

Hints & Tips Of Dataset I/O Performance

Meral Temel
Türkiye İş Bankası (İŞBANK)

14 August 2013
14032





Hints & Tips Of Dataset I/O Performance

Agenda



- ⌘ Who is İşBank ?
- ⌘ Mainframe Configuration
- ⌘ Why Do We Need To Analyze SMF42 Records ?
- ⌘ How Is The Life Of I/O In MVS ?
- ⌘ Dataset Access Types
- ⌘ What Is In SMF 42 ?
- ⌘ SMF 42 Dataset I/O Statistics
- ⌘ Sample Studies Using SMF 42 Records
- ⌘ Planned Tasks




Who Is İŞBANK ?



 **The Biggest & First Bank Of Turkey**

 **4851 ATMs**

 **1231 Branches In Turkey, 19 Branches Outside Turkey**

 **Has The Highest Profit According To All Bank Announcements 2013**

 **Member Of SHARE Inc.**



Who Is İŞBANK ?

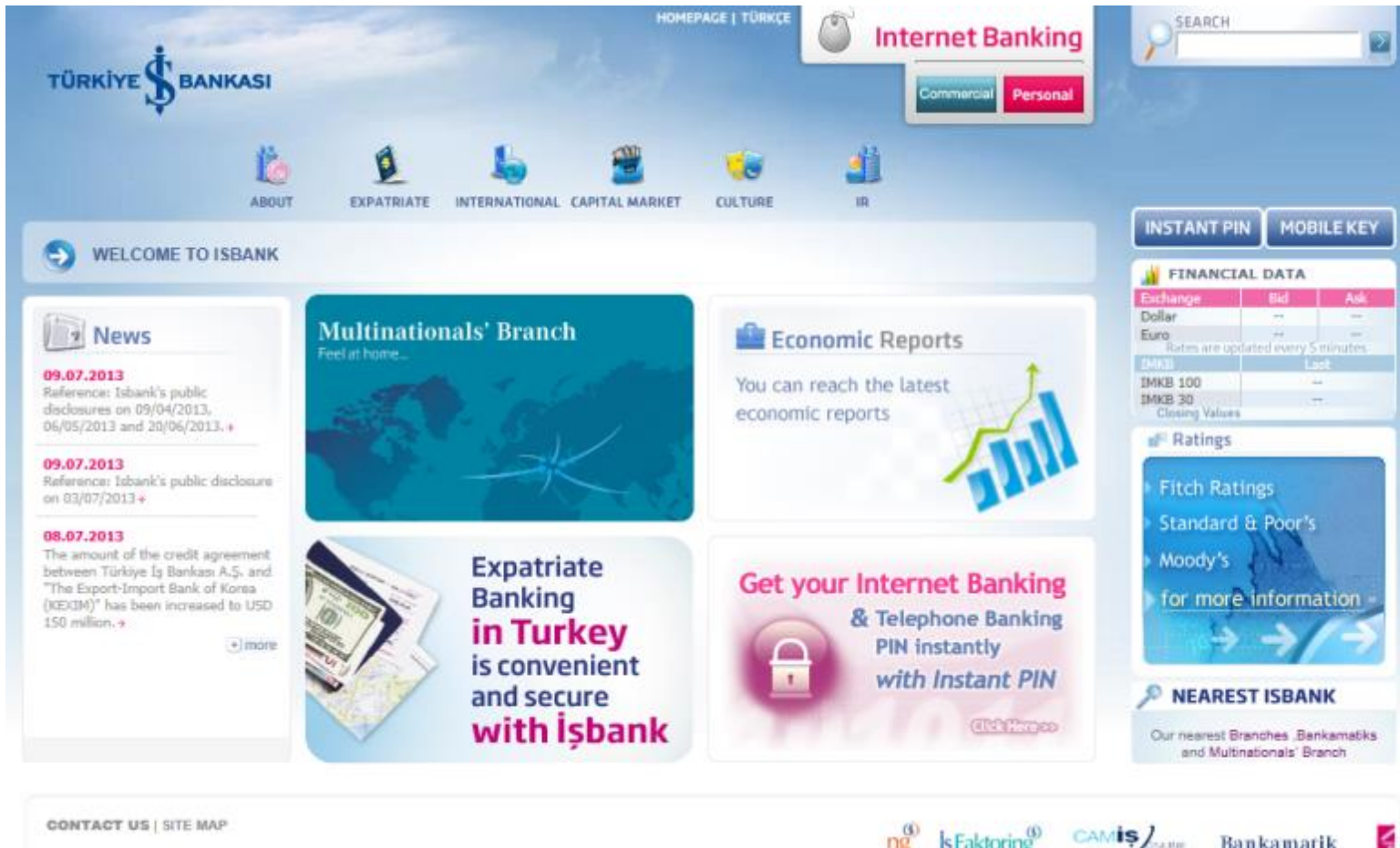


BRANCHES



Who Is İŞBANK ?

INTERNET BANKING



HOME PAGE | TÜRKÇE

TÜRKİYE İŞ BANKASI

Internet Banking

Commercial Personal

ABOUT EXPATRIATE INTERNATIONAL CAPITAL MARKET CULTURE İR

WELCOME TO İSBANK

News

09.07.2013
Reference: İsbank's public disclosures on 09/04/2013, 06/05/2013 and 20/06/2013. +

09.07.2013
Reference: İsbank's public disclosure on 03/07/2013 +

08.07.2013
The amount of the credit agreement between Türkiye İş Bankası A.Ş. and "The Export-Import Bank of Korea (KEKİM)" has been increased to USD 150 million. +

more

Multinationals' Branch
Feel at home...

