

# **RMF – The Latest and Greatest**

Peter Muench (<u>pmuench@de.ibm.com</u>) IBM Corporation

Monday, August 2, 2010 Session 7763





#### **Trademarks**



#### The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by (are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

#### For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

\*, AS/400@, e business(logo)@, DBE, ESCO, eServer, FICON, IBM@, IBM (logo)@, iSeries@, MVS, OS/390@, pSeries@, RS/6000@, S/30, VM/ESA@, VSE/ESA, WebSphere@, xSeries@, z/OS@, zSeries@, z/VM@, System i, System i5, System p, System p5, System x, System z, System z9@, BladeCenter@

#### The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries. Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both. Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries

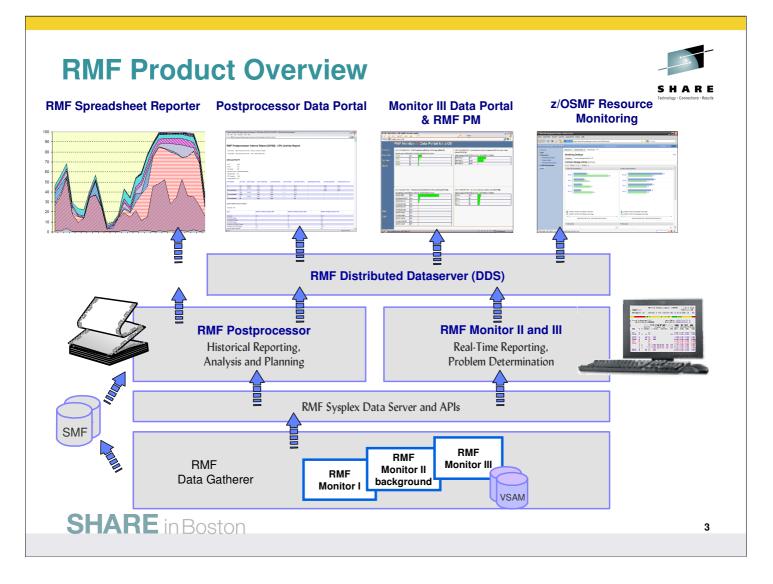
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both. ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office. IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

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

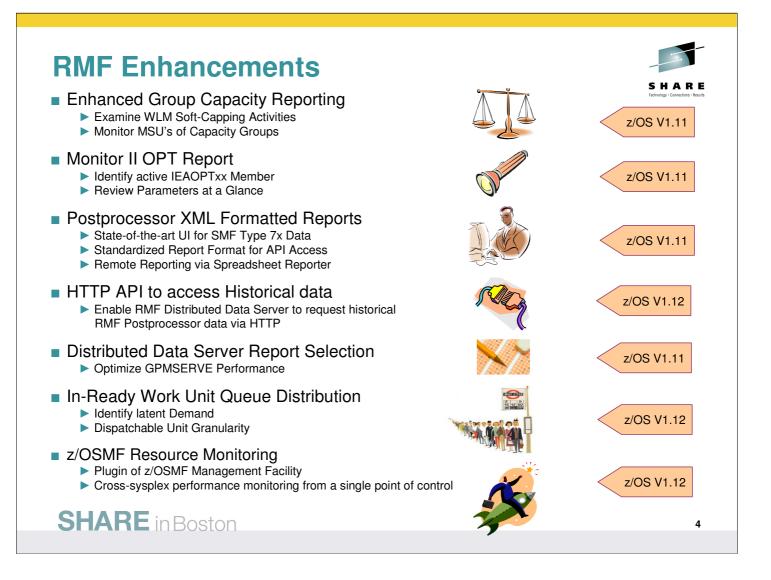
Notes: Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's 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. The customers in other countries, and the information may be subject to charace without.

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.

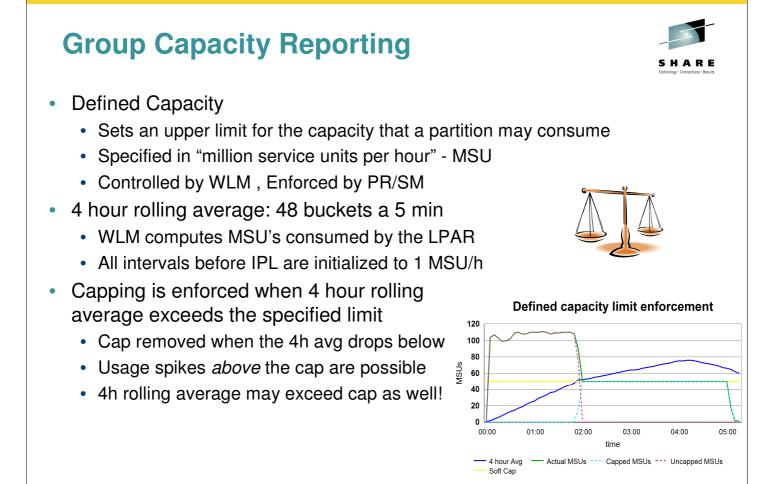


- z/OS Resource Measurement Facility (RMF) is an optional priced feature of z/OS. It supports installations in performance analysis, capacity planning, and problem determination. For these disciplines, different kinds of data collectors are needed:
  - Monitor I long term data collector for all types of resources and workloads. The SMF data collected by Monitor I is mostly used for capacity planning and performance analysis
  - Monitor II snap shot data collector for address space states and resource usage. A subset of Monitor II data is also displayed by the IBM SDSF product
  - Monitor III short-term data collector for problem determination, workflow delay monitoring and goal attainment supervision. This data is also used by the RMF PM Java Client and the RMF Monitor III Data Portal
- Data collected by all three gatherers can be saved persistently for later reporting (SMF records or Monitor III VSAM datasets)
- While Monitor II and Monitor III are realtime reporters, the RMF Postprocessor is the historical reporting function for Monitor I data
- One of the key components for the sysplex wide access of Monitor III data is the RMF Distributed Data Server (DDS). With RMF for z/OS 1.12, DDS supports HTTP requests to retrieve RMF Postprocessor data from a selection of RMF Postprocessor reports. Since the requested data are returned as XML document, a web browser can act as Data Portal to retrieve and display RMF Postprocessor data.
- With z/OS 1.12 there's another exploiter of the RMF DDS data: The z/OSMF Resource Monitoring plugin of the z/OSMF Management Facility.



In accordance with the availability of new z/OS releases and new hardware functionality, the capabilities of RMF are enhanced consecutively

- Since z/OS V1.8, the system administrator can specify a capacity limit for groups including multiple partitions
  - Beyond the existing metrics, RMF provides now extended measurements to track the MSU consumption of Capacity Groups and to control WLM Soft-Capping activities
- The new Monitor II OPT report allows the system administrator to identify the currently active IEAOPTxx member and to review the contained parameters at a glance
- As an alternative to the standard text format, Postprocessor reports can now be generated in XML format. The new format provides:
  - Browser based, state-of-the-art display of SMF Type 7x data
  - · Standardized format for access to performance data via XML parsing
- The RMF Distributed Data Server (DDS) is extended by the capability to grant instant and easy access to RMF long-term historical data by using HTTP and XML.
- The footprint of the RMF Distributed Data Server can be reduced by the possibility to exclude certain reports from the processing
- · The in-ready queue distribution is now reported on work unit granularity
- z/OS Resource Monitoring is a new plugin of the z/OSMF Management Facility. The Web browser based user interface allows a cross-sysplex performance monitoring from a single point of control:
  - · Enterprise-wide health check of all z/OS sysplexes.
  - Performance monitoring with graphically display of RMF Monitor III metrics as well as Linux metrics by means of customizable views.



