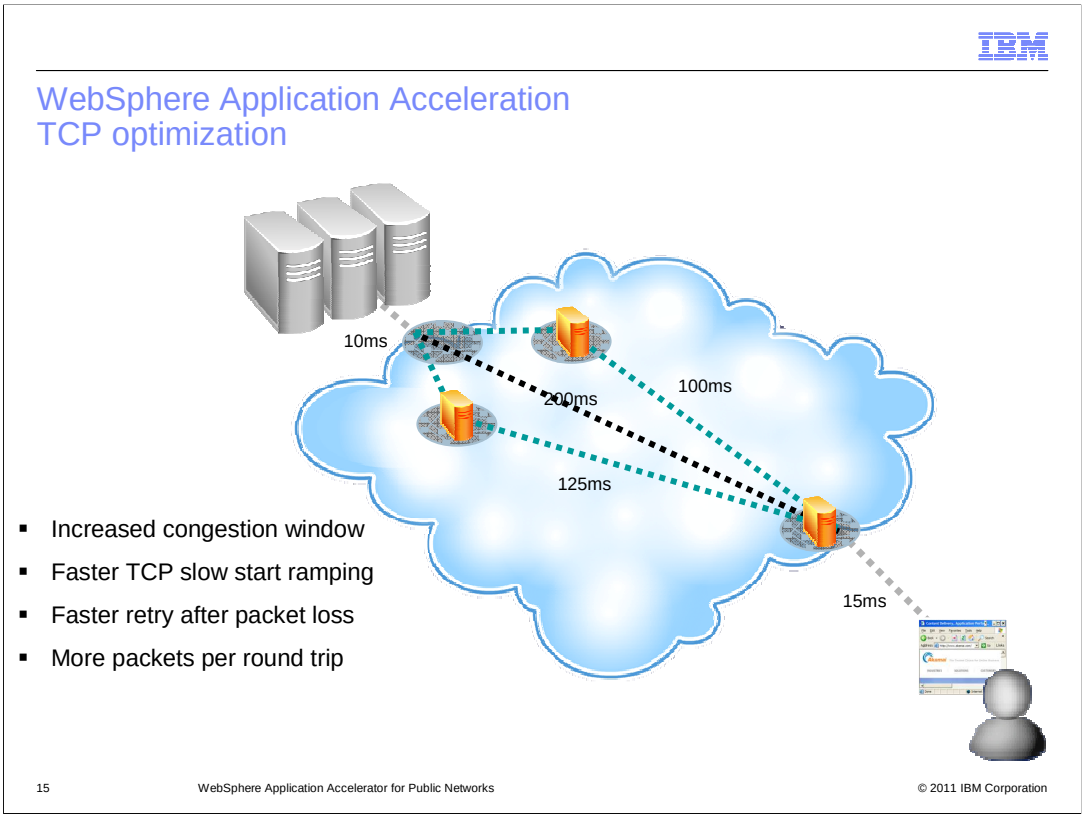
If the optimal path fails, the next best alternative disjoint path is chosen to ensure application availability. Alternative paths are cached and thus convergence time is fast.

For example, in this picture, the two best paths are highlighted in yellow. Once the top most path fails, the bottom most path is automatically used.



Persistent connections greatly reduce the TCP handshake overhead, but this situation is further optimized by maintaining connection pools at the closest point of presence to a local group of user agents. This further reduces the load on the origin server and has security benefits as denial of service attacks can be detected close to the originator and can be dropped or filtered well before the origin server.

