

z/OS 1.12 zFS Shared File System Update

Speaker Name: Ann Totten

Speaker Company IBM Corporation

Date of Presentation: Wednesday, August 4, 2010: 1:30-2:30 PM

Session Number: 7507

anntotten@us.ibm.com



SHARE in Boston

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

IBM
Language Environment
z/OS

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

UNIX is a registered trademark of The Open Group in the United States and other countries.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.



SHARE in Boston

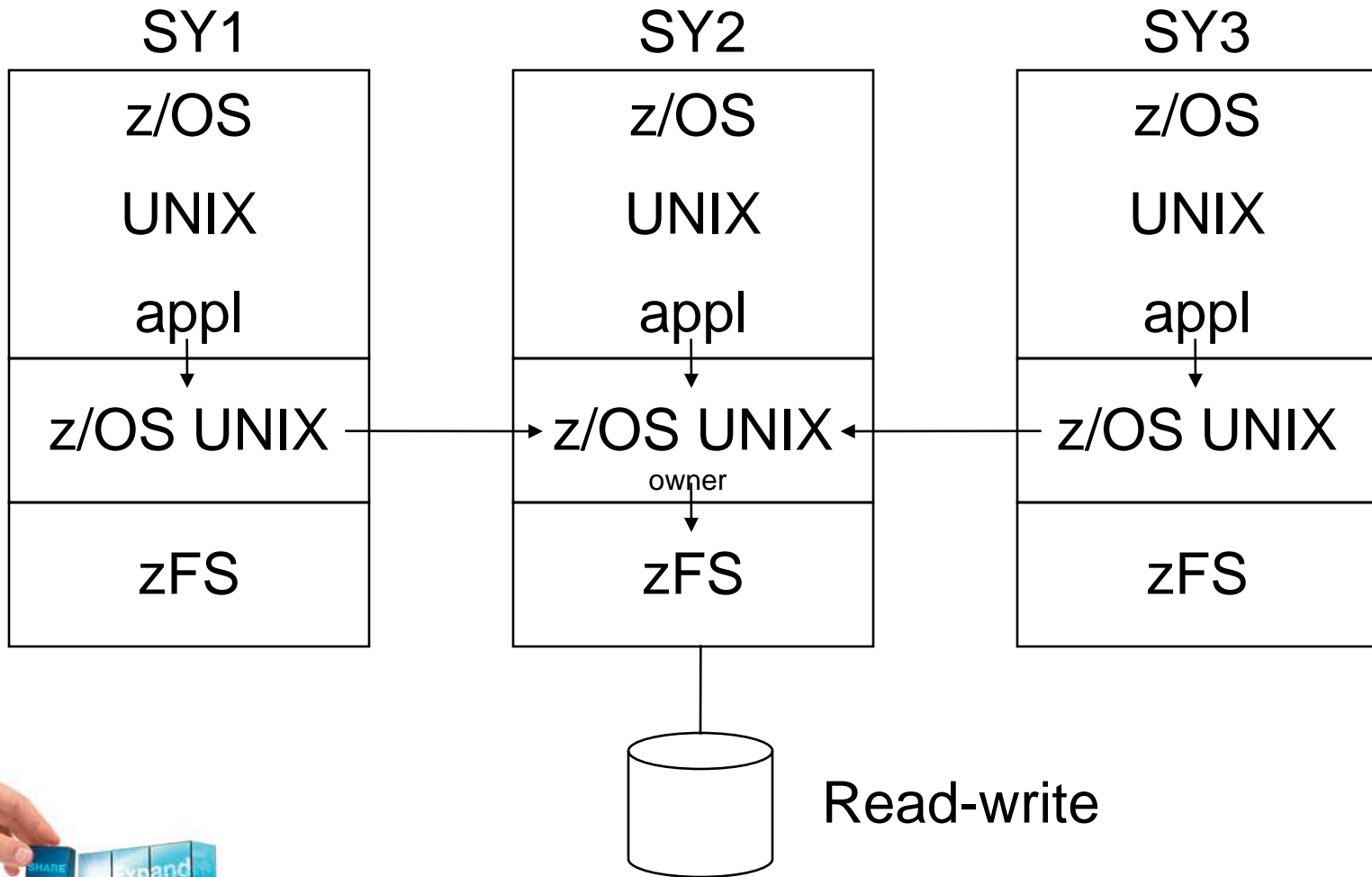
Session Topics

- Background on current shared file system environment.
- Actions required to migrate to z/OS Release 11 and 12.
- zFS shared file system changes new z/OS Release 11 and 12.
- Terminology on zFS sysplex-aware filesystems.
- System and shell commands.
- Health checks.