• The WLC pricing model allows to define an individual capacity limit for each partition within a CEC. The sum of all limits

**SHARE** in Boston

for a specific software product running in the affected partitions is denoted as Product Certificate Capacity. It is taken as base for the software charge

5

- The limits are specified by means of the Support Elements resp. the Hardware Management Console
- How does this work in detail? WLM controls the CPU consumption of each LPAR in terms of MSU's/hour. Once the limit
  is reached, WLM will advice the LPAR Hypervisor to cap this partition. The capping technique is based on a long-term
  observation: WLM considers a 4 hour history (48 intervals of five minutes each) for the capping decisions. This ensures
  that an individual partition can receive temporarily CPU cycles above the limit to get important work done. Partitions
  cannot be capped below the MSU limit. That is, a guaranteed minimum of work can be executed at any time
- immediately after IPL ,the 4 hour history is initialized with a value of 1 MSU/h. This leds to an extremely high starting credit. Despite of incipient capping, this credit can remain for very long periods. Additionally, the 4 hour rolling average can exceed the defined capacity limit as well

# **Group Capacity Reporting**

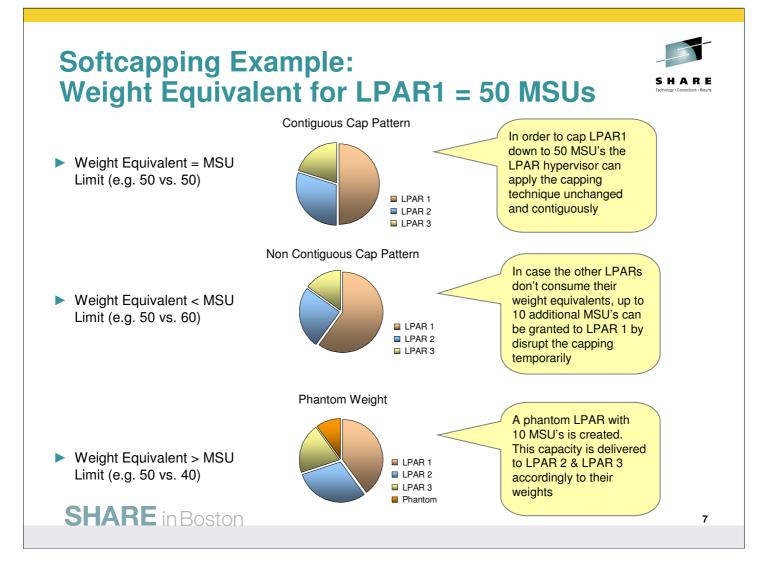


#### How does WLM Softcapping work?

- WLM exploits the existing Capping mechanism in a PR/SM environment. (used by PR/SM hypervisor to limit the CPU access for a logical partition)
- When CAPPING=YES for an individual partition is active, the maximum capacity granted to this partition is derived from the weighting factors. The partition cannot "borrow" additional capacity from other partitions (even though contention is not in effect)
- > The Capping mechanism depends on partition weight vs. defined capacity:

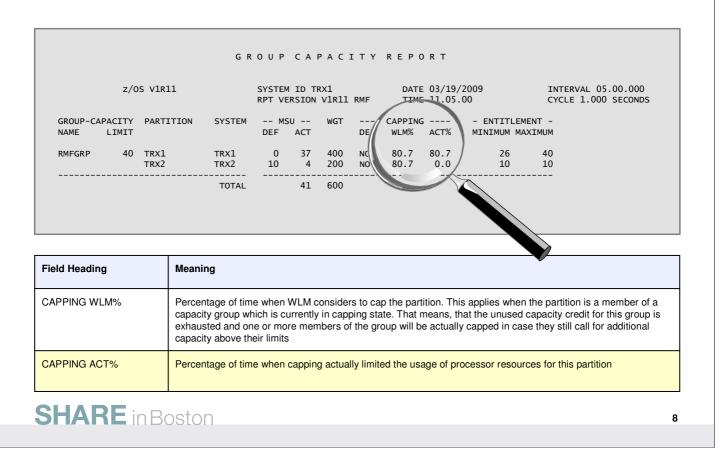
Weight vs. defined capacity	How capping is enforced
Weight equivalent = defined capacity limit	PR/SM hypervisor caps at partition weight
Weight equivalent < defined capacity limit	Noncontiguous cap pattern
Weight equivalent > defined capacity limit	Capping via "Phantom weight"

- WLM has not reinvented the capping technique: it simpy exploits the existing capping mechanism in a PR/SM environment
- The LPAR Hypervisor assigns CPU cycles to individual partitions accordingly to their weighting equivalents. However, a partition can consume processor capacity above the weighting equivalent (up to the total CEC capacity under the following circumstances:
  - WAIT COMPLETION is set to NO (this global setting delivers the complete CPU time slices to the partitions, even if not needed)
  - CAPPING for the receiving partition is set to NO
  - · There is free capacity available in the CEC (no contention)
- · As soon as contention comes up, the LPAR Hypervisor will cap the partitions down to their weighting equivalents
- WLM exploits this mechanism to initiate soft-capping by juggling either with time slices or weighting factors:
  - In case the weighting equivalent limit matches exactly the defined capacity limit, the capping mechanism can be applied unchanged (set CAPPING=YES)
  - In case the weighting equivalent is below the defined capacity, the unlimited appliance would hurt too much. Therefore, WLM uses a noncontiguous cap pattern. In other words, WLM makes the LPAR Hypervisor temporarily believe that CAPPING=YES has been specified
  - In case the weighting equivalent is above the defined capacity, theoretically capping below the weighting equivalent would be impossible. In order to adjust the weighting equivalent down to the defined capacity, WLM invents a virtual partition (phantom) with a corresponding weighting factor



# Group Capacity Reporting...





- Together with Release 1.11 the Postprocessor CPU Activity report is enhanced. The Group Capacity Report section provides a new field ACT% in the CAPPING section which indicates the percentage of interval time where capping actually limited the usage of processor resources within the partition.
- In contrast WLM%, the ACT% value represents the true capping ratio for this individual partition. RMF compares the current MSU consumption (based on the partitions dispatch times) to the actual MSU limit for this partition. The individual MSU limit is derived from the following definitions:
  - · Weighting equivalent
  - Logical processor configuration
  - Individual MSU limit
  - · Group capacity definitions
- Although an individual partition is not capped at all or only capped partially, the WLM% value can reach up to 100% under the following circumstances:
  - The partition is a member of a group which is actually in a 100% capping state
  - The partitions MSU weighting equivalent is higher than their actual MSU limit. Therefore, in case of demand, the phantom weight capping technique would be applied to this partition

Gro	bu	p	Ca	<b>IP</b>	ac	ιτγ	H	ep	ortin	g				S I Technol	
							P	ARTI	TION DA	TA REPOR	т				
	z/	'os v1	R11			SYSTEM RPT VE		RX1 V1R11		03/19/2009 11.05.00		VAL 05.00 1.000 SI		PAC	GE 2
MVS PARTI IMAGE CAP, NUMBER OF WAIT COMP DISPATCH	ACITY CONF	, IGURE	D PARTI	TIONS		TRX1 40 58 NO YNAMIC		NUMBE	R OF PHYSICAL CP AAP IFL ICF IIP	PROCESSORS	26 4 1 18 2 1		GROUP NAME LIMIT AVAILABLE	RMFC	GRP 40* 12
	PART	ITION	DATA -				L	OGICAL	PARTITION PRO	CESSOR DATA	AVERAGE	PROCESSO	R UTILIZATION	PERCENT	AGES -
NAME	S	WGT	MSU DEF		-CAPI DEF	PING WLM%	PROC NUM	ESSOR- TYPE	DISPATCH EFFECTIVE	TIME DATA TOTAL	LOGICAL PRO EFFECTIVE		PHYSICAL LPAR MGMT EF		
TRX1 H05LP45 TRX2CFA H05LP59 H05LP60 TRX2 *PHYSICAL TOTAL	A A A	400 10 100 100 10 200	50 0 0 0 60	1 1 8	NO NO YES NO NO	0.0 0.0 0.0 0.0	3.0 2 1 3 2 3.0		00.00.01.866 00.00.04.139 00.00.02.594 00.00.21.928 00.00.09.073	00.00.10.501 00.00.02.218 00.00.04.210 00.00.3.274 00.00.23.760 00.00.09.495 00.00.22.805 	1.13 0.31 1.38 0.29 3.65 1.01	1.17 0.37 1.40 0.36 3.96 1.06	0.03 0.01 0.06 0.15 0.04 1.90  2.21	0.85 0.16 0.34 0.22 1.83 0.76	0.88 0.18 0.35 0.27 1.98 0.79 1.90  6.36
Field Hea	ding		Mean	ing											
LIMIT			starte	d to b	e a me	mber o	f this	capacity	group less that	pacity group. An an four hours ago e limitations and i	. Therefore, th	ne capacit	y of this partition	on is still	artition
AVAILABI	.E		memb	ers. I	f the v		negati			be allowed by th ap one more mer					

