Advanced Technical Skills (ATS) North America

Getting Started in (z/OS) Capacity Planning (Topics in Capacity Planning) Part I

SHARE
Technology · Connections · Results

March 2011

Ray Wicks
561-236-5846
RayWicks@us.ibm.com
CPS Tools Group, Gaithersburg. MD
CPSTools@us.ibm.com

Advanced Technical Skills
TECHNICAL SALES EARTH AMERICAS

Advanced Technical Skills (ATS) North America

Bibliography

Ray has spent most of his career at IBM in the performance analysis and capacity planning end of the business in Poughkeepsie, London, and now at the Washington Systems Center. He is the major contributor to IBM's internal PA & CP tool zCP3000. This tool is used extensively by the IBM services and technical support staff world wide to analyze existing zSeries configurations (Processor, storage, and I/O) and make projections for capacity expectations.

Ray has given classes and lectures worldwide. He was a visiting scholar at the University of Maryland where he taught part time at the Honors College.

He won the prestigious Computer Measurement Group's A.A. Michelson award in 2000. His recent virtual sessions "Getting Started in Performance Analysis & Capacity Planning" workshop held for attendees in China and India was well accepted.

© 2010 IBM Corporation

Trade Marks, Copyrights & Stuff

Many terms are trademarks of different companies and are owned by them.

- **On foils that appear in this presentation are not in the handout. This is to prevent you from looking ahead and spoiling my jokes and surprises.**

Abstract

This tutorial is a **two** part introductory level session designed to introduce the student to the concepts required for Performance Analysis and Capacity Planning.

Emphasis is placed on large processor systems and examples will be largely drawn from z/OS but the concepts apply to all operating systems and hardware. The tutorial is organized to review the architecture where appropriate (albeit briefly). Topics:

- Conceptual and Perceptual structures for performance analysis and capacity planning,
- Using the Forced Flow law in PA & CP
- Performance Analysis queries for capacity planning,
- Processor performance data (ITRRs & MIPS),
- Resource Metrics for use in the Balance System model,
- Sample selection,
- Data preparation in z/OS,
- Using the utilization growth process in capacity planning,

Capacity Planning

Capacity Planning ensures that Adequate resources are available for the Workload to complete in an Appropriate time.

- **Performance Analysis is Short Term (3-7 Days)**
- **Capacity Planning is Long Term (6-24 Months)**
- **What's the Workload using now?**
- **What's Appropriate? Adequate?**
- **Service Level Objective or Service Level Agreement**
- **Things are ordered by Priority or Importance**
- **Discretionary Workloads?**

CP Questions

EASY

- **Do I have enough resource (CPU, I/O, Storage,..) to do the job today?**
- **If not, who's suffering?**
- **If I get more, who will be helped? How much?**
- **If I need more, when will it be?**
- **How much more?**
- **Can I use specialized Processing Units?**
- **What variables should I track?**
- **Do I have any latent demand?**

Harder

- **Do I want faster or more CPs?**
- **How do I establish my growth?**
- **How do I size a new application?**
- **What tools should I use?**
- **Which interval do I model?**
- **If I reduce the #CPs & keep the MIPS the same will there be a problem?**

CP Flow

Business Model
Forecast
Cost

Data Model
λ,Traffic, I/O, #clients,
etc.

Knowledge Preparation
- Understand Environment
- Workload Characterization
- Business Units
- Application Programs
- H/W Resources

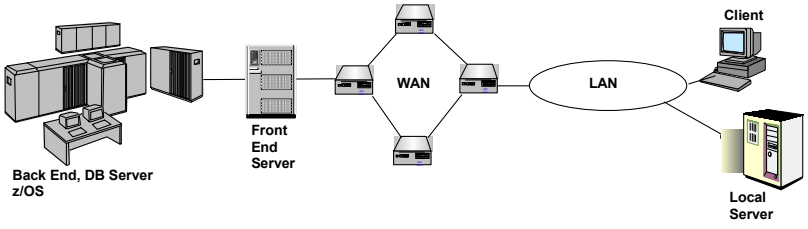
Model Preparation
- Data Acquisition
- Representative
- Calibration

Model Execution
- Workload Forecasting

Modeling Output
- Future Requirements
- H/W & S/W
- Speeds & Feeds
- Architecture

Business Negotiation
- Investment & Configuration

Topology



For each node: # Servers? Utilization? Service time? Queuing? Response Time? Traffic? Etc.?

Lots of Data

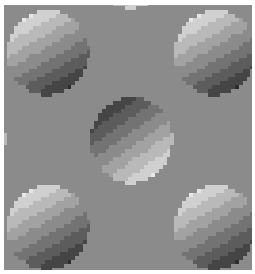
Frameworks

Conceptual Framework (Model, Paradigm)
- **What's it supposed to mean?**
- **How do things connect?**

