



z/OS Systems Working Group

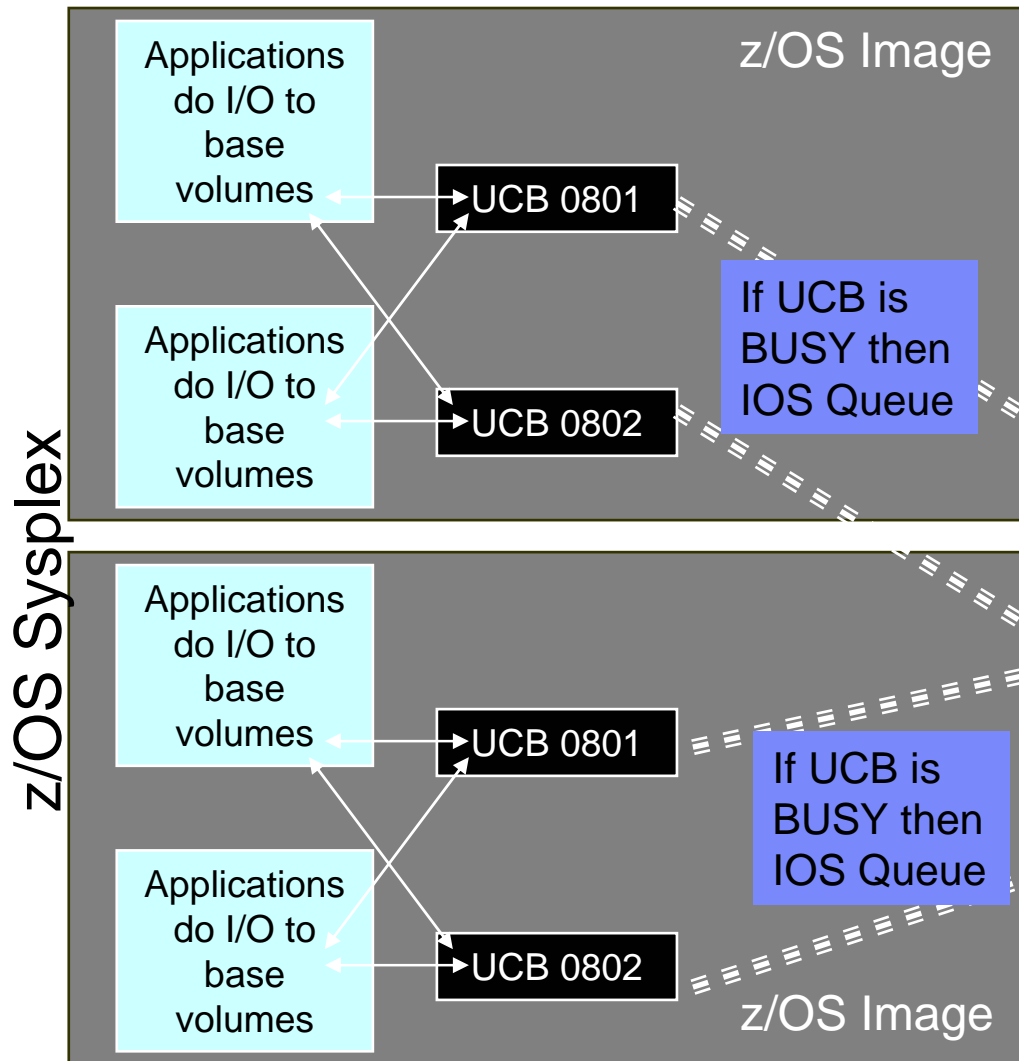
June 6, 2007

HyperPAV

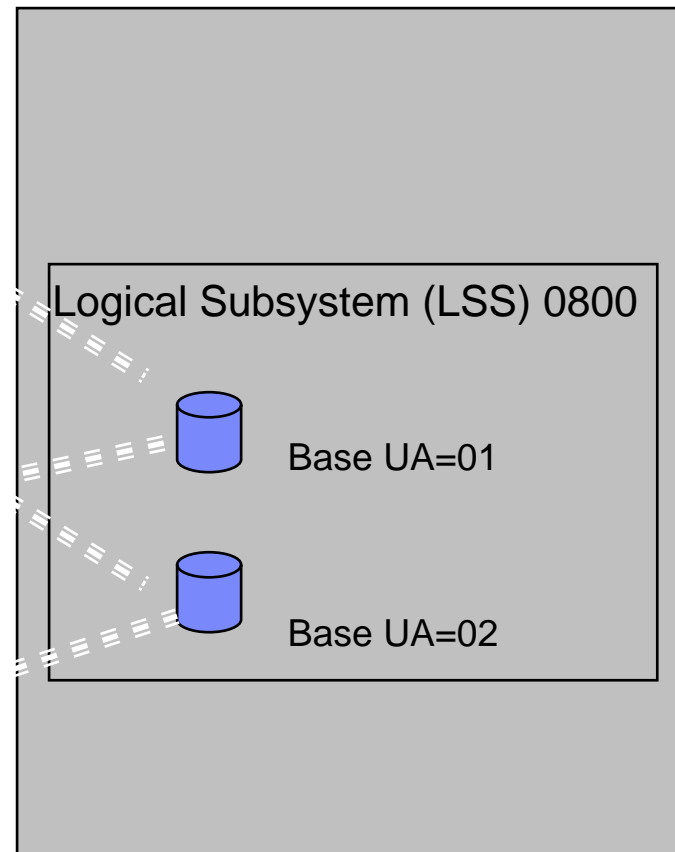
**Jacques Glorieux
IBM Belgium**

jacques_glorieux@be.ibm.com

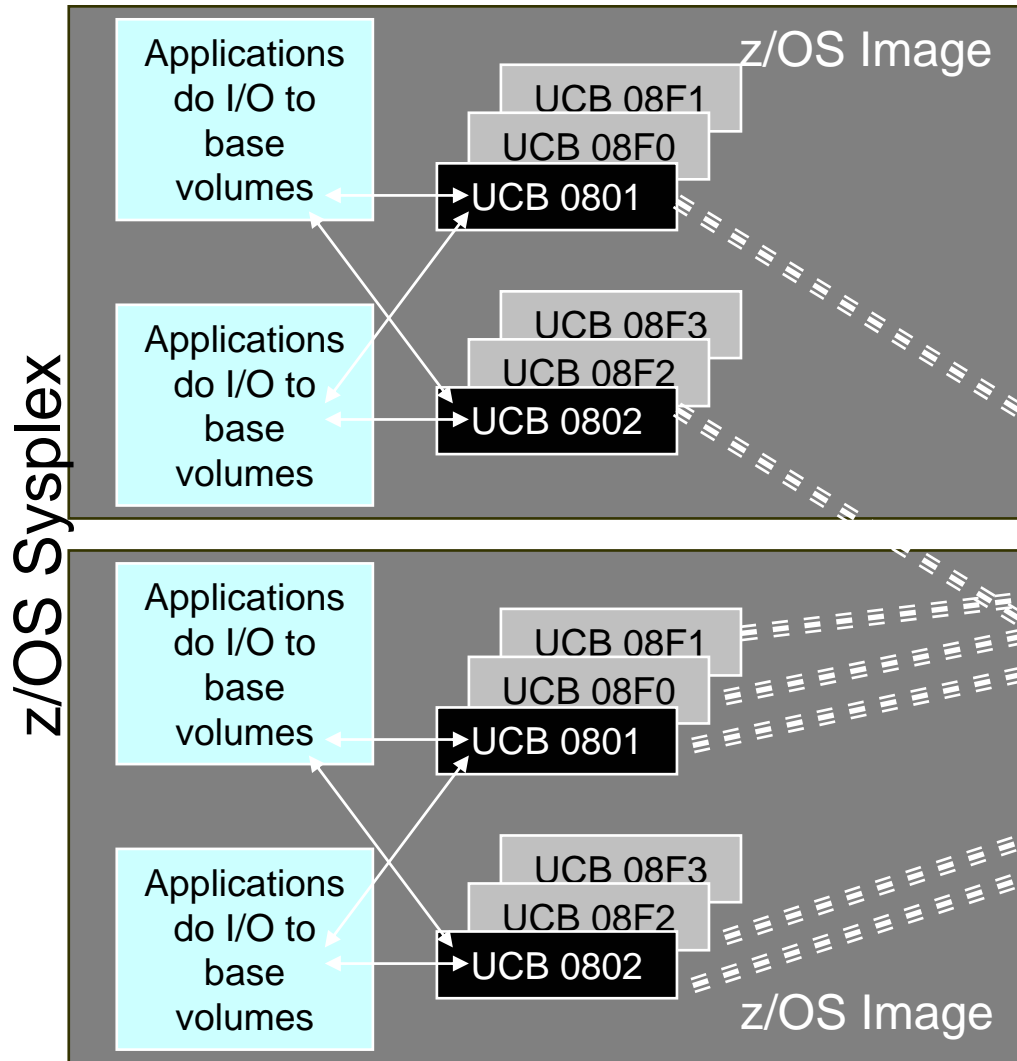
IO without PAV



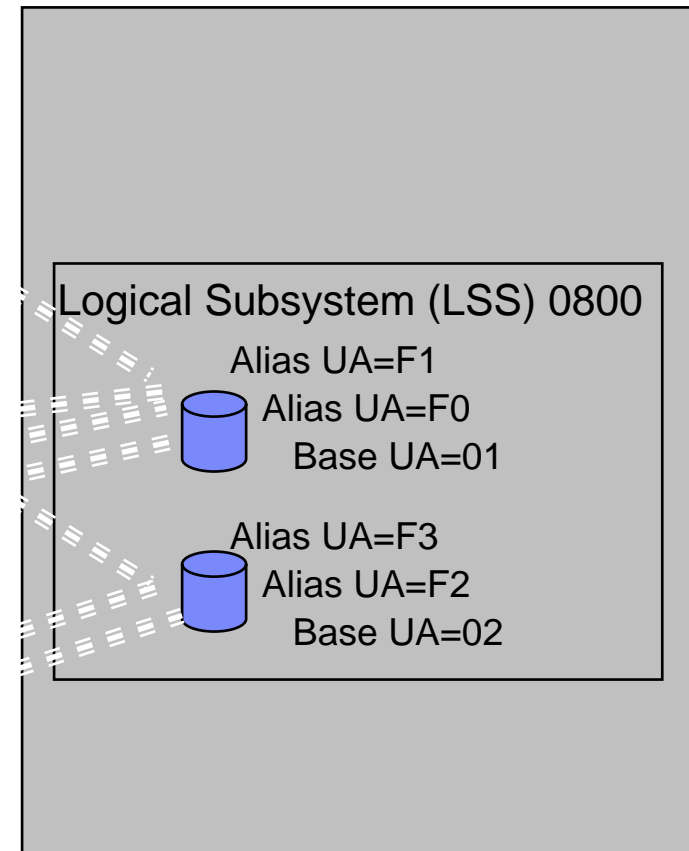
