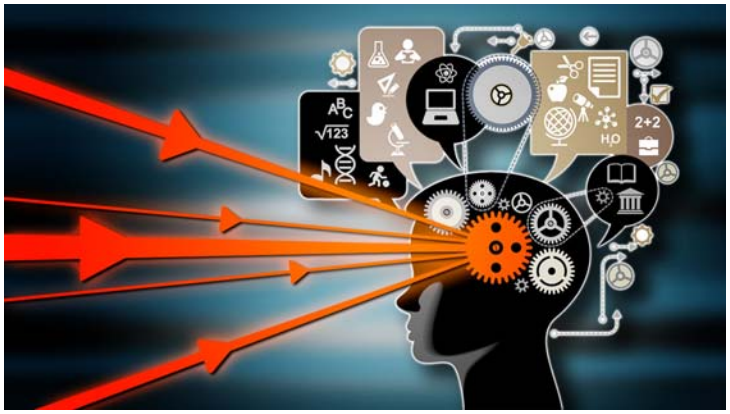
Glenn Anderson, IBM Lab Services and Training

IBM

Your Changing z/OS Performance Management World: New Workloads, New Skills



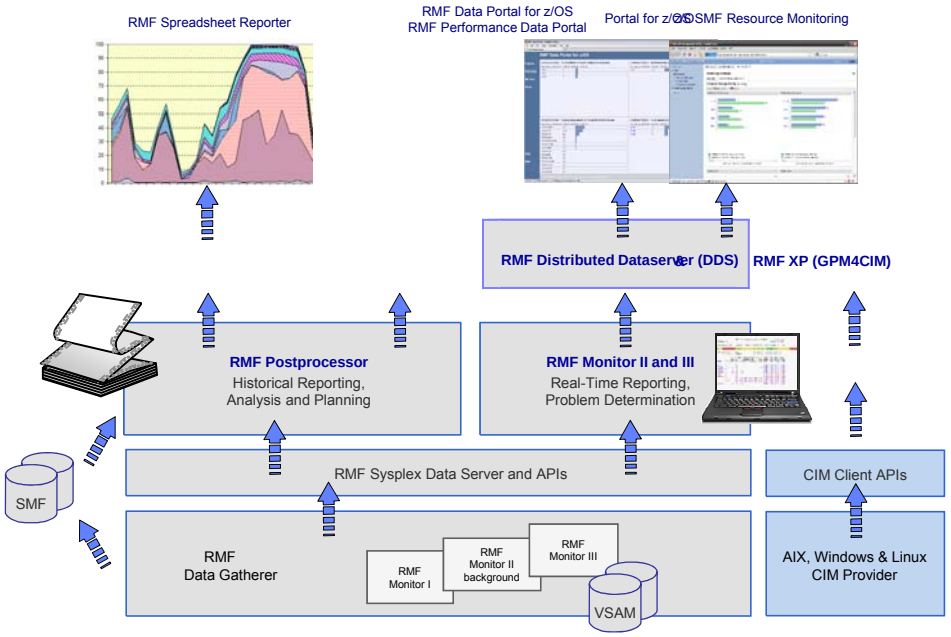
Summer SHARE
August 2015
Session 17642

Agenda

IBM

- The new world of RMF monitoring
  - RMF XP and z/OSMF
- z Systems hardware functions
  - z13 Simultaneous Multithreading (SMT) RMF data
  - When your LPAR runs at 100%
- z/OS in the new world of cloud, mobile and analytics
  - z/OS Connect
  - IDAA and WLM

RMF Product Overview and RMF XP

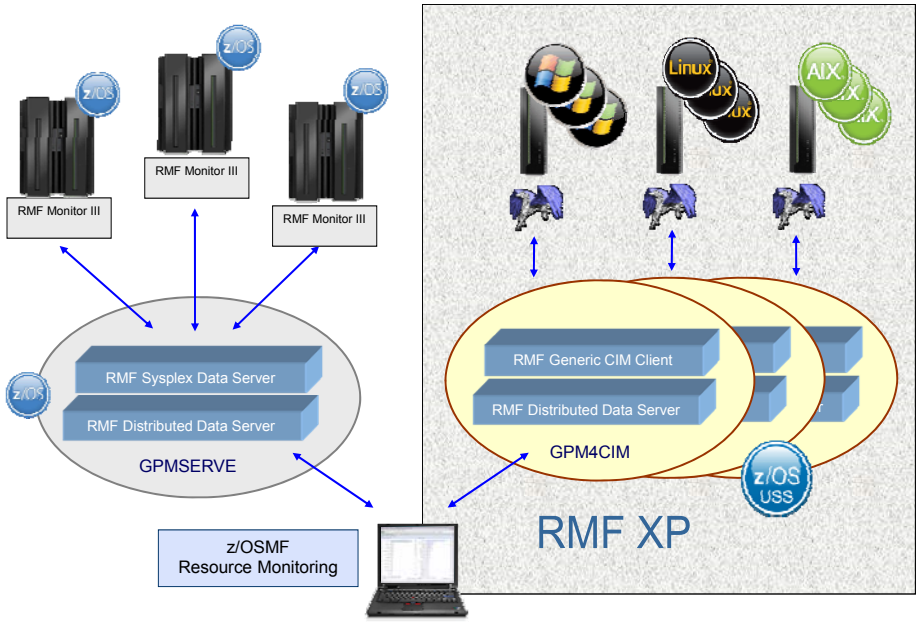


RMF XP Enhancements

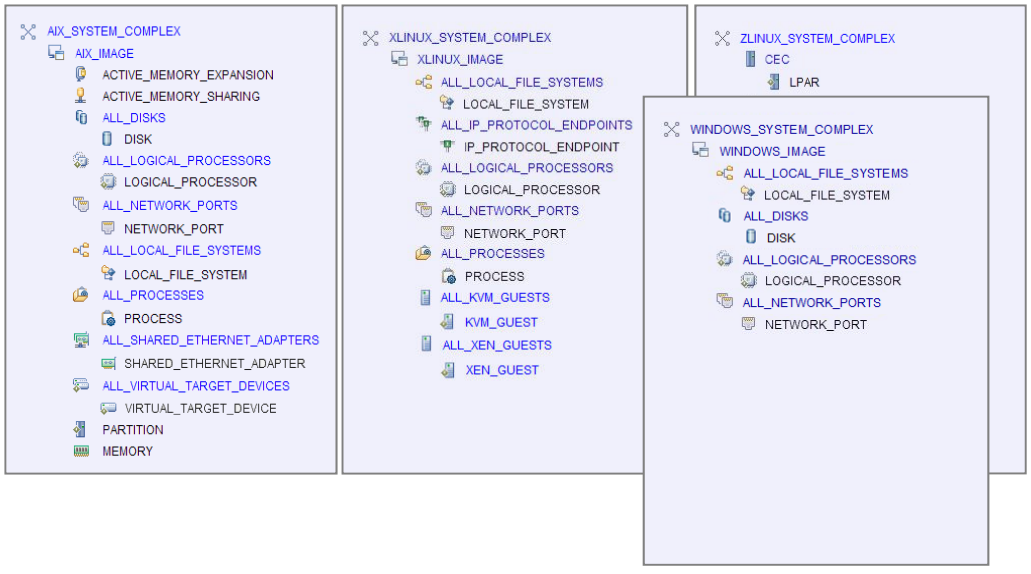
- RMF XP is the solution for Cross Platform Performance Monitoring
- RMF XP supports the Operating Systems running on
  - x Blades
  - p Blades

