

Containing MLC Costs For Mobile and New Workloads

Cheryl Watson (cheryl@watsonwalker.com)

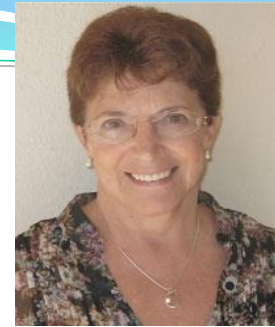
Frank Kyne (frank@watsonwalker.com)



SHARE 2016 in San Antonio - Session 18456

Welcome

- THANK YOU for attending this session
- Who are we and what we do.
 - Three-person company; started in 1987
 - Quarterly subscription-based newsletter -
 - Cheryl Watson's Tuning Letter
 - Cheryl Watson's System z CPU Chart
 - Public and private classes and consulting on z/OS new features, WLM, performance, Parallel Sysplex, high availability, software pricing, outsourcing contract reviews, and chargeback.
 - Software products - GoalTender and BoxScore
 - For more info, see our new website - www.watsonwalker.com
- What are we going to talk about?
 - Mobile Workload Pricing (MWP)
 - z Systems Collocated Application Pricing (zCAP)
- We have a LOT of material to cover, so if you have questions, please hold them to the end and we will be delighted to discuss them then.



Mobile Workload Pricing

Mobile Workload Pricing

- Understand what IS Mobile Workload Pricing (MWP)
- What are the things you need to consider when evaluating MWP?
- How does MWP impact how you configure your systems?

Introduction to MWP

- First, mobile is not a fad, it is not going away.
 - We know large z/OS customers where mobile consumes up to 50% of their z/OS capacity.
- IBM (and many others) believe mobile use will out-accelerate all other platforms over the next few years. Even if mobile is not a major player on your z/OS systems *now*, it very likely will be in your future.
- Mobile Workload Pricing (MWP) is an IBM pricing mechanism that is intended to reduce the cost of adding mobile-initiated workloads to z/OS.
- Important to note that MWP is aimed at sites that are re-using *existing* z/OS applications with mobile platforms.

Mobile Workload Pricing

- What *is* Mobile Workload Pricing?
- Headline is that it offers a **60% discount** on MSUs consumed by CICS/DB2/IMS/MQ/WAS transactions that originated on a mobile device.
 - 60 ... PERCENT ... OFF! WOW!
What else is there to say??
- Quite a bit....



Introduction to MWP

- We believe that MWP is a significant initiative from IBM - it indicates that IBM acknowledges that it must improve the cost-competitiveness of z/OS if customers are to grow and roll out new applications on this platform.
- MWP (together with zCAP and CMP) is aimed at reducing the cost of GROWTH.
 - Depending on how much of your current workload is 'mobile', and if that usage aligns with your Peak R4HA, MWP might not immediately reduce your SW bills.
 - BUT, signing up for MWP should make it cheaper to *add* mobile workload than if you had not signed up for it.
 - As the percentage of your workload that originated from mobile (and other new workloads) increases over time, at some point the bulk of your work will be priced at the new, more competitive price points, and your traditional (higher-priced) work will be a decreasing portion of the total work (and cost).
- **THE MESSAGE:** MWP, and *positioning to get the best value from it*, is a long term play. Start thinking about it now, even if you have very little mobile work yet.

Introduction to MWP

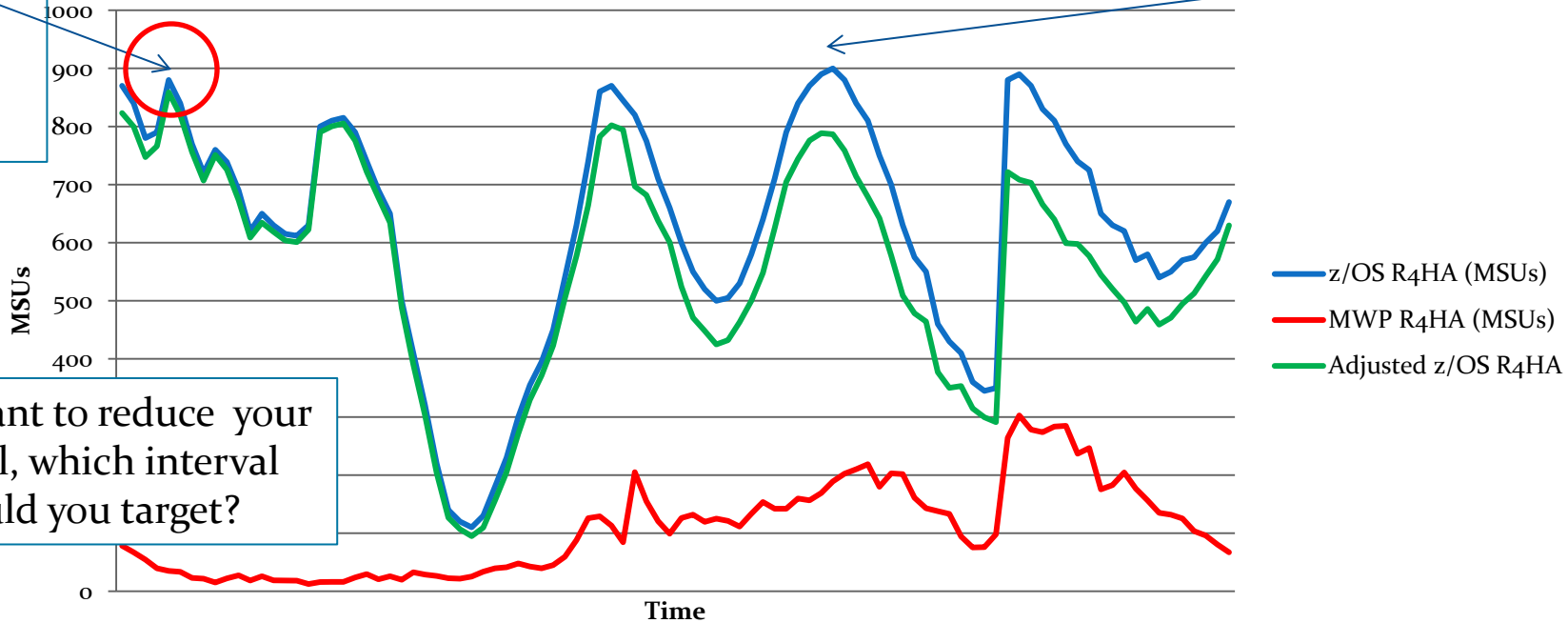
- **SHOW ME THE MONEY!**
- If you sign up for Mobile Workload Pricing (it is optional, and you must sign an agreement and supplements if you want to use it), IBM will reduce the R4HA FOR EVERY IBM Sub-Capacity MLC PRODUCT IN THAT LPAR in each interval by 60% of the corresponding R4HA of the CICS, DB2, IMS, MQ, or WAS transactions that originated from a mobile device.
 - An important point here is that it is not only the subsystem where the transaction ran (CICS, for example) that is discounted. It is EVERY sub-capacity IBM MLC product in that LPAR - SDSF, DB2, PL/1, you name it.

