

### **DNA Sequencing with the z/VM Virtual Switch**

A look at Virtual Switch security and Directory Network Authorization



V1.5 – Last updated 15 June 2017 #vmworkshop #IBMz #zVM

#### **Trademarks**

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

BladeCenter* DB2* DS6000* DS8000* ECKD * Registered trademarks	FICON* GDPS* HiperSockets HyperSwap IBM z13* of IBM Corporation	OMEGAMON* Performance Toolkit for VM Power* PowerVM PR/SM	RACF* Storwize* System Storage* System x* System z*	System z9* System z10* Tivoli* zEnterprise* z/OS*	zSecure z/VM* z Systems*
---	--	---	---	---	--------------------------------

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries. Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and

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

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the OpenStack website.

TEALEAF is a registered trademark of Tealeaf, an IBM Company.

Windows Server and the Windows logo are trademarks of the Microsoft group of countries.

Worklight is a trademark or registered trademark of Worklight, an IBM Company.

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

\* Other product and service names might be trademarks of IBM or other 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.

This information provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs) ("SEs"). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at www.ibm.com/systems/support/machine\_warranties/machine\_code/aut.html ("AUT"). No other workload processing is authorized for execution on an SE. IBM offers SE at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

#vmworkshop #IBMz #zVM

# Agenda

- An Overview of the z/VM Virtual Switch
- User-Based Vs. Port-Based Access
- New: Directory Network Authorization (DNA)
  - How it works
  - Changes in CP
  - Implications to Guest Mobility (VMRELOCATE)
  - Changes in DirMaint
  - ESM Implications
  - Migration from Old to New



### Disclaimer

The information contained in this document has not been submitted to any formal IBM test and is distributed on an "AS IS" basis without any warranty either express or implied. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

In this document, any references made to an IBM licensed program are not intended to state or imply that only IBM's licensed program may be used; any functionally equivalent program may be used instead.

Any performance data contained in this document was determined in a controlled environment and, therefore, the results which may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environments.

It is possible that this material may contain reference to, or information about, IBM products (machines and programs), programming, or services that are not announced in your country. Such references or information must not be construed to mean that IBM intends to announce such IBM products, programming or services in your country.

#vmworkshop #IBMz #zVM

© 2017 IBM Corporation

4

### **Overview of the z/VM Virtual Switch**

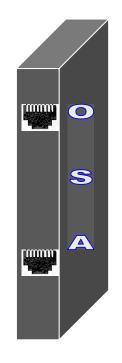
#vmworkshop #IBMz #zVM

### Switch – a box that creates a LAN

- Enable and disable a port
- Set port type: trunk or access
- Assign port to one or more VLANs
- Set port speed: 10 / 100 / 1000 / Auto
- Set port duplex mode: Simplex / Duplex / Auto
- Define an internal router
- Define SNAP (sniffer) ports



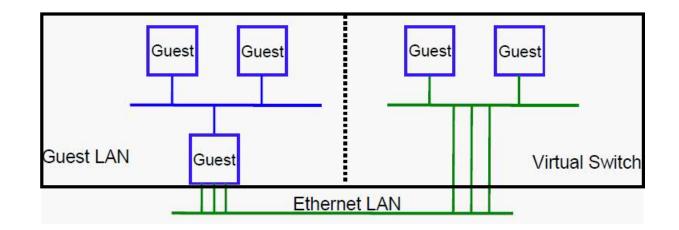
© Cisco Corp



6

#vmworkshop #IBMz #zVM

### **Guest LAN vs. Virtual Switch**

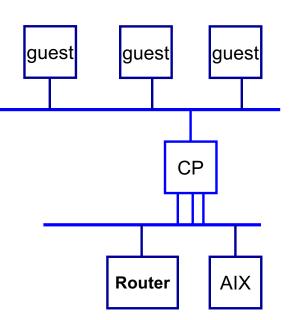


- Virtual router is required
- Different subnet
- External router awareness
- Guest-managed failover

- No virtual router
- Same subnet
- Transparent bridge
- CP-managed failover

# z/VM Virtual Switch

- A special-purpose Guest LAN –Ethernet
  - Built-in IEEE 802.1q bridge to an outside network
  - -IEEE VLAN capable
- Defining in z/VM:
- SYSTEM CONFIG (static definition)
- CP DEFINE VSWITCH command



