

 #SHAREorg



Enabling z/VM V6.2 for Ensemble Management

Alan Altmark
Senior Managing z/VM and Linux Consultant
Alan_Altmark@us.ibm.com

Session 11776





Trademarks & Disclaimer

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DB2, e-business logo, ESCON, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, System Storage, System z9, VM/ESA, VSE/ESA, WebSphere, xSeries, z/OS, zSeries, z/VM.

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 Linux Torvalds in the United States and other countries. UNIX is a registered trademark of The Open Group in the United States and other countries. Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation. SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC. Intel is a registered trademark of Intel Corporation. * 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. References in this document to IBM products or services do not imply that IBM intends to make them available in every country. Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use. The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice. Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

Complete your sessions evaluation online at SHARE.org/AnaheimEval



Agenda

- Hardware components of an ensemble
- z/VM Ensemble Components
- Virtual Switch Controllers
- DIRMAINT authorizations
- Enable SMAPI Servers
- Validating the Enablement
- Linux Ensemble Considerations

What is a zEnterprise Ensemble

- A zEnterprise node is a single zCEC with 0 to 4 zBX frames and up to two blade centers per frame
- A zEnterprise Ensemble is a collection of 1 to 8 zEnterprise Nodes managed as a single virtualized pool of server resources
- A zEnterprise node can be a member of a single ensemble
- An ensemble is the management scope for the Unified Resource Manager
- A primary / alternate pair of HMCs provide the management console for the ensemble



Enhancements for zEnterprise Ensembles

- Supported SLES and RHEL distributions
 - Optional - Guest Platform Management Provider (GPMP)
 - IEDN/INMN (OSX/OSM) NIC support
 - Legacy NIC connection to IEDN or INMN via virtual switch ok via OSDSIM support in VSWITCH
- z/VM V6
 - z/VM Management Guest to forward Linux GPMP data to Unified Resource Manager
 - z/VM SMAPI enhancements
 - z/VM Directory Maintenance server (or equivalent)
 - INMN and IEDN virtual switch controllers
 - Control point for MAC assignment and VLAN access

Enhancements for zEnterprise Ensembles

- INMN and IEDN access provided via new z/VM virtual switch types
 - Uplink is the z/VM Management Guest
 - Automatic connection to INMN
 - Ensemble membership sets ensemble-defined MAC for each IEDN NIC
- SMAPI updates SYSTEM CONFIG
- z/VM is authoritative source of virtual machine state
 - State automatically reflected in Unified Resource Manager

z/VM System Management API Infrastructure Changes



- New SMAPI servers:
 - Support for IPv6 (INMN is IPv6)
 - Resiliency and error recovery (aka “guard” functions)
 - Management Guest, instantiated by the Unified Resource Manager
- New Systems Management APIs

z/VM SMAPI Family




- **VSMGUARD** Is responsible for starting and monitoring other SMAPI servers and platform management guest.
- **VSMPROXY** Speaks to the Support Element and HMC
- **VSMREQIN** Requests from IPv4 clients
- **VSMREQIU** Requests from other guests using AF_IUCV sockets
- **VSMREQIM** Requests from IPv6 management networks
- **VSMREQI6** Requests from IPv6 clients
- **VSMEVSRV** Gathers data from *VMEVENT and *LOGREC system services
- **ZVMLXAPP** z/VM Unified Resource Manager platform management guest

