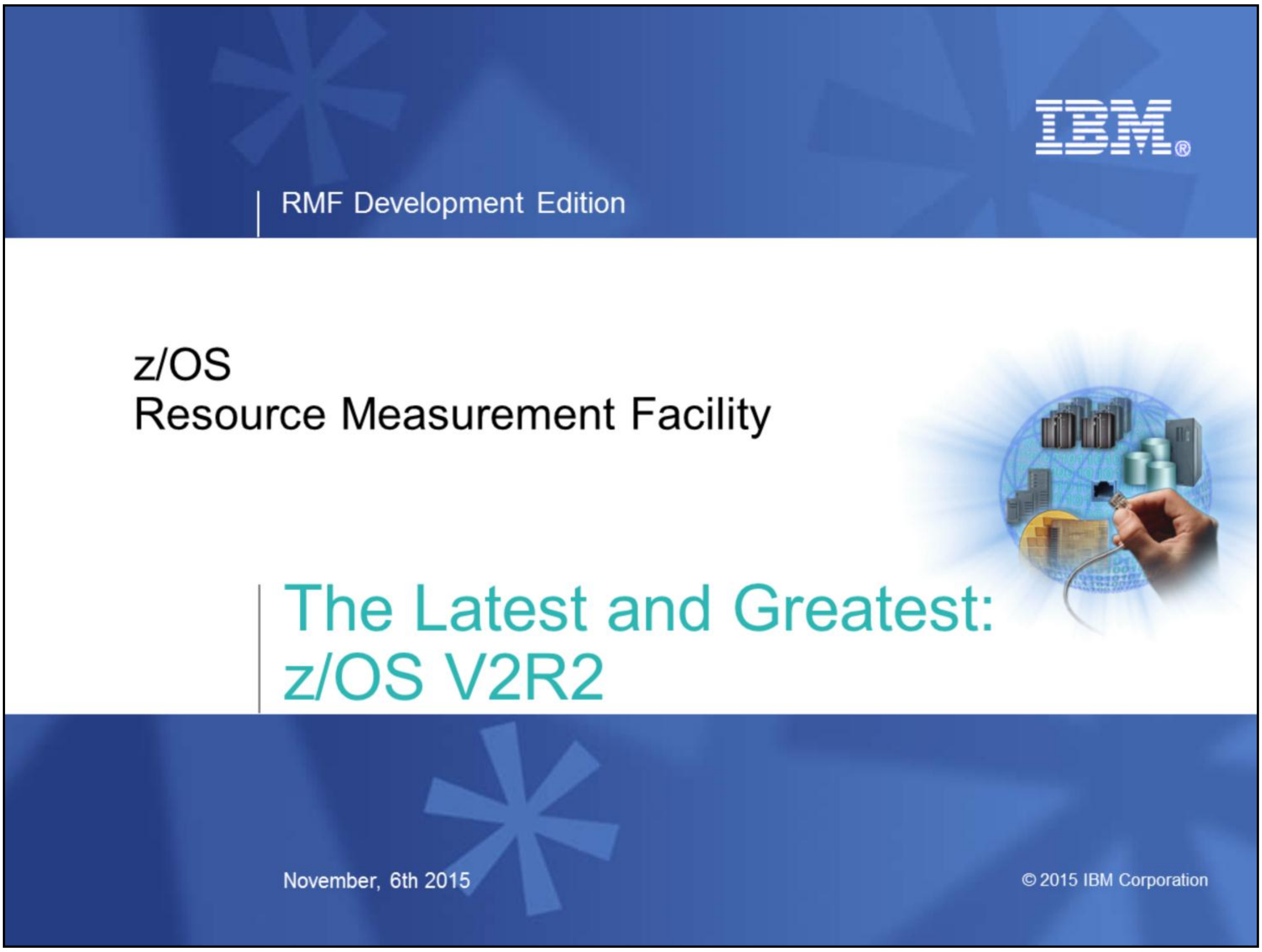
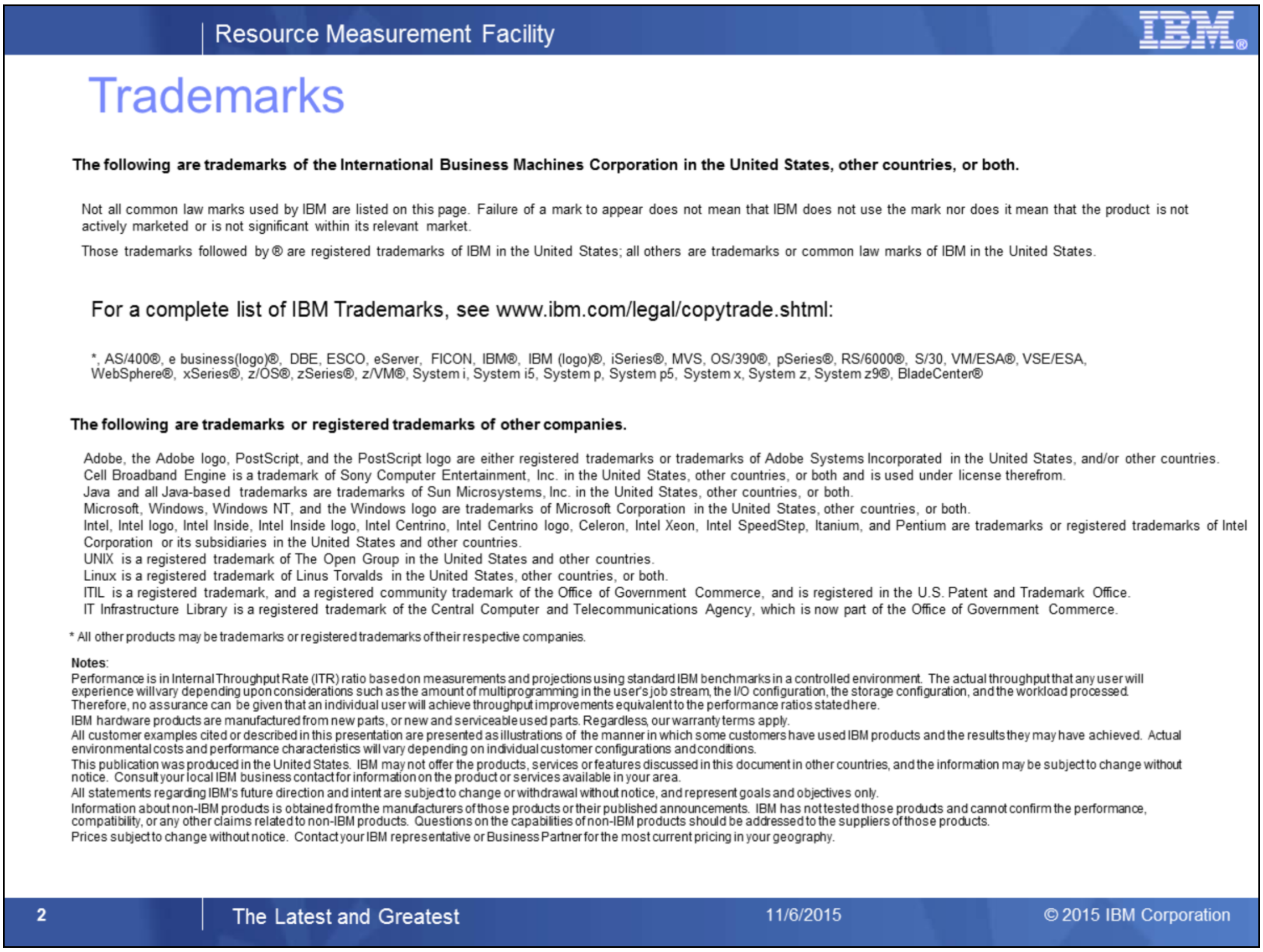