TRM

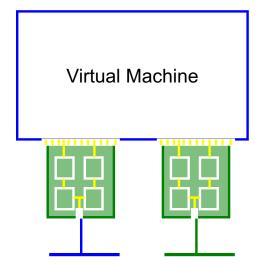
# **Virtual Switch Attributes**

- Name of the Switch
- Associated OSAs (up to 8) (may be aggregated and/or shared)
- Access Control List for z/VM userids
- 1+ controller VMs (minimal VM TCP/IP stack servers)
  - -User pre-configured DTCVSW1 and DTCVSW2
  - -Controller not involved in data transfer
  - -Starts, stops, and monitors OSAs
  - -Do not ATTACH or DEDICATE
- Similar to Guest LAN
  - -Owner SYSTEM
  - -Type QDIO
  - -Persistent
  - -Restricted

# Virtual Network Interface Cards (NICs)

- A simulated network adapter –OSA-Express QDIO
  - –HiperSockets
  - -Must match LAN type
- 3 or more devices per NIC

   More than 3 to simulate port sharing on 2nd-level system or for multiple data channels
- Provides access to Guest LAN or Virtual Switch

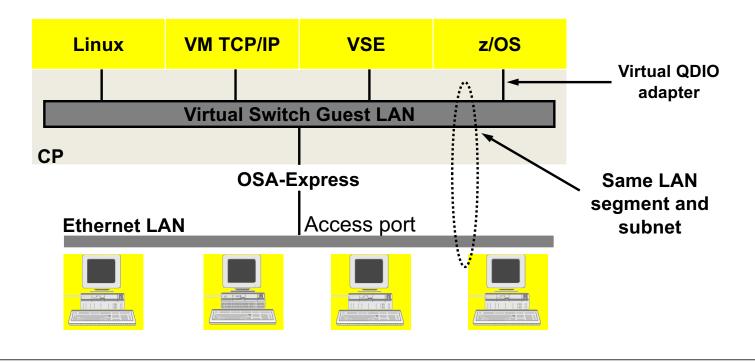


Guest LAN or virtual switch

Created by directory or CP DEFINE NIC command

TEM

### z/VM Virtual Switch – VLAN unaware



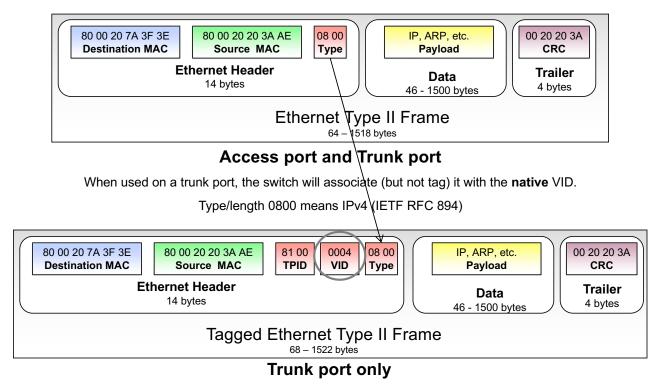
#vmworkshop #IBMz #zVM

### What's a "VLAN"? (This slide gleefully stolen from Alan Altmark. Sshhhh.)

- Defined by IEEE 802.1Q standard (not z/VM!)
  - "A subset of the active topology of a bridged LAN."
- A bridged LAN is what you get when you use a switch instead of a hub
  - Enables the application of ingress and egress rules to the frames that enter and exit the switch ports
- IEEE 802.1Q establishes a new set of rules and frame formats
  - Associated with each VLAN is a VLAN Identifier (VID).
  - VLAN-tagged frames carry the VID within the frame. Allowed only on trunk ports.
  - Untagged frames do not carry the VID, but are instead associated with a VID by the switch and then managed as though they were tagged
- VLAN-aware bridges create logical groups of end stations that can communicate as if they were on the same LAN by associating the physical port used by each of those end stations with the same VID.
- Traffic between VLANs is restricted. Bridges forward unicast, multicast, and broadcast traffic to ports that serve the VLAN to which the traffic belongs.

