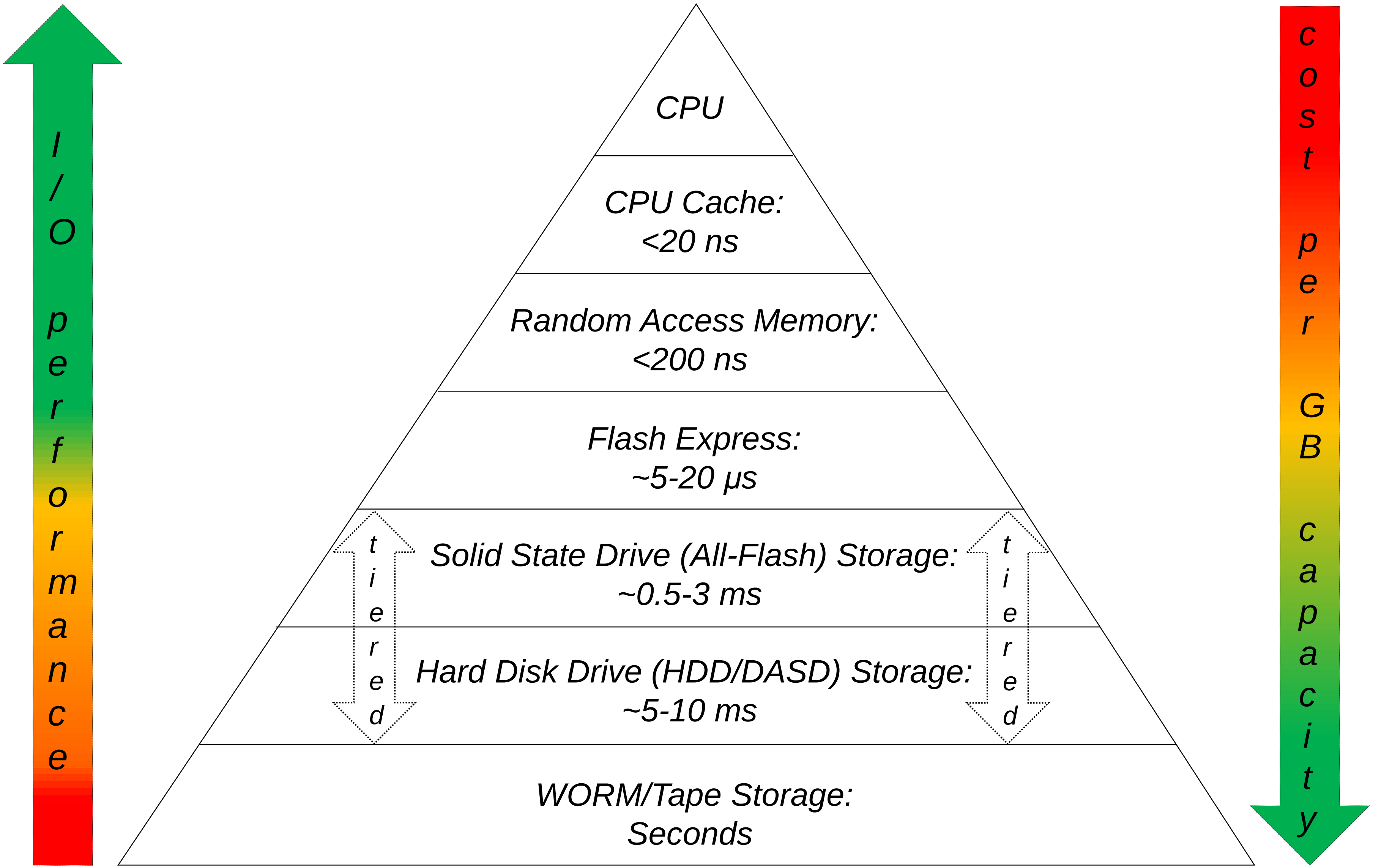
DASD I/O Performance Management Is Easy? Actually, Performed Well, It's Not Necessarily That Easy!

Michael W. Moss
Value-4IT
mossmw(at)value-4it.com
Date of presentation (02/11/2016)
Session <LJ>

Agenda

- The Data In Memory Storage Hierarchy
- DASD I/O Analysis Process & Basics
- DASD I/O Operation Breakdown
- DASD I/O Workload Analysis Overview
- DASD I/O Workload Analysis via RMF
- Advanced DASD Functions Overview
- The Disk Tuning Expert Demise?
- Collaboration With The Disk IHV
- DASD I/O Analysis Recommendations
- Real-Life z/OS DASD Case Studies
- Questions & Session Feedback

The Data In Memory Storage Hierarchy



Falling memory costs reduce physical/device level I/O activity

DASD I/O Analysis: Who/When/Why?

Why: Traditional Reactive Approach

- ❖ SLA objectives compromised
- ❖ Unacceptable user response times
- ❖ Batch Window duration issues
- ❖ Key I/O indicators being breached

Fixing a one-off I/O issue in isolation, typically doesn't resolve the overall performance problem.

Who/When: The SMF Data Owner?

- ❖ Typically the CPU/Performance team
- ❖ Their focus is typically CPU (R4HA)
- ❖ Their SMF data reporting tool is generic, not RMF DASD I/O based
- ❖ Typical daily PDB data creation

Too many anomalies for today's I/O?

Structured Overall System Approach

- ✓ Consider the complete DASD I/O subsystem, not one component
- ✓ Mange to SLA objectives
- ✓ Identify the issue in terms of I/O component (E.g. DISC, IOSQ)

A "Balanced System" is the utopia & the best I/O is a memory based I/O!

Storage Management Centric Process

- ✓ Daily aggregation of "golden" RMF data, trending for 12+ Months
- ✓ Personnel centric data classification (E.g. Manager, Technician, IHV)
- ✓ Proactive 24*7 I/O issue alerts

I/O is system wide, Application, DBA, Performance & Storage consideration

*A proactive 24*7 BAU process, ideally performed by Storage Management*

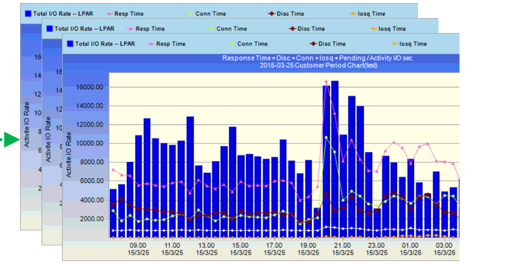
DASD Performance Management Process

Real Time Monitor

Reporting Process

Issue(s) Identified?

Full DASD I/O Analysis



** RMF/CMF Data Base*
** SAS/MXG/MICS Reports*
** Expert ISV System Tools*

Why not 1 "Golden Record" set of expert reports, for all personnel levels, I/O usage, error analysis, et al?

System z I/O Analysis
* CEC/LPAR/SYSPLEX
* CSS/FICON/zHPF

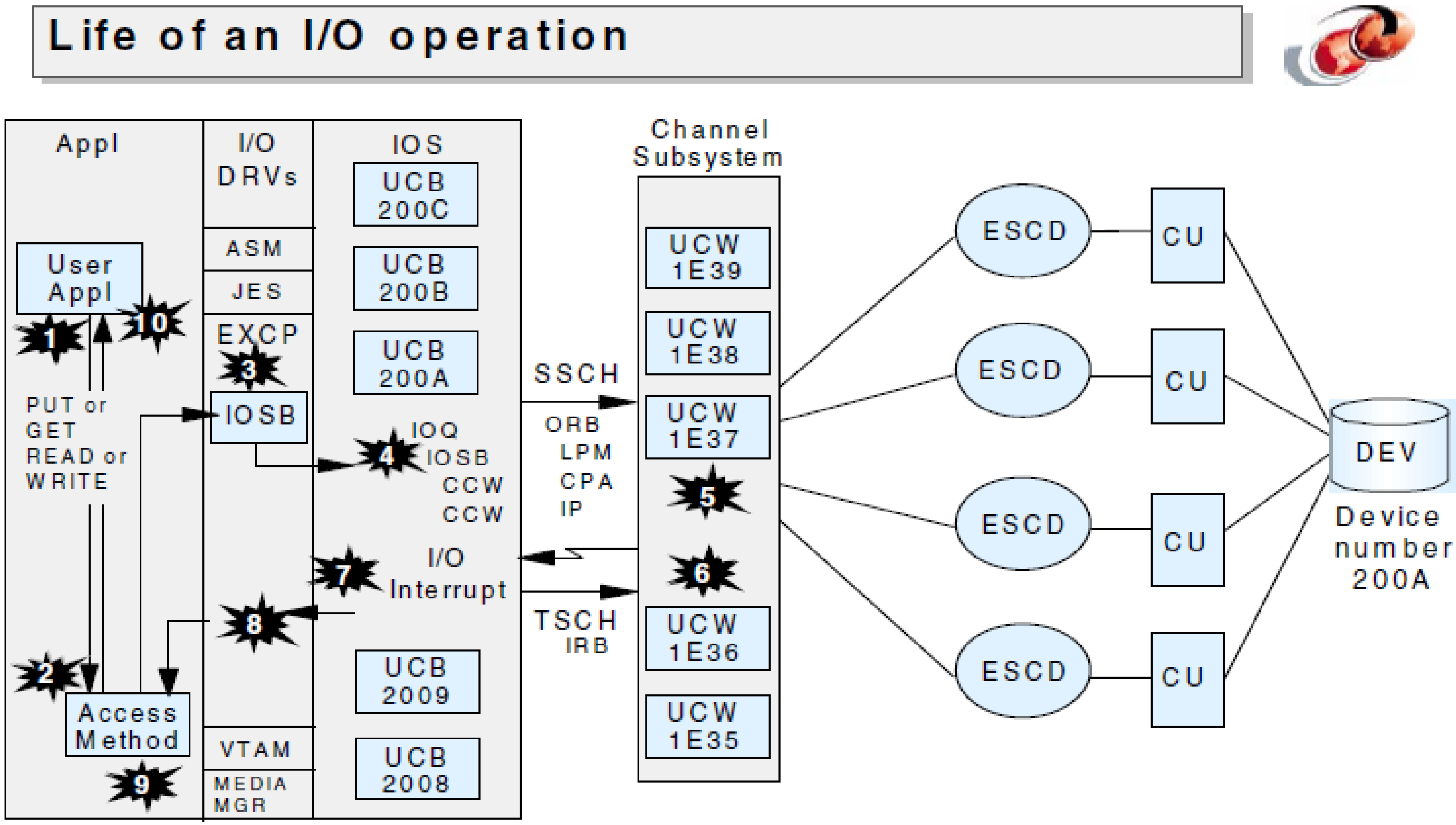
Subsystem Analysis
* SSD-Flash/Cache
* SSID/LCU

DASD Analysis
* Response Time
* I/O Busy Rate

SMTP Escalation
* Internal Experts
* 3rd Party IHV?

Conceptually, very simple, but the devil as always, is in the detail!

I/O Operation Flow: The S/390 Legacy View



Recommended Reading (~2001): z/OS Intelligent Resource Director Redbook (SG24-5952-00)

Know the big picture; DASD performance starts with the Channel (CSS)!

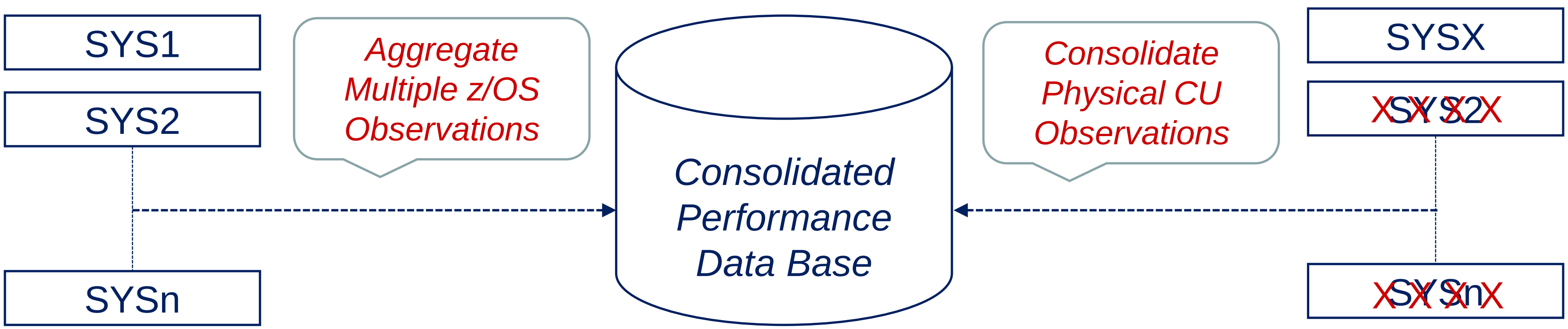
Collecting The Required Performance Data

Minimum SMF Data Required For DASD Performance Analysis

SMF (RMF) Record Number	***Description***
Type 73 (X49)	RMF Channel Path Activity
Type 74 (X4A) Subtype 1	RMF Device Activity
Type 74 (X4A) Subtype 5	RMF Cache Subsystem Device Activity
Type 74 (X4A) Subtype 7	RMF FICON Director Statistics
Type 74 (X4A) Subtype 8	RMF Enterprise Disk System Statistics
Type 78 (X4E) Subtype 3	RMF I/O Queuing & HyperPAV Activity
Type 42 (X2A) Subtype 6	*DASD Data Set Level I/O Statistics*

Note: EMC VMAX & HDS G1000 generate their own SMF records, for further detailed disk subsystem analysis.