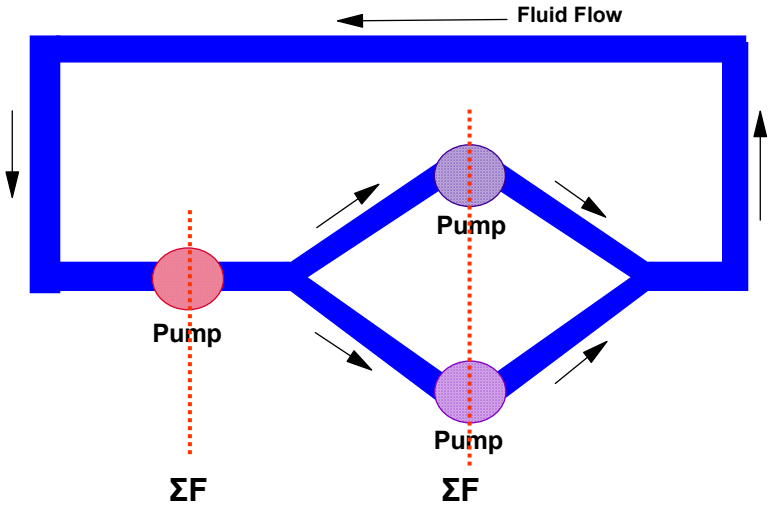
Perceptual Framework
- **What's it supposed to look like?**
- **How do things connect?**

The Danger of a Framework

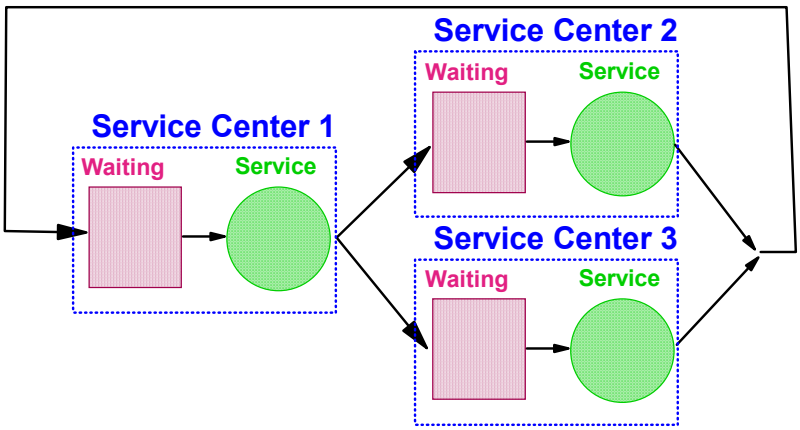
Turn 180°



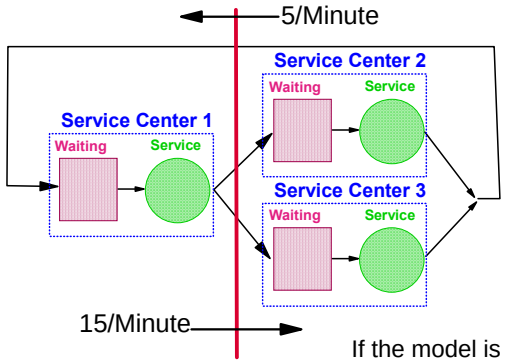
A Plumbing Problem



Conceptual Framework Forced Flow Law Flow Law

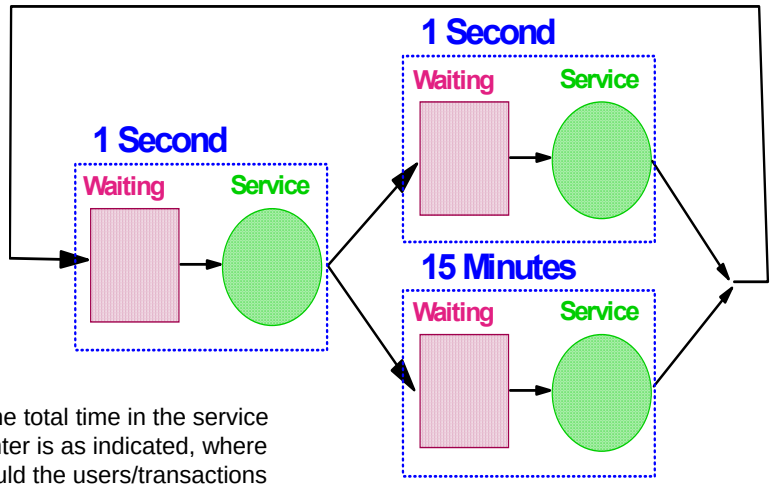


Conceptual Framework



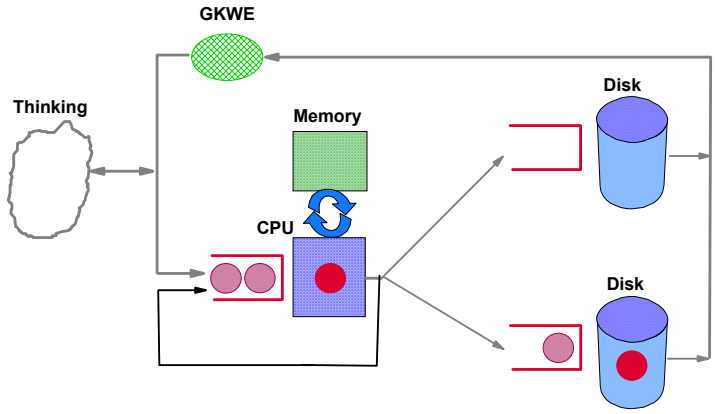
If the model is divided in two and the number of transactions crossing the boundaries is as indicated, what would the forced flow law say?

Conceptual Framework

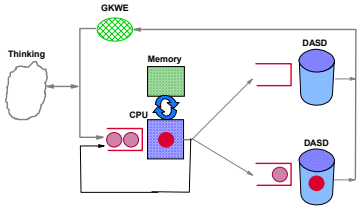


If the total time in the service center is as indicated, where would the users/transactions be?