RMF Development Edition

z/OS
Resource Measurement Facility

The Latest and Greatest: z/OS V2R2

November, 6th 2015



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:

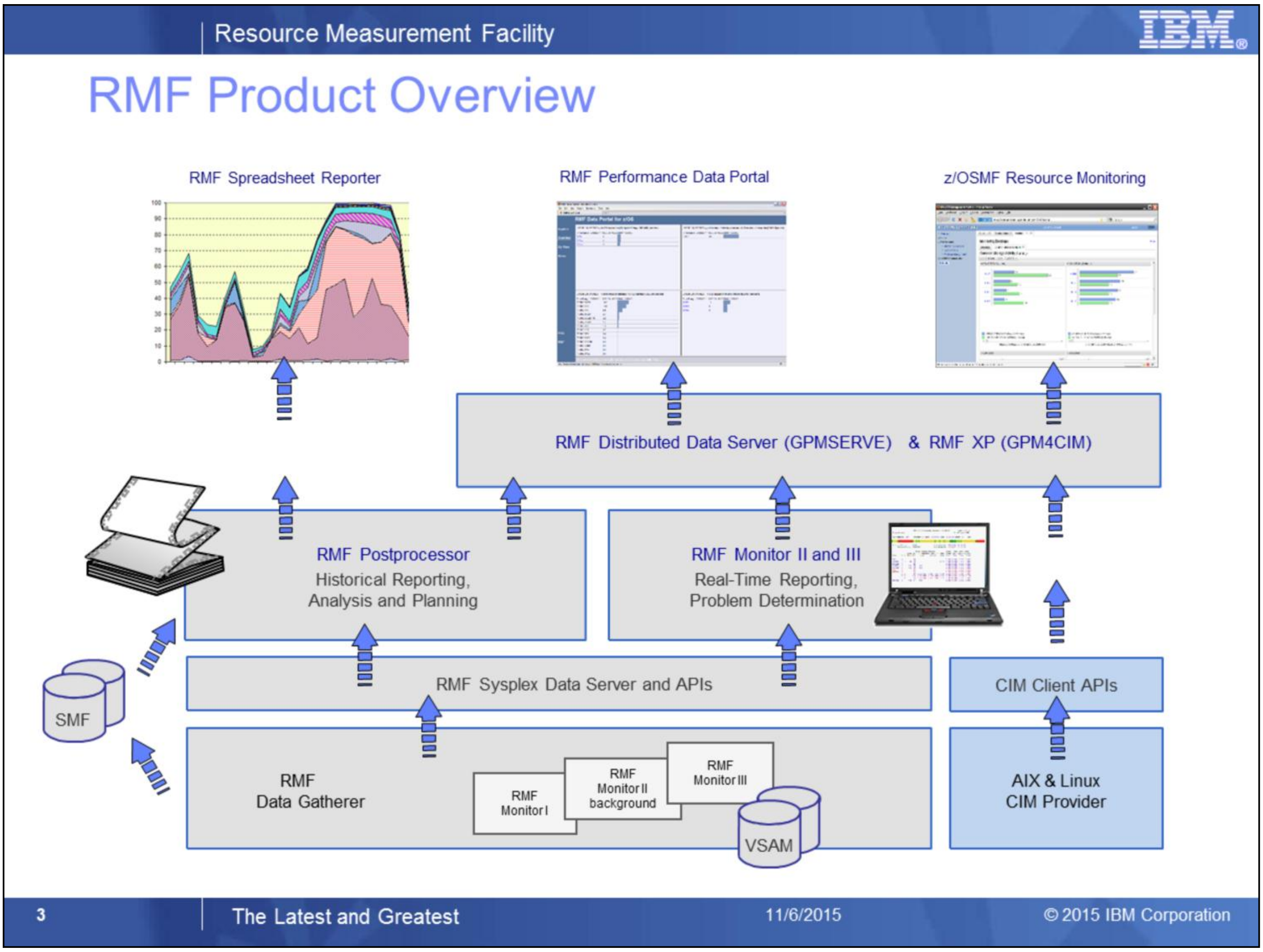
*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

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 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.



RMF Product Overview

RMF Spreadsheet Reporter

RMF Performance Data Portal

z/OSMF Resource Monitoring

RMF Distributed Data Server (GPMSERVE) & RMF XP (GPM4CIM)