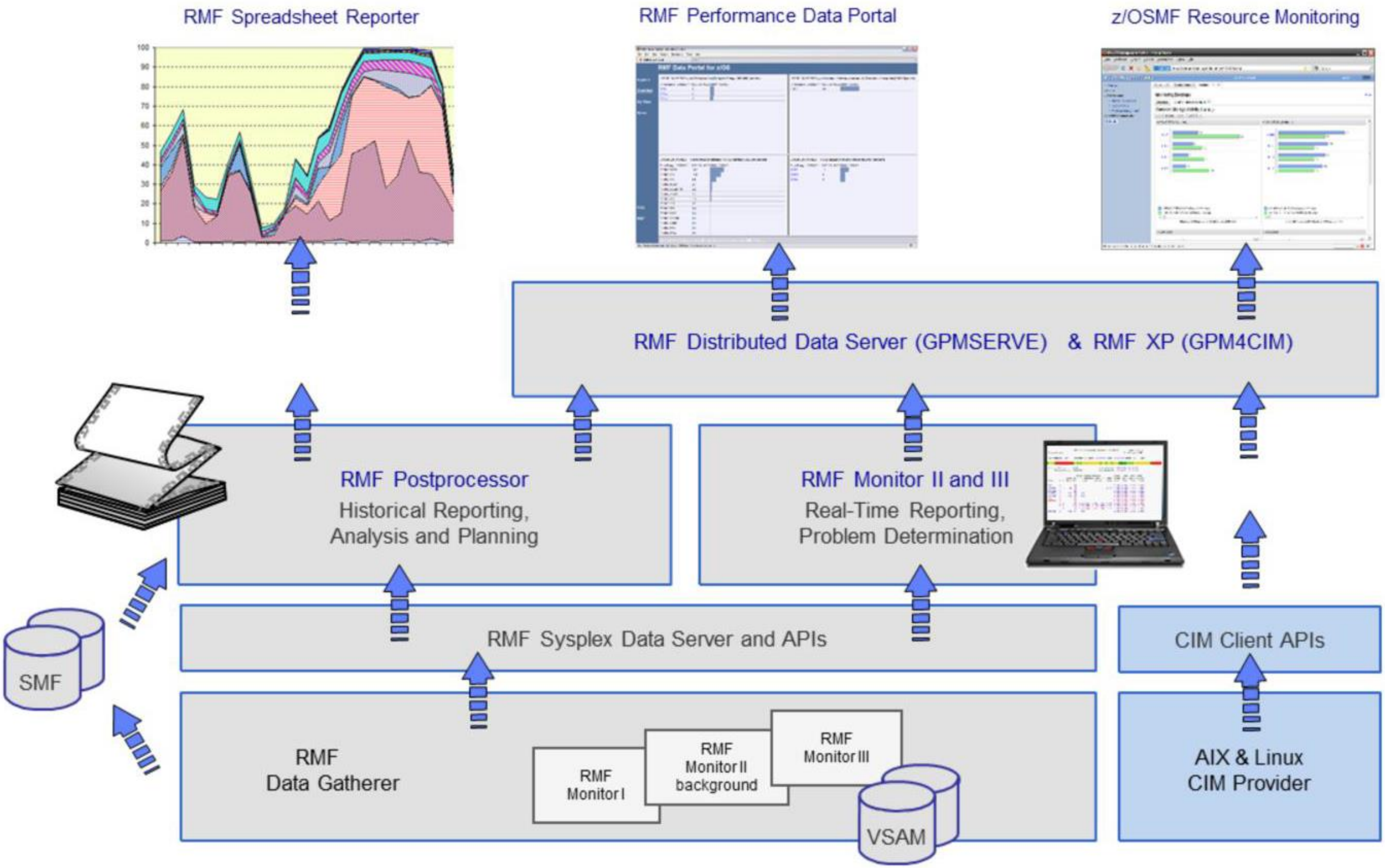
Consolidate, Deduplicate & Enhance via Analytics – Avoid Too Much Data, Not Enough Facts



Recommended Reading: z/OS MVS System Management Facilities (SMF) - ieag200

RMF (CMF) evolves & improves, hence mandatory big data analytics…

Resource Measurement Facility (RMF) Overview



RMF continues to evolve & improve since its first introduction in 1976…

DASD I/O Operation Timing: Fundamentals

z/OS Time | Hardware Service Time

DASD I/O Response Time = IOSQ Time + PEND(ing) Time + DISC(onnect) Time + CONN(ect) Time

DASD I/O Response Time

IOSQ Time: z/OS logical volume (I.E. UCB), *wait time*
Pending Time: Hardware resource (E.g. Disk, Channel, Switch), *wait time*
Disconnect Time: I/O (E.g. Read, Write, Sync) & Control Unit miss, *wait time*
Connect Time: I/O (E.g. Disk, Channel, CPU) operation, *transfer time*

Recommended Reading: z/OS RMF Performance Management Guide - SC33-7992-nn

Response time analysis includes software & hardware considerations

DASD I/O Operation: Typical Issue Analysis

DASD I/O Response Time = IOSQ Time + PEND(ing) Time + DISC(onnect) Time + CONN(ect) Time

DASD I/O Response Time: Issue Analysis			
Queueing Time	***Hardware Service Time***		
z/OS Time	***PEND(ing) Time***	***DISC(onnect) Time***	***CONN(ect) Time***
Device (USB) busy	Cross System (I.E. z/OS) device reserve	Read cache miss	FICON channel busy (E.g. Speed, Protocol)
High HyperPAV alias allocation (E.g. LCU, Port, Channel contention)	Command Response (CMR) delay (E.g. CHPID, Port)	Synchronous Remote Copy write completion	High FICON director port utilization (PEND(ing))
Device error recovery (NB. z/OS operation)	System Assist Processor (SAP) overhead	High (synchronous) write cache workload	
	Concurrent write operation domain extent conflict	Control unit busy	
	High FICON director port utilization (CONN(ect))	Multiple allegiance; PAV write extent conflicts	

Know & appreciate your enemy for DASD I/O operation response time!

DASD I/O Operation: Issue Resolution Functions

DASD I/O Response Time: Function Enhancments			
Queueing Time	**Hardware Service Time**		
z/OS Time	**PEND(ing) Time**	**DISC(onnect) Time**	**CONN(ect) Time**
Parallel Access Volume (PAV): Multiple I/O operations	Multiple Allegiance (I.E. I/O parallelism, DB2, IMS)	Adaptive Multi-stream Prefetching (AMP): Optimal sequential read performance (DB2)	Modified Indirect Data Address Word (MIDAW): Optimized processing (EF, 4K Blocks, DB2 PF)
HyperPAV: Simplify PAV with control unit alias		Intelligent Write Caching (IWC): Optimal write cache management	*High Performance FICON for z (zHPF): Optimal single disk CU port I/O*
		Sequential Prefetching in Adaptive Replacement Cache (SARC): Self optimization & tuning	FICON Express nnS: 16, 32+ Gbps performance
		zHyperWrite: Metro Mirror HyperSwap DB2 log writes	
		Easy Tier Integration (DB2): Ensure same tier for all data extents	
		List Prefetch Optimizer: Faster zHPF fetching of unordered DB2 data	

Disk controller functions can resolve many I/O performance issues!

I/O Workload Analysis: Hierarchical Approach

SLA/KPI: Not all workloads and devices are equal. Safeguard you know all of your business priorities before performing any I/O tuning. This will include WLM considerations, both CPU & I/O related. Eliminate isolated/interim/one-off/erroneous/device failure performance issues!

CPU/Memory: Safeguard a "balanced system" approach. Faster I/O processing dictates a higher CPU requirement for the increased workload. Collaboration is the key, Storage, DBA & CPU Capacity Planning personnel all play a key role as System z Data In Memory functions evolve (E.g. CPU Chip Cache, Memory Layers, Disk Memory & Functions).

Disk Subsystem: Before analysing individual DASD volumes, analyse the physical (cache, channels) & logical (LCU) performance of the disk subsystem (I.E. IBM DS8000, EMC VMAX, HDS VSP). Ultimately high DISC(onnect) times typically dictate insufficient disk cache capacity; put simply, DISC time is inversely proportional to disk cache effectiveness.

Disk Volumes: Having considered the Disk Subsystem, this will lead onto analysis of logical DASD volumes. Ultimately IOSQ is the #1 device level issue, as all device level I/O, queues at the same speed. PAV and HyperPAV are the simplest solutions for eradicating IOSQ issues, but even HyperPAV has concurrent alias allocation considerations. Sometimes, an individual file (data set) can cause I/O issues; hence DBA (Application) collaboration…

Reporting: Safeguard your Management team and peers receive your structured I/O analysis reports. As per the reality of high DISC and IOSQ times, sometimes only technology refreshes will increase I/O throughput. Faster FICON channels or protocols will reduce CONN(ect) time; higher capacity disk cache eradicates DISC(onnect) time, et al…

In an ideal world, every I/O would be a Data In Memory I/O!

I/O Workload Analysis: Keep It Simple…

Rules Of Thumb (ROTs): A rule of thumb is a principle with broad application that is not intended to be strictly accurate or reliable for every situation. A ROT only applies while that technology and/or related function is current. *Recommendation: Don't use Rules Of Thumb (ROTs) for your I/O workloads*.

Baseline Data: A minimum or starting point used for comparisons. A baseline evolves over time, while being a true reflection of your workload. When changes occur, whether technology, workload or policy related, the baseline changes can be measured in terms of success criteria & future trends. *Recommendation: Don't overlook the facts, a baseline never falsifies data.*

Balance: Because of the storage hierarchy, balancing I/O workloads requires careful thought, as sooner or later, ideally sooner, the I/O will be processed by the CPU resource. Adding cache memory resource, impacts the balance. *Recommendation: Optimize the resources you have, before adding capacity.*

256 Device LCU

Base Architecture: Each Logical Control Unit (LCU) is restricted to 256 logical volumes (disk devices). In an ideal world, all LCUs would be equal, which can only follow, by allocating each & every LCU with 256 device addresses. Currently, a physical DASD subsystem supports 4096 devices via 16 LCUs. *Observation: A 256 device LCU design isn't a constraint, just a consideration.*

Let common sense apply & don't over complicate this important process!

Cache Subsystem Summary: A Basic RMF View

C A C H E   S U B S Y S T E M   S U M M A R Y

PAGE 1

z/OS V1R13 SYSTEM ID OS04 DATE 09/28/2011 INTERVAL 14.59.996

RPT VERSION V1R13 RMF TIME 09.30.00

SSID	CU-ID	TYPE	CACHE	NVS	I/O RATE	OFF RATE	--CACHE HIT RATE- READ	DFW	CFW	--------DASD I/O RATE-------- STAGE	DFWBP	ICL	BYP	OTHER	ASYNC RATE	TOTAL H/R	READ H/R	WRITE H/R	% READ
0600	0600	3990-006	256	64	3.1	0.0	1.0	2.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	1.000	1.000	30.8
06C0	06E0	3990-006	512	32	38.6	0.0	16.5	22.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.999	0.997	1.000	42.9
0A80	0ABE	3990-006	512	32	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	N/A	N/A	N/A
6801	4000	2105-F20	8192	192	55.3	0.0	0.0	55.2	0.0	0.0	165.5	0.0	0.0	0.0	486.2	1.000	1.000	1.000	0.1
6802	4100	2105-F20	8192	192	49.5	0.0	0.2	49.3	0.0	0.0	145.1	0.0	0.0	0.0	500.8	1.000	1.000	1.000	0.4
6803	4200	2105-F20	8192	192	56.9	0.0	0.2	56.7	0.0	0.0	169.9	0.0	0.0	0.0	489.1	1.000	1.000	1.000	0.4
6804	4300	2105-F20	8192	192	50.7	0.0	0.2	50.5	0.0	0.0	149.0	0.0	0.0	0.0	465.6	1.000	1.000	1.000	0.4
6805	4400	2105-F20	8192	192	53.0	0.0	0.2	52.8	0.0	0.0	158.3	0.0	0.0	0.0	467.6	1.000	1.000	1.000	0.4
6806	4500	2105-F20	8192	192	42.0	0.0	0.3	41.8	0.0	0.0	123.6	0.0	0.0	0.0	434.6	1.000	1.000	1.000	0.6
6807	4600	2105-F20	8192	192	50.3	0.0	0.2	50.0	0.0	0.0	150.2	0.0	0.0	0.0	454.1	1.000	1.000	1.000	0.4

The RMF Cache Subsystem Summary report provides you with an entry point into the hierarchical (top-down) approach when performing I/O analysis activities. This report provides a simplistic viewpoint of what disk subsystems might be important to you, where any disk subsystems causing concern can be easily identified. However, subsequent and more detailed RMF reports would be required, providing more detail.

A great starting point, identifying all DASD subsystems & related statistics

Cache Subsystem Summary: Top 20 Device View

C A C H E   S U B S Y S T E M   S U M M A R Y

PAGE 2

z/OS V1R13 SYSTEM ID OS04 DATE 06/05/2011 INTERVAL 14.59.996

RPT VERSION V1R13 RMF TIME 09.30.00

*** DEVICE LIST BY DASD I/O RATE ***

VOLUME	DEV	SSID	%	I/O	---CACHE HIT RATE--			----------DASD I/O RATE----------					ASYNC	TOTAL	READ	WRITE	%
SERIAL	NUM		I/O	RATE	READ	DFW	CFW	STAGE	DFWBP	ICL	BYP	OTHER	RATE	H/R	H/R	H/R	READ
PRD440	077E	00B1	25.7	53.1	21.3	16.7	0.0	13.9	0.0	1.3	0.0	0.0	1.6	0.714	0.604	1.000	67.9
PPDS14	0220	00CC	6.9	29.1	1.0	15.7	0.0	0.1	0.0	12.4	0.0	0.0	2.0	0.573	0.943	1.000	6.2
PRD437	0214	00CC	4.5	19.0	12.5	2.0	0.0	4.4	0.0	0.0	0.0	0.0	2.0	0.765	0.738	0.998	89.2
PPD026	0B76	00F1	27.4	71.4	65.9	1.1	0.0	4.3	0.0	0.0	0.0	0.0	0.5	0.939	0.939	0.975	98.4
PRD554	06B9	00FE	59.1	32.4	28.9	0.0	0.0	3.5	0.0	0.0	0.0	0.0	0.0	0.893	0.893	1.000	99.9
PRD339	0231	00CC	2.4	10.3	7.2	0.0	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.698	0.698	N/A	0.0
...																	

*** DEVICE LIST BY TOTAL I/O RATE ***

VOLUME	DEV	SSID	%	I/O	---CACHE HIT RATE--			----------DASD I/O RATE----------					ASYNC	TOTAL	READ	WRITE	%
SERIAL	NUM		I/O	RATE	READ	DFW	CFW	STAGE	DFWBP	ICL	BYP	OTHER	RATE	H/R	H/R	H/R	READ
PPD026	0B76	00F1	27.4	71.4	65.9	1.1	0.0	4.3	0.0	0.0	0.0	0.0	0.5	0.939	0.939	0.975	98.4
PRD440	077E	00B1	25.7	53.1	21.3	16.7	0.0	13.9	0.0	1.3	0.0	0.0	1.6	0.714	0.604	1.000	67.9
PRD327	0200	00CC	11.8	49.8	3.0	46.8	0.0	0.1	0.0	0.0	0.0	0.0	6.3	0.998	0.973	1.000	6.2
PRD343	0515	00E4	17.5	48.9	24.8	24.0	0.0	0.1	0.0	0.0	0.0	0.0	5.5	0.998	0.996	0.999	50.8
PBV321	022C	00CC	11.2	47.3	47.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.1	1.000	1.000	1.000	99.7
PRD307	0507	00E4	15.3	42.8	4.1	38.2	0.0	0.5	0.0	0.0	0.0	0.0	13.0	0.988	0.891	1.000	10.8
...																	

This RMF Cache Subsystem Summary report consists of two top-20 lists of devices, DASD I/O Rate & Total I/O rate. These two lists allow you to identify the volumes with the highest I/O rates to the lower interface of a subsystem as well as the volumes with the highest I/O rates in total. Solving a possible problem with one of the listed devices would probably be of most benefit to the overall subsystem.

A top-20 rank report helps, but what is the performance baseline?