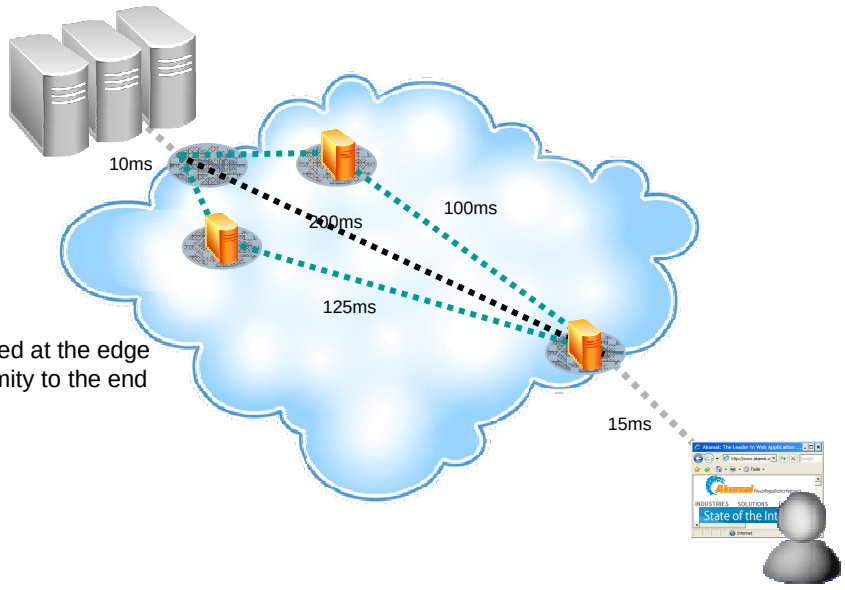
For example, in this picture, each client connects to an orange server to access the web servers. These orange servers maintain the multiple connections required to directly talk to each client. However, these orange servers are the only ones that have connections with the web servers. The orange servers pass each clients requests through these connections to the web servers.



The Akamai overlay also performs transport TCP optimizations to improve latency and throughput. Some examples are, increased congestion window, faster TCP slow start ramping and faster retry after packet loss.

For example, in this picture, the client uses a web browser to access a website. The request travels from the orange server closest to the user, to the top-most orange server, and eventually reaches the web servers. The web servers return a series of documents. When this document is received by the top-most orange server, the server performs some level of optimization to forward the documents to the orange server closest to the user. This server then forwards the document to the web browser.

WebSphere Application Acceleration
Edge caching



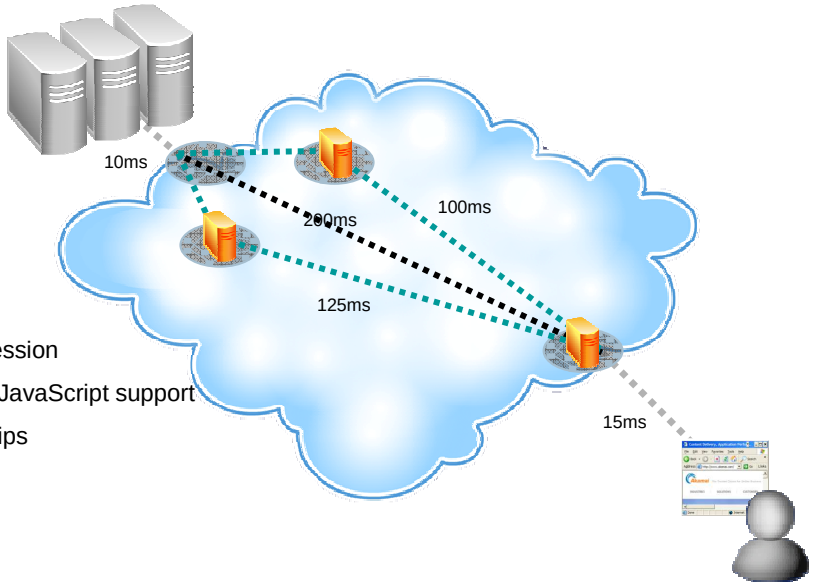
- Objects cached at the edge in near proximity to the end user

16 WebSphere Application Accelerator for Public Networks

Edge Caching caches high frequency objects in near proximity to the consumer or end user. Caching reduces the number or round trips to the origin or dramatically reduces the quantity of data transmitted from the origin to the end user. When combining IBM WebSphere DataPower technology it is possible to enable the Akamai network to cache dynamic content, over and above the typically cached static content such as images and JavaScript.

For example, in this picture, the client uses a web browser to access a website. The request travels from the orange server closest to the user, and eventually reaches the web servers. The web servers return a document. When this document is received by the orange server closest to the user, the document is cached. This server then forwards the document to the web browser. Future requests by this user or other users that use the same orange server as the original user, the cached document is returned to the user. Serving up this cached document improves user observed latency, reduces load on the web server, and reduces traffic on the larger network.