Economic Reports
You can reach the latest economic reports

Expatriate Banking in Turkey is convenient and secure with İşbank

Get your Internet Banking & Telephone Banking PIN instantly with Instant PIN

Click Here

SEARCH

INSTANT PIN MOBILE KEY

FINANCIAL DATA

Exchange	Bid	Ask
Dollar	---	---
Euro	---	---
Rates are updated every 5 minutes.		
İMKB	Last	
İMKB 100	---	
İMKB 30	---	
Closing Values		

Ratings

- Fitch Ratings
- Standard & Poor's
- Moody's

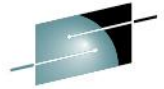
for more information

NEAREST İSBANK

Our nearest Branches, Bankamatik and Multinationals' Branch

CONTACT US | SITE MAP

İs Faktoring CAMİS Bankamatik



SHARE
Technology - Connections - Results

Who Is İŞBANK ?

ATM



İŞCEP Mobile Phone Application



İŞBANK IPAD FINANCE CENTER Application

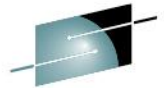


Who Is İŞBANK ?

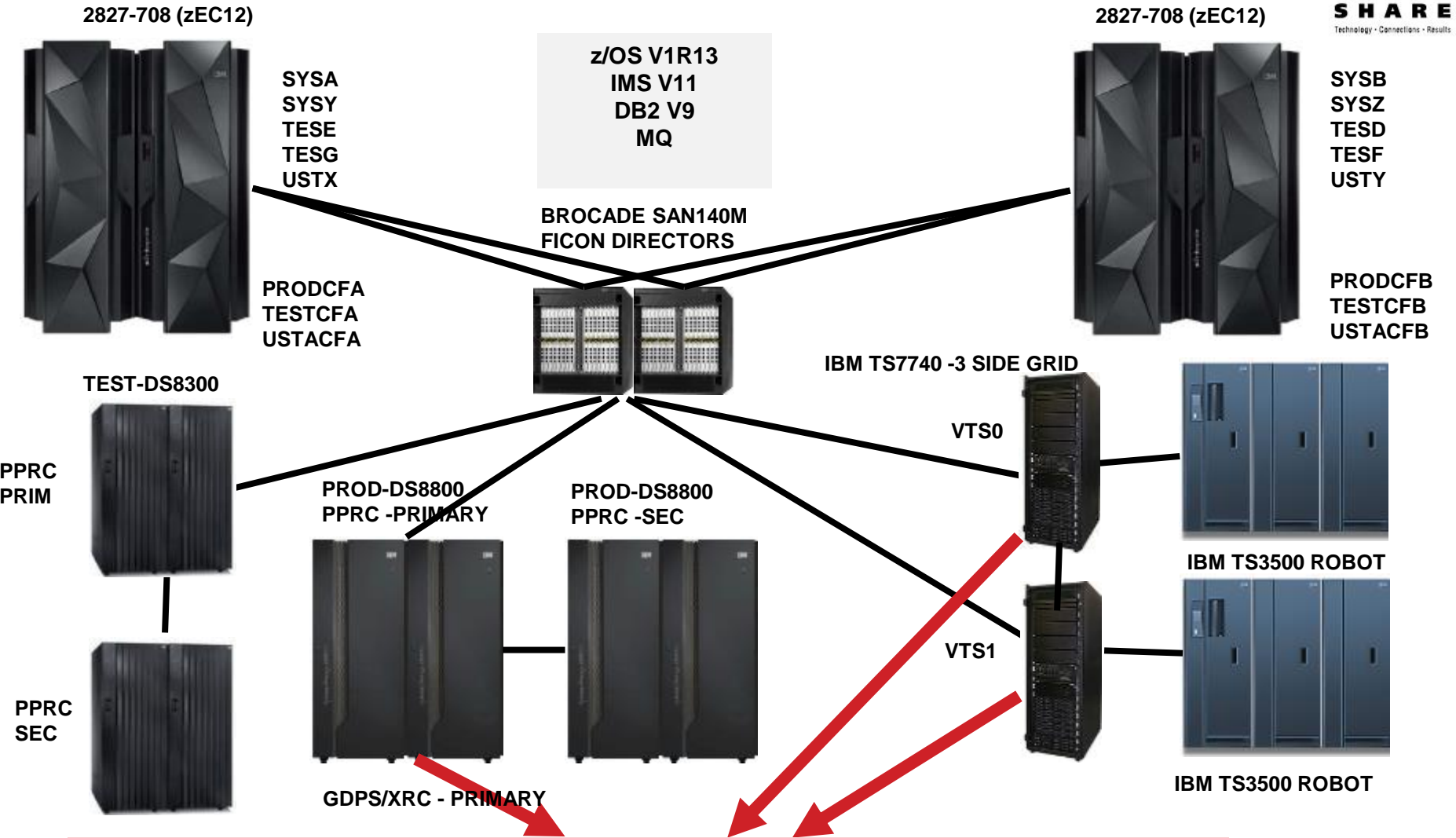
Credit Cards



İşbank – Mainframe Configuration



SHARE
Technology · Connections · Results



IZMIR DISASTER CENTER (600 Km Away From Istanbul)



Işbank – zEC12 Configuration Details

2827-708 (zEC12)



8 GCP (5.5 GHz)
3 ICF,3 zIIP,1 IFL
10063 IBM PCI
10062,8 MIPS(Average RNI)
1224 MSU
1257,9 MIPS/CP
153 MSU/CP
64777 SU/sec
192 GB Memory
- 160 GB Customer
- 32 GB HSA
18 FICONExpress 8S
(32port- 32 FICON Channel)
10 HCA3-Fanout Cards
20 Infiniband3 CF Link
12 OSA-Express4S 10 GbE
SR 1 port
4 OSA-Express4S 1000BASE-T
4 Crypto Express4S
2 FlashExpress
zAware