- Routers connect to multiple VLANs

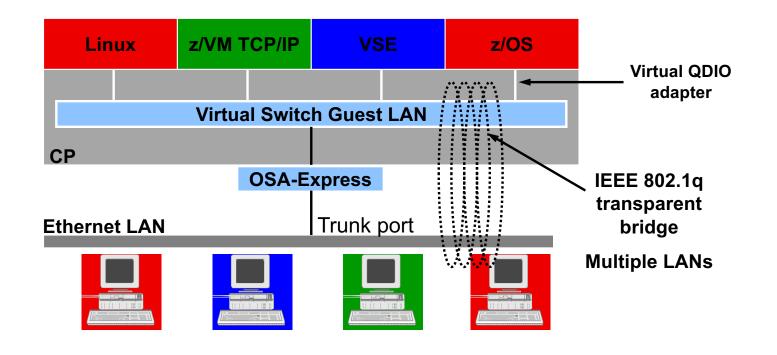
### **VLAN tags** (This slide gleefully stolen from Alan Altmark. Sshhhh.)



Value 8100 in the Type field means a VLAN tag follows, followed by the actual type/length field

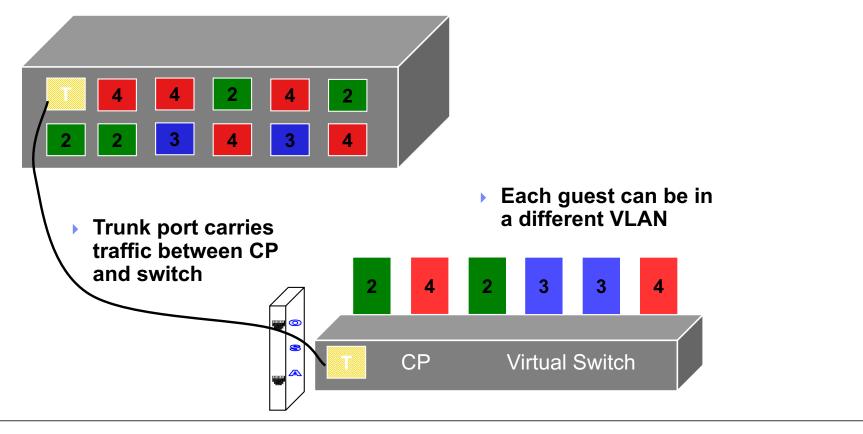
#vmworkshop #IBMz #zVM

### z/VM Virtual Switch – VLAN aware



#vmworkshop #IBMz #zVM

### **Physical Switch to Virtual Switch**



# **Security controls on the VSWITCH**

- Virtual Sniffers
  - -Guest must be authorized via SET VSWITCH or security server
  - -Guest enables promiscuous mode using CP SET NIC or via device driver controls (tcpdump -P)
  - -Guest receives copies of all frames sent or received for authorized VLANs

### Port Isolation or VEPA modes

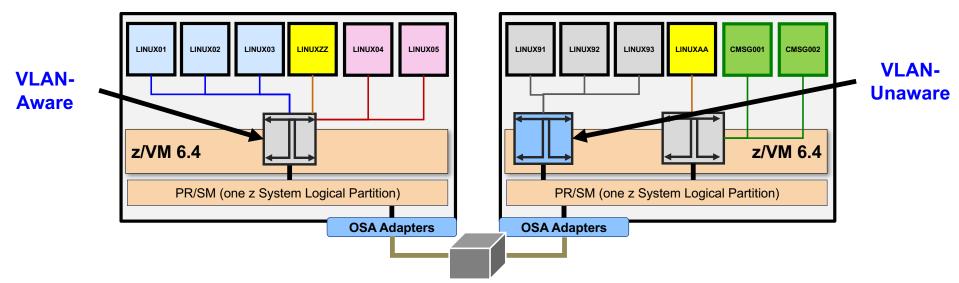
- -Isolation: Shut off OSA "short circuit" to other members of the VLAN
- -Virtual Edge Port Aggregator (VEPA) pushes separation to physical switch

#### Access lists

- -Not all guests can access a VSWITCH by default
- -Userid list defined via CP command or configuration statement

