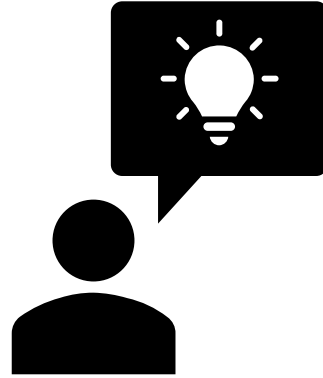# How well is your mainframe outsourcer managing capacity and performance?

Steven Thomas, Chief Technology Officer, SMT Data
Steven.Thomas@SMTData.com
www.SMTData.com

Understanding what you pay for

What do you pay for?

- CPU
- Storage
- Memory
- Software
- Backup
- Security
- SLA
- Services
- Agility
- Complexity
- Terms and Conditions
- And more…

What do you pay for?

- CPU
- Storage
- Memory
- Software
- Backup
- Security
- SLA
- Other Services
- Agility
- Complexity
- Terms and Conditions
- And more…

- CPU based pricing usually includes some software and some services
- Special terms and conditions are often also rolled into the price – e.g. location, ownership, SLA, design authority, reporting requirements and much more
- Contracts are often fuzzy on:
  - What units are used for billing (e.g. cpu seconds, MIPS, MSU)
  - How those units are measured or calculated
  - Guaranteed and max capacity (and how these affect cost and quality of service)
 - Reporting and data sharing

CP, GP, zIIP, IFL

- General Processors ('GP' or simply 'CP')
  - Can run all workloads

- Integrated Information Processors ('zIIP')
  - Physically the same as a GP, but with microcode changes to limit functionality
  - Primarily for certain types of DB2 workloads, but also other stuff
  - A 'commercial' not a 'technical' invention, allows IBM to provide a discount for certain workloads

- Integrated Facility for Linux ('IFL')
  - Like a GP, but with certain instructions disabled so it can only run Linux
 - Also a way of providing a discount for selected workload
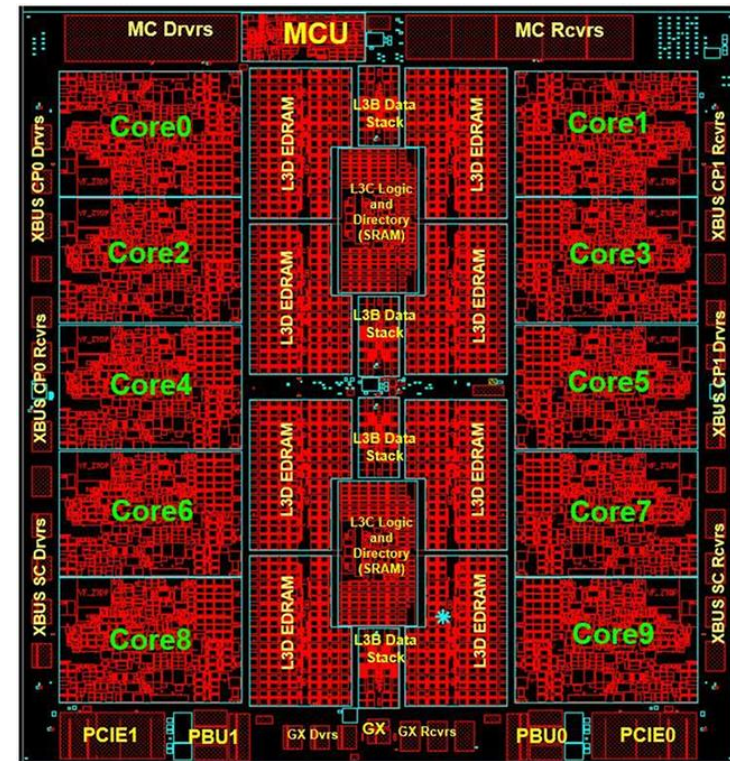
Measuring CPU – CPU Seconds

- The number of seconds that a CPU has been busy on a certain activity

- Measured and reported directly by the operating system in SMF

- Not normalized – faster hardware will take fewer CPU seconds to perform the same work

- If used for billing purposes, then the contract normally includes a price adjustment mechanism for hardware or configuration changes
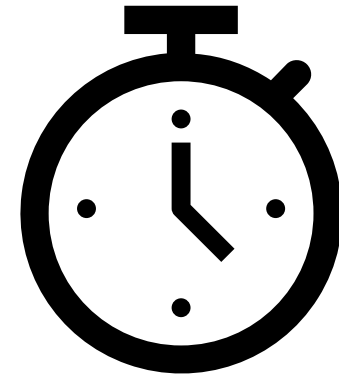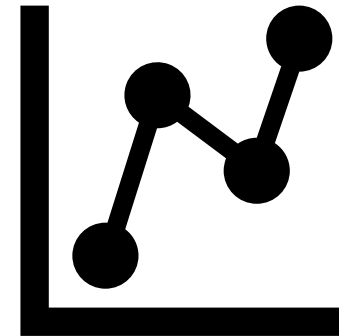
Measuring CPU - MIPS

- Originally 'Million Instructions per Second', but no longer directly related to number of instructions

- A way of normalizing CPU usage across different hardware models and configurations

- <u>In theory</u>, the same workload uses <u>roughly</u> the same number of MIPS no matter what configuration it is run on

- Cannot be measured directly

- Calculated from CPU seconds and a configuration dependent MIPS factor (more on this later)
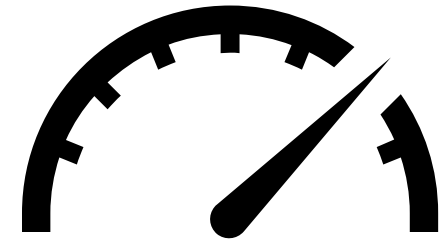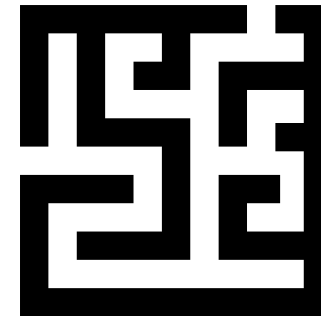
Measuring CPU - MSU

- Originally 'Million Service Units'

- Calculated based on factors determined by IBM

- Like MIPS, a way of normalizing CPU usage across different CPU models and hardware configurations

- Used by IBM (and others) to charge for software

- In theory, the same workload uses roughly the same number of MSU's no matter what configuration it is run on

