



# The Cheryl and Frank zRoadshow (Part 1 — aka 'Frank gets a word in')

Frank Kyne (frank@watsonwalker.com)

Cheryl Watson (cheryl@watsonwalker.com)

Session 19877 August 4, 2016





SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.





#### Welcome

SHARE EDUCATE · NETWORK · INFLUENCE

- THANK YOU for attending this session!
   We hope you are enjoying the week, we sure are.
- Who are we and what we do:
  - Founded in 1987 by Cheryl Watson & Tom Walker;
     Frank Kyne joined in 2014
  - Cheryl Watson's Tuning Letter quarterly, published since 1991
  - 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
  - GoalTender WLM policy analyzer
  - BoxScore After-the-fact benchmark
  - For more info see our NEW website: <a href="www.watsonwalker.com">www.watsonwalker.com</a>
- This is the first half Part Deux will be at 10:00 tomorrow.







#### Welcome



#### Topics to cover:

- Holistic Capacity Planning
- z13 Performance Tips
- IBM Continuous Delivery Strategy
- SMF Streaming and Spark
- WLM Mobile Workload Pricing Support
- WLM New Capping Support
- SHARE Live!





- One of our customers (USAA) kindly shared their z13 upgrade experiences with us for a Tuning Letter article.
- USAA upgraded 4 zEC12s to 4 z13 CPCs. Initially went to 711 models (from EC12 711s), then to 716s, then to 726s using 'banked MIPS'.
- The results of the upgrades were:
  - zEC12 to z13 INcrease in MIPS consumption of 4000 MIPS.
  - Z13 711 to 716 DEcrease in MIPS consumption of 9000 MIPS.
  - Z13 716 to 726 DEcrease in MIPS consumption of a further 4000 MIPS.
- This equates to a reduction of about 1600 MSUs in monthly SW bills, comparing z13 711 to 726.
  - Apply your own \$ per MSU to see how much this would save in your environment.





#### Lessons learned:

- We are all told "Mainframe hardware is expensive". Well, have you looked at your software bill recently? Hardware costs are more-or-less a one time charge your software bill shows up every month. To determine and optimize the real TCO you need to include all costs SW, HW, maintenance, and so on.
- The GHz of z13 is 10% *lower* than zEC12, yet the MIPS for an *Average RNI* workload on a z13 is 10% *higher* than zEC12. The increased capacity was achieved largely through processor design changes. It is likely that some workloads will really love the new design, and some less so, resulting in more variability between workload types. As a result, you need to be VERY careful when using the "Average RNI" number base your upgrade on your ACTUAL workload type (obtained using the SMF type 113 records), don't just assume that it is 'average'. DON'T UPGRADE to z13 WITHOUT USING SMF TYPE 113 (HIS) RECORDS..





- Lessons learned:
  - When moving from z196 to zEC12, there was a 5% difference in additional capacity for High vs. Low RNI workload (22% vs. 27%). When moving from zEC12 to z13, there is an 8% difference between Low and High RNI (6% vs. 14%).



Software MSUs are based on Average RNI workloads, so if you move a High RNI workload from zEC12 to z13, you might see that more MSUs are consumed to perform the same amount of work. USAA's main workload RNI was Very High RNI – IBM projected that they would see a 4-5% capacity increase, rather than the 10% indicated by the IBM MIPS number (which is based on Average RNI workload).