- In addition RMF XP supports Linux on System z
  - LPAR Mode
  - VM Guest Mode

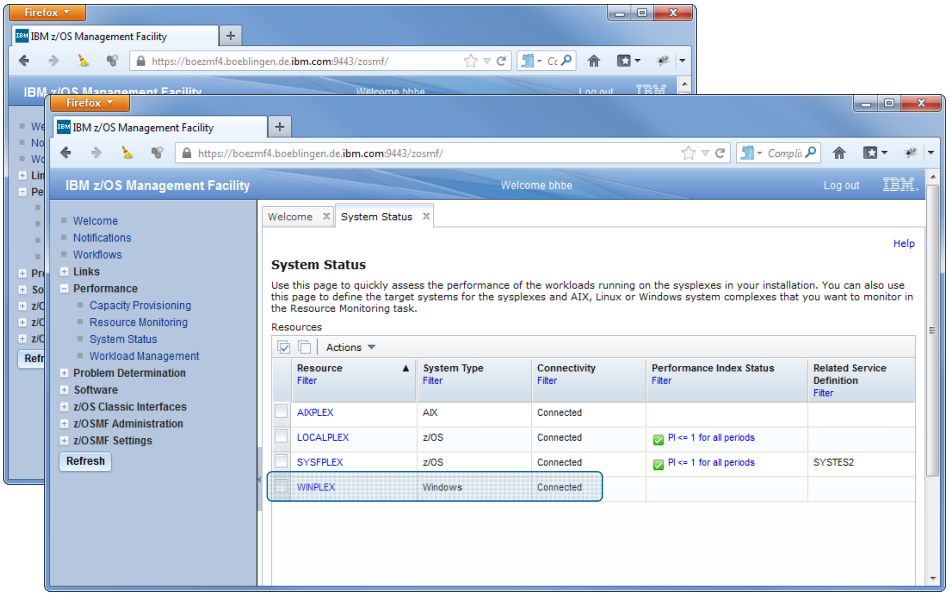
RMF XP – Component Overview



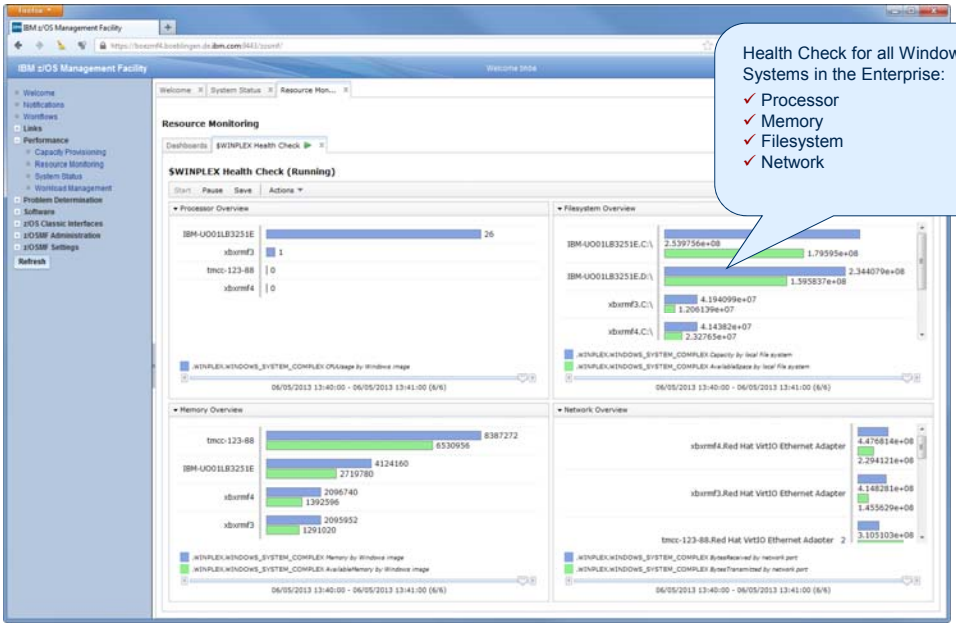
RMF XP Windows Support – Resource Model



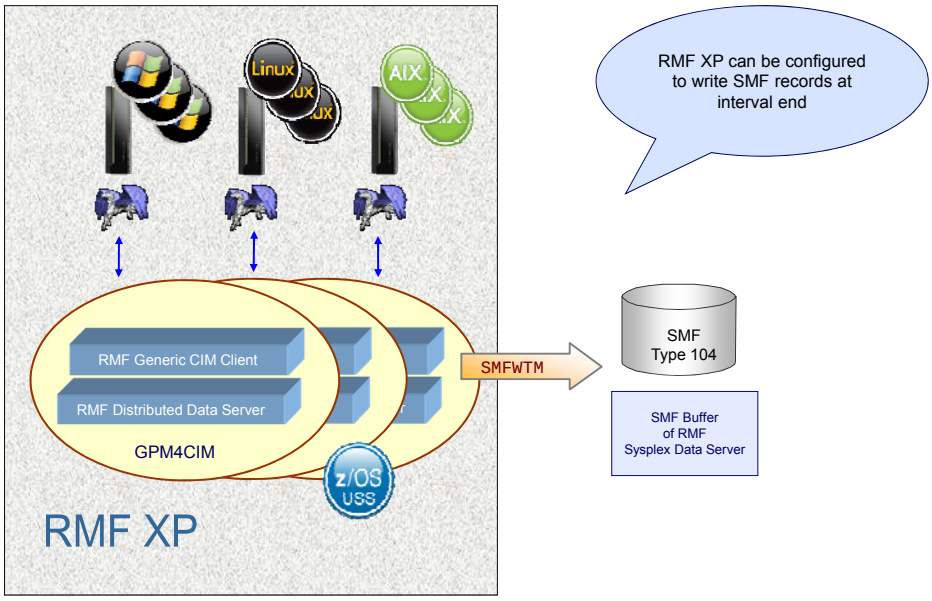
RMF XP Windows Support – z/OSMF Resource Monitoring