Cache Subsystem Summary: Activity Overview

GSE GUIDE SHARE EUROPE UK REGION

```
                                   C A C H E   S U B S Y S T E M   A C T I V I T Y
                                                                                                    PAGE    1
          z/OS V1R13             SYSTEM ID OS04             DATE 09/28/2011          INTERVAL 14.59.996
                                 RPT VERSION V1R13 RMF      TIME 09.30.00

SUBSYSTEM  3990-06      CU-ID  074C      SSID 00B1      CDATE  06/05/2011      CTIME 09.30.03       CINT  15.00
TYPE-MODEL 9396-001

TOTAL I/O   186363   CACHE I/O   181922   CACHE OFFLINE        0
TOTAL H/R    0.865   CACHE H/R    0.886

CACHE I/O    -------------READ I/O REQUESTS-------------  ----------------------WRITE I/O REQUESTS---------------------       %
REQUESTS       COUNT     RATE      HITS     RATE     H/R     COUNT     RATE      FAST     RATE      HITS     RATE     H/R      READ

NORMAL        116003    128.9     95313    105.9   0.822     15150     16.8     15150     16.8     15134     16.8   0.999      88.4
SEQUENTIAL     16655     18.5     16617     18.5   0.998     34114     37.9     34114     37.9     34114     37.9   1.000      32.8
CFW DATA           0      0.0         0      0.0     N/A         0      0.0         0      0.0         0      0.0     N/A       N/A

TOTAL         132658    147.4    111930    124.4   0.844     49264     54.7     49264     54.7     49248     54.7   1.000      72.9

----------------------CACHE MISSES----------------------          -----------MISC-----------          ------NON-CACHE I/O-----
REQUESTS       READ     RATE    WRITE     RATE   TRACKS     RATE                      COUNT     RATE                  COUNT     RATE
                                                                   DFW BYPASS            289      0.3       ICL           4435      4.9
NORMAL        20690     23.0       16      0.0    12993     14.4   CFW BYPASS              0      0.0       BYPASS           6      0.0
SEQUENTIAL       38      0.0        0      0.0      870      1.0   DFW INHIBIT             0      0.0       TOTAL         4441      4.9
CFW DATA          0      0.0        0      0.0                     ASYNC (TRKS)         7239      8.0

TOTAL         20744     RATE     23.0

---CKD STATISTICS---     ---RECORD CACHING---

WRITE             268     READ MISSES     14502
WRITE HITS        268     WRITE PROM      12938
```

This RMF report details summary cache statistics for one subsystem (00B1), detailing Cache I/O Requests & Non-Cache I/O requests. A H/R (Hit Rate) of 0.844 (84.4%) for Read I/O Requests is typical of a well performing DASD subsystem; a H/R of <80%, performance problems generally follow. A Write H/R of ~1.000 (100%) is typical. ICL (Inhibit Cache Load) is the number I/O requests that inhibited load of data into cache & data was not in cache. DFSMS Dynamic Cache Management & DB2 will selectively disable I/O operations for caching, using inhibit & bypass operation modes; hence, cache misses require thought!

Another solid detail report, but data aggregation & presentation requires…

DASD I/O Response Time: A Basic RMF View

D I R E C T   A C C E S S   D E V I C E   A C T I V I T Y

PAGE 2

z/OS V1R13 SYSTEM ID UIG1 DATE 09/28/2011 INTERVAL 14.59.998
RPT VERSION V1R13 RMF TIME 11.17.31 CYCLE 1.000 SECONDS

TOTAL SAMPLES = 900 IODF = 00 CR-DATE: 09/14/2011 CR-TIME: 10.31.31 ACT: ACTIVATE

STORAGE GROUP	DEV NUM	DEVICE TYPE	NUMBER OF CYL	VOLUME SERIAL	PAV	LCU	DEVICE ACTIVITY RATE	AVG RESP TIME	AVG IOSQ TIME	AVG CMR DLY	AVG DB DLY	AVG PEND TIME	AVG DISC TIME	AVG CONN TIME	% DEV CONN	% DEV UTIL	% DEV RESV	AVG NUMBER ALLOC	% ANY ALLOC	% MT PEND
	732A	33903	200	SYSXCA		002C	0.037	4.87	.000	.178	.000	.287	.000	4.58	0.02	0.02	0.0	0.0	100.0	0.0
	733A	33903	200	SYSAXA		002C	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
	733D	33903	200	SYSDXA		002C	1.354	1.86	.000	.133	.000	.274	.052	1.54	0.21	0.22	0.0	14.0	100.0	0.0
				LCU		002C	1.391	1.94	.000	.134	.000	.275	.051	1.62	0.07	0.08	0.0	14.0	100.0	0.0
	969F	33909	10017	SYSIMS	1	0040	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
	969F	33909	10017	SYSIMS	1	0040	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
	9691	33909	10017	SYSUSR	1	0040	1.707	.536	.000	.133	.037	.239	.000	.297	0.05	0.05	0.0	4.0	100.0	0.0
	9692	33909	10017	SYSOPE	1	0040	0.041	2.22	.000	.135	.000	.214	.000	2.01	0.01	0.01	0.0	3.0	100.0	0.0
	9693	33909	10017	SYSCDS	1	0040	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
	9694	33909	10017	SYSHM1	1	0040	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
SMS	9695	33909	10017	SYSSMS	1	0040	0.012	.593	.000	.163	.000	.256	.000	.337	0.00	0.00	0.0	1.0	100.0	0.0
SMS	9696	33909	10017	SYSST1	1	0040	0.001	1.79	.000	.512	.000	.640	.000	1.15	0.00	0.00	0.0	0.0	100.0	0.0
SMS	9697	33909	10017	SYSST2	1	0040	0.001	2.94	.000	1.41	.000	1.54	.000	1.41	0.00	0.00	0.0	0.0	100.0	0.0
	9699	33909	10017	SYSHM2	1	0040	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
DB2	969A	33909	10017	SYSSD1	1	0040	0.001	1.66	.000	.384	.000	.512	.000	1.15	0.00	0.00	0.0	0.0	100.0	0.0
DB2	969B	33909	10017	SYSSD2	1	0040	0.001	1.02	.000	.256	.000	.384	.000	.640	0.00	0.00	0.0	0.0	100.0	0.0
DB2	969C	33909	10017	SYSSD3	1	0040	0.001	2.05	.000	.640	.000	.768	.000	1.28	0.00	0.00	0.0	0.0	100.0	0.0
	969E	33909	10017	SYSDB2	1	0040	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
	969F	33909	10017	SYSIMS	1	0040	0.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	0.0
				LCU		0040	3.047	1.05	.365	.143	.036	.259	.002	.428	0.01	0.01	0.0	46.0	100.0	0.0

A restriction with RMF is that reports are for one z/OS system only, therefore z/OS system aggregation is required for complete LCU & Device activity, delivering a full view of overall disk subsystem performance. 15 Minute RMF intervals means 96 observations per day, for n systems & nnnn devices (n million records)!

RMF data collection is brilliant, aggregation & presentation requires…

DASD I/O Response Time: A Shared RMF View

```
                    S H A R E D   D I R E C T   A C C E S S   D E V I C E   A C T I V I T Y
                                                                                                    PAGE    1
      z/OS V1R13            SYSPLEX SYSDPLEX           DATE 09/28/2011       INTERVAL 14.59.997
                            RPT VERSION V1R13 RMF      TIME 08.30.00         CYCLE  1.000  SECONDS

   TOTAL SAMPLES(AVG) =   900.0   (MAX) =   900.0   (MIN) =   900.0

                         SMF            DEVICE   AVG   AVG   AVG   AVG   AVG   AVG   AVG    %     %     %     AVG
DEV  DEVICE    VOLUME PAV SYS  IODF LCU ACTIVITY RESP  IOSQ  CMR   DB    PEND  DISC  CONN   DEV   DEV   DEV   NUMBER
NUM  TYPE      SERIAL     ID   SUFF     RATE     TIME  TIME  DLY   DLY   TIME  TIME  TIME   CONN  UTIL  RESV  ALLOC

2000 33909     MVSLIB     *ALL              0.138 0.327 0.000 0.001 0.000 0.140 0.000 0.187  0.00  0.00   0.0  43.0
                      1.0H SYSD  00   002B  0.003 0.256 0.000 0.000 0.000 0.128 0.000 0.128  0.00  0.00   0.0  25.0
                         1 SYSE  00   0029  0.134 0.329 0.000 0.001 0.000 0.141 0.000 0.188  0.00  0.00   0.0  18.0

2001 33909     MVSTGT     *ALL              0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00   0.0   156
                      1.0H SYSD  00   002B  0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00   0.0  72.0
                         1 SYSE  00   0029  0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00   0.0  84.0
```

This RMF report provides a baseline level of information for each system sharing the disk device, presuming the Device Number is the same for all systems. In this instance for the MVSTGT VOLSER, the Device Number is different between z/OS systems, hence limited data is available. *Consolidate, Deduplicate & Enhance RMF Data via Analytics – Avoid Too Much & Too Little Data, Not Enough Facts!*

RMF data collection is brilliant, aggregation & presentation requires…

DASD I/O Problem Reports: An RMF Flow #1

```
                       RMF V1R13   System Information                  Line 1 of 27
Command ===>                                                       Scroll ===> HALF

Samples: 100      System: PRD1   Date: 09/28/11   Time: 10.32.00   Range: 100   Sec

Partition:     MVS1      2094 Model 716            Appl%:        63   Policy: STANDARD
CPs Online:       4      Avg CPU Util%:  73        EAppl%:       65   Date:    09/28/11
AAPs Online:      0      Avg MVS Util%:  84        Appl% AAP:     0   Time:    14.05.07
IIPs Online:      0                                Appl% IIP:     0

Group       T WFL  --Users--  RESP TRANS -AVG USG-   -Average Number Delayed For -
                %   TOT  ACT  Time  /SEC PROC  DEV   PROC  DEV STOR SUBS OPER  ENQ

*SYSTEM        34   664   26       13.95  5.1  5.0    1.9  4.1  7.0  2.6  2.0  2.0
*TSO           50   534    8       13.95  2.6  2.1    0.4  1.5  2.0  0.8  0.0  0.0
*BATCH         26    11   10        0.00  1.5  1.4    1.4  1.7  0.5  1.8  1.0  2.0
*STC           27   115    8        0.00  1.1  1.5    0.1  1.0  4.5  0.1  1.0  0.0
*ASCH                 3    0        0.00  0.0  0.0    0.0  0.0  0.0  0.0  0.0  0.0
*OMVS                 2    0        0.00  0.0  0.0    0.0  0.0  0.0  0.0  0.0  0.0
*ENCLAVE        5     4  N/A         N/A  0.2  N/A    3.7  N/A  0.0  N/A  N/A  N/A
PRIMEBAT    W  26    11   10  46.0  0.06  1.5  1.4    1.4  1.7  0.5  1.8  1.0  2.0
NRPRIME     S  26    11   10  46.0  0.06  1.5  1.4    1.4  1.7  0.5  1.8  1.0  2.0
            1  23     9    9  27.9  0.06  0.9  1.4    0.8  1.6  0.5  1.8  1.0  2.0
            2  29     0    0  54.2  0.02  0.1  0.0    0.2  0.0  0.0  0.0  0.0  0.0
            3  59     1    1  .000  0.00  0.6  0.0    0.4  0.0  0.0  0.0  0.0  0.0
PRIMETSO    W  50   527    8  .759 13.98  2.6  2.1    0.4  1.5  2.0  0.8  0.0  0.0
TSOPRIME    S  50   527    8  .759 13.98  2.6  2.1    0.4  1.5  2.0  0.8  0.0  0.0
            1  48   526    8  .403 13.98  2.1  1.9    0.3  1.3  2.0  0.8  0.0  0.0
            2  75     1    1  30.6  0.08  0.3  0.2    0.1  0.1  0.0  0.0  0.0  0.0
            3  75     0    0   126  0.02  0.1  0.0    0.0  0.0  0.0  0.0  0.0  0.0
```

The *RMF Monitor III System Information (SYSINFO)* report shows whether there is any device delay. It also measures the size of the problem by the number of users affected. A subjective viewpoint of whether there is a problem, does it impact the service enough an should we do something about it should be considered. In this case, further investigation into the *SYSTEM, NPRIME and TSOPRIME groups follows.

RMF reports provide you with a high-level system wide viewpoint…

DASD I/O Problem Reports: An RMF Flow #2

```
                         RMF V1R13   Delay Report                       Line 21 of 49
Command ===>                                                        Scroll ===> HALF

Samples: 100     System: PRD1  Date: 04/07/11  Time: 10.32.00  Range: 100   Sec

            Service      WFL USG DLY IDL UKN ----- % Delayed for -----  Primary
Name     CX Class    Cr  %   %   %   %   %   PRC  DEV STR SUB OPR  ENQ  Reason

MISTYDFS B  NRPRIME        0   0 100   0   0   0    0   0   0 100    0  Message
BHOLEQB  B  NRPRIME        0   0  49   0   0   0    0   0   0   0   49  SYSDSN
BHOLWTO1 B  NRPRIME        0   0  33   0   1   0    0   0   0  33    0  Message
BRHI     T  TSOPRIME       0   0   4   0   0   0    0   4   0   0    0  COMM
CONSOLE  S  SYSSTC         0   0   3   0  97   0    0   3   0   0    0  LOCL
VTMLCL   S  SYSSTC         0   0   2   0  98   0    0   2   0   0    0  LOCL
STARTING T  TSOPRIME       0   0   2   0   7   0    0   2   0   0    0  COMM
CATALOG  S  SYSSTC         0   0   1   0  99   0    0   1   0   0    0  LOCL
INIT     S  SYSSTC         0   0   1  49   0   0    0   1   0   0    0  LOCL
BHOLDEV3 B  NRPRIME       20  20  80   0   0   1   79   0   0   0    0  BBVOL1
BTEUDASD B  NRPRIME       22  22  77   0   1   1   76   0   0   0    0  BBVOL1
BAJU     T  TSOPRIME      24  24  75   0   1   0   75   0   0   0    0  BBVOL1
KLSPRINT B  NRPRIME       25  25  75   0   0   0   75   0   0   0    0  BBVOL1
BHIM     T  TSOPRIME      31  11  25  64   0   2   18   5   0   0    0  BBVOL1
BHOLPRO1 B  NRPRIME       52  13  12   0   0  12    0   0   0   0    0  BHOLPRO2
```

The *RMF Monitor III Delay Report (DELAY)* report identifies the address space (Name) and associated Service Class for potential delays, with an associated Primary Reason. In this instance the NRPRIME and TSOPRIME Service Class resources have significant delays, associated with BBVOL1. In this instance, we will delve deeper into the BBVOL1 resource, most likely a DASD volume…

RMF reports help you make informed decisions & where to look next…

DASD I/O Problem Reports: An RMF Flow #3

```
                         RMF V1R13  Device Delays                    Line 1 of 9
Command ===>                                                    Scroll ===> HALF

Samples: 100     System: PRD1  Date: 04/07/11  Time: 10.32.00  Range: 100   Sec

                        DLY USG CON  ------------ Main Delay Volume(s) ------------
Jobname  C Class         %   %   %     %  VOLSER    %  VOLSER    %  VOLSER    %  VOLSER

BHOLDEV3 B NRPRIME  79  15   4    72 BBVOL1
BTEUDASD B NRPRIME  76  20   4    75 BBVOL1
KLSPRINT B NRPRIME  75  23   4    75 BBVOL1
BAJU     T TSOPRIME 75  19   1    71 BBVOL1
BHIM     T TSOPRIME 18   6   1    16 BBVOL1     1 430D13
BHOL     T TSOPRIME  4  11   6     9 BBVOL1
JES2     S SYSSTC    4   3   0     6 SYSPAG
BHOLNAM4 B NRPRIME   1   0   0     1 BBVOL1
```

The *RMF Monitor III Device Delays Report (DEV)* report identifies the address space (Jobname) and associated Service Class (Class) delayed due to device contention. In this instance we can see that DASD VOLSER BBVOL1 is the Main Delay Volume for several address spaces associated with the NPRIME and TSOPRIME service classes. Up to 4 DASD volumes causing delay can be identified, in this instance DASD VOLSER 430D13 is also listed, but its delay of 1% is inconsequential. We can also determine that when an I/O transfer is actually connected (CON %), DASD VOLSER BBVOL1 isn't very busy, when finally connected. In this instance, with 4 delay observations exceeding 70%, DASD VOLSER BBVOL1 will be further analysed.

