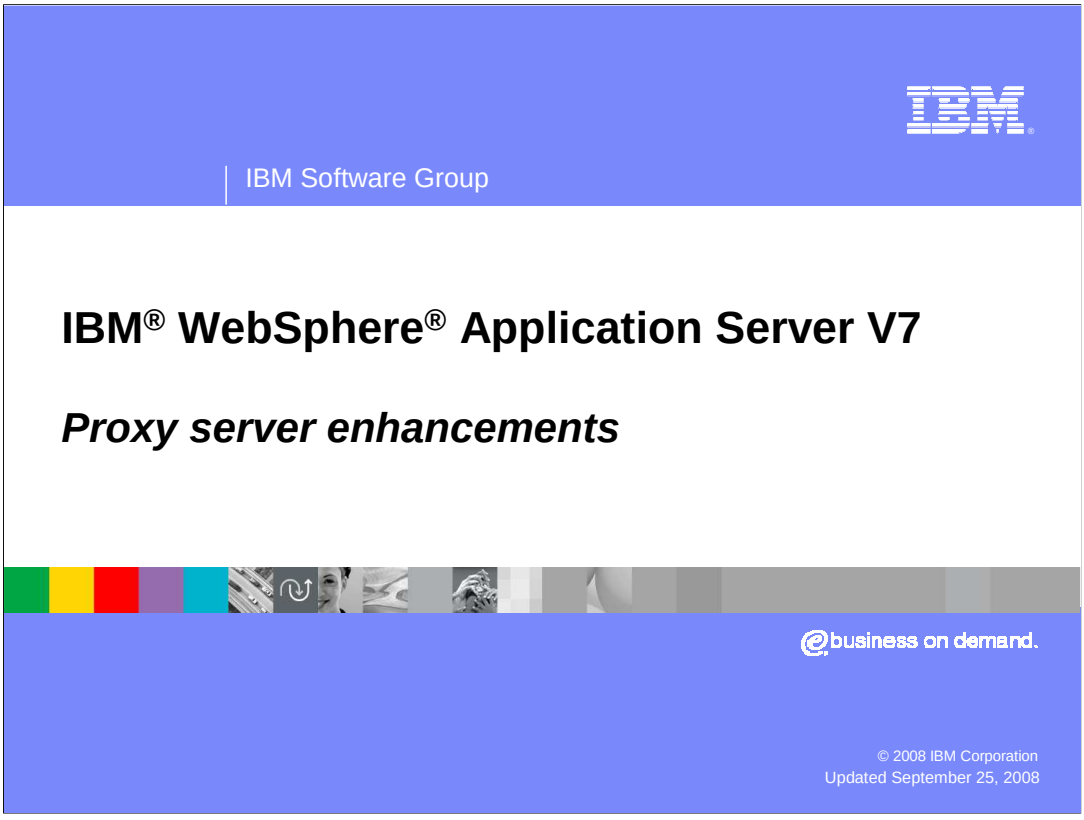


This presentation discusses enhancements to the proxy server and with a strong focus on the secure proxy server as found in IBM WebSphere Application Server V7.

Agenda

- Present proxy server
- Enhancements in V7
- Problem determination
- Summary

© 2008 IBM Corporation

This presentation starts a quick look at the earlier proxy server as a basis to understanding the secure proxy server. This presentation focuses on the secure proxy server with a mention of only a few enhancements to the existing proxy server. It is important to note that there are two proxy servers and they are different: the proxy server and the secure proxy server. Next are enhancements to the proxy server to create the secure proxy server, and a look at a few enhancements to the proxy server itself. You will finish with useful trace strings and finally a summary.