2827-708 (zEC12)



8 GCP (5.5 GHz)
3 ICF,3 zIIP
10063 IBM PCI
10062,8 MIPS(Average RNI)
1224 MSU
1257,9 MIPS/CP
153 MSU/CP
64777 SU/sec
192 GB Memory
- 160 GB Customer
- 32 GB HSA
18 FICONExpress 8S
(32port- 32 FICON Channel)
10 HCA3-Fanout Cards
20 Infiniband3 CF Link
12 OSA-Express4S
10 GbE SR 1 port
4 OSA-Express4S 1000BASE-T
4 Crypto Express4S

Why Do We Need To Analyse SMF 42s ?

Large Volumes – Device Responetime Relationship

**Batch Run Performance Problem Troubleshooting
-Including DB2/IMS Files
Merging With SMF 30s**

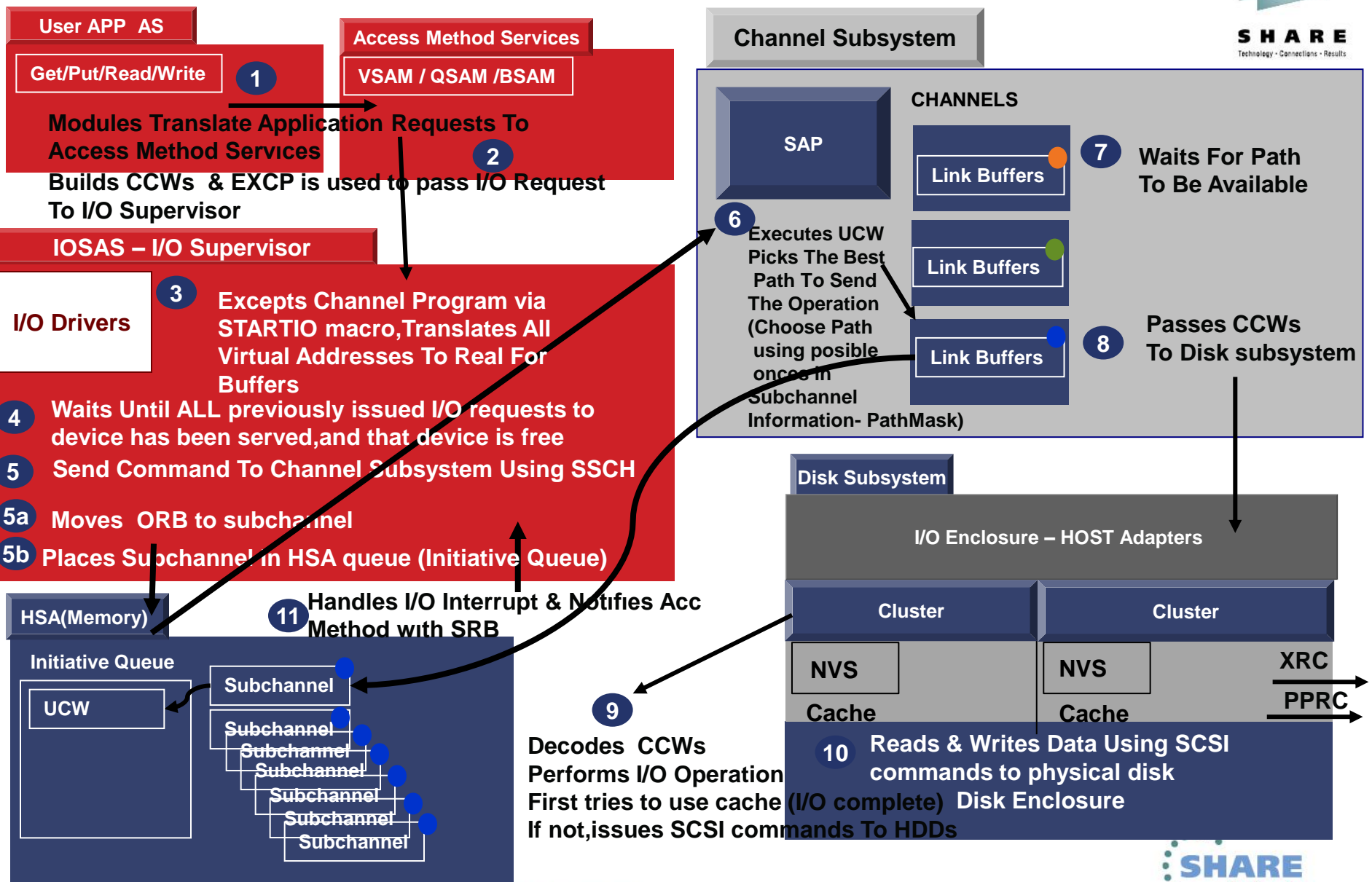
More information related to most loved ones

....And many other important items related to dataset activity

Way to understand what Is going on in each DB2/IMS I/O management changes

If you start using SSD, analysis of dataset I/O performance becomes greate periodic method to monitor performance

Life OF I/O Flow



Which Period Of Time Causes I/O ResponseTime Components ?