RMF Postprocessor
Historical Reporting,
Analysis and Planning

RMF Monitor II and III
Real-Time Reporting,
Problem Determination

SMF

RMF Sysplex Data Server and APIs

CIM Client APIs

RMF
Data Gatherer

RMF
Monitor I

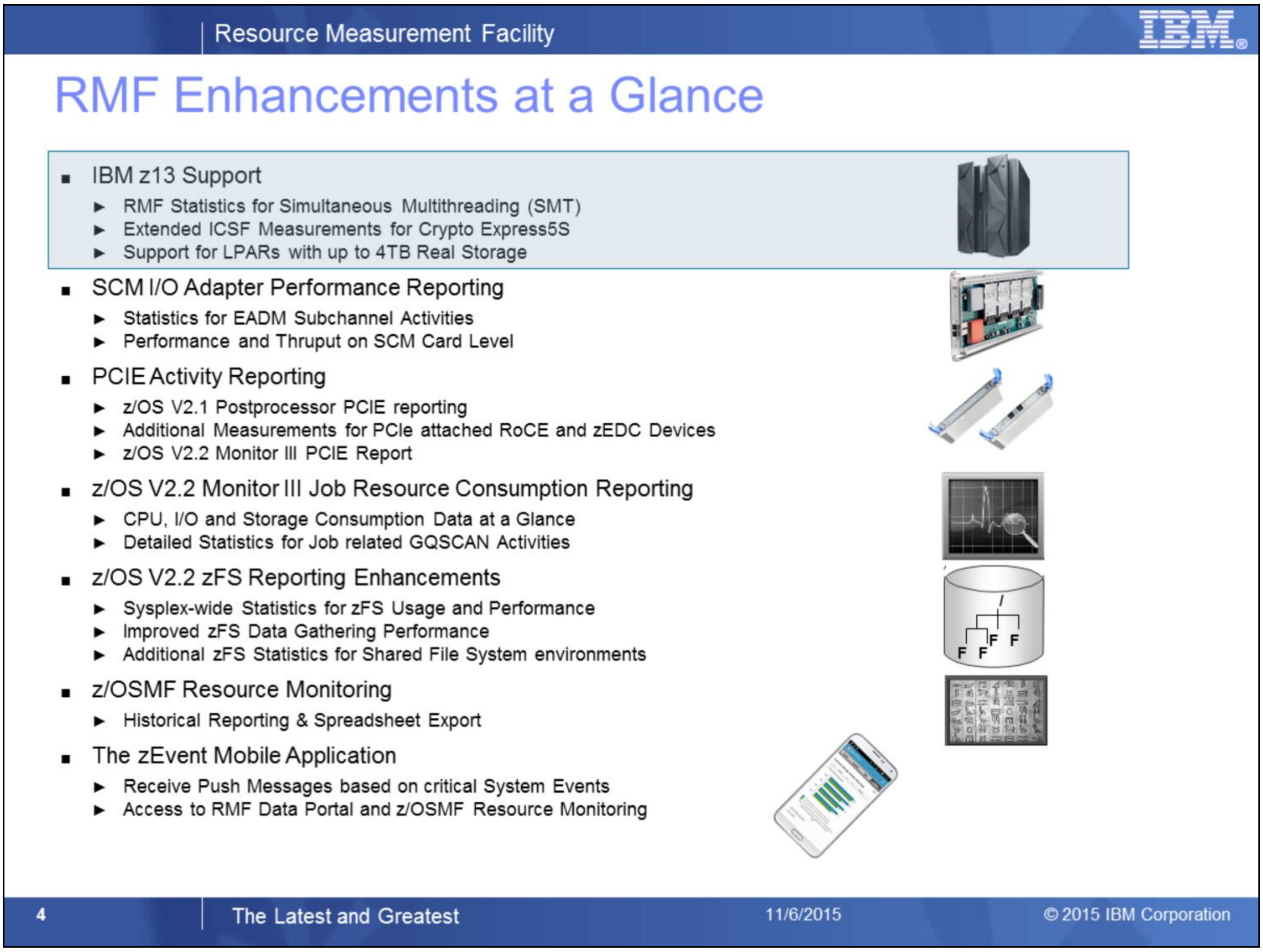
RMF
Monitor II
background

RMF
Monitor III

VSAM

AIX & Linux
CIM Provider

- 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). Beginning 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 RMF Postprocessor data.
- Since z/OS 1.12 there’s another exploiter of the RMF DDS data: The z/OSMF Resource Monitoring plugin of the z/OS Management Facility.
- RMF for z/OS 1.13 enhances the DDS layer with a new component:
  - RMF XP is the new solution for Cross Platform Performance Monitoring
  - Provides a seamless performance monitoring for all operating systems running on the IBM zEnterprise Bladecenter Extension.



RMF Enhancements at a Glance

- IBM z13 Support
  - RMF Statistics for Simultaneous Multithreading (SMT)
  - Extended ICSF Measurements for Crypto Express5S
  - Support for LPARs with up to 4TB Real Storage
- SCM I/O Adapter Performance Reporting
  - Statistics for EADM Subchannel Activities
  - Performance and Thruput on SCM Card Level
- PCIE Activity Reporting
  - z/OS V2.1 Postprocessor PCIE reporting
  - Additional Measurements for PCIe attached RoCE and zEDC Devices
  - z/OS V2.2 Monitor III PCIE Report
- z/OS V2.2 Monitor III Job Resource Consumption Reporting
  - CPU, I/O and Storage Consumption Data at a Glance
  - Detailed Statistics for Job related GQSCAN Activities
- z/OS V2.2 zFS Reporting Enhancements
  - Sysplex-wide Statistics for zFS Usage and Performance
  - Improved zFS Data Gathering Performance
  - Additional zFS Statistics for Shared File System environments
- z/OSMF Resource Monitoring
  - Historical Reporting & Spreadsheet Export
- The zEvent Mobile Application
  - Receive Push Messages based on critical System Events
  - Access to RMF Data Portal and z/OSMF Resource Monitoring

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

With the availability of the IBM z13 servers, RMF provides first day support for a couple of notable hardware features.

- Comprehensive Statistics for Simultaneous Multithreading (SMT)
- Extended ICSF Measurements for Crypto Express5S
- Support for LPARs with up to 4TB Real Storage

Storage Class Memory – aka Flash Memory – is a new tier within the memory hierarchy of the zSeries family.

RMF provides detailed usage statistics for Storage Class Memory related operations by means of a new Monitor III SCM Activity report.

Together with the z EC12 server family two new types of PCIe cards have been introduced:

- The new z Enterprise Data Compression (zEDC) Express offering provides low-cost data compression for z/OS system services and applications
- Shared Memory Communication via Remote Direct Memory Access (SMC-RDMA or SMC-R) is a zEC12 feature that provides high performance CPC to CPC communication (similar than Hipersockets for LPAR to LPAR Communication)
- The existing Postprocessor PCIe Report has been extended with additional measurements
- With z/OS V2R2 RMF a complete new Monitor III PCIe Report allows to keep track of PCIe related operations in online mode

The Monitor III Job Usage Report complements the Monitor III reporting suite with detailed statistics about address space resource consumption.

