

Using z/VM VSWITCH

David Kreuter

February 13, 2007

Tampa SHARE

Session 9124

Using vswitch on z/VM

- Definition of guest lan
- Vswitch concepts
- Vswitch implementation, management, and recovery
- VM TCPIP stack configuration
- linux stack configuration

Guest Lans

- Virtual network adapters connect IP stacks in virtual machines.
- No hardware is required.
 - It's all done by CP commands, directory statements, configuration file statements, etc.
- High speed and high volume networks.
- One z/VM system can have multiple guest lans.
 - Guest lans can connect to other guest lans ...
 - Or be isolated from other guest lans
- One IP stack can belong to multiple guest lans.
- Supports multicast, unicast, broadcast networks.
- Supports all protocols.
- VM TCPIP and linux support guest lan

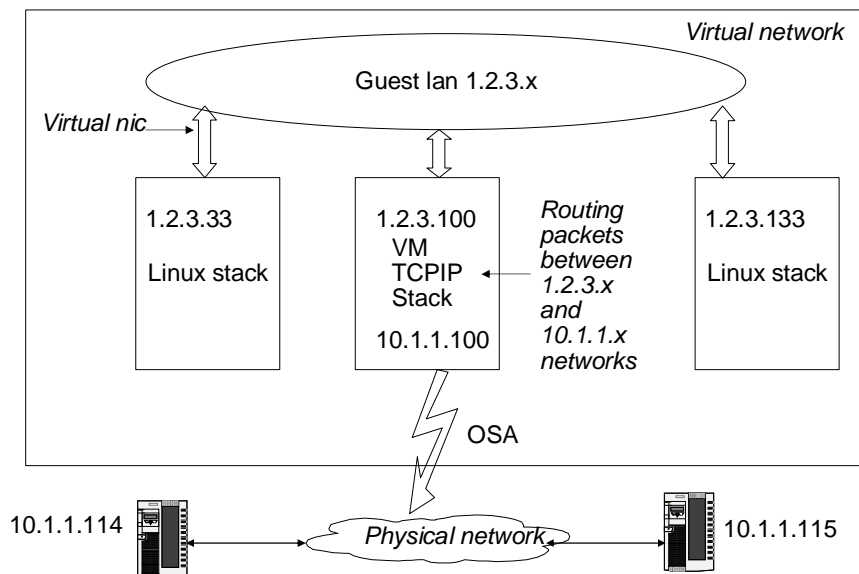
VSWITCH Concepts

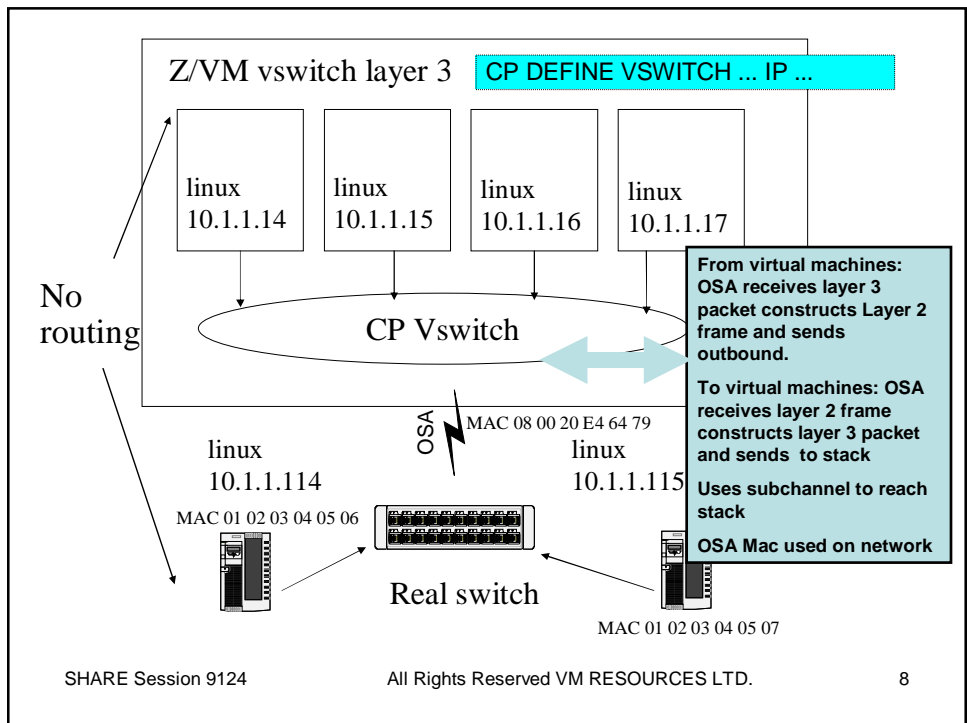
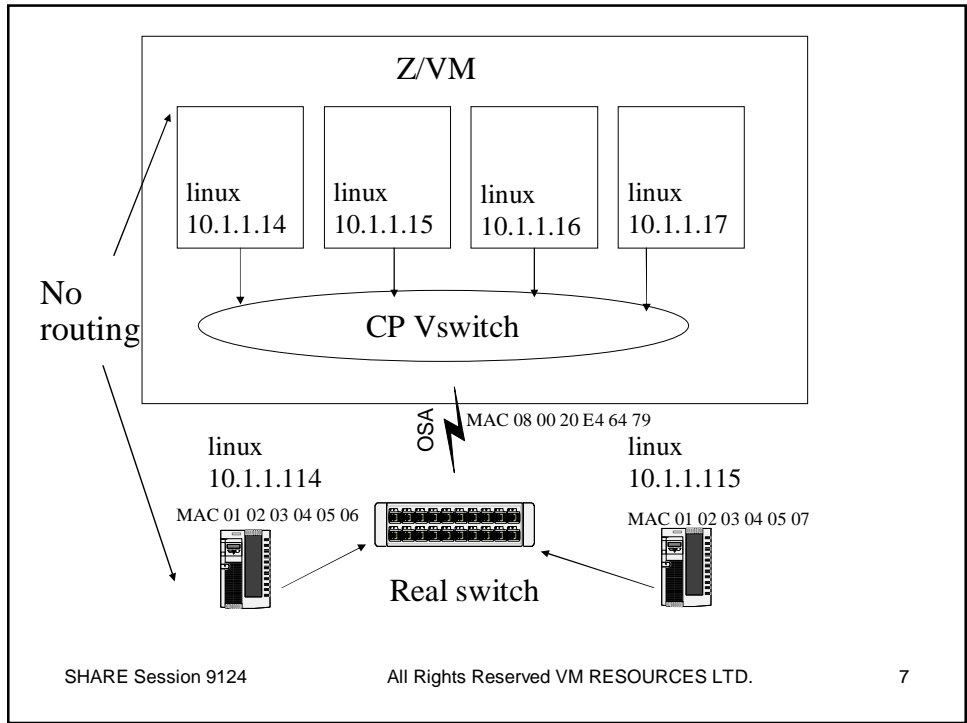
- Special kind of Guest LAN
- Like a Guest LAN Provides network of virtual network interfaces
- Connects directly to an OSA-Express QDIO Interface
- Or can run disconnected from real devices.
- Connects to external LAN segments without need for routing on z/VM.
- Operates as layer 2 or layer 3.
- Can have multiple Vswitches on one z/VM LPAR.