6.1

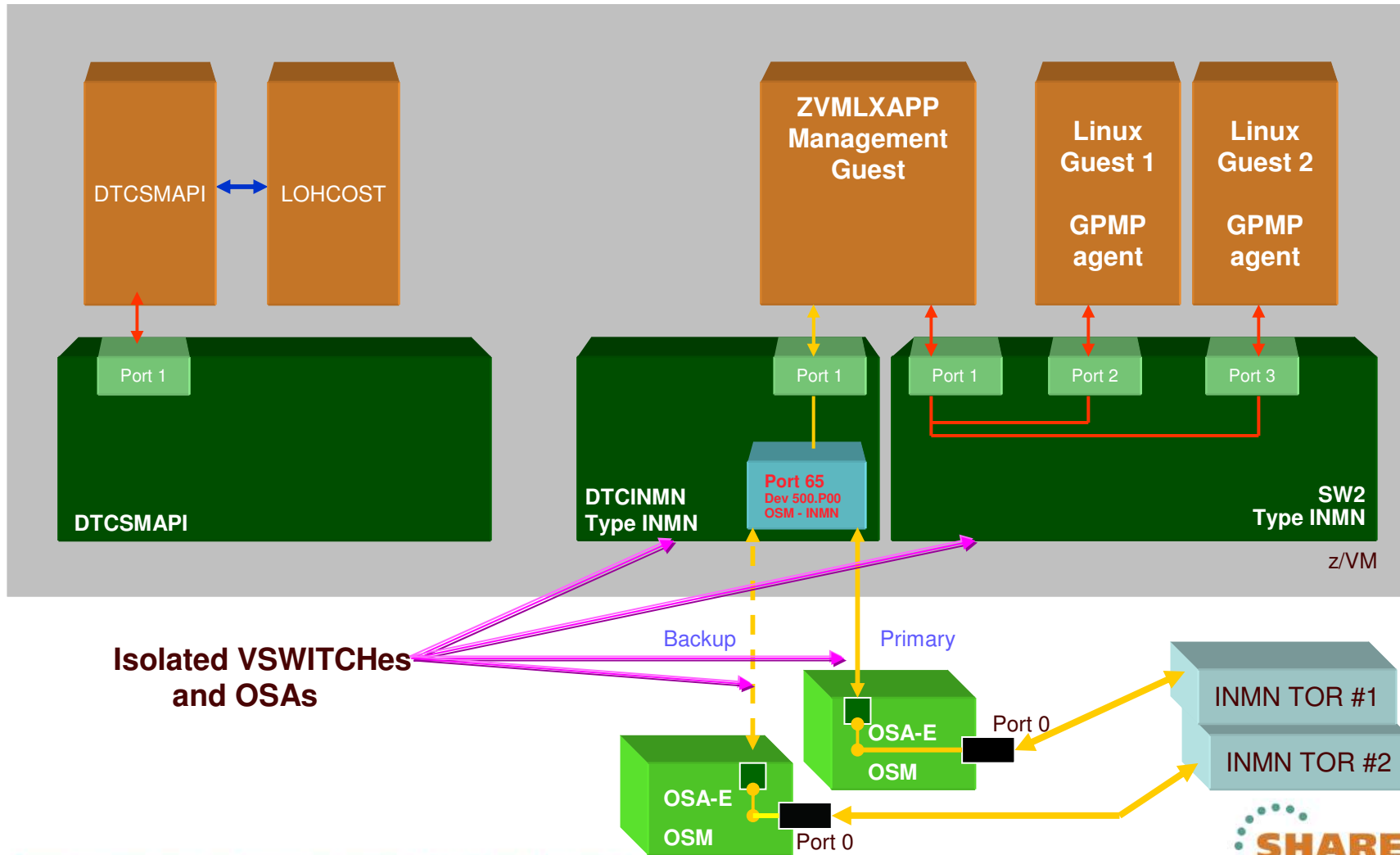
6.2

6.1

z/VM SMAPI Family

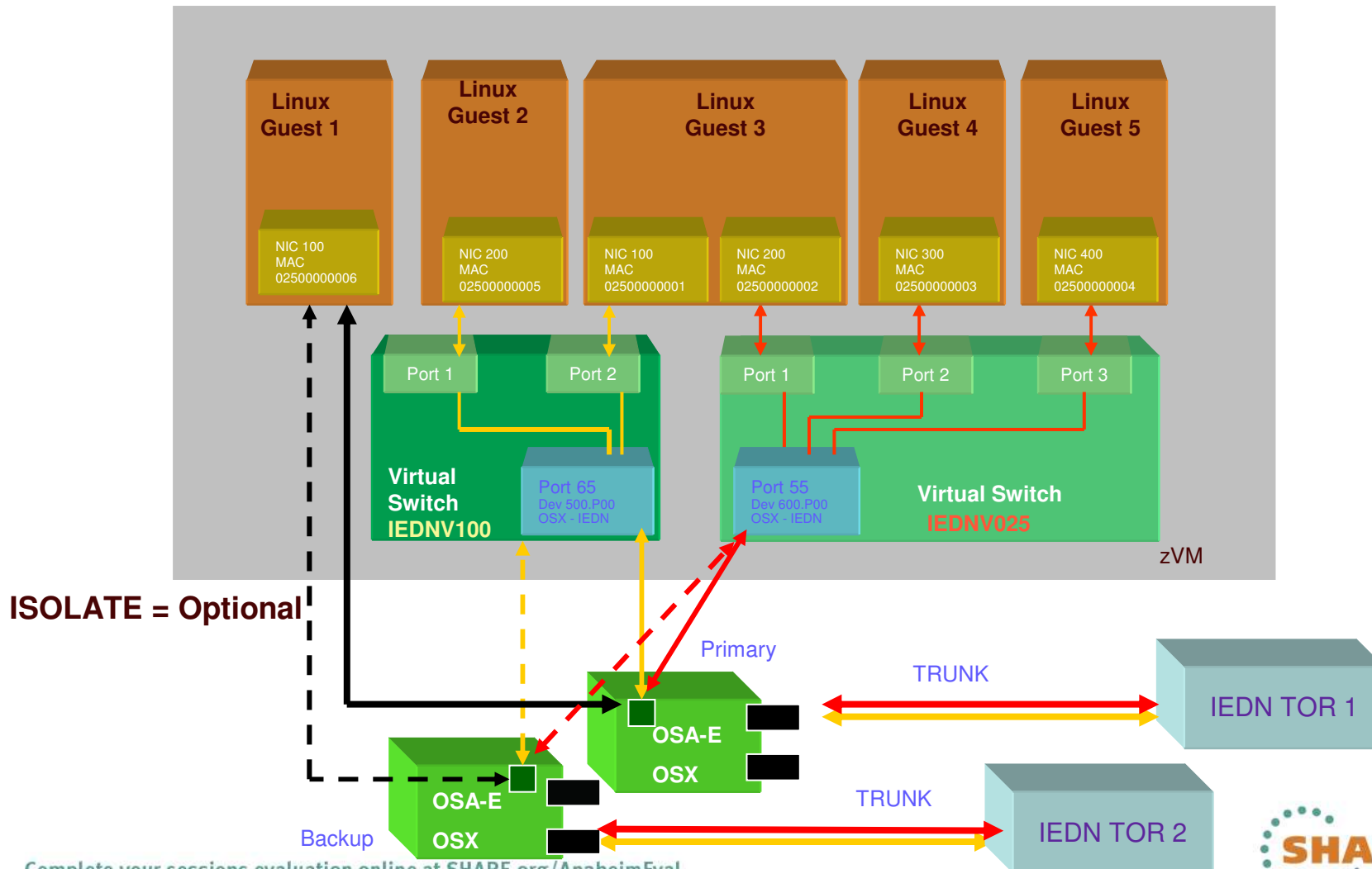
- **VSMWORK1** Request server for short transactions
- **VSMWORK2** Request server for long-running transactions
- **VSMWORK3** Request server for long-running transactions
- **LOHCOST** Caching server for Query-type operations 
- **DTCSMAPI** Private TCP/IP stack for SMAPI components that require IP connectivity 
- **PERSMAPI** Performance monitor used if managing z/VM exclusively by Unified Resource Manager 

z/VM Ensemble INMN Infrastructure



Complete your sessions evaluation online at SHARE.org/AnaheimEval

z/VM Ensemble IEDN Infrastructure



Complete your sessions evaluation online at SHARE.org/AnaheimEval

Preparation for Enablement

- Software
 - z/VM 6.2 at Service Level 1101 or higher
 - Get the latest APAR information from
 - <http://www.vm.ibm.com/service/vmrequirm.html>
- Hardware
 - System z firmware bundle 41z or higher
 - OSX and OSM CHPIDs configured and cabled
 - If not cabled, you will see error codes E080
 - Proper ports, please!