Background

- Basic shared file system support in releases before z/OS Release 11
 - Networking model provided through XCF services brings remote file data to the user.
 - This is called "function shipping".
 - Shared file system support is available through z/OS UNIX.

Shared file system environment R/W - review



Shared FS advantages

- **Greater user mobility**
- **Flexibility with file system balancing**
- **Consolidation of data**
- **One common BPXPRMxx for all systems**

Shared FS advantages for the administrator

- **Write to file systems from all systems in the shared file system group**
- **Have greater availability of data in the event of a system outage**
- **Have a common file system hierarchy on all systems**
- **Better manage file system placement**

Migration to V1R11-12, New Terminology

- **sysplex_admin_level (Interface level)**
 - **What it is:**
 - zFS XCF communication level in shared file system environment
 - **Where it is specified:**
 - IOEPRMxx member

```
sysplex_admin_level=1 | 2
```
 - **How to query:**
 - `MODIFY ZFS,QUERY,LEVEL`

Last line will show:

```
sysplex(admin-only) interface(1)
```

Toleration with zFS R9 and R10, Step 1

- Install APAR OA25026 on all z/OS V1R9 and z/OS V1R10 systems. This is a conditioning function APAR for zFS on z/OS V1R11. Make APAR OA25026 active on all systems through a rolling IPL. You are now running with zFS *sysplex_admin_level=1*.
- Running with *sysplex_admin_level=1* enables code that uses a simplification of the zFS XCF protocol between zFS members.
- Verify that you are running with the correct configuration by issuing the following on each image:

```
MODIFY ZFS,QUERY,LEVEL
```

```
IOEZ00639I zFS kernel: z/OS zSeries File System  
Version 01.10.00 Service Level OAnnnnn - HZFS3A0.  
Created on Tue Jan 6 20:02:25 EST 2009.  
sysplex(admin-only) interface(1)
```


Toleration with zFS R9 and R10, Step 2

- After APAR OA25026 is active on each z/OS V1R9 and V1R10 system, specify the *sysplex_admin_level=2* configuration option in the IOEPRMxx file(s). Make this level active on all z/OS V1R9 and V1R10 systems through another rolling IPL. This is the toleration function for zFS on z/OS V1R11. (The default for *sysplex_admin_level* is *sysplex_admin_level=1*).
- Verify that you are running with the correct configuration by issuing the following on each image:

```
MODIFY ZFS,QUERY,LEVEL
```

```
IOEZ00639I zFS kernel: z/OS zSeries File System  
Version 01.10.00 Service Level OAnnnnn - HZFS3A0.  
Created on Tue Jan 6 20:02:25 EST 2009.  
sysplex(admin-only) interface(2)
```

Toleration with zFS R9 and R10, Notes

- *sysplex_admin_level* is defaulted to 1 in APAR OA25026
- You cannot skip step 1 and do only step 2 if you are running any z/OS V1R9 or V1R10 systems that do not have APAR OA25026 installed and active.
- Your systems must be running zFS V1R9 and above to support a zFS V1R11 environment. In other words, zFS V1R11 is not compatible with V1R8 and below.
- The *sysplex_admin_level* configuration option cannot be specified dynamically through the **zfsadm config** command.
- When all images are running at *sysplex_admin_level=2*, the XCF protocol used by zFS is simplified.
- The value must be 2 on all members of the sysplex in order to bring z/OS V1R11 zFS into the shared file system environment.
- In V1R11 and later, the *sysplex_admin_level* setting is ignored and set to 3.

New Terminology in R11/R12

- **sysplex=on|off**
 - **What it is:**
 - Specifies whether zFS should run sysplex-aware.
 - On specifies sysplex-aware for read-write file systems.
 - Off specifies non sysplex-aware for read-write file systems.
 - **Where it is specified:**
 - IOEPRMxx member

sysplex=on
 - **Default:**

sysplex=off
 - **How to query:**
 - `MODIFY ZFS,QUERY,LEVEL`
 - `zfsadm configq -syslevel`