B.S. Conceptual Framework

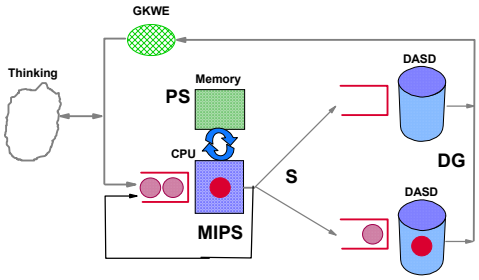


Conceptual Framework Implications



- **Users distribute themselves among nodes in proportion to the time spent at each node.**
- **The capacity of the System is determined by the slowest node (server).**
- **The resource usage (transaction rates) at various nodes are in proportion.**

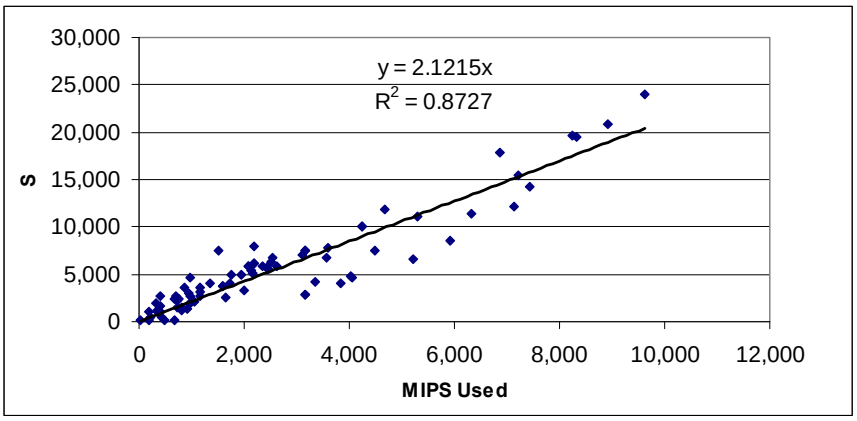
2010 z/OS B.S. Metrics



MIPS used
S = SSCH rate
DG = DASD gigabytes. The computation is nominal in that it is 2.83/act
PS = Central Storage configured

	10%	50%	90%
S/MIPS	1.201	2.349	3.707
DG/MIPS	2.236	6.593	52.539
PS/MIPS	6.766	14.055	37.306

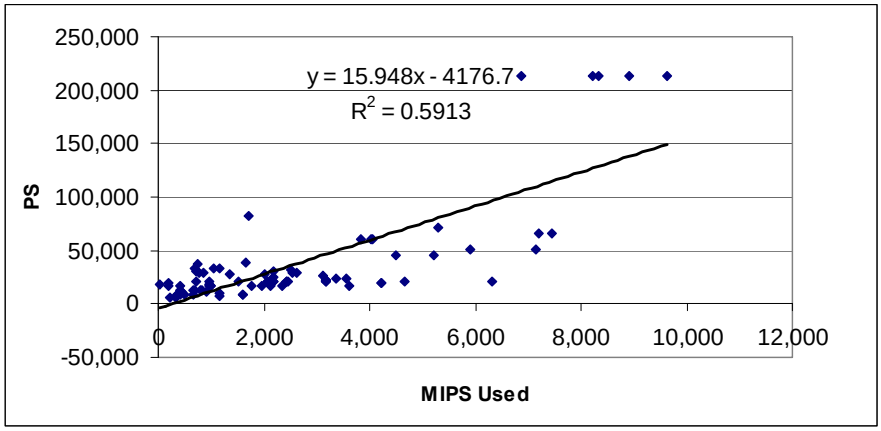
S/MIPS Partition Samples



MIPS Used and SSCH are in proportion (correlated).

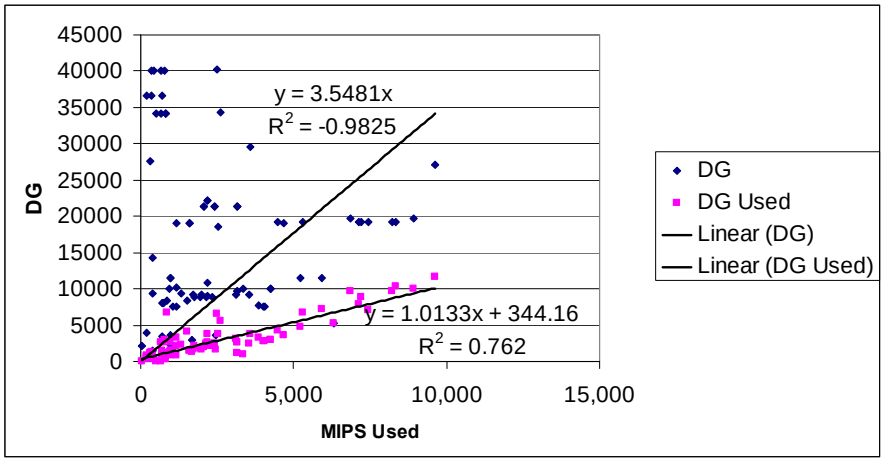
Good linear relationship.