Preparation for Enablement

- Books
 - CP Planning and Administration Guide, Chapter 15
 - April 2012
 - z/VM Systems Management Application Programming, Chapter 4
 - April 2012
 - zEnterprise Ensemble Planning and Configuration
 - -04a from Resource Link (March 2012)

Validate OSX/OSM devices are available

Q OSA TYPE ENSEMBLE

OSA	2300	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2301	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2302	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2303	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2304	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2305	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2306	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2307	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX
OSA	2308	OFFLINE	DEVTYPE	IEDN	CHPID	18	OSX

Validate OSX/OSM devices are available

Continued ...

OSA	2340	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2341	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2342	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2343	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2344	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2345	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2346	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2347	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM
OSA	2348	OFFLINE	DEVTYPE	INMN	CHPID	0A	OSM

Configure DIRMAINT Authorizations

- Command permission
- Surrogate permission

DIRMAINT Authorization

- Give SMAPI worker virtual machines permission to issue privileged DIRMAINT commands

- Update AUTHFOR CONTROL file:

```
ALL VSMWORK1 * 140A ADGHMOPS
ALL VSMWORK1 * 150A ADGHMOPS
ALL VSMWORK2 * 140A ADGHMOPS
ALL VSMWORK2 * 150A ADGHMOPS
ALL VSMWORK3 * 140A ADGHMOPS
ALL VSMWORK3 * 150A ADGHMOPS
ALL VSMGUARD * 140A ADGHMOPS
ALL VSMGUARD * 150A ADGHMOPS
```

DIRMAINT CONFIGxx DATADVH Additions

- Allow SMAPI worker virtual machines to issue requests on behalf of already-authenticated SMAPI clients
- Update CONFIGxx DATADVH file:
 - ALLOW_ASUSER_NOPASS_FROM= VSMWORK1 *
 - ALLOW_ASUSER_NOPASS_FROM= VSMWORK2 *
 - ALLOW_ASUSER_NOPASS_FROM= VSMWORK3 *
 - ALLOW_ASUSER_NOPASS_FROM= VSMGUARD *

Authorize the Management Guest

```

===== * * * Top of File * * *
      |...+....1....+....2 // .6....+....7....+....8.. // +...13....+...14..
===== DO.NOT.REMOVE                DO.NOT.REMOVE                DO.NOT.REMOVE
===== MAINT                          ALL                          ALL
===== VSMPROXY                        ALL                          ALL
===== ZVMLXAPP                        ALL                        ALL
===== * * * End of File * * *
=====>

```

VM65083 – PTF UM33623

VSMWORK1 AUTHLIST
in
VMSYS:VSMWORK1.

Note column numbers: 1, 66, 130

SFS Administrator Authority for VMSYS

- On VMSEVS 191 minidisk:

```
00000 * * * Top of File * * *
00001 ADMIN MAINT 6VMTCP10 VSMGUARD
00002 NOBACKUP
00003 SAVESEGID CMSFILES
00004 FILEPOOLID VMSYS
00005 USERS 100
00006 * * * End of File * * *
```

How to Operate This New Infrastructure?

