WebSphere Application Server Version 8 New z/OS Exploitation/Differentiation

David Follis
IBM

August 11, 2011
Session Number 9485

WebSphere Application Server Sessions

Day	Time	#	Title	Speaker	Room
Wednesday	3:00	9483	Using IBM's New Cross-Platform Installer on z/OS	Mierzejewski	Oceanic 5
Thursday	8:00	9482	WAS Version 8 – Overview	Follis	Europe 2
Thursday	9:30	9486	WAS Version 8 – Batch Update	Hutchinson	Europe 2
Thursday	11:00	9485	WAS Version 8 – New z/OS Exploitation/Differentiation Features	Follis	Europe 2
Thursday	1:30	9484	WAS Version 8 – High Availability Enhancements	Follis	Europe 2
Thursday	3:00	9488	WAS - Back to Basics Part 1	Loos	Europe 2
Thursday	4:30	9489	WAS - Back to Basics Part 2	Stephen	Europe 2
Friday	8:00	9490	WAS for z/OS - Level 2 Update	Stephen	Europe 2
Friday	9:30	9487	WAS for z/OS – PotPourri	Follis, Hutchinson, Loos, Mierzejewski, Stephen, etc.	Europe 2

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

CICS*
DB2*
GDPS*
Geographically Dispersed Parallel Sysplex
HiperSockets
IBM*
IBM eServer
IBM logo*
IMS
On Demand Business logo
Parallel Sysplex*
RACF*
System z9
WebSphere*
z/OS
zSeries*

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput 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 improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Disclaimer

- The information contained in this documentation is provided for informational purposes only. While efforts were many to verify the completeness and accuracy of the information contained in this document, it is provided “as is” without warranty of any kind, express or implied.
- This information is based on IBM’s current product plans and strategy, which are subject to change without notice. IBM will not be responsible for any damages arising out of the use of, or otherwise related to, this documentation or any other documentation.
- Nothing contained in this documentation is intended to, nor shall have the effect of , creating any warranties or representations from IBM (or its suppliers or licensors), or altering the terms and conditions of the applicable license agreement governing the use of the IBM software.
- Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput 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 improvements equivalent to the performance ratios stated here.
- All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

Granular RAS Controls

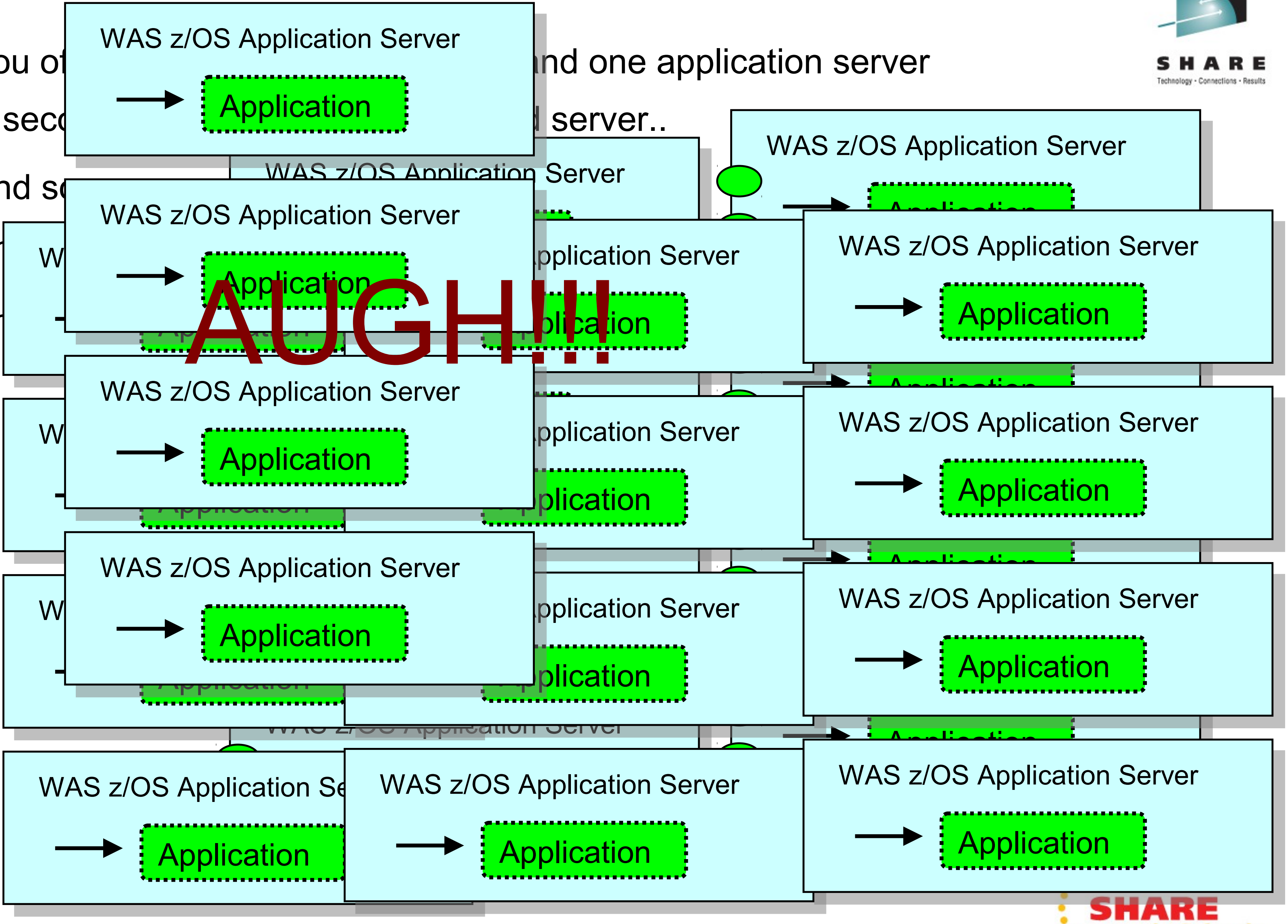
You of ... nd one application server

A seco ... server..