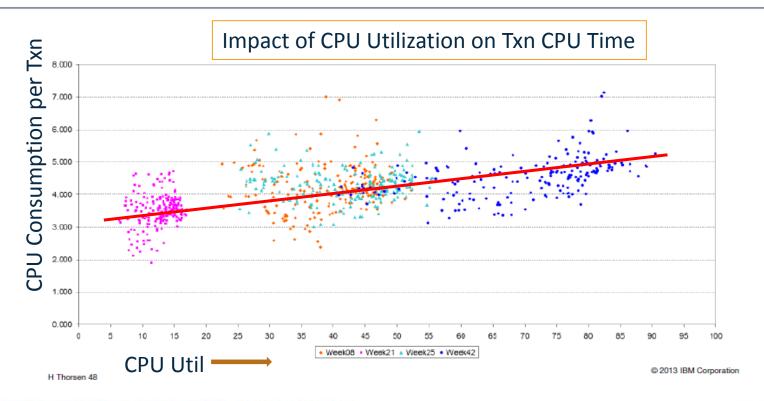


#### Lessons:

- Know your RNIs. Critical for upgrade planning, valuable for tuning.
  - See John Burg's <u>CPU MF Update</u> SHARE Session for more info.
  - If your RNI is higher than the normal High range, talk to IBM before finalizing your target configuration.
- Optimize your HiperDispatch config to maximize your number of VH CPs.
  - Investigate if a sub-capacity model (5xx, 6xx) would be suitable.
  - Use Alain Maneville's <u>LPARDesign tool</u> on the WLM web site.
- Provide enough memory to avoid CPs or zIIPs in one drawer having to access memory in a different drawer.
- Every additional 10% of CPU utilization increases CPU/Txn by between 3% (Low RNI workload) and 5% (High RNI workload). (See Gary King White Paper, <u>Running IBM System z at High Utilization</u>, for more info.) So, reducing CPC utilization reduces CPU/Txn by a similar amount.







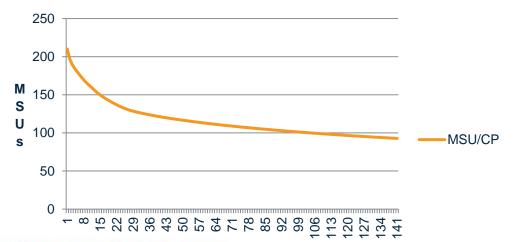




MP effect means that capacity does not grow linearly. To allow for this, IBM MSUs/CP decrease as you add more CPs. The MSUs/CP for a z13 711 is 160, vs. 131 for a 726 – an 18% difference. However, the MP effect is related to utilization – lower utilization should result in a lower MP effect. So, running an identical volume of work on a 726 should result in a smaller SW bill than running that same workload on a 711. THIS IS GOODNESS.

MSUs/CP

IBM have multiple MIPS values (Low/Avg/High) per model, but only a single MSU value.



Two customers using 100% of 1 CP would pay the same SW bill, but one might get a lot more MIPS (work) than the other.





- EVERYONE should collect and keep the HIS (SMF type 113) records for every system all the time (this is WSC Best Practice).
  - Use tools such as IntelliMagic, Pivotor, SAS/MXG, or zHISR to format them.
  - If you don't have any tools, use the IBM SMF113 Reporting Tool, available from:
    - <a href="http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools">http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools</a>

       .html#SMF113
- WLM type 99.14 SMF records provide invaluable information about the HiperDispatch topology, and are very low volume (about 250 per day) – make sure that you do not suppress them in your SMFPRMxx member (ALL type 99s are typically suppressed). And collect them on EVERY system!
  - Use WLM Topology Report tool to view them: <a href="http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm">http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm</a>
     <a href="http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm">http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm</a>
     <a href="http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm">http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm</a>
     <a href="http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm">http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm</a>
     <a href="http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm">http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm</a>
     <a href="http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm">http://www.ibm.com/systems/z/os/zos/features/wlm/WLM\_Further\_Info\_Tools.htm</a>
     <a href="http://www.ibm.com/systems/zos/gen/zos/features/wlm/wlm/wlm/systems/zos/gen/





#### References

- Achieving Significant Capacity Improvements on the IBM z13: User Experience, by Todd Havekost, SHARE 2016 in Atlanta, Session 19488
- To MIPS or Not to MIPS, That is the CP Question!, by Gary King, SHARE 2016 in San Antonio, Session <u>18643</u>
- 'Holistic Capacity Planning' article and 'HiperDispatch Questions and Answers' article in Cheryl Watson Tuning Letter 2015, No. 4.
- IBM Redbook z13 Technical Guide, <u>SG24-8251</u>