User APP AS

Get/Put/Read/Write **1**

Access Method Services

VSAM **2**

5-3 IOSQTime

Modules Translate Application Request To Access Method Services
Builds CCWs & EXCP is used to pass I/O Request To I/O Supervisor

IOSAS - I/O Supervisor

I/O Drivers **3**

Excepts Channel Program via STARTIO macro, Translates All Virtual Addresses To Real For Buffers

4 Waits Until ALL previously issued I/O requests to device has been served, and that device is free

5 Send Command To Channel Subsystem Using SSCH

5a Moves ORB to Subchannel

5b Places Subchannel in Initiative Queue

9-6 Pend Time

Channel Subsystem

SAP

Link Buffers

Link Buffers

6 Executes UCW Picks The Best Path To Send The Operation (Choose Path using possible paths once in Subchannel Information)

7 Waits For Path To Be Available

Passes CCWs To Disk subsystem

9,10+ PPRCCopyTime+XR C-WP DisconnectTime *10-Cache Miss (Subchannel is busy but device is not connected)

HSA(Memory)

Initiative Queue

UCW

Subchannel

Subchannel

Subchannel

Subchannel

Subchannel

Subchannel

11 Handles I/O Interrupt & Notifies Access Method with SRB

9 Decodes CCWs Performs I/O Operation First tries to use cache (I/O complete) Disk Enclosure If not, issues SCSI commands To HDDs

Cluster

Cluster

NVS

NVS

Cache

Cache

DeStaging

XRC

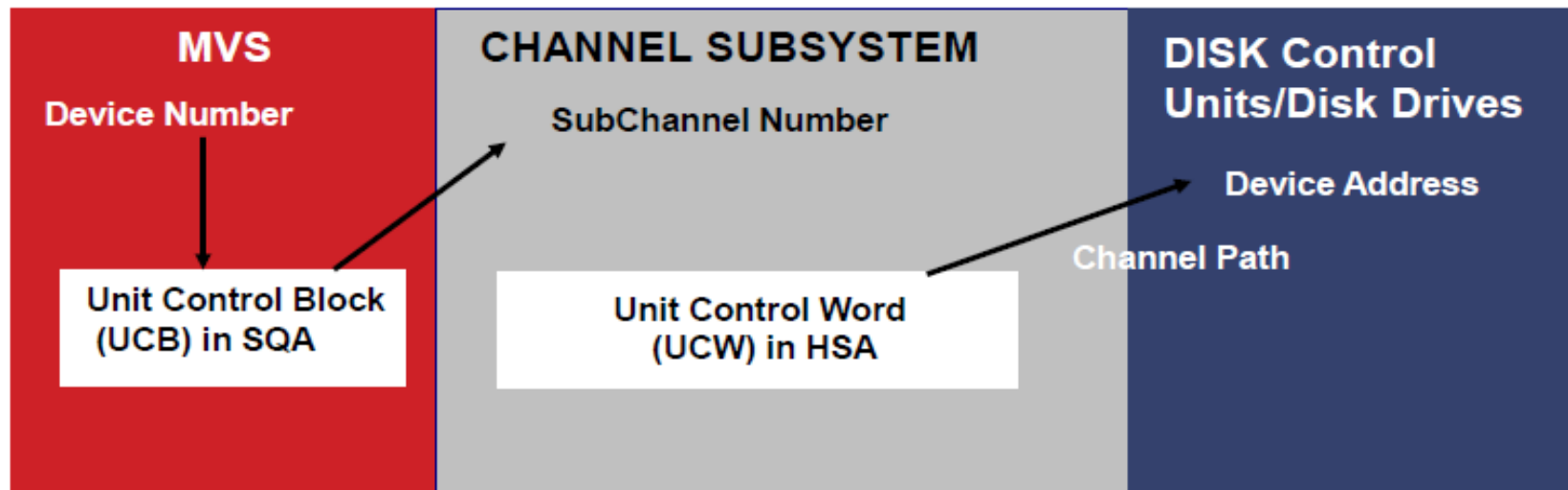
PPRC

Staging

10 Reads & Writes Data Using SCSI commands to physical disk

Life Of I/O Terms That We Need To Remember

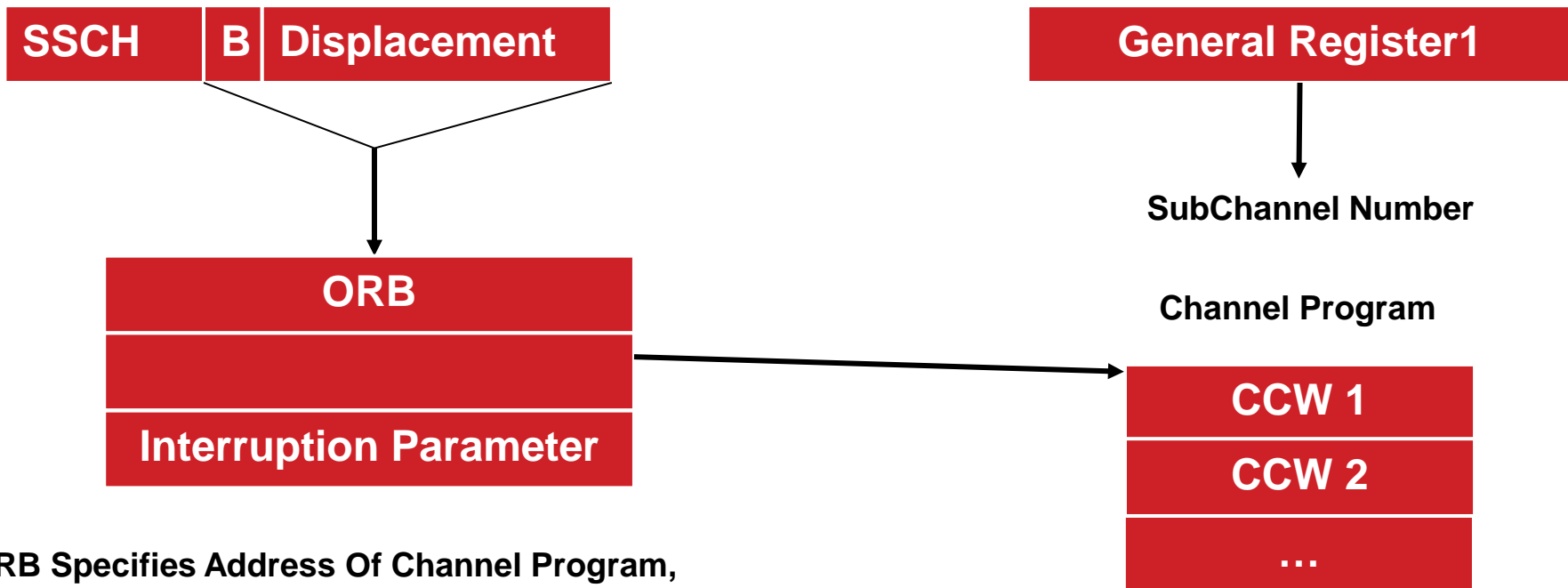
- ❑ Each Device Has UCB in SQA in MVS
- ❑ Each Device Has Corresponding Subchannel Number maintained by Channel Subsystem
- ❑ UCB contains subchannel number, device type,
List of users that have outstanding request for I/O on device
- ❑ Channel Subsystem Identifies Device Through Subchannel Number
- ❑ UCW are created during POR /dynamic IODF activation, located in HSA
- ❑ UCW(Unit Control Word) contains control blocks with subchannel description & status