WebSphere Application Acceleration Compression



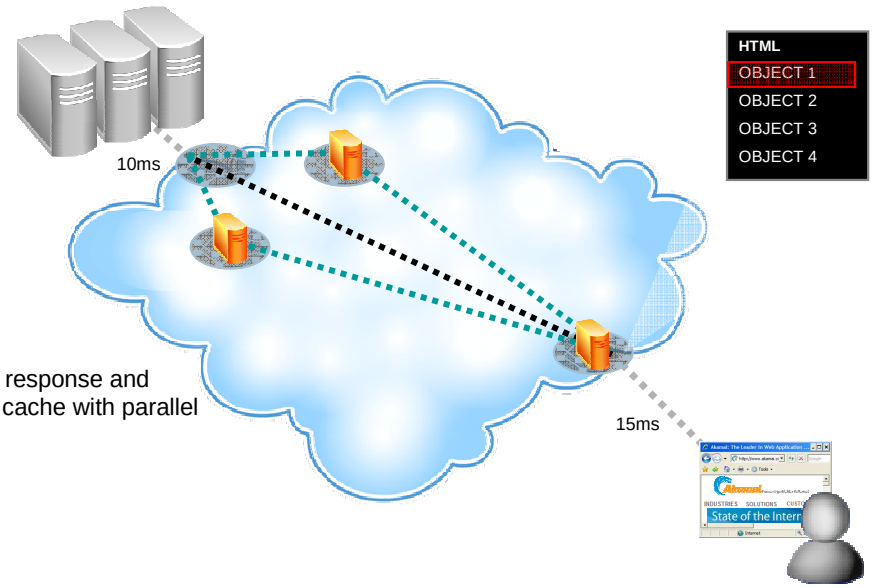
- Intelligent compression
- Full browser and JavaScript support
- Reduces round trips

17 WebSphere Application Accelerator for Public Networks

Intelligent Compression compresses content to reduce the size and therefore the latency required to transmit the objects across the Internet. A typical 100k byte page compresses down to 15k bytes. Intelligent compression ensures that already compressed objects are not double compressed and that compatibility is maintained regardless of user agent type or version.

For example, in this picture, the client uses a web browser to access a website. The request travels from the orange server closest to the user, to the top-most orange server, and eventually reaches the web servers. The web servers return a large document. When this document is received by the top-most orange server, the document is compressed and sent to the orange server closest to the user. This server forwards the document to the web browser. Since there is browser support for compression, this document is decompressed and rendered for the user.