RMF XP Windows Support – z/OSMF Resource Monitoring



RMF XP & SMF Records



RMF XP & SMF Records

One Subtype
per Metric Category

AIX on System p	ST	Linux on System x	ST	Linux on System z	ST
AIX_ActiveMemoryExpansion	1	Linux_IPProtocolEndpoint	20	Linux_IPProtocolEndpoint	40
AIX_Processor	2	Linux_LocalFileSystem	21	Linux_LocalFileSystem	41
AIX_ComputerSystem	3	Linux_NetworkPort	22	Linux_NetworkPort	42
AIX_Disk	4	Linux_OperatingSystem	23	Linux_OperatingSystem	43
AIX_NetworkPort	5	Linux_Processor	24	Linux_Processor	44
AIX_FileSystem	6	Linux_UnixProcess	25	Linux_UnixProcess	45
AIX_Memory	7	Linux_Storage	26	Linux_Storage	46
AIX_OperatingSystem	8	Linux_KVM	30	Linux_zCEC	50
AIX_Process	9	Linux_Xen	31	Linux_zLPAR	51
AIX_SharedEthernetAdapter	10			Linux_zChannel	52
AIX_ActiveMemorySharing	11			Linux_zECKD	53
AIX_VirtualTargetDevice	12				

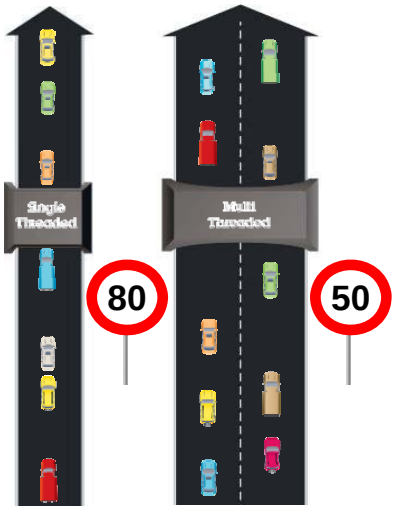
RMF XP & SMF Records

Windows on System x	ST
Windows_LocalFileSystem	60
Windows_NetworkPort	61
Windows_OperatingSystem	62
Windows_Processor	63
Windows_Storage	64

One Subtype
per Metric Category

z13 - Simultaneous Multithreading (SMT)

- "Simultaneous multithreading (SMT) permits multiple independent threads of execution to better utilize the resources provided by modern processor architectures."*
- With z13, SMT allows up to two instructions per core to run simultaneously to get better overall throughput
- SMT is designed to make better use of processors
- On z/OS, SMT is available for zIIP processing:
  - Two concurrent threads are available per core and can be turned on or off
  - Capacity (throughput) usually increases
  - Performance may in some cases be superior using single threading



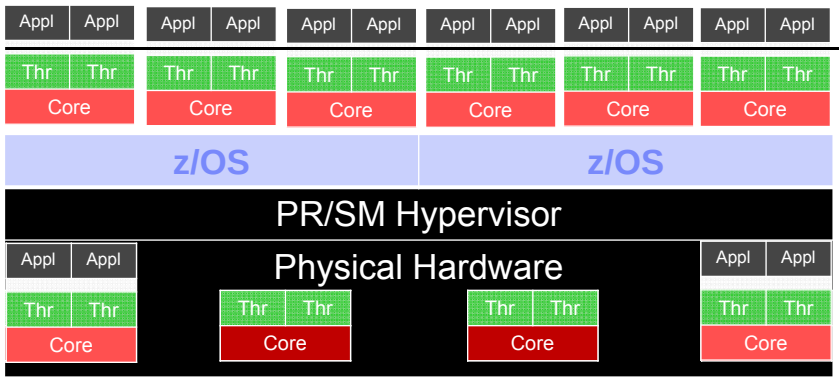
Two lanes process more traffic overall

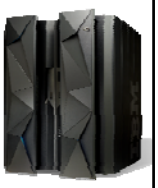
*Wikipedia®

Note: Speed limit signs for illustration only

12

z13 - SMT Exploitation





- SMT Aware OS informs PR/SM that it intends to exploit SMT
  - PR/SM can dispatch any OS core to any physical core
  - OS controls the whole core – must follow rules
    - Maximize core throughput (Drive cores with high Thread Density [2])
    - Maximize core availability (Meet workload goals using fewest cores)
- SMT is transparent to applications
- LOADxx and IEAOPTxx parmlib options to enable SMT on z/OS:
  - LOADxx: PROCVIEW **CORE|CPU**
  - IEAOPTxx: MT_ZIIP_MODE={1 | 2}

13

z13 – z/OS SMT Metrics

- Capacity Factor (CF)
  - How much work core actually completes for a given workload mix at current utilization - relative to single thread
  - MT-1 Capacity Factor is 1.0 (100%)
  - MT-2 Capacity Factor is workload dependent
- Maximum Capacity Factor (mCF)
  - How much work a core can complete for a given workload mix at most
- Core Busy Time
  - Time any thread on the core is executing instructions when core is dispatched to physical core
- Average Thread Density
  - Average number of executing threads during **Core Busy Time** (Range: 1.0 - 2.0)
- Productivity
  - Core Busy Time Utilization (percentage of used capacity) for a given workload mix
  - Productivity represents capacity in use (**CF**) relative to capacity total (**mCF**) during **Core Busy Time**.
- Core Utilization
  - Capacity in use relative to capacity total over some time interval
  - Calculated as **Core Busy Time** x **Productivity**

