So You Don’t Think You’re An Architect?

Martin Packer, IBM
martin_packer@uk.ibm.com

Abstract

If you do Performance you might well be right. But if you think that way you're missing an important trick or two.

This presentation will show you the many ways in which you can gain architectural insight from SMF data.

You might be the only person in your shop who can actually gain that kind of insight in a systematic manner so you could be very valuable if you learn how to do this.

Examples include at the system and middleware levels.

All the techniques use data you probably have already.

Wikipedia Definition Of Systems Architect

- The systems architect is an information and communications technology professional. Systems architects define the architecture of a computerized system (i.e., a system composed of software and hardware) in order to fulfill certain requirements. Such definitions include: a breakdown of the system into components, the component interactions and interfaces (including with the environment, especially the user), and the technologies and resources to be used in its design and implementation.

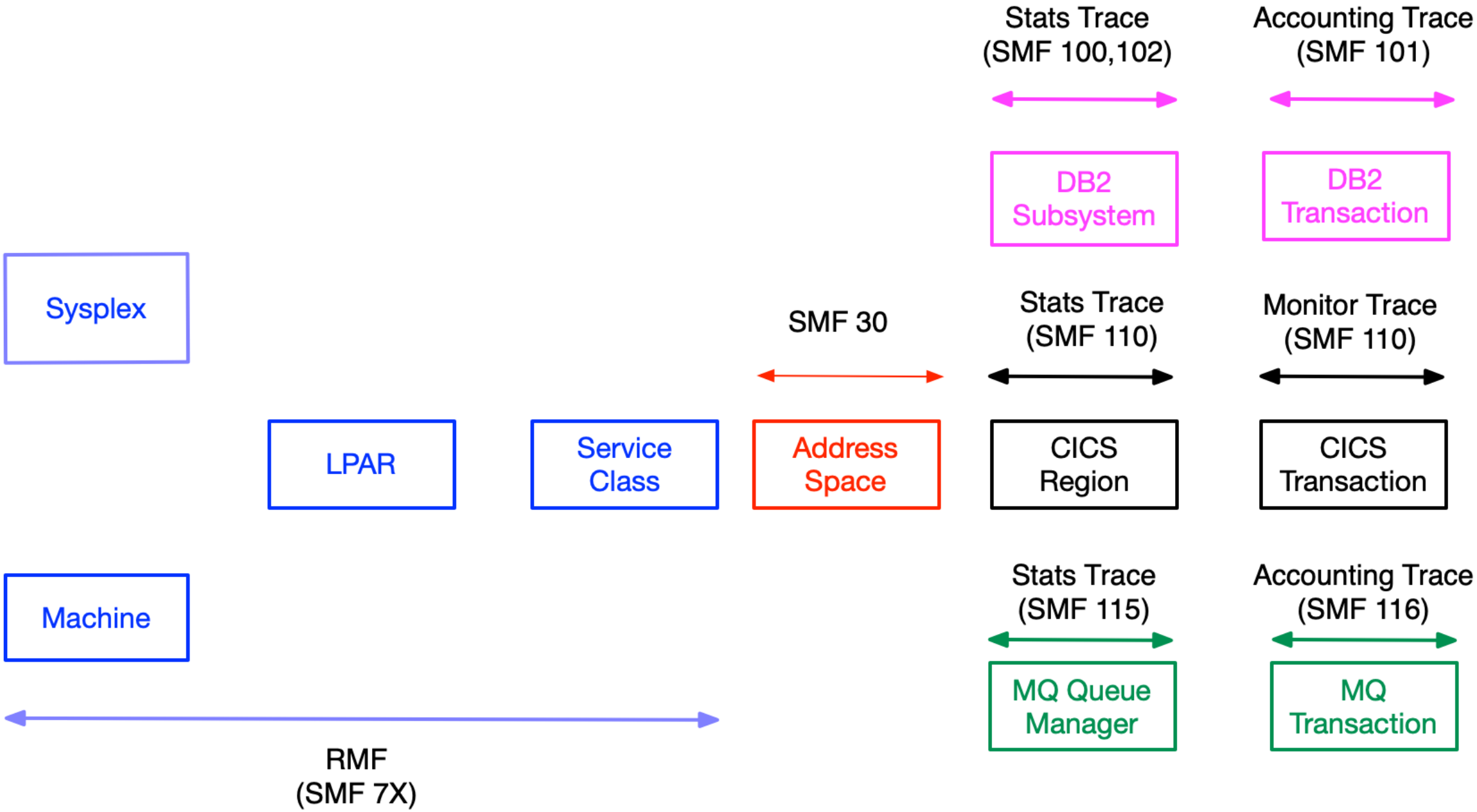
  https://en.wikipedia.org/wiki/Systems_architect

- My point of view: It's not essential to match this definition but it is valuable to bring architectural insight

What We'll Talk About

- The Top Layer - Machines, Sysplexes And LPARs
- The Middle Layer - Workloads And Middleware
- The Bottom Layer - Application Components
- How Does This Fulfil The Definition Of Architecture?