DS8000 Storage Server



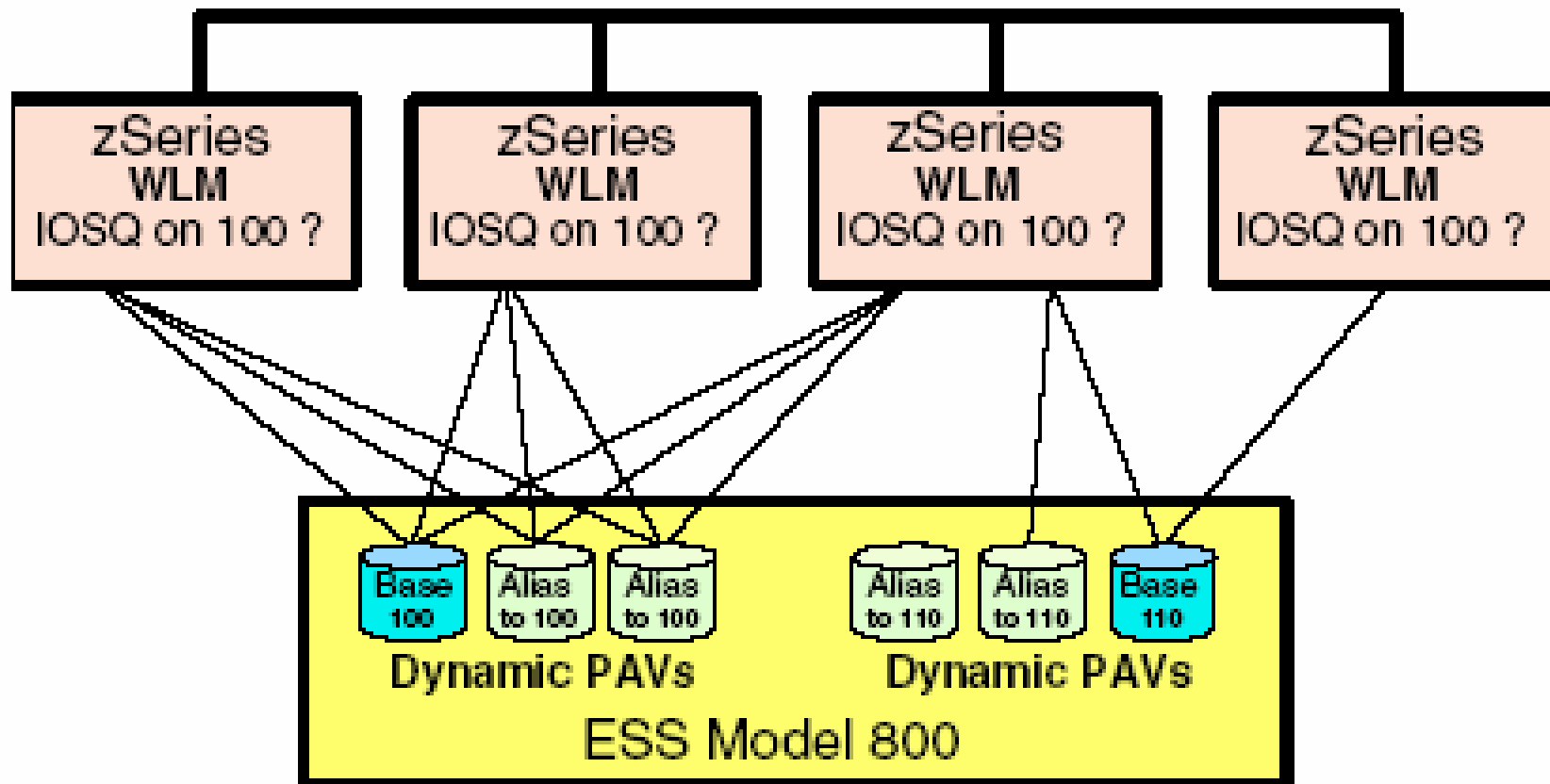
Parallel Access Volumes - Today



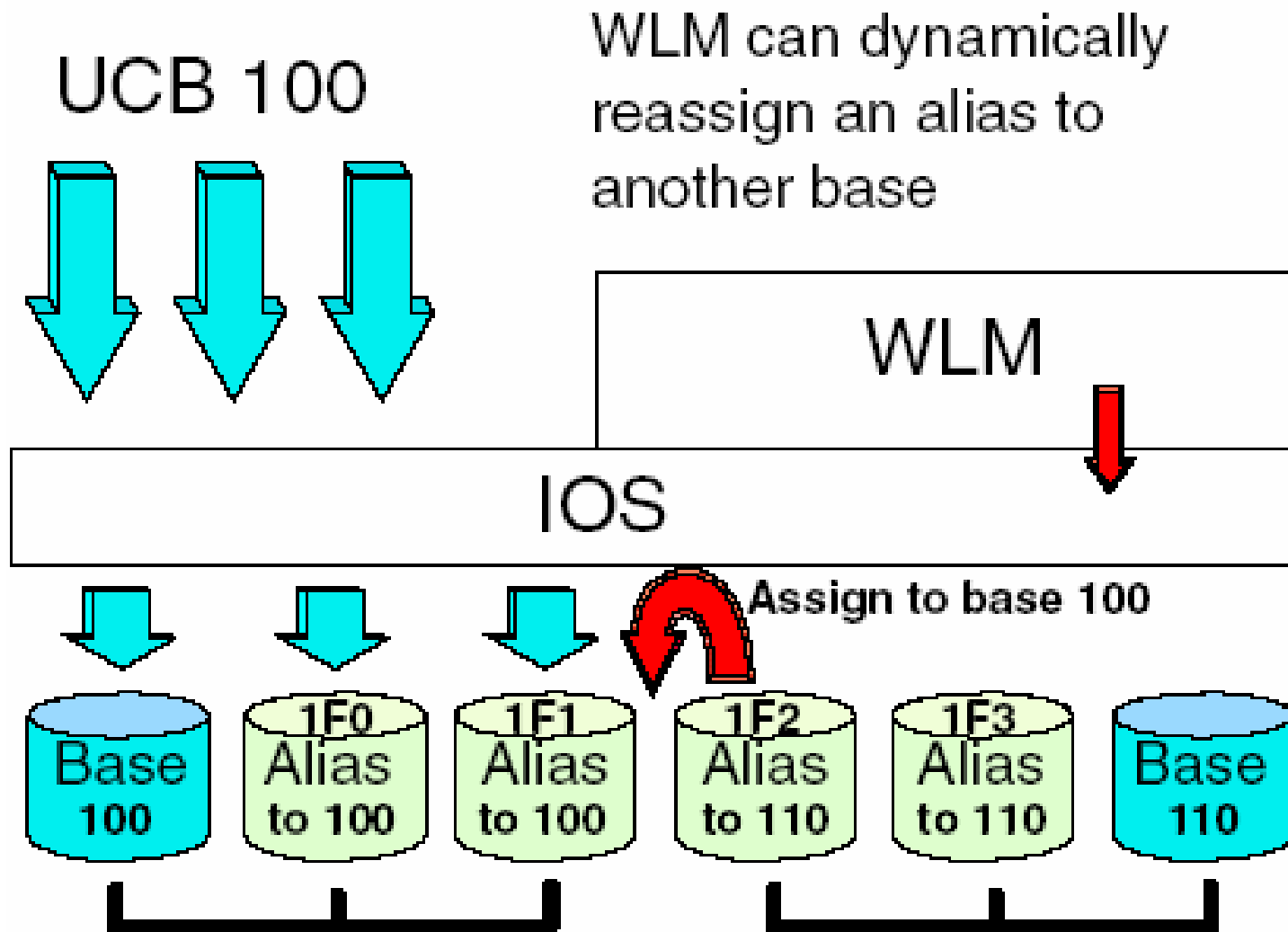
DS8000 Storage Server



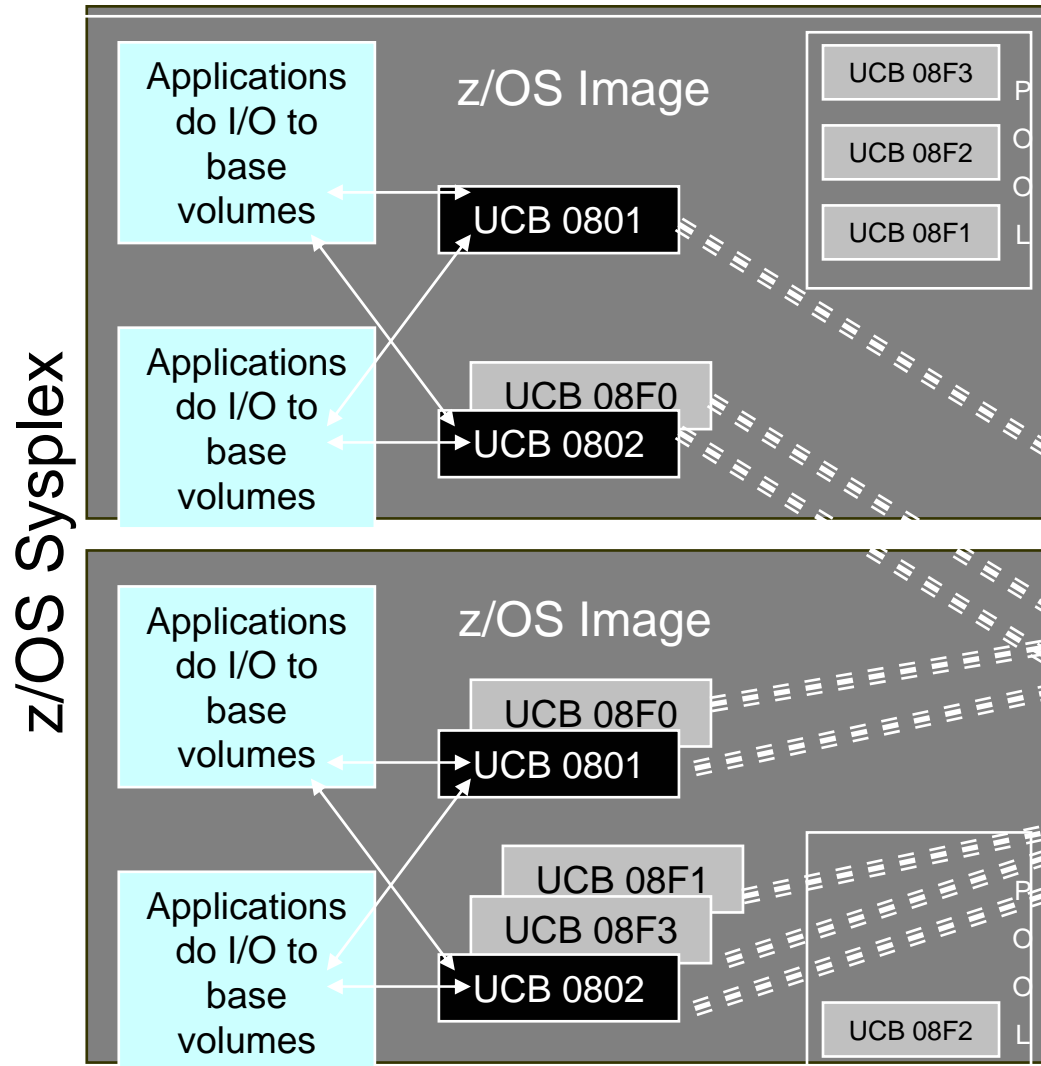
Dynamic PAV



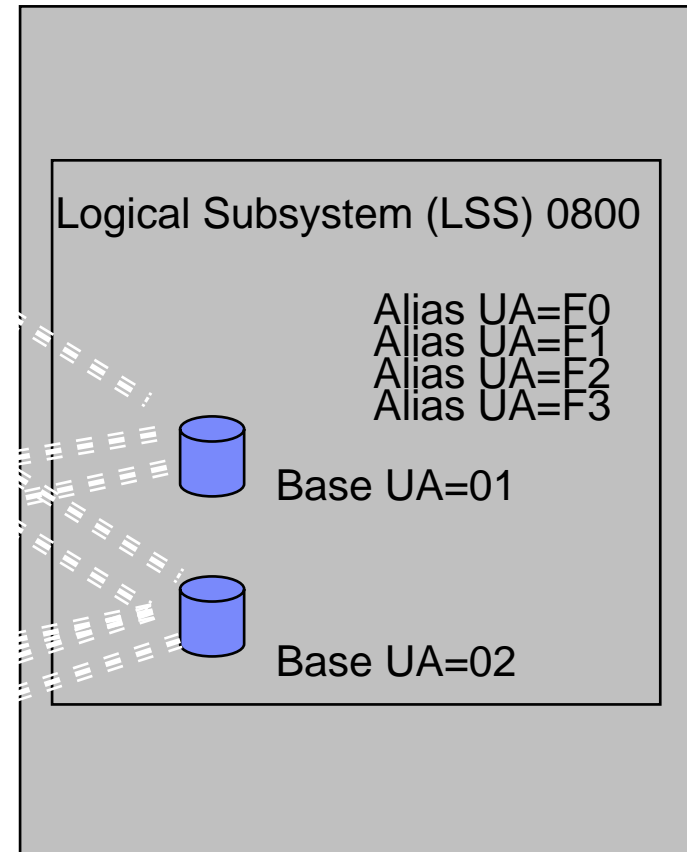
Dynamic PAV



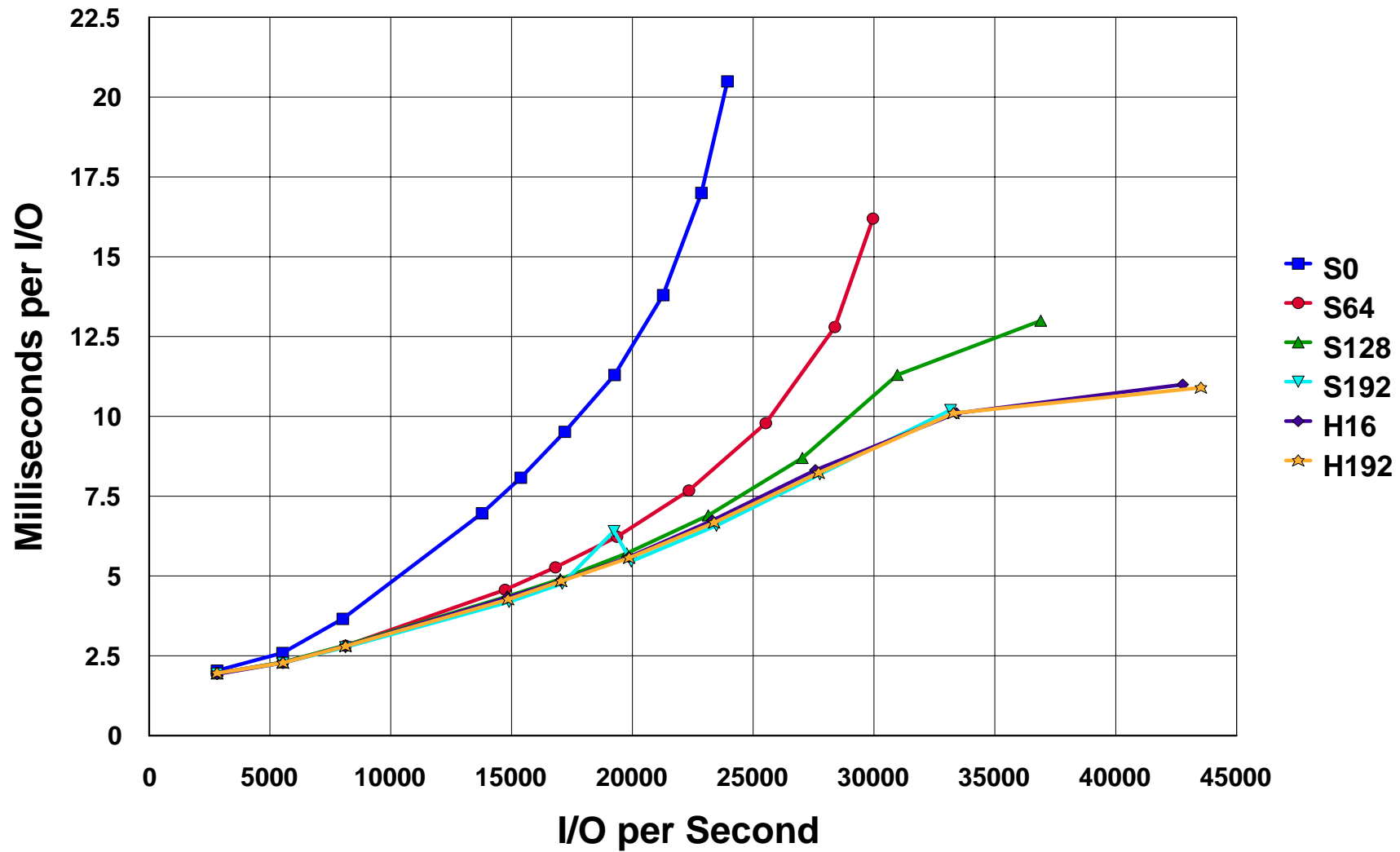