Understanding MWP

Impact of MWP on R4HA

Post-MWP basis for your SW bill ('Adjusted' R4HA)

Pre-MWP basis for your SW bill ('Real' R4HA)



If you want to reduce your SW bill, which interval should you target?

Introduction to Software Pricing

- Important points:
 - Traditional sub-capacity SW pricing is based on the combined **Peak** Rolling 4-Hour Average for every LPAR that product ran in. But with MWP, you now have *two* Rolling 4-Hour Averages - the 'real' one, and the MWP-adjusted one.
 - To reduce your SW bills, you must target the Peak R4HA interval.
 - If you are using MWP, you must target the *MWP-adjusted* R4HA.
 - A product's bill is NOT based on CPU usage of just *that* product. If you add a large CICS application (for example), the cost of *every* sub-cap product in that LPAR will increase.
 - Part of MWP's appeal is that it reduces the size of the increase for every sub-cap product, not just for the subsystem used by the mobile application.

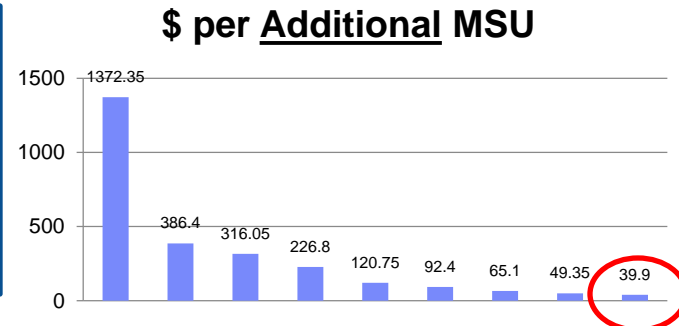
Introduction to Software Pricing

- Important points:

- Sub-capacity pricing provides a bulk discount with most sub-cap products - the cost of additional MSUs generally decreases as the total MSU consumption increases.
 - This reduces the cost/txn as you *grow* - GOOD!

• When you add a mobile workload *and* you are signed up for MWP, you get a double benefit - bulk discount means that additional MSUs cost less than your average. **AND**, your Real Rolling 4-Hour Average is reduced by 60% of the R4HA of the mobile workload.

• The flip side is that if you *reduce* the consumed MSUs, you are removing the cheapest ones. **So reducing MSUs by xx% will NOT reduce your bill by the same percent.** This indicates that it would make sense to sign up for MWP early rather than waiting until you have a large mobile workload.