Life Of I/O – TERMS To Remember

- ❑ **SSCH** : Start Subchannel Instruction - Created By IOSAS

Specifies SubChannel Number & ORB (Operation Request Block)



ORB Specifies Address Of Channel Program,
Interruption Parameter,
Storage Key,
Channel Path Mask,
Control Bits

- ❑ Channel Program Is A Set Of CCWs

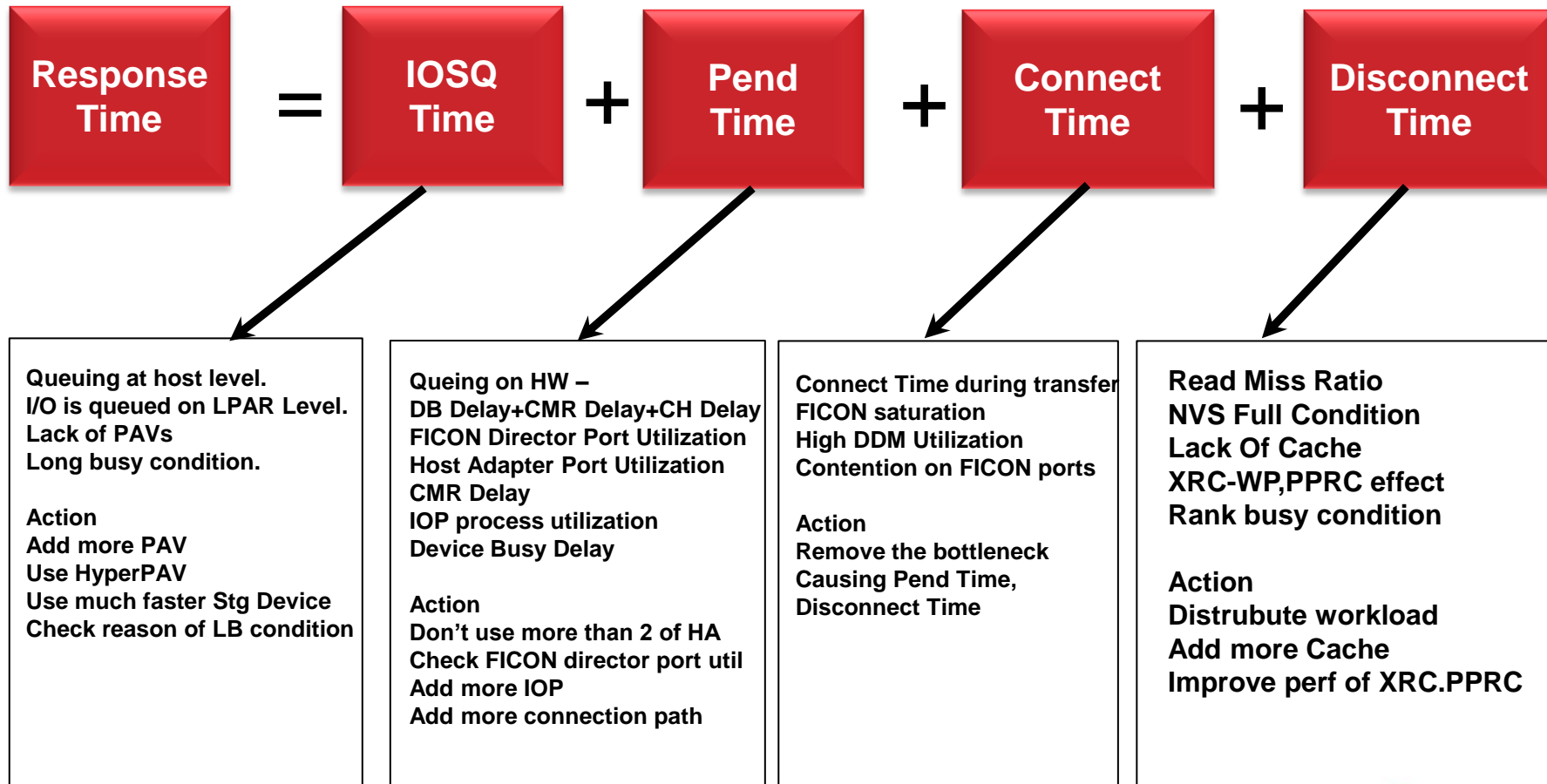
I/O Response Time Components



**New Interrupt Delay Time Is Not Included In ResponseTime
This is very usefull information**

Items That Effect Each ResponseTime Component

What Are The Possible Reasons? Where To Look ? What To Do ?