- The top resource consumers in terms of CPU, I/O and Storage can now be identified at a glance
- The report can serve as an excellent starting point for further drill-down and analysis
- Job related GQSCAN activities have been invisible in the past. With the new report detailed statistics with regard to GQSCAN usage can now be obtained

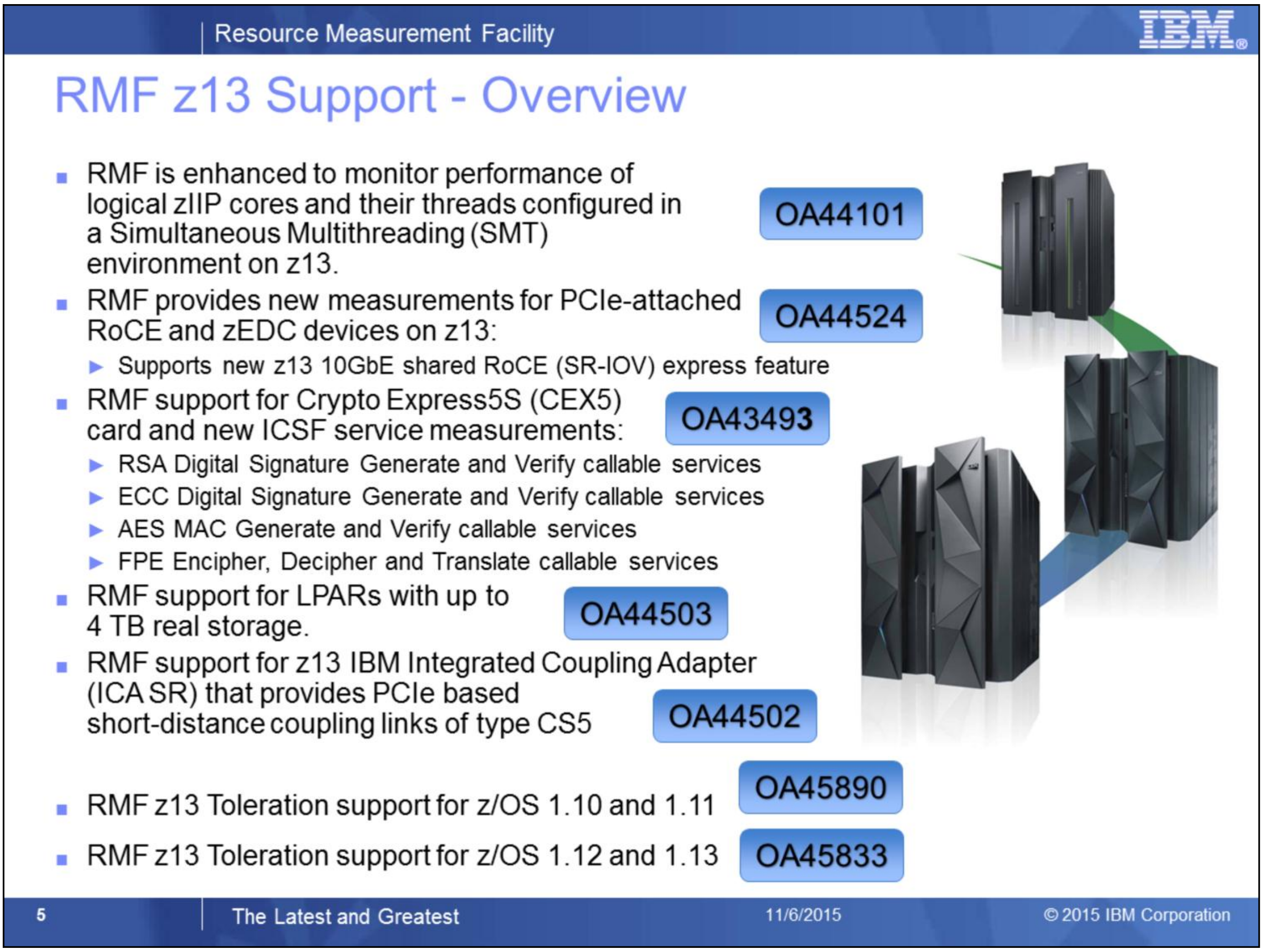
RMF z/OS 2.2 introduces new Monitor III Sysplex reports to monitor sysplex-wide z/OS Distributed File system (zFS) usage and performance.

z/OSMF Resource Monitoring is the strategic state-of-the-art frontend for RMF Monitor III performance data. Quite recently the following features have been added:

- Collect and display historical performance data on individual metric group level
- Export the data contained in the dashboards in CSV format for later usage in a spreadsheet

Last not Least the zEvent Mobile Application:

- This mobile App is currently under development and all details and capabilities – including the application name – are subject to change
- In a nutshell the app provides the following two main features to system administrators and performance analysts:
- Receive push messages based on critical system events instantly
- Access to the RMF Data Portal and z/OSMF Resource Monitoring anywhere and every time