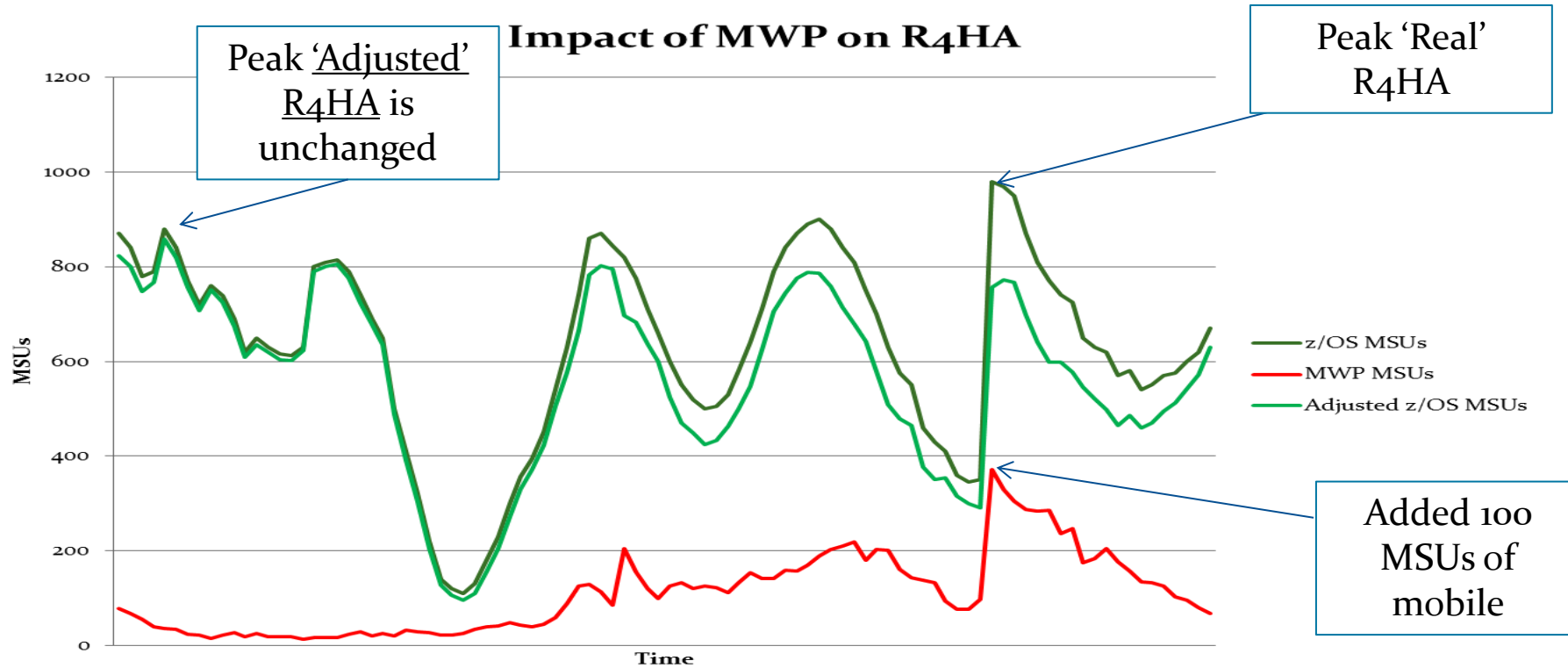
Introduction to Software Pricing

- Important points:
 - Many installations use soft-capping to control their software bills. IT Budgets are specified in terms of dollars, but soft-caps are specified in terms of MSUs.
 - This is fine when there is a predictable relationship between MSUs and \$s.
 - But where would you set the cap if the \$ per MSU for an LPAR constantly varies, depending on the workload mix at the time?
 - If you set the cap *high*, in the expectation that most of your work will come from mobile, you expose yourself to the risk that traditional work might use most of the capacity, driving up costs.
 - If you set the cap *low*, to protect yourself from exceeding the budget, you might have performance issues because the cap isn't high enough to serve an unexpected blip in mobile transactions.
 - Because the R4HA is reduced by 60% of the mobile R4HA, you *could* increase the cap without increasing your costs. But what is the right cap that gives you the perfect balance of cost vs. performance?

Introduction to Software Pricing

- Important points:
 - Traditionally, there has been a relationship between your software bill and your *CPC* utilization - when your utilization increases, your bill would also typically increase.
 - With *MWP*, only 40% of the capacity (CPU time) used by transactions that originated on a mobile device is counted when calculating your adjusted Rolling 4-Hour Average.
 - And the percent of your total capacity that is used by mobile tends to be less predictable and more variable than your traditional workloads.

Understanding MWP



Introduction to MWP

- The positive side of the disconnect between physical utilization and the R4HA that is used to calculate your bill, is that your software bill did *not* increase, despite the fact that your utilization increased by 100 MSUs.
 - If you had the same workload and had not signed up for MWP, your z/OS bill would have increased from the previous peak of 900 MSUs to about 1000 MSUs.

Introduction to MWP

- To sum up so far:
 - Mobile is not a fad.
 - MWP can potentially make it a lot less expensive to handle growing mobile workloads on your z/OS systems.
 - The cost of those financial savings is increased complexity in managing your configuration.
- Next questions are:
 - What qualifies as 'mobile'?
 - How does IBM know how much of my capacity is used by mobile?

What is a 'mobile device'?

- What is NOT a mobile device?
 - Laptop - even if it is using a web application, or if it is connected using a 3G/4G card.
 - Desktop PC. PC Server. UNIX system. z13 (not even a z13s). 370/138.....
- What IS a 'mobile device'?
 - Smartphone. Even if it initiates a txn using the exact same browser and application as a laptop - if it was initiated from a phone, that is OK, but not if a laptop initiated it.
 - Tablet.
- What else? Handheld stock-taking device? Playstation 3? Fitbit? Your car? Your refrigerator? ASK IBM. Technology is changing so quickly, it is not possible for IBM to provide an up-to-date, comprehensive, list of every device that qualifies. You might find that this lack of clarity plays to your advantage.

Managing an environment that has MWP

- How does IBM know the R4HA of your mobile workload? The terms and conditions for MWP specify that:
 - You must use MWRT (an alternative to SCRT) or a new Java version of SCRT (V23 R10 or later).
 - YOU are responsible for providing CPU usage information (CPU Seconds for each subsystem) about your mobile workload in a format specified by IBM.
- However, before you dive into your CICS Type 110 SMF records and IMS log records, pause for a minute to consider:
 - How will I identify the transactions that are eligible for MWP?
 - How much capacity (DASD GBs, CPU Secs) will I need to extract the info I need to input to SCRT/MWRT?
 - How will the method/topology that I select intersect with my soft-capping strategy?

Managing an environment that has MWP

- You basically have 3 options:
 - Run your MWP-eligible transactions in the same regions and subsystems as your traditional workloads.
 - Provide regions and subsystems that are dedicated to MWP-eligible transactions, but that run in shared LPARs.
 - Provide dedicated LPARs for the MWP-eligible transactions.
- We will look at the benefits and drawbacks of each of these, but first, a short diversion....

Workload Manager Support - MWP

- At the end of 2015, IBM delivered a number of enhancements to WLM (and RMF) that are especially of interest to anyone interested in MWP:
 - Additional classification criteria for CICS and IMS, intended to give you more flexibility in identifying transactions that originated on a mobile device.
 - The ability to assign a 'reporting attribute' of NONE (the default), MOBILE, or 2 other categories (for future use) to txns, jobs, STCs, etc.
 - Collection of CPU service unit information (for GP, zIIP, & zAAP) at the service class, reporting class, and system level for each of the 4 categories.
 - *Real time* tracking of Rolling 4-Hour Average for Total, MOBILE, CATEGORYA, and CATEGORYB transactions at the system level.

Workload Manager Support - MWP

Subsystem-Type Xref Notes Options Help

Modify Rules for the Subsystem Type Row 1 to 8 of 36
 Command ==> Scroll ==> PAGE

Subsystem Type : CICS
 Description : CICS tr

Action codes: A=After C=Copy M=Move I=Insert rule
 B=Before D=Delete row R=Repeat IS=Insert Sub-rule
 <== More

Action	Qualifier	Name	Start	Storage Critical	Reporting Attribute	Manage Region Using Goals Of
1	SI	DSTCPT2*		NO	NONE	N/A
2	TN	TPCI		NO	MOBILE	N/A
2	TN	TPSL		NO	NONE	N/A
2	TN	TPPA		NO	NONE	N/A
2	TN	TPOS		NO	NONE	N/A
2	TN	TPNO		NO	NONE	N/A
2	TN	TPDF		NO	NONE	N/A
2	TN	C*		NO	NONE	N/A

New field in WLM classification rules panels

Reporting Attribute

NONE

MOBILE

NONE

NONE

NONE

NONE

NONE

NONE

Options are:

- NONE (default)
- MOBILE
- CATEGORYA
- CATEGORYB

Workload Manager Support - MWP

- Prior to this new support, WLM collected transaction response times and transaction counts, but not CPU times, at the service class and reporting class level for CICS or IMS.
 - When you install the WLM PTF, you automatically get CPU usage info for CICS and IMS txns in service classes and reporting classes. This is great new capability even if you have no interest in MWP.
- For CICS txns, the service class and reporting attributes are assigned in the region where the transaction arrives (TOR or SOR, for example), then is passed to the AOR, FOR, etc.
 - IF you have TORs or SORs that process only MWP-eligible transactions, you could set up an SI CICS classification rule and assign all txns in those regions to MOBILE.
 - And in the STC classification rule, assign the TOR/SOR started task name to MOBILE.
 - If you want to capture the AOR and FOR region CPU time, those regions must also be dedicated.
 - This would result in both the transaction service time AND the region overhead time being assigned to MOBILE.

Workload Manager Support - MWP

- **IF** the expanded set of classification criteria in WLM allow you to identify all or most of your mobile transactions **OR** if you can direct all your MWP-eligible work to dedicated regions/subsystems, this is a hugely significant enhancement:
 - Would eliminate the need to process huge volumes of transaction-level information to extract the CPU usage information.
 - Would make it possible to determine the current MWP-adjusted R4HA for the system - this could then be used by dynamic capping products to determine if the cap should be increased or decreased.

Workload Manager Support - MWP

- WLM obtains CPU time info from:
 - Enclaves, for DDF and WAS
 - Existing SRM-maintained information for other work types (jobs, started tasks, TSO IDs..)
 - For txn server address spaces like CICS or IMS, the region CPU time is the time consumed by those address spaces that is *not* charged to transactions .
 - New, with this new WLM support - CICS and IMS report txn CPU time to WLM when they report the execution or response times. This requires:
 - CICS TS 5.3 (but you do NOT have to enable CICS CMF to gather this info)
 - For more info, see *IBM CICS Performance Series: CICS TS for z/OS V5 Performance Report*, [SG24-8298](#)
 - IMS V14 with APARs [PI46933](#) and [PI51948](#)

Workload Manager Support - MWP

- RMF PP Workload Activity Report support:

```

REPORT BY: POLICY=WLMPOL      WORKLOAD=BAT_WKL      SERVICE CLASS=BATHI      RESOURCE GROUP=*NONE
                                CRITICAL      =NONE
                                DESCRIPTION      =high priority batch

-TRANSACTIONS-  TRANS-TIME  HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---  SERVICE TIME  ---APPL %---  --PROMOTED--  ----STORAGE----
AVG      2.80  ACTUAL      0  SSSCHRT  0.0  IOC      0  CPU      6.846  CP      11.52  BLK      0.000  AVG      467.67
MPL      2.80  EXECUTION    0  RESP    0.0  CPU     561709  SRB     0.001  AAPCP   0.00  ENQ     0.000  TOTAL   1309.90
ENDED    0  QUEUED      0  CONN    0.0  MSO     581022  RCT     0.002  IIPCP   0.00  CRM     0.000  SHARED  2.80
END/S    0.00  R/S AFFIN   0  DISC    0.0  SRB     53  IIT     0.000  LCK     0.000
#SWAPS   100  INELIGIBLE  0  Q+PEND  0.0  TOT     1143K  HST     0.000  AAP     N/A  SUP     0.000  -PAGE-IN RATES-
EXCTD    0  CONVERSION  0  IOSQ    0.0  /SEC    19228  AAP     N/A  IIP     0.00  SINGLE  0.0
AVG ENC  0.00  STD DEV    0  ABSRPTN 6865  IIP     0.000  SHARED  0.0
REM ENC  0.00  TRX SERV 6864  HSP     0.0
MS ENC   0.00
    
```

TRANSACTION APPL% :	TOTAL :	CP 11.52	AAP/IIP ON CP 0.00	AAP/IIP 0.00
	MOBILE :	CP 4.82	AAP/IIP ON CP 0.00	AAP/IIP 0.00

Note: CATEGORYA or CATEGORYB are currently not reported by RMF PP

Workload Manager Support - MWP

- RMF PP Overview Report support:

RMF OVERVIEW REPORT

PAGE 001

z/OS V2R1			SYSTEM ID FPK2			START 02/26/2016-14.41.35 INTERVAL 00.01.00		
			RPT VERSION V2R1 RMF			END 02/26/2016-18.22.36 CYCLE 1.000 SECONDS		
NUMBER OF INTERVALS 220			TOTAL LENGTH OF INTERVALS 03.40.00					
-DATE	TIME	INT	RR4HA	MOBIL	CATA			
MM/DD	HH.MM.SS	HH.MM.SS						
02/26	14.42.35	00.00.59	21	8	10			
02/26	14.43.35	00.01.00	21	8	10			
02/26	14.44.35	00.01.00	21	8	10			
02/26	14.45.35	00.00.59	21	8	10			
02/26	14.46.35	00.01.00	21	8	10			
02/26	14.47.35	00.00.59	21	8	10			
02/26	14.48.35	00.01.01	21	8	10			
02/26	14.49.36	00.00.59	21	8	10			
02/26	14.50.35	00.01.00	21	8	10			

```
//RMFPP EXEC PGM=ERBRMFPP,REGION=0M
//MFPMSGDS DD SYSOUT=*
//*FPINPUT DD DSN=CLITHED.SMFJA0.PET,DISP=SHR
//PPOVWREC DD DISP=OLD,DSN=KYNEF.RMF.OVIEW
//SYSIN DD *
OVERVIEW(RECORD,REPORT)
OVW(RR4HA(LACS))
OVW(MOBIL(LACSM))
OVW(CATA(LACSM))
SYSOUT(X)
```

System
R4HA

MOBILE
R4HA

CATEGORY
A R4HA

Workload Manager Support - MWP

- Just a little more information about these enhancements:
 - Available for z/OS 2.1 + later.
 - Delivered by [OA47042](#) (WLM) and [OA48466](#) (RMF).
 - Require CICS TS 5.3 and/or IMS V14 + PTFs
 - They do NOT require any changes to your existing Report or Service Class structure.
 - CPU usage information is kept in a new field in WLM.
 - Reported in new fields in SMF Type 70 and 72.3 records. Also in Type 99.2.
 - RMF Postprocessor Workload Activity report and Overview reports enhanced.
- And now, back to our 3 options for configuring for MWP....