#### z13/z13s Memory Pricing ("It depends")

- When moving to z13, we strongly recommend that you get quotes for a variety of memory sizes, but especially 3X current memory and 5X current memory.
- On z13s, get quote for >1 TB of memory IF you are planning for an N20 anyway.
- If you are interested in Plan Ahead Memory, get price of available + Plan Ahead memory, and compare that to just buying the total amount of memory now – it might cost less to just buy it all now.
- AND you can then utilize that memory to reduce I/Os this should result in a better Relative Nest Intensity (RNI), which reduces the number of MSUs to do the same amount of work, or lets you do more work with the same number of MSUs.



#### z13 and z13s Memory Opportunities



#### z13/z13s Memory

- If you have lots of memory, <u>exploit</u> it!
- SOME exploiters of large pages and large memory are: DB2, MQ, IMS, WAS, Java, Sort, CICS, VSAM RLS (future), VSAM LSR, BLSR, SMB
  - z13 and z13s have separate TLBs for 4KB, 1MB, and 2GB pages, so maximizing use of large pages increases the effectiveness of all 3 TLBs.
- See IBM Techdocs White Paper <u>WP102461</u>, IBM z Systems: Performance Report on Exploiting Large Memory for DB2 Buffer Pools with SAP by Brenda Beane, Seewah Chan, and Paul Lekkas
- Also, Techdocs White Paper <u>WP102464</u>, System z: Advantages of Configuring More Memory for DB2 Buffer Pools, by Brad Snyder and Mark Wisniewski
- Also review Redpaper <u>REDP-5238</u> Benefits of Configuring More Memory in the IBM z/OS Software Stack
- SHARE session <u>19878</u> How to Leverage Large Memory on z by Elpida Tzortzatos



#### z13 and z13s Memory Opportunities



One of the best things you can do for z13 performance is to avoid situations where a CP in one drawer has to access memory in a different drawer.

- PR/SM can move CPs and zIIPs between drawers. And starting with z13, it can also 'move' memory between drawers.
- It can configure memory online in one drawer and config off a corresponding amount of memory in another drawer in an attempt to have all of an LPAR's memory in the same drawer as its CPs and zIIIPs.
- **But**, it is limited by what is physically installed in each drawer. 'Smart' planning when you are ordering memory can increase the amount of memory that PR/SM has to play with...



#### z13 and z13s Memory Opportunities



What you paid for

	Customer Memory	Model N30	Model N63		Model N96			Model NC9 and Model NE1			
1	(GB)	CPC drawer 1	CPC drawer 1	CPC drawer 2	CPC drawer 1	CPC drawer 2	CPC drawer 3	CPC drawer 1	CPC drawer 2	CPC drawer 3	CPC drawer 4
	64	320	320	320	320	320	320	320	320	320	320
	96	320	320	320	320	320	320	320	320	320	320
	128	320	320	320	320	320	320	320	320	320	320
	160	320	320	320	320	320	320	320	320	320	320
	192	480	320	320	320	320	320	320	320	320	320
	256	480	320	320	320	320	320	320	320	32U	320
	320	640	320	320	320	320	320	320	320	320	320
	384	640	480	480	120	320	320	320	320	320	320
	448	960	480	480	320	320	320	320	320	320	320
	544	960	480	480	320	320	320	220	320	320	320
	640	960	640	640	480	480	480	320	320	320	320
	736	1280	640	640	480	480	480	320	320	320	320
	832	1280	640	640	480	480	480	320	320	320	320
	928	1280	960	960	480	480	480	480	480	480	480
	1056	1920	960	960	640	640	640	480	480	480	480

What IBM installs





#### **z13 Performance Tips - SIIS**



- z13 really doesn't like programs that store into the instruction stream (SIIS).
   IBM has warned against this for years, but it is more of an issue on z13.
- Modern compilers are designed not to do this, so this is unlikely to be an issue for high level languages compiled with reasonably up to date compilers. However, if you have Assembler programs that do a lot of this, that could be an issue.
  - Likely to show up as programs that use a lot more CPU on z13 than on previous processors.
  - To detect this situation, use SMF Type 113 records:
    - (E163/B2)\*100 indicates the relative percentage of SIIS L1 I-cache misses compared to the total number of L1 I-cache misses as counted by B2 (Basic Counter Set).
    - Can also use SMF30COUNT data to identify specific jobstep.
- See User Experience article in Tuning Letter 2015 No. 4.