Actual MT-2 Efficiency

Estimated max MT-2 Efficiency

% Used MT-2 Core Capacity during Core Busy Time

% Used MT-2 Core Capacity during Measurement Interval

14

z13 – SMT: Postprocessor CPU Activity Report

- PP CPU activity report displayed in “old” format when SMT is active
- PP CPU activity report provides new metrics when SMT is active
  - MT Productivity and Utilization of each logical core
  - MT Multi-Threading Analysis section displays MT Mode, MT Capacity Factors and average Thread Density
- One data line in PP CPU activity report represents one thread (CPU)
  - CPU NUM designates the logical core
- Some metrics like TIME % ONLINE and LPAR BUSY provided at core granularity only

```
                                C P U   A C T I V I T Y
z/OS V2R1                SYSTEM ID CB8B            DATE 02/02/2015          INTERVAL 15.00.004
                         RPT VERSION V2R1 RMF      TIME 11.00.00            CYCLE 1.000 SECONDS
---CPU---    --------------- TIME % ------------   --- MT % ----  LOG PROC    --I/O INTERRUPTS--
NUM  TYPE    ONLINE   LPAR BUSY  MVS BUSY PARKED   PROD   UTIL   SHARE %      RATE     % VIA TPI
 0    CP     100.00   68.07      67.94      0.00   100.00  68.07  100.0  HIGH   370.1    13.90
 1    CP     100.00   46.78      46.78      0.00   100.00  46.78   52.9  MED     5.29    16.93
...
TOTAL/AVERAGE          8.66       54.17             100.00   8.66  152.9          375.3    13.95
 A    IIP    100.00   48.15      41.70      0.00    85.84  41.33  100.0  HIGH
                                 35.66      0.00
 B    IIP    100.00   38.50      32.81      0.00    85.94  33.09  100.0  HIGH
                                 26.47      0.00
...
TOTAL/AVERAGE         29.48      23.23              86.47  25.39  386.7
----------- MULTI-THREADING ANALYSIS ---------------
 CPU TYPE    MODE     MAX CF          CF       AVG TD
      CP        1      1.000       1.000        1.000
      IIP       2      1.485       1.279        1.576
```

MT-2 core capacity used

Productivity of logical core while dispatched to physical core

15

z13 – SMT: Postprocessor Workload Activity Report

```
                              W O R K L O A D   A C T I V I T Y

z/OS V2R1      SYSPLEX UTCPLXCB          DATE 02/02/2015     INTERVAL 15.00.004        MODE = GOAL
               RPT VERSION V2R1 RMF      TIME 11.00.00

REPORT BY: POLICY=BASEPOL

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---
AVG     790.12  ACTUAL            27.787  SSCHRT  3975  IOC      4233K  CPU 4659.039  CP    542.89
MPL     790.12  EXECUTION         15.761  RESP     2.8  CPU    306468K  SRB 1513.624  AAPCP   0.00
ENDED     9173  QUEUED                1  CONN     1.4  MSO         0   RCT    0.622  IIPCP   2.00
END/S    10.19  R/S AFFIN             0  DISC     1.2  SRB     97415K  IIT   21.116
#SWAPS    6087  INELIGIBLE            0  Q+PEND   0.1  TOT    408116K  HST    0.000  AAP      N/A
EXCTD    15860  CONVERSION            0  IOSQ     0.0  /SEC   453461   AAP      N/A  IIP    97.84
AVG ENC   4.00  STD DEV         8.40.915                              IIP 1308.346
REM ENC   0.00                                          BSRPTN   574
MS ENC    0.00                                          TRX SERV 574
```

MT-1 equivalent service units and service times

Percentage of maximum core capacity used

- Pre SMT:
  - ⇨ Service time : Logical processor CPU time
  - ⇨ APPL%: Percentage of logical processor capacity
- SMT mode active:
  - ⇨ Service time: MT-1 equivalent CPU time
  - ⇨ APPL% : Percentage of maximum core capacity calculated as

$$APPL\% = \frac{\text{MT–1 equivalent CPU time}}{\text{Interval length} \times \text{mCF}} \times 100$$

For MT Mode = 1 ⇨ mCF = 1

7/27/2015

Running at or near 100% - Blocked Workload Support

IBM

- **Problem**
  - Work competes for resources, serialized by locks and latches
    - Low import work may hold a resource and high important work may have to wait for it
- **WLM Blocked Workload Support**
  - Recognizes blocked work
    - Work which doesn't show any progress for an elongated period of time
  - Allows this work to use a small amount of CPU periodically
    - With the hope to resolve existing (potential) resource contentions

Blocked Workload Support: User Interface: IEAOPT

IBM

BLWLTRPCT	Percentage of the CPU capacity of the LPAR to be used for promotion ❑ Specified in units of 0.1% ❑ Default is 5 (=0.5%) ❑ Maximum is 200 (=20%) ❑ Would only be spent when enough units of work exist which need promotion
BLWLINTHD	Specifies threshold time interval for which a blocked address space or enclave must wait before being considered for promotion. ❑ Minimum is 5 seconds. Maximum is 65535 seconds. ❑ Default is 20 seconds.

Blocked Workload Support: RMF

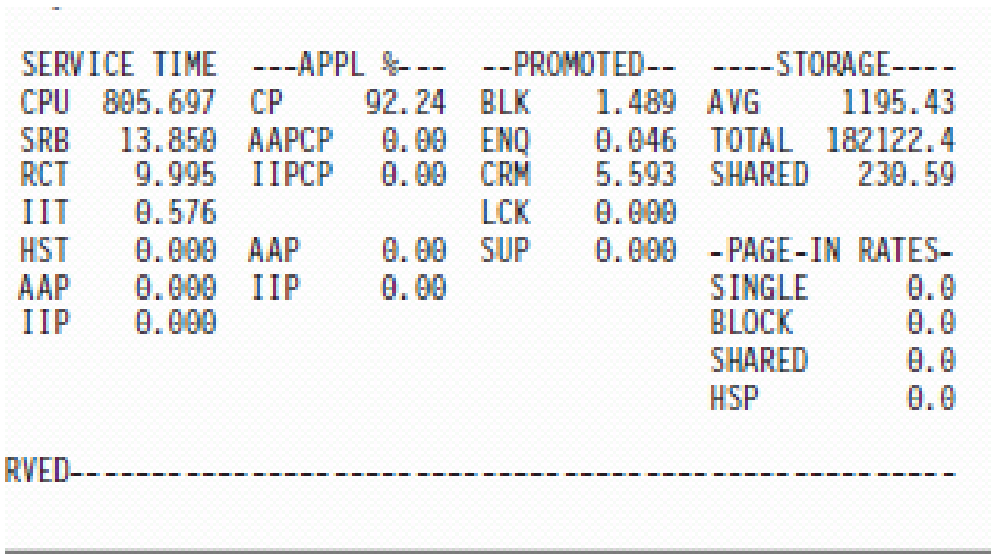
```
                                C P U   A C T I V I T Y
…

BLOCKED WORKLOAD ANALYSIS

OPT PARAMETERS: BLWLTRPCT (%)   0.5   PROMOTE RATE:  DEFINED    50000   WAITERS FOR PROMOTE:  AVG   0.001
                BLWLINTHD        60                 USED (%)      95                         PEAK     15
```

- **Extensions of RMF Postprocessor CPU Activity and WLMGL reports with information about blocked workloads and the temporary promotion of their dispatching priority**
- **SMF record 70-1 (CPU activity) and SMF 72-3 (Workload activity)**