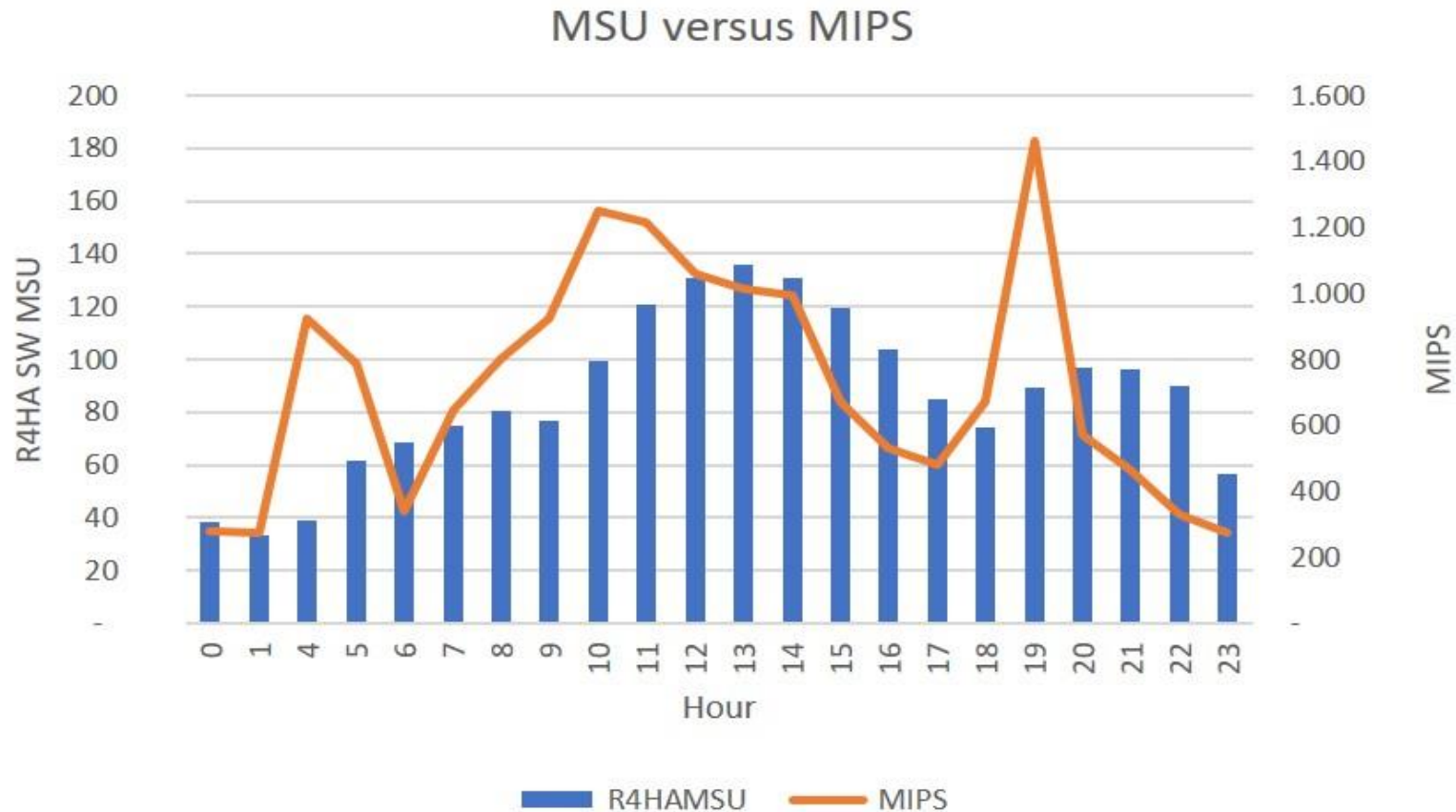- But IBM adjusts the factor for different HW for commercial reasons

Measuring CPU - MSU

The term MSU can mean different things depending on the context:

- The actual hardware MSUs consumed during a period – roughly corresponds to 5-5.5 MIPS

- The actual software MSUs consumed during a period – roughly corresponds to 8-8.5 MIPS

- The rolling four-hour average of consumed software MSUs as logged by the operating system every few minutes. This is reported in SMF.

- The hourly average of these individual rolling four-hour averages. This is reported in the SCRT report.

- The monthly peak of these hourly averages. This is the number on the SCRT report that IBM uses to bill software.

Comparing MIPS and MSU

MIPS versus MSU for billing purposes

	MIPS	**R4HA SW MSU**
Advantage	• Directly computed from CPU seconds in a (hopefully) transparent way • Independent of IBM's pricing policy • Applies to CP and zIIP • Can be computed on all levels	• R4HA Flattens out spikes • Measured directly = no room for interpretation • IBM gives a 'bonus' for upgrading to newer hardware • Can give alignment of interests with outsourcer
Disadvantage	• Can be spikey • Computation method open to discussion – must be contractually agreed	• Affected by IBM's pricing policy • Only measured at the LPAR level • Only includes CP, not used for zIIP billing • Risk of confusion (e.g. with HW MSU)

Calculating MIPS

- MIPS = CPU seconds * configuration dependent MIPS factor

- For contractual purposes, the MIPS factor is generally determined based on:
  - 'Effective' machine model – determined by model and number of CPU's
  - MIPS per CPU second delivered by that model – determined by a standard table (e.g. Watson Walker or IBM)

- Many other factors can affect the actual MIPS factor such as
  - Workload characteristics
  - Vertical affinity
  - Overall system load
  - And many others (more on this later)
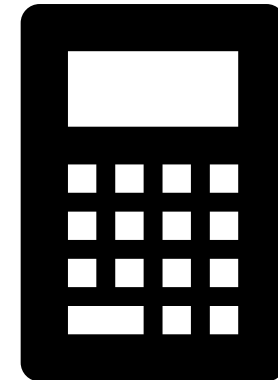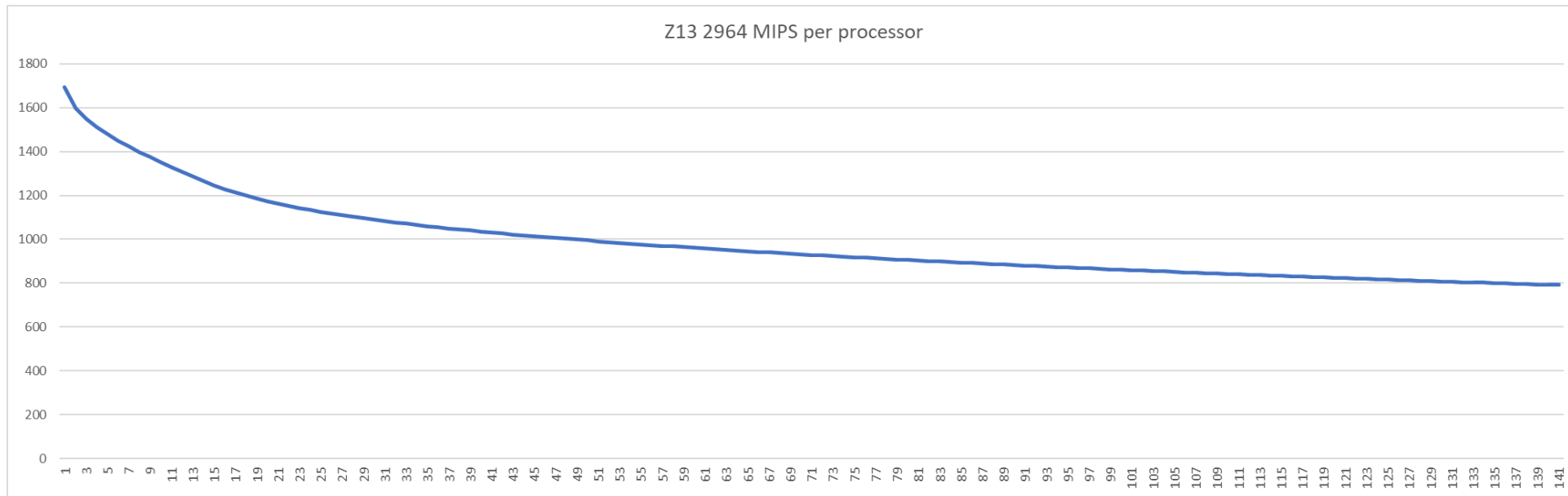
Finding the 'effective' machine model

The 'effective' model is determined by the hardware model and the number of processors:

- More processors on a machine means:
  - More overhead
  - Fewer MIPS delivered per CPU second

- Newer models generally deliver more MIPS per CPU than older ones

An example

- A z13 model 2964-733 has 33 processors and delivers about 1071 MIPS per CP

- Adding 2 more processors to this configuration (model 2964-735) reduces the MIPS rate to 1060 MIPS per CP

- Normally the outsourcer will compensate for this in the MIPS calculation so the impact of configuration changes on the customer is small

Z13 2964 MIPS per processor

SMT Data
Specialists in IT Business Intelligence

Finding the 'effective' machine model

Pitfalls:

- Calculating MIPS as if the LPAR was a stand-alone machine. This gives a much higher MIPS rate.

- Include only CP's, ignoring zIIPS, in determining the 'effective' model. This can give an unrealistically high MIPS rate.