RMF reports help you target the most likely cause of I/O operation issues

DASD I/O Problem Reports: An RMF Flow #4

```
                    RMF V1R13   Device Resource Delays            Line 1 of 16
Command ===>                                                      Scroll ===> HALF

Samples: 100     System: PRD1   Date: 04/07/11   Time: 10.32.00   Range: 100   Sec
```

Volume /Num	S/ PAV	Act Rate	Resp Time	ACT %	CON %	DSC %	PND %, Reasons	DEV/CU Type	Jobname	C	Service Class	USG %	DLY %
BBVOL1		41.8	.099	88	21	66	PND 1	3380K	KLSPRINT	B	NRPRIME	6	92
0222								3990-2	BHOLDEV3	B	NRPRIME	2	92
									BTEUDASD	B	NRPRIME	5	91
									BAJU	T	TSOPRIME	3	91
									BHIM	T	NRPRIME	1	22
									BHOL	T	TSOPRIME	3	10
									RMFGAT	S	SYSSTC	0	1
									BHOLNAM4	B	NRPRIME	0	1
									BHOLPRO1	B	NRPRIME	0	1
SYSPAG		10.9	.028	19	6	13	PND 0	3380K	BVAU	T	TSOPRIME	6	14
0225								3990-2	JES2	S	SYSSTC	0	7
430D13		1.1	.032	4	2	2	PND 0	3380K	BENK	T	TSOPRIME	0	1
0220								3990-2	BHIM	T	TSOPRIME	0	1
									BHOL	T	TSOPRIME	2	0
SYSCAT		0.3	.017	0	0	0	PND 0	3380K	INIT	S	SYSSTC	0	1
0221								3990-2					

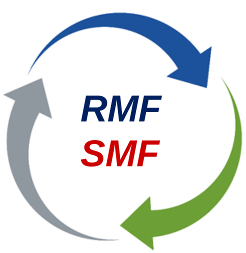
The *RMF Monitor III Device Resource Delays Report (DEVR)* report identifies all of the address spaces (Jobname) being delayed by a DASD VOLSER (Volume). In this instance, we first identified several address spaces being delayed, then the associated DASD VOLSER and DEV/CU details. Going back to S/370 I/O basics, a faster CU (3990-3) and associated DEV (3380K) would most likely resolve this issue.

*Can we really use this process for a 24*7 operation with nnnn VOLSERs?*

Advanced DASD Functions & 3rd Party IHV's

IBM introduce advanced DASD functions for their System Storage DS8000 series disk subsystem, typically as Licenced Machine Code (LMC) features with full z/OS and System z server interoperability, while these functions are captured in the associated SMF (RMF) records. Therefore RMF instrumentation data evolves accordingly, providing the requisite data to measure DASD I/O performance. The latest DS8880 advanced DASD functions include Adaptive Multi-stream Prefetching (AMP), Easy Tier Integration, FICON Express 16S, High Performance FICON for z (zHPF), HyperPAV, Intelligent Write Caching (IWC), List Prefetch Optimizer, Multiple Allegiance, Modified Indirect Data Address Word (MIDAW), Sequential Prefetching in Adaptive Replacement Cache (SARC), zHyperWrite, et al. *Therefore the DASD performance analyst, needs to fully understand these current functions & whatever follows! RMF provides limited data for Replication (GDPS/PM SMF Type 105), FlashCopy & EasyTier…*

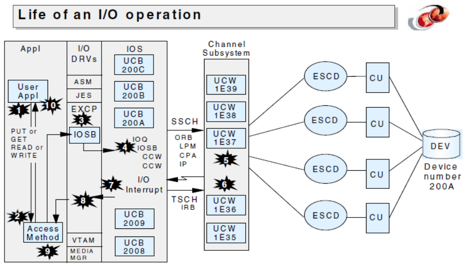
EMC (E.g. VMAX) and HDS (E.g. VSP G1000) attempt to fully emulate all major DASD functions (E.g. zHPF, HyperPAV, FICON Express, MIDAW), with Channel Subsystem (CSS) and Operating System (I.E. z/OS) interoperability, capturing any relevant data in associated RMF records. *Such function emulation will be subject to Product Roadmap development time. Each IHV will have their own proprietary internal hardware (E.g. Cache) algorithms, which might impact RMF performance monitoring. Inevitably, other measurement data (SMF) is required.*



EMC VMAX creates SMF data for SRDF activity, HDS VSP G1000 creates SMF data via an Analytics Recorder & there is even an SMF record for DS8000 GDPS Global Mirror session statistics. Observing the 80/20 rule, RMF delivers the majority of data for minimal effort; for a full I/O activity view, other SMF data is required. *Adopt a standard disk I/O measurement process, comparing IHV disk performance, as & if required.*

Change your IHV; don't modify, just evolve, a disk I/O monitoring process!

Myth: The Demise Of The Disk I/O Tuner?



With an open architecture, plus a full understanding of the engineering, a human being can easily apply tweaks & repairs for a perfectly tuned system!

For decades, Baby Boomers evolved I/O tuning techniques, knowing the basics & advanced functions, adapting these processes for 21st Century workloads.

With Licensed Machine Code, plus proprietary algorithms & engineering, only the IHV with electronic diagnosis can truly tune this “closed” system…

Millennials require meaningful concise information from big data, allowing for informed business decisions. Micro managing disk controllers is not an option!

Don’t rely upon the self-tuning promise, become the performance expert!

Collaboration: Working With Your Disk IHV…

The Data In Memory I/O: Utilize 5% Solid State Drives (SSD) & deliver 10 times performance improvement. Allocate 10% SSD, to manage 90% of the I/O workload. Storage resources are becoming intelligent/self-aware, analysing data usage, making decisions on where hot/active & cold/inactive data should be stored.

The Big Data Era: The 21st Century data centre generates increasing amounts of data, typically random access, from mobile, cloud & social platforms, requiring fast & agile processing. All-Flash or SSD disk arrays will someday overtake traditional disk for primary storage workloads from a price/performance perspective.

The IHV: Saint or Sinner?: As per the System z server with the On Off Capacity on Demand (OOCoD) capability, an ability to activate additional disk electronic memory for capacity or component failure is arguably mandatory. As most disk I/O performance issues will become disk subsystem related, working closely with your disk IHV is advisable, where both parties collaborate to deliver the best price/performance balance for business data!

Proactively sharing performance data will develop a win-win relationship…

DASD I/O Performance: Recommendations

The RMF product is cost efficient, with a similar cost to SDSF, delivering invaluable resource measurement metrics for the application enabling CPU & disk I/O hardware. SMF records generated by IHV DASD subsystems add further value to the baseline RMF data available for analysis. *Recommendation: RMF (CMF) data collection is cost-efficient & mandatory.*

Storage is no longer the party wallflower, it’s the party animal! Data and the value derived from this unique resource differentiates one business from another. Safeguarding data availability & agile performance, is just as important as providing the requisite CPU capacity to process this data. *Recommendation: Ensure you have a robust disk I/O performance process.*

Ultimately RMF/SMF data is just unformatted resource metrics. Intelligent data reduction (deduplication, consolidation), supplemented by analytics (categorisation, summarisation) & experience based (I/O tuning techniques) enrichment is required, delivering meaningful information from massive data. *Recommendation: Systems Management software simplifies this process.*

The CPU R4HA is a high priority focus for all System z users. A well-tuned I/O subsystem can deliver cost savings, while optimizing performance. Tools such as IntelliMagic Vision (Disk Magic) & Technical Storage Control Centre (EADM) can assist organizations in optimizing the disk I/O resource. *Recommendation: Disk I/O & CPU R4HA metrics should be considered equally.*

Proactive DASD I/O performance analysis delivers business benefit

Case Study 1: z13 Server Memory Upgrade

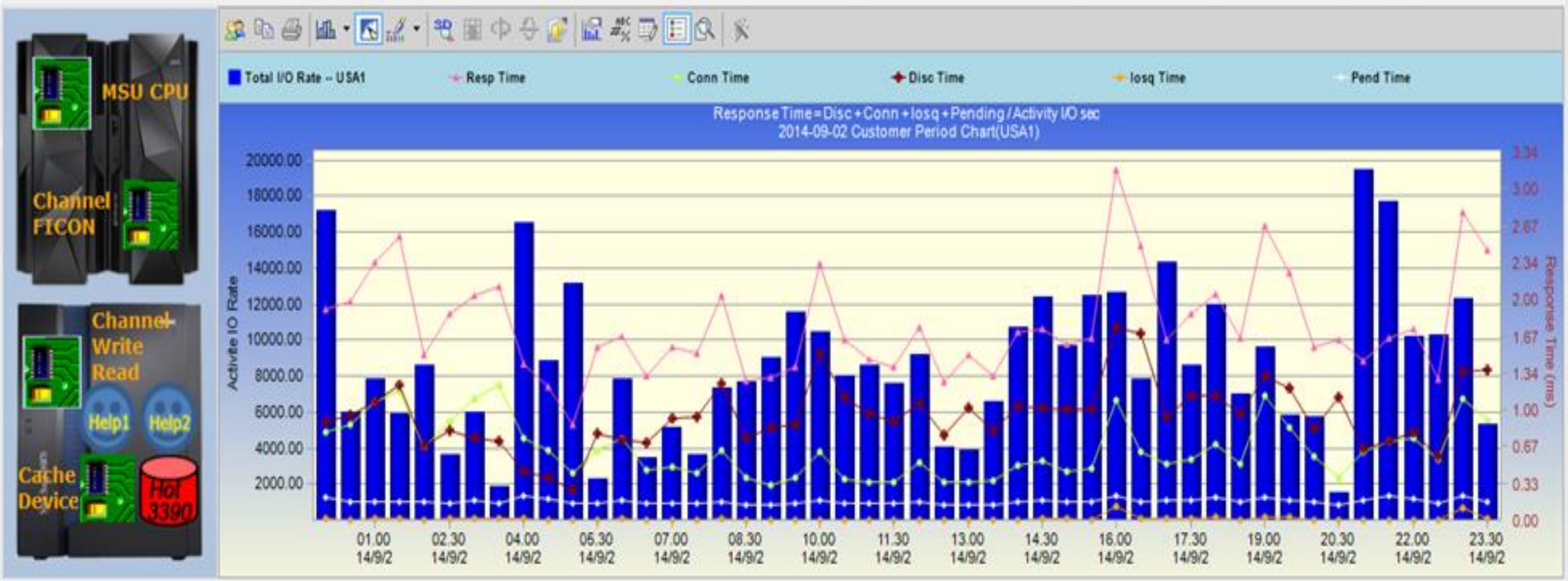
z13 Server & z/OS 2.1 Operating System
EMC VMAX (Symmetrix) DASD Subsystems
DB2 11 Data Sharing Environment
z13 Server Memory Upgrade: 25/06/2016

Customer Expectation: Higher throughput & performance for DB2

Case Study 1: z13 Server Memory Upgrade #1

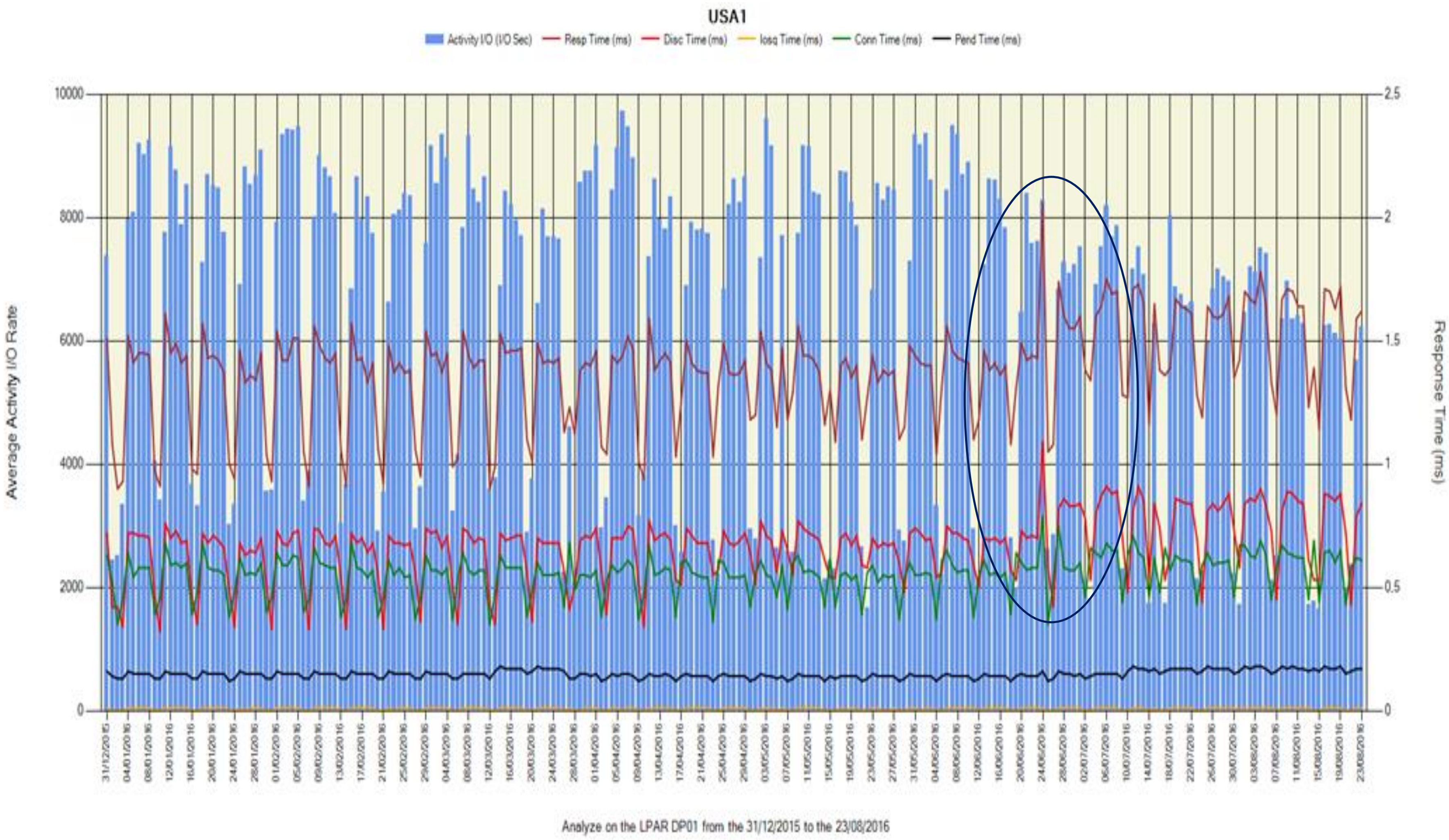
DASD I/O | LCU I/O | Summary | Alert | Error | Volume by period | Small Dcollect | Big Dcollect | Vol/Space Used | CPU I/O | SSD Tracks | SSD R/W | Back end/Front end

Date	Time	Period	LPAR	I/O	Resp	Conn	Disc	Iosq	Pend	QInt	Busy%	Vol used	Used Capa(Go)	Vol Total	Capa(Go)	Access...	Access...
9/2/2014...	00.00	Batch	USA1	17215.63	1.92	0.80	0.89	0.02	0.20	19193.87	33.53	381	9057	449	9954	45.19	1.90
9/2/2014...	00.30	Batch	USA1	6012.03	1.98	0.87	0.95	0.00	0.16	6707.61	11.49	285	6648	449	9954	21.09	0.90
9/2/2014...	01.00	Batch	USA1	7809.65	2.33	1.08	1.08	0.01	0.17	9787.28	13.72	282	6663	449	9954	27.69	1.17
9/2/2014...	01.30	Batch	USA1	5918.79	2.57	1.17	1.22	0.01	0.17	8288.31	12.44	352	8475	449	9954	16.81	0.70
9/2/2014...	02.00	Batch	USA1	8566.88	1.50	0.66	0.67	0.00	0.16	7160.32	14.19	357	8580	449	9954	24.00	1.00
9/2/2014...	02.30	Batch	USA1	3646.63	1.87	0.89	0.81	0.01	0.15	3555.41	7.49	354	8594	449	9954	10.30	0.42
9/2/2014...	03.00	Batch	USA1	5997.52	2.03	1.10	0.74	0.01	0.18	5605.57	9.30	356	8600	449	9954	16.85	0.70