Example output: `sysplex(file) interface(3)`

More on Terminology in R11/R12

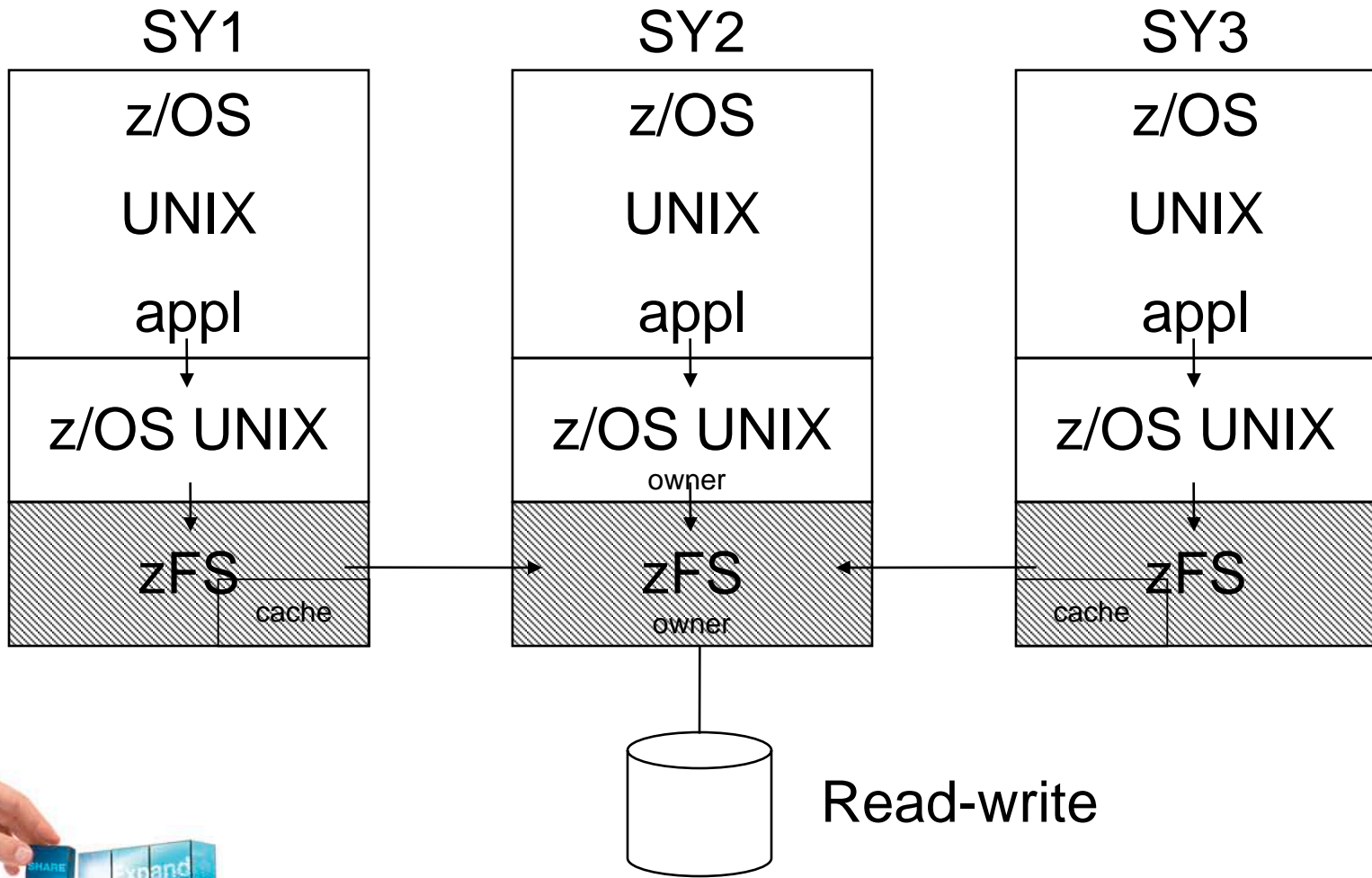
- **z/OS UNIX owner**
 - System where the original mount is issued. Administrator can still control this via `/usr/sbin/chmount` or the other externals. Significant when the zFS R/W file system is mounted non sysplex-aware.
 - Use the `D OMVS,F`, `df -v`, `ishell` to query the z/OS UNIX owner.
- **zFS owner**
 - System where the actual i/o is issued when the zFS R/W file system is sysplex-aware.
 - At mount time, this is the same system as the z/OS UNIX owner.
 - It is independent of the z/OS UNIX owner.
 - zFS will move the owner based on i/o activity.
 - User `zfsadm lsaggr` command to view the zFS owner.