- MIPS tables based on the outsourcer's own measurements and methodologies

- Many other 'creative' models are used to determine the MIPS factor

- Many contracts are silent on how the MIPS factor is determined

Finding the MIPS Rate

- Use a table (here thanks to Watson Walker) to determine the MIPS/CP for the effective model

- Here we use 'Avg RNI'. (More on RNI in a minute)

| Type-Model	# of CPs	IBM PCI	Avg RNI MIPS	Avg RNI MIPS/CP	MP	Low RNI MIPS	Low-Avg RNI MIPS	Avg-Hi RNI MIPS	High RNI MIPS
2964-732	32	34456	34455.7	1076.7	0.64	42741.7	38154.0	31471.0	28962.2
2964-733	33	35341	35341.3	1070.9	0.63	43903.2	39159.7	32275.3	29698.8
2964-734	34	36223	36222.5	1065.	0.63	45058.8	40160.4	33076.0	30432.5
2964-735	35	37099	37099.3	1060.0		46208.7	41156.0	33873.2	31163.2
2964-736	36	37972	37971.7	1054.8	0.62	47352.9	42146.6	34666.8	31891.0
2964-737	37	38840	38839.8	1049.7	0.62	48491.3	43132.2	35456.8	32616.0

Calculate MIPS

For billing purposes, MIPS are normally calculated as an average over one hour:

MIPS = (cpu seconds per hour / 3600) * MIPS per processor

e.g. if we use 1000 cpu seconds in an hour on a z13 model 733 we get:

(1000 cpu seconds / 3600) * 1070.9 = 297.5 MIPS
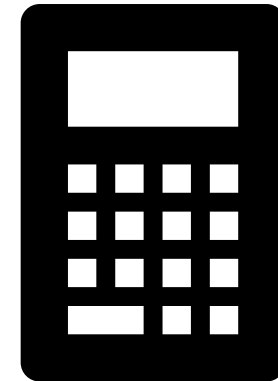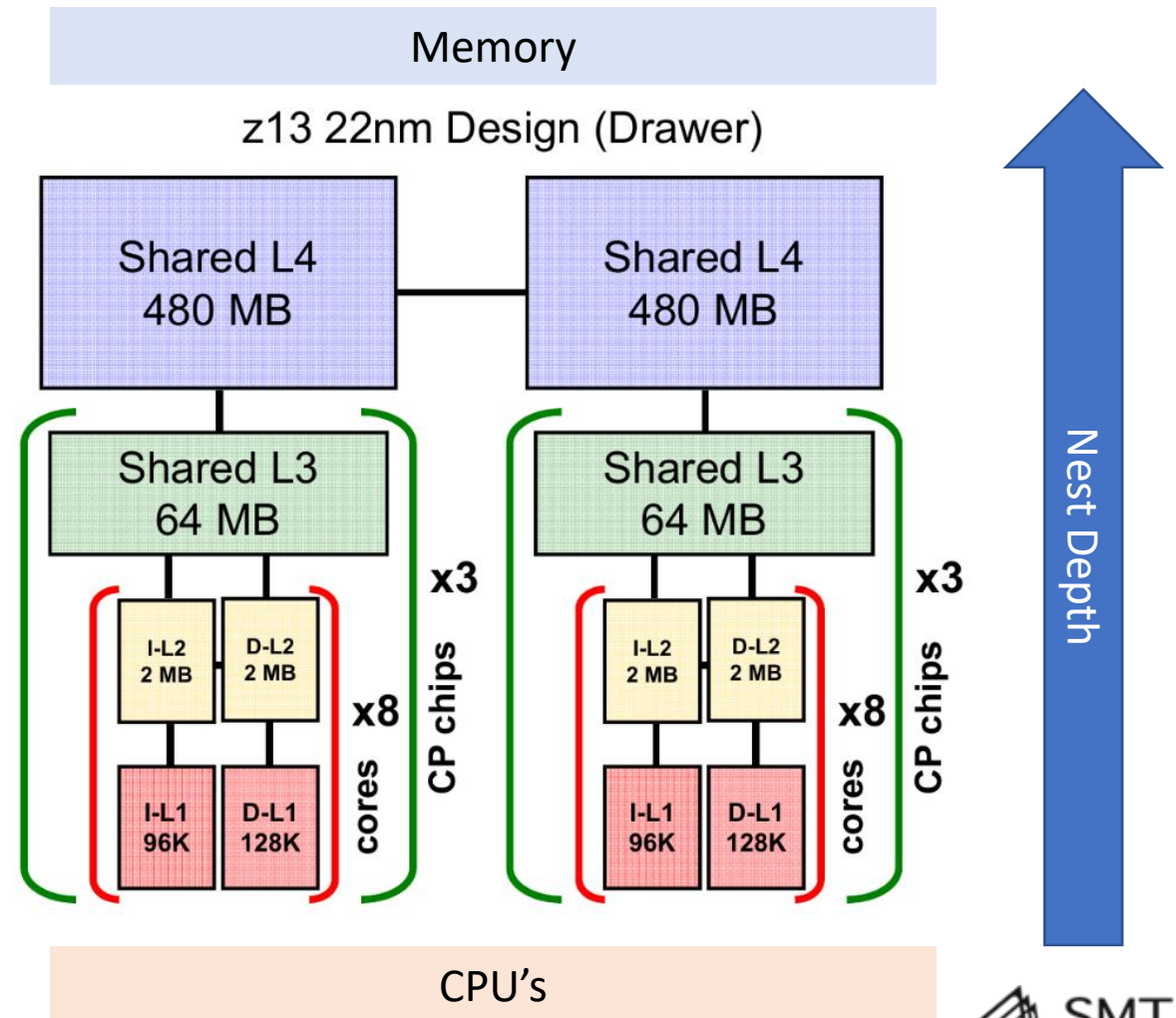
Relative Nest Intensity - RNI

There are multiple layers of cache on a mainframe.

Those layers closest to the CPU are the fastest.

The deeper into the cache hierarchy or 'nest' that an instruction must go to find its data, the more CPU time is spent waiting. This depth is referred to as the Relative Nest Intensity (RNI).

To ensure confusion:

- High RNI means deep in the memory nest and gives a low MIPS rate. Which is bad.

- Low RNI means a high MIPS rate. Which is good.

- IBM usually draws the memory nest with the deepest part at the top.

z13 22nm Design (Drawer)

Memory

Shared L4 480 MB — Shared L4 480 MB

Shared L3 64 MB — Shared L3 64 MB

I-L2 2 MB | D-L2 2 MB   x8 cores   x3 CP chips
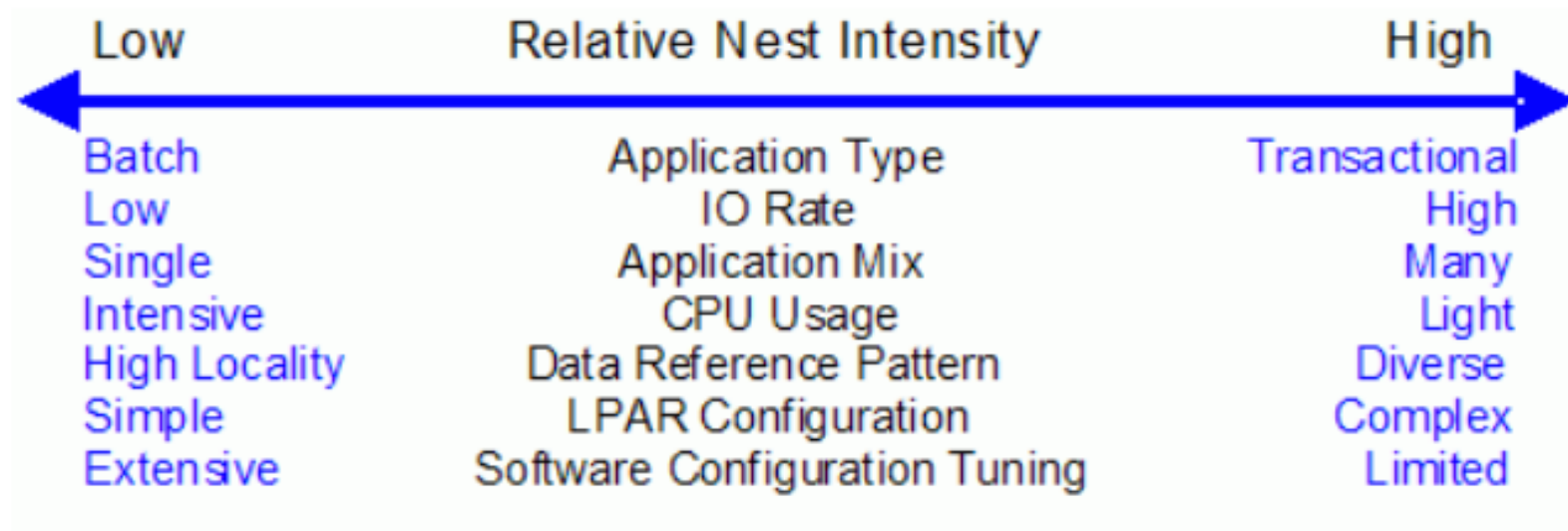
I-L1 96K | D-L1 128K