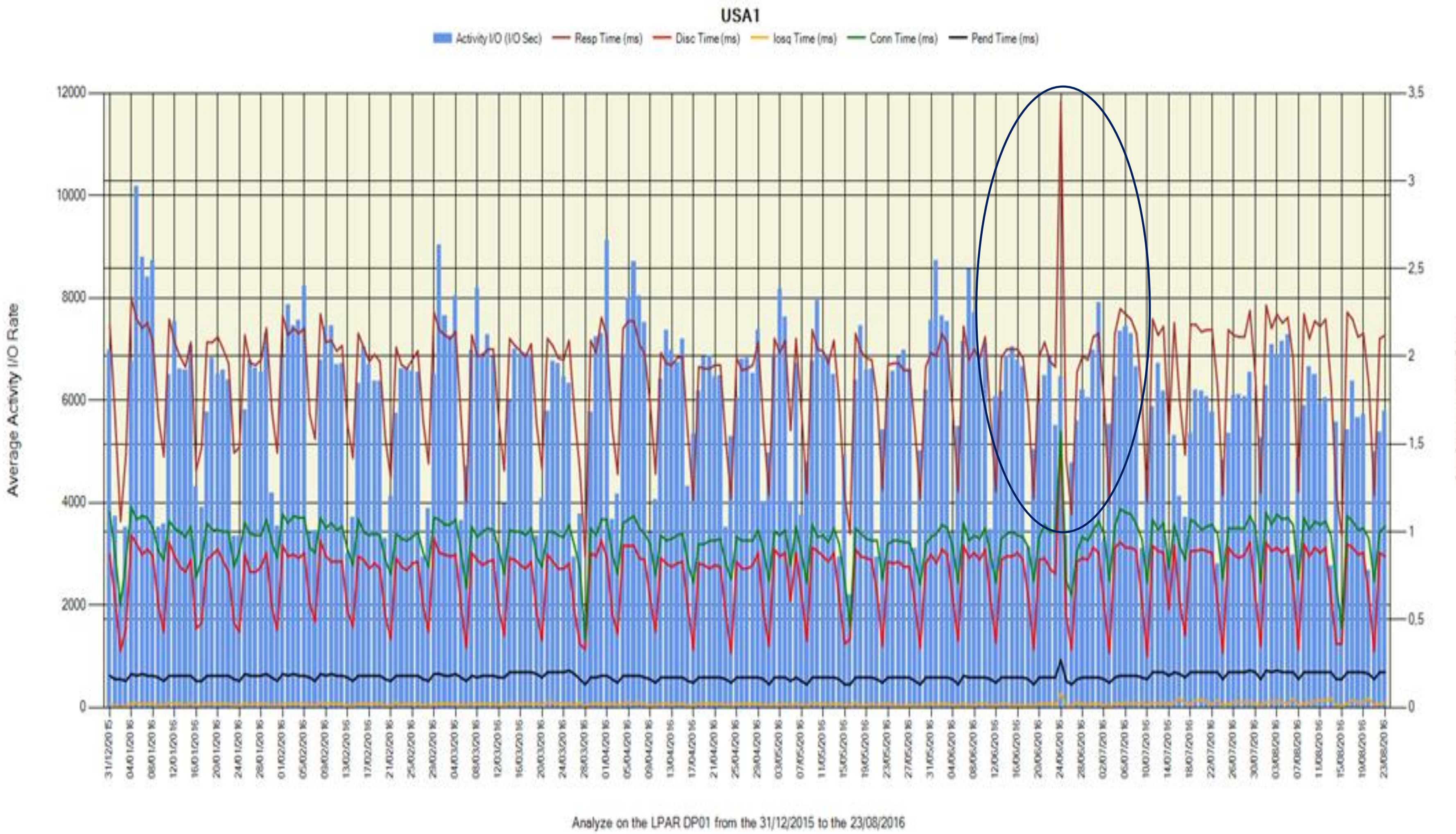
In this example DISC time is of interest, when reviewing the RESP time

Case Study 1: z13 Server Memory Upgrade #2



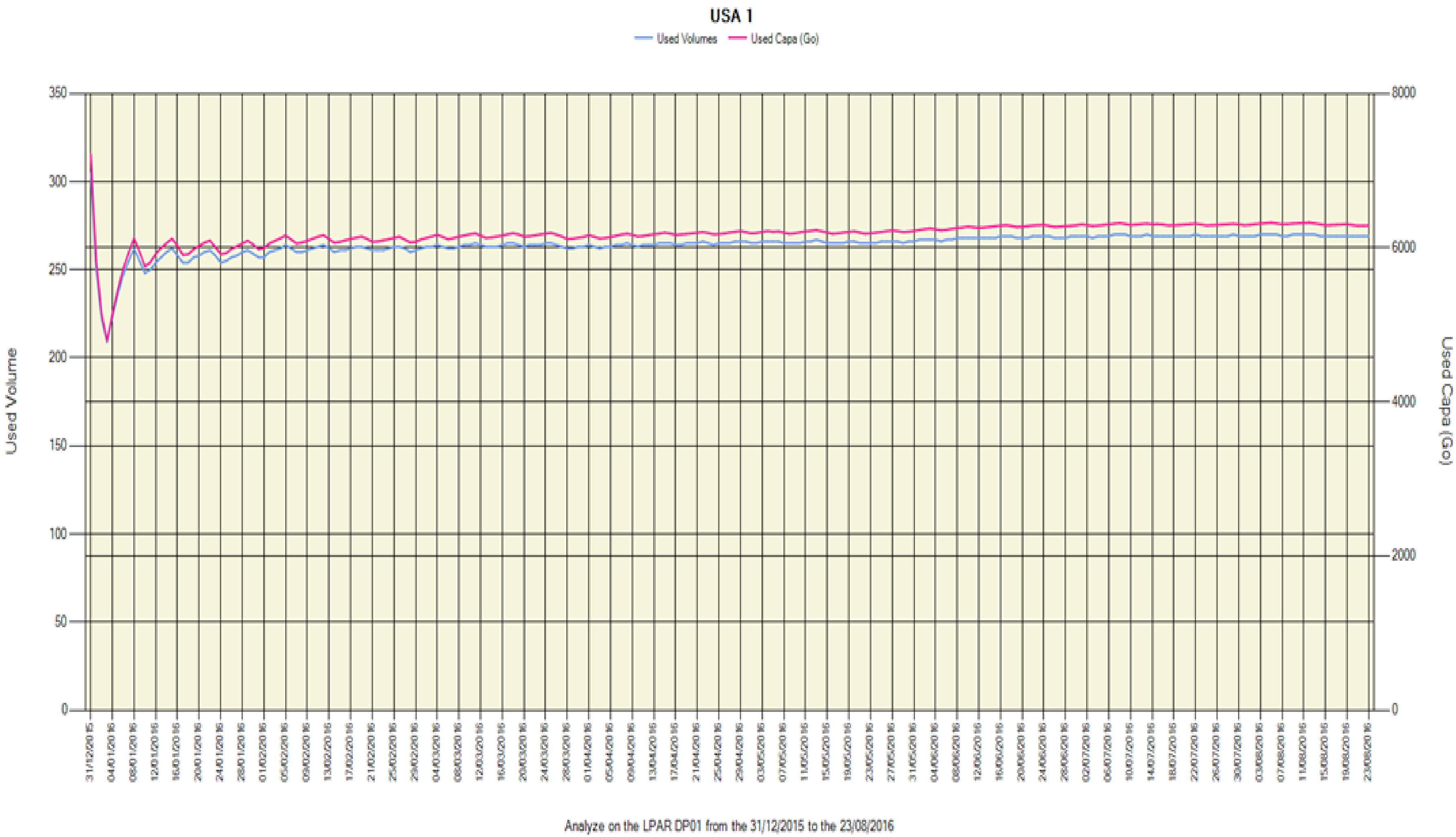
TP: I/O decrease due to DB2 buffer pool, DISC and RESP increase

Case Study 1: z13 Server Memory Upgrade #3



Batch: An anomaly on the upgrade day, otherwise the I/O profile is normal

Case Study 1: z13 Server Memory Upgrade #4

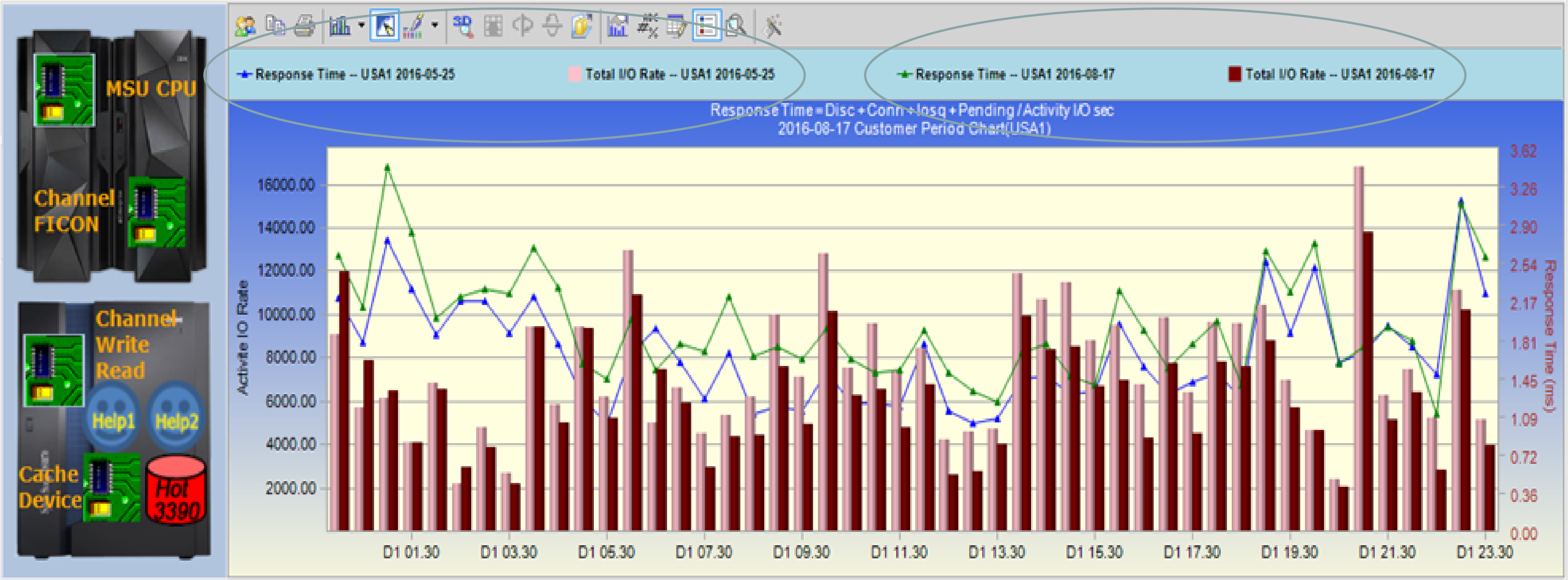


Capacity: The used capacity of 3390 disk has the expected slight growth…

Case Study 1: z13 Server Memory Upgrade #5

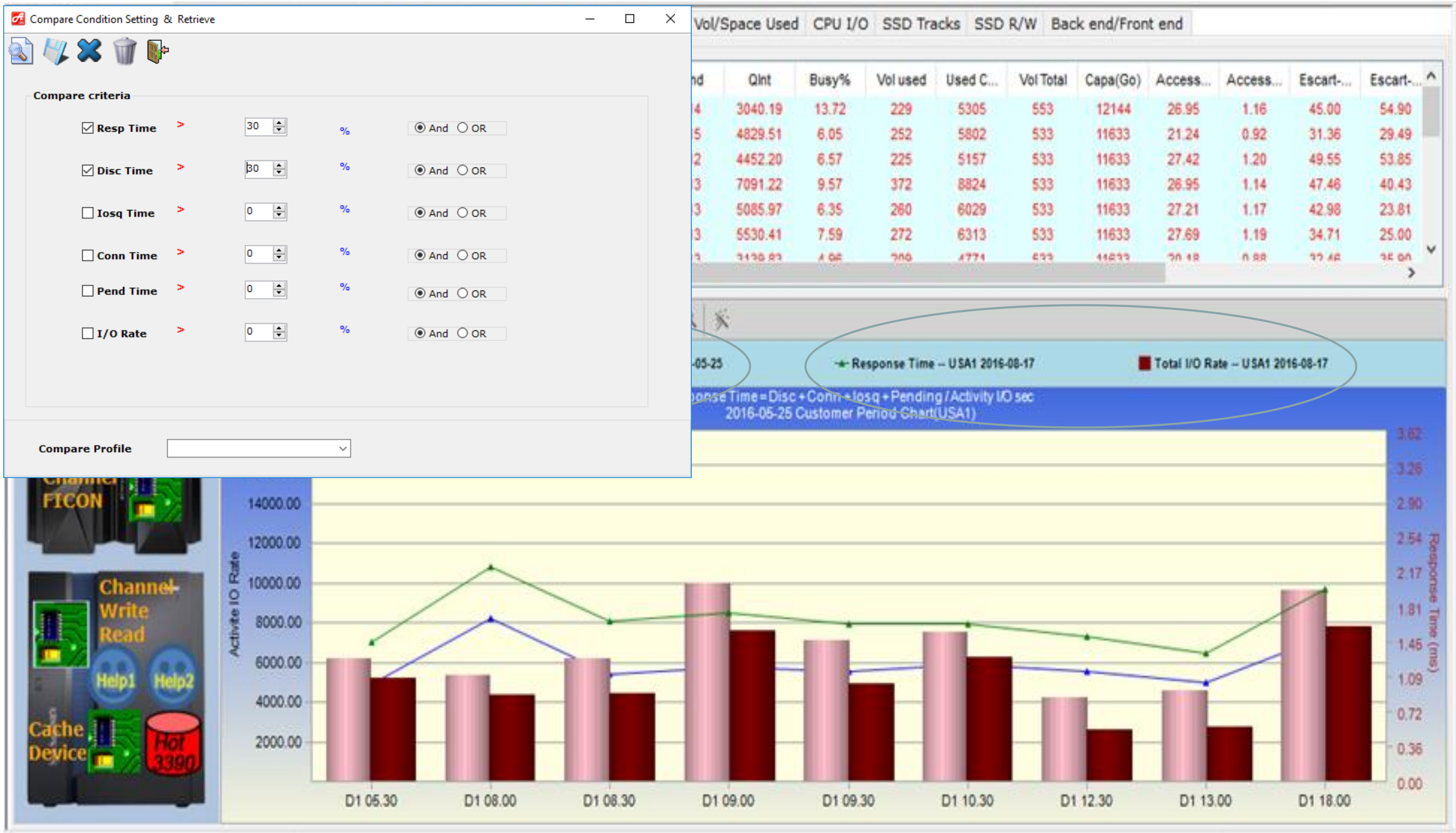
DASD I/O | LCU I/O | Summary | Alert | Error | Volume by period | Small Dcollect | Big Dcollect | Vol/Space Used | CPU I/O | SSD Tracks | SSD R/W | Back end/Front end