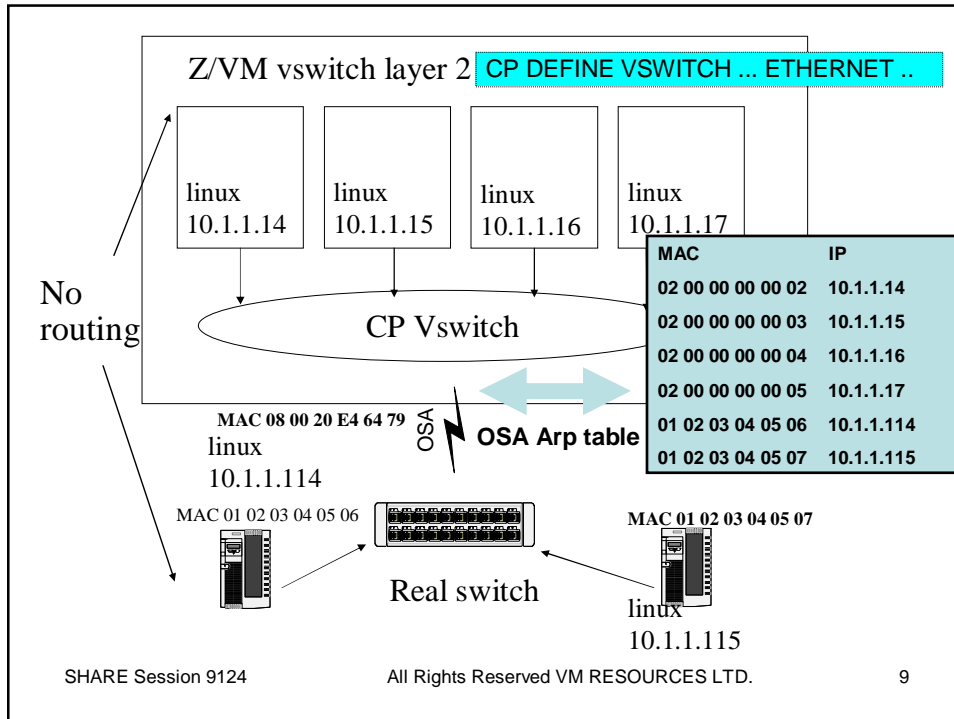
VSWITCH Presentation Goals

- Show controller command for dynamic controller management with two ranges of devices
- Show controller configuration
- Show configuration of 1st level vm tcpip stack
- Show configuration of 1st level linux stack
- Show configuration of 2nd level vm tcpip stack
- Show recovery scenarios

Typical Guest Lan



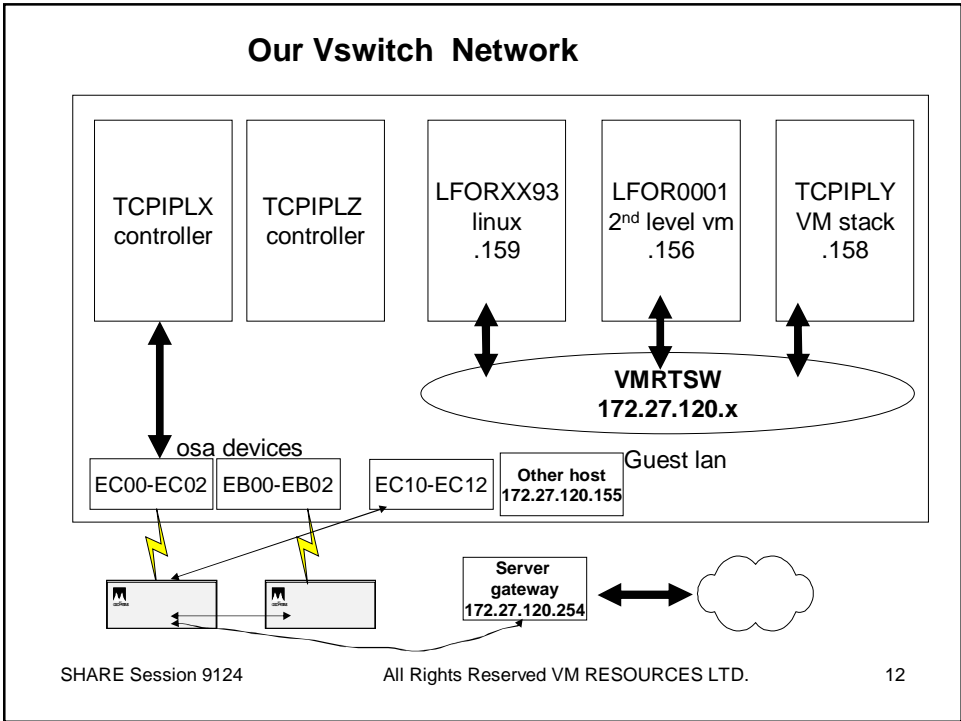
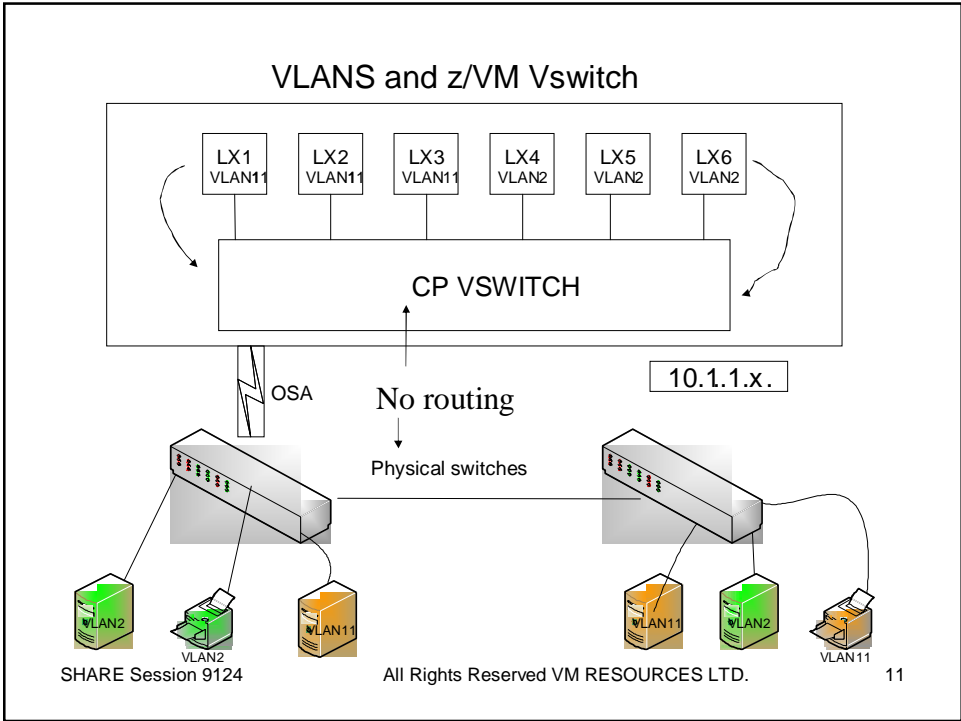