Promoted transactions: RMF workload activity report

```
                                W O R K L O A D   A C T I V I T Y                                    PAGE   1
   z/OS V1R13              SYSPLEX SVPLEX3            DATE 09/28/2011          INTERVAL 15.00.003   MODE = GOAL
                           RPT VERSION V1R13 RMF      TIME 17.00.00

                                   POLICY ACTIVATION DATE/TIME 09/14/2011 11.08.09

------------------------------------------------------------------------------------------------------------------- SERVICE CLASS(ES)

 REPORT BY: POLICY=BASEPOL    WORKLOAD=STC_WLD    SERVICE CLASS=STCLOW     RESOURCE GROUP=NONE
                                                  CRITICAL     =NONE
                                                  DESCRIPTION  =Low priority for STC workloads

 -TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---  --PROMOTED--  ----STORAGE----
 AVG     153.37  ACTUAL          3.02.885  SSCHRT  56.9  IOC      3964   CPU  805.697  CP     92.24  BLK    1.489  AVG      1195.43
 MPL     152.35  EXECUTION       3.02.391  RESP    15.1  CPU     15184K  SRB   13.850  AAPCP   0.00  ENQ    0.046  TOTAL  182122.4
 ENDED      599  QUEUED               494  CONN     1.3  MSO         0   RCT    9.995  IIPCP   0.00  CRM    5.593  SHARED   230.59
 END/S     0.67  R/S AFFIN              0  DISC     0.3  SRB    261005   IIT    0.576                LCK    0.000
 #SWAPS    3391  INELIGIBLE             0  Q+PEND   4.5  TOT     15449K  HST    0.000  AAP     0.00  SUP    0.000  -PAGE-IN RATES-
 EXCTD        0  CONVERSION         5.188  IOSQ     9.0  /SEC    17202  AAP    0.000  IIP     0.00                SINGLE      0.0
 AVG ENC   0.00  STD DEV         3.27.429                                 IIP    0.000                              BLOCK       0.0
 REM ENC   0.00                                                                                                     SHARED      0.0
 MS ENC    0.00                                        ABSRPTN   113                                              HSP         0.0
                                                       TRX SERV  112

 ----------------------------------------------SERVICE CLASSES BEING SERVED----------------------------------------------------------
 DB2LOW
```

Promoted transactions RMF workload activity report



```
SERVICE TIME  ---APPL %---  --PROMOTED--  ----STORAGE----
CPU  805.697  CP     92.24  BLK    1.489  AVG     1195.43
SRB   13.850  AAPCP   0.00  ENQ    0.046  TOTAL  182122.4
RCT    9.995  IIPCP   0.00  CRM    5.593  SHARED   230.59
IIT    0.576                LCK    0.000
HST    0.000  AAP     0.00  SUP    0.000  -PAGE-IN RATES-
AAP    0.000  IIP     0.00                  SINGLE      0.0
IIP    0.000                                BLOCK       0.0
                                            SHARED      0.0
                                            HSP         0.0

RVED-----------------------------------------------------
```

Promoted transactions RMF field definitions

CPU time in seconds that transactions in this group were running at a promoted dispatching priority, separated by the reason for the promotion:

BLK CPU time in seconds consumed while the dispatching priority of work with low importance was temporarily raised to help blocked workloads
ENQ CPU time in seconds consumed while the dispatching priority was temporarily raised by enqueue management because the work held a resource that other work needed.
CRM CPU time in seconds consumed while the dispatching priority was temporarily raised by chronic resource contention management because the work held a resource that other work needed
LCK In HiperDispatch mode, the CPU time in seconds consumed while the dispatching priority was temporarily raised to shorten the lock hold time of a local suspend lock held by the work unit.
SUP CPU time in seconds consumed while the dispatching priority for a work unit was temporarily raised by the z/OS supervisor to a higher dispatching priority than assigned by WLM.

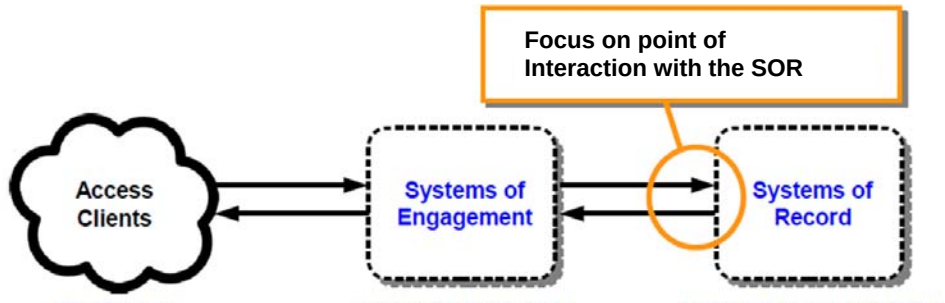
Work unit queue distribution: Mon I CPU report

```
SYSTEM ADDRESS SPACE AND WORK UNIT ANALYSIS
  ---------NUMBER OF ADDRESS SPACES---------      ----------------------DISTRIBUTION OF IN-READY WORK UNIT QUEUE------------
  QUEUE TYPES          MIN       MAX      AVG     NUMBER OF             0    10   20   30   40   50   60   70   80   90   100
                                                  WORK UNITS     (%)    |....|....|....|....|....|....|....|....|....|....|
  IN                   550     1,008    594.8
  IN READY               4       438     22.7     <=  N           55.5   >>>>>>>>>>>>>>>>>>>>>>>>>>>>
  OUT READY              0         1      0.0      =  N +   1     4.4   >>>
  OUT WAIT               0         0      0.0      =  N +   2     4.0   >>>
  LOGICAL OUT RDY        0       628     10.3      =  N +   3     3.7   >>
  LOGICAL OUT WAIT     178       634    589.4     <=  N +   5     7.1   >>>>
                                                  <=  N +  10    14.7   >>>>>>>>
  ADDRESS SPACE TYPES                             <=  N +  15     8.0   >>>>>
                                                  <=  N +  20     1.9   >
  BATCH                281       284    282.0     <=  N +  30     0.2   >
  STC                  708       763    736.4     <=  N +  40     0.0
  TSO                   97        98     97.9     <=  N +  60     0.0
  ASCH                   0         1      0.0     <=  N +  80     0.0
  OMVS                  43        97     68.0     <=  N + 100     0.0
                                                  <=  N + 120     0.0
  ---------NUMBER OF WORK UNITS------------        <=  N + 150     0.0
  CPU TYPES            MIN       MAX      AVG     >   N + 150     0.0

  CP                   444       888    555.5    N = NUMBER OF PROCESSORS ONLINE UNPARKED (22.4 ON AVG)
  AAP                   22        33     28.8
  IIP                    0         0      0.0
```