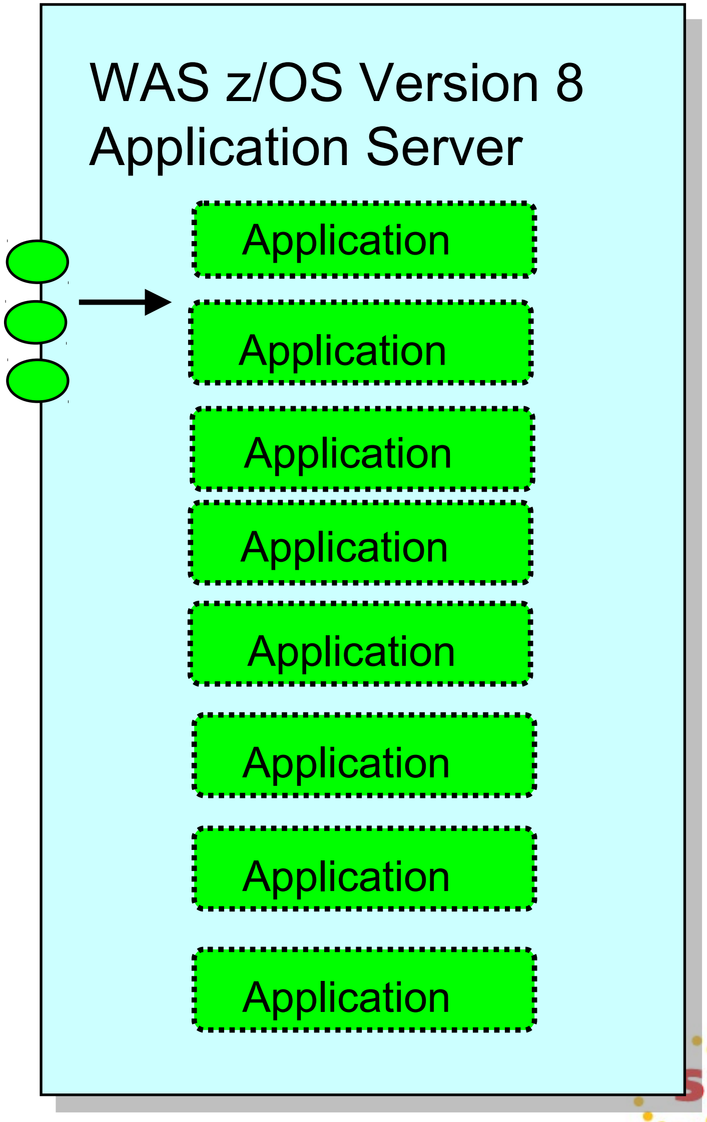
And s

Ar

Ur

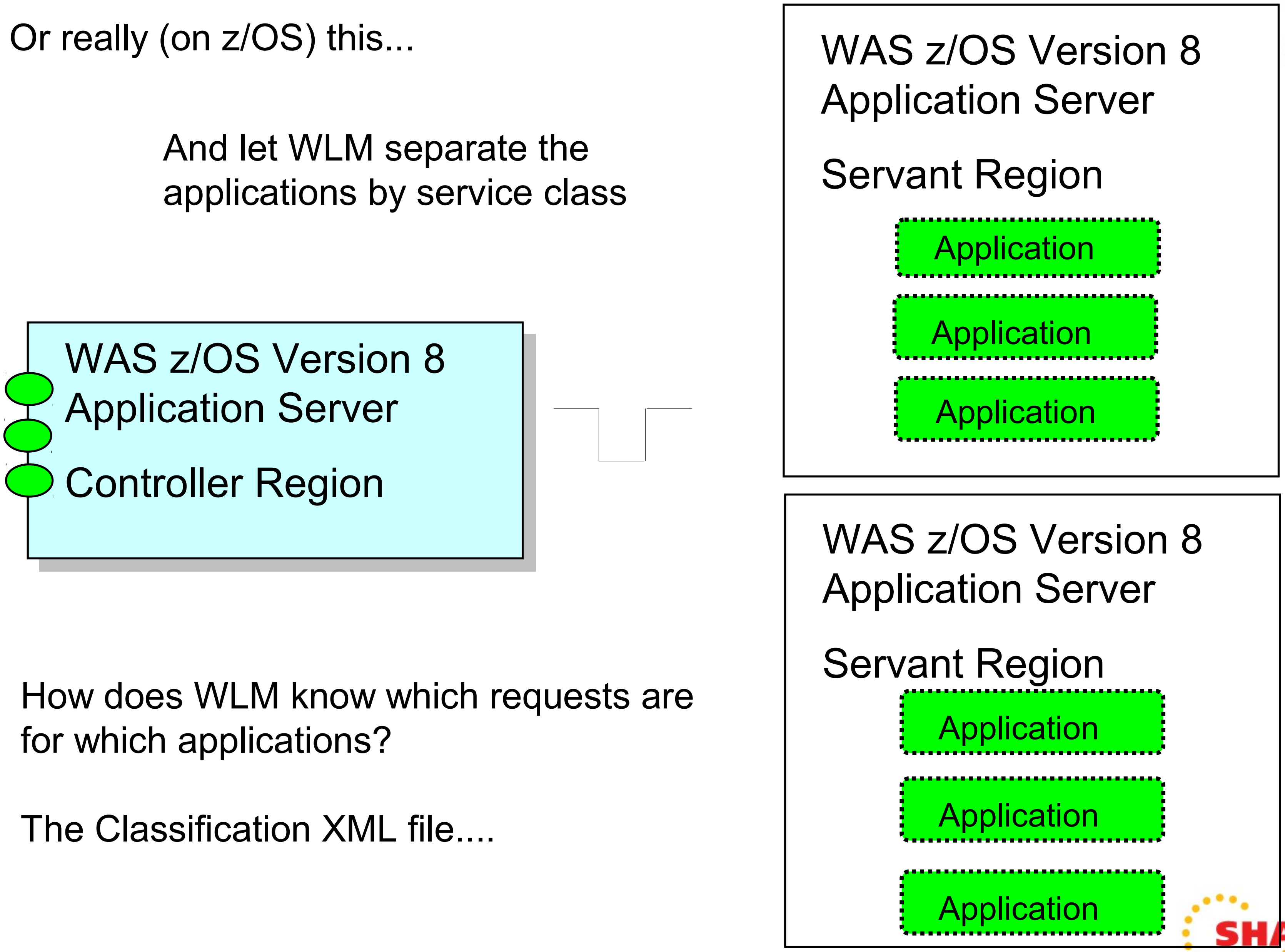


Why can't we just do this?



Or really (on z/OS) this...

And let WLM separate the applications by service class



WAS z/OS Version 8 Application Server

Controller Region

WAS z/OS Version 8 Application Server

Servant Region

WAS z/OS Version 8 Application Server

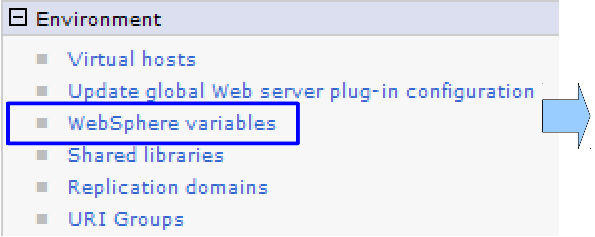
Servant Region

How does WLM know which requests are for which applications?

The Classification XML file....

How it Works

The file supplies a set of criteria to match requests to transaction class names, which then match with rules in the CB subsystem type

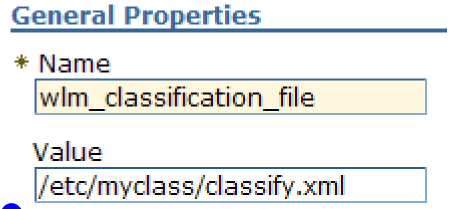


⊟ Environment

- Virtual hosts
- Update global Web server plug-in configuration
- WebSphere variables
- Shared libraries
- Replication domains
- URI Groups

Scope to cell or node

server scope for classification deprecated



General Properties

* Name

wlm_classification_file

Value

/etc/myclass/classify.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE Classification SYSTEM "Classification.dtd"
>
<Classification schema_version="1.0">
  :
   <InboundClassification type="iiop" ... >
     (classification information)
   </InboundClassification>
   <InboundClassification type="http" ... >
     (classification information)
   </InboundClassification>
   <InboundClassification type="sip" ... >
     (classification information)
   </InboundClassification>
   <InboundClassification type="mdb" ... >
     (classification information)
   </InboundClassification>
   <InboundClassification type="sib" ... >
     (classification information)
   </InboundClassification>
  :
</Classification>
```

Transaction Name based on request match → Rules in CB subsystem type → Service Class Reporting Class

From that we get goals and importance based on specific transactions based on criteria in the classification XML file

Another big issue with multiple applications in one server....

TIMEOUTS

Timeout configuration is at the server level, per protocol

For example, HTTP has these variables:

```
protocol_http_timeout_output
protocol_http_queue_timeout_percent
server_region_http_stalled_thread_dump_action
```

And for all protocols:

```
server_region_request_cputimeused_limit
server_region_request_cputimeused_dump_action
```

Could we put some of this stuff in the classification XML?

YES!

But before we get started.. here's a simple XML file, just the HTTP section:

```
<InboundClassification type="http" schema_version="1.0" default_transaction_class="M">
   <http_classification_info transaction_class="M">
        <http_classification_info transaction_class="Q" uri="/gcs/admin"/>
        <http_classification_info transaction_class=“R” uri=“/gcs/admin/1*”/>
   </http_classification_info>
</InboundClassification>
```

Adding more tags is going to make this hard(er) to read, especially on a chart

SO......

Examples from here on are NOT COMPLETE and WILL NOT PARSE

One more thing....   attribute inheritance

Suppose we have this XML fragment:

```
<InboundClassification type="http"    . . .>
   <http_classification_info  . . .>
        <http_classification_info uri="/gcs/admin" . . . />
        <http_classification_info uri="/gcs/admin/1*” . . ./>
   </http_classification_info>
</InboundClassification>
```

The outer 'http_classification_info' does NOT specify any classification values
Host/Port/URI

So it matches ALL HTTP requests

The inner rules match a specific URI or one with a wildcard

Now what if we add some dispatch timeout tags?

```
<InboundClassification type="http"    . . .>
   <http_classification_info  dispatch_timeout=”300”. . .>
      <http_classification_info uri="/gcs/admin"
dispatch_timeout=”500”.
      <http_classification_info uri=“/gcs/admin/1*” . . ./>
   </http_classification_info>
</InboundClassification>
```

A request which matches /gcs/admin will get a timeout of 500 seconds

All other HTTP requests get 300 seconds.

Got it?

Now...finally.. we will go through the new tags in the XML file

There are three questions for each tag:

1) What is the tag value?

2) How does the tag interact with existing environment variables?

3) Are there Modify commands that interact with these variables?

Dispatch Timeout

1) What is the tag value?

    dispatch_timeout

2) How does the tag interact with existing environment variables?

   The environment variables are used as default values if no value is found in the XML file. Variables affected are:

    control_region_wlm_dispatch_timeout
    protocol_http_timeout_output
    protocol_https_timeout_output
    protocol_sip_timeout_output
    protocol_sips_timeout_output
    control_region_mdb_request_timeout

3) Are there Modify commands that interact with these variables?

    No.

Timeout on the queue

1) What is the tag value?

    queue_timeout_percent

2) How does the tag interact with existing environment variables?

    The environment variables are used as default values if no value is found in the XML file. Variables affected are:

    control_region_iiop_queue_timeout_percent
    protocol_http_queue_timeout_percent
    protocol_https_queue_timeout_percent
    protocol_sip_queue_timeout_percent
    protocol_sips_queue_timeout_percent
    control_region_mdb_queue_timeout_percent

3) Are there Modify commands that interact with these variables?

    No.

Timeout dump action

1) What is the tag value?

   stalled_thread_dump_action

2) How does the tag interact with existing environment variables?

   The environment variables are used as default values if no value is found in the XML file. Variables affected are:

   server_region_iiop_stalled_thread_dump_action
   server_region_http_stalled_thread_dump_action
   server_region_https_stalled_thread_dump_action
   server_region_sip_stalled_thread_dump_action
   server_region_sips_stalled_thread_dump_action
   server_region_mdb_stalled_thread_dump_action

3) Are there Modify commands that interact with these variables?

   No.

CPU Timeout and CPU Timeout Dump Action

1) What is the tag value?

   cputimeused_limit
   cputimeused_dump_action

2) How does the tag interact with existing environment variables?

   The environment variables are used as default values if no value is found in the XML file. Variables affected are:

   server_region_request_cputimeused_limit
   server_region_cputimeused_dump_action

3) Are there Modify commands that interact with these variables?

   No.

IIOP Outbound Request Timeout

1) What is the tag value?

   request_timeout

2) How does the tag interact with existing environment variables?

   Actually its a Java property in this case. The property is used as default values if no value is found in the XML file. The property is:

   com.ibm.CORBA.RequestTimeout

3) Are there Modify commands that interact with these variables?

   No.

Timeout Output Recovery (SESSION or SERVANT)

1) What is the tag value?

    timeout_recovery

    Only applies to HTTP/HTTPS/SIP/SIPS sections

2) How does the tag interact with existing environment variables?

    The environment variables are used as default values if no value is found in the XML file. Variables affected are:

    protocol_http_timeout_output_recovery
    protocol_https_timeout_output_recovery
    protocol_sip_timeout_output_recovery
    protocol_sips_timeout_output_recovery

3) Are there Modify commands that interact with these variables?

    No.

SMF Recording

1) What is the tag value?

   SMF_request_activity_enabled
   SMF_request_activity_CPU_detail
   SMF_request_activity_timestamps
   SMF_request_activity_security

2) How does the tag interact with existing environment variables?

   The environment variables are used as default values if no value is found in the XML file. Variables affected are:

   server_SMF_request_activity_enabled
   server_SMF_request_activity_CPU_detail
   server_SMF_request_activity_timestamps
   server_SMF_request_activity_security

3) Are there Modify commands that interact with these variables?

   YES!

The existing commands:

```
MODIFY server,SMF,REQUEST,ON | OFF
MODIFY server,SMF,REQUEST,CPU,ON | OFF
MODIFY server,SMF,REQUEST,TIMESTAMPS,ON | OFF
MODIFY server,SMF,REQUEST,SECURITY,ON | OFF
```

These now tell the server to IGNORE the XML contents and turn the indicated SMF recording ON or OFF

When you want to go back to honoring the XML contents

Use the new RESET option

```
MODIFY server,SMF,REQUEST,RESET
MODIFY server,SMF,REQUEST,CPU,RESET
MODIFY server,SMF,REQUEST,TIMESTAMPS,RESET
MODIFY server,SMF,REQUEST,SECURITY,RESET
```

Remember that MODIFY server,DISPLAY,SMF will show the current settings

Dispatch Progress Monitor (DPM)

1) What is the tag value?

   dpm_interval
   dpm_dump_action

2) How does the tag interact with existing environment variables?

   The DPM Interval is only set via Modify. If no value for the dump action is found in the XML, the default will be taken from this environment variable:

   server_region_dpm_dump_action

3) Are there Modify commands that interact with these variables?

   YES!

   And MODIFY server,DISPLAY,DPM will show the current settings

DPM Modify Command Options

MODIFY server,DPM,CLEAR_ALL

Override XML, DPM Action=None, All DPM Intervals to zero

MODIFY server,DPM,RESET

Honor the XML content

MODIFY server,DMP,DUMP_ACTION=NONE (or other values)

Ignore DPM dump actions in the XML

MODIFY server,DPM,DUMP_ACTION=RESET

Honor DPM dump actions in the XML

MODIFY server,DPM,HTTP=500 (or other protocol, other values)

Ignore DPM intervals in the XML for HTTP and use 500 instead

MODIFY server,DPM,HTTP=RESET

Honor HTTP DPM intervals in the XML

MODIFY server,DPM,INTERVAL=0 (or other values)

Ignore DPM intervals in the XML and set all intervals to zero

MODIFY server,DPM,INTERVAL=RESET

Honor all DPM intervals in the XML

All New Function: Message Tagging!