HyperPAV



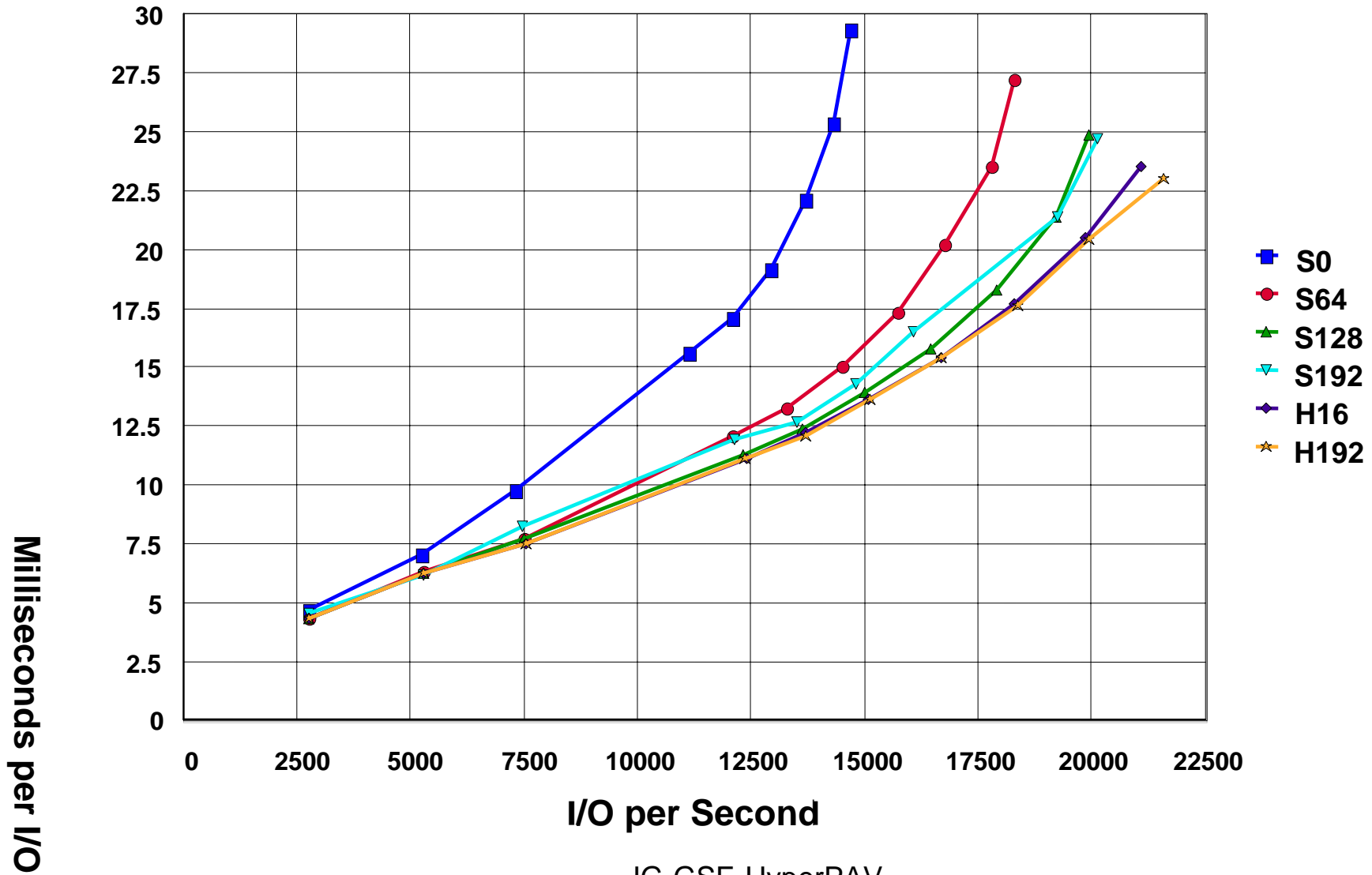
DS8000 Storage Server



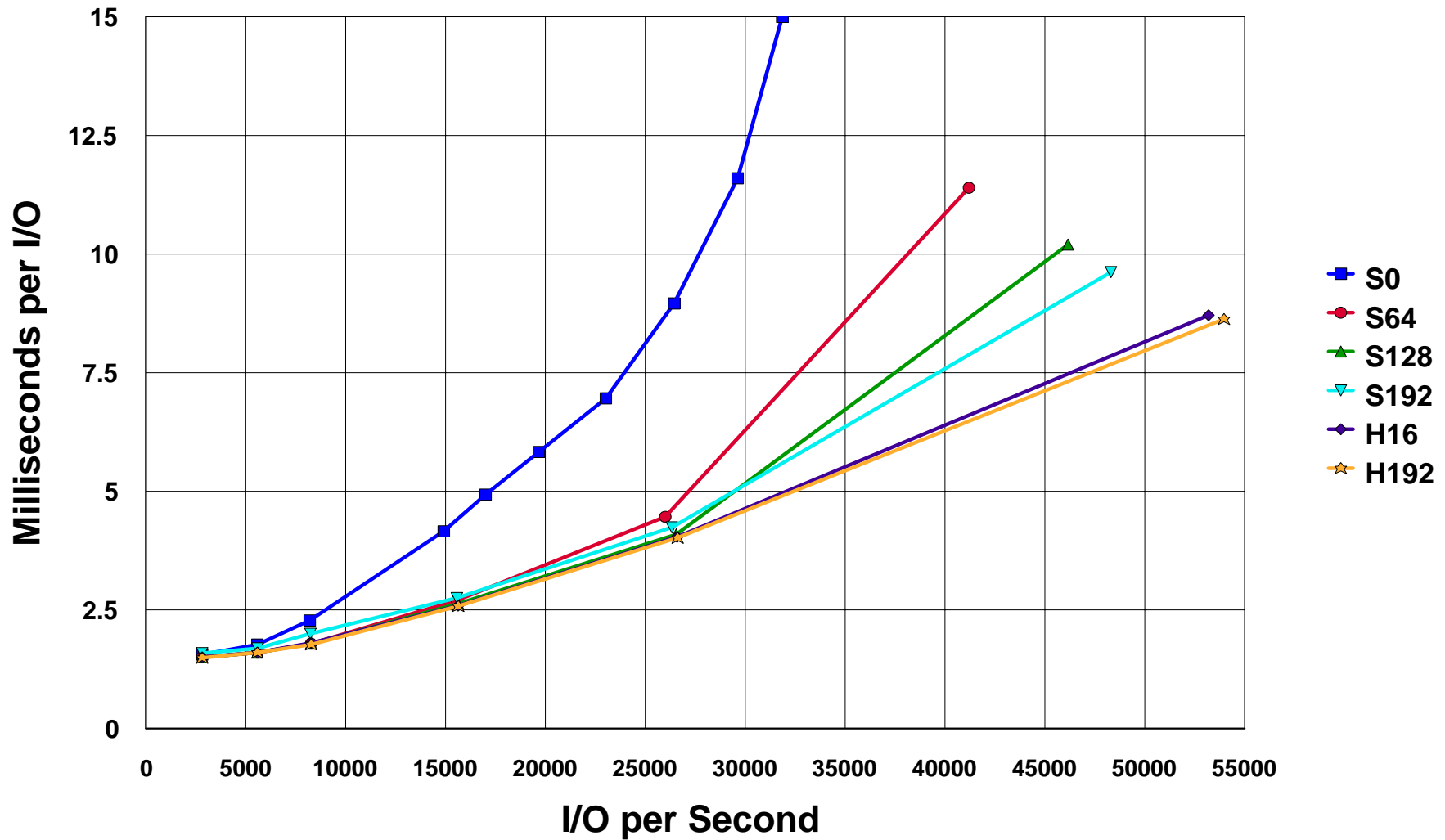
Static PAV vs HyperPAV Cache Standard



Static PAV versus HyperPAV – Cache Hostile

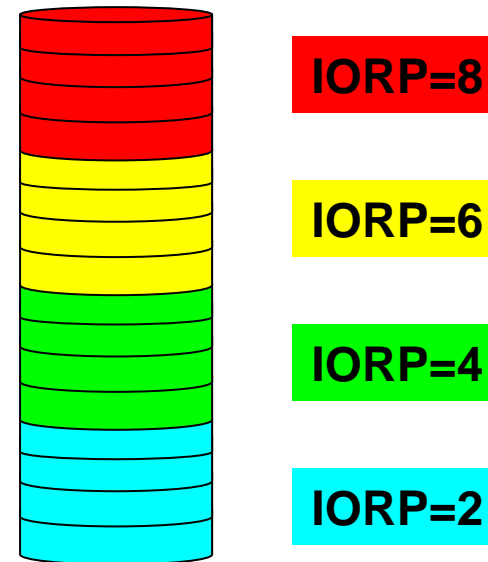


Static PAV vs HyperPAV Cache Friendly



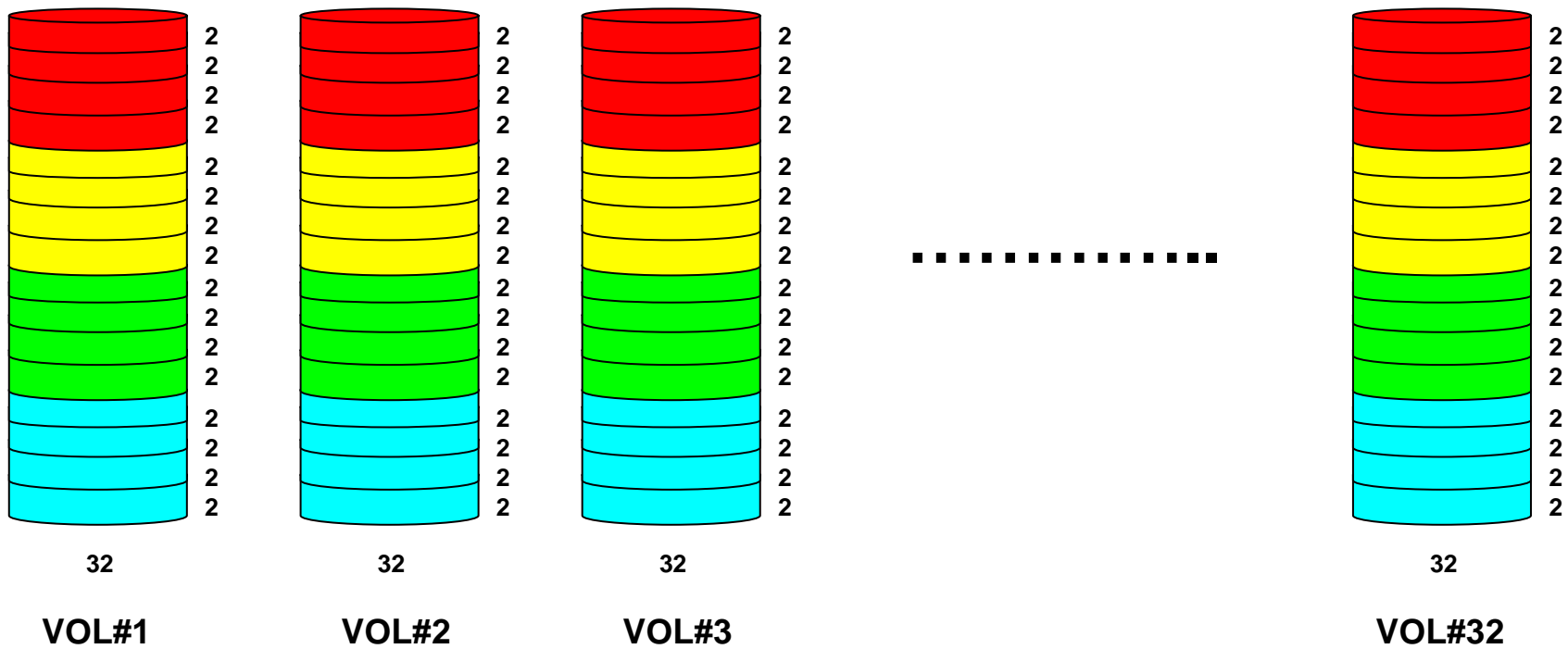