PS/MIPS

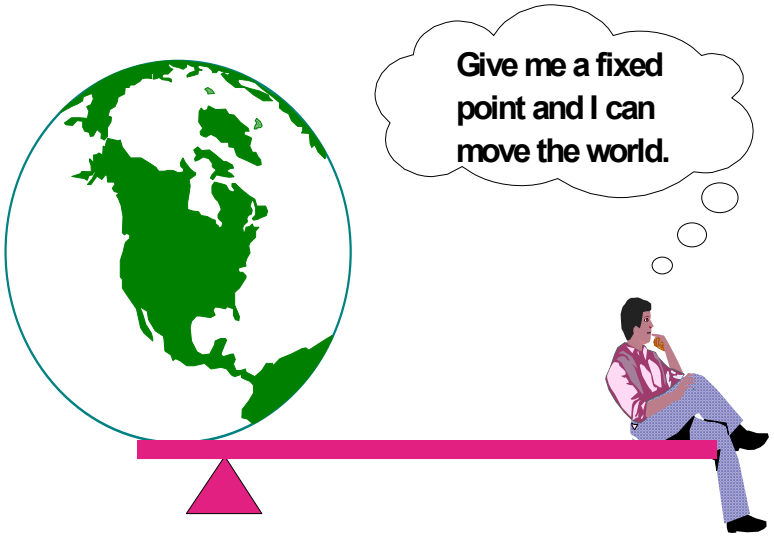


MIPS Used and PS are not in proportion (correlated).

DG/MIPS



Archimedes



Desperate Capacity Planning

What part of a 500 MIPS machine would DB2 use at 100 I/Os per second?
How many MIPS for DB2?

S/MIPS ♏ [1.201, 2.349, 3.707]
100/MIPS = 1.201 → MIPS = 83.26
100/MIPS = 2.349 → MIPS = 42.57
Answer:
Between 8.5% and 16.6% (????) of a 500 MIPS Machine

More Metrics Defined

MIPS used
S = SSCH rate

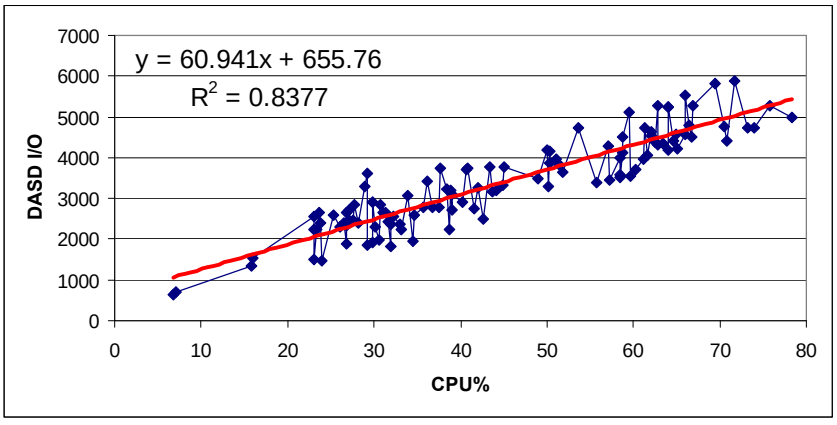
DG = DASD gigabytes. The computation is nominal in that it is 2.83/act
DG = Nacts * 2.83
CS = Central Storage configured
D as in CS = 4000 + 0.04(MIPS)^D
where CS (or PS) is configured Central Storage (or RAM)
NNTacts = acts with rate >=2
AD = Access Density = S/DG

Full Metrics

2010 83 Partitions

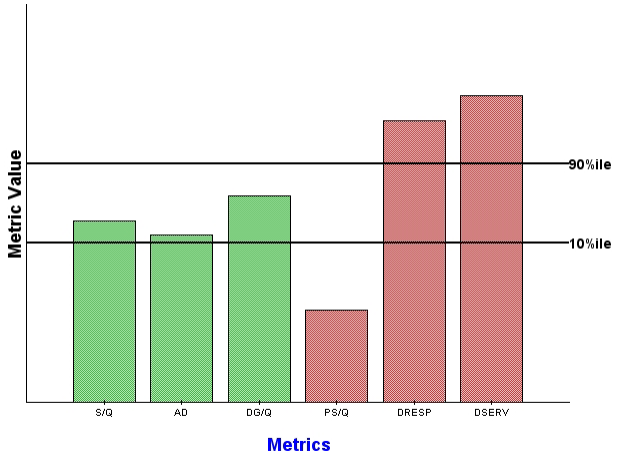
	10%	50%	90%
MIPS	403	2004	6247
S	1123	4221	11779
S/MIPS	1.201	2.349	3.707
DG/MIPS	2.236	6.593	52.539
PS/MIPS	6.766	14.055	37.306
D	1.534	1.630	1.894
DASD Resp	0.952	1.865	3.626
DASD Serv	0.681	1.564	2.980
Resp/Serv	1.227	1.827	1.232
Nacts	1305	4084	12105
NNTacts	73	264	882
DASD GB	3693	11558	34256
Used DG	585	2114	7065
AD	0.059	0.360	0.909

If the Model Works, What Should I See?



If the model is valid, the forced flow law would prescribe that the variation of CPU service would be proportional to the I/O service. In this graph that translates into a linear relationship (R^2>0.7?).