```
<InboundClassification type="http"    . . .>
   <http_classification_info uri=”/gcs/*” message_tag=”GCS”. . ./>
</InboundClassification>
```

The tag 'GCS' will show up in messages, traces, and printlns issued by a thread dispatching any request that matches /gcs/*

For example:

```
Trace: 2011/03/21 22:15:48.298 02 t=6BEE88 c=0.6 key=S2 tag=GCS (0401D00A)

BossLog: { 0233} 2011/03/24 14:05:52.951 03 SYSTEM=SY1 CELL=WAS00 NODE=NDN1
CLUSTER=BBOC001 SERVER=BBOS001  PID=0X010063 TID=0X315663000000043 t=6C6938
c=UNK ./bbgrjtr.cpp+717  tag=GCS ... BBOO0220E: SECJ6237E: Authorization
failed. The SAF user MSTONE1 does not have READ access to any of the following
SAF profiles in the EJBROLE class:  [All#Role].
com.ibm.ws.security.zOS.authz.SAFAuthorizationTableImpl
```

To stop tagging of WTOs, set ras_tag_wto_messages=0

In case message tagging breaks automation

Another New Function: Classification Only Trace!

Classification XML like this:

```
<InboundClassification type="http"    . . .>
   <http_classification_info uri=“/gcs/admin/” classification_only_trace=”1”/>
</InboundClassification>
```

Then turn on some tracing (e.g. WebContainer)

WebContainer tracing is captured

ONLY from threads dispatching a request that matches a URI of /gcs/admin

MODIFY server,TRACERECORD,OFF	- turns all tracing off
MODIFY server,TRACERECORD,ON	- ignores the XML, trace on?
MODIFY server,TRACERECORD,RESET	- honor the XML
MODIFY server,DISPLAY,TRACERECORD	- shows current setting

Dynamic Updates to the XML

Re-read the current XML file

```
MODIFY server,RECLASSIFY
```

Read a new XML file

```
MODIFY server,RECLASSIFY,FILE=/some/path/and/file.xml
```

Display the file last read and when