Sysplex Aware Capability Expanded in R11/R12



- **sysplex-aware**
 - A zFS file system can be mounted sysplex-aware or non sysplex-aware. When a file system is mounted sysplex-aware, it means that the file system is locally mounted on every system. When a R/W file system is mounted non sysplex-aware, this is functionally equivalent to z/OS UNIX file system sharing.
 - Term can be used to describe the way that zFS was started. If started with `sysplex=filesystem` or `sysplex=on`, then zFS is said to be “sysplex-aware”. That is, zFS is running capable of mounting file systems sysplex-aware.
- **Unchanged:**
 - Read-only mounted file systems are always sysplex-aware.
 - HFS mounted R/W in a shared file system environment are always non sysplex-aware.

zFS shared file system, R/W sysplex-aware



Benefits of Sysplex-Aware zFS

- **The zFS read-write file systems that are owned and accessed by any of the systems that are running zFS sysplex-aware typically see a performance benefit.**
- **The zFS sysplex sharing allow sysplex clients to cache data to reduce the number of calls sent to servers.**
- **zFS ownership move support allows ownership of aggregates to dynamically move to the highest usage system.**

zFS sysplex-aware on a file system basis

- **Enhanced function**
- **Provided via**
 - **zFS APAR OA29619**
and
 - **z/OS UNIX APAR OA29712**
- **Provides another option to implement zFS sysplex-aware.**
- **Allows you to specify that individual zFS file systems should be sysplex-aware rather than all zFS file systems being sysplex-aware.**

New Terminology provided by APAR OA29619

- **sysplex=filesystems**
 - **What it is:**
 - Specifies that zFS should run sysplex-aware and support sysplex-aware behavior on a file system basis.
 - **Where it is specified:**
 - IOEPRMxx member
sysplex=filesystems
 - **How to query:**
 - `MODIFY ZFS,QUERY,LEVEL`
 - `zfsadm configq -syslevel`
 - Example output: `sysplex(filesys,rwshare) interface(3)`
 - **IBM recommendation for R11 and later releases:**
sysplex=filesystems

New Terminology provided by APAR OA29619

- **sysplex_filesys_sharemode**

- **What it is:**

- Specifies if zFS read-write file systems will be mounted sysplex-aware when a MOUNT is issued on that system without specifying either NORWSHARE or RWSHARE in the MOUNT PARM.

- **Where it is specified:**

- IOEPRMxx member

`sysplex_filesys_sharemode=rwshare`

- **Can be dynamically changed:**

`zfsadm config -sysplex_filesys_sharemode rwshare`

- **Default if not specified:**

- norwshare

- **How to query:**

`MODIFY ZFS,QUERY,LEVEL`

`zfsadm configq -sysplex_filesys_sharemode`

New Terminology provided by APAR OA29619

- Mount parameter, NORWSHARE or RWSHARE
 - **What it is:**
 - Specifies whether a zFS read-write mounted file system will be mounted sysplex-aware or non-sysplex aware.
 - **Where it is specified:**
 - On the mount statement. Example:

```
# /usr/sbin/mount -f USSZFS.WORK01.ZFS -t zfs -o RWSHARE /zfs_wkld/work1
# df -Pkv /zfs_wkld/work1
```

Filesystem	1024-blocks	Used	Available	Capacity	Mounted on
USSZFS.WORK01.ZFS	1693440	17281	1676159	2%	/zfs_wkld/work1

ZFS, Read/Write, Device:1386, ACLS=Y

RWSHARE

```
File System Owner : NP4          Automove=Y          Client=N
Filetag : T=off    codeset=0
Aggregate Name : USSZFS.WORK01.ZFS
```

What else has changed in this output?

- How do you know if this file system is mounted as zFS sysplex aware?

```
# /usr/sbin/mount -f USSZFS.WORK01.ZFS -t zfs -o RWSHARE -d NP5 /zfs_wkld/work1
```

```
# df -Pkv /zfs_wkld/work1
```

Filesystem	1024-blocks	Used	Available	Capacity	Mounted on
USSZFS.WORK01.ZFS	1693440	17281	1676159	2%	/zfs_wkld/work1

```
ZFS, Read/Write, Device:1386, ACLS=Y
```

RWSHARE

```
File System Owner : NP5          Automove=Y          Client=N
```

```
Filetag : T=off  codeset=0
```

```
Aggregate Name : USSZFS.WORK01.ZFS
```

- What about the File System Owner?
 - Still the z/OS UNIX owner
- What about Automove setting?
 - Unchanged, z/OS UNIX still uses this when processing mount takeover
- What about Client=?
 - Yes, this is one way to tell if a system is accessing the z/FS directly.