Overall Evaluation 1



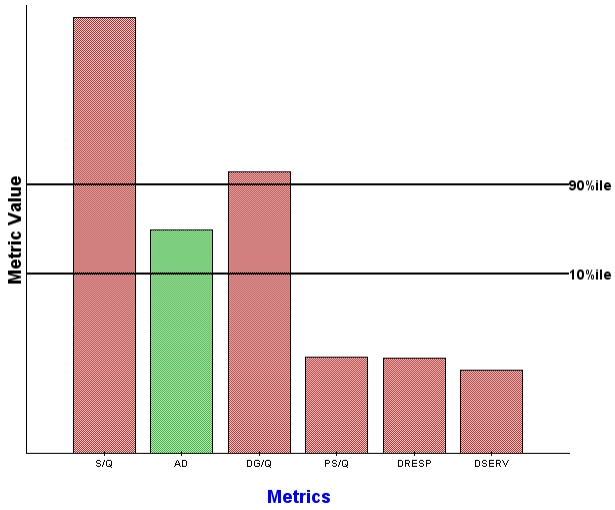
The total power used in the system image, in MIPS (Q), was 1959.9. The DASD I/O rate per second, S, was 5087.4. The S/Q (MIPS Used) currently is 2.6. The 10th percentile and 90th percentile of this metric are 1.3 and 5.8.

Access Density (AD) is the ratio of the DASD I/O rate (S) and the total DASD gigabytes (DG) in the enterprise. This value can often be invalid if the DASD models are not actually there, the device is emulated. DG is developed from a table which states a nominal DG by device type. For this model, DG=25371.8 thus making AD=0.2 It is an indicator of whether you have enough gigabytes on average as compared to other systems "out there". The expected range for AD in 10th percentile/90th percentile is 0.1/1.2.

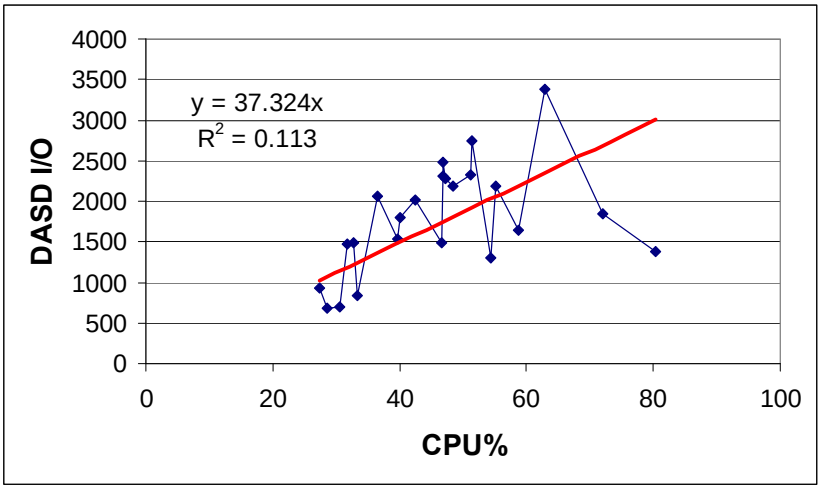
The ratio of DASD gigabytes(DG) to MIPS used(Q) is a companion metric to Access Density. This metric is one test of the theoretical expectation that larger systems need more DASD bytes. The ratio DG per MIPS used here is 12.9. The expected range in 10th percentile/90th percentile is 2.3/20.4.

The ratio of processor storage used (PS) to MIPS used has an expected range in 10th percentile/90th percentile of 8.4/58.4. The value of PS/Q=4.9. This value is below the 10th percentile.

Overall Evaluation 2

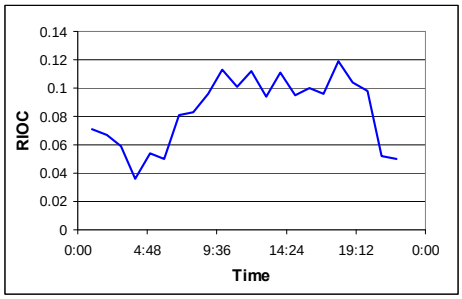


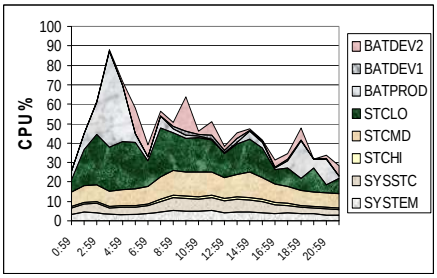
When It Doesn't Work



If the model Doesn't work, you would see a non-linear relationship.

RIOC=F(Workload Composition)



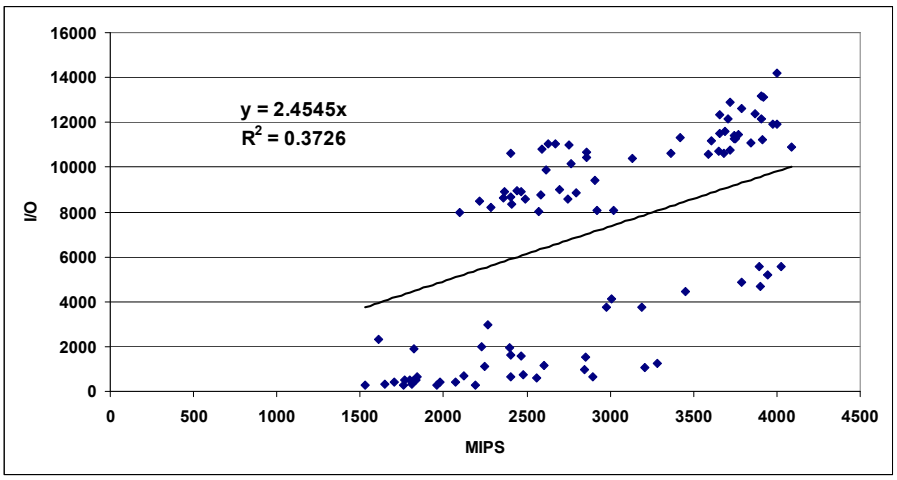


Compare those intervals in the RIOC plot where the RIOC is stable or unstable with the workload mixture for the same intervals in the CPU% plot. The workload mixture often explains the RIOC variation.