Participates in VLAN

- Supports Virtual Local Area Networks (VLANs) as per IEEE 802.1Q.
- CP provides virtual switch function.
- Hosts (Virtual Machines with IP stacks) on separate VLANs are isolated from each other.
- VLAN support operates in a layer 2 or 3 vswitch.

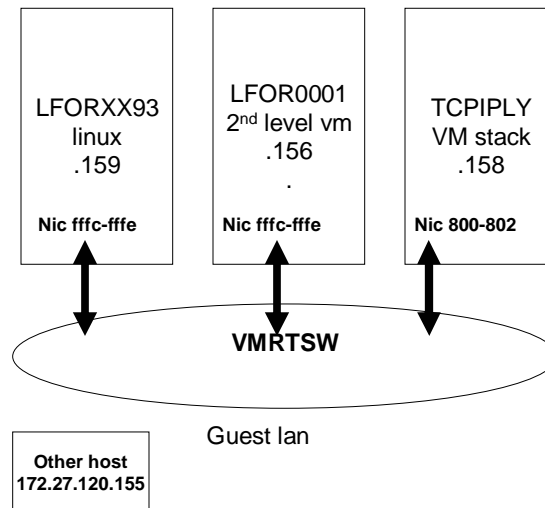
SHARE Session 9124 All Rights Reserved VM RESOURCES LTD. 10



Our Vswitch Network: nic devices

The virtual machines all have nic devices. QDIO type devices require 3 addresses: read, write and data. Nic devices are coupled to the guest lan VMRTSW. Hint: for linux cloning use the same nic address for all cloned linuxes.