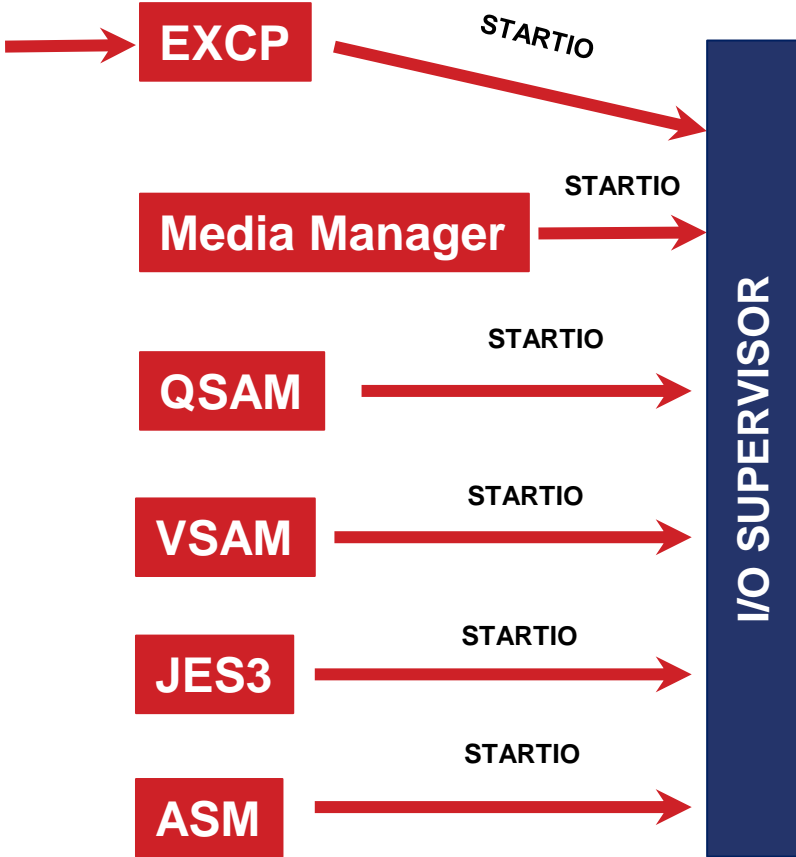


Dataset Access Method Types

Insert
Custom
Session
QR if
Desired.



Application Programs
DFHSM
JES2
DFDSS
DFSORT



SSCH CHANNEL SUBSYSTEM



SMF42 Subtypes

- Subtype 1 – Storage Class Related Information-BufferManagement Facility
- Subtype 2 – Cache Control unit
- Subtype 3 – SMS Configuration Changes
- Subtype 4 – SDM Statistics /Extended Seq Dataset Concurrent Copy
- Subtype 5 – Storage class VTOC and VVDS I/O statistics.
- Subtype 6 – records DASD data set level I/O statistics.
- Subtype 7/8/26 – NFS.
- Subtype 9 – is written each time a B37/D37/E37 abend
- Subtype 10 – is written at the time of volume selection failure because of insufficient space when allocating a data set.
- Subtype 11 – is written for extended remote copy (XRC) session statistics whenever the SMF timer interval ends.
- Subtype 14 – (ADSM)
- Subtype 15/16/17/18/19 – VSAM-RLS
- Subtype 20 who/what STOW INITIALIZE to delete all the members from a PDSE
- Subtype 21 who/what when a member is deleted from a PDS or a PDSE
- Subtype 22/23 DFSMSrmm for audit/security records.
- Subtype 24 is written when a member is added or replaced from a PDS or a PDSE
- Subtype 25 who/what when a member is renamed from a PDS or a PDSE

DASD DATASET LEVEL I/O STATISTICS

There are two events that cause subtype 6 to be generated:

- 1- Close, or
- 2- Immediately after the recording of the type 30 interval record. There is one type 42 subtype 6 record for each type 30 interval record.

For the SMSPDSE and SMSPDSE1 address spaces, type 42 subtype 6 records are recorded at the BMFTIME interval immediately after the type 42 subtype 1 records

Dataset Level I/O Statistics



S42DSIOR - Average Responsetime

S42DSIOC - Average Connecttime

S42DSIOP - Average Pendtime

S42DSIOD - Average Disconnecttime

S42DSIOQ - Average ControlUnitQtime

S42DSMXR – Maximum I/O Responsetime

S42DSMXS – Maximum Dataset Servicetime

Average IOSQtime in MXG variables



Dataset Level I/O Statistics- Related To IO Count



Total IO Count

Sequential IO Count

Sequential Blocks Read/Write

Random Blocks Read/Write

DASD Rate And Several MXG fields

Dataset Level I/O Statistics- Related To Cache



Read Cache Hit %

Write Cache Hit %

Dataset Level I/O Statistics- Related To AS



Jobname

System

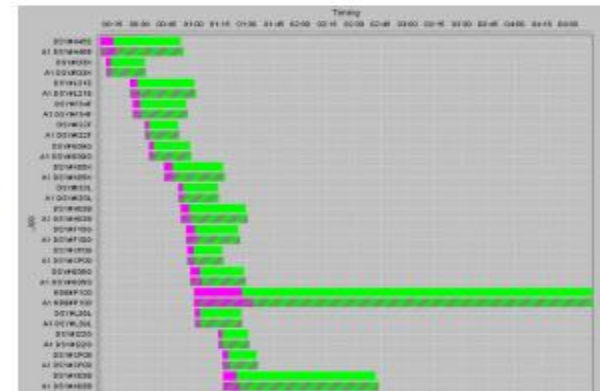
Service Class

Workload Name

Now Used By zBNA

IBM System z Batch Network Analyzer (zBNA)