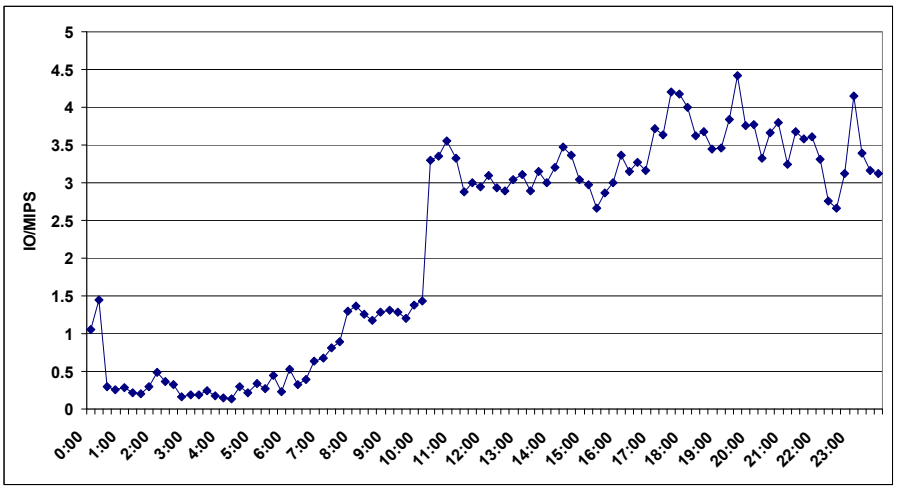
Use of Conceptual Framework like BS Model

- **Establishes Relationships**
  - Balanced System, resource ratios
- **Builds Expectations**
  - Linear graph
- **High lights Exceptions**
  - Non Linearity
  - Outliers
- **Generates Questions**
  - Especially if not as expected
- **But... it may cause you to see Framework (model) interactions that just aren't there!**