CPU's

Nest Depth

RNI

Low RNI is 'good'. Less time is spent waiting for cache and memory, so you get more MIPS out of your machine. In other words you use fewer MIPS to do the same amount of work.
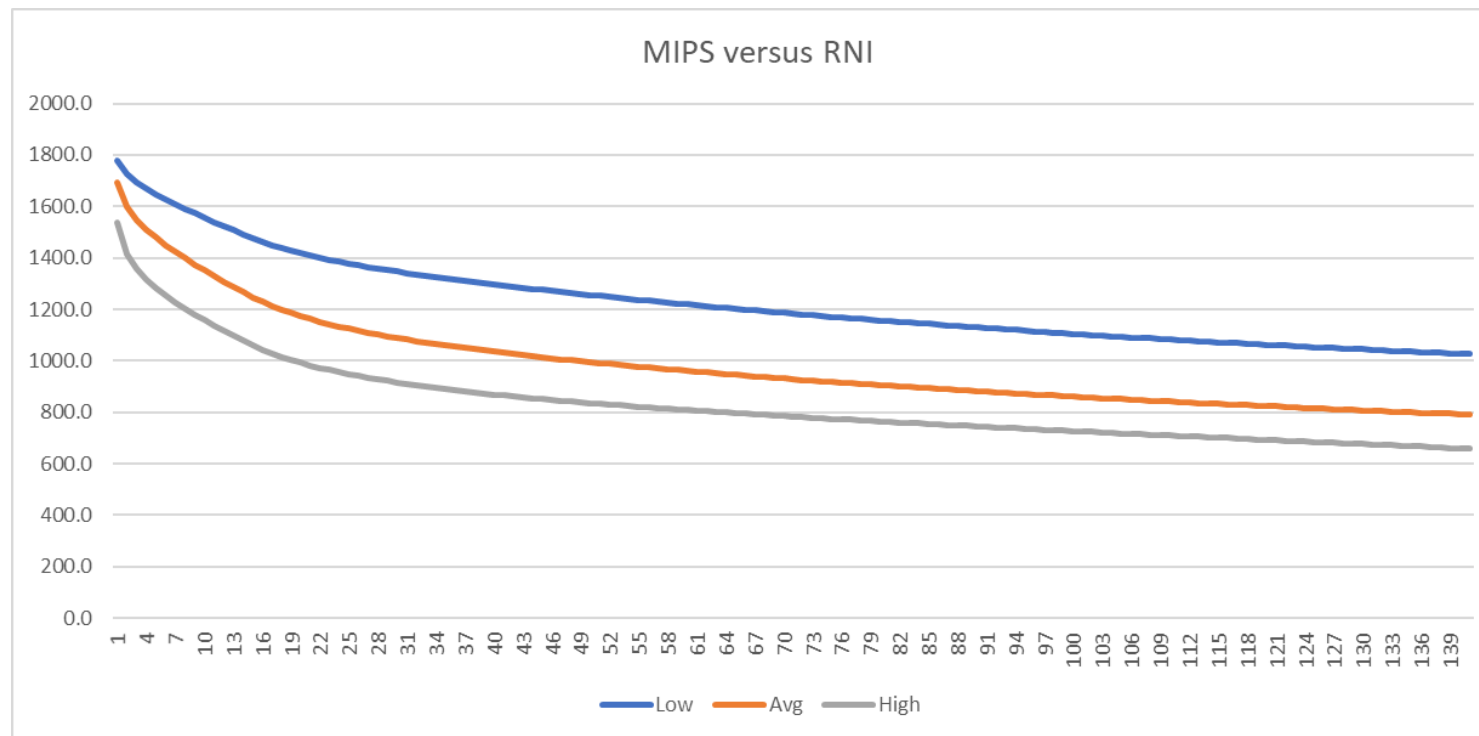
RNI

Many factors affect RNI. Some are in control of the Customer and some in control of the outsourcer. More on this later...

Low	Relative Nest Intensity	High
Batch	Application Type	Transactional
Low	IO Rate	High
Single	Application Mix	Many
Intensive	CPU Usage	Light
High Locality	Data Reference Pattern	Diverse
Simple	LPAR Configuration	Complex
Extensive	Software Configuration Tuning	Limited

MIPS Calculation - Best Practice

- Have a clear definition of MIPS in your contract:
  - Use the total number of CP + zIIP to determine the model.
    - E.g. a z13 model 2964-733 with 33 CP's and 2 zIIP's has a MIPS rate like a model 2964-735 with 35 CP's and no zIIP's
    - Ignore processors that are dedicated or offline
  - Use a MIPS table from an independent source (e.g. Watson Walker)
  - Use 'average' RNI for contractual purposes
  - Base MIPS measurements on CPU seconds as measured in SMF 70
- Ensure reporting or access to data that will allow you to follow up on all the above (and on other things that can affect MIPS rate)
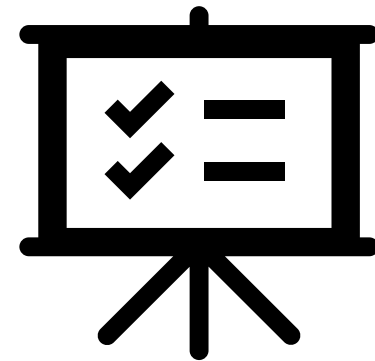
Examples of billing models

- Peak MIPS per hour per month

- Peak R4HA MSU (or MIPS) per month

- Peak hour during working hours only

- Percentile Peak (e.g. 95th Percentile Peak)

- Peaks calculated independently for each LPAR (best for outsourcer)
  - Or combined peak across multiple LPARs (best for customer)

- Peaks for CP and zIIP calculated together or separately

- Fixed capacity (e.g. based on weight, lcpu's, capping)

- Total or average usage per month

- Bands = penalties for using too much or too little

- Combinations of the above e.g. CP usage based on MIPS peak and zIIP usage based on fixed capacity

Billing models best practice

- Work towards alignment of interests with your outsourcer

- Ensure transparency through access to data and reporting

- Understand your usage patterns when considering alternate billing models

- Avoid complexity that doesn't add value

- Understand the relationship between cost and quality of service
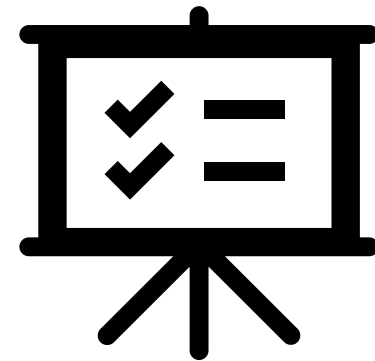
Weight and capping

- Capping determines the LPAR's maximum capacity
  - If there is no capping, then the maximum capacity is determined by the number of logical processors assigned to the LPAR
  - If the max capacity is higher than the weight, then the LPAR can use more than the guaranteed capacity (if the machine has spare capacity)
  - There are various types of capping and they can be 'soft' or 'hard'
    - Soft capping is based on your R4HA, so you can go over your capping for short periods
    - Hard capping does not allow you to go over your limit
- LPAR 'Weight' determines the LPARS' guaranteed capacity (if it is less than the capping)
- Workload Manager (WLM) determines what gets prioritized when you don't have enough capacity

Weight and capping

- Weight, capping and WLM can affect your cost and quality of service
  - Ensure that there is a clear mechanism in your contract to regulate these.
  - Ensure that you have transparency via access to data (SMF 70, 72, and WLM policies) or other reporting.
- Pitfalls:
  - Group capping across multiple customers
  - Peak based billing and no capping, especially during batch periods
  - Too low guaranteed capacity or capping compared to high priority peak usage
  - Assuming the outsourcer has 'spare' capacity during your peak
  - WLM policies not aligned with business priorities
  - Leaving all this up to outsourcer