```
MODIFY server,DISPLAY,WORK,CLINFO
```

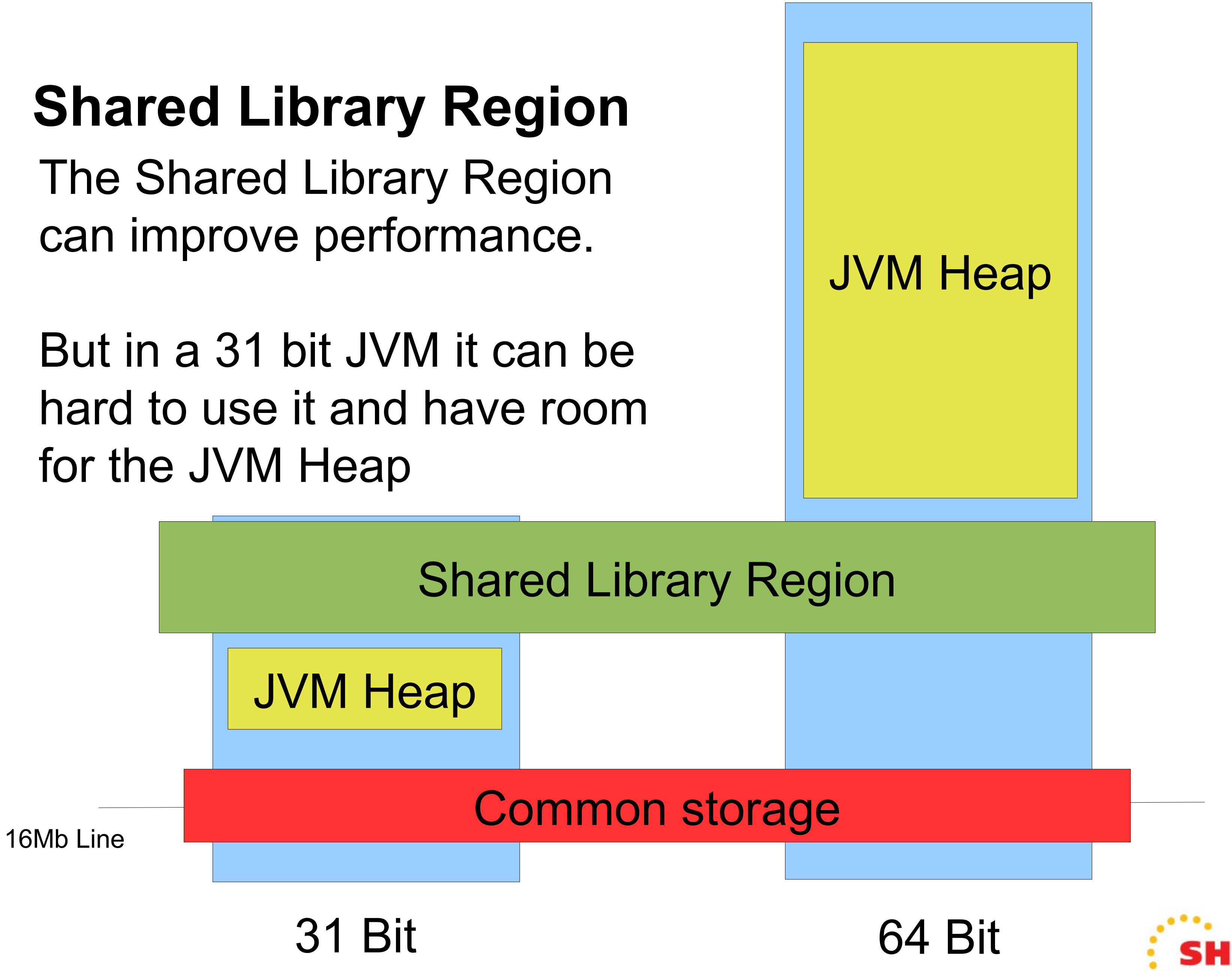
```
BBOJ0129I: The /tmp/wlm4.class.xml workload classification
file was loaded at 2009/07/14 19:33:35.297 (GMT).
```

Other Hidden Gems (Service Stream)

Shared Library Region

The Shared Library Region can improve performance.

But in a 31 bit JVM it can be hard to use it and have room for the JVM Heap



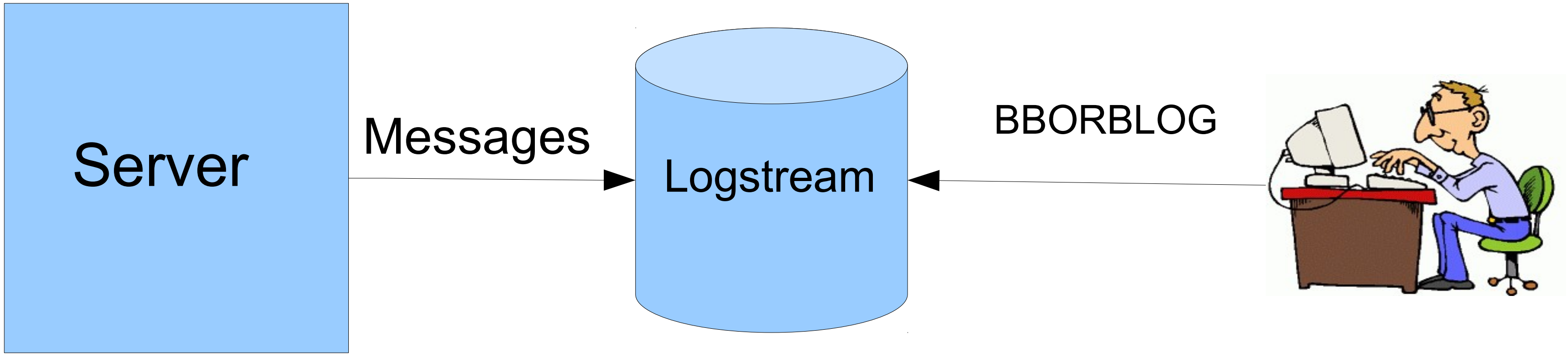
Shared Library Region

- WP101320 describes how to use BPXK_DISABLE_SHLIB in OA33516 to turn off the shared library region in the current process
- Frees up below-the-bar storage for the JVM Heap
- Leaves it enabled for other processes that have the room
- To tell if it was set, you needed to look at a dump
- WAS V6.1 (PM36368) and WAS V7 (PM32677) help

BBOO0341I VARIOUS RESOURCE MONITORING DATA: (64):(**67108864**):():():():()

- Shows the shared library region size for this process

'Error' logstream browser options



'Error' messages that are not WTOs are written to SYSOUT
Or to the 'Error' log Logstream

Viewed with the BBORBLOG utility

PK91010 adds **outdsn** and **reclen** to control output dataset name and record length

Other Hidden Gems (Version 8)

Routing for Modify Javacore etc.

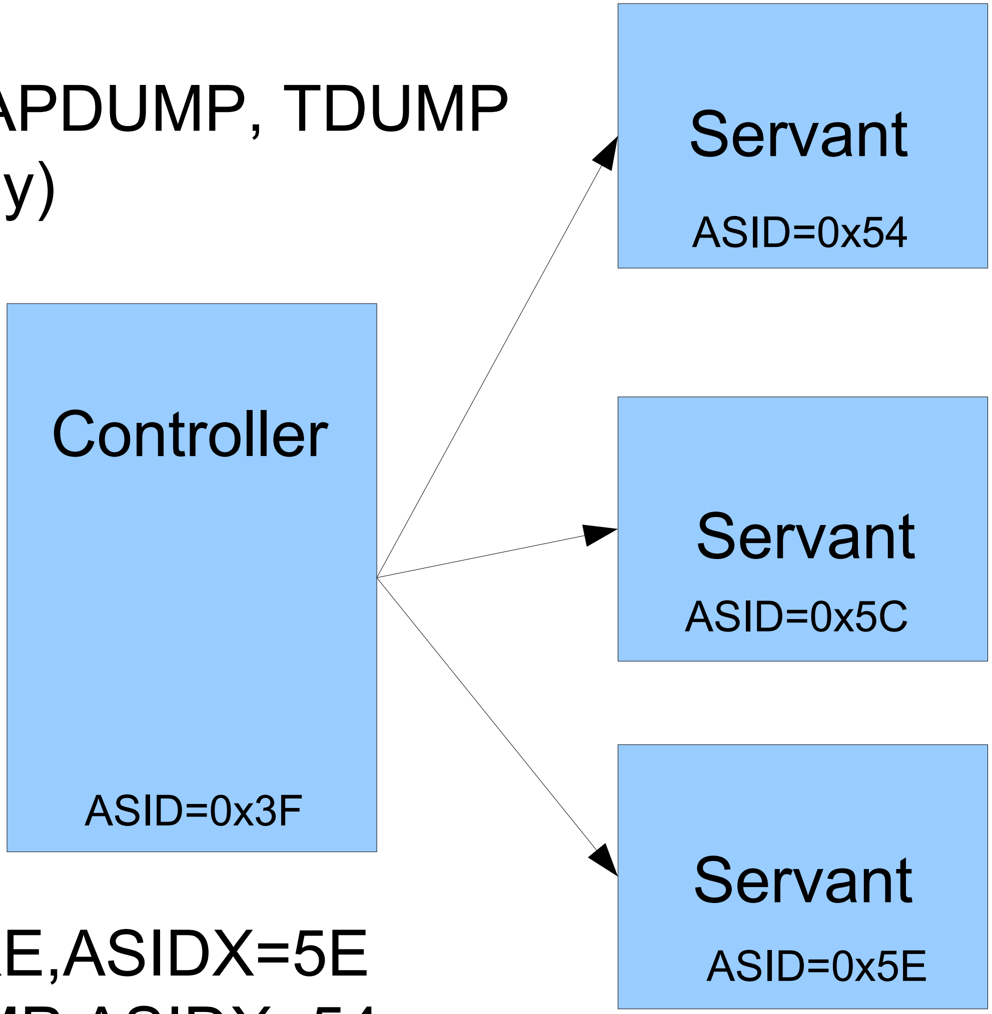
Modify JAVACORE, HEAPDUMP, TDUMP
Dumps all regions (mostly)

Until V8!

Specify target ASID:

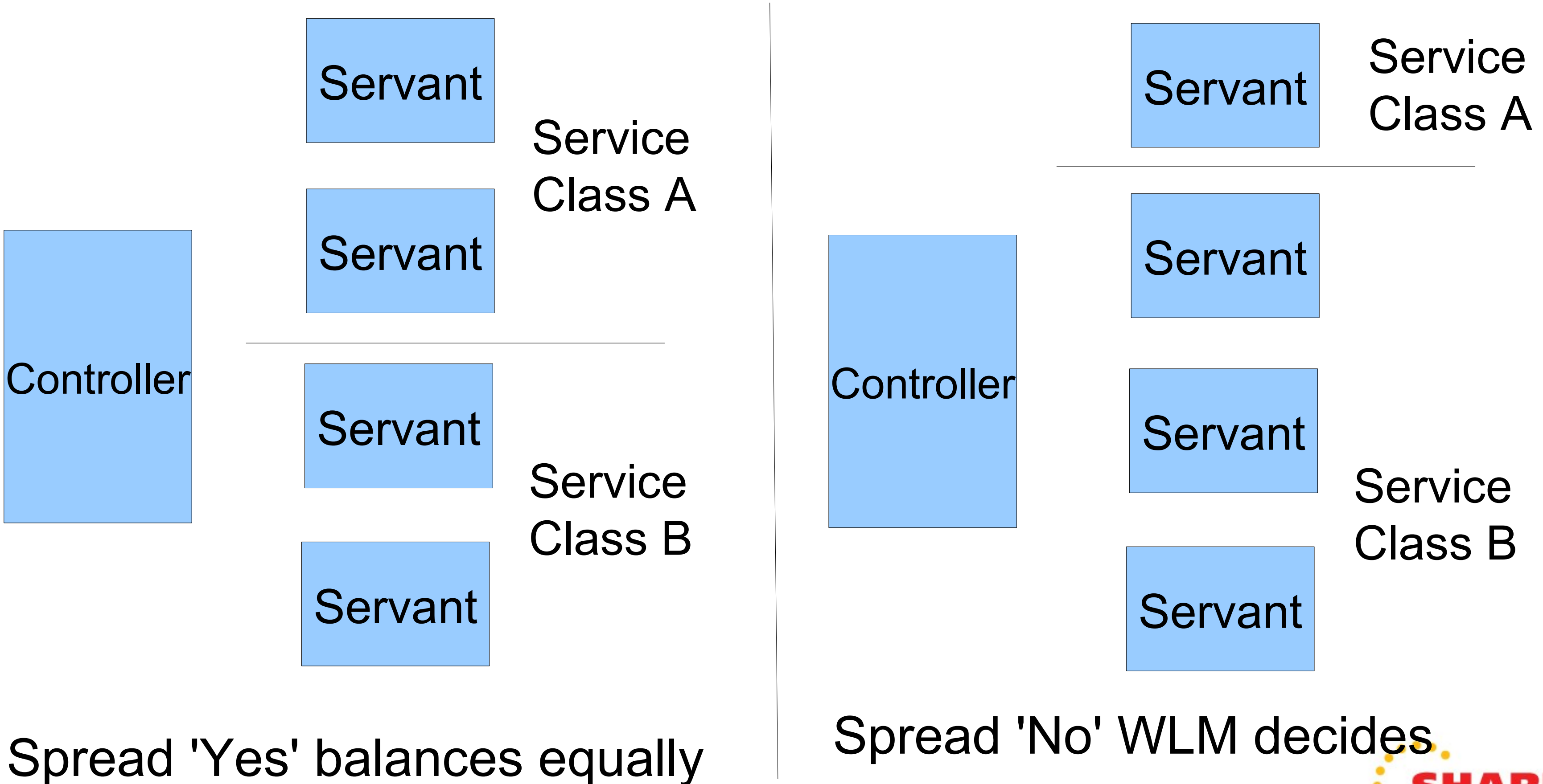
For example:

Modify server,JAVACORE,ASIDX=5E
Modify server,HEAPDUMP,ASIDX=54
Modify server,TDUMP,ASIDX=**3F**



Configure WLM AE_SPREADMIN

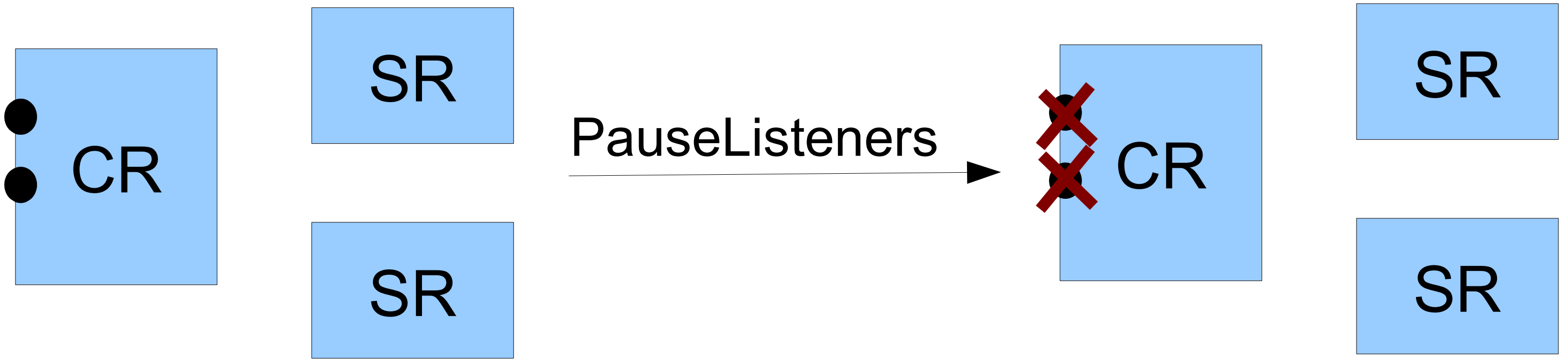
- Configure wlm_ae_spreadmin=0 or 1
  - Sets AE_SPREADMIN(NO|YES) on IWM4SLI API call



Spread 'Yes' balances equally

Spread 'No' WLM decides

Displaying 'Paused' state

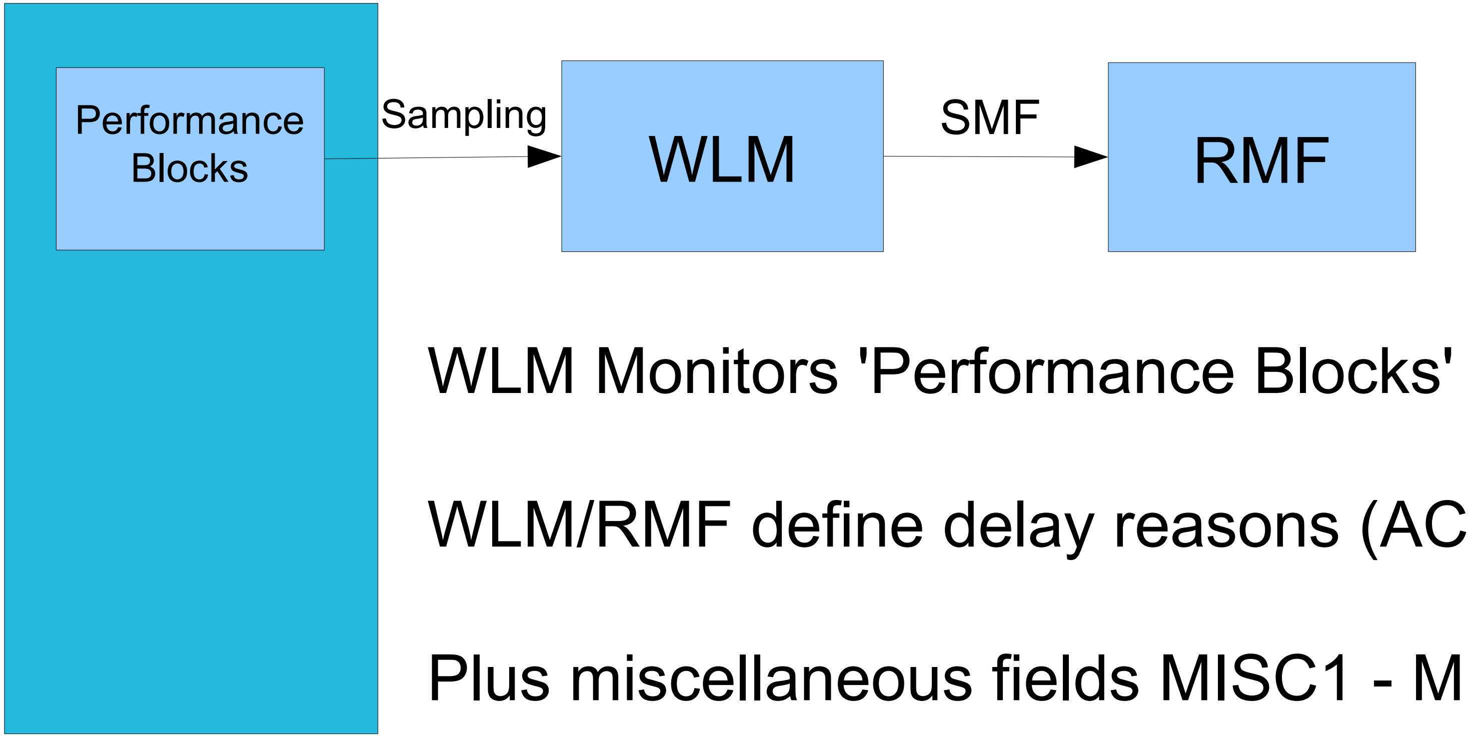


Modify PAUSELISTENERS closes ports etc.
But how do you know? In **Version 8!**

Modify server,DISPLAY,SERVERS output updated:

```
BBOO0182I SERVER           ASID SYSTEM LEVEL                STATE
BBOO0183I WAS00   /ZWASAXXX 6Fx  SY1    8.0.0.0 (ff1106.32) ACTIVE
BBOO0183I BBON001 /BBON001  58x  SY1    8.0.0.0 (ff1106.32) ACTIVE
BBOO0183I BBOC001 /BBOS001  5Bx  SY1    8.0.0.0 (ff1106.32) PAUSED/STOPPING
BBOO0183I BBODMGR /BBODMGR  57x  SY1    8.0.0.0 (ff1106.32) ACTIVE
```

Delay Monitoring Misc. States



WLM Monitors 'Performance Blocks' inside WAS

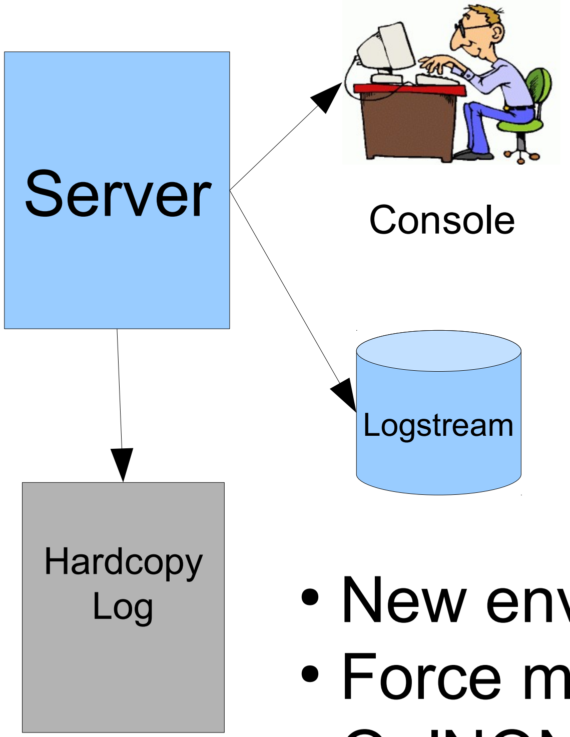
WLM/RMF define delay reasons (ACTIV_APPL)

Plus miscellaneous fields MISC1 - MISC15

WLM API IWM5MGDD allows WAS to 'explain' MISC fields

Reasons like 'RRS' will show up in RMF

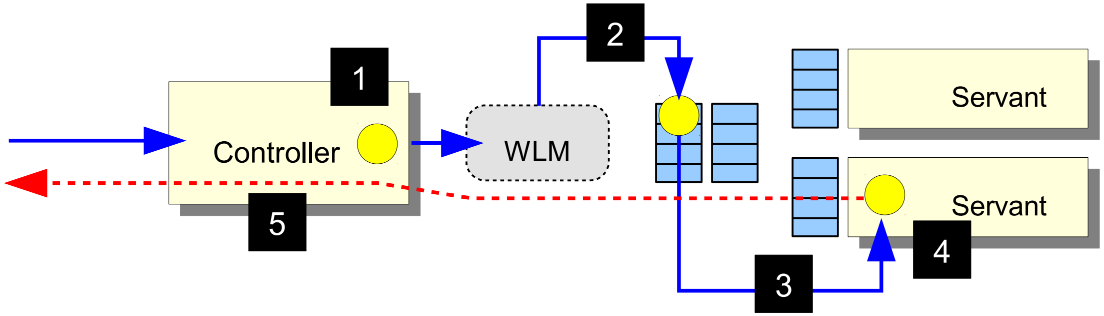
Message routing



- Messages are written as:
  - WTOs to the console
  - WTOs to the log
  - Writes to SYSOUT or Logstream
- The destination for a message is determined by the code that issues it

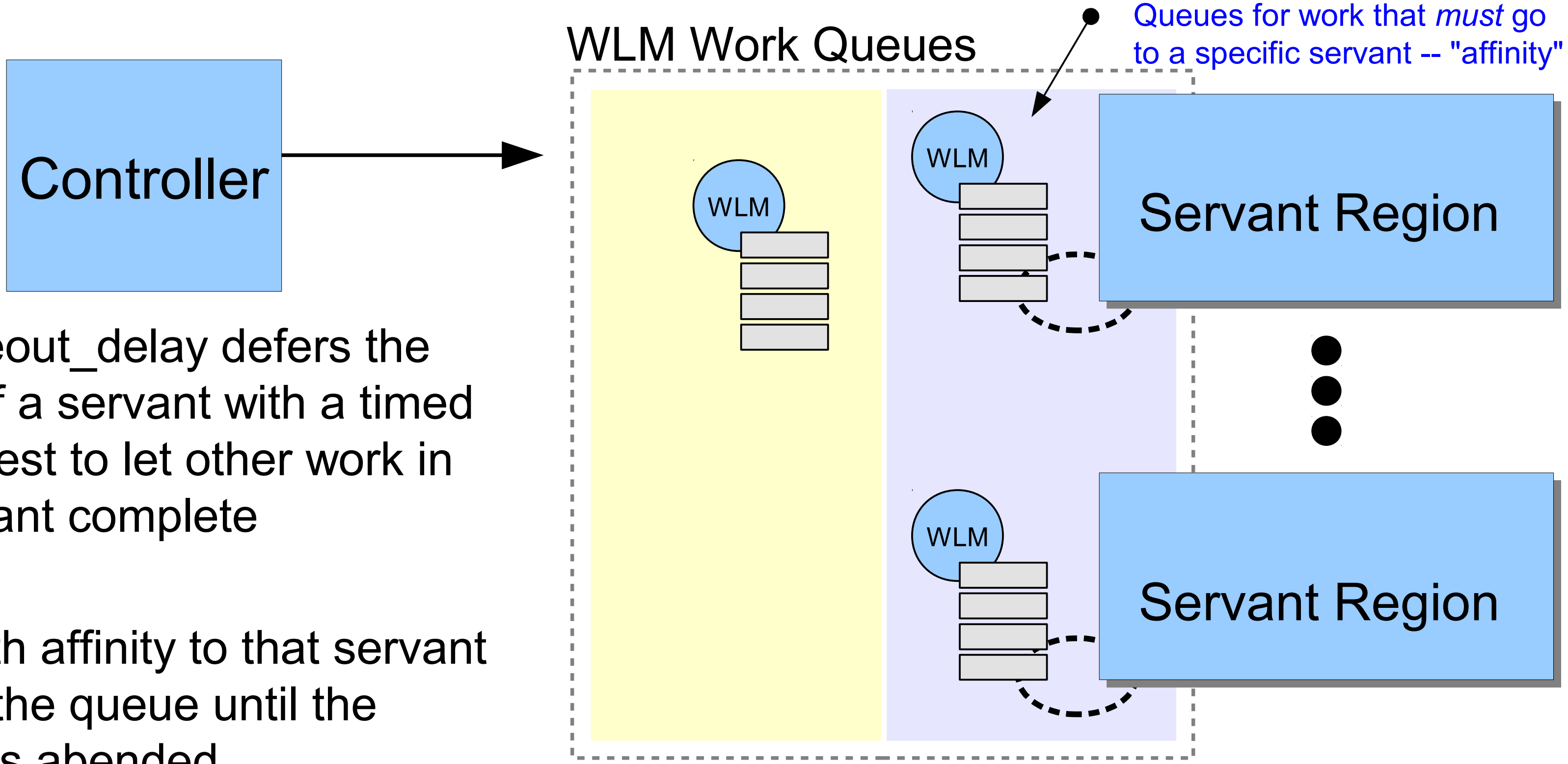
- New environment variables override the code
- Force messages (by ID) to a chosen target
- Or 'NONE' to suppress entirely
- Update dynamically with MODIFY
- Use DISPLAY to see current configuration

SMF 120-9 Updates for affinity routing



- Some requests establish an affinity to a servant region
- Later requests use that affinity and must run in the same servant
  - HTTPSession and Stateful Session Beans are examples
- The SMF 120.9 record already indicates if a request ran in a particular servant because of an affinity
- In **Version 8** we added an affinity token to the SMF record
- Find the request that created the affinity and all the later requests that used it

Timeouts and Affinity Routing



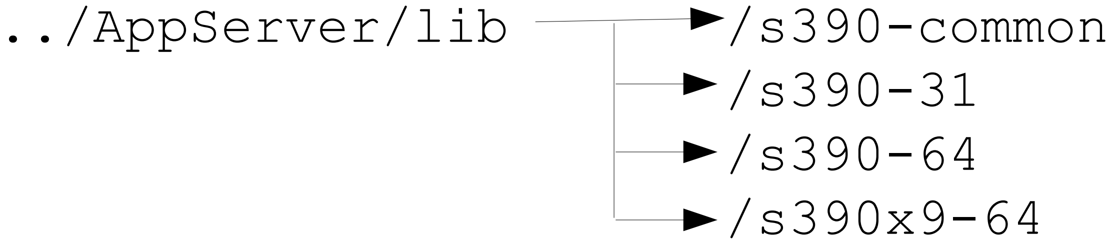
Queues for work that *must* go to a specific servant -- "affinity"

The timeout_delay defers the abend of a servant with a timed out request to let other work in the servant complete

Work with affinity to that servant waits in the queue until the servant is abended

In **Version 8** we changed the behavior so work with affinity to a dying servant is rejected.
This frees up the client thread (maybe in IHS) to try again and run in another servant after the bad one dies

Native DLLs optimized for z196