SMF Records We're Going To Meet

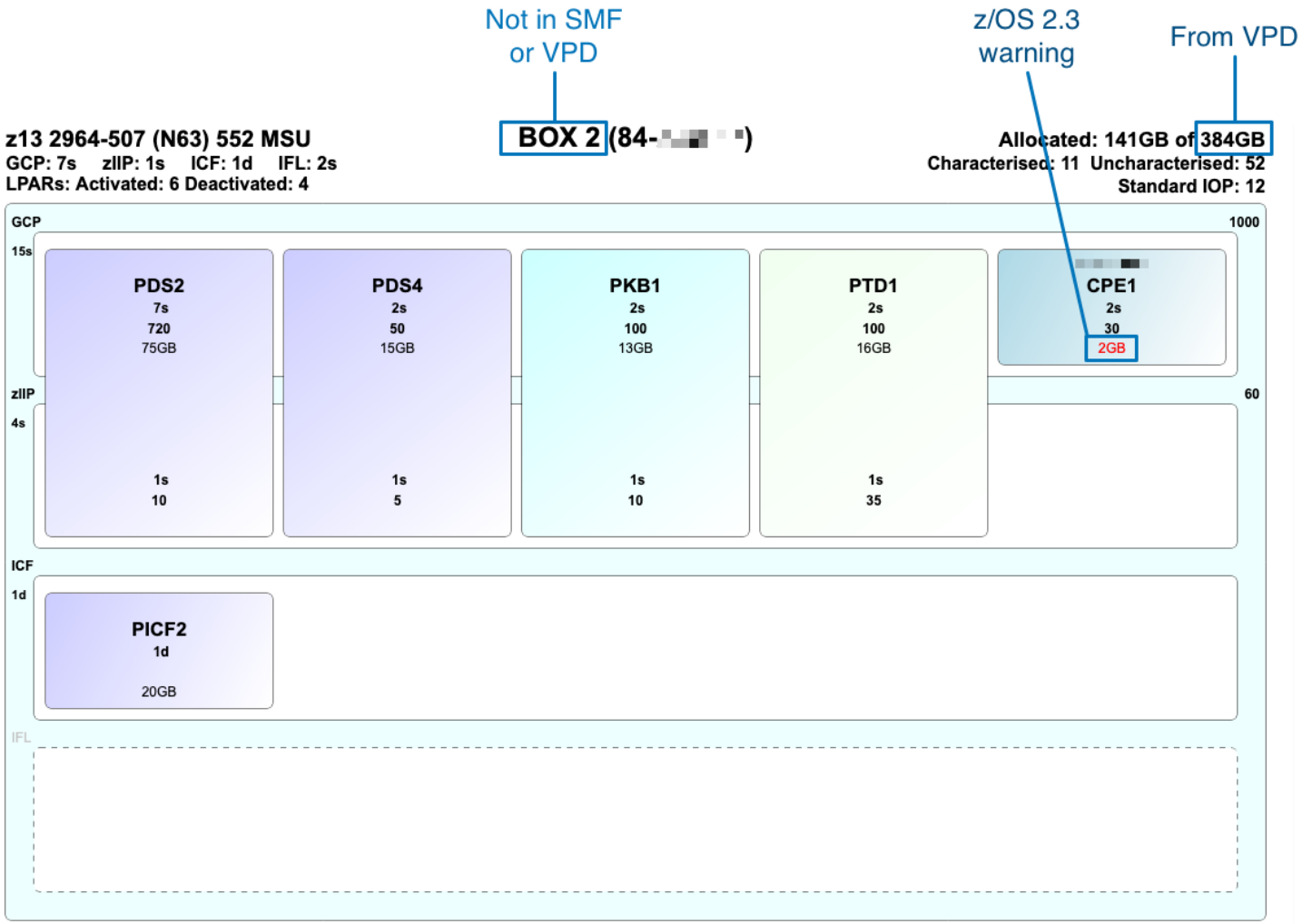


The Top Layer - Machines, Sysplexes And LPARs

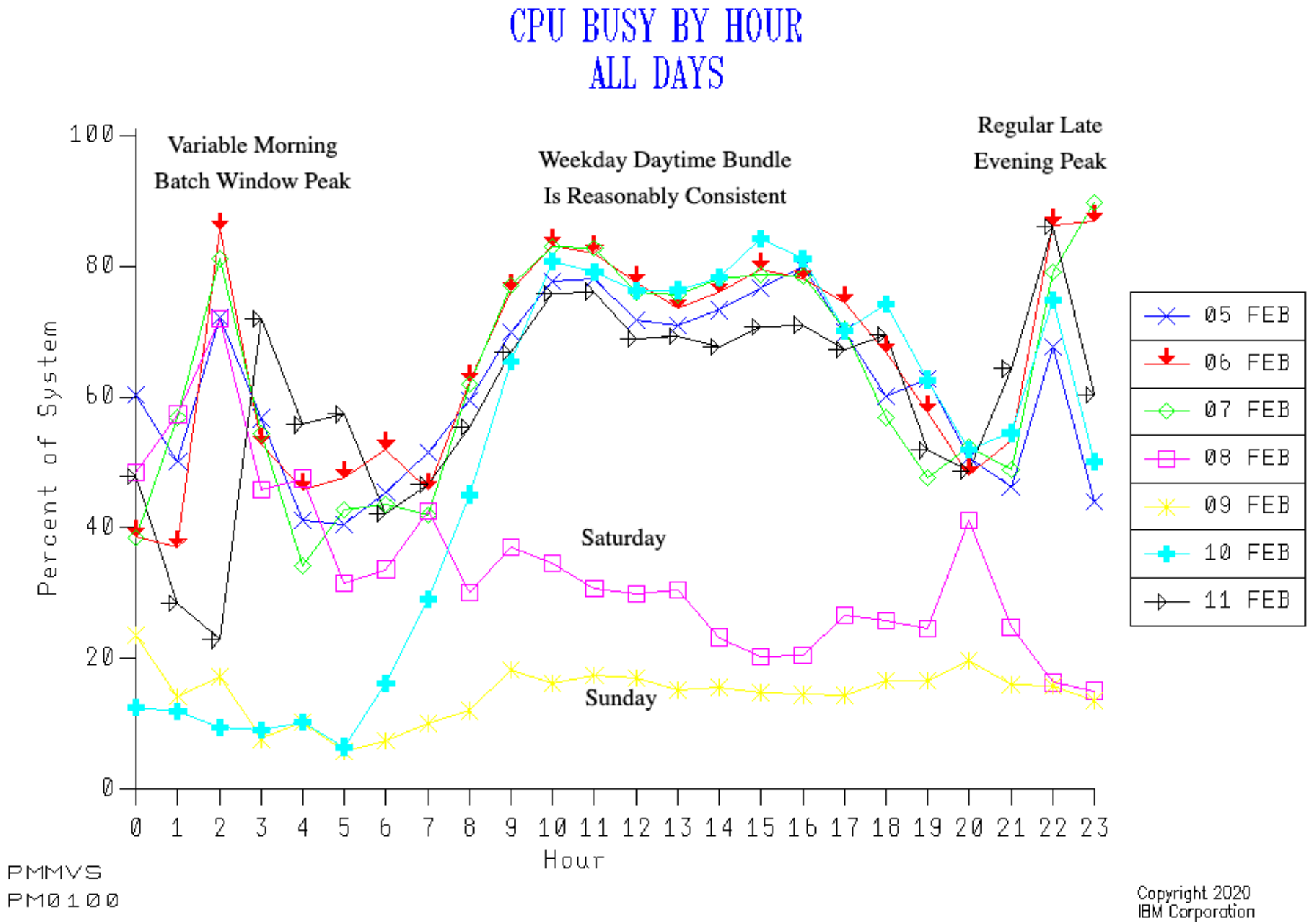
Machines - From SMF 70-1

- What machines exist?
- What models are they?
- What is their configuration?
- How do they behave?
- If I want to upgrade one how far can I go without
  - Adding hardware?
  - Taking an outage?

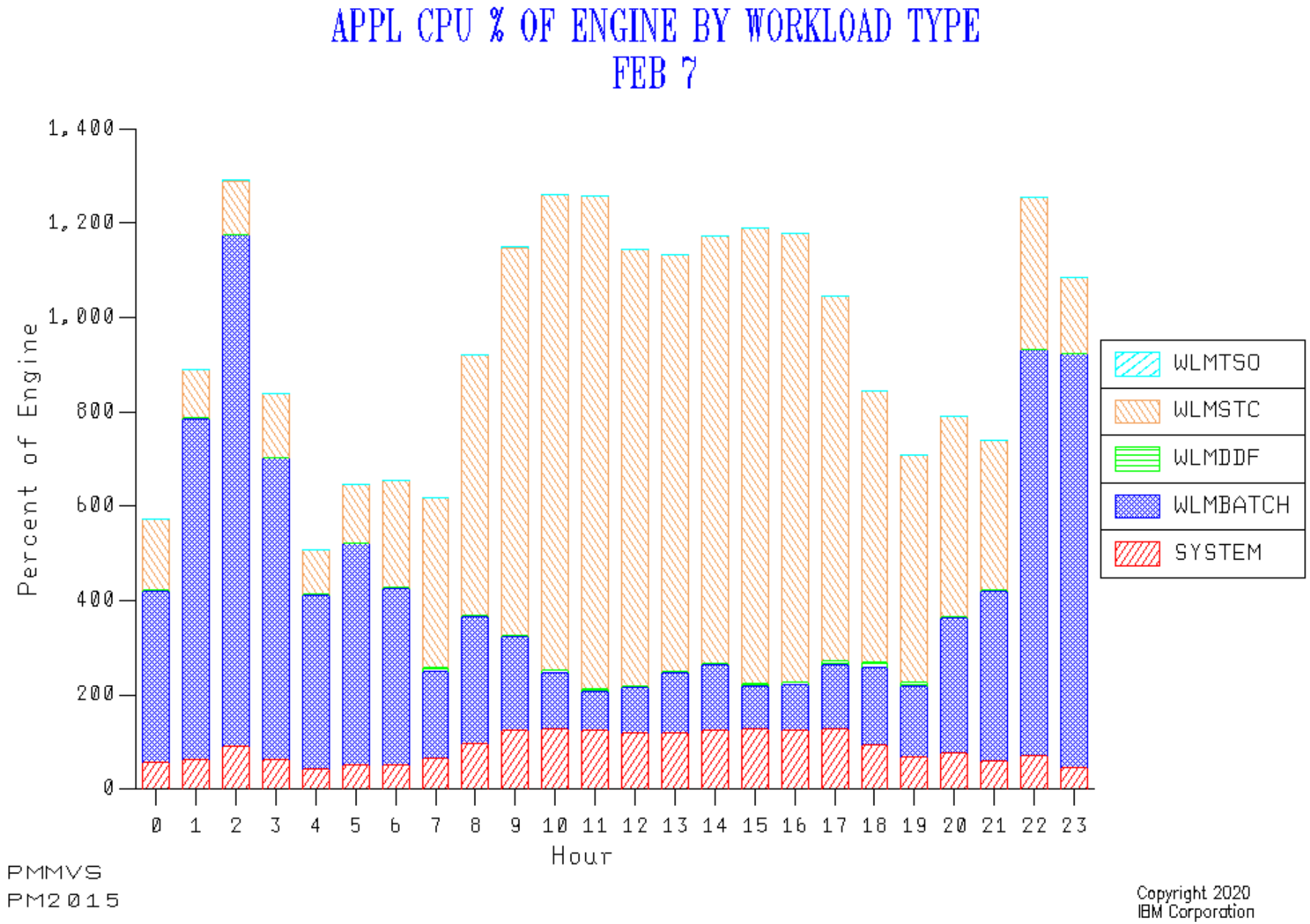
Here's An Example Machine - Showing Its LPARs In All Processor Pools