- The header information of the Partition Data report is enhanced to display the long-term average of CPU service units which would be allowed by the limit of the capacity group but are not used by its members. If the value is negative, the capacity group is in WLM capping state.
- Note: Since the AVAILABLE value has system scope, it cannot be added to the Group Capacity Report. The values might differ slightly due to asynchronous data collection intervals on the members of the capacity group

# Group Capacity Reporting...



10

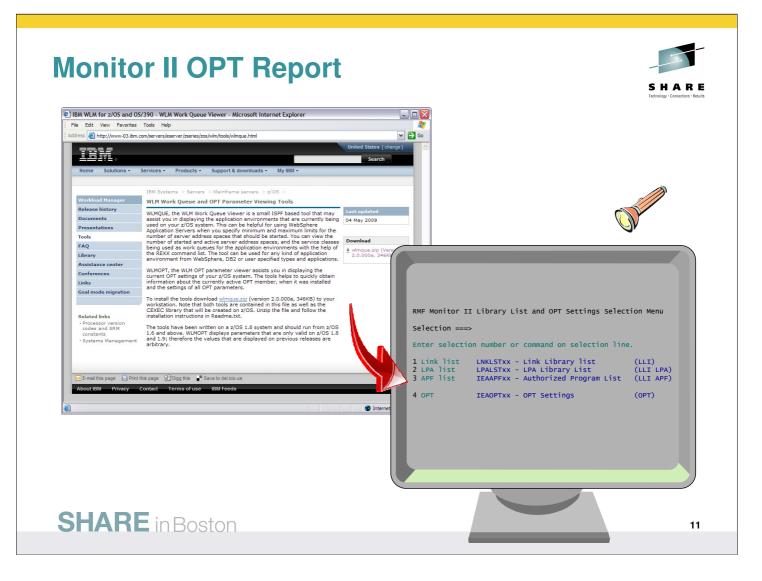
The Monitor III CPC Capacity report is enhanced:

- New metric: 4h unused group capacity average. This is the average available capacity for the group during the last 4 hours. If the value is negative the members have consumed capacity beyond the defined group limit. Therefore WLM sets the group into the capping state.
- New metric: Remaining time until group capping. This is the projected time until the usage of processor resources for one or more members of the group might be limited.
- Both new metrics are available as hidden fields in the ISPF version of the CPC report
- The threshold applied when calculating the *remaining time until capping* is changed for partitions being member of a capacity group: The MSU limit for the projection is now derived from the image capacity <u>or</u> group capacity definitions.

The following new metrics will be provided by the RMF Distributed Dataserver:

	Resource	New metric
	Sysplex	<ul> <li>available capacity (MSU/h) for group by partition</li> <li>remaining time until group capping in seconds by partition</li> </ul>
	LPAR	<ul> <li>available capacity (MSU/h) for group</li> <li>remaining time until group capping in seconds</li> </ul>
Sł		

- The Monitor III CPC Capacity report computes the projected time until WLM soft capping will start. WLM soft capping takes place to prevent a partition from using more than the defined capacity. The defined capacity limit is measured in millions of service units per hour (MSUs/h). WLM enforces the defined capacity limit by tracking the partition's CPU usage and continually averaging it over the past 4 hours. When this average (4h MSU Avg) goes beyond the defined limit, WLM starts to cap the partition. Remaining time until capping is calculated under the assumption that the system is going to be used as in the immediate past. If the partition is member of a capacity group, the value reported might not meaningful because the defined group limit is not considered. With Release 1.11, Monitor III computes a threshold which is derived from either the defined image capacity or group capacity. In other words, without the consideration of a group membership, the projection could indicate no capping at all while capping in fact is instantly going to be started.
- Remaining time until group capping is the projected time until WLM sets the capacity group into the capping state. That is, the usage of processor resources for one or more members of the group might be limited. WLM soft capping takes place to prevent you from using more than the defined group capacity limit over a long period of time. This is under the assumption that all group members (indicated by the 4h unused group capacity average) continue to use the system as they have done in the immediate past. The maximum number RMF reports is 14400 seconds or 4 hours. If RMF reports 14400, it means the remaining time until the group capping starts is at least 4 hours resp. group capping will never take place under the current circumstances



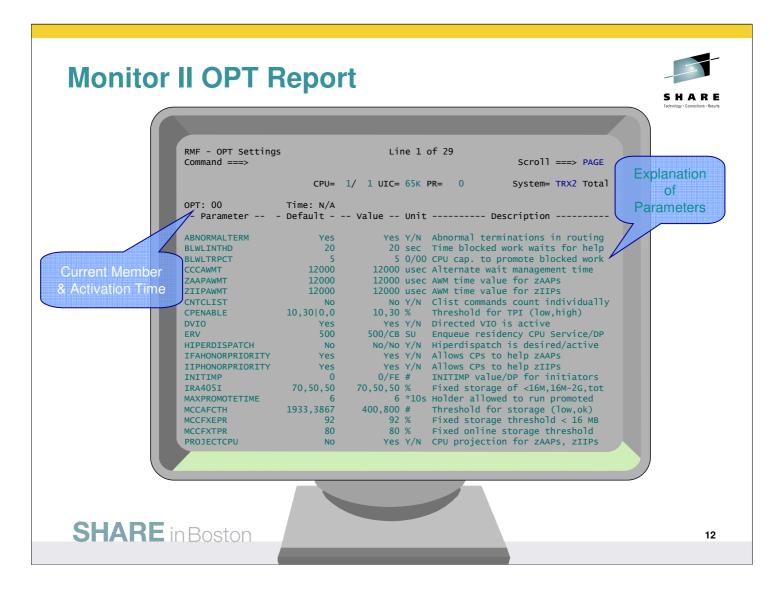
From the WLM website you can download the wlmqueue package which contains two different tools:
 WLMQUEUE
 WLM Work Queue Viewer, a small ISPF based tool that assists you in

displaying the application environments that are currently being used on your z/OS system

WLMOPT

the WLM OPT parameter viewer assists you in displaying the current OPT settings of your z/OS system

- With z/OS V1R11, the capabilities of the OPT parameter viewer are integrated in the RMF Monitor II report suite
- The new Monitor II OPT Settings report is invoked by:
  - the command OPT with no parameters.
  - the RMF Monitor II Primary Selection Menu, selection L.4
  - selection 4 of the RMF Monitor II Library List and OPT Settings Selection Menu



- The Monitor II OPT Settings report helps you to quickly obtain information about the current OPT member, when it was
   activated and the settings of all OPT parameters
- The report consists of the following columns:

• OPT	The suffix of active IEAOPTxx member in the PARMLB, containing the OPT parameter settings
• Time	Timestamp when the IEAOPTxx member was activated. If the system programmer did not change the active IEAOPTxx parmlib member since the last IPL, 'N/A' is shown
<ul> <li>Parameter</li> </ul>	The name of the WLM OPT parameter
Default	The default value(s) of the parameter. When more than one default exists, the values are separated by "/"
Value	The actual value(s) of the parameter. The actual value may differ from the value originally specified. When separated by "/" the second value is provided by SRM
• Unit	The unit of the parameter value
<ul> <li>Description</li> </ul>	Basic description of the purpose of the parameter. For detailed information refer to the z/OS MVS Initialization and Tuning Reference