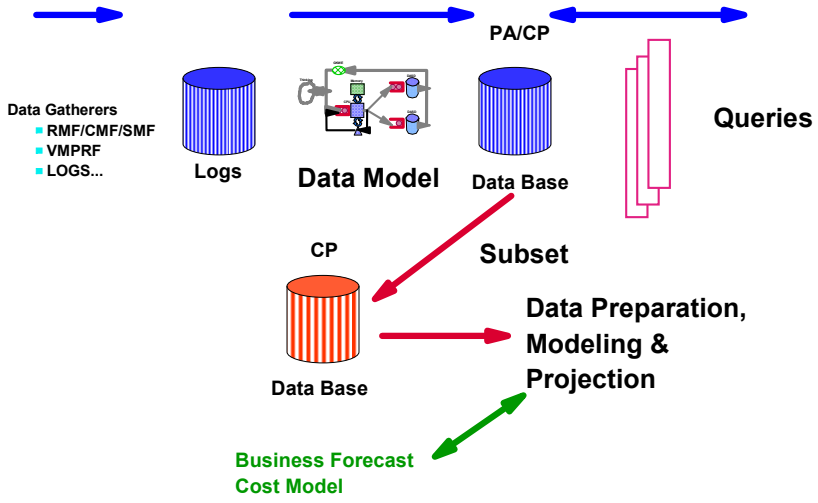
Another Example

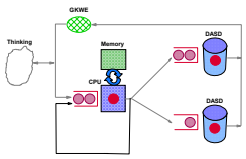


Same Data



Data Acquisition & Analysis



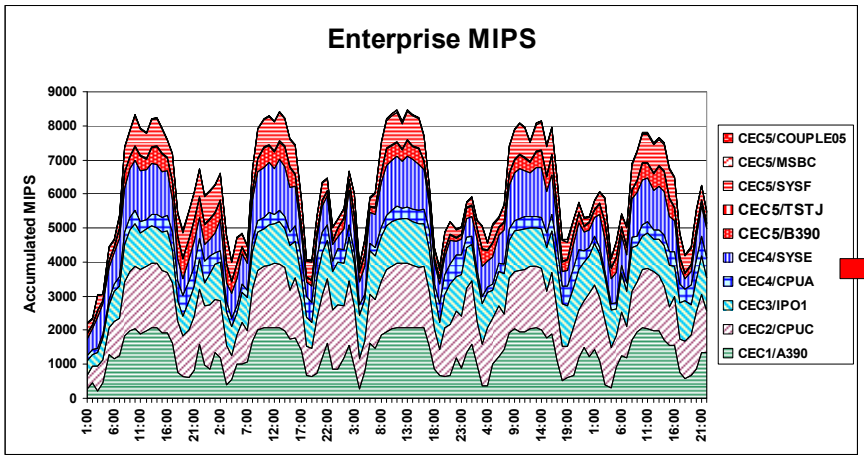


Performance Analysis Queries of Interest to CP

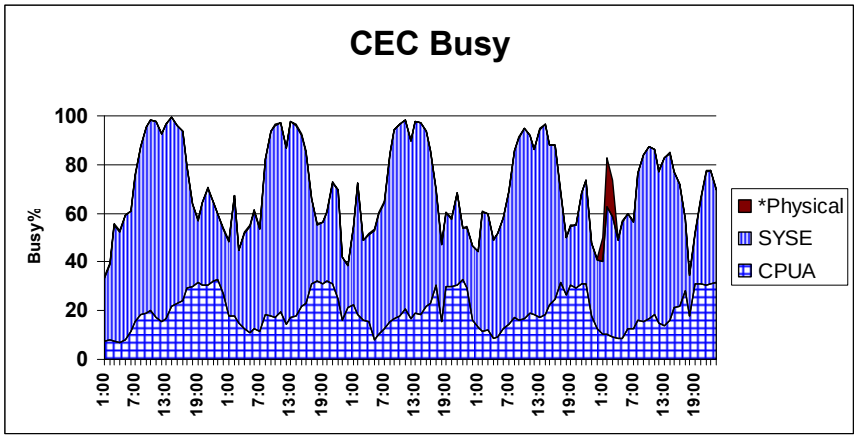
- Enterprise Level MIPS Usage
- CPC Level Busy
- System Image (Partition) Level
  - Workload Busy CPU and I/O Intensity
  - Contention Indicators
    - Average in and Ready (AINR)
    - OCPU1 = %AINR >= 1+#CPs
    - OCPU2 = %AINR >= 2+#CPs
    - Average Out and Ready (AOUTR)
    - Delays by Workload
    - Performance Index >1

Enterprise Level

View at least 24x5



CPC Level View



CPU%

# Engines	Maximum ST in Seconds	Seconds if CPU%=42%
1	1	0.42
2	2	0.84
3	3	1.26
4	4	1.68
5	5	2.10

0.84 Secs or 840 Ms.

Weights

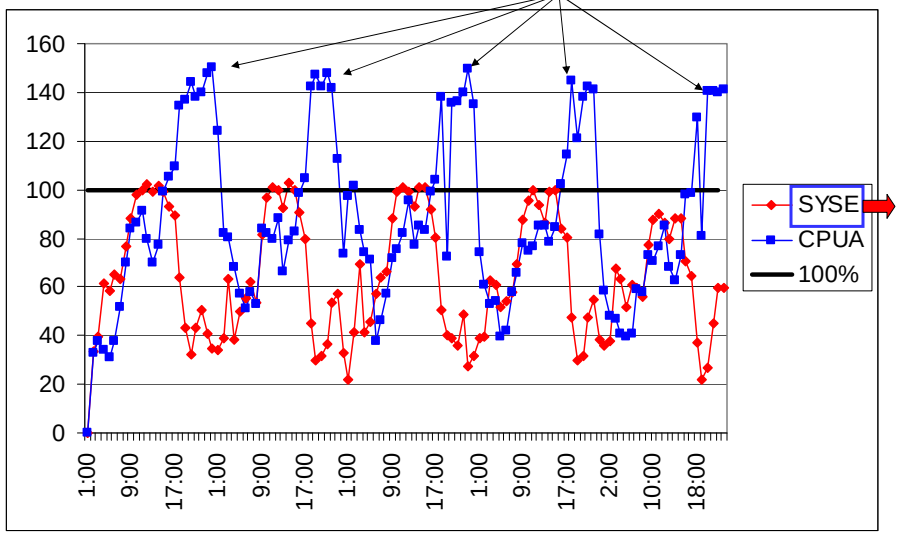
Weight as Percent: When utilization of CPC gets to 100%, this partition should get this percent of the Shared CPU pool.

10 CP CPC

Part	#LCPS	SHARED	Weight	Weight %	Logical CPU%	Physical CPU%	Used (#CPs)	Entitlement (#CPs)
LPAR1	2	Y	300	25%	50%	10%	1	2.5
LPAR2	3	Y	400	33%	67%	20%	2	3.3
LPAR3	6	Y	500	42%	100%	60%	6	4.2

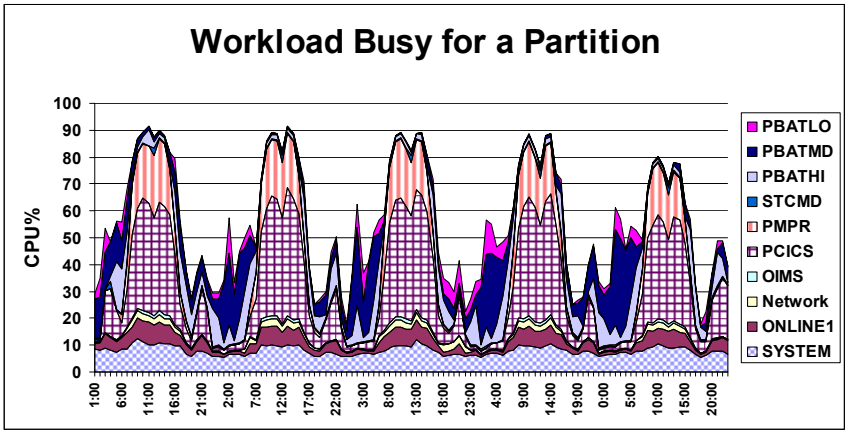
% Of Share

Warning: When if SYSE demanded more, CPUA may suffer. But notice the time.