This Might Be A "Performance" Chart But I Consider The Pattern An Architectural Artefact



This Gives You Some Idea What Drives That Pattern - And What This LPAR Is **For**

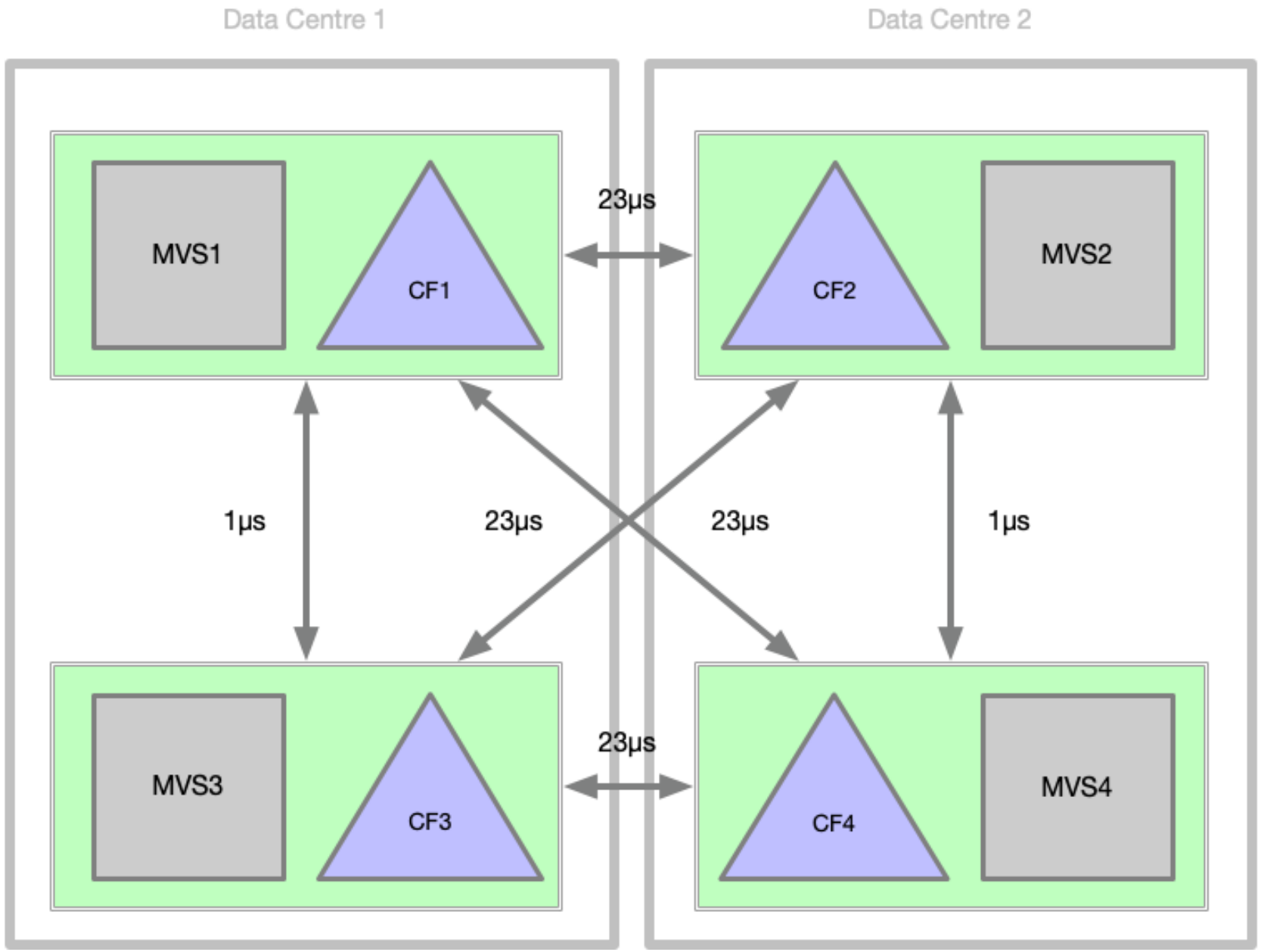


Sysplexes - From SMF 70-1 And SMF 74-4

- What Sysplexes?
- Which Systems?
- Which Coupling Facilities?
- What Structures?
- What's the connectivity
  - Over what distance?

- SMF 70-1 Partition Data Report lists systems in each sysplex
  - Doesn't show which coupling facilities are in which sysplex
- SMF 74-4 lists coupling facilities in each sysplex
  - Can relate to 70-1 LPAR information
- SMF 74-4 lists link types in detail and their distances
  - Often observed to be multiple groups of same-distance links
  - SMF 73 reveals shared CHPIDs

Put It All Together And You Get Something Like This



Here's Another, More Compact, Way Of Depicting Sysplexes

Parallel Sysplexes

	BOX 1 84-[redacted]	BOX 2 84-[redacted]
SPX[redacted]P	PICF1 PDS1 PDS3	PICF2 PDS2 PDS4
SPX[redacted]2	CPE2	
SPX[redacted]P		PKB1
SPX[redacted]P		PTD1
SPX[redacted]1		CPE1

Inactive LPARs (from SMF 70-1) can be instructive - especially if they have similar names to other machines' active LPARs

LPAR Number	LPAR Name	Notes
4	N1PRODXX	Counterpart: S1PRODXX / MVSX
5	N1PRODYY	Counterpart: S1PRODYY / MVSY
8	N1PRODZZ	Counterpart: S1PRODZZ / MVSZ
9	N1SUPTAA	Counterpart: S1SUPTAA / MVSA
11	CFCAP1	Used to cap ICF pool when activated?