## **Postprocessor XML Formatted Reports**



#### Rationale:

- RMF Postprocessor reports are limited to a page width of 132 characters
- No state-of-the-art display capability of Postprocessor reports
- No easy access to RMF Postprocessor data for application programs
  - $\Rightarrow$  cumbersome to parse the text output
  - ⇒ each report has its own layout





RMF Postprocessor reports can now be generated in XML Format



In the past, users of the RMF Postprocessor had to struggle with a couple of shortcomings:

- The standard text output was limited to to a width of 132 characters. Therefore some reports appeared somehow squeezed. This applies in particular to Overview Reports with a high number of OVW conditions
- The basic text format didn't exploit todays presentation capabilities like resizing windows, scrolling back and forth etc.
- Depending on the report type, the report layouts were quite different. Hence, API programs need to supply specific logic in order to extract metrics out of the reports

Starting with z/OS V1R11 RMF, the Postprocessor is able to generate a subset of the reports alternatively in XML format

#### **Postprocessor XML Formatted Reports** The following single-system reports are available in XML format (both Interval and Duration reports) ⇒ CPU Activity report (including Partition Data, LPAR Cluster and Group Capacity reports) CRYPTO Activity report FICON Director Activity report ⇒OMVS Kernel Activity report ⇒ ESS Disk Systems Activity report z/OS 1.12 ⇒ Device Activity report With z/OS 1.12, the first sysplex report is available in XML format z/OS 1.12 ⇒ Workload Activity report Overview reports are also available in XML format Summary and Exception reports are not available in XML format • Interval reports based on data collected during a Monitor II background session • are not available in XML format **SHARE** in Boston 14

- It is RMF's objective to introduce the XML format for the majority of Postprocessor reports over the next releases.
- With the first stage in RMF for z/OS 1.11, those reports have been selected where no similar Monitor III report is available since Monitor III reports are already XML enabled by means of the Data Portal. These are the following Postprocessor single-system reports:
  - CPU Activity report
  - CRYPTO Activity report
  - FICON Director Activity report
  - OMVS Kernel Activity report,
  - · ESS Disk Systems Activity report
- In addition, Postprocessor Overview reports belong also to the first stage since multiple Overview conditions have caused nasty line-feeds in the old text format
- RMF for z/OS 1.12 extends the list of Postprocessor XML reports with another single-system report and the frist sysplex report:
  - · Device Activity report (single-system report)
  - · Workload Activity report (sysplex report)

#### **Postprocessor XML Formatted Reports**



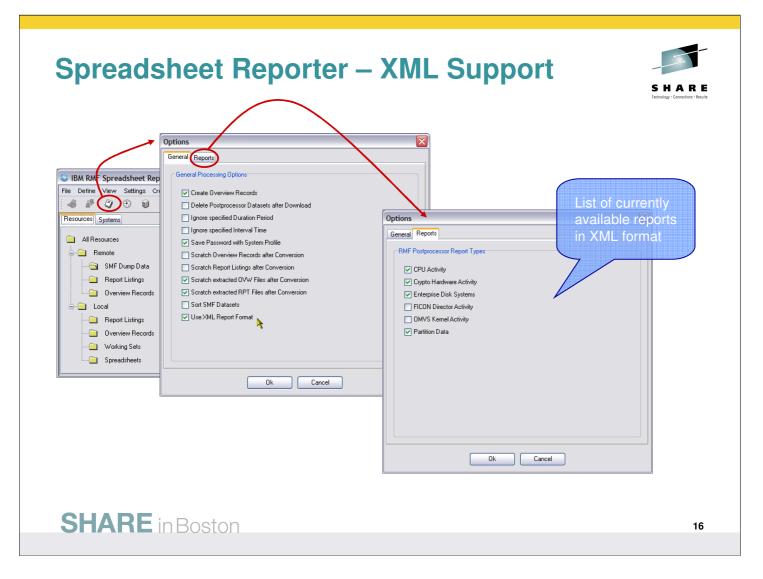
- The generation of Postprocessor reports in XML format is controlled by the new ddnames XPRPTS, XPXSRPTS and XPOVWRPT
- If the XML output is routed to permanent data sets rather than to SYSOUT, define the data set with RECFM=VB and LRECL between 256 and 8192. Specify an appropriate BLKSIZE.

ddname	Contents	Allocations	Notes
XPRPTS	Combined single-system report in XML format	One ddname for one data set to contain all single system reports for each interval during the session.	There is no dynamic allocation of this ddname, you have to define it explicitly if you want to get all reports in XML format into one data set or output class. If you define this ddname, no MFRnnnnn files are created. If you define this ddname and PPRPTS, no XML output in file XPRPTS is created.
XPOVWRPT	Combined Overview report in XML format	One ddname for one data set to contain all overview reports for each system included in the input data.	There is no dynamic allocation of this ddname, you have to define it explicitly if you want to get all overview reports in XML format into one data set or output class. If you define this ddname, no PPORPnnn files are created.
XPXSRPTS	Combined sysplex-wide report in XML format	One ddname for one data set to contain all sysplex reports for each interval included in the input data.	There is no dynamic allocation of this ddname, you have to define it explicitly if you want to get all reports in XML format into one data set or output class. If you define this ddname, no MFRnnnnn files are created. If you define this ddname and PPXSRPTS, no XML output in file XPXSRPTS is created.

#### SHARE in Boston

15

- The request for XML output from the RMF Postprocessor is triggered by means of the existence of a DD card for XPRPTS (for single-system reports) and/or XPOVWRPT (for Overview reports) in the JCL for the Postprocessor job.
- If one of these DD statements are defined in the JCL, then XML output will be generated.
- The following rules apply if requests for text and XML are given simultaneously:
  - If the user specifies a DD card for cumulative text output, i.e. PPRPTS, as well as XPRPTS, then only text output will be generated into PPRPTS and no XML is produced.
  - If single-system reports are requested on the REPORTS control statement which are not XML enabled yet and the DD card XPRPTS is defined, no output is generated for them.
- Beginning with RMF for z/OS 1.12 the new ddname XPXSRPTS was introduced to support XML formated Postprocessor sysplex report.
- For the XPXSRPTS ddname similar rules apply as for XPRTS, that is:
  - If the user specifies the PPXSRPTS ddname for cumulative text output together with XPXSRPTS, no XML output is generated.
  - If sysplex reports other than the WLMGL report are specified in the SYSRPTS control statement, no output is generated.



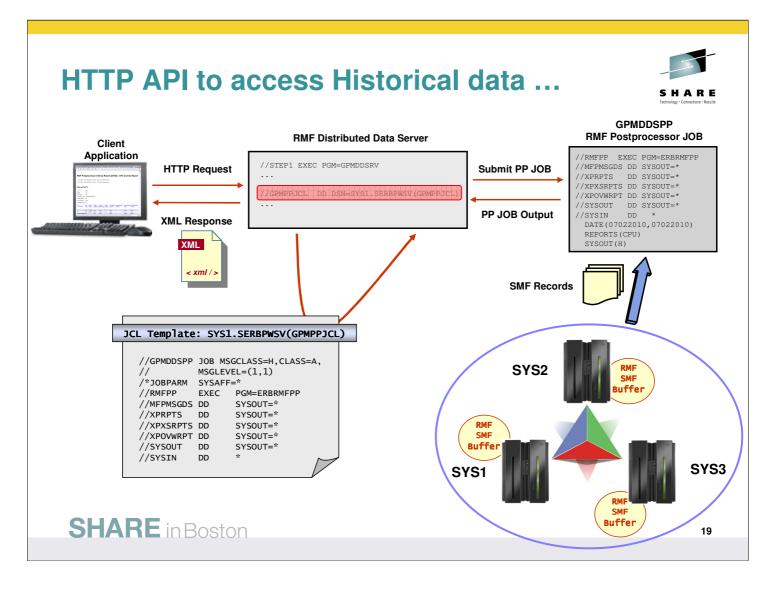
- The RMF Spreadsheet Reporter provides built-in support for the new Postprocessor XML formatted reports
- You can request the new XML format by means of the general option "Use XML Report Format". This will cause the following results:
  - The report selection list on the Options dialogs Reports tab displays just the report types which can be generated in XML format
  - Under the cover, the generated JCL will contain the new DD names for the XML format (XPRPTS and XPOVWRPT)
  - The default filetype for local report listings changes from .lis to .xml