#### z13 Performance Tips - Compression



- Hardware compression (not to be confused with zEDC) uses about 50% less CPU on z13 than on zEC12.
  - Hardware compression is still important, even if you have zEDC.
  - zEDC is better than hardware compression FOR THE DATA TYPES THAT SUPPORT IT.
  - VSAM KSDS, IMS, and DB2 do NOT support zEDC.
- If you looked at database compression in the past and discounted it because of CPU cost, go back and look at it again.
  - If you encrypt your data, compressing it first reduces compression cost.
- If you use database compression today in IMS or DB2, expect to see CPU savings when you move to z13.
  - DB2 V12 will support use of zEDC for LOB databases.



### z13 Performance Tips – zEDC Compression SH△RE



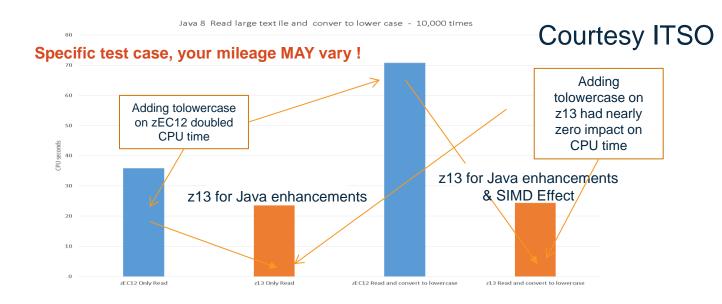
- Recently closed DFSMS New Funciton APAR OA50637 reduces CPU time to compress/decompress BSAM/QSAM data sets with zEDC.
- Improvement is obtained by fixing buffers used to communicate with zEDC.
- Customer measurements have seen CPU savings between 20% and 50%, and elapsed time savings of 15%-22%.
- Based on this, you might consider specifying a PERMFIX value in SMFPRMxx member if you use zEDC for your SMF log streams.
  - Our customers typically see 8x-10x compression for SMF data.



#### z13 Performance Tips – SIMD



If you have lots of Java, SIMD might deliver significant CPU savings...

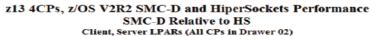


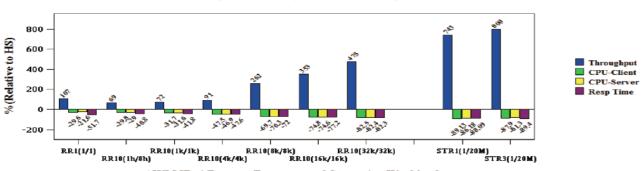


#### z13 Performance Tips – SMC-D



z13s/z13 GA2 introduced Internal Shared Memory (ISM). First exploiter of ISM is Comms Server, with SMC-D.





**AWM IPv4 Request Response and Streaming Workloads** 

The *worst* case compared to HiperSockets shows a 69% throughput improvement and a 21% decrease in CPU consumption. When compared to OSA Exp4, the results are even more impressive, with a *minimum* throughput improvement of 593%! The implementation effort is *not* significant – add a definition to HCD and change one keyword in your TCP Profile For more info, see **Gus Kassimis** Session 19215 Shared Memory Communications protocol - Direct Memory Access (SMC-D) - Going Beyond HiperSockets.