- To start the SMAPI servers, XAUTOLOG VSMGUARD
 - Add it to the PROFILE EXEC of AUTOLOG1 or AUTOLOG2
- VSMGUARD will start the SMAPI servers and the Management Guest will start automatically.
- ZVMLXAPP can be restarted via the Unified Resource Manager as a task of the z/VM Hypervisor

Validating the configuration

```
q vmlan
VMLAN maintenance level:
  Latest Service: VM64780
VMLAN MAC address assignment:
  System MAC Protection: OFF
  MACADDR Prefix: 020000 USER Prefix: 020000
  MACIDRANGE SYSTEM: 000001-FFFFFF
                   USER: 000000-000000
VMLAN Unified Resource Manager status:
  Hypervisor Access: YES      Status: MANAGED ←
  ID: 52BD737254BF11E0B85A0010184CB262
  MAC Prefix: 023C90
VMLAN default accounting status:
  SYSTEM Accounting: OFF      USER Accounting: OFF
VMLAN general activity:
  PERSISTENT Limit: INFINITE   Current: 5
  TRANSIENT Limit: INFINITE    Current: 0
Ready; T=0.01/0.01 16:19:45
```

Validating the INMN configuration

```
q vswitch dtcinmn
VSWITCH SYSTEM DTCINMN  Type: INMN      Connected: 2      Maxconn: INFINITE
PERSISTENT  RESTRICTED  ETHERNET          Accounting: OFF
VLAN Unaware
MAC address: 02-3C-90-00-00-01      MAC Protection: Unspecified
State: Ready
IPTimeout: 5          QueueStorage: 8
Isolation Status: ON
Uplink Port:
RDEV: 236D.P00 VDEV: 236D Controller: DTCENS1
RDEV: 234D.P00 VDEV: 234D Controller: DTCENS1  BACKUP
```

Validating the SW2 configuration

```
q vswitch sw2
VSWITCH SYSTEM SW2      Type: INMN      Connected: 1      Maxconn: INFINITE
  PERSISTENT RESTRICTED  ETHERNET        Accounting: OFF
  VLAN Unaware
  MAC address: 02-3C-90-00-00-03  MAC Protection: Unspecified
  State: Ready
  IPTimeout: 5                QueueStorage: 8
  Isolation Status: ON
  Uplink Port:
  NIC: ZVMLXAPP  VDEV: 0200
```


Validating the DTCENSx controllers

```

q controller
Controller DTCVSW2   Available: YES   VDEV Range: *           Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP   LINKAGG   ISOLATION
              NO_ENSEMBLE NO_INMN
    SYSTEM VSWITCH1 Primary           Controller: <list>     VDEV: 2100
Controller DTCVSW1   Available: YES   VDEV Range: *           Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP   LINKAGG   ISOLATION
              NO_ENSEMBLE NO_INMN
    SYSTEM VSWITCH1 Backup            Controller: <list>     VDEV: 2120
Controller DTCENS1   Available: YES   VDEV Range: *           Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP   LINKAGG   ISOLATION
              ENSEMBLE   INMN
    SYSTEM DTCINMN   Primary           Controller: DTCENS1     VDEV: 236D
    SYSTEM DTCINMN   Backup            Controller: DTCENS1     VDEV: 234D
Controller DTCENS2   Available: YES   VDEV Range: *           Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP   LINKAGG   ISOLATION
              ENSEMBLE   NO_INMN
  
```

Validating your SFS configuration

```
q auth vmsys:vsmwork1.
```

```
Directory = VMSYS:VSMWORK1.
```

```
Grantee R W NR NW
```

```
MAINT X X X X
```

```
VSMWORK1 X X X X
```

```
VSMGUARD X X X X
```

```
VSMPROXY X - X -
```

```
VSMREQIM X - X -
```

```
VSMREQIN X - X -
```

```
VSMREQIU X - X -
```

```
VSMREQI6 X - X -
```

```
VSMWORK2 X - X -
```

```
VSMWORK3 X - X -
```

Validating your SFS configuration

q auth vmsys:vsmwork1.data.