Following up

Monitoring your configuration (SMF 70)

Overall machine configuration

- Hardware model – to determine the MIPS rate
- Total number of processors (CP's and zIIP) – to determine the MIPS rate
- Total weight of all LPARS on the machine – to determine the guaranteed capacity
- Any group capping – to understand the max capacity

Model	Nick Name	Total MIPS	CP MIPS	IIP MIPS
2964-732	z13	43,958	32,713	11,245

Daily Total # Physical Processors by Type
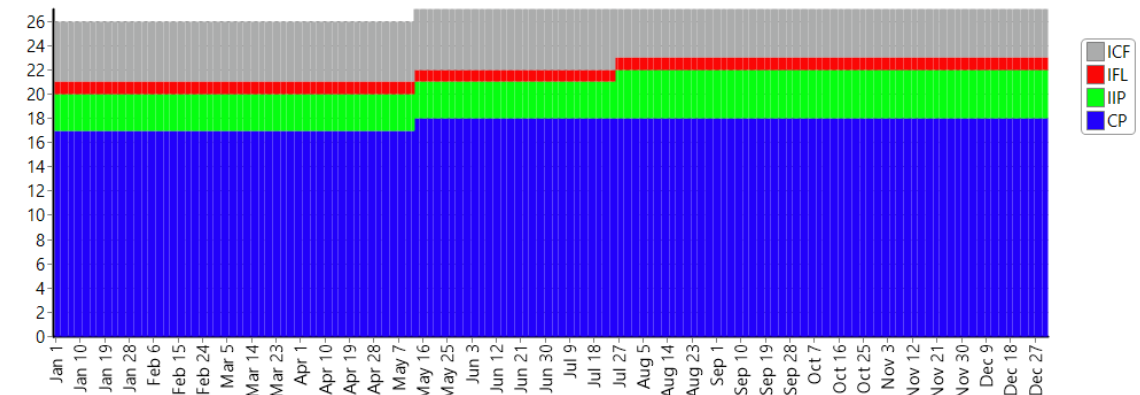
Monitoring your configuration (SMF 70)

LPAR configuration

- Number of logical processors per LPAR (CP's and zIIP) –to determine the guaranteed and maximum capacity as well as the LCP/PCP ratio

- LPAR weight – to determine the guaranteed capacity

- LPAR Capping – to determine the max capacity

Daily # of Logical Processors by Type (for Systems Measured Directly)
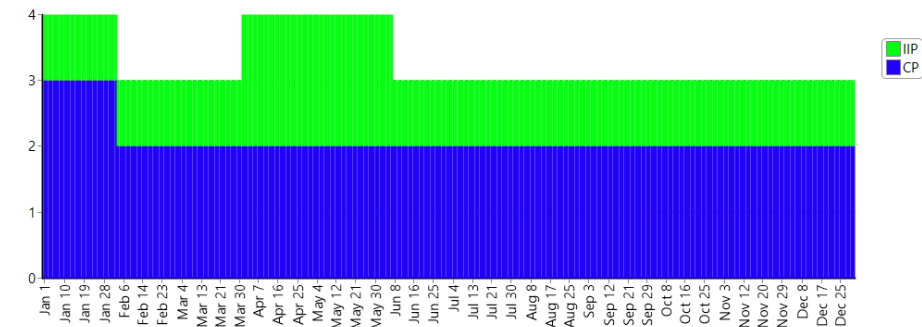
Hourly MIPS Usage & Guaranteed Capacity (Weight) for CP and zIIP Processors
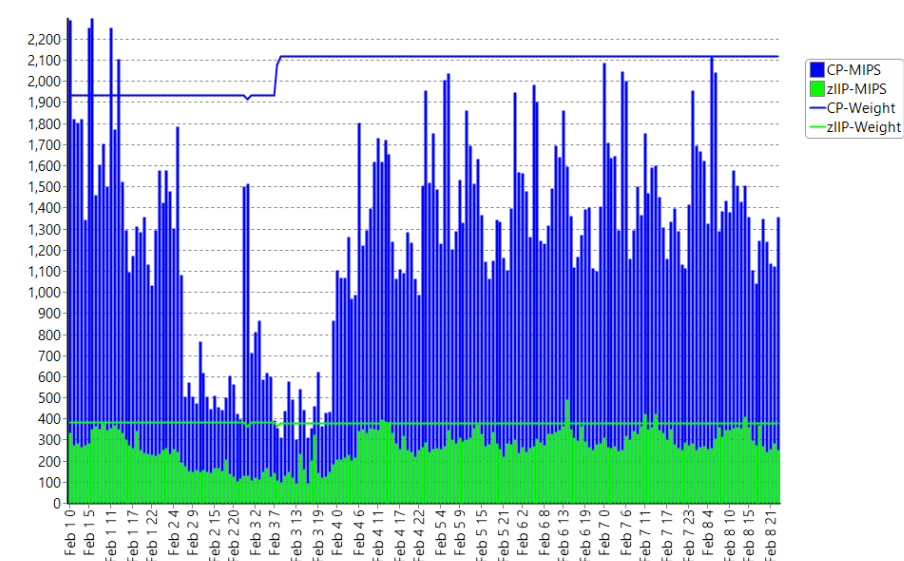
Monitoring zIIP Eligible

- zIIP usage is typically much cheaper than CP usage

- But if there is not enough zIIP capacity, then work that is 'elibible' to run on a zIIP will be run on a CP

- Be aware of your billing model. If you are paying for peak usage, then zIIP eligible workload running on a CP in off peak periods may not impact your cost
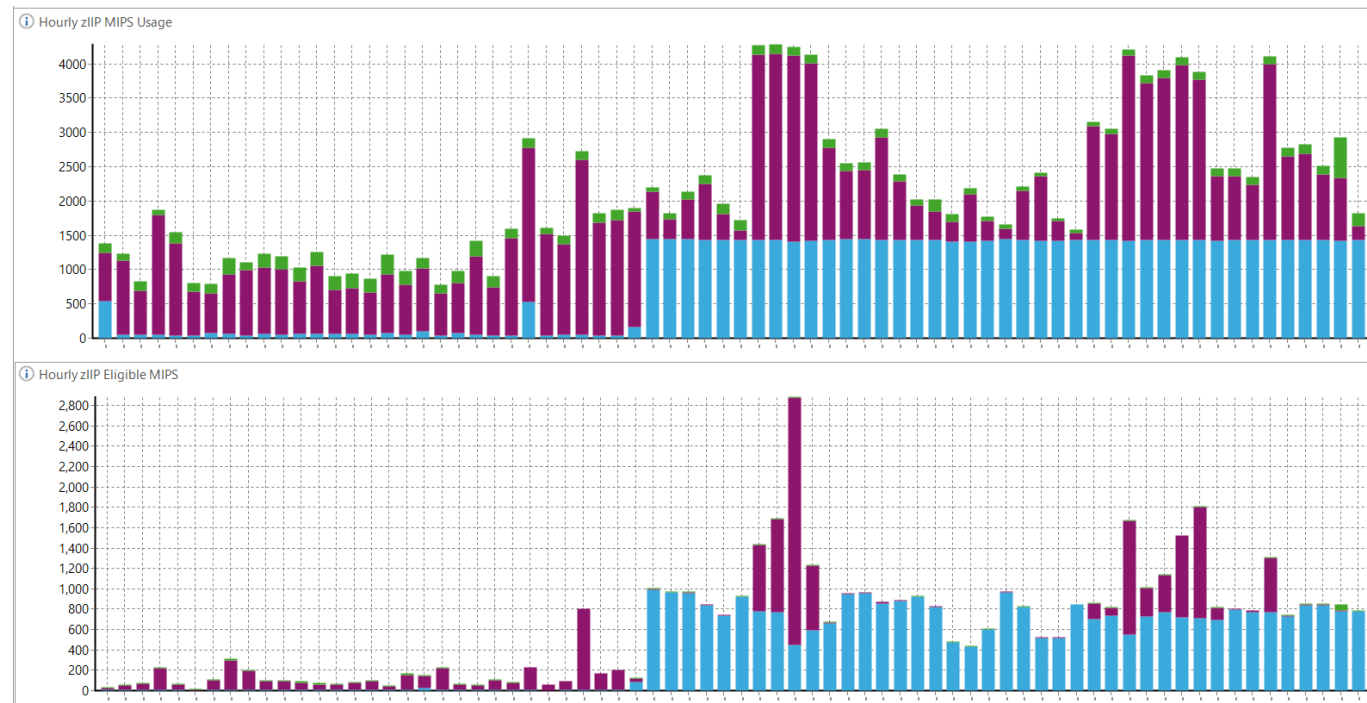
Monitoring your vertical affinity

- Logical cpu's are assigned a 'vertical affinity'

- Vertical affinity determines how likely a logical cpu is to be scheduled on the same physical location.