Experiments with the PAI/O Driver for z/OS

- z9 Processor
- z/OS 1.8
- DS8300T
- 8 FICON Express4 Channels
- 1 LCU
- 32 3390-3
- 16 101-cyl datasets per volume



Source : <http://www.perfassoc.com>

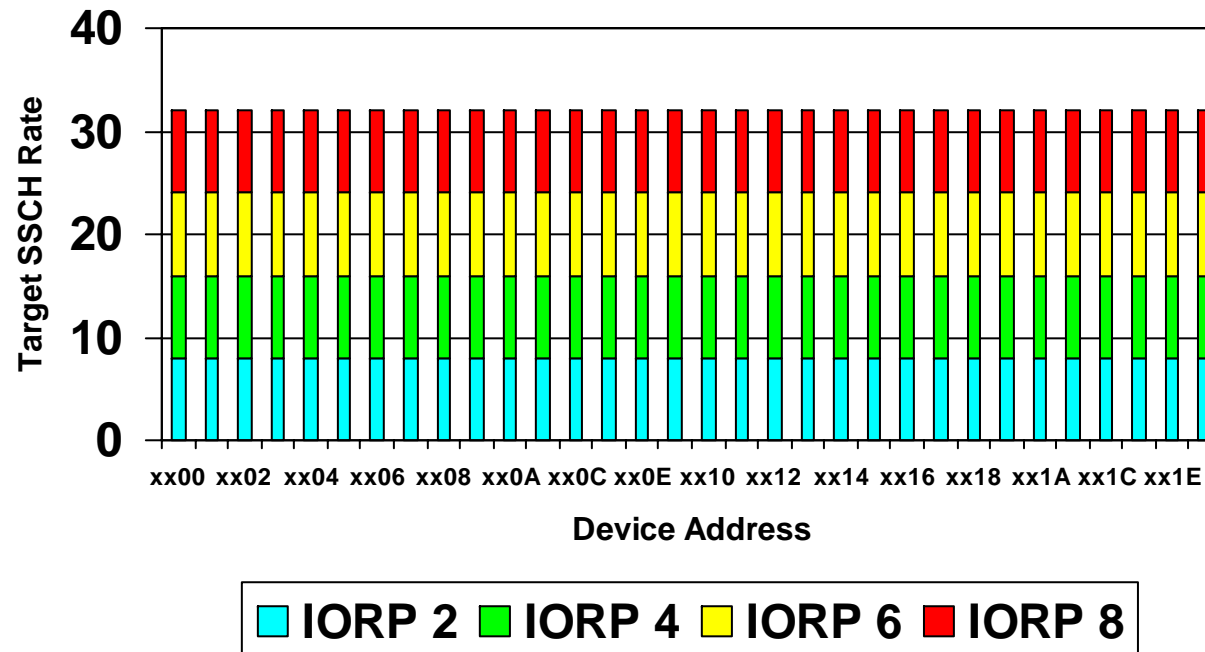
Experiments with the PAI/O Driver for z/OS ...