*Directory = **VMSYS:VSMWORK1.DATA***

Grantee R W NR NW

MAINT X X X X

VSMWORK1 X X X X

VSMGUARD X X X X

VSMPROXY X X X X

VSMREQIM X X X X

VSMREQIN X X X X

VSMREQIU X X X X

VSMREQI6 X X X X

VSMWORK2 X X X X

VSMWORK3 X X X X

Validating your SMAP configuration

```

netstat
VM TCP/IP Netstat Level 610      TCP/IP Server Name: TCPIP

Active IPv4 Transmission Blocks:

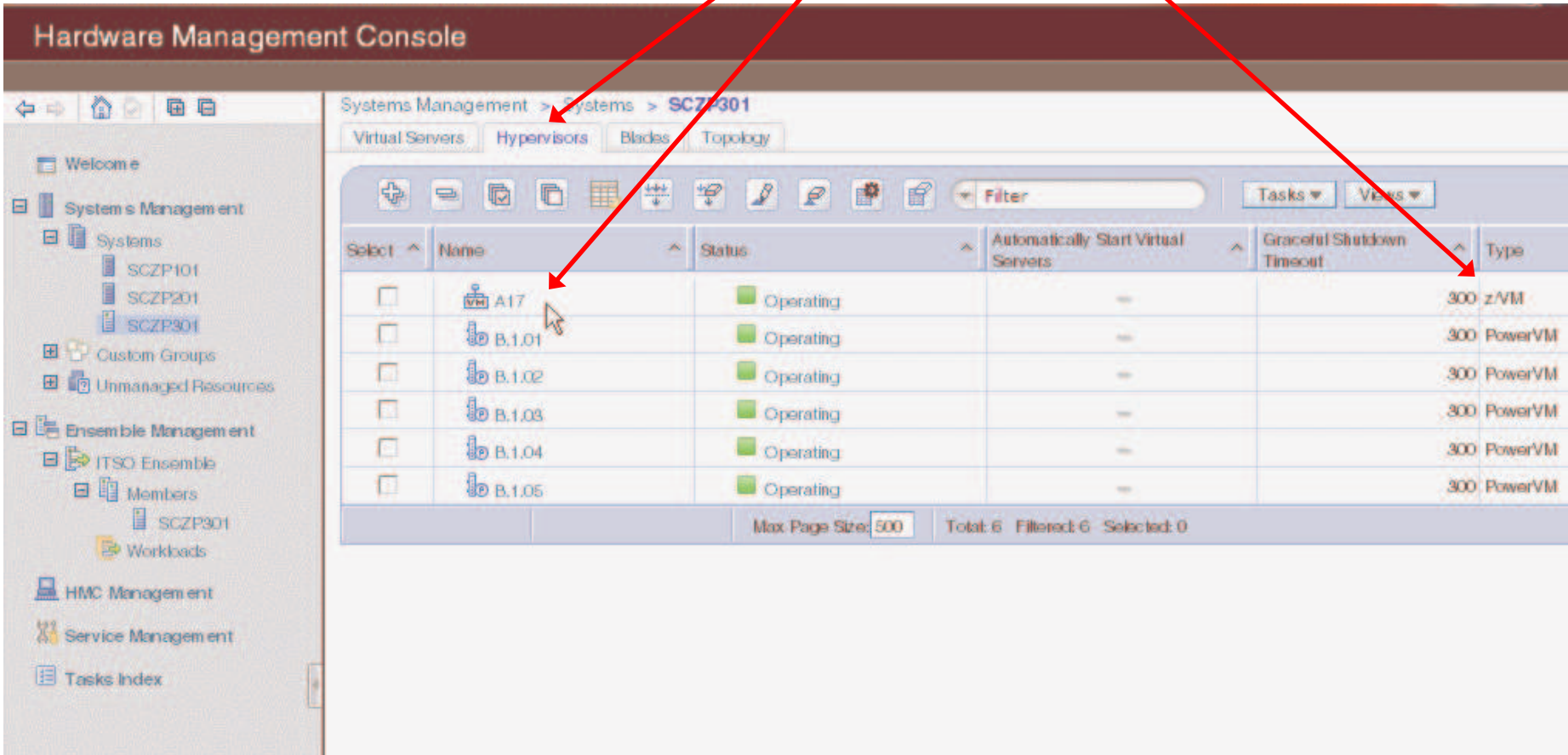
User Id  Conn      Local Socket          Foreign Socket         State
-----  --  -----  -
INTCLIEN 1006      *. .TELNET            *. .*                  Listen
INTCLIEN 1007      9.12.4.189. .TELNET    9.76.158.39. .50358    Established
VSMREQIN 1002      *. .44444             *. .*                  Listen
VSMPROXY 1003      *. .55555 ←          *. .*                  Listen

Active IPv6 Transmission Blocks:

User Id  Conn      State
-----  --  -----
VSMREQI6 1001      Listen
  Local Socket: *. .44445
  Foreign Socket: *. .*