IPL Timing Is An Interesting Topic - In SMF 30 MSTJCLxx Reader Start Time is a good signifier (Also SMF 70 with APAR OA54916)

Date	Time	System
Machine A		
Sunday, 2 February	00:25	MVSA
=	00:31	MVSB
Machine B		
Sunday, 2 February	02:01	MVSC
=	02:10	MVSD
Machine C		
Sunday, 2 February	01:18	MVSE
=	01:23	MVSF

The Middle Layer - Workloads And Middleware

Workloads / Service Classes / Report Classes

- Their rhythm is interesting
  - Particularly CPU and zIIP
  - Also transaction rates
  - All from SMF 72-3
- Ensure the descriptions are useful and accurate
  - They appear in SMF 72-3
- SMF 30 allows you to tie together Service and Report Class
  - Is one a subset of the other?
- SMF 30 says what runs in each Service / Report Class
  - Except Independent Enclaves
- SMF 30 gives you e.g. DFSMShsm and CATALOG spikes

When I Said "DFSMShsm and Catalog" an Interesting One Is DUMPSRV
For DUMPSRV generally it's EXCP rate that spikes

Date	Time	EXCP Relative To Average	EXCP Rate
Monday, 8 October	6:00		31
Tuesday, 9 October	5:00		33
	21:00		208
Sunday, 14 October	4:00		30

SMF 30 Usage Data Section Is **Very** Useful

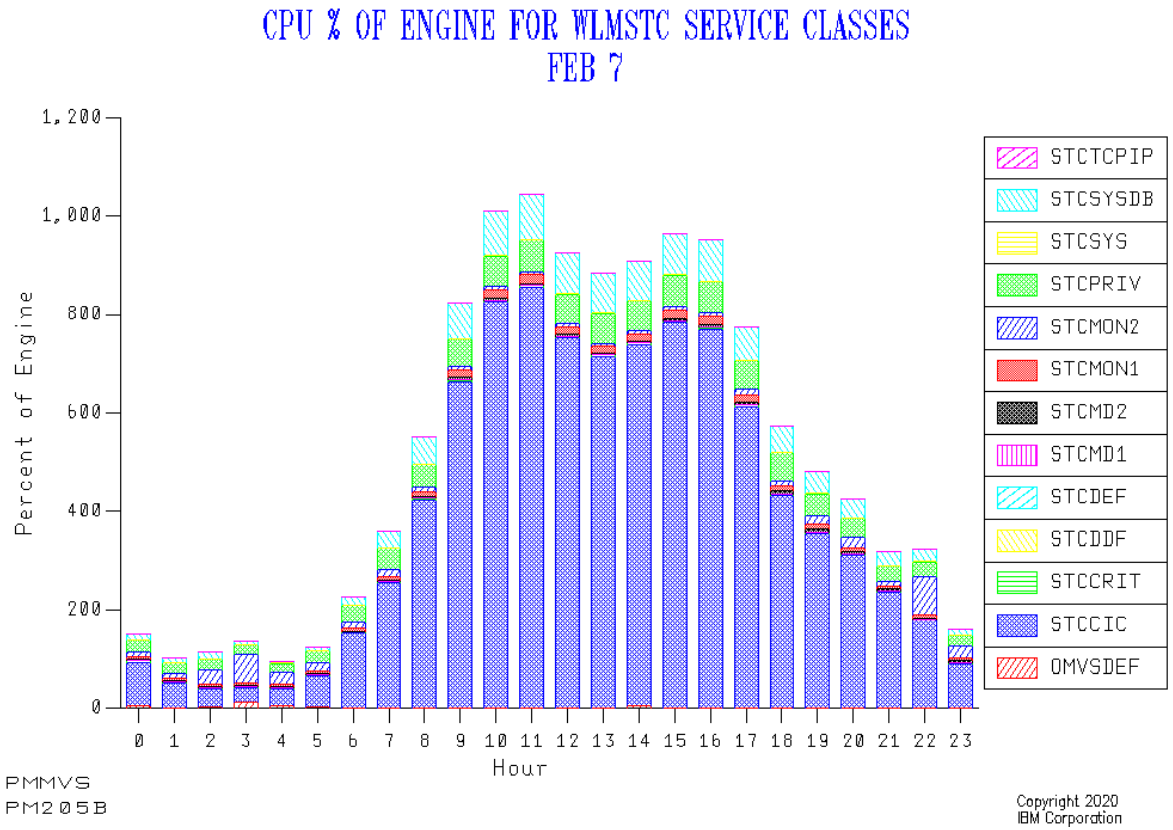
- Each e.g. CICS region, Db2 subsystem, MQ queue manager writes at least one
  - Always contains software level
  - Often contains a subsystem name
    - Db2 subsystem itself e.g. "DB2P"
    - MQ subsystem with connection type encoded e.g. "MQ01BATC"
- If you want to know whether a batch job uses Db2 or MQ you can do it
  - Without Db2 or MQ instrumentation
  - With the subsystem name

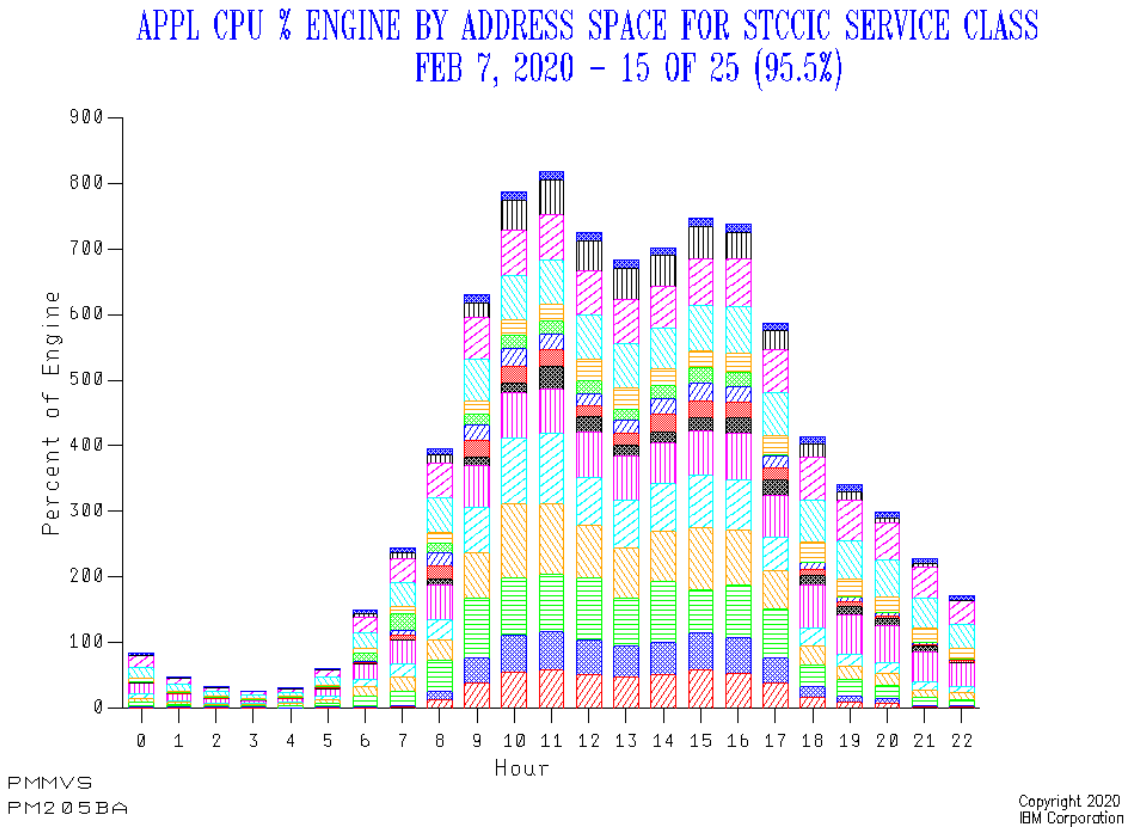
You Can Get IBM (And Some Other Vendors') Software Levels - From SMF 30 Usage Data

Note: Some other products appear in SMF 89

Product	Version	Identifier	SYSA	SYSB	SYSE	SYSP
z/OS	2.3	5645-001	X	X	X	X
CICS	5.2	5655-Y04	X			X
DB2	V11	5615-DB2	X	X	X	
MQ	V8	5655-W97	X		X	
WAS	V8 NOTUSAGE	5655-W65	X		X	