- High vertical affinity is good, because it increases the likelihood of finding your data already in cache

- High vertical affinity gives a low RNI

- Vertical affinity is determined by the system based on LPAR weight

z13 22nm Design (Drawer)
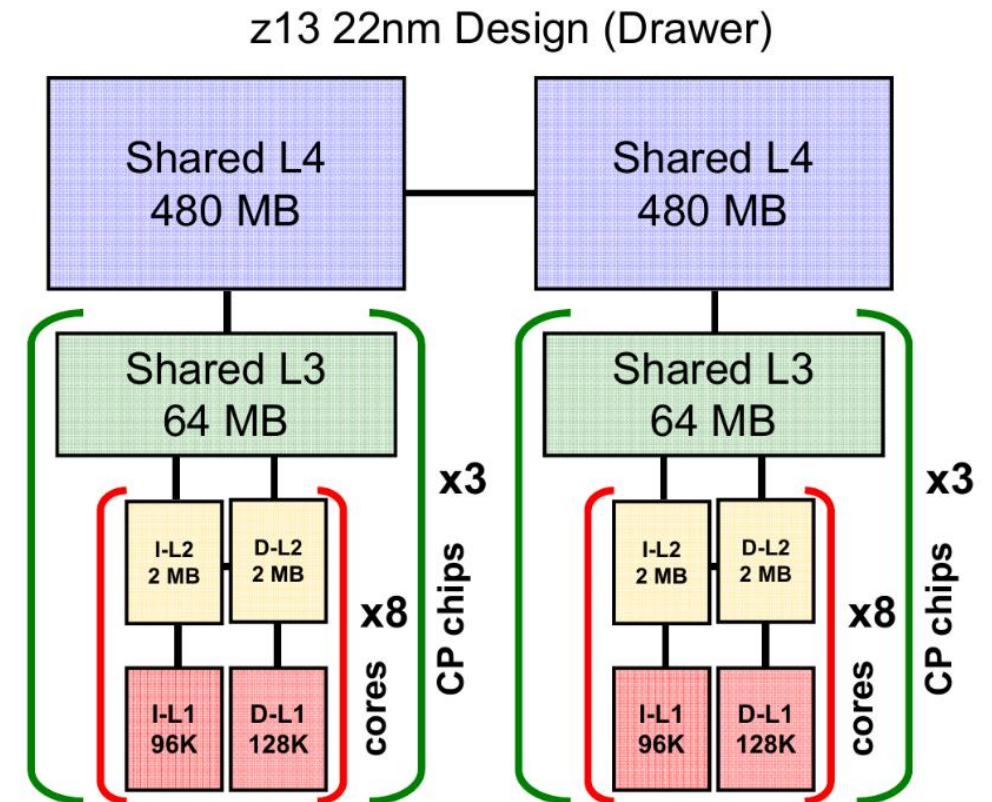
Monitoring your vertical affinity

- Vertical Affinity can be monitored based on the SMF70 record

- Small adjustments in weight or number of LCP's can have a major impact on vertical affinity

- But the tradeoffs can be complex for the outsourcer

- Use the vertical affinity reporting as a basis for dialog with the outsourcer
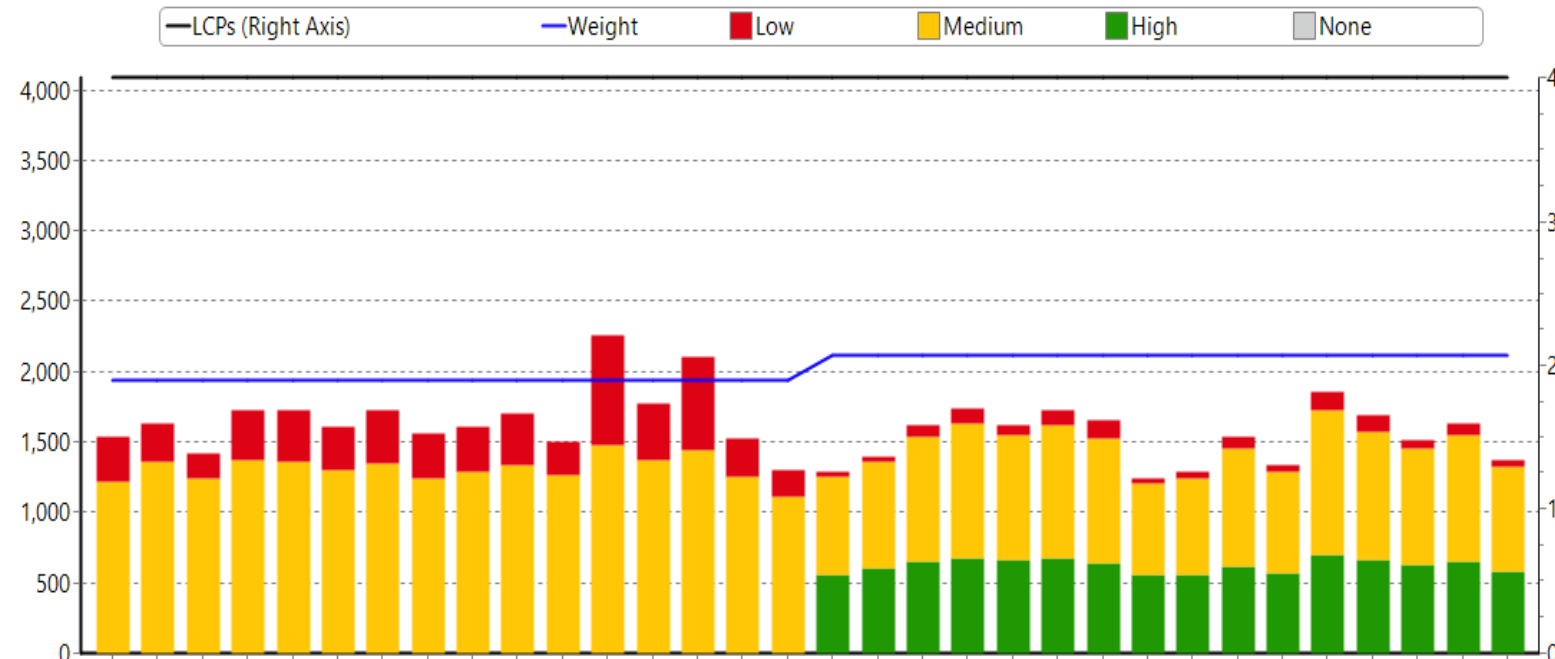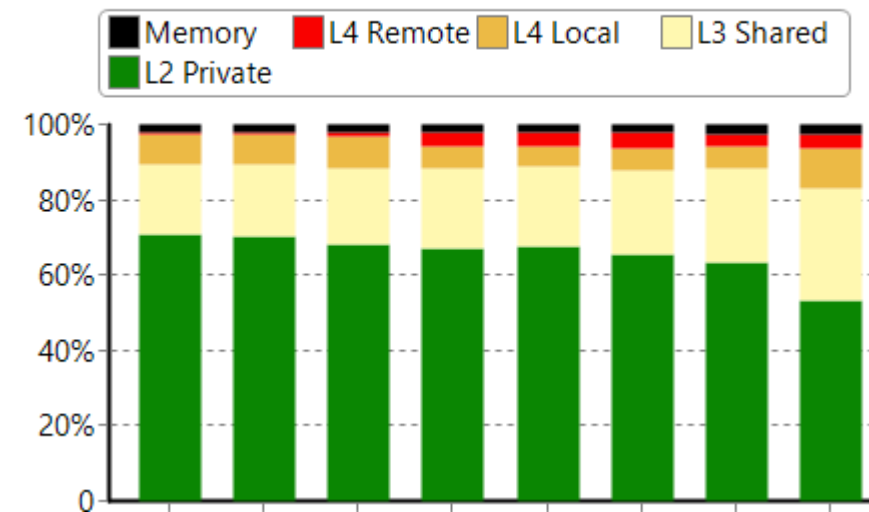
Monitoring your RNI