Source : <http://www.perfassoc.com>

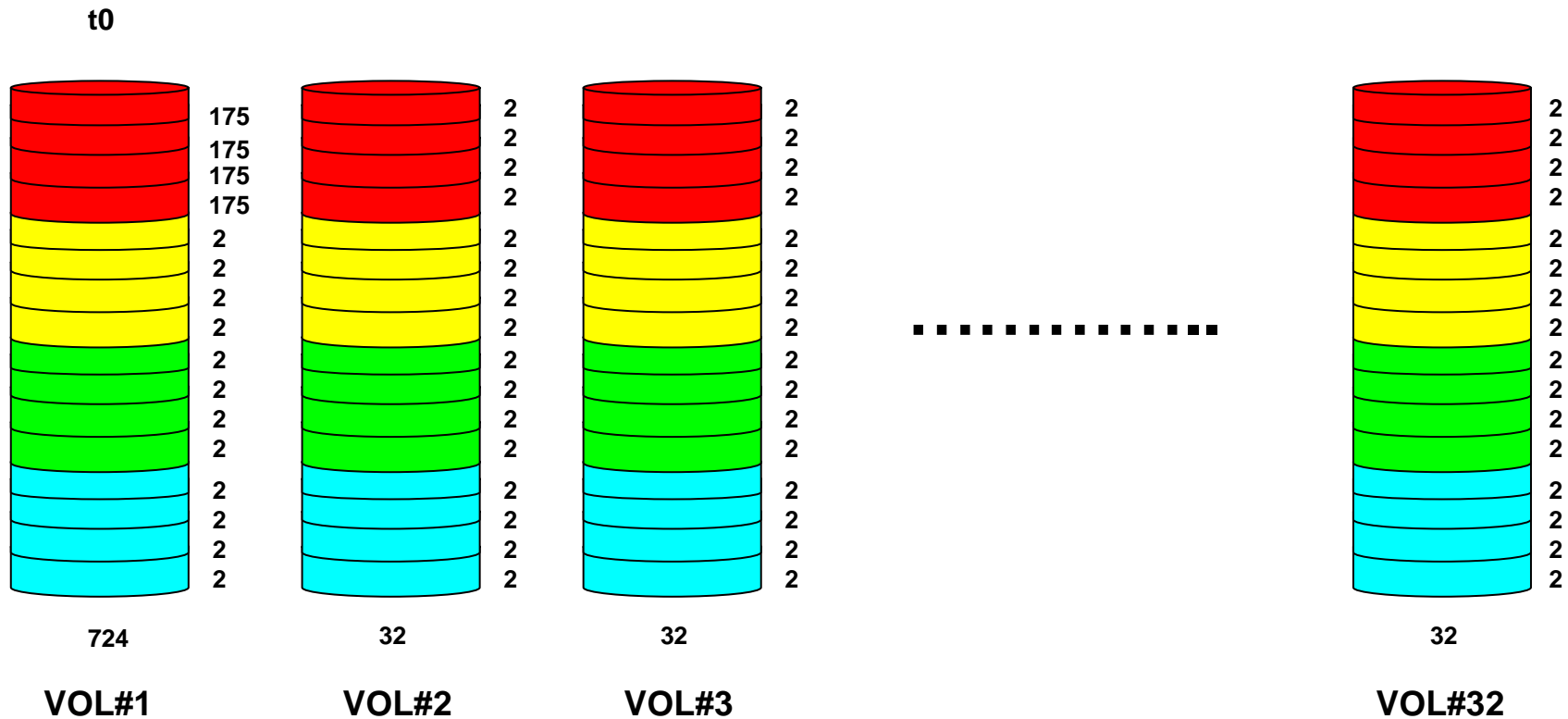
Experiments with the PAI/O Driver for z/OS ...

Device Address Target SSCH Rates



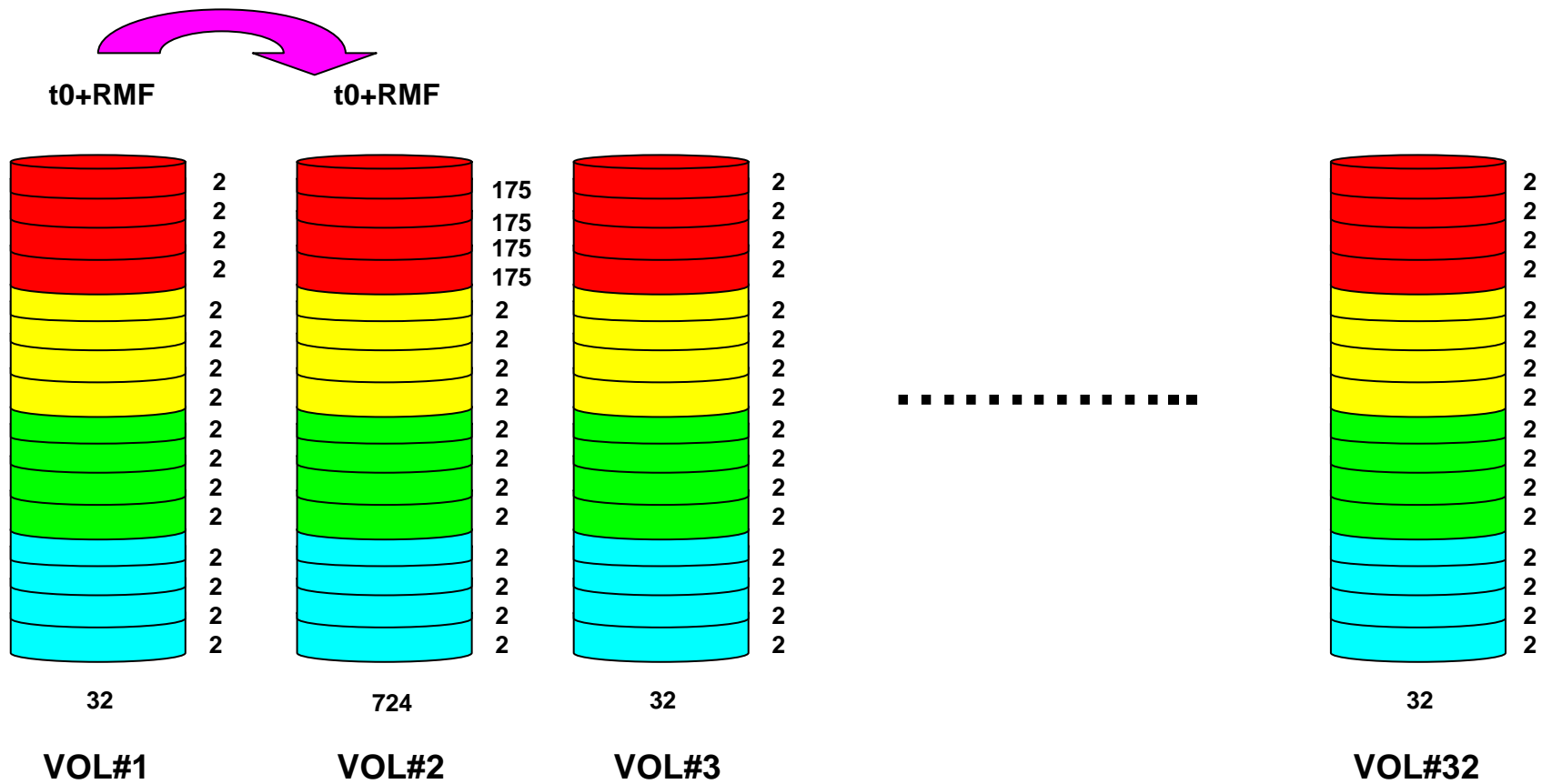
Source : <http://www.perfassoc.com>

Experiments with the PAI/O Driver for z/OS ...



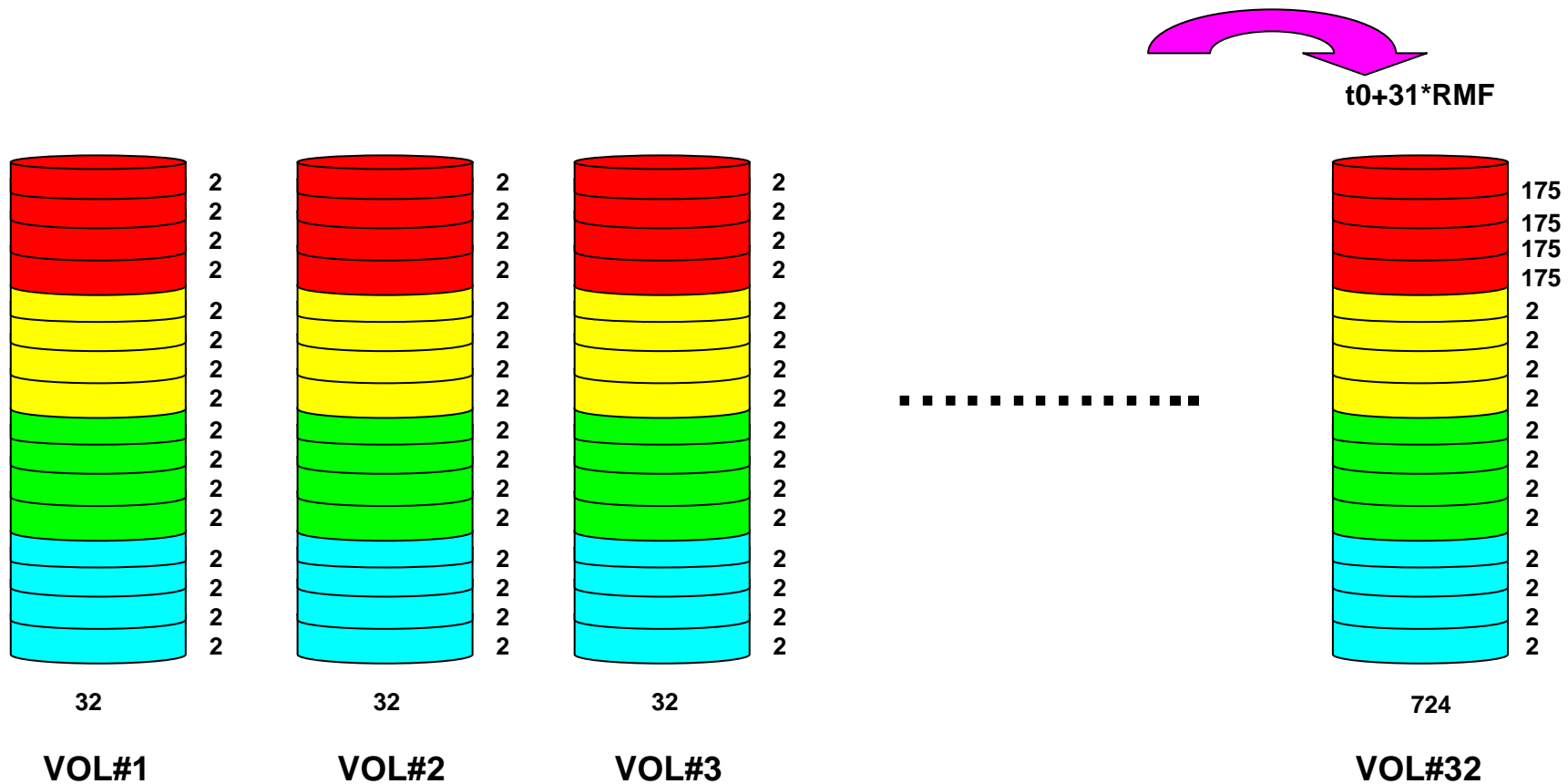
Source : <http://www.perfassoc.com>

Experiments with the PAI/O Driver for z/OS ...