Section

Present proxy server

© 2008 IBM Corporation

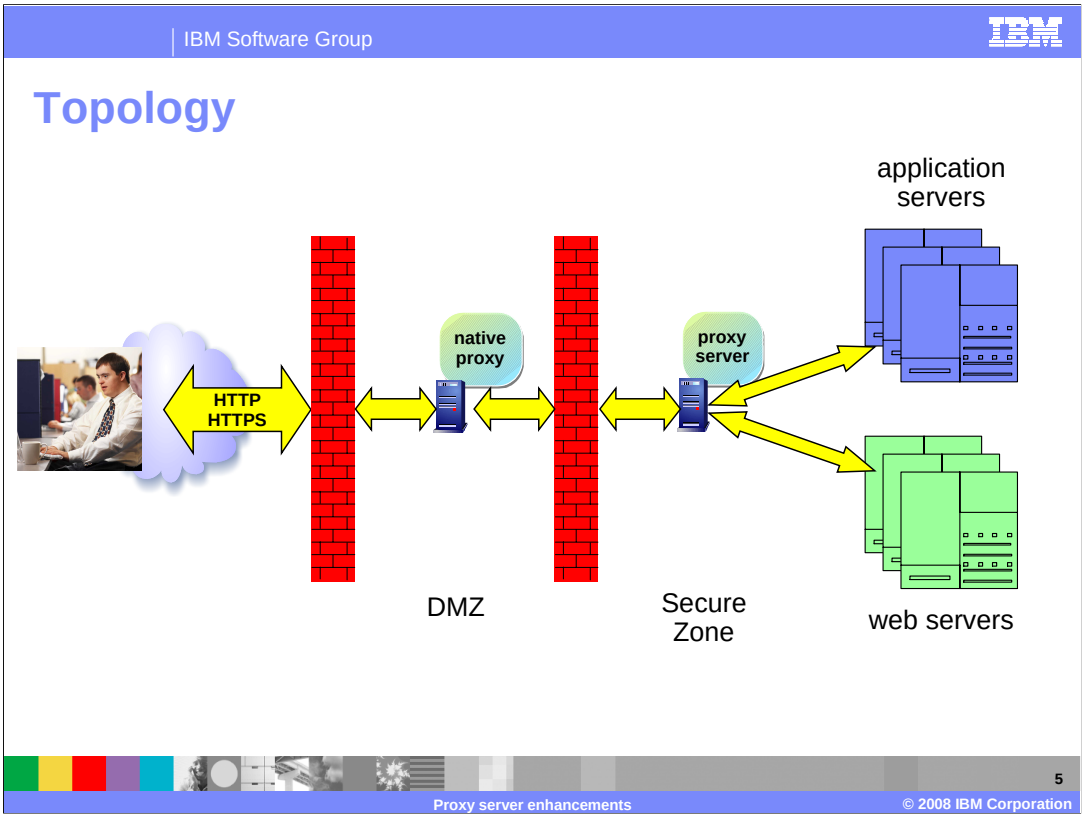
The following section describes the proxy server delivered with WebSphere V6.

Present proxy server

- Java base
  - Leverage IBM core competencies
  - Remove buffer overrun security exposure
  - Multiplexed non-blocking network traffic
  - Leverage WebSphere Application Server infrastructure
- Product consolidation
  - Focused on a single code base

© 2008 IBM Corporation

IBM has years of experience in creating Java based products that are highly efficient . Programming the proxy server in Java avoids several of the pitfalls found in the C language such as a buffer overflow security exposure caused by programming errors in manipulating pointers. Java has some inherent performance advantages such as its ability to scale to the number of clients and resources available. Java has efficient support for multiplexed non-blocking network traffic. As an extension to the WebSphere family the proxy server interface cleanly and efficiently to a WebSphere application server. To reduce footprint, the Proxy server has been architected to not have any dependencies on the basic Java EE containers provided by the Application Server. It does however use services such as Dynamic Caching for caching and high availability manager for On Demand configuration. The proxy server has become the base for IBM edge servers.