- IBM System z Batch Network Analyzer
 - A free, “as is” tool to analyze batch windows
 - Available to Customers, Business Partners and IBMers
 - PC based, providing graphical and text reports
 - Includes Gantt charts and support for Alternate Processors (what if scenarios)
 - Next release* targeted for 4Q2013
 - Support for SMF 42 records to understand the “Life of a Data Set”
 - Support for SMF 14/15 records to understand zEDC compression candidates



- Available Now on Techdocs

<https://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/PRS5132>

EWCP Project Opening – Kathy Walsh Hot Topics

SSD Eligible Datasets Analysis Using Flashda



Flashda is IBM's free program that is used in SSD analysis

<http://www-03.ibm.com/systems/z/os/zos/downloads/flashda.html>

Flashda uses SMF42 and SMF74

SAS program – I requested from Barry to write Windows SAS version of this program and we are lucky that he accepted and I used it for our data



Sample FlashDA program output

Sorted Ascending Order Of (I/O Count*ReadOnly Disconnect Time)

DEVICE NUMBER	DATA VOLSER	SET NAME	TOTAL	AVERAGE	TOTAL	AVERAGE	AVERAGE	AVERAGE	AVERAGE	AVERAGE	TOTAL	TOTAL		
			READ-ONLY DISCON.TIME (MILLISEC)	READ-ONLY DISCON.TIME (MILLISEC)	DISCONNECT TIME (MILLISEC)	DISCONNECT TIME (MILLISEC)	RESPONSE TIME (MILLISEC)	I/O CONNECT TIME (MILLISEC)	PENDING TIME (MILLISEC)	CONTROL UNIT QUEUE (MILLISEC)			CLASS	STORAGE SIZE
5931	IMSCPL	IMSPDB1.CPBCRD.DATA	891730.0	3.576	891730.0	3.576	4.348	0.388	0.064	0.000	IMSK	8192	249331	0
5237	IMSCPQ	IMSPDB1.CPCUSF03.DATA	725548.1	2.248	719590.7	1.720	2.360	0.309	0.039	0.000	IMSK	8192	322795	95461
5C07	IMSCPU	IMSPDB1.CPBCRD.DATA	676921.0	3.840	676921.0	3.840	4.608	0.448	0.064	0.000	IMSK	8192	176280	0
5112	IMSCP5	IMSPDB1.CPCUSF01.DATA	644580.3	2.623	645229.8	1.895	2.821	0.618	0.037	0.000	IMSK	8192	245746	94753
5013	IMSCP4	IMSPDB1.CPCRD08.DATA	642336.1	2.530	650086.7	1.931	2.743	0.487	0.047	0.000	IMSK	8192	253912	82812
5B0B	IMS671	IMSPDB1.MAASDS01.D	627421.9	0.374	624240.1	0.367	0.767	0.162	0.089	0.000	IMS	2048	1678230	22159
5C07	IMSCPU	IMSPDB1.CPCRD16.DATA	623554.3	0.932	632733.8	0.899	1.416	0.272	0.000	0.000	IMSK	8192	668738	35443
5A0A	IMS622	IMSPDB1.DATA.KPF00300	520087.3	2.113	522177.2	2.107	2.659	0.190	0.000	0.000	IMS	4096	246121	1718
560F	IMSCPE	IMSPDB1.CPCRD10.DATA	492303.7	1.832	502466.6	1.479	2.073	0.274	0.037	0.000	IMSK	8192	268664	71026
610F	IMSCP8	IMSPDB1.CPBCRD.DATA	464744.5	3.821	464744.5	3.821	4.651	0.575	0.000	0.000	IMSK	8192	121639	0
1211	IMS669	IMSPDB1.DATA.AL102000	440971.1	0.507	454356.3	0.511	0.908	0.128	0.000	0.000	IMS	4096	870295	18921
5C07	IMSCPU	IMSPDB1.CPCRD15.DATA	435866.2	1.033	449709.7	0.853	1.386	0.264	0.019	0.000	IMSK	8192	421928	105424
5E0A	IMO009	IMSPDB1.BLOKEG01	433243.9	1.209	433306.2	1.209	1.677	0.213	0.033	0.000	IMSOSA	1690	358489	58
120F	IMS667	IMSPDB1.DATA.AL342500	429446.7	0.489	423499.5	0.471	0.906	0.155	0.014	0.000	IMS	4096	878310	21199
1210	IMS668	IMSPDB1.DATA.AL104000	421128.8	0.502	423137.0	0.491	0.890	0.128	0.076	0.000	IMS	4096	838978	23475
1114	IMS2X1	IMSPDB1.CPBCRD.DATA	404116.4	3.316	404116.4	3.316	4.024	0.332	0.062	0.000	IMSK	8192	121859	0
5E0B	IMO010	IMSPDB1.BLOKEG01	403840.3	1.821	403849.3	1.821	2.364	0.225	0.025	0.000	IMSOSA	1690	221742	5
6315	IMSCP2	IMSPDB1.CPBCRD.DATA	397109.2	3.328	397109.2	3.328	3.776	0.192	0.000	0.000	IMSK	8192	119320	0
5921	IMSCPI	IMSPDB1.CPCUSF02.DATA	394262.7	2.246	399674.1	1.713	2.371	0.364	0.038	0.000	IMSK	8192	175574	57810
6C24	DPUS82	DB2PROBD.DSNCDB.DFIDCUST.	393074.5	1.431	394286.8	1.431	1.837	0.130	0.000	0.000	DBPSC	4096	274605	860
6824	IMS453	IMSPDB1.DATA.DIBCEK01	388021.8	0.405	390096.9	0.406	0.662	0.047	0.000	0.000	IMS	2048	957001	2807