RMF z13 Support - Overview

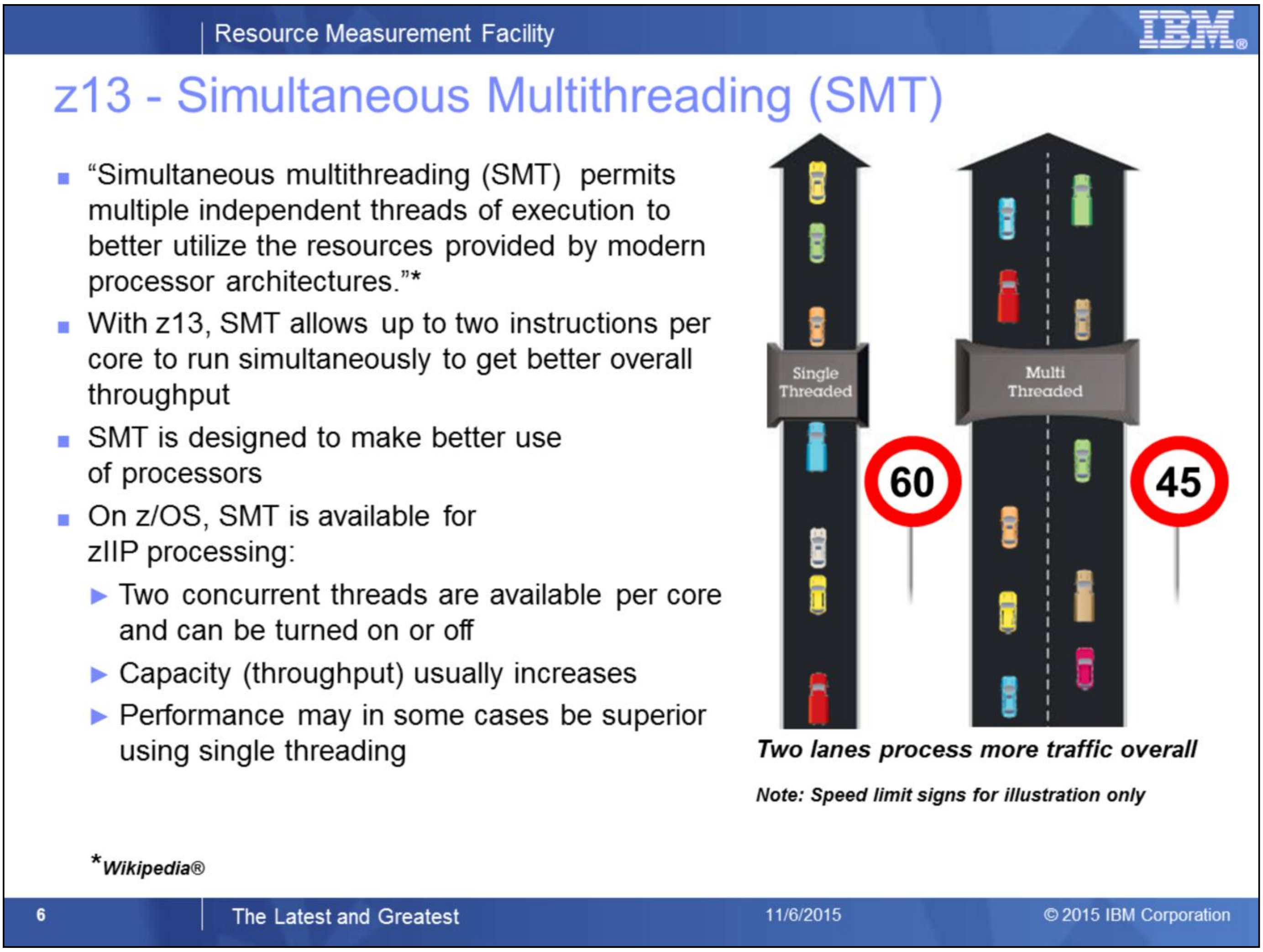
- RMF is enhanced to monitor performance of logical zIIP cores and their threads configured in a Simultaneous Multithreading (SMT) environment on z13. **OA44101**
- RMF provides new measurements for PCIe-attached RoCE and zEDC devices on z13: **OA44524**
  - Supports new z13 10GbE shared RoCE (SR-IOV) express feature
- RMF support for Crypto Express5S (CEX5) card and new ICSF service measurements: **OA43493**
  - RSA Digital Signature Generate and Verify callable services
  - ECC Digital Signature Generate and Verify callable services
  - AES MAC Generate and Verify callable services
  - FPE Encipher, Decipher and Translate callable services
- RMF support for LPARs with up to 4 TB real storage. **OA44503**
- RMF support for z13 IBM Integrated Coupling Adapter (ICA SR) that provides PCIe based short-distance coupling links of type CS5 **OA44502**
- RMF z13 Toleration support for z/OS 1.10 and 1.11 **OA45890**
- RMF z13 Toleration support for z/OS 1.12 and 1.13 **OA45833**

With various new function APARs, RMF exploits the new functionality of the IBM z13:

• OA44101:	RMF support for the Simultaneous Multithreading (SMT) environment on z13 PTF available for z/OS 2.1
• OA44524:	RMF PCIE enhancements for RoCE and zEDC devices on z/OS 2.1.
• OA43493:	RMF support for the Crypto Express5S (CEX5) card and new ICSF service measurements. The support is available for z/OS 1.13 and z/OS 2.1.
• OA44503:	RMF support for z/OS 2.1 LPARs on z13 with up to 4TB real storage.
• OA44502:	RMF support for z13 IBM Integrated Coupling Adapter (ICA SR) that provides PCIe based short-distance coupling links of type CS5. PTFs available for z/OS 1.13 and z/OS 2.1.

RMF tolereation support for IBM z13:

• OA45890:	z13 toleration for z/OS 1.10 and z/OS 1.11
• OA45833:	z13 toleration for z/OS 1.12 and z/OS 1.13 .



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

Note: Speed limit signs for illustration only

**Wikipedia®*

Simultaneous multithreading (SMT) allows two active instruction streams (threads) per core, each dynamically sharing the core's execution resources. SMT will be available in IBM z13 for workloads running on the Integrated Facility for Linux (IFL) and the IBM z Integrated Information Processor (zIIP).

SMT utilizes the core resources more efficiently: When a thread running on a core encounters a cache miss and can no longer make progress, the core switches to run a different thread that is ready to execute.

Each thread runs slower than a non-SMT core, but the combined 'threads' throughput is higher. The overall throughput benefit depends on the workload.

Basic idea of the SMT approach: better exploitation of processor resources with out cloning all peripheral components of a physical CPU, e.g. Cache, Register(?)

The SMT mode might not be suited for single threaded workloads (typically batch workloads).

Highway example: More overall thruput on a physical core, but speed on the individual lane is lower:

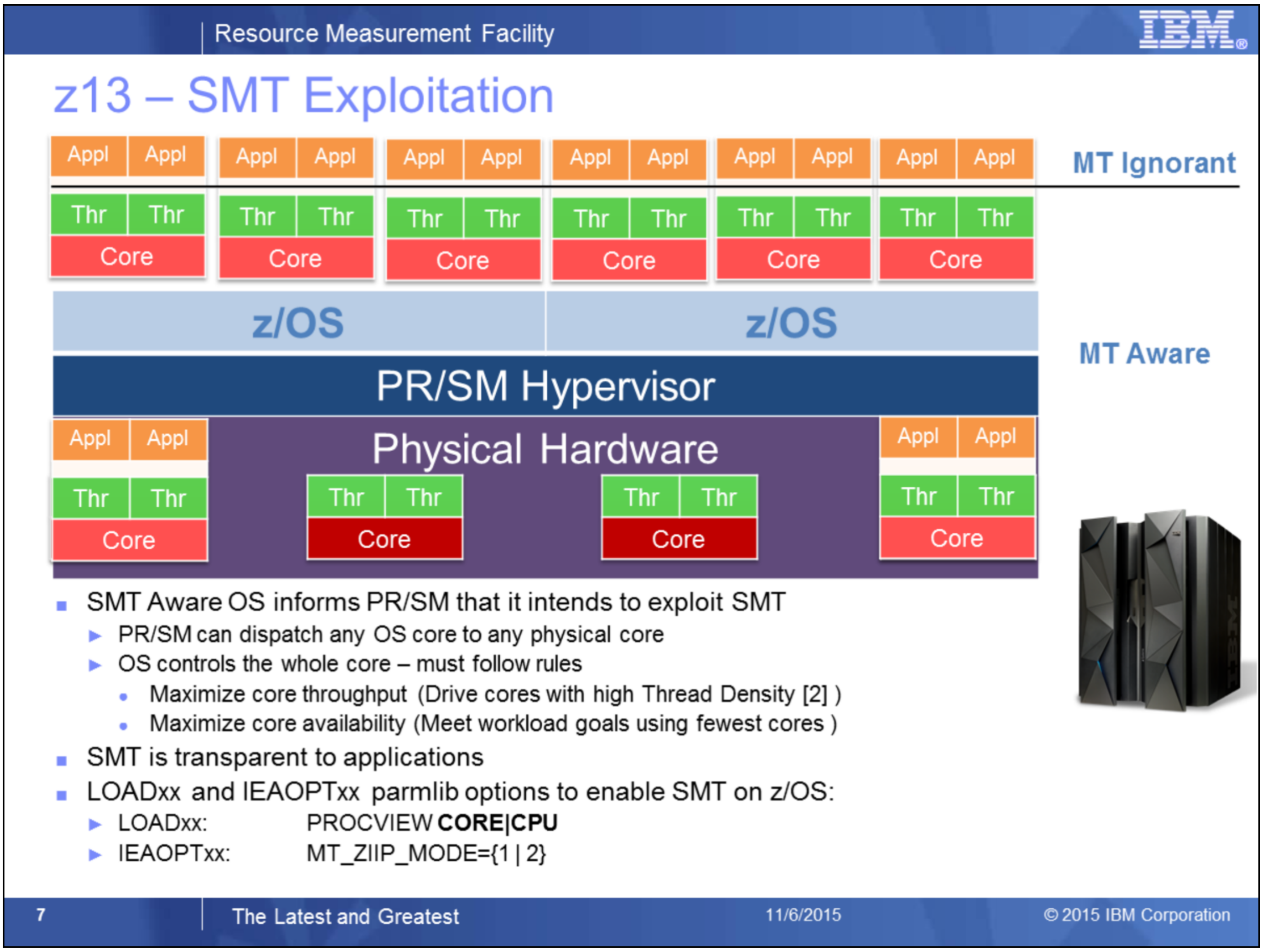
- As always, administration and switching of two lanes causes overhead
- Threads must share the peripheral processor resources, queueing and waiting cycles can occur