Considerations for shared regions

- **Benefits:**

- EASY to set up – just use existing regions and subsystems.

- **Drawbacks:**

- Transaction-level SMF records do not capture region management time – about **80% is captured**, at best.
- MQ does not provide transaction-level CPU usage info in its SMF records, so you are limited to collecting whatever MQ charges back to CICS/IMS/etc.
- Calculating CPU usage in real time is expensive, maybe impossible (unless you can use WLM MWP support).

- **Drawbacks:**

- If it is not possible to identify MWP-eligible txns to WLM, you **MUST** process transaction-level SMF data to identify CPU consumption of MWP-eligible transactions. This could be a LOT of data.
- Identifying the source of the transaction from the SMF records *might* not be possible.
- How do you identify the original source of transactions that are called by other txns?
- Maintenance effort for programs that extract CPU usage info is not insignificant – every time a new MWP-eligible application is deployed or modified, you need to update your programs. And not every application will use the same mechanism for identifying where it originated.

Dedicated regions in shared LPARs

- **Benefits:**

- Massive reduction in volume of SMF data to be processed compared to processing txn-level SMF Recs.
- IBM will allow ALL A.S. CPU time to be discounted – both region-level AND txns (100% instead of 80%).
 - If using WLM MWP support, you can get discount for Txn & region CPU time AND you get real-time MWP R4HA.
 - If *not* using WLM MWP, IBM will accept data extracted from SMF Type 30
- Might be easier to identify the transaction source in the network and route it to the dedicated regions – removes the need to identify txn source in WLM classification rules or transaction-level SMF records.

- **Benefits:**

- Because CPU time is obtained from WLM/RMF or Type 30 records, maintenance effort should be a lot lower than if you are gathering this info from transaction-level SMF or log records.

- **Drawbacks:**

- Requires additional regions/subsystems, meaning more work to set up and manage, plus the resources required for more address spaces.
- Additional address spaces might reduce CPC cache hit rate, increasing the RNI.
- Requires data sharing if you want to extend this to database manager.

Dedicated LPARs

- **Benefits:**

- All of the benefits of dedicated regions, plus....
- Easier to manage LPAR capacity, because nearly all work in the LPAR has the same average price per MSU if you have MWP LPARs and Traditional LPARs.
- Easier to fence off additional capacity provided for MWP work from traditional workloads.
- IBM will accept data from just the Type 70 and Type 89 records – no need to collect, keep, and post-process transaction-level or even address space-level SMF records.
- Could run minimal SW stack to reduce costs

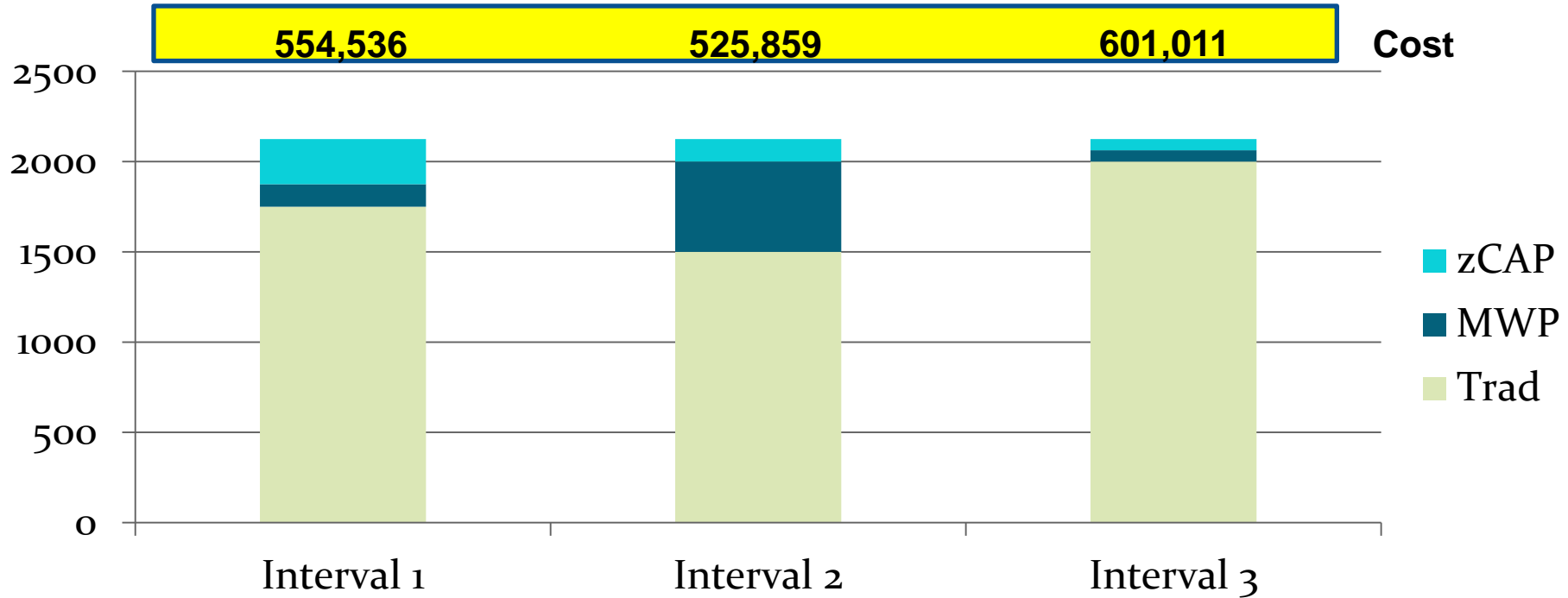
- **Benefits:**

- There might be security advantages to isolating transactions originating on a mobile device into their own LPARs.

- **Drawbacks:**

- Setting up new systems means more work to set up and manage, plus the resources (CPU, DASD, memory) required for more LPARs.
- Additional LPARs negatively impacts overall CPC cache hit rates.
- Requires data sharing, assuming that you want to share data between MWP and traditional applications.