Source : <http://www.perfassoc.com>

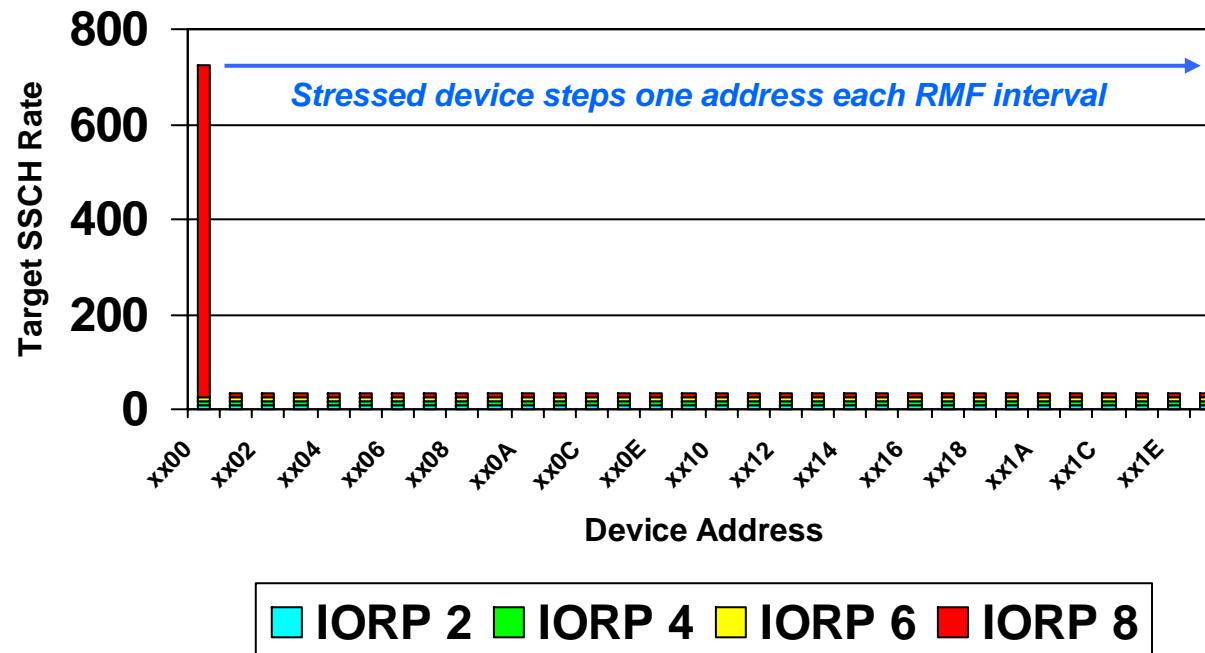
Experiments with the PAI/O Driver for z/OS ...



Source : <http://www.perfassoc.com>

Experiments with the PAI/O Driver for z/OS ...

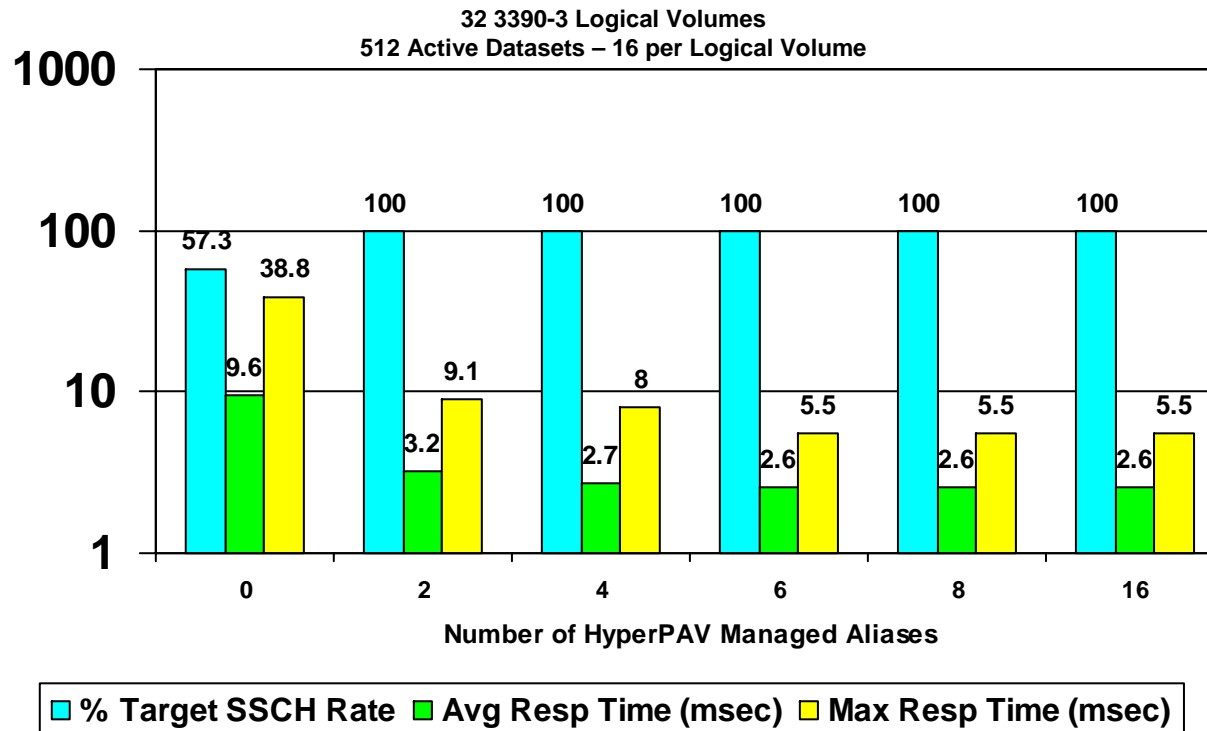
Device Address Target SSCH Rates



Source : <http://www.perfassoc.com>

Experiments with the PAI/O Driver for z/OS ...

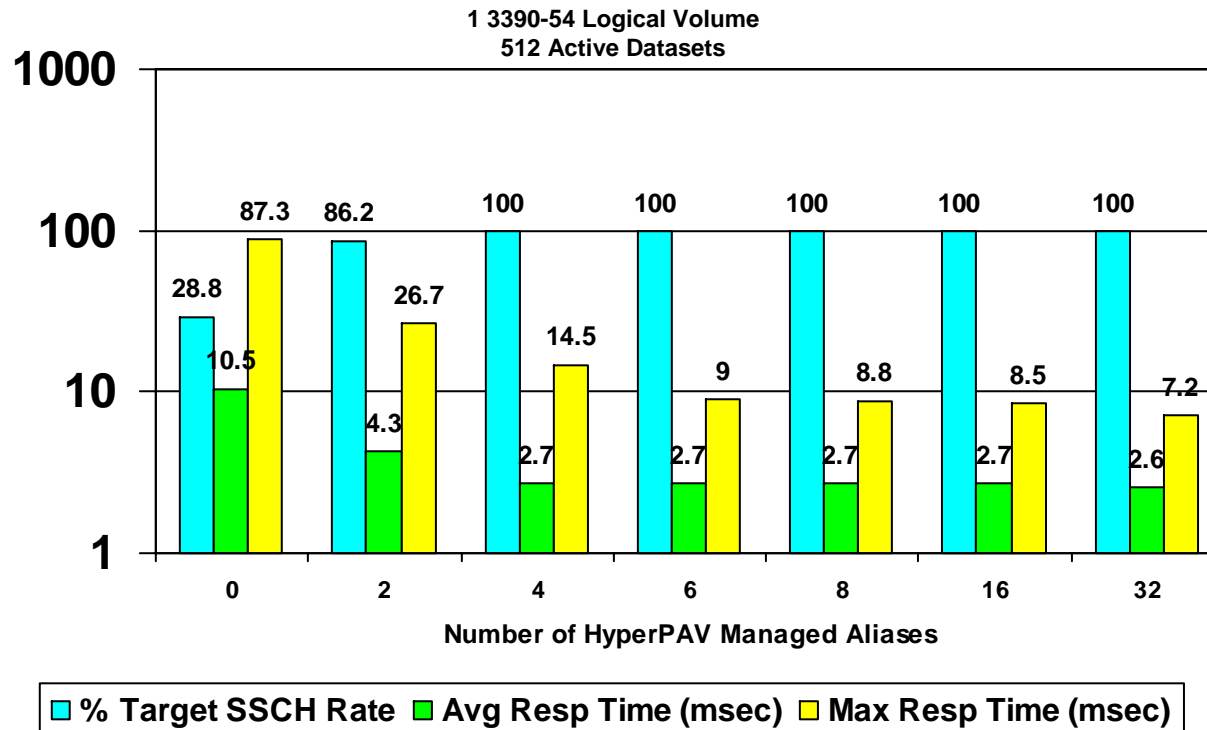
HyperPAV Agility Experimental Results



Source : <http://www.perfassoc.com>

Experiments with the PAI/O Driver for z/OS ...