Date	Time	Period	LPAR	I/O	Resp	Conn	Disc	Iosq	Pend	QInt	Busy%	Vol used	Used C...	Vol Total	Capa(Go)	Access...	Access...
8/17/201...	00.00	Batch	USA1	11942.80	2.61	1.26	1.10	0.04	0.21	16105.45	19.15	437	10485	549	11990	27.33	1.14
8/17/201...	00.30	Batch	USA1	7849.40	2.12	1.15	0.78	0.01	0.19	7651.14	10.69	343	8196	549	11990	22.88	0.96
8/17/201...	01.00	Batch	USA1	6507.43	3.45	1.52	1.19	0.47	0.27	12583.37	21.56	314	7341	549	11990	20.72	0.89
8/17/201...	01.30	Batch	USA1	4109.92	2.83	1.30	1.34	0.00	0.18	6278.41	9.15	224	5132	549	11990	18.35	0.80
8/17/201...	02.00	Batch	USA1	6530.83	2.02	1.02	0.79	0.01	0.20	6520.04	9.89	411	9957	569	12502	15.89	0.66
8/17/201...	02.30	Batch	USA1	2946.96	2.22	1.29	0.66	0.07	0.20	2719.02	4.61	256	5950	569	12502	11.51	0.50
8/17/201...	03.00	Batch	USA1	3876.73	2.29	1.41	0.67	0.01	0.21	3434.11	6.82	244	5711	567	12450	15.89	0.68



I/O Service: Comparing before vs. after upgrade, I/O service is worse…

Case Study 1: z13 Server Memory Upgrade #6



I/O Service: Post upgrade service time issues via RESP/DISC (30) criteria

Case Study 1: z13 Server Memory Upgrade #7

Date	Time	LPAR	SG	Vo...	Type	Dua	L...	I/O	Resp	Disc	Conn	Iosq	Pend	AVG ...	AVG ...	AVG I...
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	5.14	3.13	2.46	0.54	0.00	0.13	0.020	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	10.78	2.98	2.43	0.44	0.00	0.12	0.011	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.44	3.55	2.37	1.03	0.00	0.15	0.040	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.79	2.97	2.36	0.47	0.00	0.15	0.034	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.19	3.84	2.36	1.33	0.00	0.16	0.046	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.42	2.97	2.35	0.46	0.00	0.17	0.067	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.34	2.83	2.04	0.64	0.00	0.15	0.044	0.000	0.000
5/25/2...	10:30	USA1		U...	33903*3,339	2...	0...	4.57	4.70	2.03	2.50	0.00	0.17	0.042	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.42	3.47	1.91	1.42	0.00	0.14	0.035	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	4.01	2.71	1.80	0.76	0.00	0.15	0.036	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	2.22	2.41	1.60	0.65	0.00	0.16	0.046	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	12.55	2.12	1.54	0.46	0.00	0.12	0.014	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.49	2.23	1.50	0.58	0.00	0.16	0.047	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	1.06	2.21	1.45	0.60	0.00	0.17	0.055	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	5.07	4.33	1.36	2.81	0.00	0.16	0.037	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	39.63	2.52	1.23	1.15	0.00	0.14	0.030	0.000	0.000
5/25/2...	10:30	USA1		U...	33903*3,339	2...	0...	1.16	1.87	1.21	0.54	0.00	0.12	0.017	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	51.69	1.64	1.17	0.35	0.00	0.12	0.009	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	12.46	2.60	1.17	1.26	0.00	0.17	0.053	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	2.83	3.00	1.17	1.63	0.00	0.20	0.088	0.000	0.000
5/25/2...	10:30	USA1		S...	33903*3,339	2...	0...	2.60	1.81	1.14	0.51	0.00	0.17	0.047	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	37.21	1.79	1.11	0.56	0.00	0.12	0.014	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	22.90	1.59	1.08	0.38	0.00	0.12	0.017	0.000	0.000
5/25/2...	10:30	USA1	D...	D...	33909*32.8K	2...	0...	31.46	1.50	1.07	0.30	0.00	0.12	0.009	0.000	0.000

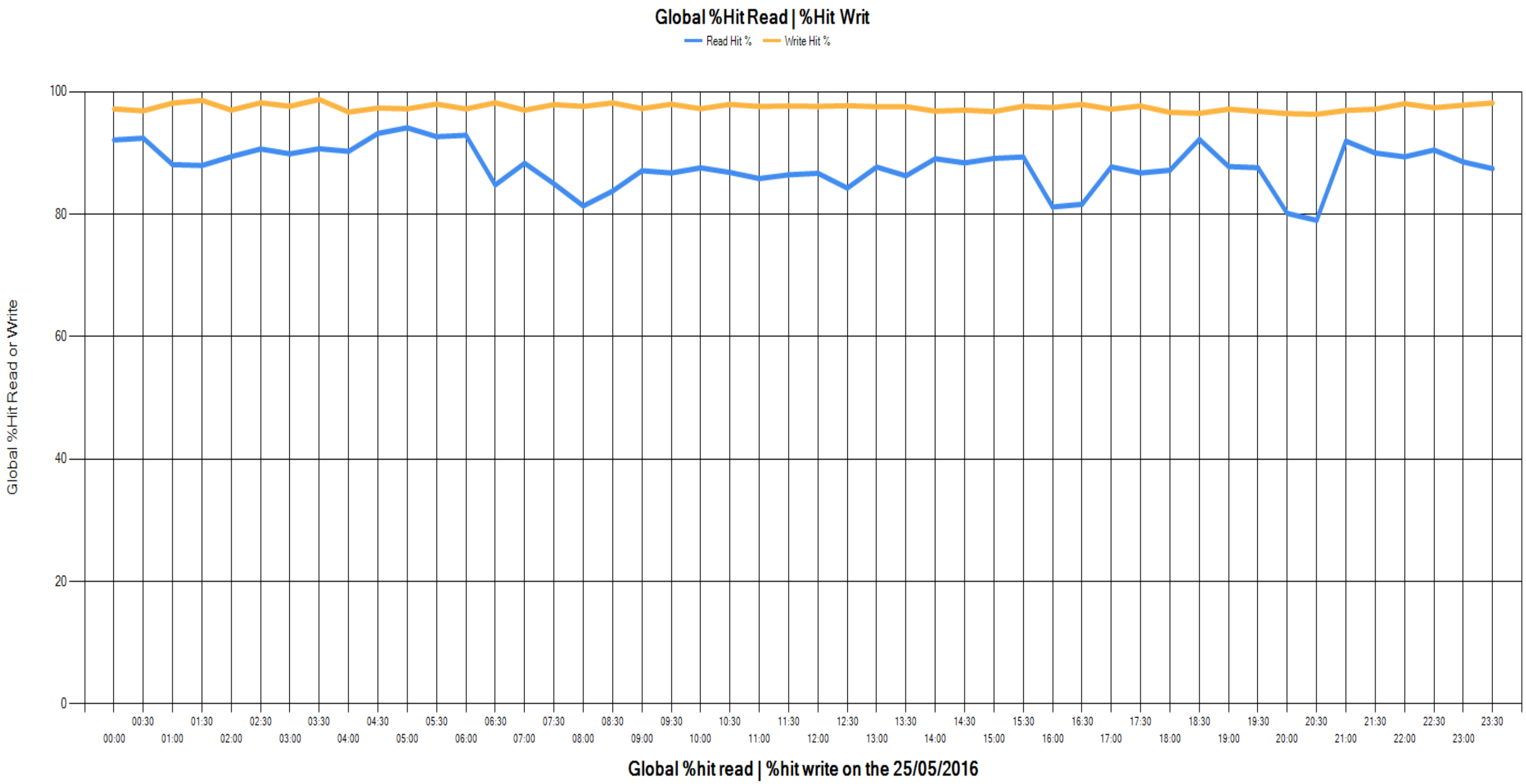
I/O Service: Site specific TP time of 10:30 (normalized peak) pre upgrade

Case Study 1: z13 Server Memory Upgrade #8

Date	Time	LPAR	SG	V...	Type	D...	LCU	I/O	Resp	Disc	Disc	Iosq	Pend	AVG ...	AVG ...	AVG I...	QInt
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.75	7.88	6.88	2.46	0.00	0.20	0.082	0.000	0.000	12.40
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.15	7.73	6.54	2.43	0.00	0.20	0.091	0.000	0.000	7.74
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.60	5.09	4.26	2.37	0.00	0.20	0.085	0.000	0.000	7.13
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.39	3.55	2.86	2.36	0.00	0.18	0.066	0.000	0.000	4.21
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	3.22	3.41	2.63	2.36	0.00	0.17	0.047	0.000	0.000	8.98
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.23	3.03	2.37	2.35	0.00	0.19	0.074	0.000	0.000	3.15
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.18	2.82	2.27	2.04	0.00	0.18	0.063	0.000	0.000	2.88
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.08	2.73	2.16	2.03	0.00	0.18	0.063	0.000	0.000	2.52
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.20	2.72	2.10	1.91	0.00	0.18	0.062	0.000	0.000	2.72
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	7.94	8.78	2.07	1.80	0.00	0.17	0.055	0.000	0.000	17.84
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.25	2.57	2.02	1.60	0.00	0.19	0.070	0.000	0.000	2.76
8/17/2...	10:30	USA1		U...	33903...	2...	0...	4.83	4.86	1.99	1.54	0.00	0.20	0.070	0.000	0.000	10.57
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	10.20	2.69	1.93	1.50	0.00	0.17	0.054	0.000	0.000	21.46
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	33.51	2.96	1.88	1.45	0.00	0.16	0.039	0.000	0.000	68.32
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.28	2.65	1.88	1.36	0.00	0.19	0.066	0.000	0.000	2.63
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	3.52	2.49	1.86	1.23	0.00	0.15	0.029	0.000	0.000	7.05
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	31.65	2.87	1.83	1.21	0.00	0.16	0.035	0.000	0.000	62.92
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.29	2.31	1.72	1.17	0.00	0.18	0.059	0.000	0.000	2.44
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	1.63	2.28	1.68	1.17	0.00	0.17	0.045	0.000	0.000	3.02
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	12.05	2.36	1.65	1.17	0.00	0.17	0.044	0.000	0.000	21.84
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	31.56	2.10	1.60	1.14	0.00	0.15	0.022	0.000	0.000	55.30
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	5.04	4.67	1.60	1.11	0.00	0.19	0.053	0.000	0.000	9.01
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	24.94	2.10	1.55	1.08	0.00	0.15	0.028	0.000	0.000	42.52
8/17/2...	10:30	USA1	D...	D...	33909...	2...	0...	37.24	2.06	1.54	1.07	0.00	0.15	0.027	0.000	0.000	62.82

I/O Service: Site specific TP time of 10:30 (normalized peak) post upgrade

Case Study 1: z13 Server Memory Upgrade #9

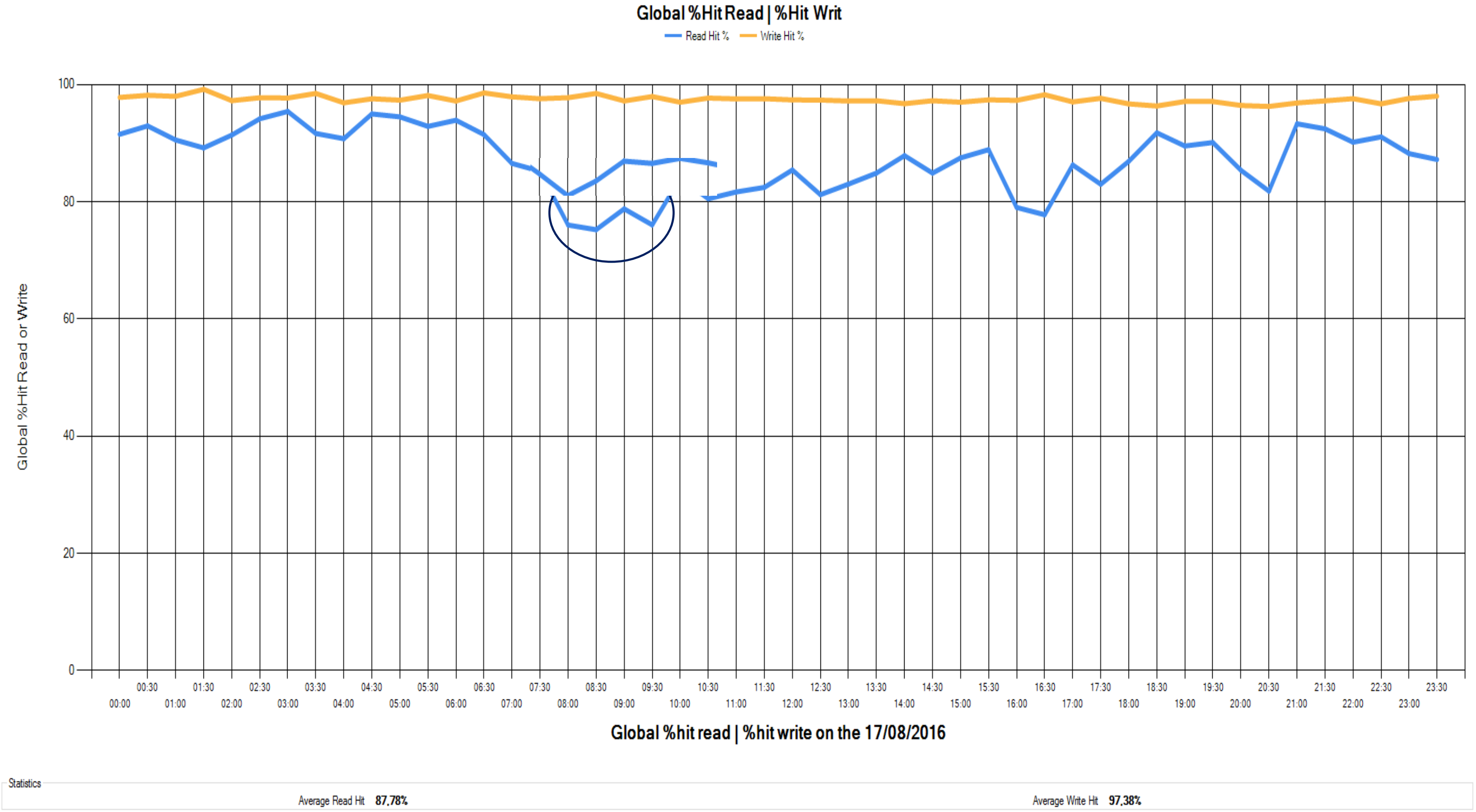


Statistics

Average Read Hit **88,27%** Average Write Hit **97,33%**

I/O Service: 80%+ workload cache read hit metric is the low watermark?

Case Study 1: z13 Server Memory Upgrade #10



I/O Service: A sustained cache read hit <80% severely degrades service!