									A RMF Spreadsheet	Reporter	Java TM Technology Edition	(SYSE)
									efine View Settings	•	<u>.</u>	
								8	# Ø 🖲 🗑	1	2	
C:\Documents	and So	ttings\Adm	hinistrator\Apr	plication Dat		F Spreadsheet	Reporter\RmfList	in - Microsoft I	nternet			
File Edit View		-				spreadsheet	Reporter Killicist	III - MICLOSOIC I			MVS1.D127.T112039.lis	
IC EUR VIEW	ruvon	10013	пар							-	BCLM.D215.T133732.lis	
											B SCLM.D215.T133808.lis	
											B SYSF.D212.T114702.xml	
	store	ocess	or Dura	tion Re	port	ISYSD1	CPU Act	ivity Rer	ort		SYSF.D215.T113455.xml	考 Start
											部 SYSF.D215.T113918.xml 해 Sample.lis	් View
z/OS V1R9 St	art: 07/3	1/2009-08.1	15.00 Interval:	000:44:59 hou	irs						2pr Sample.lis	New
												Rename
V1R11 RMF Er	1d: 07/31											
		12003-03.0	0.00 0,000.10		100							🧃 Delete
		1/2005-05.0	0.00 0,000 1		105							Delete Properties
		12003-03.0	o.oo oyac. n									-
	/ITY	1/2003-03.0	0.00 0,000 1									-
	/ITY 2097	12005-05.0	0.00 0,000.10									-
CPU		12003-03.0	o.ou oyuc. K									-
CPU Model	2097	12005-05.0	o.oo oyac. K									-
CPU Nodel H/W Model	2097 709 E26											-
CPU Model H/W Model Sequence Code	2097 709 E26 000000									-		-
CPU Nodel H/W Model Sequence Code	2097 709 E26											-
CPU Nodel HW Model Sequence Code HiperDispatch	2097 709 E26 000000 NO	0000070B8	32 Time% LPAR	Time% MVS	Time%	LOG PROC	HiperDispatch	VO Interrupts	I/O Interrupts%			-
CPU Nodel HW Model Sequence Code HiperDispatch	2097 709 E26 000000 NO	0000070B8	32			LOG PROC Share%	HiperDispatch Priority	1/O Interrupts Rate	I/O Interrupts% via TPI			-
CPU Aodel HW Model Sequence Code HiperDispatch PU Number	2097 709 E26 000000 NO CPU Type	D000070B2 Time% Online	32 Time% LPAR Busy	Time% MVS Busy	Time%	Share%		Rate	via TPI			-
CPU Aodel AW Model Sequence Code AliperDispatch PU Number	2097 709 E26 000000 NO CPU Type CP	000007088 Time% Online	32 Time% LPAR Busy 3.19	Time% MVS Busy 3.15	Time% Parked	Share% 28.7		Rate 9.81	via TPI			-
CPU Model 4W Model Sequence Code HiperDispatch PU Number	2097 709 E26 000000 NO CPU Type CP CP	D000070B2 Time% Online	32 Time% LPAR Busy 3.19 3.14	Time% MVS Busy 3.15 3.11	Time%	Share% 28.7 28.7		Rate 9.81 13.34	via TPI 0.25 0.61			-
CPU Model H/W Model Sequence Code HiperDispatch IPU Number	2097 709 E26 000000 NO CPU Type CP CP CP CP	00000070B8 Time% Online 100.00 100.00	32 Time% LPAR Busy 3.19 3.14 3.17	Time% MVS Busy 3.15 3.11 3.13	Time% Parked	Share% 28.7 28.7 57.4		Rate 9.81	via TPI			-
CPU Model HW Model Sequence Code HiperDispatch PU Number OTAL/AVERAGE	2097 709 E26 000000 NO CPU Type CP CP CP CP CP CP AAP	000007088 Time% Online	32 Time% LPAR Busy 3.19 3.14	Time% MVS Busy 3.15 3.11	Time% Parked	Share% 28.7 28.7		Rate 9.81 13.34	via TPI 0.25 0.61			-
CPU ACTIN CPU Model H/W Model Sequence Code HiperDispatch CPU Number PU Number OTAL/AVERAGE 2	2097 709 E26 000000 NO CPU Type CP CP CP CP CP CP AAP	00000070B8 Time% Online 100.00 100.00	32 Time% LPAR Busy 3.19 3.14 3.17 0.00	Time% MVS Busy 3.15 3.11 3.13 0.00	Time% Parked	Share%           28.7           28.7           57.4           14.7		Rate 9.81 13.34	via TPI 0.25 0.61			-

- The new XML formatted reports can be instantly displayed within the Spreadheet Reporter environment
- Once a listing is associated with the filtype *xml* the *View* action opens a browser window with the formatted report
- Report headers are displayed as name/value pairs while the report body is arranged in tabular format
- Internally the XML is visualized by means of the following stylesheets which are located in the Spreadsheet Reporters Listing directory:
  - ddsml-pp.xsl
  - ddsml.css

HTTP	API	to	<b>a</b>	CCE	ess	н	iste	orica	al d	ata		S. C.	SHARE Inchniege - Generations - Results
	ne o p	•	HTT All R Web	P API MF Po brows	to retri	eve F cessor be us	CMF Po XML ed as P	Distribut ostprocess formattec ostproces	sor XM l report ssor Dat	L reports s support ta Portal	ed	9S)	
	File Edit Viev	stpr	OCES 02/2010-11	s Help SOR Int	erval F	Repoi		D] : CPU			<b>*</b>		
	CPU ACTI CPU Model HW Model Sequence Code HiperDispatch CPC Capacity Change Reason	2097 722 E40 000000 NO 1631											
	CPU Number 0 1 TOTAL/AVERAG	CPU Type CP CP E CP	Time% Online 100.00 100.00	Time% LPAR Busy 3.50 3.52 3.51	Time% MVS Busy 3.47 3.48 3.47	Time% Parked	LOG PROC Share% 32.2 32.2 64.4	HiperDispatch Priority	I/O Interrupts Rate 4.61 5.07 9.68	VO Interrupts% via TPI 0.19 0.46 0.33	~		
SHAR	EinB	OS	ton										18

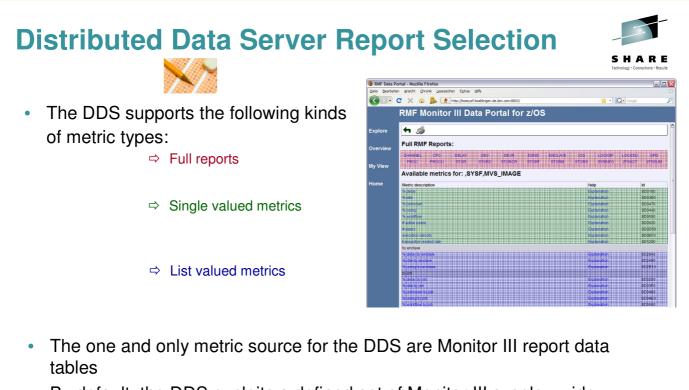
- In most installations, the access to historical data is needed for in depth performance analysis. This allows to keep track whether a critical situation has been persistent or not.
- The existing HTTP API of the RMF Distributed Data Server (DDS) already provides a sysplex-wide access of the data collected by RMF Monitor III. With RMF for z/OS 1.12 this API is extended by the capability to grant instant and easy access to RMF long-term historical data as reported by the RMF Postprocessor.
- Application programs can exploit the extended DDS HTTP API by sending standard URL requests for historical RMF Postprocessor data to the DDS.
- Since DDS returns the requested data as XML document, a web browser can act as an RMF Postprocessor Data Portal to access historical RMF data which is generated by the RMF Postprocessor in XML format.