#### z13 Performance Tips – GDPS K Systems



- For GDPS/PPRC customers, the GDPS 'K' systems play a vital role in recovering your systems following a failure.
- For this reason, IBM recommend giving those LPARs a very high relative weight.
- However, on z13 in particular, this can have a negative impact on your production systems' capacity.
- Even though they normally use nearly no capacity, their high weight means that they
  get most of the vertical highs, resulting in the production systems being stuck with
  mediums and lows.
- Don't just blindly implement recommended values. Consider the capacity available to K systems in a disaster – will all the other LPARs on the CPC be in a Freeze? If so, high weights are not necessary, as long as the K system has enough CPs.
- This is more of an issue on z13 than on previous generations because z13 is a lot more sensitive to vertical highs.





- A real challenge for the z/OS platform is how to maintain the stability and reliability that customers require for their mission critical applications, but at the same time address the perception that z/OS is always behind other platforms, always too slow to deliver new function?
- To help address the challenge from other platforms, IBM is moving to a new 'agile' delivery strategy for z/OS-based products. This will facilitate the roll out of some new features without waiting for the next release.
  - For example, z/OSMF delivers a package of new functions roughly every six months. IBM would like to expand that concept to other z/OS-based products.
- IBM's name for this strategy is 'Continuous Delivery'.





- We are 100% supportive of any change by IBM that:
  - Makes z/OS more competitive with other platforms.
  - Makes new functions available sooner for those customers that wish to exploit them.
  - Does not impact the stability of the platform.
  - Continues to support the tried and tested and understood APAR/PTF/RSU/Release paradigm and tools.
  - Does not reduce customer's flexibility about when and how they maintain and upgrade their software.
  - Does not add even more complexity to z/OS-based software management..
  - Does not increase costs for z/OS-based software or support.



- In March this year, IBM published a Redpaper titled 'IBM z/OS Continuous Delivery', REDP-5340, that provides some information about this new Continuous Delivery strategy.
  - The Redpaper was a good attempt to describe how this will work for z/OS products. However, it was missing information that we would have liked to see – for example, how will customers find out about new functions that are delivered mid-release (this is in the process of being addressed)? This indicates that this is still an evolving strategy.
- Other than the Redpaper, there was no formal announcement that we are aware of about this change in IBM's software delivery strategy.

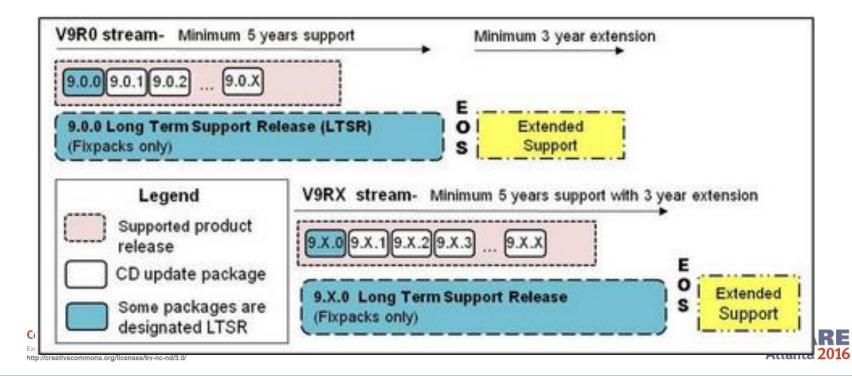




- Then, on May 24, IBM made a formal announcement 'IBM Software Support Lifecycle Policy is enhanced with a continuous delivery support model for eligible, on-premise IBM software products'. http://www-01.ibm.com/common/ssi/rep\_ca/9/897/ENUS216-269/ENUS216-269.PDF
- This announcement, in our opinion, raises a number of questions that you must consider. Specifically, based on our reading of the announcement, we have these concerns:
  - It is not clear which products it pertains to (but we heard offline that it does NOT apply to z/OS).
  - It is somewhat vague about whether all products will employ the same strategy (but we believe that they will not).
  - The announcement was not clear about how this new strategy will affect how long products will be supported for (it appeared that at least some products will move to a release being supported for just 2 years, after which extended support will be available for a fee).
  - It doesn't say anything about how customers will be notified about new functions.
  - It doesn't say anything about how product sets will support this for example, will every CICS-related product use the same methodology?
  - It appears that not every product will necessarily adhere to this new strategy customers need to consider how this will affect their SW Management strategies.