How to determine if a zFS is sysplex-aware?

- **MODIFY ZFS,QUERY,FILESET (system command)**

```
IOEZ00438I Starting Query Command FILESETS.
```

File System Name	Aggr #	Flg	Operations
USSZFS.WORK01.ZFS	38	AMS	2

The Flg column indicates the aggregate status as attached (A) and either mounted (M), quiesced (Q), cloning (C), deleting a clone (D), or growing (G). If the aggregate is sysplex-aware for read-write, it will also indicate **sysplex-aware (S)**.

How to determine if a zFS is sysplex-aware?

zfsadm aggrinfo -long (shell command)

```
# zfsadm aggrinfo -long -aggr USSZFS.WORK01.ZFS
USSZFS.WORK01.ZFS (R/W COMP): 1676159 K free out of total 1693440
version 1.4
auditfid D6C5F8F2 F0F72168 0010
sysplex-aware
    209519 free 8k blocks;           7 free 1K fragments
    16952 K log file;              56 K filesystem table
    248 K bitmap file
```

How to determine what system is the zFS owner?

zfsadm lsaggr (shell command)

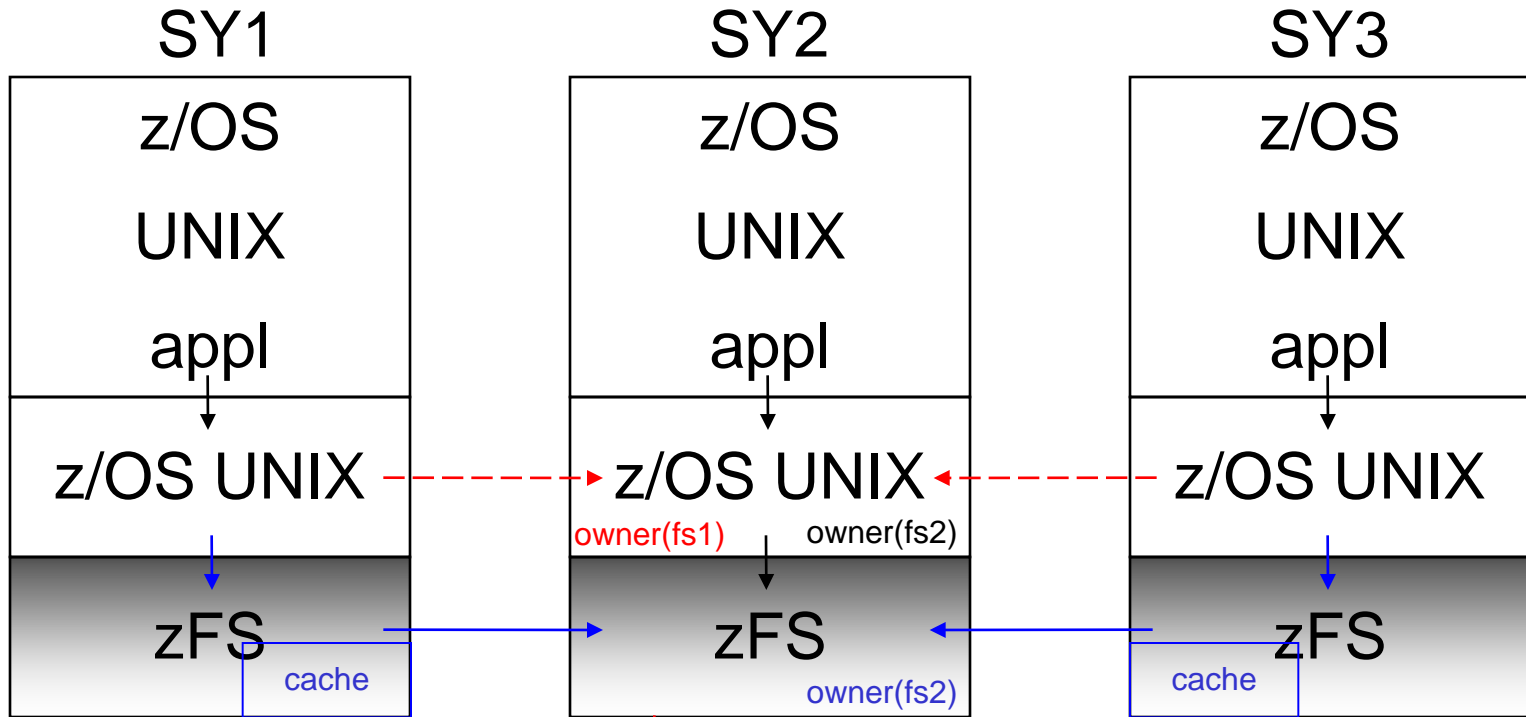
```
# zfsadm lsaggr | grep USSZFS.WORK01.ZFS
```

```
USSZFS.WORK01.ZFS
```