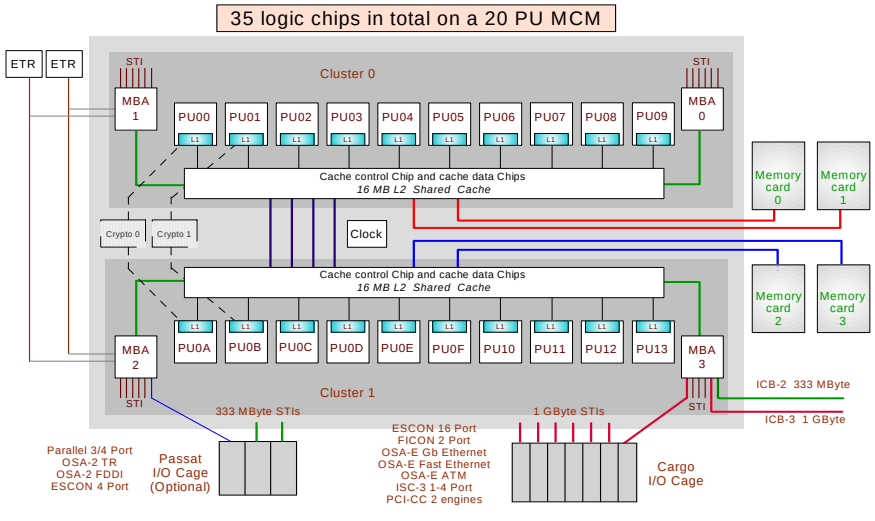
100% means partition received the amount of processor resource entitled by weight.

Partition Workload View

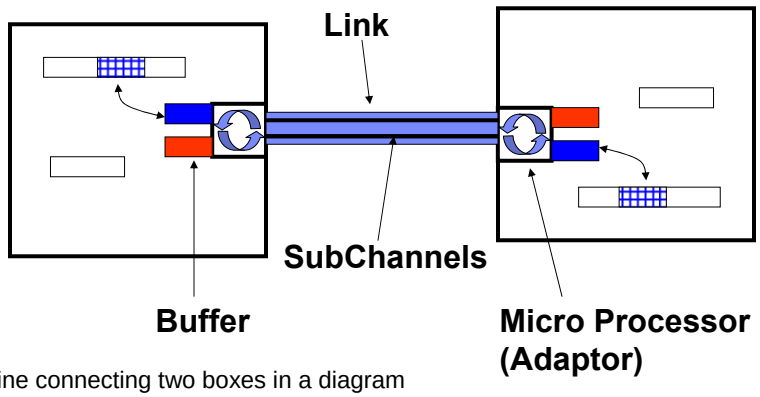


Who's using the CPU? Which workload dominates? What time period? Workloads stacked by importance.

Z900 Structure

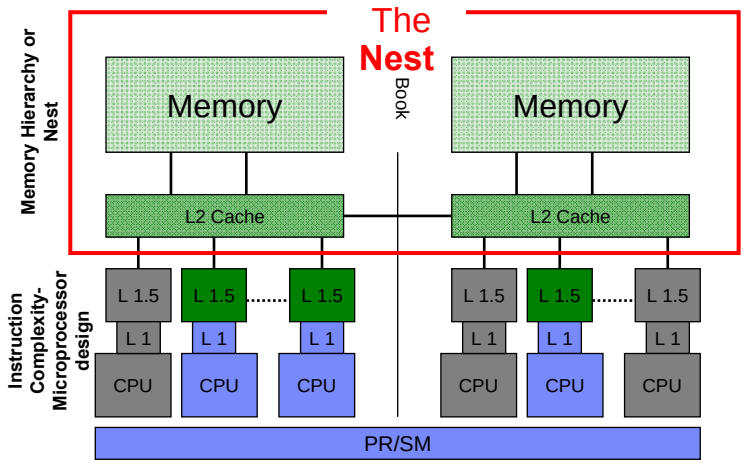


Connections



Every line connecting two boxes in a diagram implies micro processors on each end to do the talking? (What happens if they speak different languages?) Data is moved from a buffer to micro processor buffer onto link into m-processor buffer into storage buffer.

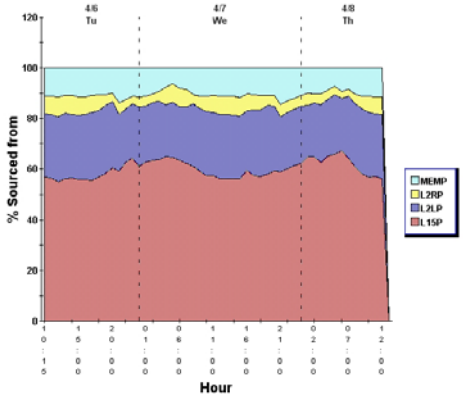
Z10 Memory – a simple view



Reference: John Burg's presentation at SHARE 8/4/2010

http://www/ibm.com/support/techdocs/atsmastr.nsf/Webindex/TC000066

If not from L1, from Where? (SMF 113 Data)



Here's the plot of percent sourcing from different levels of cache. As the sourcing moves from the highest level of cache (percent=L15P) to the slowest memory source (percent=MEMP), the performance degrades. Level 1 cache is the fastest and closest to the processing unit.

The sourcing shown in the graph is for data not found in level 1 cache. You can check the level 1 cache miss % by graphing variable L1MP.

- L1.5P=%sourced from level 1.5 cache
- L2LP=%sourced from level 2 cache same book
- L2RP=%sourced from level 2 cache different book
- MEMP=%sourced from memory

Remember than as more and more of the instructions and data has to be fetched from more distance caches, the machine effectively runs slower.