In Previous Example Most Of WLMSTC is CICS - Where Almost All CPU Is In 15 Regions Some Evidence Of "Cloned" Regions - Which Would Be Better Illustrated With A Line Graph

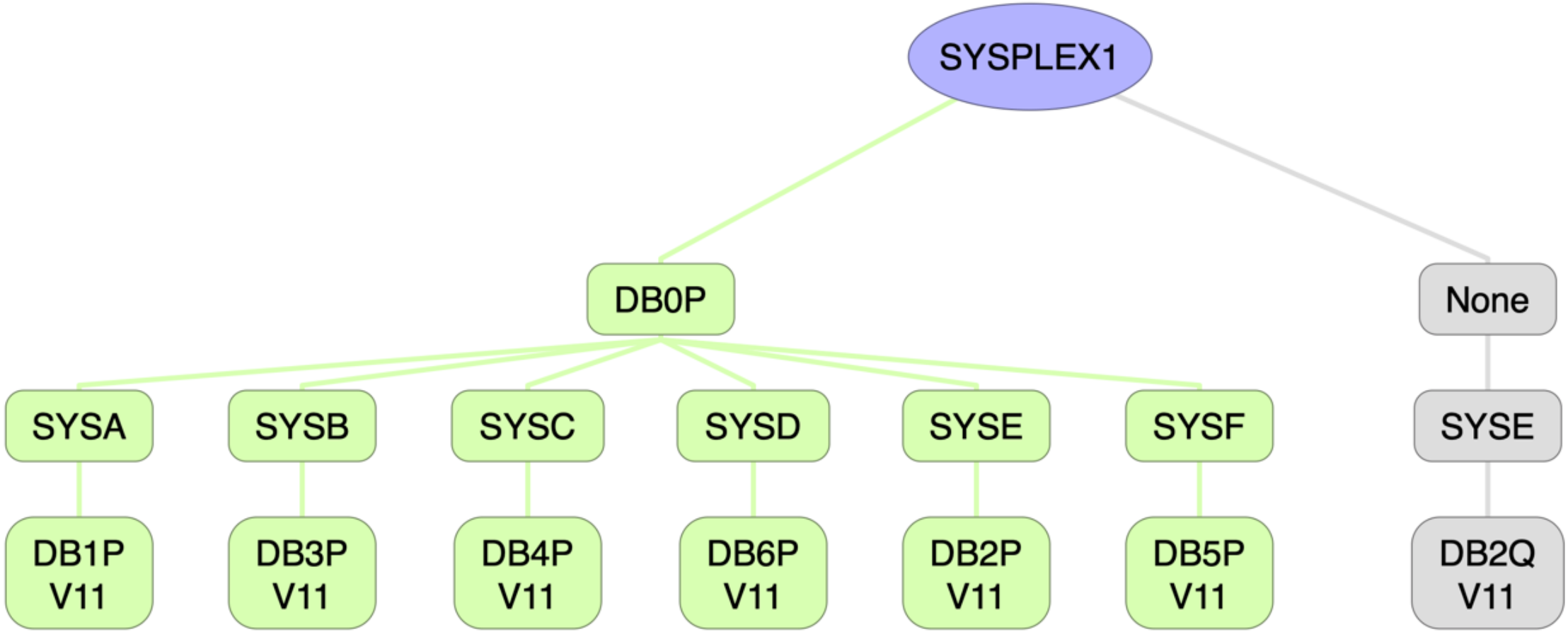




Db2 Without Db2 Instrumentation

- We can tell quite a bit about Db2 subsystems Without Using Db2 Instrumentation
  - Useful because e.g. Accounting Trace isn't always available
  - Provides a scalable method
- Which subsystems exist?
  - In which Datasharing Groups?
- What attaches to them?
- When is the subsystem maintained?
  - "Spikes" analysis could tell you when utilities run

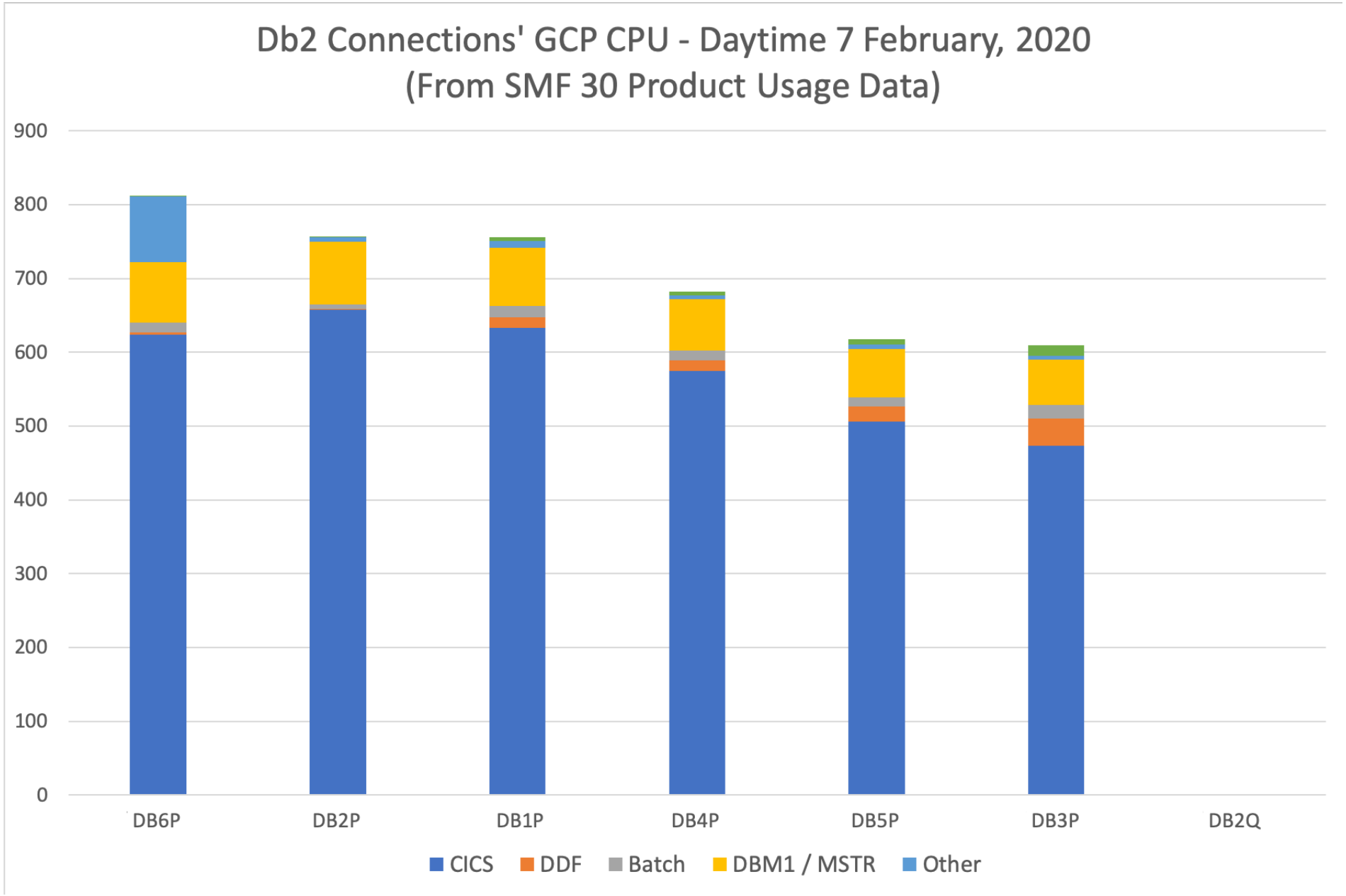
Db2 Subsystems and Datasharing



DB2 Subsystems and Datasharing - Notes

- Datasharing groups can be inferred from structure-level SMF 74-4
  - Including duplexing scheme
    - Can drive "white space" Coupling Facility memory discussion
- Membership of a datasharing group usually done with 74-2
  - IRLM XCF traffic in IXCLOnnn and especially DXR<*groupname*> groups
  - Hope Db2 subsystem and IRLM address spaces follow a naming convention
- Sometimes there's a "DB2P" subsystem and the rest don't follow that convention
  - What I term the "origin story"