```

Validating your configuration



Hardware Management Console

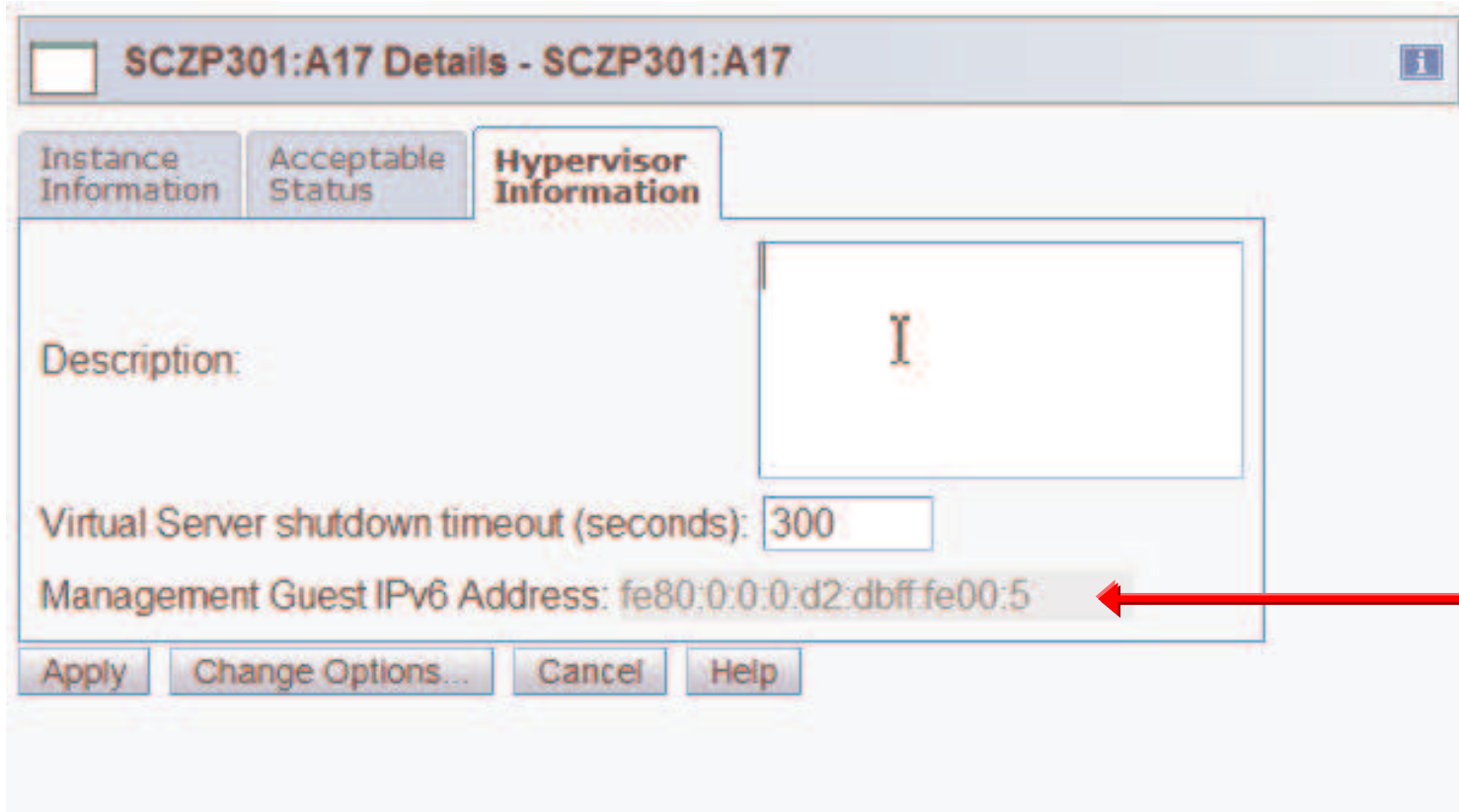
Systems Management > Systems > SCZP301

Virtual Servers | **Hypervisors** | Blades | Topology

Select	Name	Status	Automatically Start Virtual Servers	Graceful Shutdown Timeout	Type
<input type="checkbox"/>	A17	Operating	-		300 z/VM
<input type="checkbox"/>	B.1.01	Operating	-		300 PowerVM
<input type="checkbox"/>	B.1.02	Operating	-		300 PowerVM
<input type="checkbox"/>	B.1.03	Operating	-		300 PowerVM
<input type="checkbox"/>	B.1.04	Operating	-		300 PowerVM
<input type="checkbox"/>	B.1.05	Operating	-		300 PowerVM

Max Page Size: 500 Total: 6 Filtered: 6 Selected: 0

Validating your configuration



SCZP301:A17 Details - SCZP301:A17

Instance Information | Acceptable Status | **Hypervisor Information**

Description: I

Virtual Server shutdown timeout (seconds): 300

Management Guest IPv6 Address: fe80:0:0:0:d2:dbff:fe00:5

Apply | Change Options... | Cancel | Help

- The IPV6 IP address will display when the management guest is activated and z/VM is part of the Ensemble

Implementation Tips

- Access to the IEDN and OSX connections must be configured using Unified Resource Manager
 - Ensemble will reject “out-of-band” connection attempts
- VMSYS filepool needs to be backed up with the rest of your system!
 - This is where access rights and status are kept
 - If you lose it, you will start over
- No miracles. Unified Resource Manager doesn’t solve connectivity problems.
 - E.g. FCP devices must be able to access LUNs without zManager if they are going to be able to do it with zManager)
- Console output from VSMGUARD, VSMWORK1, and VSMREQUIU can hold clues if you have trouble.
 - But don’t believe everything you see

Implementation Tips