HyperPAV Large Volume Experimental Results



Source : <http://www.perfassoc.com>

HyperPAV – System Requirements

- Hardware
 - DS8000 Bundle Version 6.2.4
 - DS8000 FCP/FICON Host Adapters
 - Licensed Features
 - FICON/ESCON Attach (Turbo)
 - DS8000 #0702 and #7090
 - 239x-LFA #7090 or 2244-OEL #7090
 - PAV
 - DS8000 #0780 and #78xx
 - 239x-LFA #78xx or 2244-PAV #78xx
 - HyperPAV
 - DS8000 #0782 and #7899
 - 239x-LFA #7899 or 2244-PAV #7899
- Software
 - z/OS 1.6 plus
 - IOS support
 - OA13915
 - DFSMS support (DFSMS, SMS, AOM, DEVSERV)
 - OA13928, OA13929, OA14002, OA14005, OA17605, OA17746
 - WLM support
 - OA12699
 - GRS support
 - OA14556
 - ASM support
 - OA14248
 - RMF
 - OA12865
 - Optionally Fixes for GDPS/Omegamon

HyperPAV – z/OS Options and Commands

- **SYS1.PARMLIB(IECIOSxx)**

- HYPERPAV=YES|NO|BASEONLY

- YES – Attempt to initialize LSSes in HyperPAV mode

- NO – Do not attempt to initialize LSSes in HyperPAV mode

- BASEONLY – Attempt to initialize LSSes in HyperPAV mode, but only start I/Os on base volumes

- **Enhanced Commands**

- SETIOS HYPERPAV=YES|NO|BASEONLY

- SET IOS=xx

- D M=DEV

- D IOS,HYPERPAV

- DEVSERV QPAV,dddd

HyperPAV – D M=DEV

```
SY1  d m=dev(0710)
SY1  IEE174I 23.35.49 DISPLAY M 835
DEVICE 0710  STATUS=ONLINE
CHP                10    20    30    40
DEST LINK ADDRESS  10    20    30    40
PATH ONLINE        Y     Y     Y     Y
CHP PHYSICALLY ONLINE Y   Y   Y   Y
PATH OPERATIONAL   Y     Y     Y     Y
MANAGED            N     N     N     N
CU NUMBER          0700 0700 0700 0700
MAXIMUM MANAGED CHPID(S) ALLOWED:  0
DESTINATION CU LOGICAL ADDRESS = 07
SCP CU ND          = 002107.000.IBM.TC.03069A000007.00FF
SCP TOKEN NED      = 002107.900.IBM.TC.03069A000007.0700
SCP DEVICE NED     = 002107.900.IBM.TC.03069A000007.0710
HYPERPAV ALIASES IN POOL  4
```

D M=DEV(dddd) where dddd is a base volume in a HyperPAV LSS

HyperPAV – D M=DEV

```
SY1  D M=DEV(0718)
```

```
SY1  IEE174I 23.39.07 DISPLAY M 838
```

```
DEVICE 0718      STATUS=POOLED HYPERPAV ALIAS
```

D M=DEV(dddd) where dddd is an alias device in a HyperPAV LSS

HyperPAV – D M=DEV

```

- SY1 d m=dev
- SY1 IEE174I 23.42.09 DISPLAY M 844
- DEVICE STATUS: NUMBER OF ONLINE CHANNEL PATHS
-   0 1 2 3 4 5 6 7 8 9 A B C D E F
- 000 DN 4 DN DN DN DN DN DN DN . DN DN 1 1 1 1
- 018 DN DN DN DN 4 DN DN DN DN DN DN DN DN DN DN
- 02E 4 DN 4 DN 4 8 4 4 4 4 4 4 4 DN 4 DN
- 02F DN 4 4 4 4 4 4 DN 4 4 4 4 4 DN DN 4
- 030 8 . . . . . . . . . . . . . . .
- 033 4 . . . . . . . . . . . . . . .
- 034 4 4 4 4 DN DN DN DN DN DN DN DN DN DN DN DN
- 03E 1 DN DN DN DN DN DN DN DN DN DN DN DN DN DN DN
- 041 4 4 4 4 4 4 4 4 AL AL AL AL AL AL AL AL
- 048 4 4 DN DN DN DN DN DN DN DN DN DN DN DN DN 4
- 051 4 4 4 4 4 4 4 4 UL UL UL UL UL UL UL UL
- 061 4 4 4 4 4 4 4 4 AL AL AL AL AL AL AL AL
- 071 4 4 4 4 DN DN DN DN HA HA DN DN . . .
- 073 DN DN DN . DN . DN . DN . DN . HA . HA .
- 098 4 4 4 4 DN 8 4 4 4 4 4 DN 4 4 4 4
- 0E0 DN DN 1 DN DN DN DN DN DN DN DN DN DN DN DN DN
- 0F1 1 DN DN DN DN DN DN DN DN DN DN DN DN DN DN DN
- FFF . . . . . . . . . . . HA HA HA HA
- ***** SYMBOL EXPLANATIONS *****
- @ ONLINE, PHYSICALLY ONLINE, AND OPERATIONAL INDICATORS ARE NOT EQUAL
- + ONLINE # DEVICE OFFLINE . DOES NOT EXIST
- BX DEVICE IS BOXED SN SUBCHANNEL NOT AVAILABLE
- DN DEVICE NOT AVAILABLE PE SUBCHANNEL IN PERMANENT ERROR
- AL DEVICE IS AN ALIAS UL DEVICE IS AN UNBOUND ALIAS
- HA DEVICE IS A HYPERPAV ALIAS

```

D M=DEV shows HA for HyperPAV aliases

RMF - DASD Activity Report

DIRECT ACCESS DEVICE ACTIVITY

PAGE 1

z/OS **Column PAV changed** SYSTEM ID S5A DATE 03/20/2006 INTERVAL 14.58.578
 RPT VERSION V1R5 RMF TIME 03.15.01 CYCLE 1.000 SECONDS

TOTAL SAMPLES = 25 CR-DATE: 03/19/2006 CR-TIME: 22.47.41 ACT: ACTIVATE

STORAGE GROUP	DEV NUM	DEVICE TYPE	VOLUME SERIAL	PAV	LCU	ACTIVITY RATE	AVG RESP TIME	AVG IOSQ	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
THRASH1	1000	33903	MM1000	1.0H	000D	0.092	0.6	0.0	0.0	0.0	0.2	0.0	0.4	0.00	0.00	0.0	0.0	100.0	0.0
THRASH1	1001	33903	MM1001	1.0H	000D	0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.0	0.0	100.0	0.0
MBOCA01	1002	33903	MM1002	1.0H	000D	0.107	0.7	0.0	0.0	0.0	0.2	0.3	0.2	0.00	0.01	0.0	0.0	100.0	0.0
MBOCA01	1003	33903	MM1003	1.3H	000D	155.552	3.2	0.0	0.1	0.0	0.2	0.5	2.5	31.08	37.26	0.1	15.7	100.0	0.0
MBOCA01	1004	33903	MM1004	1.2H	000D	153.981	3.1	0.0	0.1	0.0	0.2	0.5	2.4	30.52	36.80	0.0	15.6	100.0	0.0
MBOCA01	1005	33903	MM1005	1.2H	000D	154.219	3.1	0.1	0.1	0.0	0.2	0.5	2.4	30.39	36.41	0.1	15.5	100.0	0.0
MBOCA01	1006	33903	MM1006	1.2H	000D	152.245	3.2	0.0	0.1	0.0	0.2	0.5	2.5	31.27	37.63	0.0	15.4	100.0	0.0
MBOCA01	1007	33903	MM1007	1.2H	000D	160.826	3.2	0.1	0.1	0.0	0.2	0.5	2.4	31.22	37.45	0.0	17.2	100.0	0.0
MBOCA01	1008	33903	MM1008	1.3H	000D	154.879	3.1	0.1	0.1	0.0	0.2	0.5	2.4	29.15	35.13	0.0	15.6	100.0	0.0
MBOCA01	1009	33903	MM1009	1.3H	000D	155.022	3.2	0.0	0.1	0.0	0.2	0.5	2.5	30.60	36.72	0.0	15.5	100.0	0.0
MBOCA01	100A	33903	MM100A	1.3H	000D	154.051	3.3	0.1	0.1	0.0	0.2	0.5	2.5	29.04	34.66	0.0	15.5	100.0	0.0
MBOCA01	100B	33903	MM100B	1.0H	000D	0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.0	0.0	100.0	0.0
MBOCA01	100C	33903	MM100C	1.0H	000D	0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.0	0.0	100.0	0.0
MBOCA01	100D	33903	MM100D	1.0H	000D	0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.0	0.0	100.0	0.0
MBOCA01	100E	33903	MM100E	1.0H	000D	0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.0	0.0	100.0	0.0
MBOCA01	100F	33903	MM100F	1.0H	000D	0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.0	1.0	100.0	0.0
			LCU		000D	77.555	1.7	0.0	0.1	0.0	0.1	0.3	1.3	15.20	18.25	0.0	7.3	100.0	0.0

'H' means HyperPAV

RMF – I/O Queuing Report

I/O QUEUING ACTIVITY															PAGE	1													
z/OS V1R8			SYSTEM ID S5E			DATE 03/27/2006			INTERVAL 15.00.001																				
			RPT VERSION V1R8 RMF			TIME 03.45.00			CYCLE 1.000 SECONDS																				
TOTAL SAMPLES =		900	IODF = 25		CR-DATE: 03/21/2006		CR-TIME: 22.47.41		ACT: ACTIVATE																				
- INITIATIVE QUEUE -															----- IOP UTILIZATION -----					-- % I/O REQUESTS RETRIED --					----- RETRIES / SSCH -----				
IOP	ACTIVITY	AVG Q	% IOP	I/O START	INTERRUPT	CP	DP	CU	DV	CP	DP	CU	DV	CP	DP	CU	DV												
	RATE	LNTH	BUSY	RATE	RATE	ALL	BUSY	BUSY	BUSY	ALL	BUSY	BUSY	BUSY	ALL	BUSY	BUSY	BUSY												
00	503.132	0.05	3.64	503.029	536.330					2.02	2.00	0.02	0.00	0.00	0.00	0.00	0.00												
01	123.609	0.00	1.92	123.582	113.414					5.42	6.36	0.06	0.00	0.00	0.00	0.00	0.00												
SYS	626.740	0.04	2.78	626.611	649.744					2.89	2.86	0.03	0.00	0.00	0.00	0.00	0.00												
LCU	CU	DCM GROUP	CHAN	CHPID	% DP	% CU	AVG CUB	AVG CMR	CONTESTION	DELAY Q	CLS	HPAV	WAIT	MAX															
		MIN MAX DEF	PATHS	TAKEN	BUSY	BUSY	DLY	DLY	RATE	LNTH	DLY	WAIT	MAX																
000B	1000		90	991.48	0.00	0.00	0.0	2.2																					
			91						OFFLINE																				
			92	544.85	0.00	0.00	0.0	3.5																					
			*	1536.3	0.00	0.00	0.0	2.7	0.000	0.00	27.8	1.241	182																
000C	1100		10	185.92	0.00	0.00	0.0	0.0																					
			11						PATH OFFLINE																				
		2 2 6		185.12	0.00	0.00	0.0	0.0																					
			*	371.04	0.00	0.00	0.0	0.0	0.000	0.00	0.9																		

New:
 ▶ HPAV Wait
 ▶ HPAV Max

HyperPAV - Migration

- **No HCD or DS8000 Logical Configuration Changes required**
 - On existing LSSs, assuming PAV and FICON are used today
- **HyperPAV deployment can be staged**
 - Load/Authorize HyperPAV feature on DS8000
 - Can run without exploiting this feature if necessary using z/OS PARMLIB option
 - Enable HyperPAV feature on z/OS images that want to utilize HyperPAV via PARMLIB or SETIOS command
 - Eventually... enable HyperPAV feature on all z/OS images in the sysplex and authorize licensed function on all attached DS8000s
 - Reduce the number of aliases defined

IBM HyperPAV – Designed for Easy Migration

- **WLM Dynamic alias tuning compatibility**
 - Ignores HyperPAV control units
 - Avoids SYSPLEX communications for devices on HyperPAV control units
 - Eliminates need for multi-system interlock (DST)
 - Mixed environment tolerated (within a system and within a SYSPLEX)
 - Manages non-HyperPAV_aliases for all systems in the SYSPLEX
 - Control units in mixed mode (some systems in HyperPAV, some in Base-PAV mode) only manages aliases in Base PAV mode