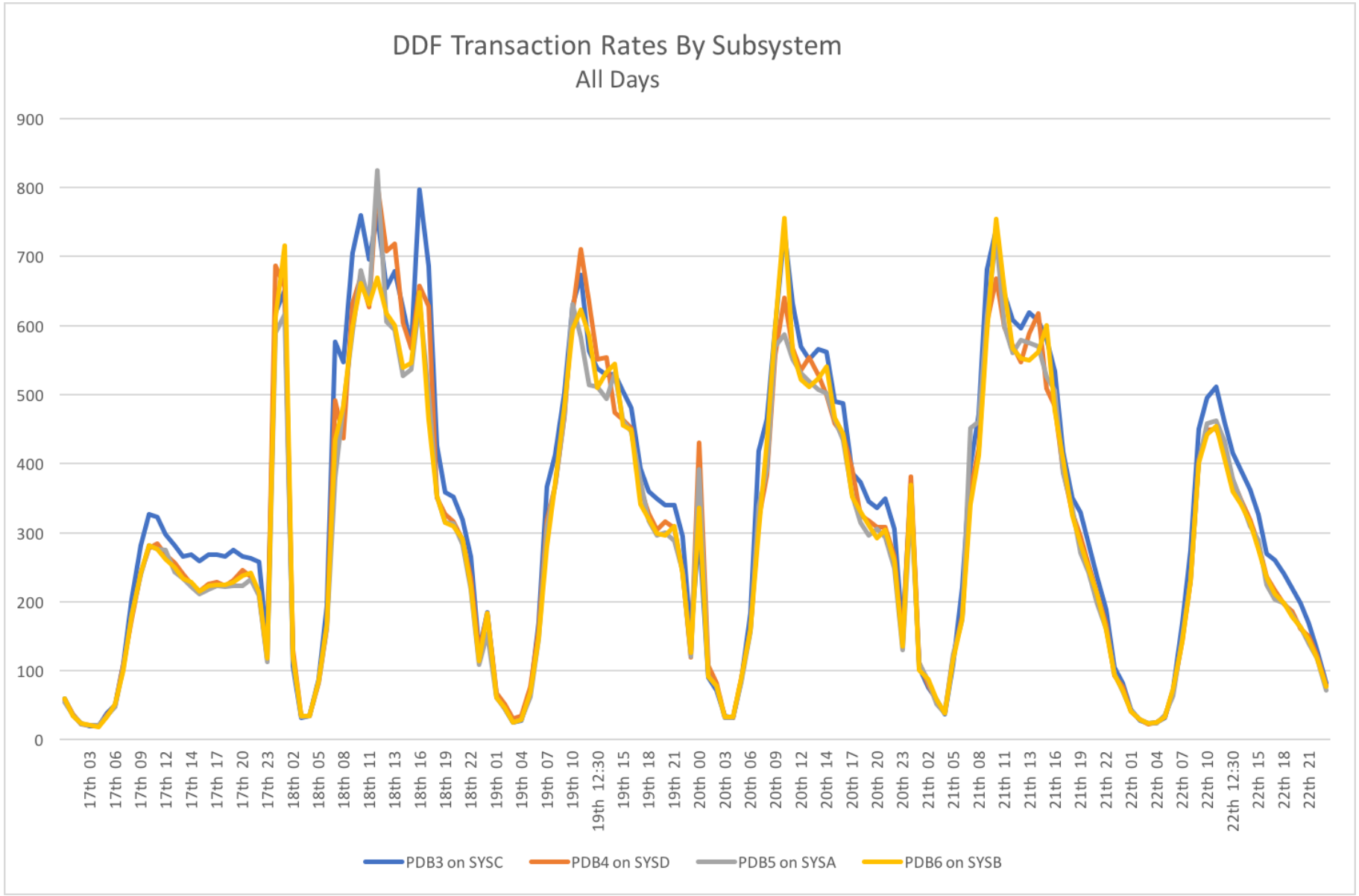
Db2 Subsystem Roles



Db2 Subsystem Roles - Notes

- Good "broad brush" approach
- Graph labelled "Db2 Connections' GCP CPU"
  - Using caller's SMF 30
    - Because most Db2 subsystems have "NO89" specified Usage Data Section won't have Db2 CPU
    - Obviously a superset of the true "in Db2" CPU
- Supplemented with Db2 subsystem's address spaces' CPU
  - Shows prevalence of eg Prefetch and Deferred Write (DBM1) and logging (MSTR)
- Similar technique for MQ
    - "NO89" not applicable to MQ
    - CHIN address space is very light compared to eg CICS
      - Fair bet CHIN involved in **all** MQ queues
      - So understates role of external connections
- Useful to show Absolute (this view) and Relative views side by side

DDF Scale And Rhythm Is In SMF 30 Independent Enclave Information
These Db2 Subsystems Are Regular And Moderately Well Balanced, Except On 18th