 To get access to RMF Postprocessor data provided by the DDS, the GPMSERVE started task is extended with an additional DD card:

//GPMPPJCL DD DISP=SHR,DSN=SYS1.SERBPWSV(GPMPPJCL)

- This DD statement points to a JCL template which contains all JCL statements necessary to start a Postprocessor job from the RMF Distributed Data Server. The JCL template is stored in SYS1.SERBPWSV(GPMPPJCL).
- You must adapt or replace the GPMPPJCL member to suit your installation, ensuring that the DDS is able to run RMF Postprocessor jobs. If you do not want to request Postprocessor data with the DDS, you can omit the GPMPPJCL ddname.
- The dataset or the member specified in the GPMPPJCL DD card can be replaced with private ones containing the JCL template for running RMF Postprocessor from the DDS.
- The JCL template that is used by DDS to start the Postprocessor JOB does not contain any Postprocessor control statements. These will be dynamically created by DDS depending on the HTTP request parameters.
- For a detailed description of the DDS HTTP API please refer to the z/OS RMF Programmer's Guide.
- · Notes:
  - The Postprocessor API functionality is only available with JES2 installed.
  - The DD card specifications for the XML output datasets must not be changed. To be able to retrieve the Postprocessor output, DDS needs these SYSOUT data sets to reside on spool.



 By default, the DDS exploits a defined set of Monitor III sysplex-wide reports and single system reports

reperte and emgre eyetem reperte	<b>G</b> D- (	3 × 🗄	۵ 👗	http://boesysf.	boeblingen.de.lbm.	.com:8803/				😭 • [	• Google	P	8
		RMF N	lonitor	III Da	ta Port	al for	z/OS						
	Explore	<b>†</b> 🏄	j										^
	Overview	Full RM											=
	My View	CACHDET XCFSYS	CACHSUM	CFACT	CFOVER	CFSYS	SPACED	SPACEG	SYSSUM	XCFGROUP	XCFOVW	XCEPATH	
SHARE in Boston	ing them	Availabl	e metrics	for: ,SY	SDPLEX,S	SYSPLE	x						

- The performance data that can be retrieved by the RMF Distributed Data Server comes originally out of Monitor III report
  data tables
- · In order to produce these data tables, a defined set of reporter modules are called every Monitor III Interval
- This set includes the following Monitor III reports:
  - Sysplex-wide reports
    - CACHDET, CACHSUM
    - CFACT, CFOVER, CFSYS
    - SPACED, SPACEG
    - SYSSUM
    - XCFGROUP, XCFOVW, XCFPATH, XCFSYS
  - Single system reports
    - CHANNEL
    - CPC
    - DELAY
    - DEV, DEVR
    - DSND
    - ENCLAVE
    - 10Q
    - · LOCKSP, LOCKSU
    - OPD
    - PROC, PROCU
    - STOR, STORC, STORCR, STORF, STORM, STORS
    - SYSINFO
    - · ZFSACT, ZFSSUM

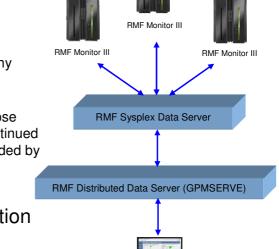
#### **Distributed Data Server Report Selection**



21

- The DDS retrieves Monitor III report data from all RMF instances in the Sysplex
- With z/OS V1.11 RMF it is possible to exclude specific Monitor III reports from being retrieved by the DDS
  - Users can exclude reports which are not required by any monitoring clients in their environment
  - In case of abnormal termination of reporter modules, operation of the DDS is affected. With this function, those reporters can be excluded temporarily. This allows continued operation of the DDS until a fix for the problem is provided by IBM
- A new DDS option is introduced for PARMLIB member GPMSRVxx. This option defines a list of Monitor III reports to be excluded from DDS processing

#### SHARE in Boston



Why do we care about report types and DDS internals?

- In large enviroments, it can be reasonable to reduce the DDS processing to a minimum. This can be achived by focussing onto those reports which are definitely needed
- In the rare event of an reporter module abend, this reporter module can be excluded from the calling list until a solution is available

#### **Distributed Data Server Report Selection**



22

- New option EXCLUDE\_REPORTS in DDS parmlib member
- The option is not mandatory
- If not specified, all Monitor III reports supported by the DDS are available
- If specified, the option contains a comma-separated list of report names which should be excluded from DDS processing
- Specify multiple options if the list of reports does not fit on one line
- In this example, the Cache Detail, Cache Summary, zFS Activity, and zFS Summary reports are excluded

'* '* NAME:	GPMSRV00	
* NAME: *	GEMSKVUU	
	: PARMLIB MEMBER FOR THE RMF DISTRIBUTED	
*	DATA SERVER HOST ADDRESS SPACE (GPMSERVE)	
*		
* COPYRIGHT:	LICENSED MATERIALS - PROPERTY OF IBM	
*	"RESTRICTED MATERIALS OF IBM"	
*	5694-A01	
*	(C) COPYRIGHT IBM CORP. 1998, 2009	
*		
···	STATUS=HRM7760	**
	******	**
'*************************************	******	
'*************************************	**************************************	**
'*************************************	orts to be excluded	**
'*************************************	<pre>************************************</pre>	**
**************************************	orts to be excluded	**
**************************************	<pre>************************************</pre>	**
**************************************	<pre>************************************</pre>	**
**************************************	<pre>************************************</pre>	**

- The GPMSRV PARMLIB member will be extended to include an additional option. The definition of this option is not mandatory. If it is omitted the DDS will act just like before, i.e. all RMF Monitor III reports will be available
- The option contains a list of report names separated by comma. Because of the number of reports these lists can easily exceed the maximum length of a single line in the PARMLIB member. Therefore multiple option entries can be specified
- Unavailable (release-specific extensions) or misspelled report names will produce a trace message, saying that the report name is invalid. Defining a report name multiple times in the option lists, will produce the same message. The spelling of all report names is case-insensitive

### **Distributed Data Server Report Selection**



 When a client requests a report from the DDS that has been deactivated by EXCLUDE\_REPORTS the following message is generated:

GPM0737I The requested report <reportname> is deactivated

 When a client requests a metric from the DDS that is based on a deactivated report the following message is generated:

GPM0738I The requested metric is based on the deactivated report <reportname>

Never deactivate the report types needed by other components !!