Relationship between MWP and Capping



Same total used capacity, three very different costs

Which topology is right for you?

- There is no one 'right' answer for everyone.
 - But the new WLM MWP support gives you a lot more flexibility.
- The important thing is that you make an informed decision based on all the considerations.
- Start working with your subsystem colleagues, network administrators, application architects NOW, to make sure that it is possible to identify the transaction source AND (if possible) to do this in a consistent manner (to reduce the maintenance cost). If you don't have much mobile yet, GREAT - you can get your architecture agreed and implemented now, rather than having to re-do it all later.
- If you use (or are considering) a product to manage soft-caps dynamically, talk to them now to determine their plans for managing this environment.

Implementing MWP

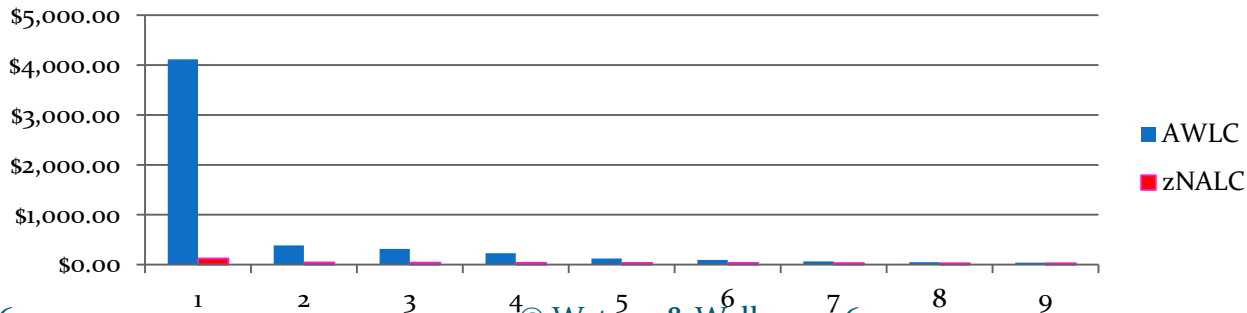
- In order to be able to avail of MWP, you must:
 - Have a zBC12 or zEC12 or later in your enterprise.
 - The MWP-eligible workloads must run on a z114/z196 or later.
 - Be running z/OS (V1 or V2) and one or more of CICS (V4 or V5), DB2 (V9, V10, or V11), IMS (V11, V12, or V13), MQ (V7 or V8), or WAS (V7 or V8).
 - Be using a sub-capacity pricing option - AWLC, AEWLC, or zNALC.
 - Sign the MWP supplement.
 - Must agree with IBM which applications will be eligible, and how you will gather the usage data for those applications. And, especially, exactly how you will identify the MWP-eligible transactions.
 - Also, any time you add new MWP transactions/applications, you must inform IBM and complete a new supplement.
 - *Provide your own mechanism to create the MWP input to IBM reporting tool.*
 - Use MWRT or SCRT 23.10 or later to report your utilization to IBM.

zCAP Pricing

z Systems Collocated Application Pricing (zCAP)

- **z Systems New Application License Charging (zNALC) has been available since 2007.**
 - zNALC significantly reduces software costs for applications that meet certain criteria.
 - However it requires that the applications are run in dedicated zNALC LPAR(s)
 - zNALC LPARs can be in same *sysplex* as traditional workloads and can share data with traditional workloads.
 - z/OS in the zNALC LPARs will be priced using special zNALC prices.

AWLC to zNALC z/OS comparison



z Systems Collocated Application Pricing

- While zNALC significantly reduces the cost of z/OS for those LPARs, some customers don't want to have to set up dedicated LPARs for those workloads. To address that concern, in 2015 IBM introduced a new pricing option called z Systems Collocated Application Pricing (zCAP).
- zCAP is conceptually similar to MWP in that discounts are based on the middleware CPU consumption of applications that meet the criteria for zCAP and that are described in your zCAP agreement and supplements with IBM.
- However, because the applications are NEW (and have unique names), they should be a lot easier to identify than MWP transactions, which use existing applications (meaning that you don't have the complexity of trying to determine the source of the transaction).

z Systems Collocated Application Pricing

- What is a 'new' workload?
 - Must be a new application to z/OS in your enterprise.
 - Does not have to be new 'in the universe' - for example, SAP has been around for many years, but if you are not using SAP on z/OS now, then it is eligible to be considered 'new' for zCAP purposes.
 - If you move an application from another platform in your enterprise to z/OS, that also counts as being 'new' for zCAP purposes.
 - The zCAP definition of 'new' is a lot more flexible than the official zNALC definition of new. Application must use at least one of CICS/DB/IMS/MQ/WAS, but that is all.
- The objective is to provide you with more flexibility to help you add new z/OS applications.
- Organic growth of existing applications does *not* count as 'new' for zCAP purposes.
- For gray areas, speak to IBM and make a case for why the application should be considered 'new'.
- Also, in the words of IBM's David Chase, 'newness does not wear off. Applications that qualified as 'new' 5 years ago are still considered new today'.

z Systems Collocated Application Pricing

- Like MWP, you have to identify the MSUs used by the zCAP-eligible workload (CICS/DB2/IMS/MQ/WAS).
 - Then you subtract 50% of that amount from the z/OS R4HA.
 - And you subtract 100% of that amount from *all MLC products* in the LPAR (CICS, DB2, IMS, MQ, WAS, COBOL, NetView, etc.) EXCEPT for the subsystems that are used by the application.
 - Then you pay for the MSUs for the subsystems used by the zCAP-eligible workload using the same pricing metric that is being used by the LPAR the application is running in.
- Let's look at two scenarios...
 - First one is where a new application is the only user of a 'zCAP-defining' subsystem (CICS/DB2/IMS/MQ/WAS).
 - Second one is where the new application uses an existing subsystem.