### An External Security Manager will provide granular control (of course)

### Zoning with the z/VM Virtual Switch



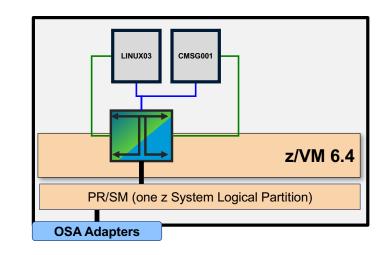
- Guests with a common VLAN can talk to one another (the OSA "hairpin turn")
- Guests that pass a security decision point (firewall) can communicate
- Separate traffic at VSwitch (VLAN), OSA (Port Isolation), or at physical switch (VEPA mode)

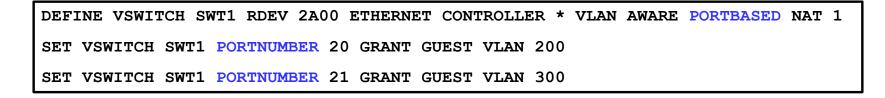
#### Port to Port – Virtual Machine Access to the Virtual Switch

#vmworkshop #IBMz #zVM

### **Does a guest need to be VLAN-aware?**

- "It depends" on Virtual Switch configuration
  - TRUNK frames are VLAN-tagged
  - ACCESS VLAN-unaware
- Linux guests often define ports to represent a particular *eth.n* connection
  - Don't necessarily care about VLAN id
- The Port-Based Vswitch allows a guest with multiple vNICs to connect via ACCESS ports to a single VLAN-aware virtual switch
  - Associate Port with VLAN
  - Guest doesn't need to care





### **Solution – VSWITCH PORTBASED definition**

A guest may have multiple vNICs connect to a VSwitch – as ACCESS Port connections.

The PORTNUMBER option makes this possible.

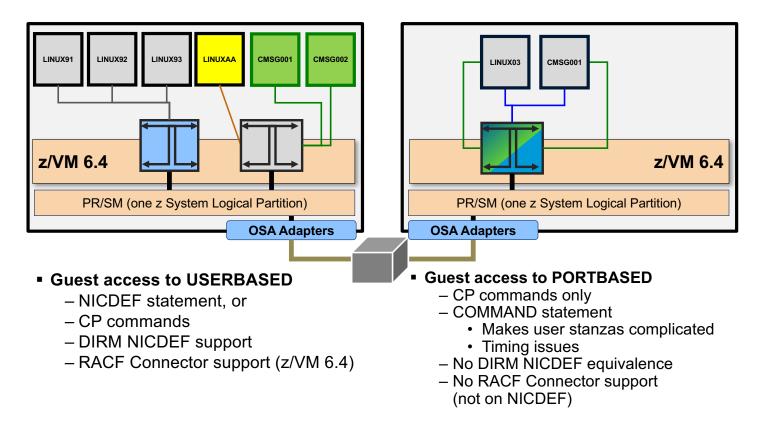
DEFINE VSWITCH LAGGSWT1 ETHERNET RDEV NONE CONTROLLER * VLAN AWARE PORTBASED NAT 1 GROUP SPG1
SET PORT GROUP SPG1 JOIN 5700.P00 5703.P01 5600.P00 5603.P01
SET VSWITCH LAGGSWT1 PORTNUMBER 15 USERID SUSE80 VLAN 1
SET VSWITCH LAGGSWT1 PORTNUMBER 16 USERID SUSE80 VLAN 192
SET VSWITCH LAGGSWT1 PORTNUMBER 17 USERID SUSE80 VLAN 300
SET VSWITCH LAGGSWT1 PORTNUMBER 18 USERID SUSE80 VLAN 700

### **Solution – VSWITCH PORTBASED definition**

- PORTNUMBERs don't have to start with (or increment by) 1 just have to be different.
- · Can migrate user-based definitions to port-based with no impact
- RACF doesn't care either way, either no substantive change to security policy