HyperPAV - Designed to Use Less

- HyperPAV needs fewer aliases and can significantly increase the number of addresses available for storage growth.
- Example Scenario:
 - **For customers using a ratio of aliases to bases of 3 to1 today, HyperPAV could enable them to use only 1/10th as many aliases**
 - **Without HyperPAV: 192 aliases per LSS with 64 bases today**
 - **With HyperPAV: 20 aliases per LSS with 236 bases for real storage use**
 - **This could triple the number of addresses that could be used for addressable storage - a 350% growth in addressable storage!**
- Use RMF and new PAV Analysis Tool to help determine alias usage/needs. Needs can vary from customer to customer

Predicting Number of Aliases Needed

- **Difficult to predict ratio of Aliases to Base addresses required to minimize the IOSQ Time**
- **If ratio is too large**
 - Limits the amount of physical volumes that can be addressed due to the addressing limit of 64K
- **If ratio is too small**
 - Higher IOSQ and Response Time
 - Impacts business commitments

Purpose of the PAV Analysis Tool

- **To help customer understand the Alias utilization in a dynamic PAV environment across LSS**
- **Allow customers to determine the right Alias to Base ratio for their workload**

How the Tool Works

- **Data collection**

- Obtains PAV-alias measurements periodically
- Writes data to QSAM dataset

- **Data analysis**

- SAS application
- Calculates measurements for each PAV-alias
- Displays PAV-alias usage over time in 3-D graphical representation

Tool Availability

- **Available now**
 - Requires z/OS 1.4.1 or higher release
- <http://www-03.ibm.com/servers/eserver/zseries/zos/unix/bpxa1ty2.html>
 - Look under PAV Analysis Tool
 - pav_analysis.doc
 - Documents how to use the Tool
 - Download and linkedit the 3 Object Codes
- **Tool will work for IBM and non-IBM DASD subsystems**

Data Collection

- **Job to run the Tool**

```
//STEP EXEC PGM=PAVMONTR,PARM='hhmmssth',TIME=NOLIMIT
//SYSPRINT DD SYSOUT=*
//SYSABEND DD SYSOUT=*
//PVDATA DD DISP=(,CATLG),DSN=XXX.PAV.OUTPUT,UNIT=SYSDA,
// LRECL=0,BLKSIZE=0,RECFM=FB,DSORG=PS,SPACE=(CYL,(30,30))
```

- PARM='00010000' means that PAVMONTR will rerun every 1 minute to collect the data

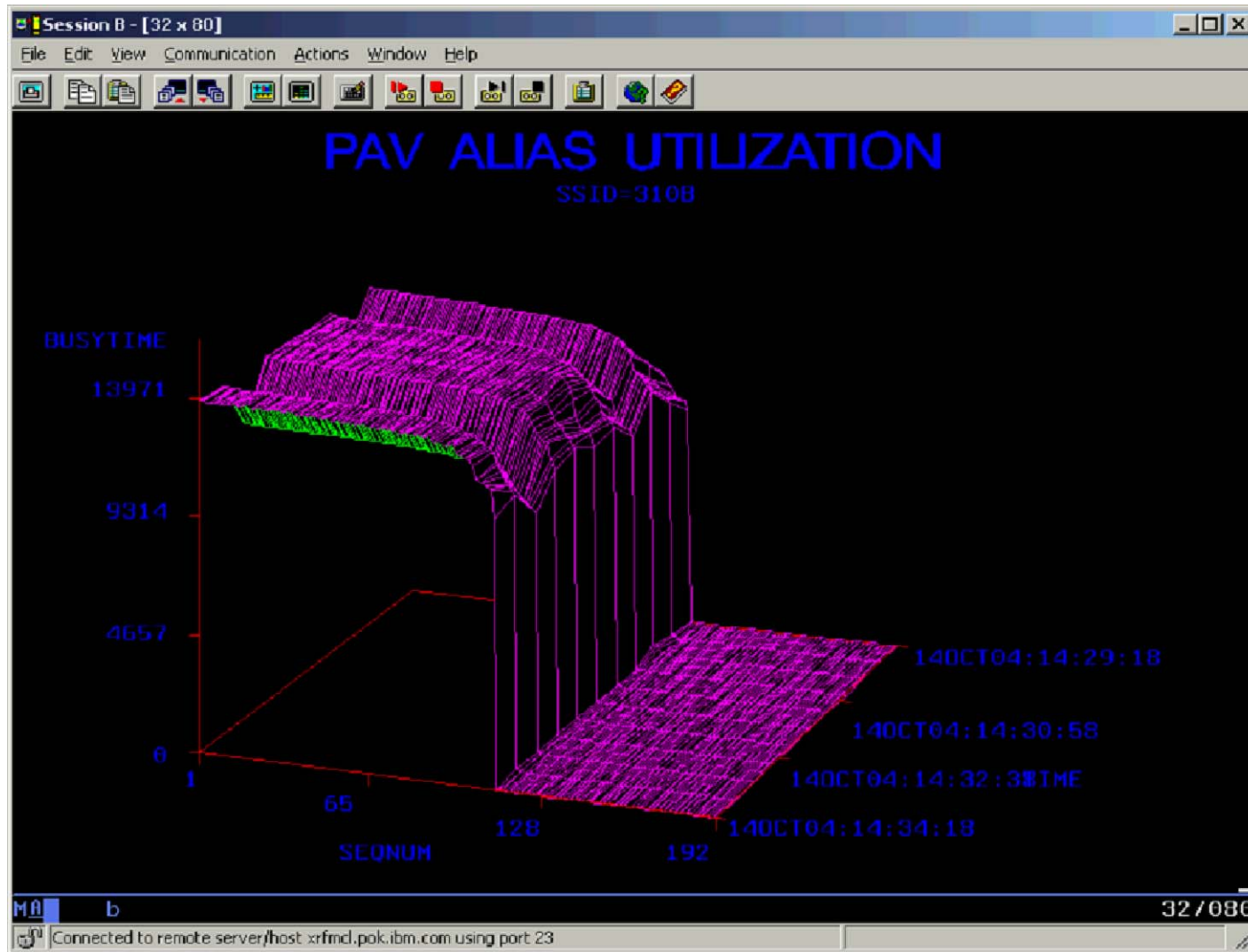
- **WTOR to allow the user to stop the Tool**

"PAV Monitor is running. Answer this WTOR to Terminate it"

Data Analysis

- **Output dataset is used as an input to the SAS program**
 - SAS program is documented in the pav_analysis.doc file
 - Data layout is also documented in the pav_analysis.doc file, so users can write their own data reduction program

PAV Analysis Tool Sample Output



The Answer

- **Currently Available z/OS Address Constraint Relief: Use one, two, or all three**
 - **Migrate to larger z/OS volume sizes**
 - Fewer objects to define and manage, less processing for fewer I/O resources, reduces storage allocation for devices
 - Considerations
 - Data migration to larger devices may be difficult, time consuming
 - **Leverage subchannel sets of new IBM System z9 servers**
 - Enables additional 64K set of subchannels, logical 5th digit for PAV-alias devices
 - Allows customer to define PAV-alias devices in a separate number space that does not consume device numbers used for PAV-base devices
 - Considerations
 - Must install new IBM z9 servers - may not be in cycle for server upgrade
 - Requires z/OS 1.7 – may be some impact to ISV applications, definitions, etc.
 - Migration of aliases will require updates to IODF
 - **Implement IBM System Storage DS8000 HyperPAV**
 - IBM DS8000 HyperPAV feature is not available from other storage vendors
 - Requires FICON attach, PAV and new HyperPAV feature
 - No new server hardware needed
 - No migration of data required
 - No impact to ISV applications
 - Cost effective, designed for easy migration/implementation

Using larger volume sizes

- **Benefits**

- Fewer objects to define and manage
- Less processing for fewer I/O resources
 - CF CHPID, VARY PATH, VARY DEVICE
 - Channel path recover, link recovery, reset event processing
 - CC3 processing
 - ENF Signals
 - RMF, SMF
- Number of physical resources: CHPIDs, Switches, CU ports, fibers
- Each device consumes real storage:
 - 768 bytes of real storage for UCB and related control blocks
 - 256 bytes of HSA
 - 1024 bytes/device * 64K devices = 64MB
 - 31 bit common storage constraints
- EOV processing to switch to the next volume of a sequential data set significantly slows the access methods

- **Considerations**

- Data migration to larger devices may be challenging, time consuming

IBM System z9 Subchannel Sets

- **Description**

- Install IBM System z9 servers to leverage subchannel sets
- Additional 64K set of subchannels, logical 5th digit for PAV-alias devices
- Allows customer to define PAV-alias devices in a separate number space that does not consume device numbers used for PAV-base devices

- **Benefit**

- On average 2x increase in the number of logical volumes
- Minimal impact to ISV software
- Minimal impact to customer I/O configuration definitions
- Allows customers to defer the decision to increase volume sizes
- Cloned symmetric I/O configurations across CECs suggests all systems exploit at the same time, but not required
 - Not required that all systems have equal performance

- **Considerations**

- **z/OS 1.7 and a New IBM System z9 Server are required for subchannel set exploitation**
- **Migration of aliases will require updates to IODF**
 - Once aliases are moved, duplicate 4-digit device numbers become possible

IBM DS8000 HyperPAV

▪ Description

- Heritage PAV design designed and implemented for ESCON attachments
 - Relatively static binding of PAV-alias to PAV-base
 - Dynamically Tunable by WLM
 - Complex rules of thumb for allocating the number of PAV-aliases
- HyperPAV maintains a shared pool of PAV-aliases
 - PAV-aliases bound to PAV-bases on an I/O request basis
 - Up to 10X increase in efficiency

▪ Benefit

- Up to 350% increase in the amount of data that can be addressed
- Minimal impact to ISV software
- Minimal impact to customer I/O configuration definitions
- Allows customers to defer the decision to increase volume sizes
- Cloned symmetric I/O configurations across CECs suggests all systems exploit at the same time, but not required
 - Not required that all systems have equal performance

▪ Considerations

- Requires FICON, IBM DS8000 microcode level 2.4 with HyperPAV and PAV features, z/OS 1.6 (HBB7709) plus
 - OA13915 – IOS support
 - OA13928, OA13929, OA14002, OA14005 – DFSMS support
 - OA12699 – WLM support
 - OA14556 – GRS support
 - OA14248 – ASM support
 - RMF - OA12865

The Benefits of IBM HyperPAV

- **Enables reduction in the number of required aliases**
 - **Can give back alias device numbers that can now be used for addressable storage**
 - **Provides relief from 64k device limit**
- **Helps z/OS to react more quickly to I/O load changes compared to WLM's Dynamic Alias Management**
- **Can increase I/O Parallelism**
- **Supports equivalent performance with fewer aliases**

End of Presentation