16 Buckets representing the work unit count with regard to the number of online processors

Work unit count on CPU type level

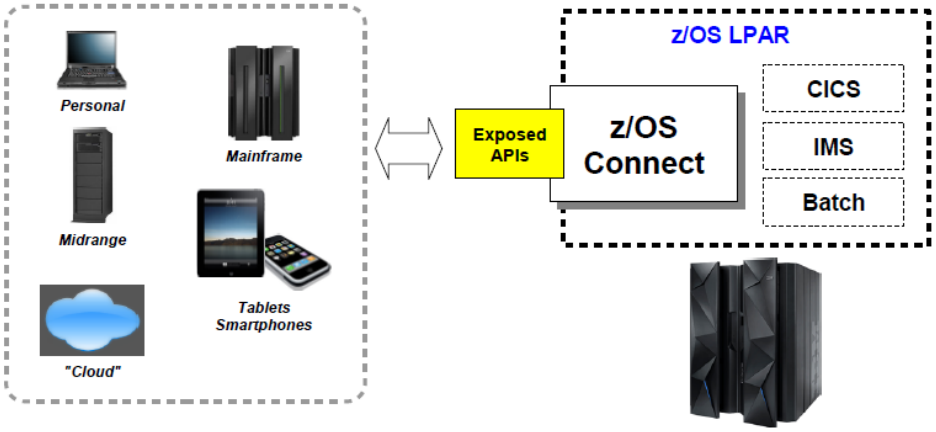
Systems of Engagement and Systems of Record



Mainframe as a Service

IBM

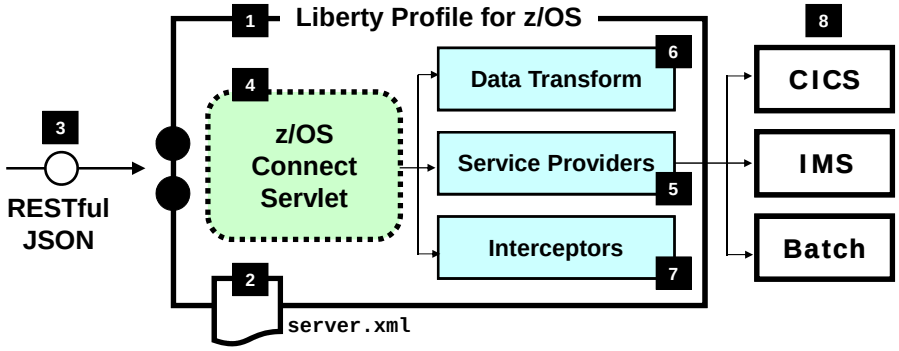
Another use-case for z/OS Connect is as a standard gateway into the z/OS LPAR to expose programs as a service:



z/OS Connect provides a way to do this with a single entry point (HA *is* possible) and common protocol (REST/JSON)

What is z/OS Connect?

IBM

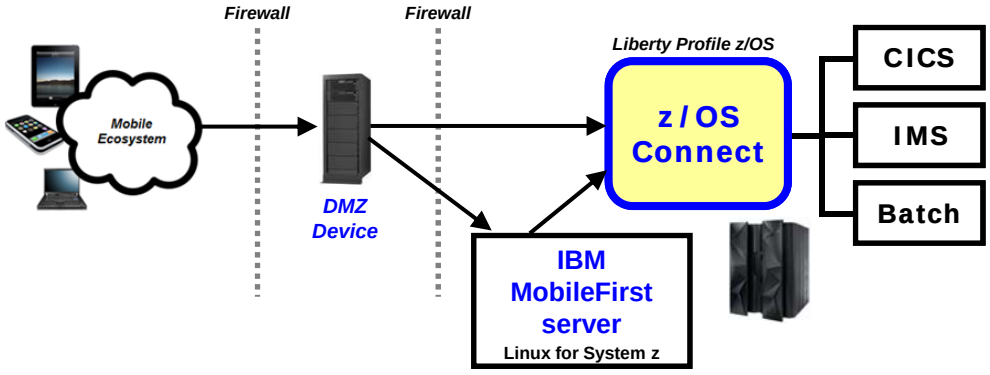


1. z/OS Connect is software function that runs in Liberty Profile for z/OS.
2. z/OS Connect is described and configured in the Liberty `server.xml` file
3. z/OS Connect is designed to accept RESTful URIs with JSON data payloads
4. One part of z/OS Connect is a servlet that runs in Liberty Profile z/OS.
5. A 'Service Provider' is software that provides the connectivity to the backend system
6. z/OS Connect provides the ability to transform JSON to the layout required by backend
7. 'Interceptors' are callout points where software can be invoked to do things such as SAF authorization and SMF activity recording
8. Initially the backend systems supported will be CICS, IMS and Batch

26

Context Within Overall Mobile Architecture

The message here is that z/OS Connect is a *piece* of the Mobile architecture, but in most cases will not be the only component:



Users of z/OS Connect would access through normal corporate firewall infrastructure

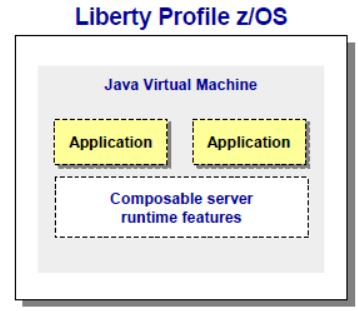
IBM MobileFirst Platform to provide application management, security and operational governance for mobile applications

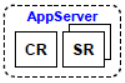
z/OS Connect would be behind the secure firewall, and on LPARs along with backend systems

27

Liberty Profile z/OS

Liberty Profile is IBM's dynamic and composable server runtime. First shipped with Version 8.5, it is available on many platforms, including z/OS:





Not this ... this is the "traditional WAS" model

- **Single JVM per server model**
  - As opposed to the multiple JVM model of traditional WAS z/OS (the CR/SR model)
- **Simple configuration structure**
  - One XML file serves as the main configuration file
- **Dynamic**
  - Changes to the configuration file or to the applications are detected and dynamically loaded
- **Composable**
  - You tell Liberty Profile what features and functions you want and only that code is loaded
- **On z/OS can run from UNIX shell or as a z/OS started task**
  - On z/OS we anticipate most will run as started task

Liberty Profile is the basis for z/OS Connect, so any discussion of z/OS Connect necessarily involves Liberty

Audit (SMF) Interceptor

IBM

The audit interceptor writes SMF 120.11 records with the following information captured:

Liberty Profile z/OS

z/OS Connect

- **System Name**
- **Sysplex Name**
- **Jobname**
- **Job Prefix**
- **Address Space Stoken**

Server Identification Section

- **Arrival Time**
- **Completion Time**
- **Target URI**
- **Input JSON Length**
- **Response JSON Length**
- **Method Name**
- **Service Name**
- **Userid**

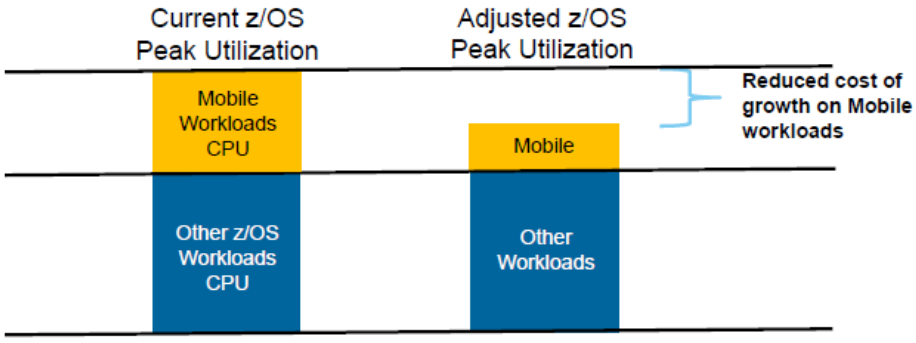
z/OS Connect User Data Section

29

System z mobile pricing model

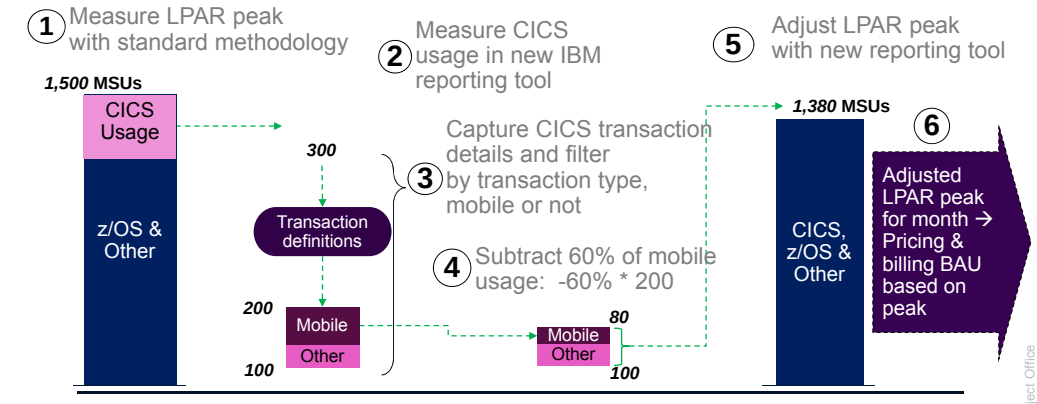
IBM

- Reduce z/OS peak MSUs attributable to mobile workloads -- up to 60%
- No Infrastructure Changes Required… (such as separate LPARs)



- Customers must tag and track z/OS CPU seconds from mobile workloads.
- New MWRT tool replaces SCRT and will subtract mobile CPU seconds from peaks.

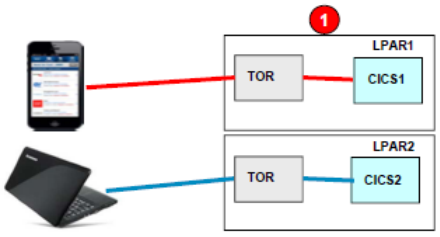
Mobile Workload Pricing helps alleviate spikes caused by increased mobile usage…



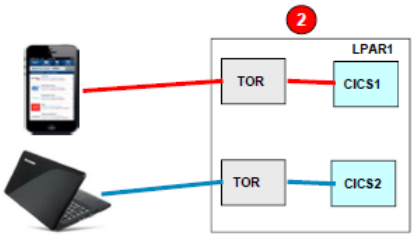
* Figures are for illustrative purposes only. Tracking process and records will vary by customer

Competitive Project Office

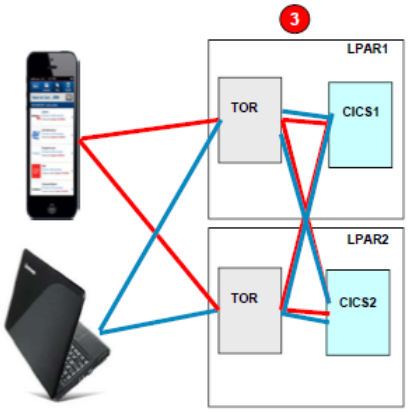
Which tracking mechanism can be used for MWP?



Use an **Individual LPAR** for mobile-only workloads



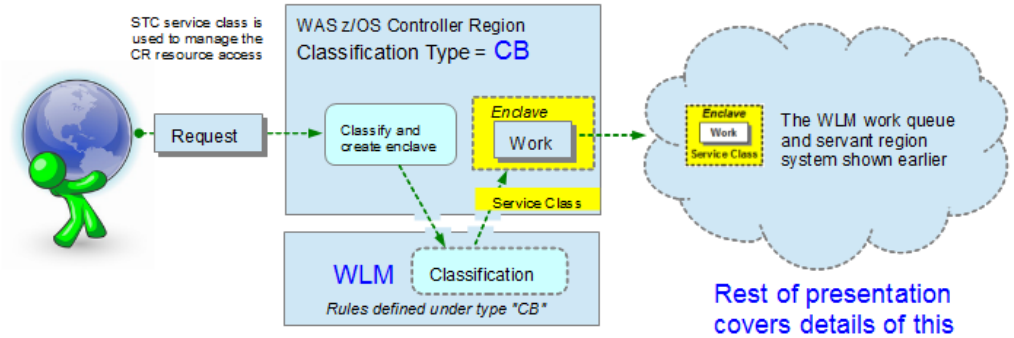
Use **Individual Regions** for mobile-only workloads



Use **Same Regions** for mobile and non-mobile workloads

The WLM Enclave

An "enclave" is a way to identify and manage individual pieces of work *within* the many parts of a running z/OS system



Key points from this chart

- An "enclave" is simply a way for WLM to understand priorities at a work unit level
- WAS does this automatically ... if you do no other configuration it'll still do this with default values