```
../AppServer/lib ──► /s390-common
                 ├─► /s390-31
                 ├─► /s390-64
                 └─► /s390x9-64
```

/s390-common

Native modules required for any bitmode and hardware

/s390-31

Native modules required for 31 bit

/s390-64

Native modules required for 64 bit and any pre-z196 hardware

/s390x9-64

Native modules required for 64 bit optimized for z196

Thread Hang Recovery – improved diagnostics

Timeout processing in Version 7

1) Dispatch begins
2) Dispatch timer expires
3) WAS issues message BBOJ0113I
4) Try to interrupt the request
5) Repeat until we give up
6) Collect configured documentation (e.g. callstack)
7) Notify the controller
8) Controller begins process of abending the servant
9) Controller issues message BBOO0327I

Thread Hang Recovery – improved diagnostics

Timeout processing in **Version 8**

1) Dispatch begins
2) Dispatch timer expires
3) WAS issues message BBOJ0113I
4) **WAS issues message BBOJ0123I**
5) **If configured, gather pre-interrupt documentation**
6) **If configured issue message BBOJ0122I with ODI info**
7) Try to interrupt the request
8) Repeat until we give up
9) **Issue message BBOJ0124I to indicate we gave up**
10) Collect configured documentation (e.g. callstack)
11) Notify the controller
12) Controller begins process of abending the servant
13) Controller issues message BBOO0327I

What are all these new messages?