Dataset I/O Statistics Section

Offsets	Name	Length	Format	Description
0	0 S42DSIOR	4	binary	Average response time.
4	4 S42DSIOC	4	binary	Average I/O connect time. There is one SMF subtype 6 record per data set, and each record has its own S42DSIOC field.
8	8 S42DSIOP	4	binary	Average I/O pending time.
12	C S42DSIOD	4	binary	Average I/O disconnect time.
16	10 S42DSIOQ	4	binary	Average control unit queue time.
20	14 S42DSION	4	binary	Total number of I/Qs. There is one SMF subtype 6 record per data set, and each record has its own S42DSION field.
24	18 S42DSCND	4	binary	Number of cache candidates.
28	1C S42DSSHTS	4	binary	Number of cache hits.
32	20 S42DSWCN	4	binary	Number of write candidates.
36	24 S42DSWHI	4	binary	Number of write hits.
40	28 S42DSSEQ	4	binary	Number of sequential I/O operations. Operations counted here are not accumulated in S42DSCND and S42DSWCN.
44	2C S42DSRLC	4	binary	Number of record level cache I/O operations.
48	30 S42DSICL	4	binary	Number of inhibit cache load I/O operations.
52	34 S42SDA0	4	binary	Average I/O device-active-only time.
56	38 S42DSMXR	4	binary	Maximum data set I/O response time.
60	3C S42DSMXS	4	binary	Maximum data set service time.
64	40 S42DSRDD	4	binary	Average disconnect time for reads.
68	44 S42DSRDT	4	binary	Total number of read operations.

Some Of The Dataset Level Performance Improvement Items

- **Distributing Datasets Through Different Ranks (Via Storage Group Distribution)**
- **Using more Buffers**
- **Compression – I/O vs CPU**
New zEC12 GA2 has zEDC Special Processor For Compression !
- **LBI support**
- **Responsetime specific checkings**
- **Blocksize**
- **ENQs**
- **Striping**
- **Access Methods –Dataset types**

Access Method Statistics Section

Offsets	Name	Length	Format	Description
0	0 S42AMSRB	4	binary	Sequential read: number of blocks.
4	4 S42AMSRR	4	binary	Sequential read: input/output delay.
8	8 S42AMSWB	4	binary	Sequential write: number of blocks.
12	0C S42AMSWR	4	binary	Sequential write: input/output delay.
16	10 S42AMDRB	4	binary	Direct read: number of blocks.
20	14 S42AMDRR	4	binary	Direct read: total input/output delay.
24	18 S42AMDWB	4	binary	Direct write: number of blocks.
28	1C S42AMDWR	4	binary	Direct write: total input/output delay.
32	20 S42AMZRB	4	binary	Number of directory reads.
36	24 S42AMZRR	4	binary	Directory read: input/output delay.
40	28 S42AMZWB	4	binary	Number of directory writes.
44	2C S42AMZWR	4	binary	Directory write: input/output delay.

Dataset Header Section



49	31	S42DSCOD S42FIRST *	1	binary	Entry descriptor flag Bit Meaning 1... First data set entry since Open. .xxx xxxxx Reserved.
50	32	S42DSFL1 S42DSGSR S42DSLRSR S42DSRLS S42DSNSR S42DSEXC S42DSFXD S42DSPL S42DSEF S42DSEFC	1	binary	Data set descriptor flags Bit Meaning 11xx xxxxx VSAM buffer flags 11.. GSR 10.. LSR 01.. RLS 00.. NSR ..X. Reserved1 Open for EXCP processing 1... Non-VSAM fixed length records1.. Program library1. Extended format1 Compressed format
51	33	*	1	EBCDIC	Reserved.
52	34	S42DSIOO	4	binary	Offset to data set I/O statistics section.
56	38	S42DSAMO	4	binary	Offset to access method statistics section.
60	3C	S42DSVOL	6	EBCDIC	Volume serial number.
66	42	S42DSDEV	2	binary	Device number.
68	44	S42DSSC	8	EBCDIC	Storage class name.
76	4C	S42DSBSZ	4	binary	Block size. For concatenated data sets, this data is taken from the first data set.
80	50	S42DSTRP	2	binary	Number of stripes
82	52	*	6	EBCDIC	Reserved.



Dataset Header Section



Offsets	Name	Length	Format	Description
0	0 S42DSNXT	4	binary	Offset to the next data set header section (0 if the last data set).
4	4 S42DSN	44	EBCDIC	Data set name.
48	30 S42DSTYP	1	binary	Data set type. Value Meaning 0 Other 1 Physical sequential 2 PDS 3 PDSE 4 Direct access 5 ISAM 6 EXCP 7 Extended physical sequential data set 10 HFS 16 KSDS data component 17 KSDS index component 18 Variable RRDS data component 19 Variable RRDS index component 20 Fixed length RRDS 21 Linear data set 22 ESDS



THANKS TO ...



BARRY MERILL MERILLE CONSULTANT (MXG)

LEE LA FREESE INTELLIMAGIC

JEFF BURGER - IBM