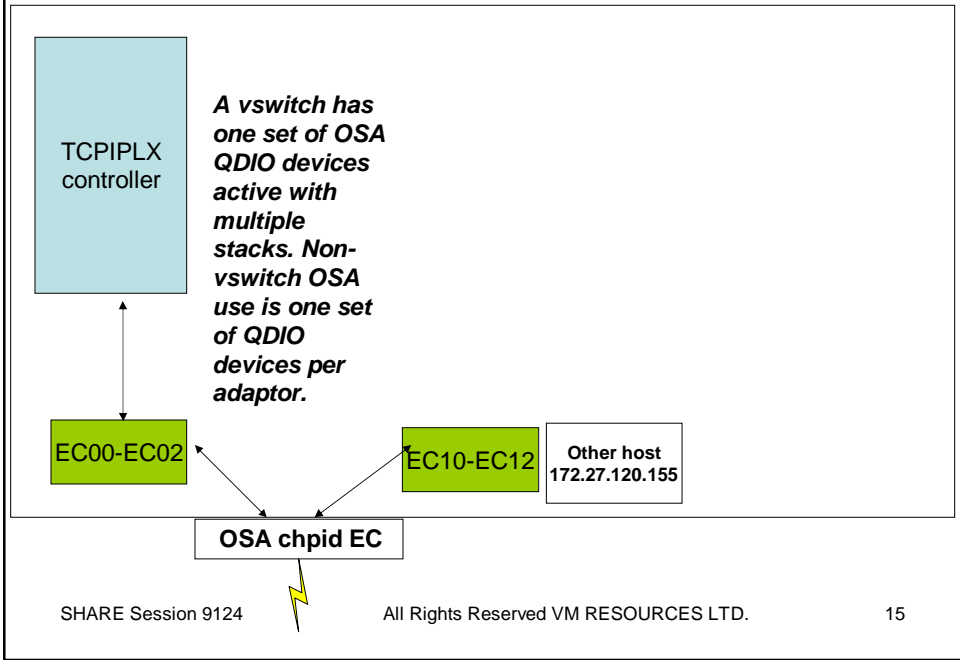
Participants on vswitches use virtual nic devices.



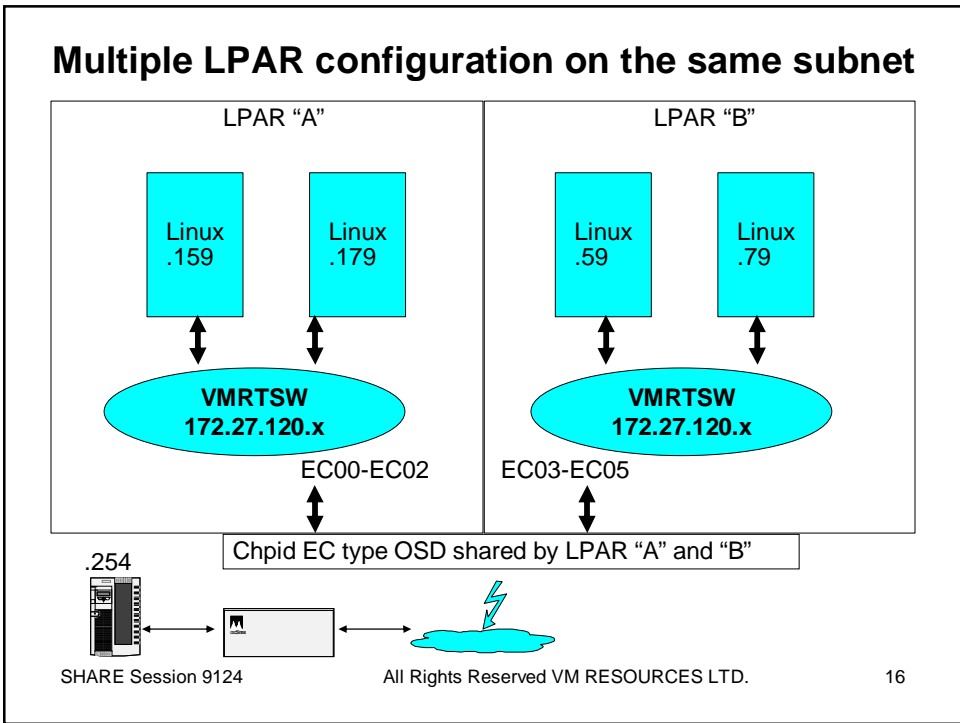
OSA and QDIO Mode

- QDIO mode is a z series high speed and high volume data transfer mechanism
 - Initiated as an I/O but ...
 - Once started remains active
 - And does not use standard I/O instructions
- OSA in QDIO mode supports:
 - Layer 3: IP mode: forwards IP broadcasts and multicasts; uses IP destinations from the IP packet. Supports VLAN.
 - Layer 2: Ethernet mode: uses MAC addresses from the LAN frame. Used by z/VM vswitch and the linux QETH drivers. Support VLAN along with multicast, broadcast and all protocols.
- Guest lans support virtual QDIO mode.