#### ▲ Caution:

- ➡ z/OS Capacity Provisioning requires RMF DDS and CIM server to obtain performance data from the CPC, SYSINFO and SYSSUM report.
- OMEGAMON XE on z/OS can be optionally configured to require RMF DDS to obtain performance data from the following reports: CFACT, CFOVER, CFSYS, XCFOVW, XCFGROUP, XCFPATH, XCFSYS, LOCKSP, LOCKSU

- The DDS is not only used by internal RMF UI's. Other components like Capacity Provisioning and optionally OMEGAMON XE on z/OS are using the DDS as data source as well.
- Please consider that you should not deactivate the report types that those components are using:
  - Capacity Provisioning: CPC, SYSINFO, SYSSUM
  - OMEGAMON XE on z/OS: CFACT, CFOVER, CFSYS, XCFOVW, XCFGROUP, XCFPATH, XCFSYS, LOCKSP, LOCKSU

## **Work Unit Queue Distribution**

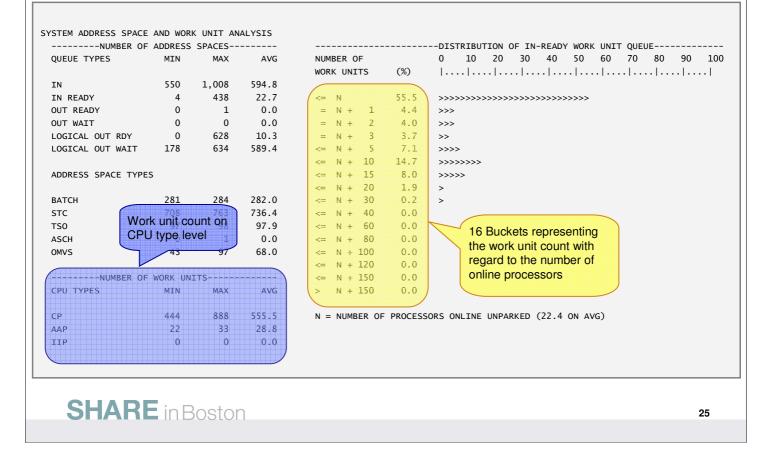


- Problem
  - RMF reporting for CPU contention was based on address spaces level
  - · Did not consider multiple work units (WEBs) within one address space
  - Performance analysts need statistics about the CPU contention on WEB granularity
- Solution
  - The RMF CPU Activity reporting is enhanced
- Benefit
  - The enhanced CPU Activity reporting helps the customer to obtain information about the In-Ready distribution based on WEBs and the number of work units per CPU type

- In the past, RMF has reported the number of address spaces waiting for or using a CPU (see In-ready statistics within the System Address Space Analysis of the Postprocessor CPU Activity report). However, with more than one task per address space, this does not really reflect the demand for CPU. Performance analysts need to know how many work units (WEBs) are running or waiting for a CPU.
- For this reason the CPU Activity report has been extended and new Overview Conditions have been implemented.

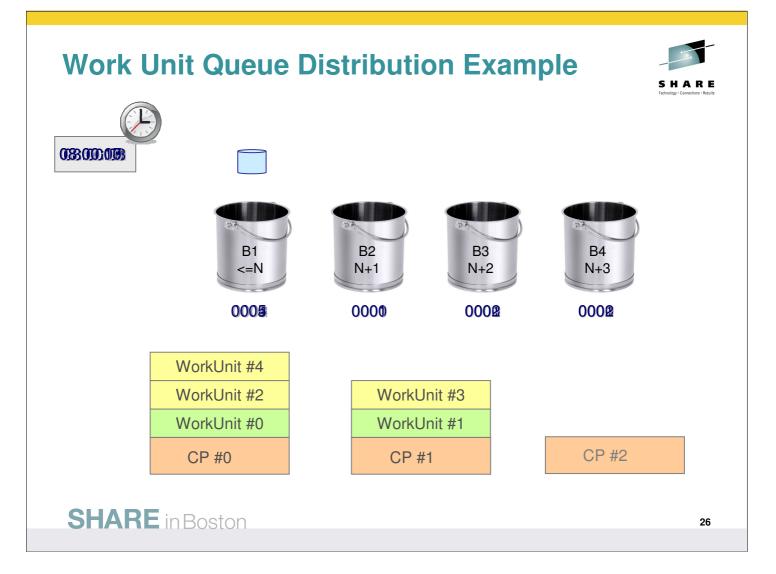
#### Work Unit Queue Distribution: Monitor I CPU Report



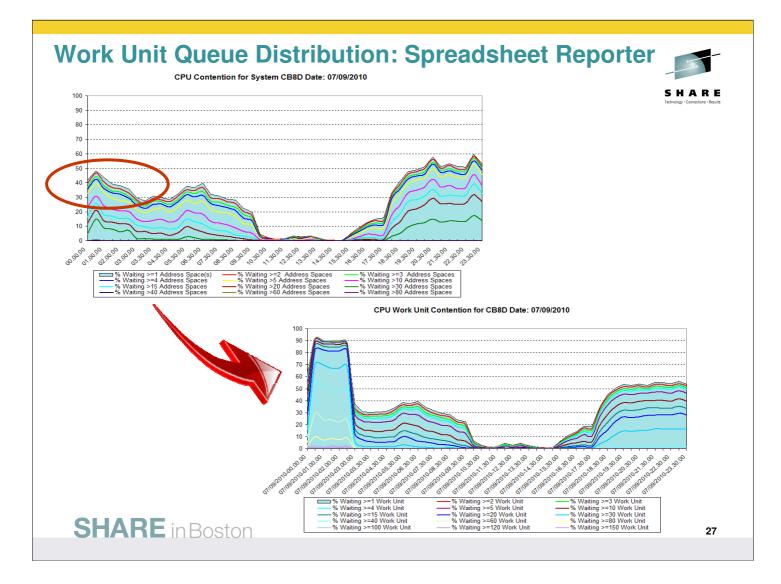


The SYSTEM ADDRESS SPACE section of the Monitor I CPU Activity Report has been changed as follows:

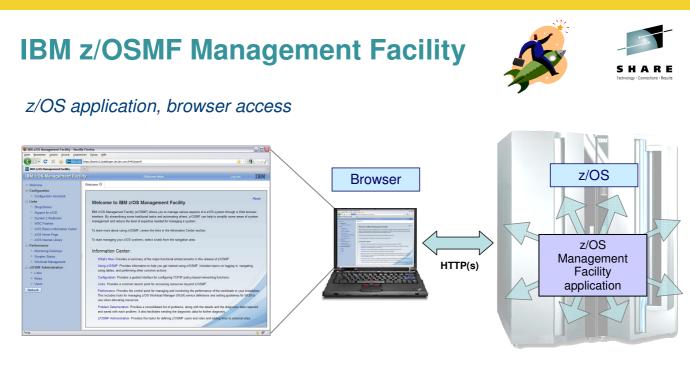
- · The In-Ready Queue Distribution is now based on the real number of dispatchable work units
  - We display 16 buckets of ascending queue lengths with regard to the number of actual online processors when the sample was taken
  - This allows to detect latent demand for the processor resource and to distinguish uncritical spikes from severe shortages
  - · The number of work units within the buckets includes standard CPs as well as specialty engines
- The new NUMBER OF WORK UNITS statistics allow to determine the CPU demand on processor type level
  - The MIN, MAX and AVG queue lenghts are shown for CPs, AAPs and IIPs



- This example points out the concept of buckets with regard to the number of online processors. For simplification reasons, there are just four buckets (instead of 16) and five work units as maximum
- · As long as there are less or equal work units than processors, all these samples contribute to bucket number one
- · In case there are more work units than processors the corresponding bucket is increased by one for each sample
- When an additional processor becomes available proposed that the number of work units remains unchanged the leftadjacent bucket will be increased



• The new **OneWebCont** sheet in the System Overview Spreadsheet of the RMF Spreadsheet Reporter allows to analyze cpu contention on work unit granularity.



- z/OS Management Facility is a Web 2.0 application on z/OS
  - Manages z/OS from z/OS
  - Browser communicates with z/OSMF via secure connection, anywhere, anytime

SHARE in Boston

28

- IBM z/OS Management Facility (z/OSMF), a new zero priced product, will simplify, optimize and modernize the z/OS system programmer experience starting in V1R11 with problem data management and TCP/IP Policy based configuration. With z/OS V1R12 the Workload Management and the Resource Monitoring Application have been added
- z/OSMF will deliver solutions in a task oriented, Web browser based user interface with integrated user assistance. And z/OSMF will make the day to day operations and administration of the mainframe z/OS systems easier to manage for both new and experienced system programmers. The focus is to help improve system programmer productivity, and make the functions easier to understand and use
- This chart explains the structure for z/OSMF and how it fits into the z/OS environment.
- z/OSMF is on the right hand side and it manages z/OS from z/OS itself. It is not an external application, nor does it have an external client. z/OSMF is an application on z/OS with direct access to z/OS data and information, and it has a browser interface from the workstation. z/OSMF contains the GUIs and the application code. Everything is installed on the z/OS server and there are no client side install requirements.
- In the middle of the screen is a workstation with a browser and it communicates with z/OSMF via HTTP(s). z/OSMF is a Web 2.0 based solution. It incorporates a browser interface that communicates with the z/OS system. The browser can be anywhere... in the data center ... or around the world. You just need a secure connection.
- And on the left is a screen capture of the z/OSMF welcome page once you log into z/OSMF.