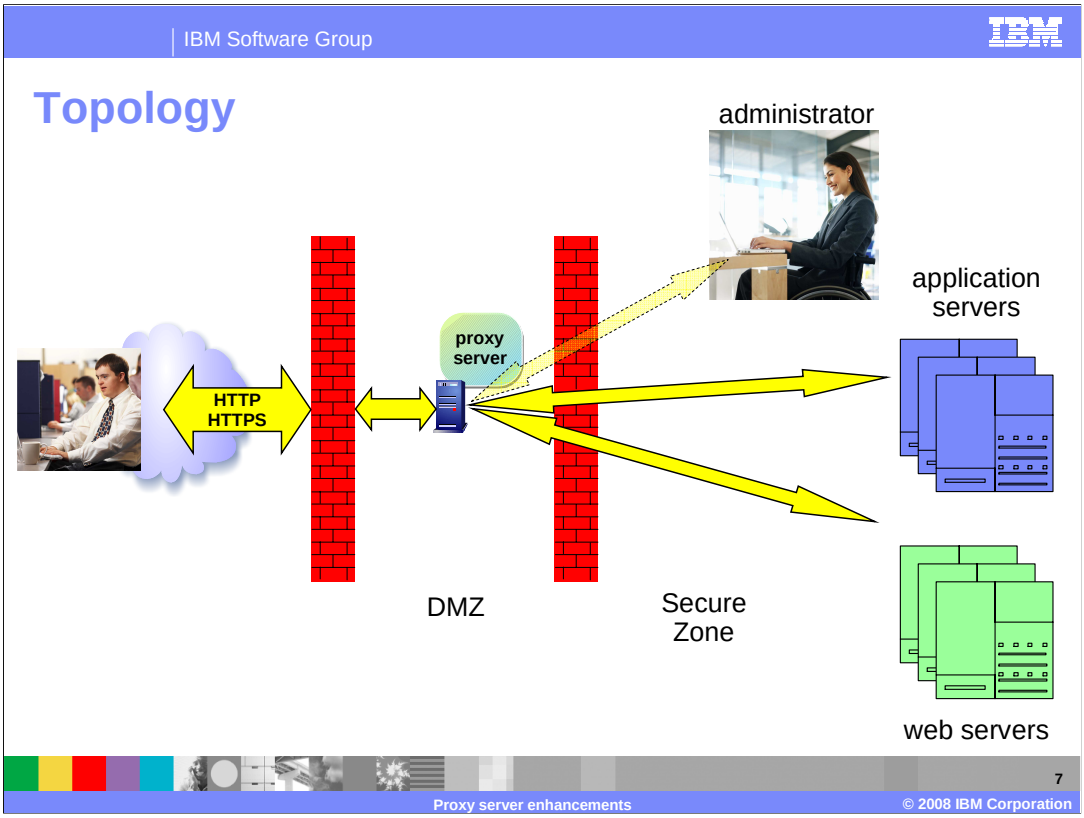
This is a typical topology using a proxy server before WebSphere V7. In many deployments, the proxy server was not suitable to place in the DMZ because of security issues. So a secure native proxy was placed in the DMZ and the proxy server was placed in the secure zone. While the proxy server still had many advantages as just mentioned this configuration requires an extra network hop. The secure proxy server removes the need to place the proxy server in the secure zone as shown in the next section.

Section

Enhancements in V7

© 2008 IBM Corporation

The following section describes the secure proxy server.



This is a typical topology with the secure proxy server. Notice that the proxy server is now in the DMZ. Six changes were made to a classic proxy server to make the secure proxy server secure.

1) The ability to compile software on a machine in the DMZ is a hacker's toolkit and so a JRE replaces the JDK.

2) How the machine is administered affects how secure the machine is since it is exposed to the world through a not as protective firewall. A shown here the administrator is considered a secure person but has restricted access to administer the proxy server from a remote system inside the secure zone. If this path is not secure, then a hacker might gain administrative access to the proxy server.

3) Routing tables are also used by hackers and static routing tables are very secure. The proxy server's dynamic routing, while it can be secure, must be one-way and tunneled over HTTP.

4) The proxy server will need to start with privileged access on many platforms to bind to ports below 1024, like ports 80 and 443. However, it will need to switch to an unprivileged user such that if an exploit succeeds and the process is taken over, he does not have any privileges on the system. This user is commonly defined as user "nobody" on many UNIX variants.

5) To avoid having to query a remote system for errors and rely on security through that remote system, the error pages should be statically defined and served.

6) The proxy server does not contain unnecessary components such as the Web container. A special installer is used to create the proxy server.

Secure proxy server security levels

- Security settings have four components
- Can select from three preset levels
- Can set any of the four components individually

	High	Medium	Low
Administration	Local SOAP	Local SOAP	Remote SOAP
Routing	Static	Dynamic	Dynamic
Start up permissions	Unprivileged	Unprivileged	Privileged
Error page handling	Local	Local	Local

© 2008 IBM Corporation

As this table shows there are four security components that can be configured. They are administration, routing, user level, and error page handling. They can be set individually or you can choose from one of the three preconfigured security levels: high, medium, and low.

If you select the high security level, the proxy server can only be administered by using local SOAP; that is, connecting to proxy server's SOAP port from the same host as the proxy server. Routing can only de defined with preset, static tables. After a proxy server connects to the required low numbered ports, the proxy server changes to an unprivileged user ID. And finally, error pages must be defined on the same host as the proxy server.