WebSphere Application Acceleration Prefetching

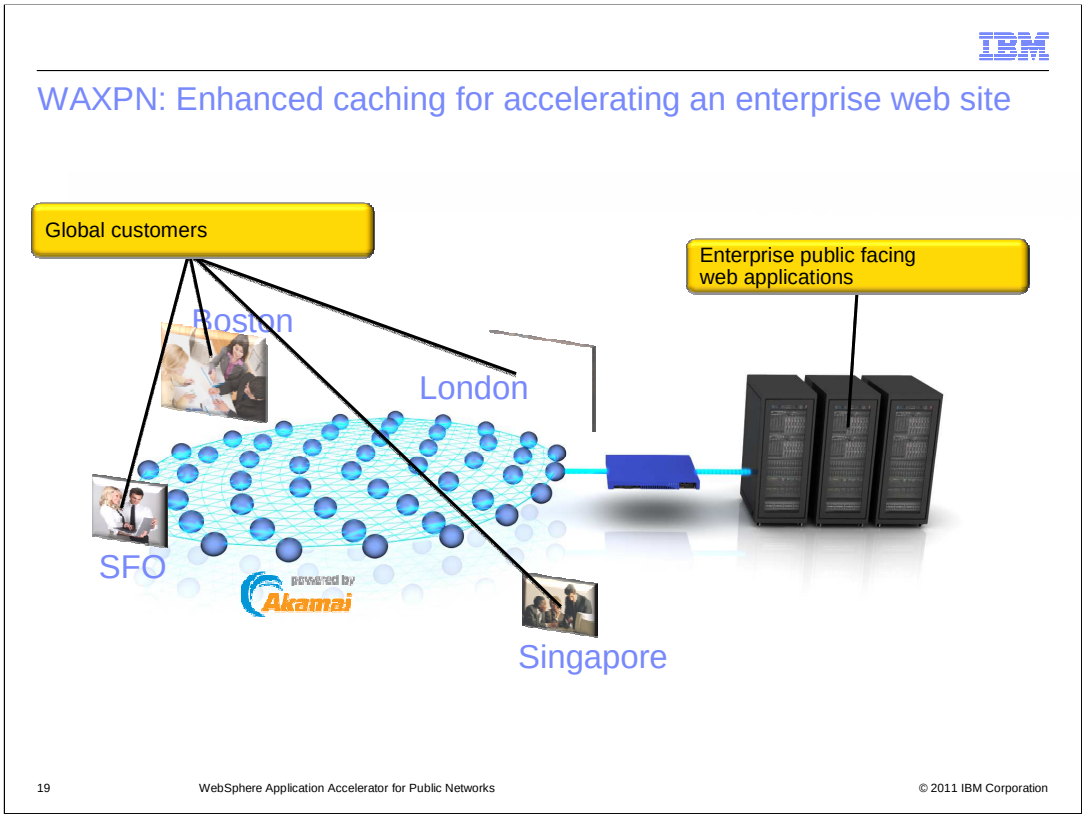


- Parses the response and primes the cache with parallel requests

18 WebSphere Application Accelerator for Public Networks

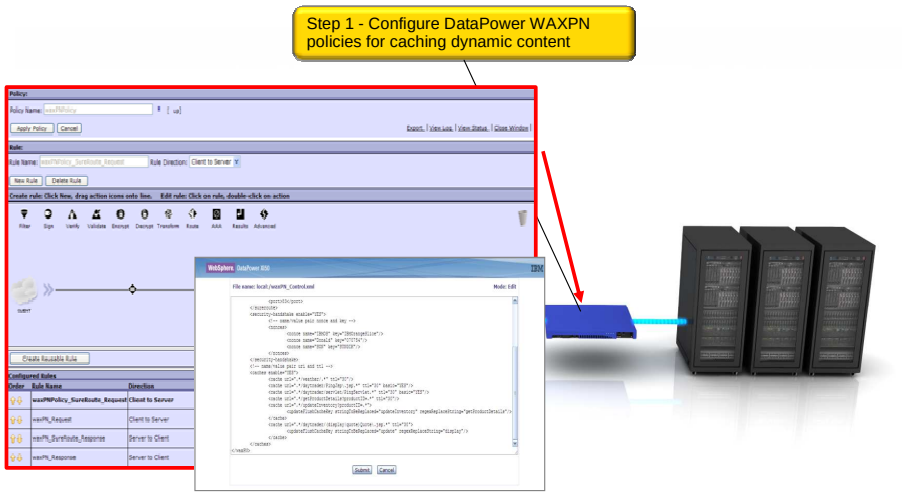
Prefetching parses web content at the closest point of presence to the end user and pre fetches objects for what will be requested in the near future. This primes the cache, increasing the cache hit count and further reduces the latency experienced by the end user.

For example, in this picture, the client uses a web browser to access a website. The request travels from the orange server closest to the user, and eventually reaches the web servers. The web servers return the HTML document. The orange server closest to the user parses the HTML and discovers there are four objects referenced by this HTML document. Therefore, at the same time the orange server sends the HTML document the web browser, the orange server requests each of the four objects from the web server. By the time the web browser parses the HTML document and requests the four objects from the web server, the orange server already has the response from the web server with these objects. Therefore, the orange server is able to serve these back to the web browser immediately, thereby reducing latency.



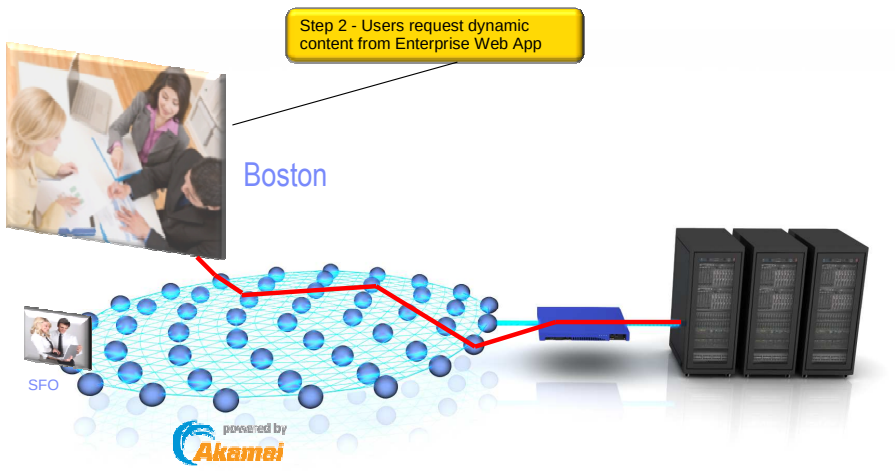
The WAXPN toolkit provides fine grain control for caching static content and also capabilities to cache dynamic content.