Case Study 1: z13 Server Memory Upgrade

z13 Server & z/OS 2.1 Operating System
EMC VMAX (Symmetrix) DASD Subsystems
DB2 11 Data Sharing Environment
z13 Server Memory Upgrade: 25/06/2016

Increased DB2 Server Memory Throughput
Decreased Disk Subsystem I/O Throughput
Higher Disconnect - Lower Cache Usage
Investigate DB2 Sequential Prefetch Tuning
Investigate EMC VMAX zBoost Upgrade

Customer Expectation: Higher throughput & performance for DB2

Case Study 2: DB2 Performance Problems

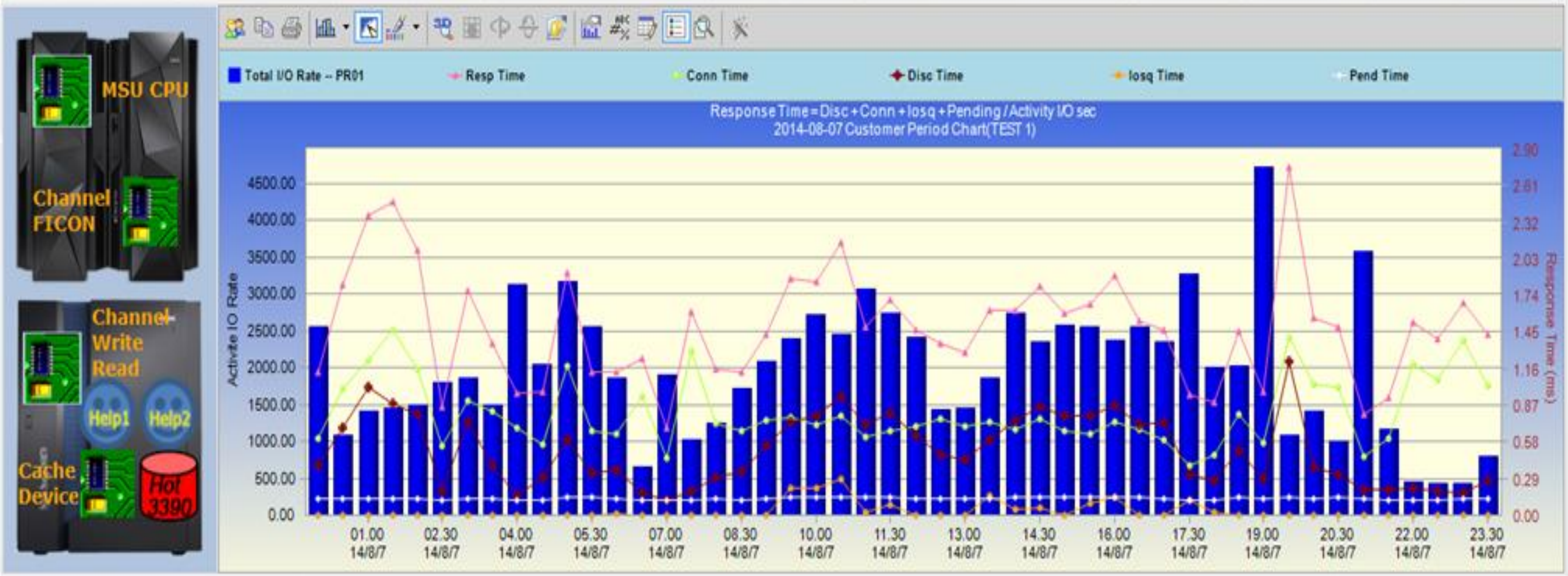
zEC12 Server & z/OS 1.13 Operating System
HDS VSP G1000 DASD Subsystems
DB2 10 Data Sharing Environment

Customer Expectation: Identify DB2 performance problem

Case Study 2: DB2 Performance Problem #1

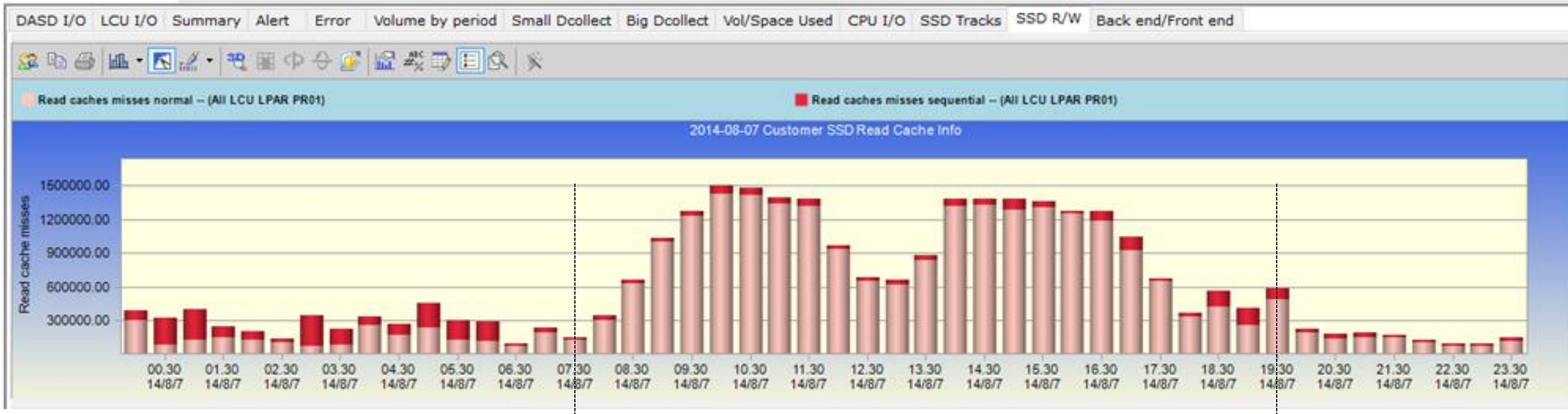
DASD I/O | LCU I/O | Summary | Alert | Error | Volume by period | Small Dcollect | Big Dcollect | Vol/Space Used | CPU I/O | SSD Tracks | SSD R/W | Back end/Front end

Date	Time	Period	LPAR	I/O	Resp	Conn	Disc	Iosq	Pend	Qint	Busy%	Vol used	Used Capa(Go)	Vol Total	Capa(Go)	Access densi...	Access...
8/7/2014...	00.00	Batch	PR01	2544.50	1.13	0.60	0.40	0.00	0.13	1333.25	11.49	217	1823	3621	30577	11.73	1.40
8/7/2014...	00.30	Batch	PR01	1081.69	1.82	1.00	0.69	0.00	0.13	887.50	2.43	130	1099	3621	30577	8.32	0.98
8/7/2014...	01.00	Batch	PR01	1417.53	2.38	1.23	1.01	0.00	0.13	1627.69	2.83	187	1568	3621	30577	7.58	0.90
8/7/2014...	01.30	Batch	PR01	1448.91	2.49	1.47	0.88	0.00	0.13	1475.75	3.62	207	1755	3621	30577	7.00	0.83
8/7/2014...	02.00	Batch	PR01	1500.01	2.10	1.16	0.80	0.00	0.13	1404.94	2.69	202	1713	3621	30577	7.43	0.88
8/7/2014...	02.30	Batch	PR01	1791.04	0.86	0.55	0.19	0.00	0.12	557.53	5.99	82	690	3621	30577	21.84	2.60
8/7/2014...	03.00	Batch	PR01	1865.49	1.78	0.91	0.74	0.00	0.13	1628.46	3.01	155	1312	3621	30577	12.04	1.42

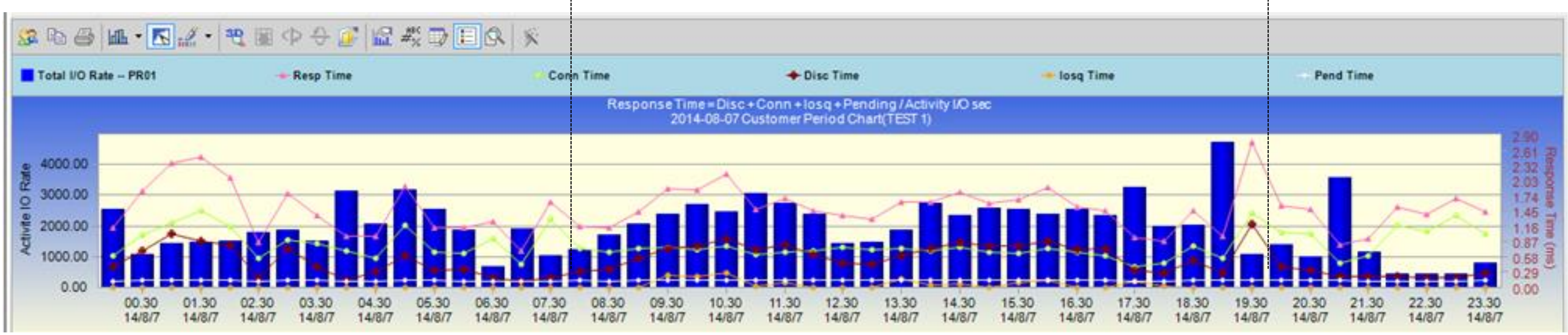


In this example DISC time is high all-day, with intermittent high IOSQ

Case Study 2: DB2 Performance Problem #2

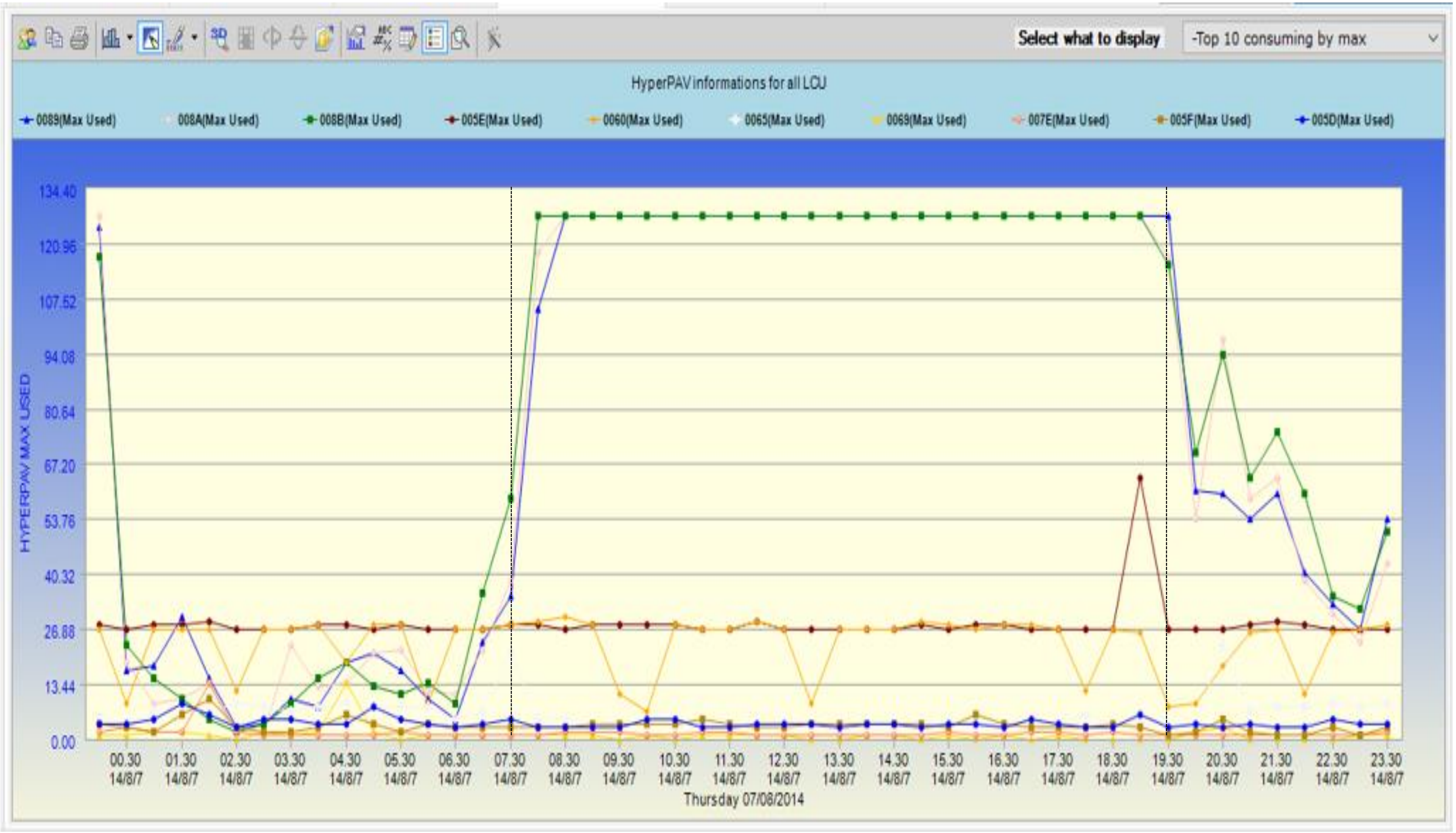


As per the high DISC time below, follows normal read cache misses



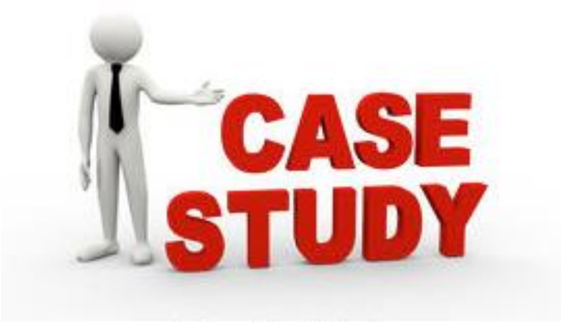
This highlights disk controller cache memory capacity constraints

Case Study 2: DB2 Performance Problem #3



Analysing the high IOSQ shows PAV Alias constraints for LCU 0089/A/B

Case Study 2: DB2 Performance Problem



zEC12 Server & z/OS 1.13 Operating System
HDS VSP G1000 DASD Subsystems
DB2 10 Data Sharing Environment

High Disconnect - Cache Capacity Issues
Investigate Cache/Disk Subsystem Upgrade
High IOSQ (HyperPav 128) - LCU Issues
Consider HyperPav Reconfiguration (32/64)
Consider Data Set Balancing Across LCUs

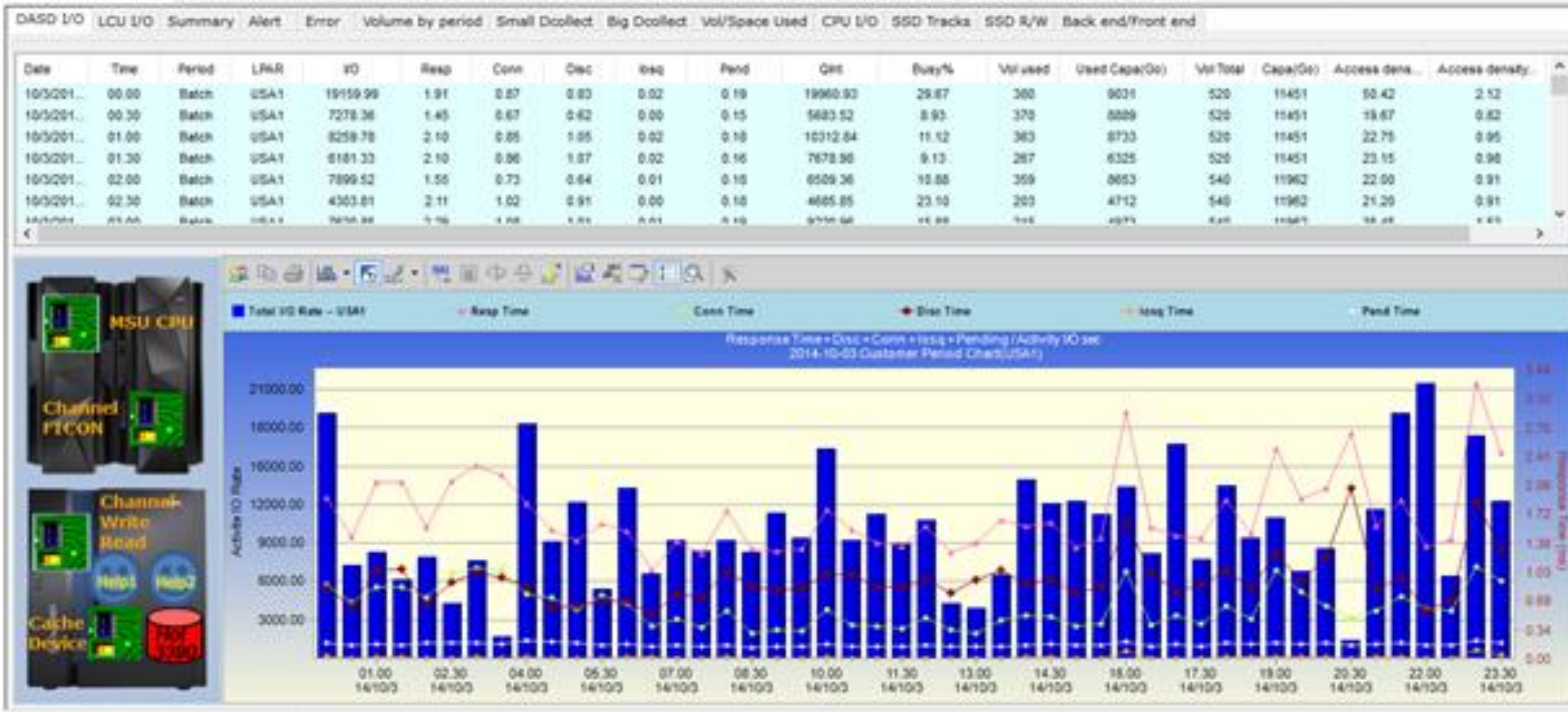
Customer Expectation: Identify DB2 performance problem