If you select the lowest security level then the proxy server can be administered from a remote host. The routing tables are dynamic as before. The proxy server continues to run as a privileged user. Even for Low security level, error pages are generated on the local host. To set error pages to remote host, you must select "custom" as the security level.

Note that the term remote SOAP can refer to running an administrative agent from another host inside the DMZ and outside.

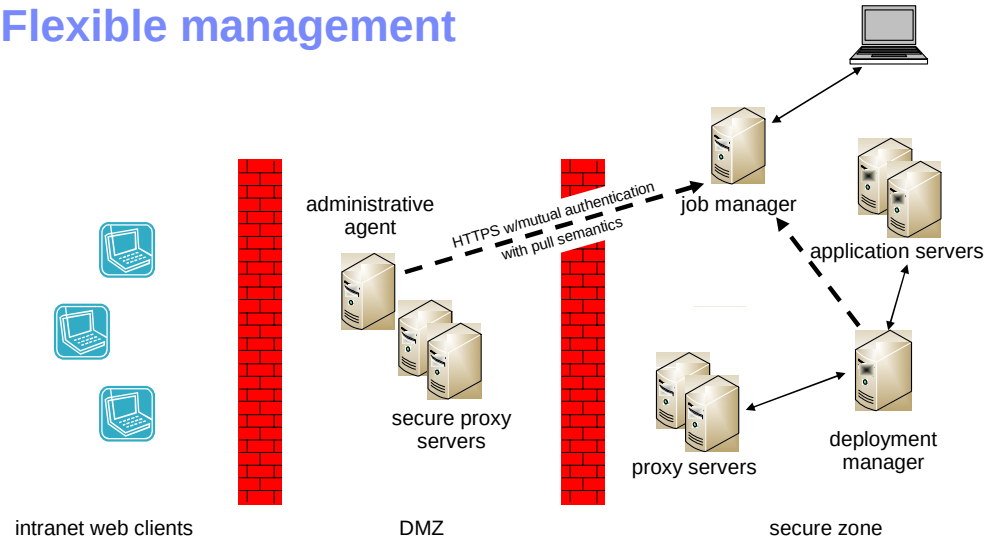
Secure proxy server administration

- Local SOAP (higher security level)
  - Command line interface
  - Administrative agent with FTP or command line interface
  - Flexible management with administrative agent
- Remote SOAP (lower security level)
  - Command line interface
  - Administrative agent
  - Flexible management with administrative agent

© 2008 IBM Corporation

How the proxy server can administered depends on the security level. At the lower level or remote SOAP, you can administer the proxy server from the secure zone depending on the setting of the inner fire wall or from the DMZ. The tools available are the command line interface (scripting), the administrative agent, and flexible management using the administrative agent. At the higher security level the option available to you are more restricted. You can still use the command line interface from a terminal connected to the same host that the proxy server is running on. You can use an administrative agent from inside the secure zone but instead of directly managing the secure proxy server you have to generate a command file. To complete the operation transfer the command file by some mechanism to the host running the secure proxy server and use the command line interface to perform the actions. This process is shown later in this presentation. If you place an administrative agent on the same host as the secure proxy server, you can administer the proxy server with flexible management. This later is not as secure but you should note that the administrative agent does use *pull* semantics. Hence a hacker can not “send” commands to the administrative agent to then be enforced on the secure proxy server. One final note is that you use the administrative agent to create or manage a secure proxy and not the administrative client: the classical proxy and the secure proxy are different.

Flexible management



© 2008 IBM Corporation

This slide shows secure proxy servers administered with flexible management from the secure zone. Again notice the administrative agent uses pull semantics. This configuration can be used to manage a large number of DMZ Secure Proxy Servers.

Consolidation

- Base for additional strategic proxy function in IBM
  - WebSphere plug-in
  - WebTrafic express
  - EdgeSEAL
  - WebSEAL
  - Branch proxy
  - Virtual Enterprise
  - Web services gateway
  - Session initiation protocol stateless proxy