WAXPN: Caching dynamic content (1 of 5)



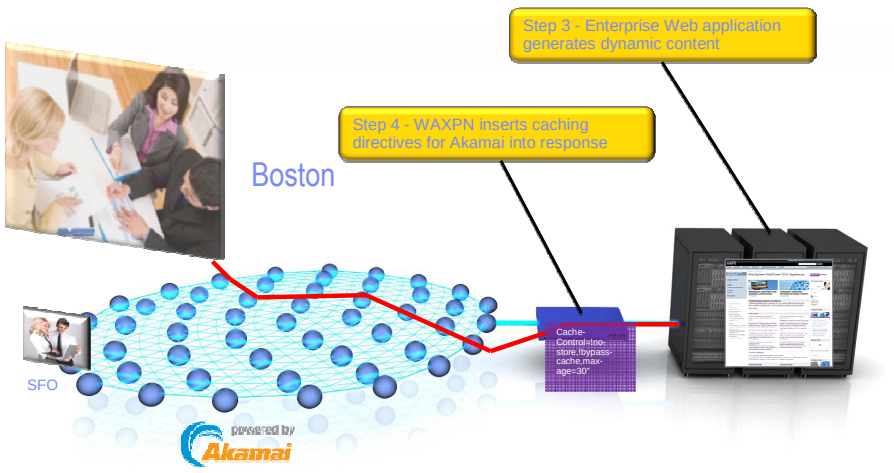
In this first step, a caching policy is created on DataPower.

WAXPN: Caching dynamic content (2 of 5)



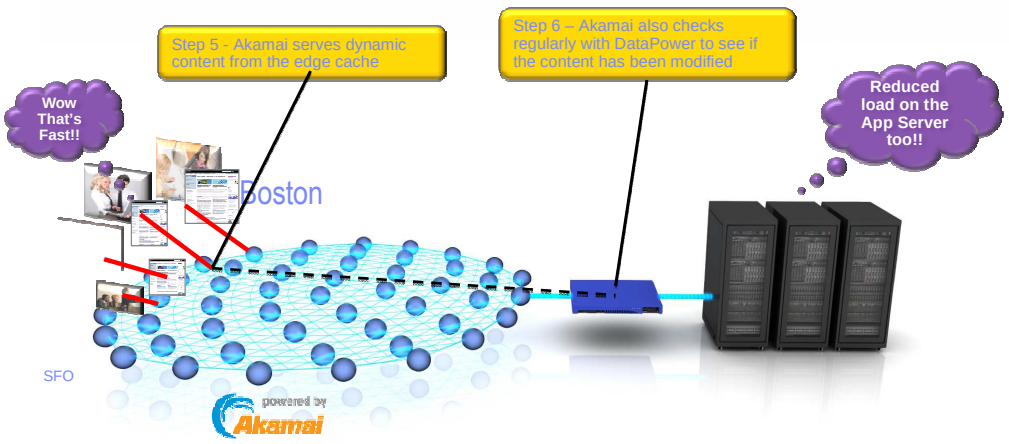
Once the caching policy has been deployed, user requests for specified content are processed as they flow through DataPower.

WAXPN: Caching dynamic content (3 of 5)



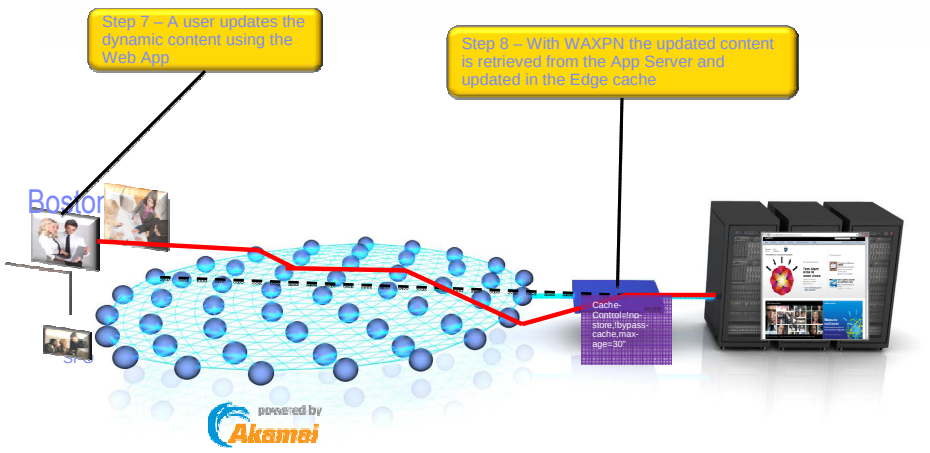
The response is generated by the backend application servers. The DataPower appliance adds special caching instructions to the response based on the policy specified.