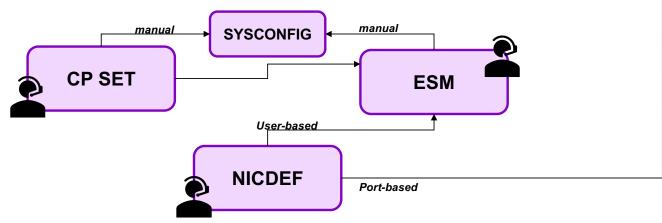
q vswitch
VSWITCH SYSTEM DTCSMAPI Type: QDIO Connected: 0 Maxconn: INFINITE
PERSISTENT RESTRICTED ETHERNET Accounting: OFF
USERBASED
VLAN Unaware
MAC address: 02-00-10-00-00-19 MAC Protection: Unspecified
State: Defined
IPTimeout: 5 QueueStorage: 8
Isolation Status: OFF
VSWITCH SYSTEM LAGGSWT1 Type: QDIO Connected: 10 Maxconn: INFINITE
PERSISTENT RESTRICTED ETHERNET Accounting: OFF
PORTBASED
VLAN Aware Default VLAN: NONE Default Porttype: Access GVRP: Enabled
Native VLAN: 0001 VLAN Counters: OFF

### And then the drama started ...



### And then the drama started ...

- Kind of complicated, isn't it?
  - NIC device properties in the User Directory
  - NIC network properties in System Config, CP SET, and/or an ESM
- Kind of hard to manage, isn't it?
  - No port-based support for NICDEF all CP SET
  - Now try deploying a switch through IBM Wave, or OpenStack, or even DirMaint ...



© 2017 IBM Corporation

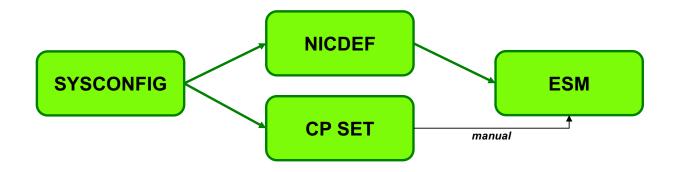
555

### **Directory Network Authorization (DNA)**

#vmworkshop #IBMz #zVM

#### **Directory Network Authorization** (APARs VM65925, VM65926, and VM65931 – August 2017)

- Centralize everything on the NICDEF statement
  - New parameters included
  - NICDEF takes precedence over CP SET VSWITCH (still supported)
  - NICDEF LAN authorizes network access
  - No meaningful distinction between user-based and port-based anymore



### How do I enable DNA? (And what if I change my mind?)

System Configuration VMLAN statement:

```
+-- ENABLE ---v
>>--VMLAN [...]- DNA --+-----+---------><
+-- DISABLE --^
```

• Or CP SET VMLAN:

• New message if there's a conflict with state vs. configuration:

HCP3224I NICDEF network configuration is ignored due to the current setting of VMLAN DNA.

### What do I do to use DNA?

- Step 1: Apply PTFs when available. (C'mon, that was the easy part.)
  - z/VM DIRECTXA processes the new directory statement changes
  - NICDEF statement defines properties of virtual NIC
  - NICDEF also now supports network attributes defined by CP SET VSWITCH
- Then, secure it. Recommend change access be restricted for:
  - LAN, MACID, PORTNUMBER, VLAN, PROMISCUOUS
  - Impacts security policy, should be administrator-only
  - DirMaint will have these changes when you apply that PTF

NICDEF	vdev	[TYPE HIPERS   QDIO]
		[DEVices devs]
		[LAN owner name]
		[CHPID xx]
		[MACID xxyyzz]
		[PORTNUMber nnnn]
		[PORType ACCESS TRUNK]
		[VLAN vidset]
		[PROmiscuous NOPROmiscuous]

### Tell me about ... LAN

- NICDEF LAN now authorizes the connection to a virtual network
  - Unless VMLAN DNA is DISABLED in the System Configuration file
  - LAN statement still optional

NICDEF	vdev	[TYPE HIPERS   QDIO] [DEVices <i>devs</i> ]
		[LAN owner name]
		[CHPID xx]
		[MACID xxyyzz]
		[PORTNUMber nnnn]
		[PORType ACCESS TRUNK]
		[VLAN vidset]
		[PROmiscuous NOPROmiscuous]

#vmworkshop #IBMz #zVM

© 2017 IBM Corporation

TEM

### Tell me about ... PORTNUMber

- New, Optional configuration setting
  - Decimal value 1-2048
  - System will assign a port 2176-4095 if not specified
  - Virtual NIC is connected to nnnn of the specified network
  - If specified port is already in use, the COUPLE will fail
  - If PORTNUMBER is omitted, VSWITCH will use a system-assigned port number