- If you vary all devices offline in the SYSTEM CONFIG and then vary on only the ones you know about, zManager-defined FCP devices could be a problem.
 - You may want to have a predefined range of devices for this
- Unified Resource Manager is not a RACF security administration application.
 - Enable DIRMAINT-RACF Connector (USE_RACF=YES)
 - VSWITCH and RDEV authorizations must be handled separately
- If ZVMLXAPP does not start, the other SMAPI service machines will not be started.
- Depending on the size and volume of the virtual server directories to be managed, you may find that the SMAPI servers will run out of memory.
 - The default is 128MB. You can increase up to 512MB.

Next Steps: Use Unified Resource Manager

- Create IEDN Virtual Switches and give guests access to the IEDN
- Define disk storage resources
 - System and user
- Define virtual server containers for Linux guests or migrate existing guest
- Manage guest resources

Managing guest priorities from zManager

- *“Too many cooks spoil the broth.”*
 - Only one resource manager at a time
 - If you are managing a guest with VMRM don't add it to a managed workload in zManager

References

- z/VM CP Planning and Administration Guide
- z/VM System Management Application Programming Reference
- z/VM CP Commands and Utilities Reference
- z/VM Directory Maintenance Facility Commands Reference

- IBM zEnterprise Ensemble Performance Management Guide
- IBM zEnterprise Ensemble Planning and Configuration Guide
- IBM zEnterprise Unified Resource Manager Redbook

Complete your sessions evaluation online at SHARE.org/AnaheimEval

Questions?

Alan C. Altmark

Senior Managing IT Consultant

IBM STG Lab Services

z/VM & Linux on System z

IBM

*1701 North Street
Endicott, NY 13760*

Mobile 607 321 7556

Fax 607 429 3323

Email: alan_altmark@us.ibm.com

Session 11776