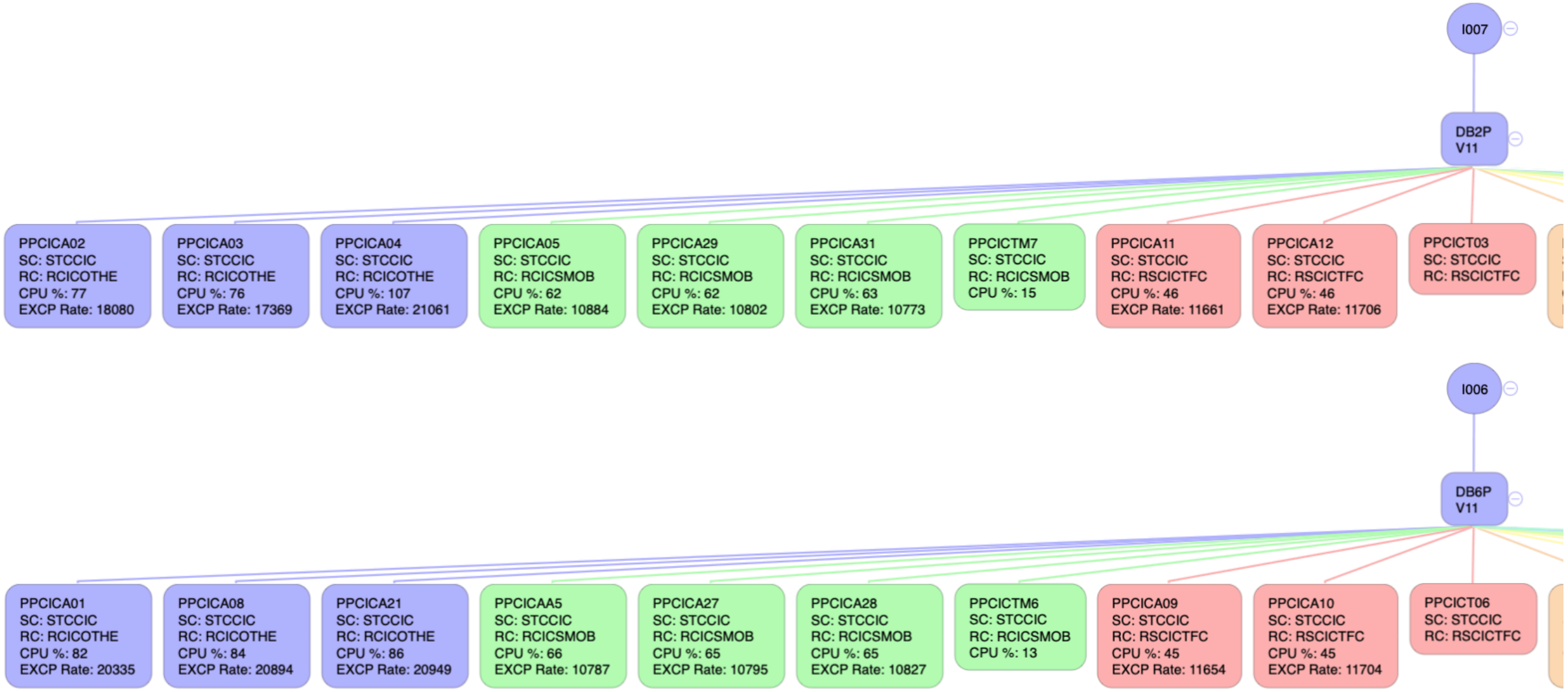
While We’re Talking About DDF

- SMF 30 - for DIST address space gives
  - Enclave Transaction Rate (COMMIT Rate)
  - Enclave Transaction CPU
  - Enclave Transaction Response Time
- SMF 72-3 - for DDF Service & Report Classes
  - Transaction Rate
  - CPU Time
  - Response Times
    - In 14 buckets around goal time
- So you needn't go to Db2 Accounting Trace
  - But there's so much more if you do

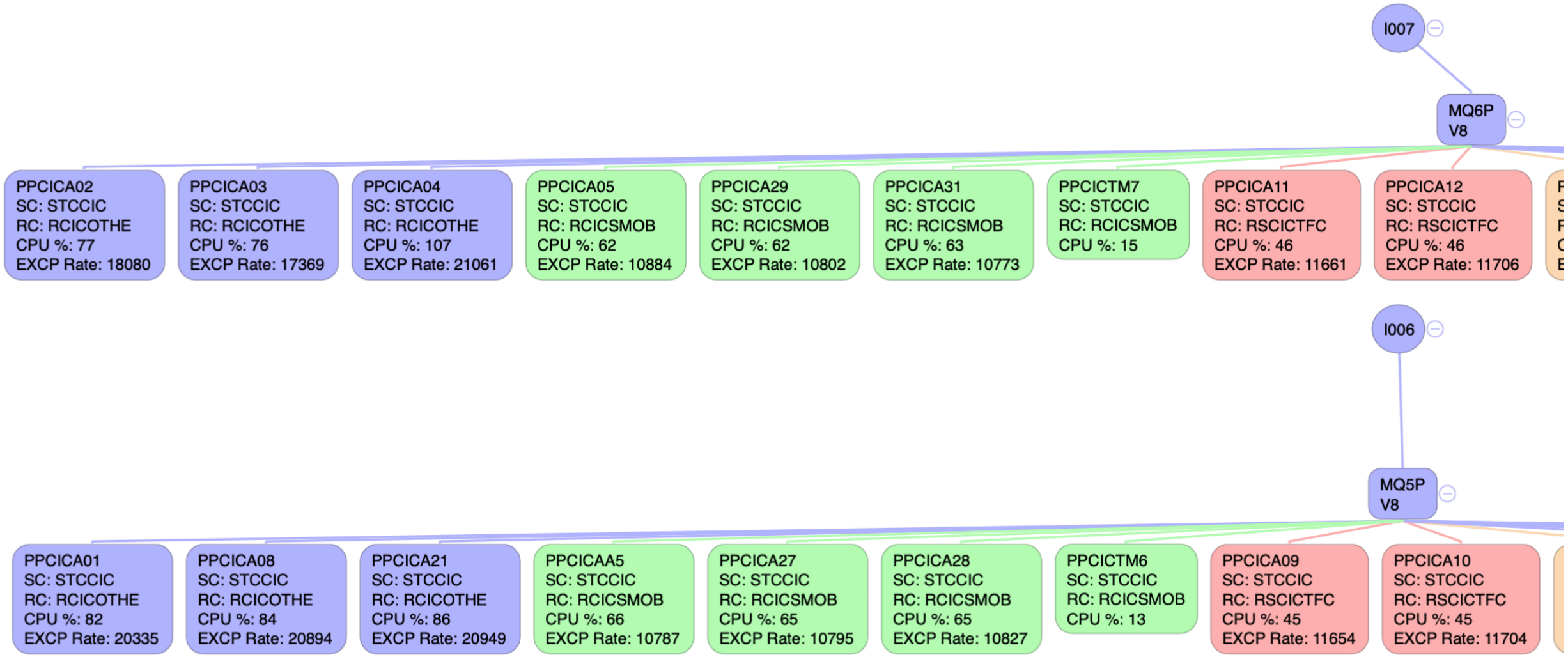
CICS Architectural Topics

- Roles
  - Concepts like "TOR", "AOR", "DOR", "FOR", "QOR" fuzzy
  - Liberty Java licence evident in SMF 30 Usage Data Section
    - Note: Its software level is not evident - so take your cue from "DFHSIP" section level
- Cloning
- Restarts
- Virtual Storage
  - Below the line
  - Above the line
  - Above the bar
- Version in SMF 30 Usage Data Section
  - Can answer questions like "which regions aren't Version 5 Release 5?"
- zIIP usage for CICS is a useful (and rare) thing to spot
  - Indicates something "modern" Java or System XML
  - CICS Transaction Gateway almost pure Java

CICS Connections - Db2 (Part Of the Estate)