## IBM z/OSMF Management Facility ...



#### Applications / R12 plugins

IBM z/OS Management Facility Ele Edit View History Bookmarks     C	<ul> <li>Configuration category</li> <li>Configuration Assistant for z/OS Communication Server – Simplified configuration and setup of TCP/IP policy-based networking functions (with z/OS V1.11)</li> </ul>
Welcome Configuration Configuration Assistant Links ShopzSeries Support for z/OS	Links to resources - provides common launch point for non-z/OSMF resources
System z Redbooks     WSC Flashes     z/OS Basics Information Center     z/OS Home Page     z/OS Internet Library     Performance     Sysplex Status     Monitoring Desktops     Workload Management     Problem Determination     Incident Log     z/OSMF Administration	<ul> <li>Performance category (with z/OS V1.12)</li> <li>Sysplex status – single view of sysplex performance</li> <li>Monitor desktops – dynamic real time metrics for system performance, customizable for correlations</li> <li>Workload Manager Policy Editor Simplified creation, editing, installation, and activation of WLM service definitions and policies</li> </ul>
= Links = Roles = Users Refresh	<ul> <li>Problem Determination category</li> <li>Incident Log Simplified capture, packaging, and sending of SVC dump diagnostic data. (Available with z/OS V1.10)</li> </ul>
SHARE in	• z/OSMF Administration Authorization services, add users, define roles, add links.
SHARE	n Boston 29

- Once you've logged in , this screen shows you the full scope of what z/OSMF provides in this first release. And really the first user that logs in as a z/OSMF administrator. So when z/OSMF is first set up, always the first ID is that of an administrator – that is a requirement for setup. And the reason for this it allows the first person to get in and to add and enable others.
- So this is the welcome screen for a user or an administrator. You'll notice that the navigation pane on the left is populated with multiple categories with tasks under each category. You have the configuration category with the Configuration Assistant, the Links category with a bunch of links, Problem Determination with incident log, and z/OSMF Administration.
- The initial functions include z/OSMF Administration, Incident Log, Configuration Assistant for z/OS Communication Server and Links (they are in alphabetical order)
  - <u>Configuration Assistant from Communication Server</u>
    - Simplified configuration and setup of TCP/IP policy-based networking functions
  - Another category on the navigation bar is Links which:
    - Provides common launch point for accessing resources beyond the IBM z/OS Management Facility.
    - Some links are pre-defined in the product.
    - The administrators can define additional links to share commonly used resources for their installation.
  - Performance category (R12) Workload Manager Policy Editor
    - Facilitate the creation and editing of WLM service definitions, installation of WLM service definitions, and activation of WLM service policies
  - Performance category (R12) Resource Monitoring
    - The *z/OSMF Resource Monitoring* plugin allows cross-sysplex performance monitoring from a single point of control. From the z/OSMF task tree, you can select the following subtasks:
      - The Sysplex Status task provides an enterprise-wide health check of all z/OS sysplexes.
      - For further analysis, the *Monitoring Desktops task* can graphically display RMF Monitor III metrics as well as Linux metrics by means of customizable views.
  - Under the Problem Determination category on the navigation bar, you will find the Incident Log task.
    - This first release will help all system programmers with problem data management tasks, providing experienced teams with
      procedural advantages through an incident log summary and detail views of z/OS dump incidents. The Incident Log provides a
      consolidated list of SVC Dump related problems, along with details and diagnostic data captured with each incident. It also
      facilitates sending the data for further diagnostics.
  - z/OSMF Administration gives the customer a:
    - Display welcome page and deployed z/OSMF applications in Task list
    - Provides Authorization services for the administrator
    - Allows Dynamic addition of links to non-z/OSMF resources
- When the user logs in with the administrator's userid, all tasks are available to work with.
- For all other users, after initial login, the user becomes an 'Authenticated Guest' and can access information about z/OSMF and some permitted tasks.



- The starting point for the monitoring is the *Sysplex Status* task. This task provides a comfortable way to assess the health status of all systems in your installation at a glance.
- The table contains the list of z/OS sysplexes and Linux images. You can add, modify and delete items in the table using the *Actions* menu.
- When you start the task for the first time, one default entry is provided as *LOCALPLEX* that points to the DDS in the sysplex in which z/OSMF is running. If you have a running DDS in this sysplex the *Sysplex Status* task detects it automatically without requiring an explicit declaration of its host name or IP address. Also, the *Sysplex Status* task can automatically detect the DDS movement within the sysplex during the z/OSMF session.
- In addition to the LOCALPLEX you can add all target sysplexes you want to monitor to the table. Now let's explore the column output as shown in the chart:
  - The PI-Status column gives you a red-yellow-green indicator for the sysplex health. This indicator is based on the WLM service class period goals and actuals.
  - If all service class periods on the system are meeting their goals (that is, have a *performance index* (PI) of less than or equal to 1), the PI Status is green.
  - If service classes with importance of 3, 4, and 5 with the PI greater than 1 are detected, the indicator is yellow.
  - If at least one important service class (that is, importance of 1 or 2) misses the WLM goal, the indicator is red
- In this case it becomes essential to figure out the reasons by drilling down into the details on the respective sysplex. This can be done within the *Monitoring Desktops* task.



- A *Monitoring Desktop* is a customizable view containing different performance metrics that you can group and arrange flexibly.
- You can create and save your own desktops or open and modify the predefined desktops that are with the task in z/OSMF.
- · When you start a Monitoring Desktop it begins retrieving the online data from the DDS periodically.
- · By default, the most current snapshot is displayed
- · It's possible to browse through the data collected since the start of the current session.
- A desktop contains multiple metric groups; their size and column count are adjustable. You add one or multiple metrics to each group.
- Now let's create a new *Monitoring Desktop*. Suppose, you want to inspect the common storage usage and identify jobs with the highest storage consumption.
  - To do this, you can create a new desktop using the drop down menu on the Desktops tab. You can now add a
    metric you want to monitor to the desktop by using the "Add Metric" dialog.
  - The first thing you need to select in this dialog is a resource you're interested in. The resources are presented in a tree structure.
  - Starting with the sysplex as top level resource, you can drill down into the contained resources, for example, the MVS images or coupling facilities.
  - Once you select a resource (for example, storage on a system), you can select a performance metric that you want to monitor.
  - The metrics of the selected resource are also presented in a tree. For a better navigation, you can use the quick-filter feature.
- Resource Monitoring provides sophisticated filtering capabilities:
  - Suppose you select the metric % CSA utilization by job, which will list all the jobs active on the system and their CSA consumption in %.
  - Because the number of jobs in the system can be pretty high, you can specify filters and work scopes: that is, display only jobs that match a certain name pattern, belong to a specific WLM service class, or meet a threshold value.



# <section-header><section-header><section-header><section-header><section-header><section-header><list-item><list-item><list-item><list-item><section-header><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item>



# **Appendix: Function Reference**



Function	Availability
Enhanced Group Capacity Reporting	z/OSV1.11 RMF
Monitor II OPT Report	z/OSV1.11 RMF
Postprocessor XML Formatted Reports	z/OSV1.11 RMF
HTTP API to access Historical data	z/OSV1.12 RMF
Distributed Data Server Report Selection	z/OSV1.11 RMF
In-Ready Work Unit Queue Distribution	z/OSV1.12 RMF
IBM z/OSMF Management Facility: Resource Monitoring Plugin	z/OSV1.12 zOSMF

SHARE in Boston

35