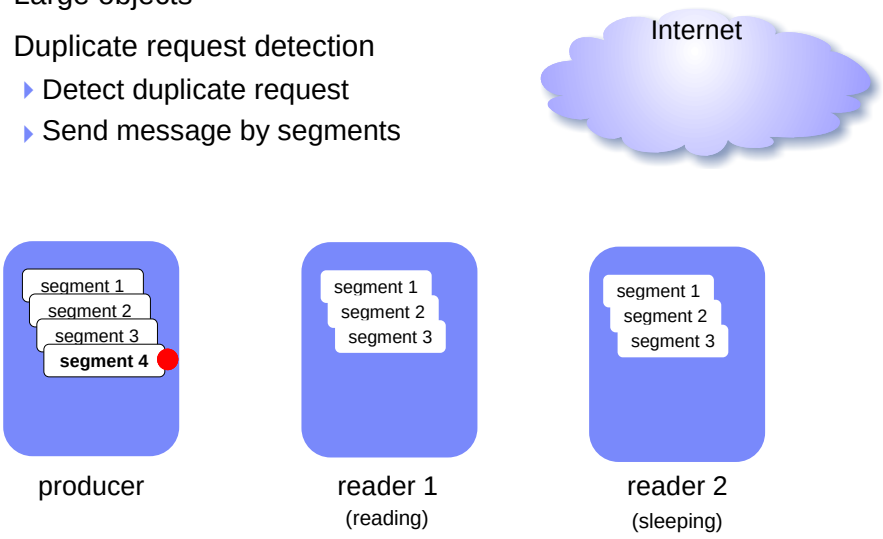
© 2008 IBM Corporation

The proxy server consolidates the functions of the WebSphere plug-ins and Web traffic express. The functionality of Tivoli EdgeSeal is available in the proxy server through a plug-in.

The proxy server is used as the base for additional strategic proxy function in IBM, such as branch proxy and Tivoli WebSEAL. The proxy server is the base for the on demand router for Virtual Enterprise and provides base edge services for Web services gateway and session initiation protocol (SIP) stateless proxy.

Caching

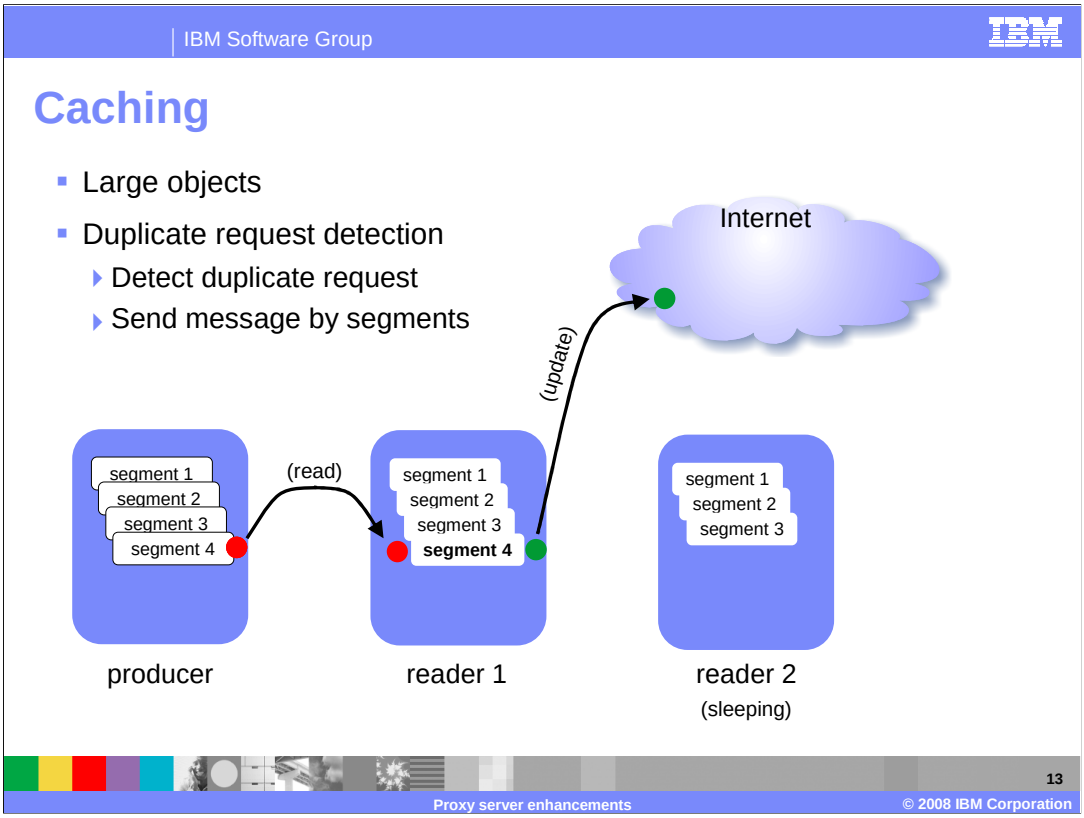
- Large objects
- Duplicate request detection
  - Detect duplicate request
  - Send message by segments



© 2008 IBM Corporation

In order to cache large files, the proxy needs a way to break the files into manageable chucks called segments. The first segment is stored in the cache with its key and subsequent segments are stored in the cache with a variation of the key, References to subsequent chunks are stored with the first segment. This way the entire file can be referenced or flushed with a single key.