Alltogether, since with SMT more logical processing units are available, SMT helps to slow down the demand for processor resources.

Potential migration problem from zEC12 to z13: in case of single threaded workloads it is not recommended to turn on MT-2 mode.

But the overall thruput in MT-1 mode on z13 is always better than on zEC12 even though the processor speed has decreased.

This is achieved by other architectural optimizations, e.g. cache etc.

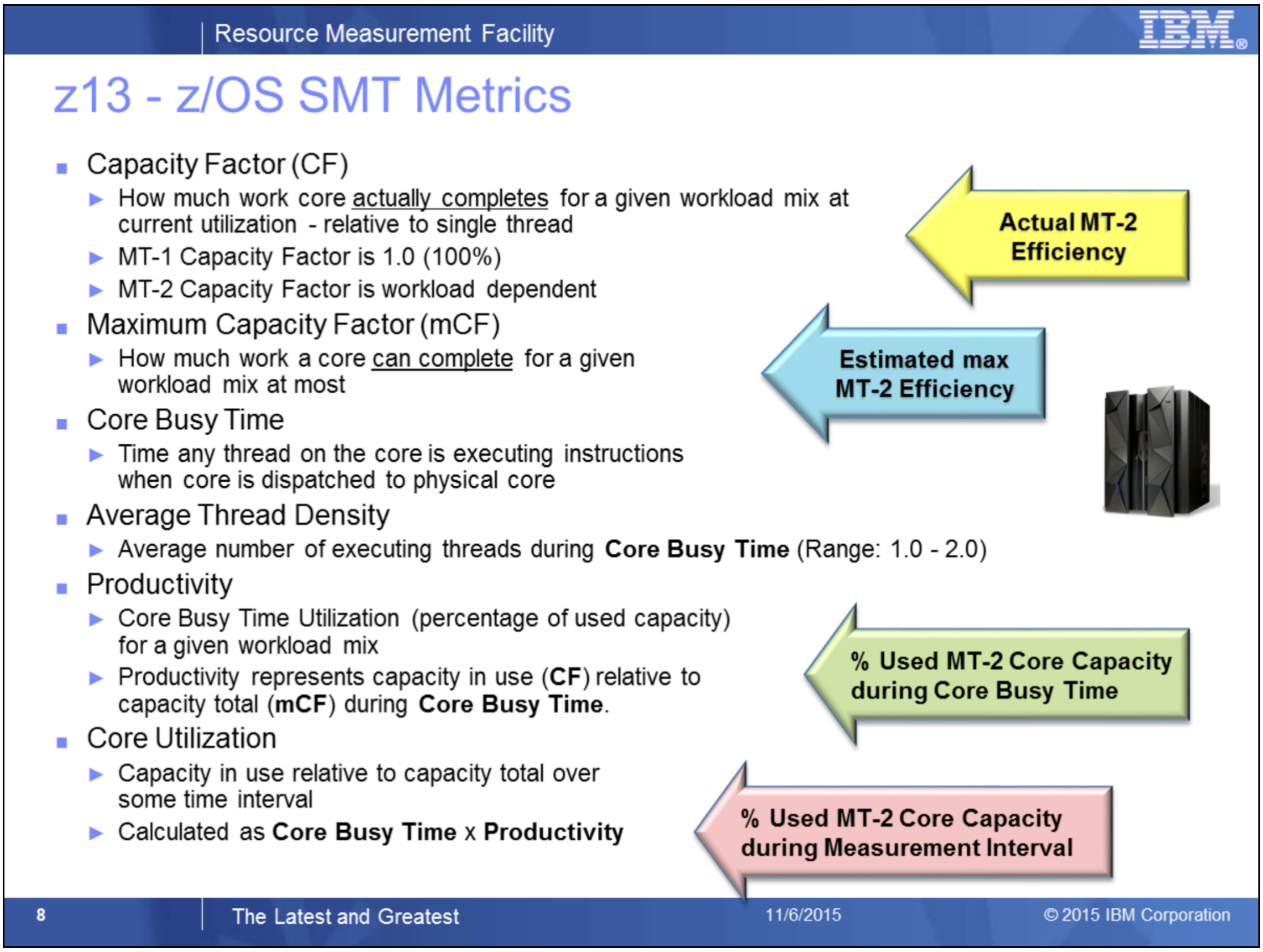


z13 – SMT Exploitation

Appl Appl Appl Appl Appl Appl Appl Appl Appl Appl Appl Appl
MT Ignorant
Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
Core Core Core Core Core Core
z/OS
z/OS
MT Aware
PR/SM Hypervisor
Physical Hardware
Appl Appl Appl Appl
Thr Thr Thr Thr Thr Thr Thr Thr
Core Core Core Core