- RNI and the underlying cache usage statistics can be monitored with SMF 113

- An increased RNI can be due to changes in your usage patterns or applications

- But can also be due to configuration changes

- Use the SMF 113 as a basis for dialog with the outsourcer

**Hourly RNI (Relative Nest Intensity) and % Level 1 Cache Miss**

Legend: —RNI —% L1 Miss
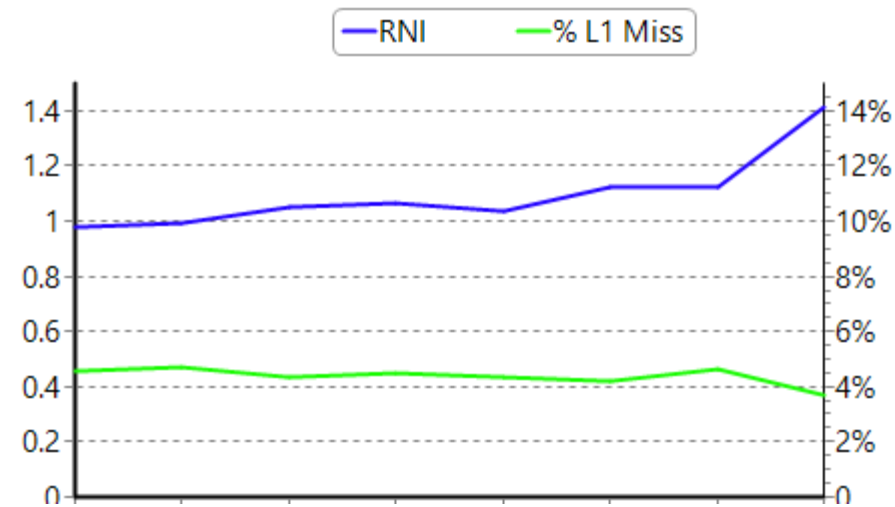
Hourly L1 Cache Sourcing in % of Total

Legend: Memory, L4 Remote, L4 Local, L3 Shared, L2 Private

Monitoring your response time

- SMF 110 gives insight into CICS transactions, including response times

- Response times in other subsystems can also be monitored:
    - SMF101 for DB2
    - IMS log 56FA

- Slow response times can be related to delays (that can be seen in WLM)

- This is the response time within CICS. Response time as seen by the user includes a lot of other factors

Hourly Number of CICS Transactions matching Selected Criteria - Note: Trx with R.T. < 0.5 Sec NOT included

Legend: 0.5 - 1.0 sec | 1.0 - 1.5 sec | 1.5 - 2.0 sec | 2 - 3 sec | 3 - 5 sec | 5 - 10 sec | 10 - 20 sec | > 20 sec
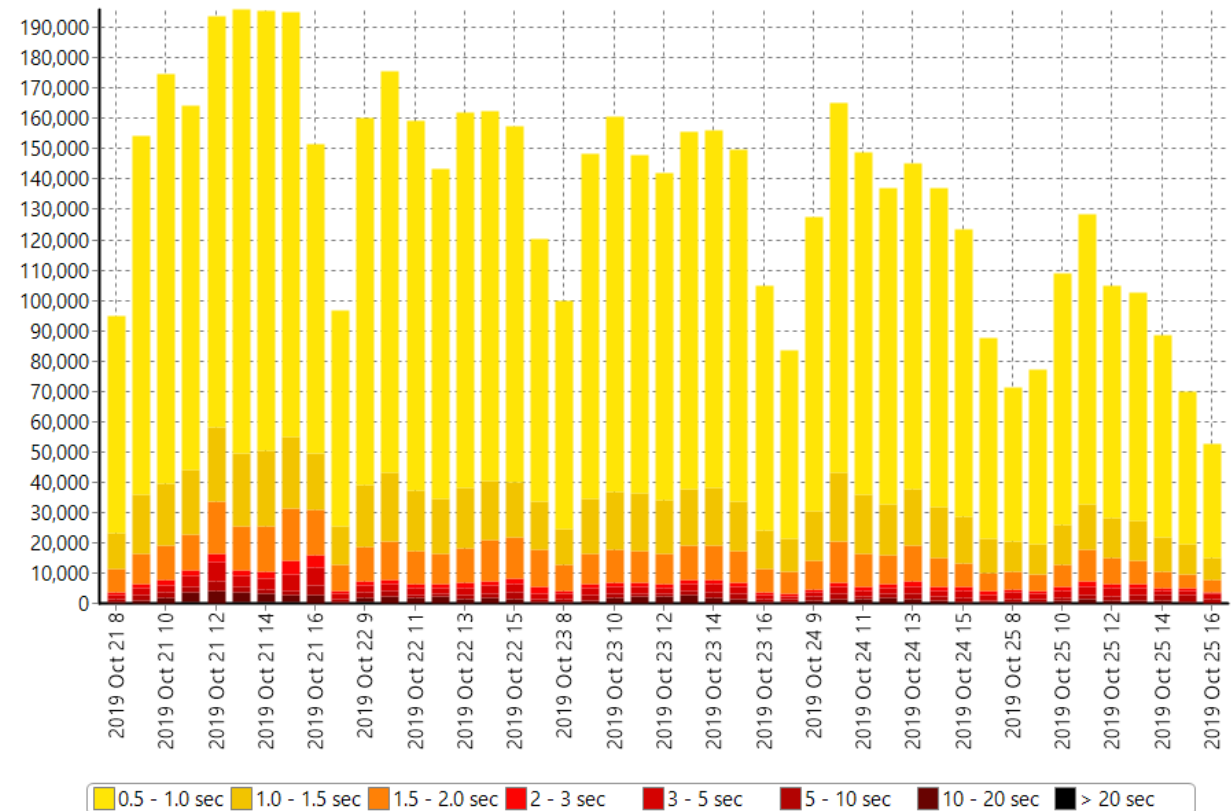
# Monitoring your quality of service - WLM

- SMF 72 gives insight into Workload Manager (WLM)

- WLM prioritizes resources based on 'goals' and 'importance'.

- Performance Index is an indication of whether the goals were met (1), missed (>1), or overachieved (<1). Ideally this number should be low and close to 1.

- Velocity is an indication of how often work is delayed due to resource constraints. 100% = no delays. Higher numbers are good, but don't expect 100%.