WAXPN: Caching dynamic content (4 of 5)

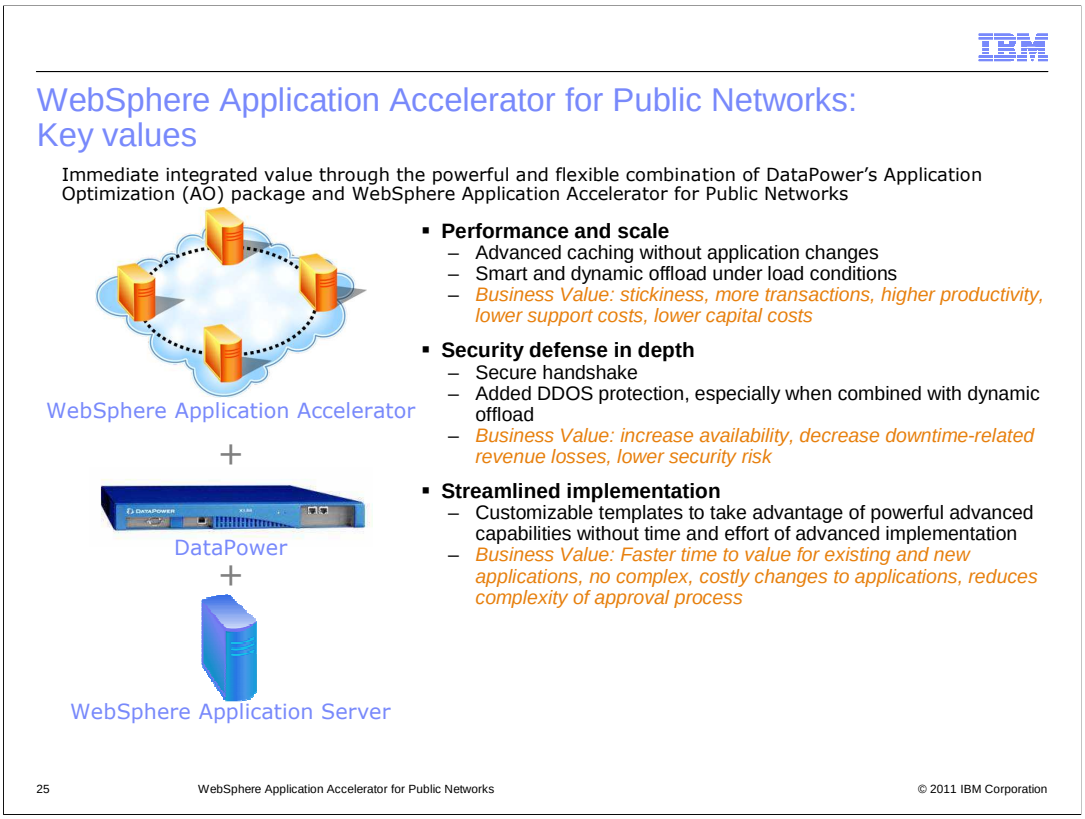


Akamai EdgeServers will then serve the content from cache as specified by DataPower in the response.

WAXPN: Caching dynamic content (5 of 5)



Dynamic content can be served from cache until it is modified through the web application. DataPower keeps track of HTTP PUT, DELETE and POST operations. Akamai EdgeServers check back regularly with DataPower to see if the dynamic content has been modified…. If not serving from cache continues. If so, the request is allowed to flow through DataPower to the application server to get the current version and the process starts again.