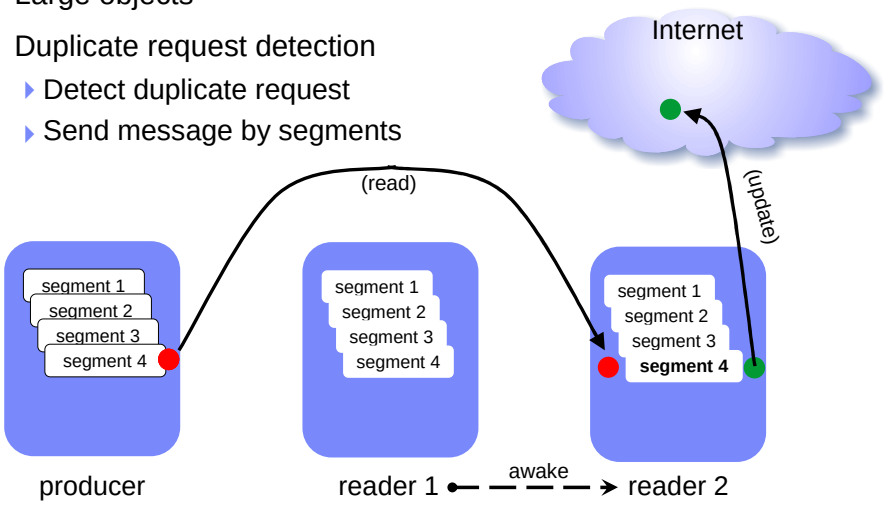
If the proxy server expects to be able to cache the incoming request from the client, the nominal behavior for subsequent duplicate requests is pause the duplicate request. The duplicate request is resumed when the cached entry is created. The duplicate request then sends the response, instead of pushing the request to the backend server. The unfortunate side effect is that if, for whatever reason, the item is not able to be cached then the duplicate request has paused unnecessarily. The duplicate request detection allows a reader to read a segment at a time and push it out to a client, and be updated when the entry is cached. In order to do this, they read a cache key that is attached to the main entry and be waiting for updates. This slide shows the beginning of the life cycle of segment 4. Here the producer has just retrieved segment four from the backend. At the same time reader one is waiting for segment four and reader two is sleeping.



Reader one now reads segment four from the producer and updates its client. Reader two is still sleeping.

Caching

- Large objects
- Duplicate request detection
  - Detect duplicate request
  - Send message by segments



© 2008 IBM Corporation

The next reader is then awakened and reads segment 4. In this case, this is reader two which retrieves the segment and updat4s his client. Since this is the last reader, the system returns to the state in the original picture and waits for segment five if there is one.

Workload management

- Random load balancing algorithm added
- Static routing
  - For secure proxy servers
- Custom advisors
  - Customer generated callback by framework
  - Returned value of false marks a server down
- Time of day routing
  - Applies to generic servers and application servers
  - Can be used to route around servers

© 2008 IBM Corporation

The Random load balancing algorithm will load balance the members of the target cluster in a random fashion. If disabled, the members of the target cluster are load balanced in the default “weighted round robin fashion”. This option is supports the migration from the WebSphere plug-in to the proxy server.

Static routing supports the proxy server in the DMZ. Intelligent application availability is not available with static routing. The code is not aware when servers and applications are stopped in the cluster so it is possible for cluster members to be selected that are down or not running the application the cluster is being sent to. This will cause requests that go to those servers down or not running the application to fail and have to be retried to another cluster member where the application is running.

Custom advisors allow you more intelligent determination of target application server availability. Custom Advisors are code modules you write to implement an interface that is called by the Custom Advisor Framework. When this method is called by the Custom Advisor Framework, it is expected to connect to the server, send a request, receive a response, and determine if the response is good or bad. If the result of this method is false, the workload manager marks the server down, until it is marked up again on a subsequent call to your server. This provides a proactive measure to determine if applications are down at the application protocol level by being able to interpret the actual application protocol response message.