 As ONE example of how this is being implemented, this is the announced MQ strategy (see the MQ FAQ <a href="http://www.ibm.com/support/docview.wss?uid=swg27047919">http://www.ibm.com/support/docview.wss?uid=swg27047919</a>):





- IBM announced the Open Beta for CICS TS 5.4 on July 12, 2016 announcement letter 216-036.
- The announcement states: 'IBM is committed to a continuous delivery roadmap that delivers regular capability updates for current releases of CICS TS. The following key features are supported on this continuous delivery release of CICS TS V5.3.:' and then lists New Functions that have already been shipped.
- However, it does not say anything about dual streams or any of the other complexity described in the MQ announcement. It would appear that, in terms of how the code will be delivered and maintained, the CICS CD strategy is a continuation of existing CICS practices.





- We believe that it is <u>VITAL</u> that customers carefully review this announcement AND any additional information.
- Speak to your ISVs to see how/if they will support this new strategy.
- Discuss with your colleagues and developers and then send your feedback (positive or otherwise) to IBM and to your User Group representatives <u>as soon</u> <u>as possible</u>. This strategy still appears to be evolving, so it will be much easier for IBM to fine tune it *now*, than to have to re-engineer their processes again later. If you decide that you don't like it, but don't tell IBM, you can't really complain later.
- This is a very active area in IBM at the moment, so make sure to remain vigilant and monitor announcements in this space.





- For 50 years, we have had SMF data to provide us with an unparalleled insight into what happened on our systems.
- But the data was always historical reporting on things that happened some time in the past.
  - This doesn't detract from the value of the data, but it places bounds on how we can use that information.
- Now, IBM will soon release the ability to easily access SMF records in near-real time (this is actually an excellent example of the new Continuous Delivery strategy).
- This is a real game changer now we can potentially alter system behavior based on what is happening NOW.





- The new capability is based on System Logger support in SMF, so requires that SMF is running in LOGSTREAM mode.
- You control, at the record type level, if record is written to one or more of: DEFAULT Logstream, user-defined Logstream, or a wrap-around, in-memory buffer.
- There is a new system service (IFAMaaaa) for your program or IBM /ISV products to connect to buffer, retrieve the 'next' record, and disconnect.
  - No need for APF authorization, no ability to change or suppress SMF records, access is controlled using SAF.
- SMF Streaming support will be available for z/OS 2.1, 2.2 with PTFs for APAR OA49263.

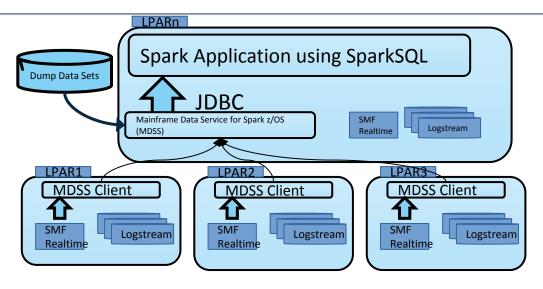




- Now that you have access to SMF records as soon as they are sent to SMF, the next step is to have a way to run analytics against the records quickly enough to be able to take action in real-time.
- Spark is an open-source analytics tool (written in Scala) that runs on z/OS, using inmemory database (for lightning-fast performance). It supports queries written in Scala, Java, or Python, and runs nearly entirely on zIIPs (cheap MIPS and no impact on your MLC bill).
- IBM's vision for Spark is that it would be used in business applications for things like more targeted marketing.
- However, the bit we really like is that the <u>IBM z/OS Platform for Apache Spark</u> product that you can download from Shopz <u>includes Rocket Software's Mainframe</u> <u>Data Service for Apache Spark on z/OS (MDSS)</u>. This allows you to run SQL queries against historical and streaming SMF data:
  - MDSS is built on Rocket's DVS, meaning that you can use Spark to run SQL queries against SMF data <u>from sequential data sets (even on tape)</u>.
  - AND, MDSS comes with mappings for the most popular IBM SMF records.