- 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}

- The use of SMT mode can be enabled on an LPAR by LPAR basis via operating system (OS) parameters
- Once the OS switches to SMT mode, the only way back to single thread (ST) mode is via a disruptive action (re-activate the partition or re-IPL it).
- With the SMT enabled mode it is possible to dynamically switch between MT-1 (multi thread) and MT-2 mode for the processor types that support MT-2
- z/OS introduces new options for the LOADxx and IEAOPTxx parmlib members that are used to enable/disable SMT support and specify the MT mode of a processor class:
    - LOADxx parmlib option PROCVIEW CORE|CPU enables/disables SMT for the life of the IPL
      - PROCVIEW CORE on z13 enables SMT support
      - IPL required to switch between PROCVIEW CPU and CORE
    - New IEAOPTxx parameter to control the MT mode for zIIP processors
      - MT_ZIIP_MODE=1 specifies MT-1 mode for zIIPs (one active thread per online zIIP core)
      - MT_ZIIP_MODE=2 specifies MT-2 mode for zIIPs (two active threads per online zIIP core)
      - When PROCVIEW CPU is specified the processor class MT mode is always 1
      - SET OPT=xx operator command allows to switch dynamically between MT-1 and MT-2 mode
      - MT-2 mode requires HiperDispatch to be in effect
    - z/OS SMT Terminology:
  - z/OS logical processor (CPU) ➔ Thread
    - A thread implements (most of) the System z processor architecture
    - z/OS dispatches work units on threads
    - In MT mode two threads are mapped to a logical core
  - Processor core ➔ Core
  - PR/SM dispatches logical core on a physical core
    - Thread density 1 (TD1) when only a single thread runs on a core
    - Thread density 2 (TD2) when both threads run on a core



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

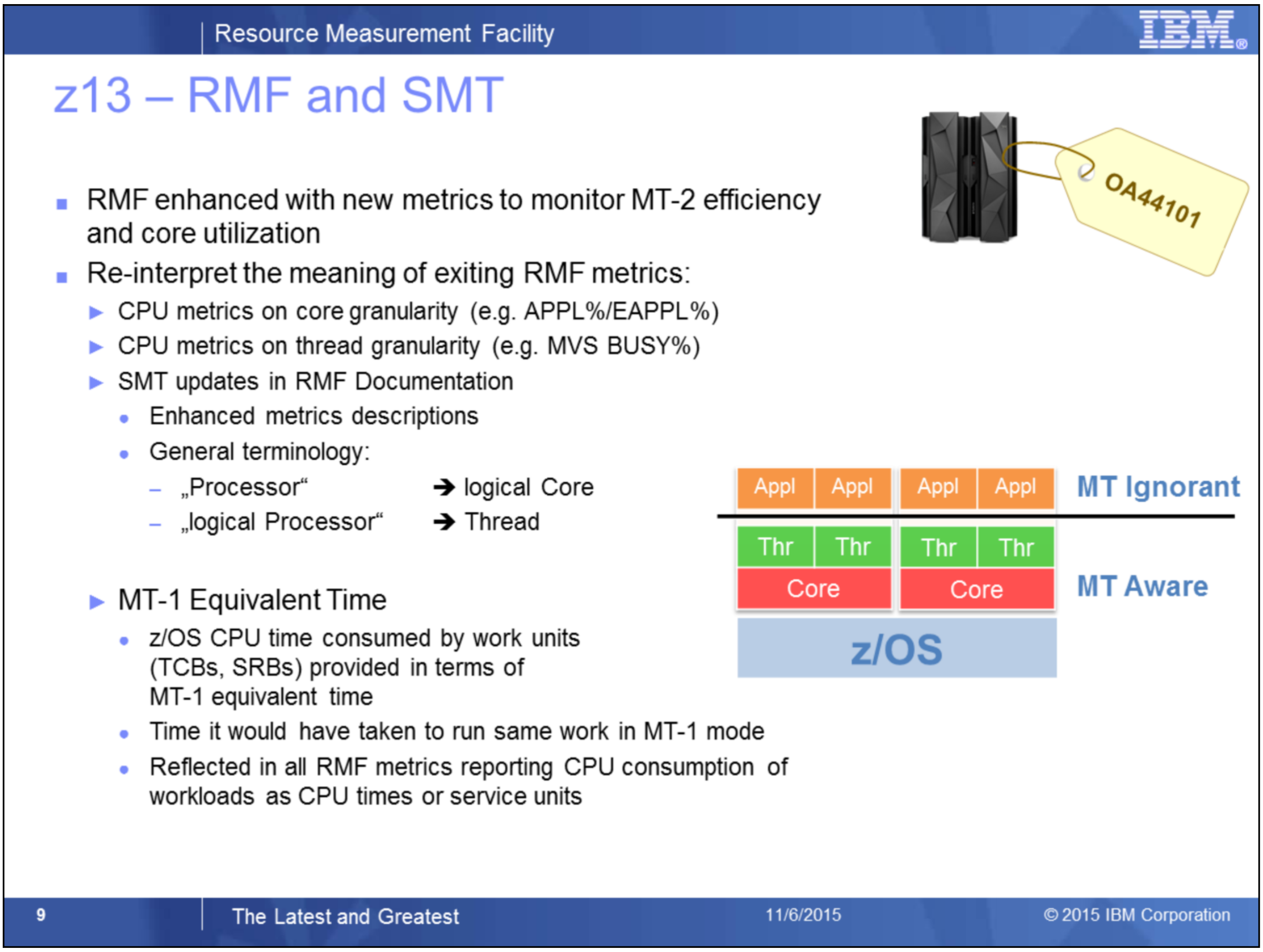
% Used MT-2 Core Capacity during Measurement Interval

z/OS SMT introduces several new metrics to describe how efficiently the core resources could be utilized and how efficiently they are actually utilized.

Is there any possibility to control the efficiency of the SMT mode?

Can I keep track of the actual speed on the 2 lanes compared to the possible maximum speed?

- The capacity factor (CF) is calculated from the number of executed instructions on both threads and the actual number of processing cycles that have been executed on the core. Usually the CF in MT-2 mode should be higher than 1.0 (except in special cases of single threaded workloads)
- The maximum capacity factor (mCF) can be achieved in case both threads are working together with maximum efficiency and no waiting cycles at all. Both, CF and mCF are subject of change: while the CF can change rapidly from interval to interval, the mCF changes are occuring slowly over time
- Core Busy Time is a metric on core level. The logical core must be dispatched on a physical core and at least one thread must be in execution state
- Average Thread Density represents the average number of executing threads during Core Busy Time
- Productivity indicates how well the available core resources have been exploited. It is calculated from the relation of CF and mCF during Core Busy Time
- Core Utilzation indicates how well my workload is suited for MT-2 mode. Since Productivity can never reach 100% the value is always below the Core Busy Time



z13 – RMF and SMT

- RMF enhanced with new metrics to monitor MT-2 efficiency and core utilization
- Re-interpret the meaning of exiting RMF metrics:
  - CPU metrics on core granularity (e.g. APPL%/EAPPL%)
  - CPU metrics on thread granularity (e.g. MVS BUSY%)
  - SMT updates in RMF Documentation
    - Enhanced metrics descriptions
    - General terminology:
      - „Processor“ ➔ logical Core
      - „logical Processor“ ➔ Thread
  - MT-1 Equivalent Time
    - z/OS CPU time consumed by work units (TCBs, SRBs) provided in terms of MT-1 equivalent time
    - Time it would have taken to run same work in MT-1 mode
    - Reflected in all RMF metrics reporting CPU consumption of workloads as CPU times or service units

The RMF support for SMT provides new SMT related metrics to allow capacity planning and performance analysis in SMT environments.