WebSphere Application Accelerator for Public Networks Key Values are:

Performance and Scale

Advanced caching without application changes

Smart and dynamic offload under load conditions

Business Value: stickiness, more transactions, higher productivity, lower support costs, lower capital costs

Security Defense in Depth

Secure handshake

Added DDOS protection, especially when combined with dynamic offload

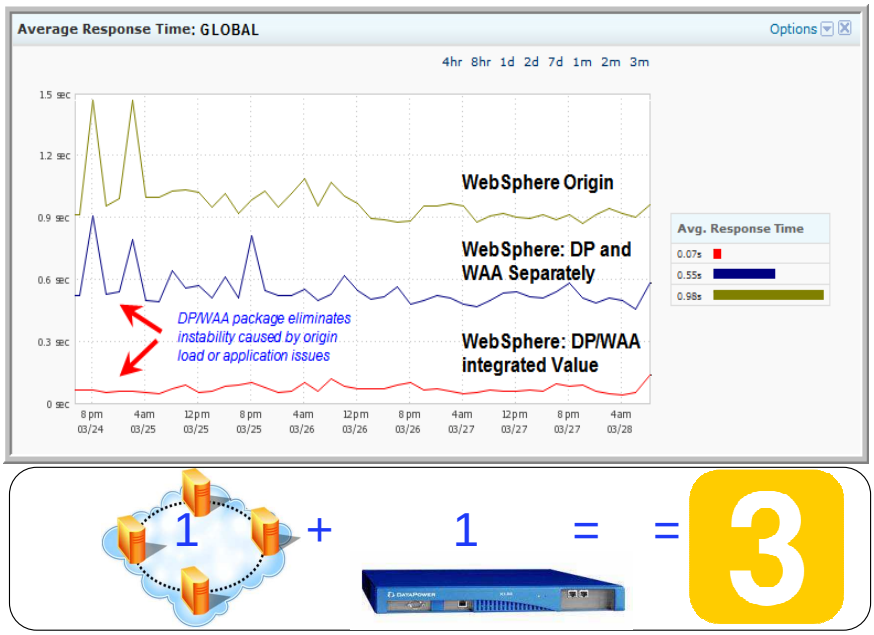
Business Value: increase availability, decrease downtime-related revenue losses, lower security risk

Streamlined Implementation

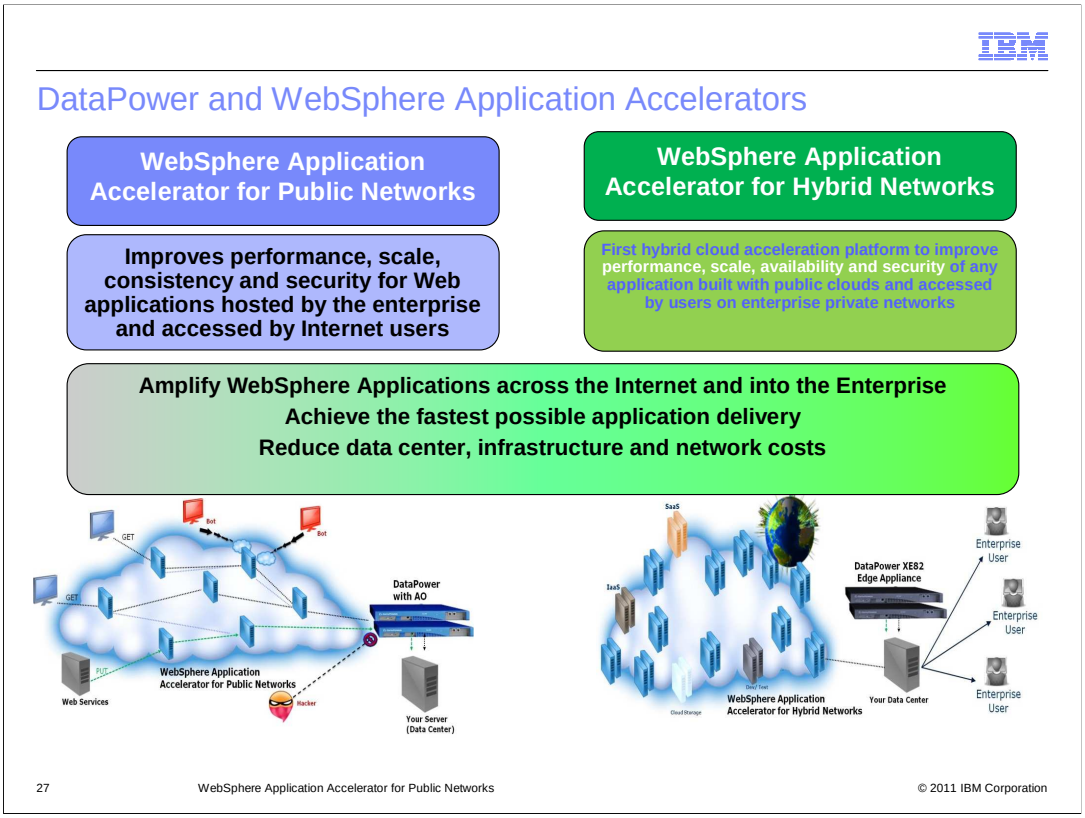
Customizable templates to take advantage of powerful advanced capabilities without time and effort of advanced implementation