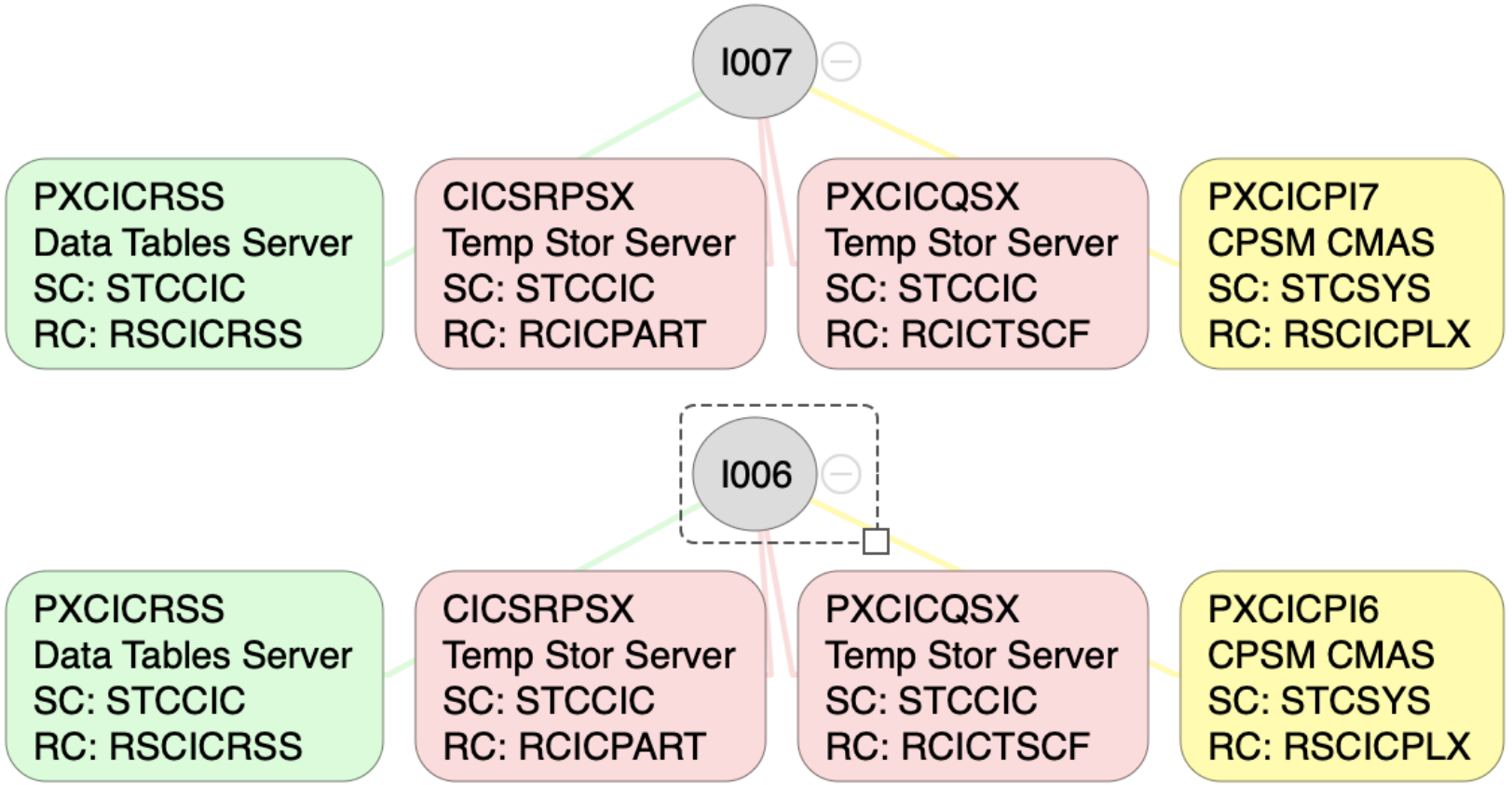
CICS Connections - MQ (Part Of the Estate)



Some Observations From The Last Two Slides

- Report Class used to identify applications
  - Which I've coloured
- Some extremely high EXCP rates
  - Which turn out to be for VSAM RLS and corroborated by SMSVSAM SMF 42-6
  - Some variation between "cloned" regions
- Some **average** CPU consumption figures that might worry you
  - But the QR TCB% is probably only a small subset of this
    - But "when to clone?" is a good question
- Some regions look like TORs
  - "T" instead of "A" in the name
  - Low CPU and I/O rate
- I could have added 74-2 XCF message rate and group
- These slides are static pictures
  - Time of day could prove interesting

Some CICS-Related Address Spaces Don’t Connect To Either Db2 Or MQ
In this example there are no Named Counter servers nor CICS regions nor CTG address spaces



An Interesting Situation: Daily CICS Restarts, With 30-50 Minute Down Times

Note: Regions taken down on a timer pop

Name	Start Time	Stop Date	Stop Time	Up Time	Down Time Since Last
Wednesday					
PPCICA01 PPCICA16	01:01	Next Day	**00:29**	23h 28m	
PPCICA08-10 PPCICA21 PPCICA74 PPCICAA5	01:07	Next Day	**00:30**	23h 22m	
PPCICA27,28 PPCICA84,85	01:12	Next Day	**00:29**	23h 17m	
CICSRPA6	01:22	Next Day	**00:30**	23h 7m	
Thursday					
PPCICA01 PPCICA16	01:01	Next Day	**00:29**	23h 28m	31m
PPCICA08-10 PPCICA21 PPCICA74 PPCICAA5	01:07	Next Day	**00:29**	23h 22m	37m
PPCICA27,28 PPCICA84	01:12	Next Day	**00:30**	23h 17m	42m
PPCICA85 CICSRPA6	01:19	Next Day	**00:29**	23h 10m	49m
Friday					
PPCICA01 PPCICA16	01:01				31m
PPCICA08-10 PPCICA21 PPCICA74 PPCICAA5	01:07				37m
PPCICA27,28 PPCICA84,85	01:13				
CICSRPA6	01:22				52m

IMS

- Comprises 3 different types of address space:
  - System - program name DFSMVRC0
  - Dependent regions - program name DFSRRC00
  - IRLM - program name DXRRLM00 (same as for Db2)
- Each (sub)system has three mandatory address spaces
  - Control Region
  - DBRC
  - DL/I SAS
- Dependent regions look identical to each other

- Usage Data Section doesn't tie together system / dependent region
  - No system encoding in Product Qualifier
- Hope for a naming convention
- WLM Service / Report Class might help

zCX Container Extensions

- Run Docker containers in a z/OS address space (instance)
  - Potentially multiple containers per instance
    - Architectural question of which containers in which instance
- JCL program name is GLZBAIN
- Memory is long-term page fixed
- Highly zIIP-eligible
  - Might need to manage with eg IIPHONORPRIORITY

- SMF does not know what is in each zCX instance
  - Nor how containers individually behave
  - No prioritisation **within** a zCX instance
    - Not using independent enclaves
- Use Docker's own instrumentation
  - But keep an eye on the address space's overall resource usage - with SMF 30
- Refer to Redbook: Getting started with z/OS Container Extensions and Docker

I/O Subsystem

- Lots of configuration information in RMF SMF
  - 73 channel path
  - 78-3 control unit to channel
  - 74-7 switch (if CUP enabled)
  - 74-1, 74-5, 74-8 controller and disk
- You can't see distance
  - But you might infer it from e.g. Disconnect Time
    - Also zHyperLink
- An interesting question I keep asking is "why is this channel rated at 16 only going at 4?"
  - SMF73GEN gives maximum and negotiated data rate

The Bottom Layer - Application Components

Middleware Traces Help You Dig Deeper

- Db2 and MQ Accounting Trace
  - Plan / Connection / Correlation ID identify actor
    - Connection Type e.g. "QWHCCICS"
    - Examples:
      - For Batch, Correlation ID is (usually) job name
      - For CICS, characters 4-7 of Correlation ID is transaction name
    - Even more detailed identification for Db2 with Package-Level Accounting
      - For example, stored procedures called
  - Voluminous if not expensive
    - Enable overnight for batch
    - Enable for, say, 30 minutes off the shoulder of the peak
  - DDF Information extremely specific
- CICS Monitor Trace
  - More voluminous / expensive and very detailed
    - Mostly can fall back on Db2 or MQ Accounting
      - But many transactions not recorded that way

Data Set Names Are Useful

- Full panoply of information in SMF
  - SMF 14/15 for non-VSAM
  - SMF 62/64 for VSAM
  - SMF 42-6 for almost all data sets
  - SMF 30 has EXCPs but also has DD information
- Db2 is a special case
  - Owning the I/O to its databases and logs
- Db2 42-6 information can be particularly interesting:
  - Database is 3rd qualifier
  - Tablespace / Index Space is 4th qualifier
  - Data set number in last qualifier has multiple uses:
    - Hotspot identification
    - Partition counting
      - Is it significant if I keep seeing tablespaces with 54 partitions?

How Does This Fulfil The Definition Of Architecture?

So You Probably Aren’t An Architect But The Role You Can Play Is Very Valuable

- Understand what's really going on
- Especially valuable to **real** systems architects
- Keep up to date view of the estate
- Inform your own capacity and performance role
- Personally, dabbling in architecture keeps me interested
  - Not that Performance and Capacity are boring
  - It just provides another angle
  - Gives me another way to find patterns among customers

- I'm seeing increasingly complex estates
  - Hundreds of CICS regions, Scores of Db2 subsystems, Dozens of MQ queue managers
  - Keeping track of it is really tough
    - Actually true for smaller estates