A Time of Day action is a rule based on a time of day range to specify a list of cluster member to use for selection. You can use this feature for capacity planning reasons or when you want to take some of the servers down for maintenance at a designated time.

Workload management (continued)

- Affinity for generic servers
  - Active affinity: allows the proxy server to manage the affinity of a non Java-EE or non-WebSphere application
  - Passive affinity: allows the proxy server to use session cookies set by non-Java EE or non-WebSphere applications
    - Limited number of cookies
    - Limited number of cookies per domain
    - Limited cookie size

© 2008 IBM Corporation

Active affinity is when the proxy server will manage the affinity for the target generic server. Use active affinity when the target generic server has no means to manage it's own affinity. The proxy server will set the cookie to maintain affinity and assure the request gets back to the correct server during the configured expiration time.

Passive Affinity is when the proxy server uses a cookie set by the target generic server to maintain affinity. You specify the cookie name and map the cookie values to each generic server cluster member.

If an ActiveAffinity or PassiveAffinity cookie routing action has been defined, that cookie exists and the affinity sever is available, the workload management selection code is preempted. In this case the server that has affinity established is selected.

Response URL rewriting

- Allows transformation of HTML and CSS tags in the response before sending response to the client
- Used to protect the identity of the content servers from the Web clients by rewriting the URLs
- Clients only interact with the public proxies

© 2008 IBM Corporation

In WebSphere Application Server V7, URL rewriting of absolute and relative URL links in an HTTP response body is supported. This applies to HTML and CSS content. Also, rewriting of a Set-Cookie domain and path attributes is possible. Furthermore, rewriting of the location header is possible. By default, URL rewriting is off but can be enabled by creating a URL rewriting action under a virtual host and associating it with a Rule Expression.

URL rewriting on the response body has to modes: active and passive. By default active is turned on. Active means to scan the response body and rewrite the URLs before sending the body data to the client. Passive means to store the rewrite prefix in a Set-Cookie header and not perform the rewrite before sending the body data to the client. The rewrite will occur on the subsequent request. The proxy will append the prefix to the received URI before sending the request to the server.

Section

Problem determination

© 2008 IBM Corporation

The next section presents the trace strings to help solve a problem with a secure proxy server.

Trace strings

- WebSphere Proxy=all
- com.ibm.ws.dwlm.*=all
- GenericBNF=all
- com.ibm.ws.http.*=all
- com.ibm.ws.odc.*=all
- WLM*=all

© 2008 IBM Corporation

These are standard proxy flags. Note that BNF refers to TCP headers and odc refers to dynamic configuration.

Section

Summary

© 2008 IBM Corporation

The following slide summarizes the secure proxy server.

Summary

- The proxy server:
  - Can be securely deployed to a DMZ
  - Provides a base for other proxy services
  - Supports new features in WebSphere Application Server V7, like flexible management
  - Includes new caching enhancements
  - Supports workload management

© 2008 IBM Corporation

This presentation has discussed the secure proxy server. The secure proxy server can be securely deployed to a DMZ. The secure proxy server is a base for other proxy servers such as Tivoli WebSEAL, and Branch Proxy. The secure proxy server supports WebSphere V7 Flexible management.

General improvements to all WebSphere V7 proxy servers include caching large objects, duplicate request optimization, Increased routing choices and active/passive affinity for generic servers.

Feedback

Your feedback is valuable

You can help improve the quality of IBM Education Assistant content to better meet your needs by providing feedback.

- Did you find this module useful?
- Did it help you solve a problem or answer a question?
- Do you have suggestions for improvements?

Click to send e-mail feedback:

mailto:iea@us.ibm.com?subject=Feedback_about_WASv7_SecureProxy.ppt

This module is also available in PDF format at: ../WASv7_SecureProxy.pdf

© 2008 IBM Corporation

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

Trademarks, copyrights, and disclaimers

The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

IBM WebSphere

A current list of other IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements or changes in the products or programs described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectual property rights, may be used instead.

Information is provided "AS IS" without warranty of any kind. THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (for example, IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products.

IBM makes no representations or warranties, express or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

© Copyright International Business Machines Corporation 2008. All rights reserved.

Note to U.S. Government Users - Documentation related to restricted rights-Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract and IBM Corp.