Assigning a Service Class to the Enclave

```
Subsystem Type CB - WebSphere z/OS CN and TC Classifications
Classification:
  Default service class is CBDEFLT 5
  Default report class is RWASDEF

    Qualifier  Qualifier      Starting       Service  Report
  # type       name           position       Class    Class
  - ---------- -------------- ---------      -------- -------
  1 CN         DFDMGR*                       CBCLASS  DFDMGR
                          1
  1 CN         DFSR01*                       CBCLASS  DFSR01
  2   TC       DFTRAN1          2            DFTRAN1  DFSR01T
  2   TC       DFTRAN2                       DFTRAN2  DFSR01T
                                3
  1 TC         DFTRAN3                       DFTRAN3  DFTRAN3
                                4
```

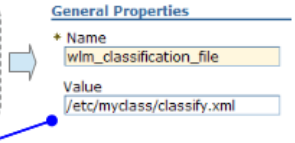
Enclaves created in WAS CR are classified by rules in CB subsystem type:

1. CN of `DFDMGR*` matches the Deployment Manager. Work there goes to `CBCLASS`.
2. Work in `DFSR01*` cluster *without* a transaction classification gets `CBCLASS` as well.
3. Work in `DFSR01*` cluster *with* TC of `DFTRAN1` or `DFTRAN2` get service classes as shown
4. Work that matches the TC of `DFTRAN3` *regardless of WAS CN* gets service class `DFTRAN3`
5. Anything that doesn't match any specific rules gets the default service class of `CBDEFLT`

The transaction class name file

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

Scope to cell or node
server scope for classification deprecated



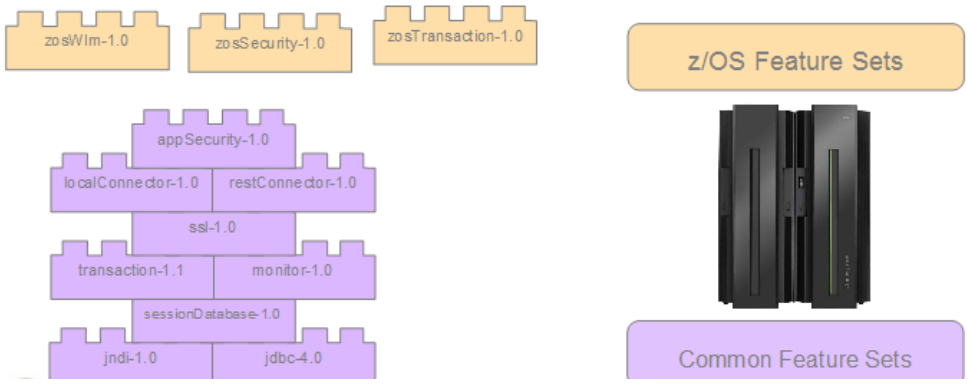

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE Classification SYSTEM "Classification.dtd"
>
<Classification schema_version="1.0">
  :
   <InboundClassification type="iiop" ... >
     (classification information)
   </InboundClassification>
   <InboundClassification type="http" ... >
     (classification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

What is the WAS for z/OS Liberty profile?

- The WAS for z/OS Liberty profile is Liberty with *optional*, independently enabled *extensions* that exploit z/OS facilities
  - Only enable exploitation of z/OS features you need
  - Only configure the z/OS functions you use
- Focus of v8.5 is basic integration and exploitation



Feature – z/OS Workload Manager

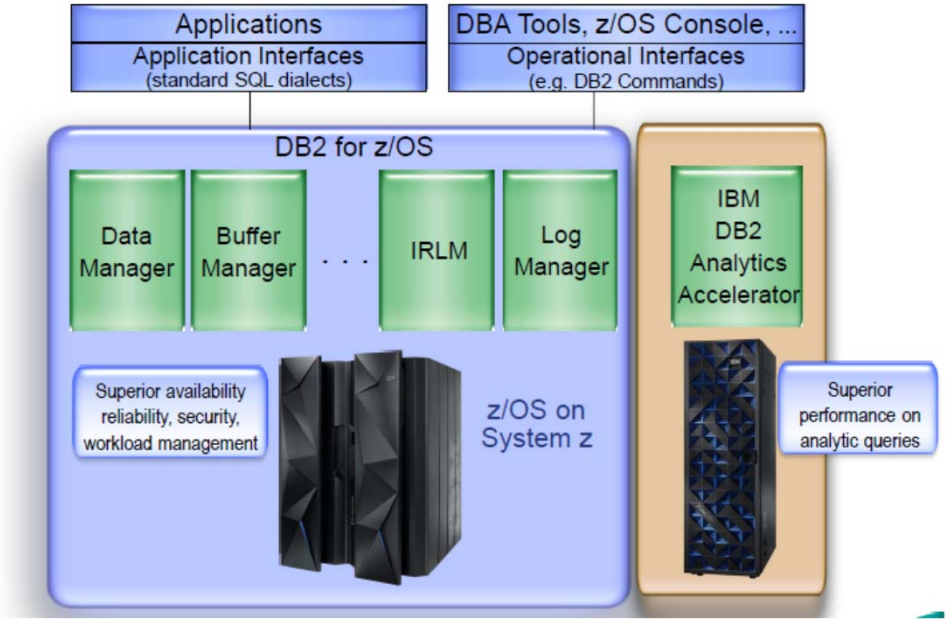
IBM

- Adds support to classify HTTP requests with z/OS WLM
  - Classification associates response time goals and importance to work run in WebSphere
  - z/OS workload manager will manage the resources available on the system in a way that ensures the most important work runs while attempting to meet response time goals
  - RMF reports provide information about completed transactions, response times, etc by service class

```
<server description="mvsworkloadManagement">

    <featureManager>
        <feature>zoswlm-1.0</feature>
    </featureManager>

    <wlmClassification/>
        <httpClassification transactionClass="WLPTRADE" resource="/tradelite/**" />
        <httpClassification transactionClass="WLPDFLT" />
    </wlmClassification>
</server>
```

IBM DB2 Analytics Accelerator

IBM



WLM and IDAA Interaction

IBM

- DB2 detects WLM service class and importance level and sends it to the accelerator with each query
- The accelerator maps the importance level to a Netezza priority and alters the session prior to the query execution, using the corresponding priority. Also, threads scheduled will have their priorities adjusted

WLM Importance Level	Netezza Priority	
	Version 3	Version 4
System	Critical	Critical
Importance 1	Critical	Critical
Importance 2	High	Critical
Importance 3	Normal	High
Importance 4	Normal	Normal
Importance 5	Normal	Low
Discretionary	Low	Low

Summary

IBM

- The new world of RMF monitoring
  - RMF XP and z/OSMF
- z Systems hardware functions
  - z13 Simultaneous Multithreading (SMT) RMF data
  - When your LPAR runs at 100%
- z/OS in the new world of cloud, mobile and analytics
  - z/OS Connect
  - IDAA and WLM