RMF supports SMT environments by extending the

- Postprocessor CPU activity report
- Monitor III CPC capacity report
- Overview Conditions based on SMF 70.1

RMF new function APAR OA44101 provides the SMT support for z/OS 2.1.

The architecture introduced with SMT requires a reinterpretion of existing RMF metrics:

- CPU metric data can now be on core or thread level granularity
- z/OS charges CPU time consumed by work units (TCBs, SRBs) in terms of MT-1 equivalent time. The MT-1 equivalent time is the time it would have taken to run the same work in MT-1 mode. All RMF metrics reporting CPU consumption of workloads as CPU time or service units reflect MT-1 equivalent time.

How can RMF help to monitor MT-modes efficiently:

- A couple of new MT related metrics are introduced
- Some existing metrics must be reinterpreted

First of all we need to distinguish between CPU metrics on core granularity and metrics on thread granularity.

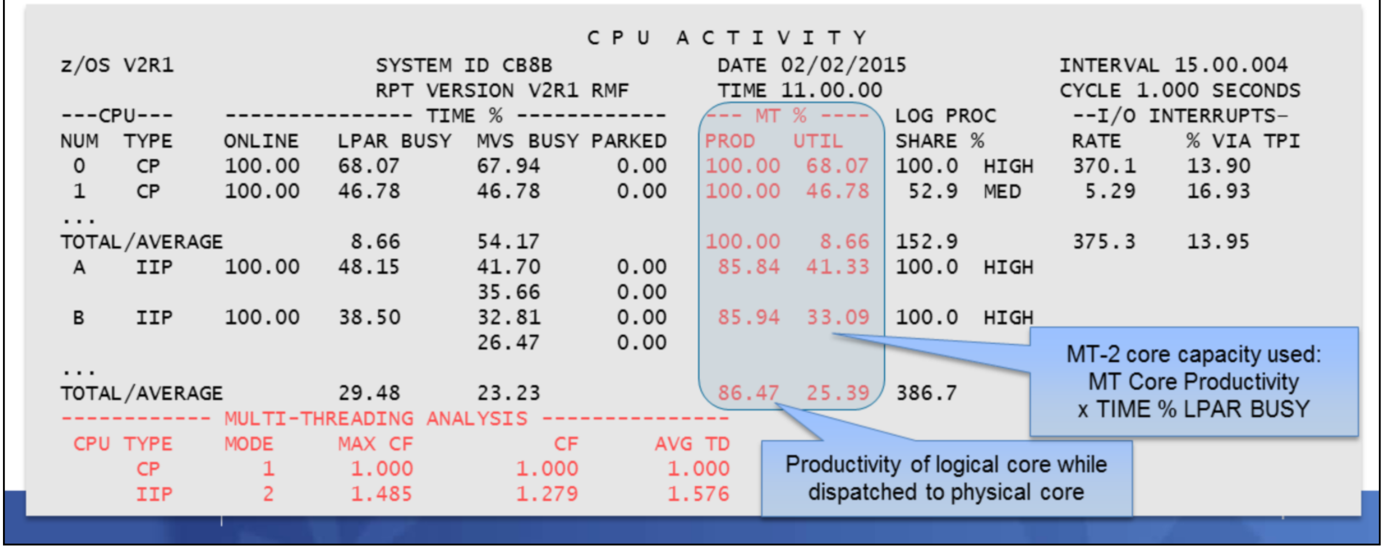
A comprehensive description can be found in the updated RMF V2R1 documentation.

In general, RMF does NOT report the measured wallclock times, but so called MT-1 equivalent times.

Hence, it is guaranteed that the reported CPU times and service units can be compared between the z13 systems and its predecessors.

z13 – SMT: Postprocessor CPU Activity Report

- PP CPU activity report displayed in “old” format when SMT is inactive
- PP CPU activity report provides new metrics when SMT is active
  - MT Productivity and Utilization of each logical core
  - 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: MT Core Productivity x TIME % LPAR BUSY

Productivity of logical core while dispatched to physical core

The CPU Activity section reports on logical core and logical processor activity. For each processor, the report provides a set of calculations that are provided at a particular granularity that depends on whether multithreading is disabled (LOADxx PROCVIEW CPU parameter is in effect) or enabled (LOADxx PROCVIEW CORE parameter is in effect).

If multithreading is disabled for a processor type, all calculations are at logical processor granularity.

If multithreading is enabled for a processor type, some calculations are provided at logical core granularity and some are provided at logical processor (thread) granularity. The CPU Activity section displays exactly one report line per thread showing all calculations at logical processor granularity. Those calculations that are provided at core granularity are only shown in the same report line that shows the core id in the CPU NUM field and which is representing the first thread of a core.

The following calculations are on a per logical processor basis when multithreading is disabled and on a per logical core basis when multithreading is enabled

- Percentage of the interval time the processor was online
- LPAR view of the processor utilization (LPAR Busy time percentage)
- Percentage of a physical processor the logical processor is entitled to use
- Multithreading core productivity (only reported when multithreading is enabled)
- Multithreading core utilization (only reported when multithreading is enabled)

The following calculations are on a per logical processor basis regardless whether multithreading is enabled or disabled:

- MVS view of the processor utilization (MVS Busy time percentage)
- Percentage of the online time the processor was parked (in HiperDispatch mode only)
- I/O interrupts rate (general purpose processors only)
- Percentage of I/O interrupts handled by the I/O supervisor without re-enabling (general purpose processors only)

The PROCVIEW CPU/CORE setting determines the content of the CPU report.

In case of CORE, new columns on CPU resp. thread level have been added.

The efficiency of MT-2 mode for zIIP engines can be monitored by means of a new MULTI-THREADING ANALYSIS section.

Remark: Since Parking is executed on core level, thread level reporting doesn‘n make sense. However, the values are extracted from local MVS control blocks for the home partition.