- SMF 72 can also tell us what resources are constrained
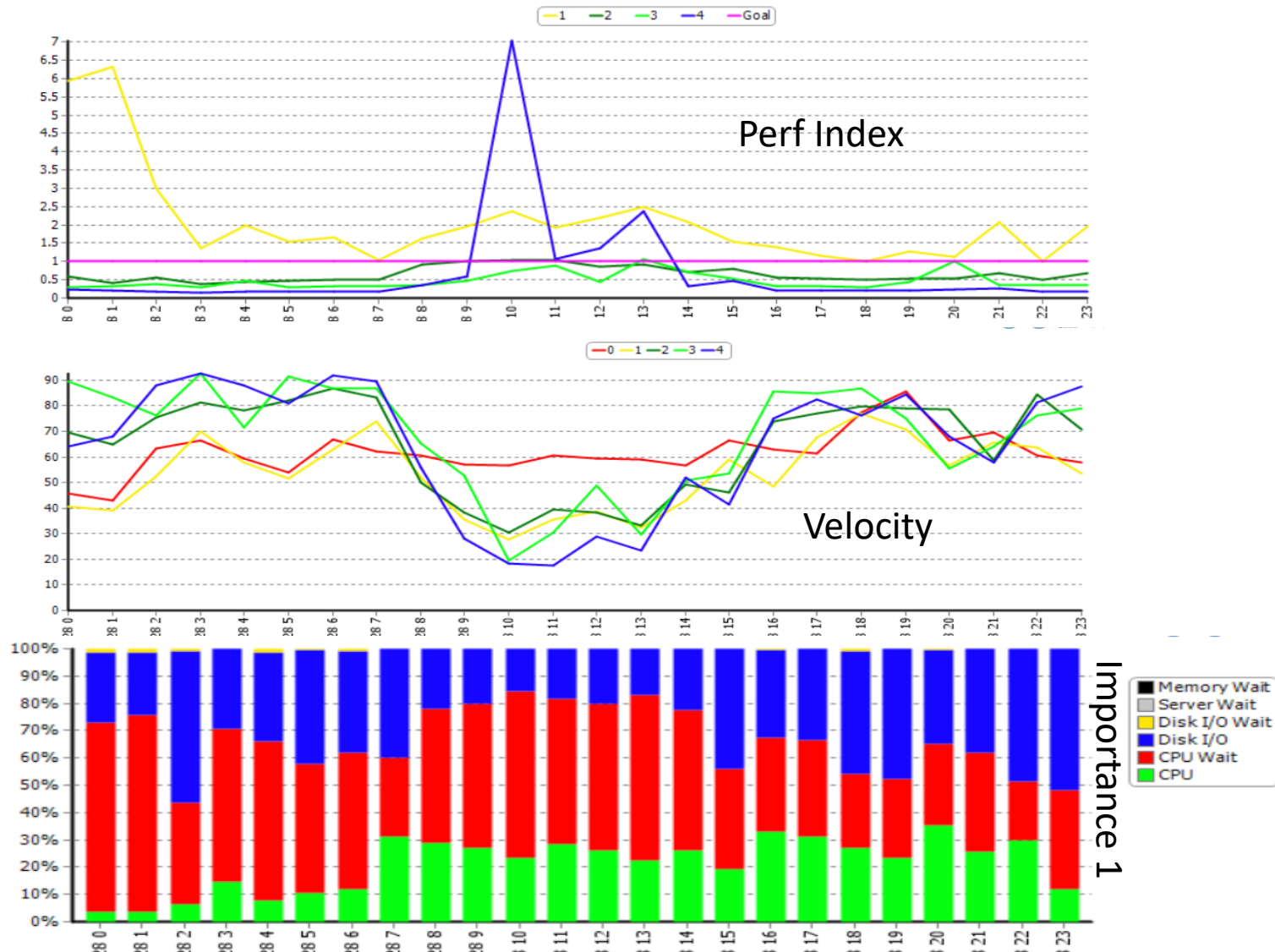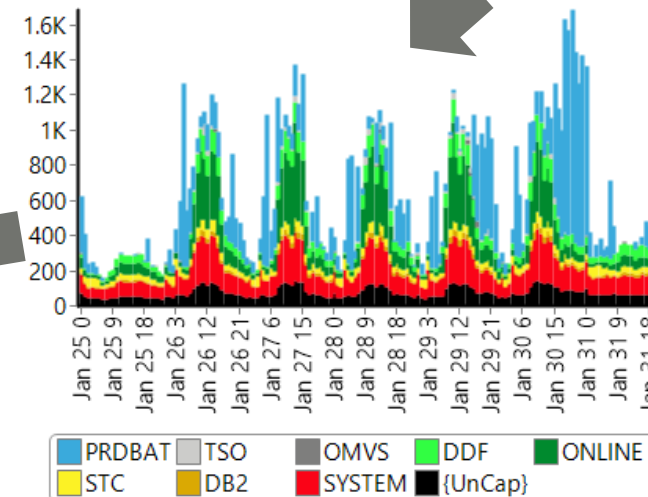
Identify Applications for Tuning

- Identify peaks

- Drill down to understand cost drivers e.g: LPAR, workload, job, transaction, DB2 package, user, etc.
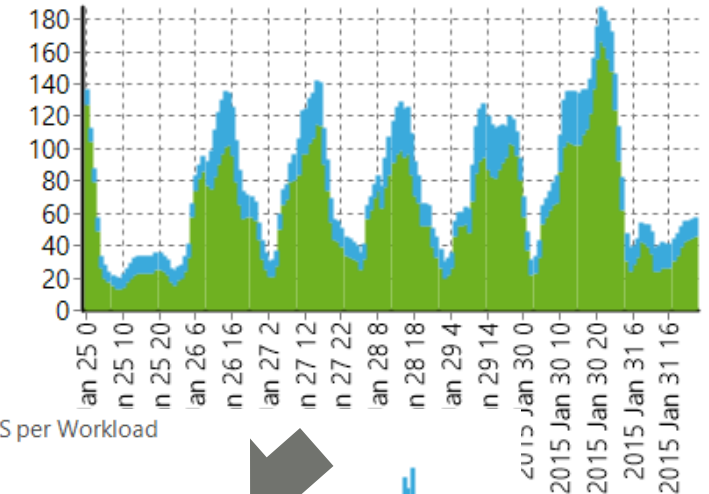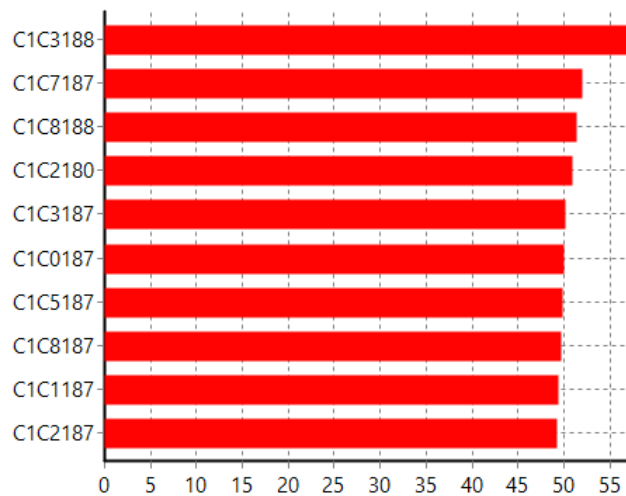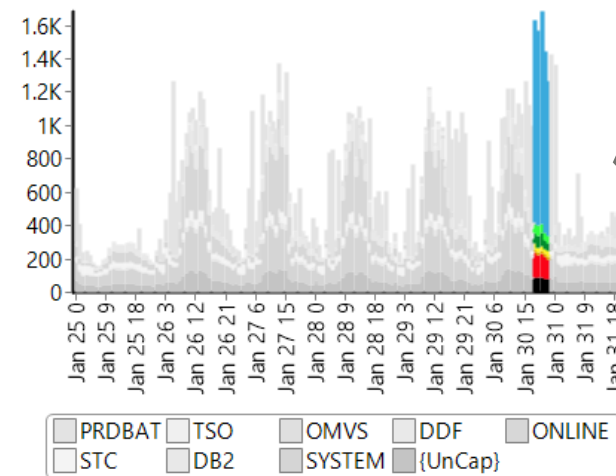
Summary

- An outsourced environment is usually a shared environment, so there will be compromises

- Have a clear and common understanding with your outsourcer of:
  - What you are paying for?
  - How it is measured or calculated?
  - How you monitor it?
  - And what you can do about it?

- Have access to good reporting tools – ideally 'self service'.

- Have access to an independent 'trusted advisor' if you don't have skills in-house.
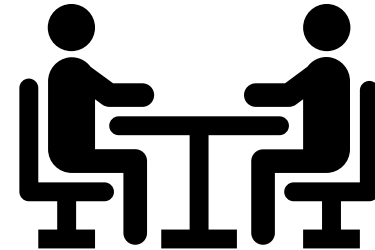
Please submit your session feedback!

- Do it online at http://conferences.gse.org.uk/2019/feedback/LL

- This session is LL

1. What is your conference registration number?

💡 **This is the three digit number on the bottom of your delegate badge**

2. Was the length of this presentation correct?

💡 1 to 4 = "Too Short" 5 = "OK" 6-9 = "Too Long"

1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○

3. Did this presentation meet your requirements?

💡 1 to 4 = "No" 5 = "OK" 6-9 = "Yes"

1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○

4. Was the session content what you expected?

💡 1 to 4 = "No" 5 = "OK" 6-9 = "Yes"

1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○