Our Vswitch Network: osa devices



Multiple LPAR configuration on the same subnet



A Few Words on VSWITCH

- The VSWITCH table of MACs, IP addresses, and virtual stacks is maintained by CP.
- The controller machine does *not* have DEVICE/LINK statements for the vswitch OSA devices.
- The OSA devices are automatically attached by CP to the controller machine when the VSWITCH is created.
 - One active set of OSA devices per vswitch.
- Virtual machines must be explicitly granted permission to join the vswitch..
 - Or access can be controlled by RACF.

Let's take a look

- Vswitch will be defined to use two sets of devices: EC00-EC02 and EB00-EB02:
 - EC00-EC02 will become active; EB00-EB02 will be standby.
 - *No load balancing*
- CP will look for controller (VM TCPIP stack machine):
 - Explicitly defined by CP command or SYSTEM CONFIG file statement
 - Or available machine (connected to *VSWITCH service)
- Will show two types of recovery:
 - Detaching EC00-EC02
 - Forcing off the active vswitch controller
- DEFINE VSWITCH is Class B
- DEFINE VSWITCH configuration file statement
- Guest lan user defines NIC with type QDIO

Defining the VSWITCH from MAINT

```
q ec00-ec02 eb00-eb02
```

```
OSA EC00 FREE , OSA EC01 FREE , OSA EC02 FREE , OSA EB00 FREE  
OSA EB01 FREE , OSA EB02 FREE
```

```
define vswitch vmrtsw ip controller * rdev ec00 eb00
```

```
VSWITCH SYSTEM VMRTSW is created
```

```
HCPSWU2830I VSWITCH SYSTEM VMRTSW status is ready.
```

```
HCPSWU2830I TCPIPLX is VSWITCH controller.
```

```
OPERATOR: HCPSWU2830I VSWITCH SYSTEM VMRTSW status is ready.
```

```
OPERATOR: HCPSWU2830I TCPIPLX is VSWITCH controller.
```

```
q ec00-ec02 eb00-eb02
```

```
OSA EC00 ATTACHED TO TCPIPLX EC00  
OSA EC01 ATTACHED TO TCPIPLX EC01  
OSA EC02 ATTACHED TO TCPIPLX EC02  
OSA EB00 ATTACHED TO TCPIPLX EB00  
OSA EB01 ATTACHED TO TCPIPLX EB01  
OSA EB02 ATTACHED TO TCPIPLX EB02
```

Create a vswitch called *vmrtsw* as a layer 3 using rdevices *ec00-ec02* and *eb00-eb02*. Choose any available controller

netstat devlink tcp tcpiplx

```
VM TCP/IP Netstat Level 510
```

```
Device VSWITCHDEV Type: VSWITCH-IUCV Status: Connected  
Queue size: 0 CPU: 0 IUCVid: *VSWITCH Priority: B  
Link VSWITCHLINK Type: IUCV Net number: 1  
BytesIn: 876 BytesOut: 1474  
Device VMRTSWEC00DEV Type: VSWITCH-OSD Status: Ready  
Queue size: 0 CPU: 0 Address: EC00 Port name: UNASSIGNED  
IPv4 Router Type: NonRouter Arp Query Support: Yes  
Link VMRTSWEC00LINK Type: QDIOETHERNET Net number: 0  
Transport Type: IP  
Broadcast Capability: Yes  
Multicast Capability: Yes  
Device VMRTSWEB00DEV Type: VSWITCH-OSD Status: Inactive  
Queue size: 0 CPU: 0 Address: EB00 Port name: UNASSIGNED  
IPv4 Router Type: NonRouter Arp Query Support: No  
Link VMRTSWEB00LINK Type: QDIOETHERNET Net number: 0  
Transport Type: IP  
Broadcast Capability: Unknown  
Multicast Capability: Unknown
```

Controllers: TCPIPLX and TCPIPLZ

- In their PROFILE TCPIP's this statement:

```
VSWITCH CONTROLLER ON
```

... but no need for HOME, GATEWAY, START statements ... unless there are other adapters