Case Study 3: DB2 STOGROUP Performance

z13 Server & z/OS 2.1 Operating System
IBM DS8870 DASD Subsystems
DB2 11 Environment

Customer Expectation: Identify DB2 performance problem

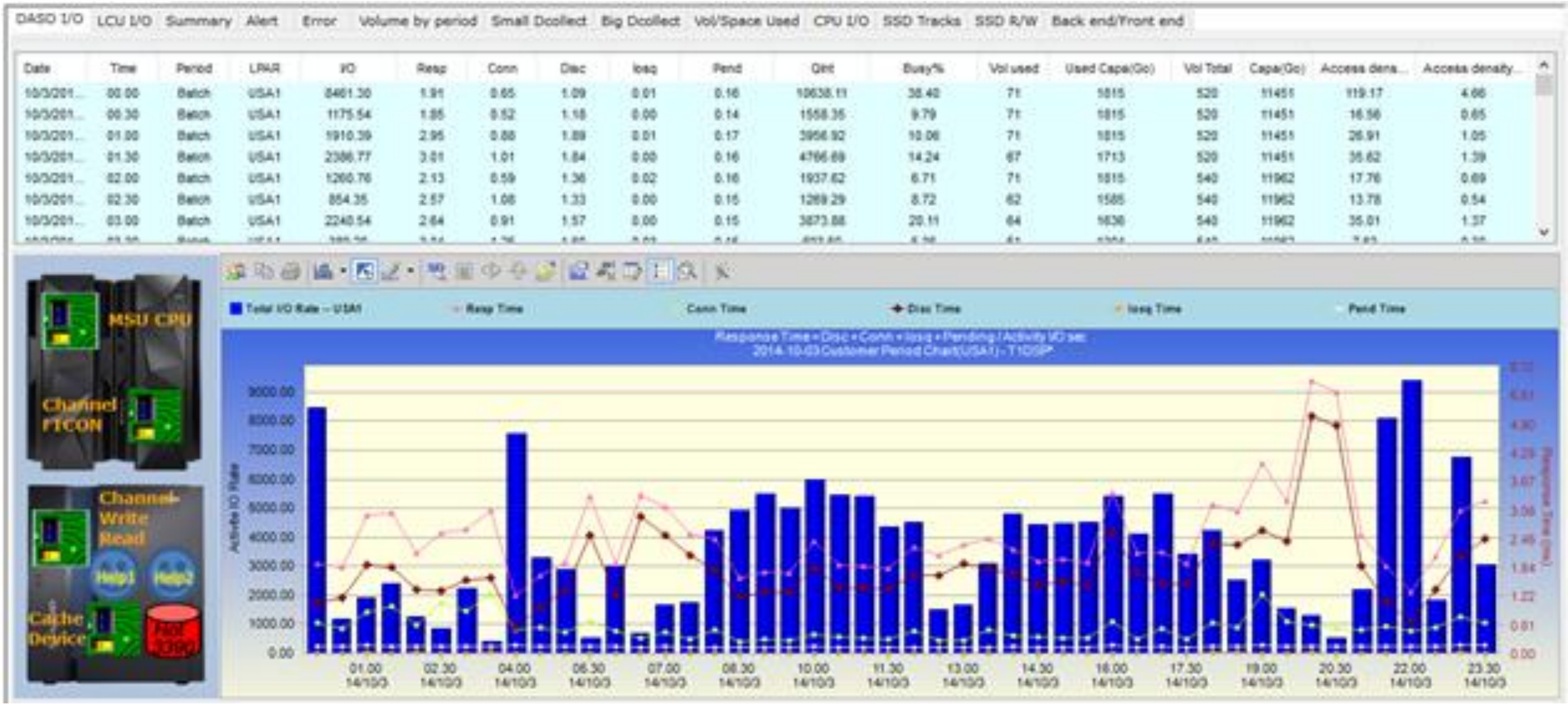
Case Study 3: DB2 STOGROUP Performance #1



	TP			Batch			TP+Batch			
	Average	Min	Max	Average	Min	Max	Average	Min	Max	
Resp Time	1.59	1.27	2.92	1.85	1.05	3.28	1.73	1.05	3.28	ms
Disc Time	0.94	0.78	1.62	0.91	0.52	2.03	0.93	0.52	2.03	ms
Iosq Time	0.01	0.00	0.08	0.02	0.00	0.11	0.01	0.00	0.11	ms
Conn Time	0.48	0.30	1.03	0.75	0.37	1.09	0.62	0.30	1.09	ms
Pend Time	0.16	0.13	0.19	0.18	0.14	0.22	0.17	0.13	0.22	ms
Activity I/O	10359.54	3921.84	16728.25	10038.11	1428.13	21516.57	10185.43	1428.13	21516.57	I/OSec

Analysing the entire disk subsystem; I/O service times are acceptable

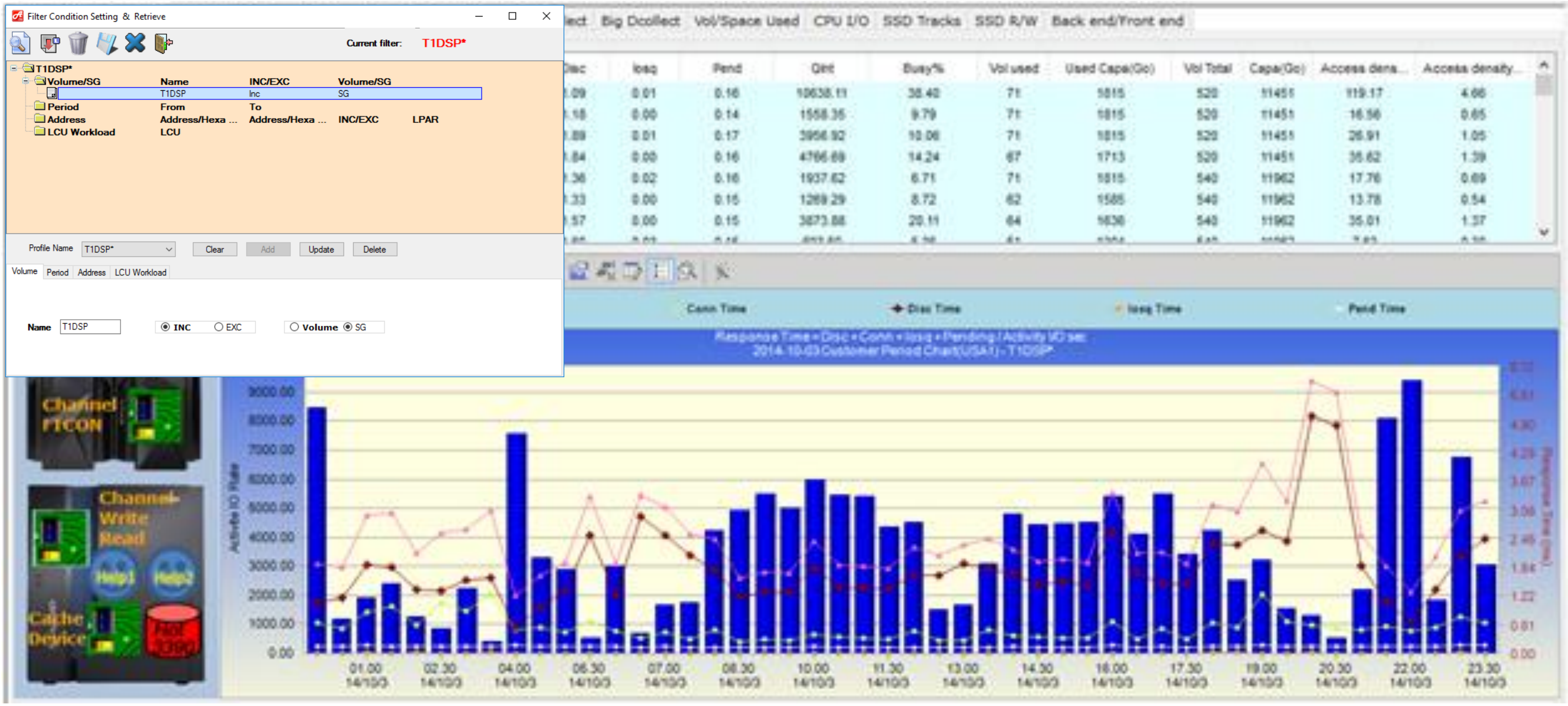
Case Study 3: DB2 STOGROUP Performance #2



	TP			Batch			TP+Batch			
	Average	Min	Max	Average	Min	Max	Average	Min	Max	
Resp Time	2.18	1.60	3.43	2.29	1.23	5.83	2.23	1.23	5.83	ms
Disc Time	1.65	1.22	2.59	1.49	0.55	5.06	1.58	0.55	5.06	ms
Iosq Time	0.01	0.00	0.02	0.01	0.00	0.03	0.01	0.00	0.03	ms
Conn Time	0.40	0.26	0.68	0.63	0.31	1.26	0.50	0.26	1.26	ms
Pend Time	0.14	0.12	0.16	0.16	0.13	0.18	0.15	0.12	0.18	ms
Activity I/O	4320.43	1518.24	5964.70	3001.17	389.20	9417.03	3605.83	389.20	9417.03	I/OSec

Analysing global response times indicates potential high DISC time

Case Study 3: DB2 STOGROUP Performance #3



	TP			Batch			TP+Batch			
	Average	Min	Max	Average	Min	Max	Average	Min	Max	
Resp Time	2.18	1.60	3.43	2.29	1.23	5.83	2.23	1.23	5.83	ms
Disc Time	1.65	1.22	2.59	1.49	0.55	5.06	1.58	0.55	5.06	ms
Iosq Time	0.01	0.00	0.02	0.01	0.00	0.03	0.01	0.00	0.03	ms
Conn Time	0.40	0.26	0.68	0.63	0.31	1.26	0.50	0.26	1.26	ms
Pend Time	0.14	0.12	0.16	0.16	0.13	0.18	0.15	0.12	0.18	ms
Activity I/O	4320.43	1518.24	5964.70	3001.17	389.20	9417.03	3605.83	389.20	9417.03	I/OSec

Analysing the DB2 STOGRP T1DSP indicates high DISC time

Case Study 3: DB2 STOGROUP Performance #4

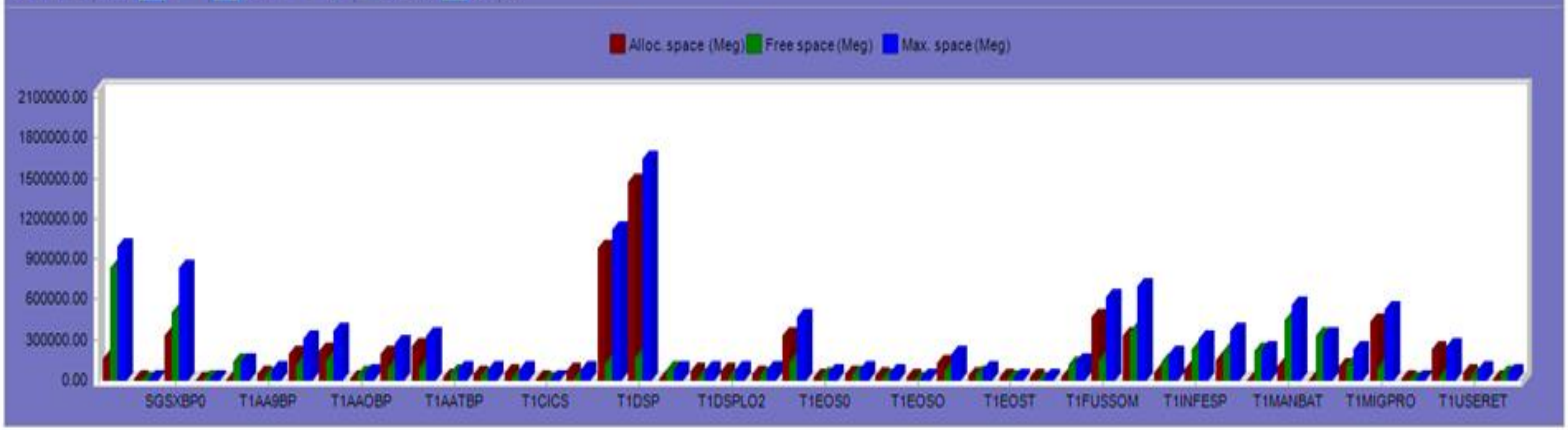
DASD I/O | LCU I/O | Summary | Alert | Error | Volume by period | Small Dcollect | Big Dcollect | Vol/Space Used | CPU I/O | SSD Tracks | SSD R/W | Back end/Front end

Detail Report | LPAR Report | Stogroup Report | Volser Report | Time Report

DCOLLECT-TP01-D141002-H221

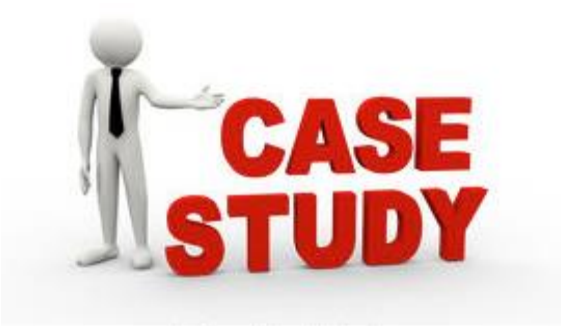
Report by Stogroup

No.	St	Alloc. space (Meg)	Free space (Meg)	Max. space (Meg)	%Free	Largest extent	Avg. frag. index
1	T	322	222435	222760	99.9	27827	0
2	T	915	333221	334140	99.7	27837	1
3	S	2164	137057	139225	98.4	27587	5
4	S	5102	6250	11352	55.1	1402	88
5	T	6536	4816	11352	42.4	1017	154
6	T	7141	76393	83535	91.5	19108	72
7	T	7816	698	8514	8.2	328	0
8	T	8320	47369	55690	85.1	16509	170
9	T	8973	46716	55690	83.9	13799	68
10	S	9968	4222	14190	29.8	934	214



DCOLLECT data; DB2 STOGRP T1DSP has highest fragmentation

Case Study 3: DB2 STOGROUP Performance



z13 Server & z/OS 2.1 Operating System
IBM DS8870 DASD Subsystems
DB2 11 Environment

High Disconnect - High DASD Fragmentation
Consider DB2 STOGROUP Reorgs
Consider SSD For DB2 STOGROUP(s):
Eliminating Tablespace Reorg Requirements

Customer Expectation: Identify DB2 performance problem

Questions & Session feedback

- Please submit your feedback at

  http://conferences.gse.org.uk/2016/feedback/lj

- Session is LJ