NICDEF	vdev	[TYPE HIPERS   QDIO]
		[DEVices devs]
		[LAN owner name]
		[CHPID xx]
		[MACID xxyyzz]
		[PORTNUMber nnnn]
		[PORType ACCESS TRUNK]
		[VLAN vidset]
		[PROmiscuous   NOPROmiscuous]

### Tell me about ... PORTTYPE

- New, Optional configuration setting
  - For a VLAN-Aware Virtual Switch only
  - Port type TRUNK is valid for any number of VLAN groups
  - Port type ACCESS is only valid for a single VLAN (frames are untagged)
  - If keyword omitted, COUPLE will resolve it (with the VSWITCH default)

1	NICDEF	vdev	[TYPE HIPERS   QDIO]
1			[DEVices devs]
			[LAN owner name]
			[CHPID xx]
			[MACID xxyyzz]
			[PORTNUMber nnnn]
			[PORType ACCESS TRUNK]
			[VLAN vidset]
			[PROmiscuous NOPROmiscuous]

#vmworkshop #IBMz #zVM

### Tell me about ... VLAN Specification

- New, Optional keyword (defaults to VSWITCH VLAN)
  - For a VLAN-Aware Virtual Switch only
  - List of VLAN IDs and/or ranges
  - Each VLAN ID is a decimal number 1-4096
  - For ACCESS, it must be precisely one number
  - For TRUNK, it may be as complicated as you want (e.g. VLAN 1 200-204 529)

NICDEF vo	lev [TYPE	HIPERS   QDIO]
	[DEVi	ces <i>devs</i> ]
	[LAN	owner name]
	[CHPI	D xx]
	[MACI	D xxyyzz]
	[PORT	NUMber nnnn]
	[PORT	ype ACCESS   TRUNK]
	[VLAN	vidset]
	[PROm	iscuous NOPROmiscuous]

### Tell me about ... Promiscuous Mode

- New, Optional keyword (defaults to NOPROMiscuous)
  - Authorizes virtual NIC to trace network traffic
  - Use sparingly

NICDEF	vdev	[TYPE HIPERS   QDIO]
		[DEVices devs]
		[LAN owner name]
		[CHPID xx]
		[MACID xxyyzz]
		[PORTNUMber nnnn]
		[PORType ACCESS TRUNK]
		[VLAN vidset]
		[PROmiscuous NOPROmiscuous]

#vmworkshop #IBMz #zVM

### How do I know what I'm using?

- CP QUERY VMLAN Command
- QUERY VMLAN output is changed to include the current system-wide DNA setting:

```
22:17:30 VMLAN maintenance level:
22:17:30 Latest Service: VM65925
22:17:30 VMLAN MAC address assignment:
          System MAC Protection: OFF
22:17:30
22:17:30
          MACADDR Prefix: 02110A USER Prefix: 02110A
22:17:30 MACIDRANGE SYSTEM: 000001-FFFFFF
22:17:30
                     USER: 000001-7FFFFF
22:17:30 VMLAN default accounting status:
22:17:30
          SYSTEM Accounting: OFF
                                     USER Accounting: OFF
22:17:30 VMLAN general activity:
22:17:30
          PERSISTENT Limit: INFINITE Current: 6
22:17:30 TRANSIENT Limit: INFINITE Current: 0
22:17:30
            Trace Pages: 8
22:17:30 VMLAN Directory Network Authorization: ENABLED
22:17:30 IVL Domain: None
```

### What about SET VSWITCH?

- SET VSWITCH still supports dynamic changes to PORTTYPE, VLAN and PROMISC.
- LOGON and DEFINE with NICDEF attributes overrides prior SET VSWITCH settings.
- Note that SET VSWITCH with GRANT/REVOKE affects ALL connections for the given user!
- ESM changes are not reflected to CP dynamically, so user must UNCOUPLE/COUPLE to adopt ESM updates.