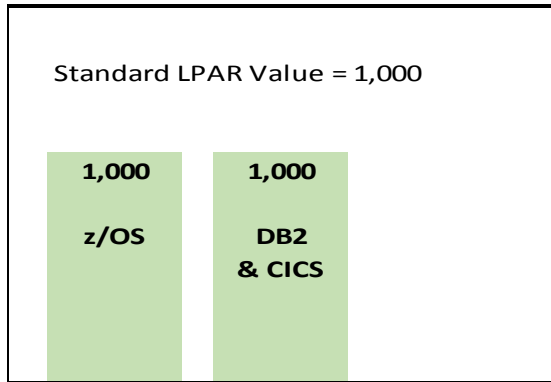
z Systems Collocated Application Pricing

Net New MQ Example = 100 MSUs of new MQ workload *

1. Existing LPAR

MSUs used for subcap billing:

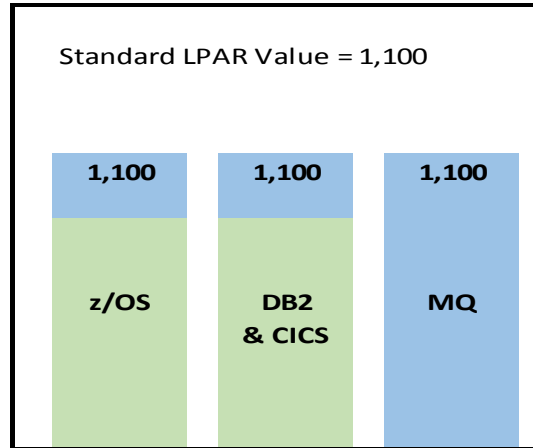
z/OS	1,000
DB2 and CICS	1,000



2. New MQ, standard rules

MSUs used for subcap billing:

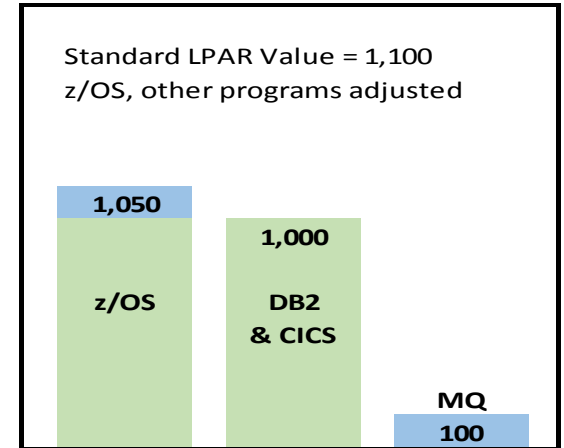
z/OS	1,100
DB2 and CICS	1,100
MQ (LPAR value)	1,100



3. New MQ with zCAP pricing

MSUs used for subcap billing:

z/OS	1,050
DB2 and CICS	1,000
MQ (usage value)	100



* Assumes workloads peak at same time
March 5, 2016

© Watson & Walker 2016 Example courtesy of David Chase, IBM

z Systems Collocated Application Pricing

- Consider what would have happened if you had used zNALC for this application...
 - You would have paid a discounted price for z/OS based on a 100 MSU R4HA.
 - You would have paid for 100 MSUs of MQ
 - CICS and DB2 would not be in the zNALC LPAR, so you would not pay anything for them.
- Because you are using zCAP in this example:
 - The MSU value used for CICS & DB2 was reduced by 100% of the capacity used by the new application because it didn't use either of those products - so adding a new MQ application had no impact on how much you paid for CICS or DB2 (you continued to pay for 1000 MSU of CICS/DB2).
 - You reduced the total z/OS R4HA number by 50% of the capacity used by the new application (50 MSU reduction) so you paid for 1050 MSUs of z/OS.
 - You only paid for 100 MSUs of MQ, even though it lived in an LPAR that was using 1100 MSUs.
- So the net financial effect may be similar to zNALC, but without the need for a separate LPAR.

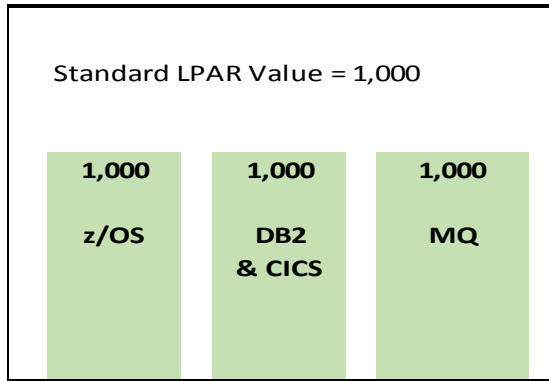
z Systems Collocated Application Pricing

Incremental MQ growth Example = 100 MSUs of MQ growth *

1. Existing LPAR

MSUs used for subcap billing:

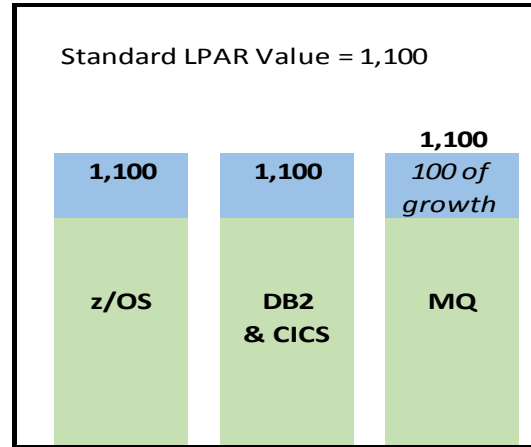
z/OS	1,000
DB2 and CICS	1,000
MQ	1,000



2. MQ growth, standard rules

MSUs used for subcap billing:

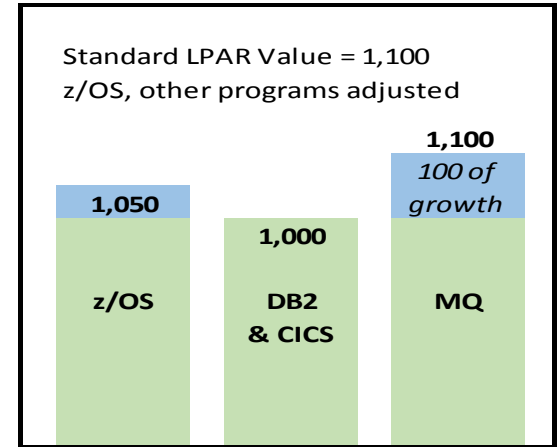
z/OS	1,100
DB2 and CICS	1,100
MQ w/growth	1,100