- Available now from Shopz, product number 5655-AAB.
  - Package is FREE, Service and Support (chargeable) is optional.
  - Code runs nearly 100% on zIIPs, so nearly no impact on SW bills.
  - There is a discount for memory bought for use by Spark.
  - There is a discount for zIIPs bought for use by Spark
- We will have an article about it in the September Tuning Letter.



## WLM Mobile Workload Pricing Support



- WLM APAR <u>OA47042</u> provides new capabilities aimed at MWP.
  - There is a new Reporting Attribute in WLM that lets you identify work as being MWP-eligible. WLM then tracks CPU (general, zIIP, zIIP on CP) usage of those 'transactions' at the service class, report class, and system level. It also tracks the R4HA of the MOBILE workload.
- For anyone using MWP, this is huge.
  - Potentially eliminates the need to process SMF 110, IMS logs, and other transaction-level (very high volume) SMF records.
  - IF you have regions dedicated to MWP, can also use this capability to capture region overhead time.
  - Lets you calculate the MWP-adjusted R4HA in real time this is vital to anyone that wants to effectively manage soft-caps on a dynamic basis.
  - AND, as soon as the APARs are installed (even if you don't assign ANY txns to MOBILE), CICS and IMS transaction service classes now report txn response times, txn numbers, AND CPU TIMES – woohoo!



#### WLM Mobile Workload Pricing Support



#### Requirements:

- WLM APAR <u>OA47042</u>, RMF APAR <u>OA48466</u>, and z/OSMF APAR PI47638.
- CICS TS 5.3 or later.
- IMS V14 or later with APARs <u>PI46933</u> and <u>PI51948</u>.

#### References:

- Session <u>19098</u>, Containing MLC Costs For Mobile Workloads
- Session <u>19668</u>, Workload Management (WLM) Update for Mobile Pricing, Capping, z13 and z/OS 2.2, by **Andreas** Henicke



### **WLM Capping Support**



 Just when you were despairing that there are no more ways to hog-tie your systems, WLM comes to the rescue....

Initial capping

Resource groups

Group capping



Absolute capping

Soft capping



## WLM New Capping Support



- z13 GA2 provides a new group absolute cap capability.
  - Cap amount and group name are specified on the HMC. Cap is for a group of LPARs and is specified in terms of hundredths of a CP.
  - Just like absolute capping introduced on zEC12 GA2, but for a group of LPARs.
  - Aimed primarily at Linux environments, but all PU types supported.
  - Limit is enforced by PR/SM, but WLM (with APAR <u>OA47752</u> for z/OS 2.1 & 2.2) is aware of the limit and uses it when calculating potential LPAR capacity.
  - CONSIDER THIS AS AN ALTERNATIVE TO DUMMY LPARS WITH HIGH WEIGHTS THAT MESS UP HIPERDISPATCH TOPOLOGY.



### WLM New Capping Support



- WLM APAR <u>OA49201</u> introduces ability to have WLM cap an LPAR based on actual MSU (not Rolling 4-Hour Average).
- Controlled using AbsMSUcapping keyword in IEAOPTxx and existing Soft Capping MSU value specified on HMC.
  - Can mix AbsMSUCapping YES and NO systems in the same LPAR group.
- Works for both individual LPAR and LPAR Groups.
- Presumably aimed at situations where you want to limit the actual capacity available to an LPAR or group of LPARs.



#### **SHARE Live!**



- Did you miss some key presentations?
- Did you hear about some fantastic presentation that you did not attend?
- Maybe you had multiple sessions that you wanted to see at the same time and had to choose just one of them?
- Did you attend a session that you would like your colleagues to see/hear?

 If you answered yes to any of the above, consider signing up for SHARE Live! See <a href="http://www.share.org/salive">http://www.share.org/salive</a>



#### Thank you



- Thank you for coming and for your support.
- Have a safe trip home.
- Don't forget to follow up on all those
   "I must look at that when I get home" items.



- Please complete an evaluation Session 19877
- See you in San Jose!