### What about VMRELOCATE?

- VMRELOCATE is not directly affected, but be advised:
  - Systems without this PTF will not recognize new NICDEF features
  - Common User Directory interpreted differently on each member
  - Be cautious

lf you have a …	on a	Then be warned that
System-assigned port	Port-based VSwitch	No relocation to a pre-DNA system. Ports > 2048 unsupported
User-defined port	User-based Vswitch	It may not have the same port if relocated to a pre-DNA system
User-defined port	User-based Vswitch	Relocation my fail if the port belongs to a different user on target system

### What about DirMaint?

- The Directory Maintenance Facility (DIRMAINT) for z/VM 6.4 is updated through the PTF for APAR VM65926.
- Changes include:
  - DIRM NICDEF support for new options (command and menu)
  - NICDEF command is now a DirMaint-Class-A command, and not Class G
  - Update to DVHRVN (RACF Connector Exit) to transmit changes in new options

TEM

#vmworkshop #IBMz #zVM

# What about RACF?

- RACF for z/VM 6.4 is updated through the PTF for APAR VM65931.
- RACF handles user-based and port-based virtual switches the same way already, so there was no need for major processing changes.
- The RPIDIRCT utility now supports the new NICDEF processing
  - Tolerates old format, too
  - Processes PROMISCUOUS into ACC(CONTROL)
  - Processes VLAN ids into discrete profiles
  - Updates to support of SPECIAL statement
  - Some smaller bug fixes

### **Questions?**

#vmworkshop #IBMz #zVM

### Conclusion

- VM65925 simplifies and streamlines hypervisor network security
  - Centralized location for network security policy
  - Can eliminate excessive use of COMMAND statements
  - The ESM still always wins, though
- DNA is enabled by default, but changes won't be immediate
  - Clues off new options in the User Directory
  - Some changes to DirMaint to reflect NICDEF's importance
  - No changes in RACF
- Be mindful of guest relocation
  - Changes in port behavior could complicate security policy requirements
  - User-based switches subject to similar constraints now

### **Best Practices for VLAN-aware VSWITCH**

- Use NICDEF to assign VLANs and port numbers (NEW)
- Define VSWITCH with "VLAN AWARE NATIVE NONE"
  - Guest that has not been given access will get errors
  - $-\operatorname{No}$  chance of untagged frames escaping from z/VM
- Use ESM and groups to manage VLAN assignments
  - Simplifies VLAN changes
  - Overrides VLAN specification on NICDEF
  - CP will use NICDEF if ESM defers

### **Best Practices for all VSWITCHes**

- Use ETHERNET mode
- Do not specify PORTTYPE TRUNK on DEFINE VSWITCH
  - This controls the default guest port type, not the OSA!
- Do not specify CONTROLLER
- Do not put CONTROLLER ON in your own TCP/IP stacks – For VSWITCH controllers only!
- Specify MACPROTECT ON and LIMIT TRANSIENT 0 on VMLAN statement in SYSTEM CONFIG

### **Useful diagnostic commands**

#### CP QUERY VMLAN

- to get global VM LAN information (e.g. limits)

- to find out what service has been applied

#### CP QUERY LAN ACTIVE

- to find out which users are coupled

- to find out which IP addresses are active

#### CP QUERY NIC DETAILS

- to find out if your adapter is coupled
- to find out if your adapter is initialized
- to find out if your IP addresses have been registered
- to find out how many bytes/packets sent/received

#### CP QUERY PORT GROUP

- To determine the members of a particular groupname
- To determine which groups are active or inactive

### Most popular VSWITCH configuration problem?

- Not issuing the COUPLE for your virtual NIC.
- Measure twice, cut once.
  - QUERY VIRTUAL NIC (a Class G command)

### For More Information ...

- <u>http://www.vm.ibm.com/virtualnetwork/</u>
- <u>http://www.vm.ibm.com/virtualnetwork/linkag.html</u> -- z/VM Link Aggregation Development
- http://www.ibm.com/servers/eserver/zseries/os/linux/
- http://www.linuxvm.org/
- With special thanks to:
  - Sue Farrell, Dennis Musselwhite, and Rick Tarcza (z/VM Networking Development)
  - Patty Rando (z/VM DirMaint Develpoment)
  - Alan Altmark (z/VM networking nerd and security enthusiast)

Contact Information:

Brian W. Hugenbruch, CISSP IBM z Systems Virtualization Security bwhugen at us dot ibm dot com @Bwhugen



#vmworkshop #IBMz #zVM