3. MQ growth with zCAP pricing

MSUs used for subcap billing:

z/OS	1,050
DB2 and CICS	1,000
MQ w/growth	1,100



* Assumes workloads peak at same time
March 5, 2016

© Watson & Walker 2016

Example courtesy of David Chase, IBM

z Systems Collocated Application Pricing

- In this example, the new application used a product (MQ) that was already being used by existing applications:
 - The MQ cost went up by the 100 MSUs that the application was using. (MQ cost now 1100 MSUs)
 - The R4HA value used for CICS & DB2 was reduced by 100 because the new application didn't use CICS or DB2. (CICS/DB2 costs are still 1000 MSUs).
 - The total z/OS MSU number was reduced by 50% of the capacity used by the new application (50 MSU reduction). (z/OS cost now 1050 MSUs).
 - The R4HA for every other MLC product would be reduced by the 100 MSUs so they would remain at 1000 MSUs.
- So, again, the net effect is similar to zNALC, but without the need for a separate LPAR.
 - With zNALC you would pay for 100 MSUs of z/OS at the very-reduced zNALC rate. With zCAP, you would pay for 50 MSUs of z/OS at your incremental price for z/OS (with the price depending on where you are on the pricing curve for z/OS).
 - The relative costs of MQ would depend on if you use AWLC or Value Unit Edition (IPLA, only available with zNALC) and where you are on the pricing curve.

z Systems Collocated Application Pricing

- As with MWP, you are responsible for identifying the capacity used by the new workload and translating that into a CSV file that is input to MWRT (or the spectacular, fantasmagorical, **z/OS-based**, new SCRT).
 - If the new application is the only user of a subsystem (as in the 1st example), it is acceptable to use data from the Type 89 SMF records.
 - If the application is using an existing subsystem product (MQ, in example 2), but runs in its own dedicated region, IBM will accept data from the Type 30 records for that region.
 - If the application is using an existing subsystem AND an existing region, then you need to use transaction-level information to determine the MSUs used by the new application.

zCAP Requirements

- zCAP is only available for new applications that run in a z114/z196 or later with AWLC, AEWLC, CMLC, or zNALC sub-capacity pricing.
- Supports both z/OS V1 and V2, and current and recent versions of CICS/DB2/IMS/MQ/WAS.
- Data must be submitted to IBM using MWRT 3.3.0 or later (current version is 3.3.5) or SCRT 23.10 or later.
- There is a new contract Addendum and accompanying Supplement:
 - Addendum for z Systems Collocated Application Pricing (Z126-6861)
 - Terms and conditions to receive zCAP benefit for AWLC, AEWLC, zNALC billing
- Supplement to the Addendum for zCAP (Z126-6862)
 - Customer explains how they measure their zCAP application CPU time
 - Agreement to and compliance with the terms and conditions specified in the zCAP contract Addendum is required

zCAP Summary

- zCAP has a similar objective to zNALC - reduce the cost of adding 'new' applications to z/OS. However, the zCAP definition of 'new' is much more flexible than the zNALC definition.
- zCap provides more flexibility in terms of where you locate new applications - you can now select a topology (dedicated LPAR or existing LPAR) that makes both financial *and* technical sense.
- It is not possible to make a blanket statement about which option (zNALC or zCAP) will have lower costs. Recommend that you work with your IBMer to price the following options:
 - Straight AWLC/AEWLC.
 - zCAP.
 - zNALC with AWLC/AEWLC for subsystems.
 - zNALC with IPLA for subsystems.
 - Don't forget to factor in cost of dedicated LPAR for zNALC.

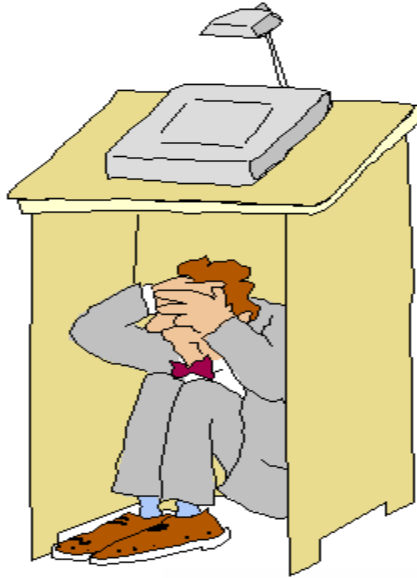
zCAP References

- List of zCAP-defining products:
 - <http://ibm.com/systems/z/swprice/reference/exhibits/zcap.html>
- zCAP [Announcement Letter \(ENUS215-174\)](#)
- For additional information, contact your IBM account team.
 - Also, if you are attending any conferences, watch for sessions by David Chase. They are very informative and he is very good at answering questions.

Other interesting sessions

- Monday 11:15, Session 18214, *Challenges of MSU Capping w/o Impacting SLAs* by **Don Zeunert**
- Monday 15:15, Session 18626, *Workload Management (WLM) Update for z13, z/OS V2.2 and V2.1*, by **Andreas Henicke**.
- Tuesday 13:45, Session 18153, *Country Multiplex Pricing (CMP) is here - What you need to know to minimize your MLC costs*, by **John Baker**
- Wednesday 10:00, Session 18345, *SHARE Live!: Achieving Significant Capacity Improvements on the IBM z13: User Experience*, by **Todd Havekost**
- Wednesday 10:00, Session 18416, *The MLC Cost Reduction Cookbook*, by **Scott Chapman**
- Thursday 08:30, (NOW!!) Session 18597, *RMF: The Latest and Greatest*, by **Peter Muench**.
- Thursday 15:15, Session 18480, *Can Country Multiplex Pricing Save Money??*, by **Al Sherkow**
- Friday 10:00, Session 18017, *SHARE Live!: The Cheryl and Frank zRoadshow*, by **Cheryl Watson** and **Frank Kyne**.
- Friday 10:00, Session 18212, *Understanding MLC Cost Impact of Performance and Capacity Management*, by **Don Zeunert**

Questions?



Thanks!

Thank you, and we look forward to seeing you again soon.