Business Value: Faster time to value for existing and new applications, no complex, costly changes to applications, reduces complexity of approval process

The power of Akamai and DataPower



The slide shows the performance and response time consistency improvement possible with WAXPN. In this use case WAXPN provides significantly faster and consistent response.



DataPower and WebSphere Application Accelerators

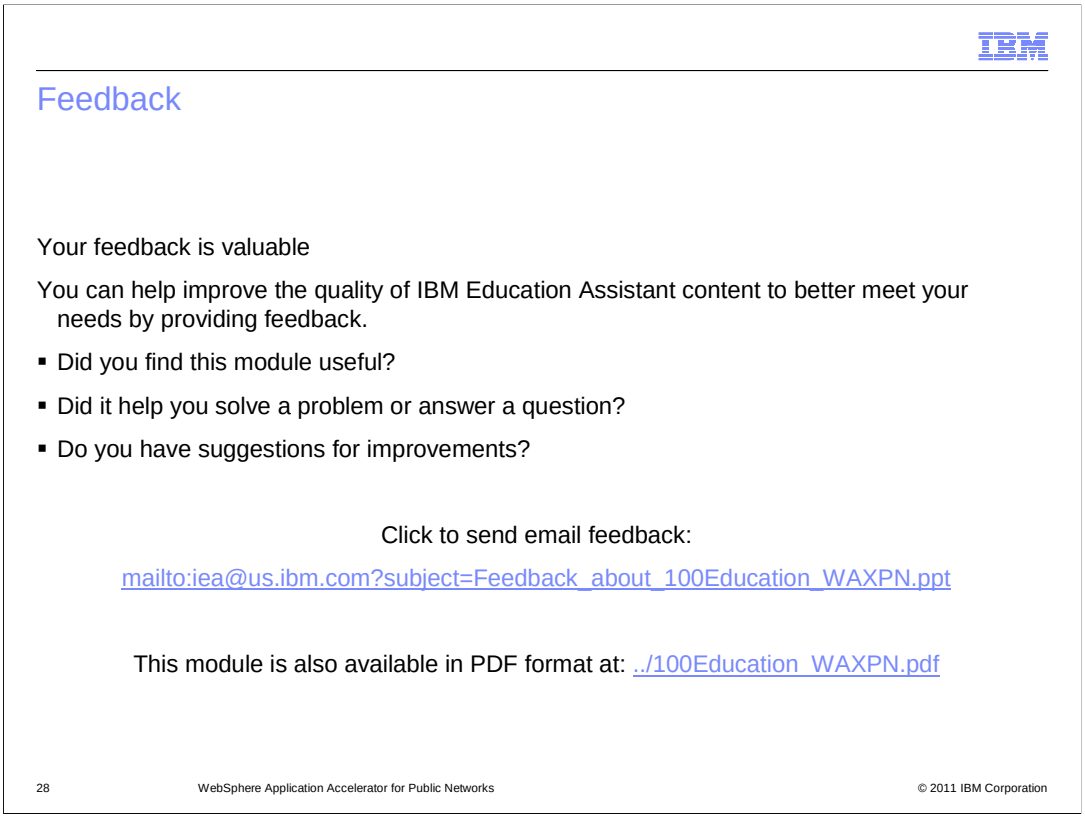
Help you achieve the fastest possible application delivery

Reduce data center, infrastructure and network costs

Bring benefits of Akamai EdgePlatform inside the firewall

Speed access by enterprise users to applications hosted in the public cloud

Software integrated within DataPower appliance extends Akamai's value to within the enterprise data-center and private network while bringing a new level of security thru authentication, authorization and accounting to govern application access for cloud-built applications



You can help improve the quality of IBM Education Assistant content by providing feedback.

Trademarks, disclaimer, and copyright information