NP4

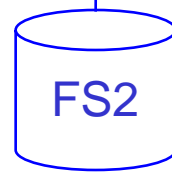
R/W

zFS Release 11/12 sysplex=filesystems support



- - -> Non-sysplex aware
 - -> Sysplex-aware

Read-write
NORWSHARE



Read-write
RWSHARE



zFS Release 11/12 sysplex=filesystem support

- Recommendation is to run sysplex=filesystem for greater flexibility in deciding which zFS file systems should be sysplex-aware.
- If you are running z/OS SMB server, the Fast Response Cache Accelerator support of the IBM HTTP Server for z/OS V5.3 or another server that uses the Register File Interest API – mount those file systems NORWSHARE.
- **Restriction:** While you can run some systems with zFS as sysplex-aware and some with zFS as non-sysplex aware, it is not recommended and **will cause a reduction in performance.**

zFS Release 11/12 sysplex=filesystem support

- What zFS file systems should be mounted sysplex-aware?

Typically, if you make a zFS read-write file system sysplex-aware, you see a performance improvement in most shared file system environments when accessing the data from a system that is not the zFS owner.

- R/W file systems where there are updates from more than one image.
- There is a tool available to aid in determining which zFS read-write file systems should be sysplex-aware. The tool and documentation are located at URL:
- <ftp://ftp.software.ibm.com/s390/zos/tools/wjsfsmon>
It can also be found at the z/OS UNIX tools and Toys site
- <http://www-03.ibm.com/systems/z/os/zos/features/unix/tools/>

Shutdown procedures and sysplex=filesys support

- In general, there are no changes to shutdown procedures when running with the new support, but note that when you issue
 - **MODIFY OMVS,SHUTDOWN**
 - or
 - **MODIFY OMVS,STOPPFS=ZFS**

You may notice that the new z/OS UNIX owner is different from the new zFS owner. This is not a problem as there is no requirement that these match.

Health Checks

- **ZOSMIGV1R11_ZFS_INTERFACELEVEL**
- **Description:**
Verifies that the system is running sysplex_admin_level=2 for zFS V1R11 toleration support.
- **Reason for check:**
zFS should be running at sysplex_admin_level=2 for all members of the sysplex. Once this is done, zFS V1R11 may be brought into the sysplex.
- Applies to z/OS V1R9 and V1R10.

Health Checks

- **USS_CLIENT_MOUNTS**

- **Description:**

This check will generate an exception when a file system is found that is function shipping but could be mounted locally.

- **Reason for check:**

File systems should not function ship if they can be mounted locally to avoid performance degradation.

- Check applies to z/OS V1R10 and later.

Publications

- z/OS UNIX System Services Planning (GA22-7800)
General Administration of z/OS UNIX file systems
- z/OS UNIX Command Reference (SA22-7802)
confighfs command for HFS
- z/OS MVS System Messages Volume 9 (IGF-IWM) (SA22-7639)
IGWxxxxt messages for HFS
- z/OS UNIX System Services Messages and Codes (SA22-7807)
z/OS UNIX return codes, z/OS UNIX reason codes, X'5Bxxxxr' reason codes for HFS
- z/OS Distributed File Service zSeries File System Administration (SC24-5989)
zFS Concepts and zfsadm command for zFS
 - <http://www-03.ibm.com/systems/z/os/zos/bkserv/r11pdf/#dfs>
- z/OS Distributed File Services Messages and Codes (SC24-5917)
IOEZxxxxt messages and X'EFxxxxr' reason codes for zFS
- z/OS Distributed File Service zSeries File System Implementation (SG24-6580)
 - Redbook available (updated March 2009 to include z/OS V1R10)
 - <http://www.redbooks.ibm.com/abstracts/sg246580.html?Open>
- z/OS Version 1 Release 8 Implementation (SG24-7265)
 - Redbook available (contains zFS updates for z/OS V1R8)
 - <http://www.redbooks.ibm.com/abstracts/sg247265.html?Open>
- z/OS DFSMS™ Access Method Services for Catalogs (SC26-7394)
IDCAMS utility
- z/OS DFSMS™ Storage Administration Reference (SC26-7402)
ADRDSSU utility for backup