```
BBOJ0113I: The Interruptible Thread Infrastructure is attempting to advance work
running under request ffff18b2

BBOJ0123I: The Interruptible Thread Infrastructure is attempting to advance work
running under request ffff18b2,
 request details: ThreadDetails: ASID = 0129, TCB = 0X006C62D8, Request =
ffff18b2, Is JVM Blocked = false, Tried to interrupt = false, Given up = false,
Internal Work Thread = false, Hung Reason = Not Hung, SR Dispatch Time =
2011/02/25 20:36:56.474373, CTL Receive Time = 2011/02/25 20:36:56.352540, CTL
Queued to WLM Time = 2011/02/25 20:36:56.471058, Request Timeout limit = 63,
Elapsed Execution Time = 65, CPU Time Used Limit = 3500000, Outbound Request
Timeout Limit = 30, ODI Details = [JVM INTERRUPTIBLE THREAD, Monitor
 ACTIVE]

BBOJ0122I: The Interruptible Thread Infrastructure about to drive a ODI to
advance work running under request ffff18b2, ODI details: Monitor ACTIVE

BBOJ0124I: The Interruptible Thread Infrastructure timed out a request and it
has become unresponsive, request ffff18b2, request details: ThreadDetails: ASID
= 0129, TCB = 0X006C62D8, Request = ffff18b2, Is JVM Blocked = false, Tried to
interrupt = true, Given up = true, Internal Work Thread = false, Hung Reason =
Dispatch Timer Popped, SR Dispatch Time = 2011/02/25 20:36:56.474373, CTL
Receive Time = 2011/02/25 20:36:56.352540, CTL Queued to WLM Time = 2011/02/25
20:36:56.471058, Request Timeout limit = 63, Elapsed Execution Time = 65, CPU
Time Used Limit = 3500000, Outbound Request Timeout Limit = 30, ODI Details =
[JVM INTERRUPTIBLE THREAD, Monitor ACTIVE]
```

What are all these new messages?

```
BBOJ0117I: JAVA THREAD STACK TRACEBACK FOR THREAD WebSphere WLM Dispatch Thread t=006c62d8:
   Hung Thread Recovery--pre-interrupt
   Traceback for thread WebSphere WLM Dispatch Thread t=006c62d8:
   com.ibm.ejs.ras.CB390TraceEventListener.writeTrace(Native Method)
   com.ibm.ejs.ras.CB390TraceEventListener.processEvent(CB390TraceEventListener.java:390)
     .
     .
     .


BBOJ0117I: JAVA THREAD STACK TRACEBACK FOR THREAD WebSphere WLM Dispatch Thread t=006c62d8:
   Thread Hang Recovery--thread could not be encouraged to complete
   Traceback for thread WebSphere WLM Dispatch Thread t=006c62d8:
   com.ibm.ejs.ras.CB390TraceEventListener.writeTrace(Native Method)
   com.ibm.ejs.ras.CB390TraceEventListener.processEvent(CB390TraceEventListener.java:390)
     .
     .
     .
```