- DIRECTORY statement required:

```
IUCV *VSWITCH MSGLIMIT 65535
```

Allow these virtual machines to join the vswitch guest lan (class B) ... or SYSTEM CONFIG statement

```
set vswitch vmrtsw grant lfor0001
```

Command complete

```
set vswitch vmrtsw grant lforxx93
```

Command complete

```
set vswitch vmrtsw grant tcpiply
```

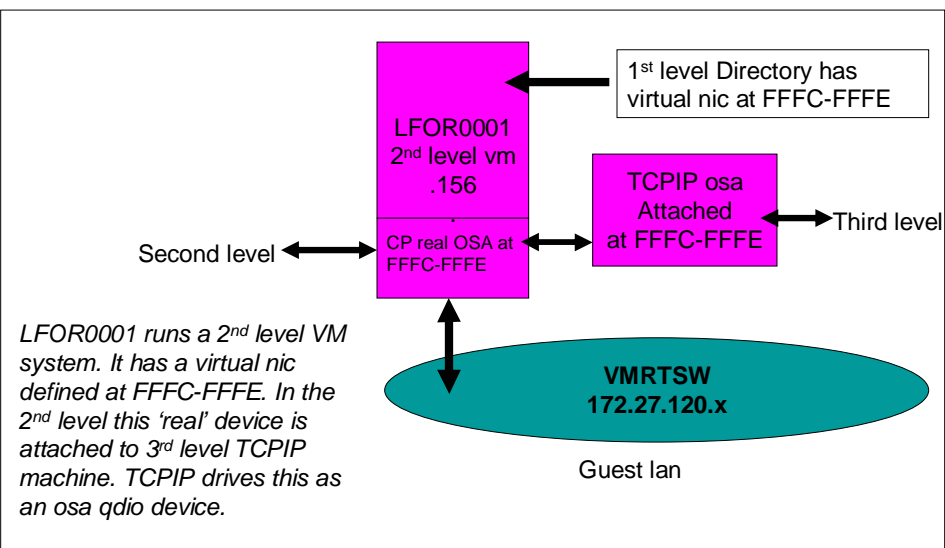
Command complete

Ask which machines have access

```

query vswitch
access
VSWITCH SYSTEM VMRTSW  Type: VSWITCH Connected: 3    Maxconn:
INFINITE
  PERSISTENT RESTRICTED    NONROUTER      Accounting: OFF
  VLAN unaware
  State: Ready
  ITimeout: 5              QueueStorage: 8
  Portname: UNASSIGNED RDEV: EC00  Controller:
TCPIPLZ  VDEV: EC00
  Portname: UNASSIGNED RDEV: EB00 Controller: TCPIPLZ  VDEV: EB00
BACKUP
Authorized userids:
LFORXX93 LFOR0001 SYSTEM TCPIPLY
  
```

Zoom in on the 2nd level STACK



Definitions for lfor0001

- First level directory:

NICDEF FFFC TYPE QDIO DEVICES 3 LAN SYSTEM VMRTSW

- Second level 'real' devices:

Q FFFC-FFFE

OSA FFFC ATTACHED TO TCPIP FFFC

OSA FFFD ATTACHED TO TCPIP FFFD

OSA FFFE ATTACHED TO TCPIP FFFE

PROFILE TCPIP

LFOR0001: TCPMAINT

```
DEVICE DEVFFFC OSD FFFC NONROUTER
LINK OSASERV QDIOETHERNET DEVFFFC MTU 1500
HOME
172.27.120.156 OSASERV
GATEWAY
172.27.0.0 = OSASERV 1500 0.0.255.0 0.0.120.0
DEFAULTNET 172.27.120.254 OSASERV 1500 0
START DEVFFFC
```

SYSTEM DTCPARMS

```
:nick.TCPIP :type.server
: :class.stack
:Attach.FFFC-FFFE
```

Lforxx93 Definitions

- Directory:

NICDEF FFFC TYPE QDIO DEVICES 3 MACID 01FF01 LAN SYSTEM VMRTSW

Macid is optional. It is appended to the MACID prefix. The MACID prefix is set in the SYSTEM CONFIG file in the VMLAN statement (VMLAN MACPREFIX xxxxxx). Default is 020000. Used by layer 2 vswitch support.

- Setup the card in the linux machine via yast or by hand

Setup the card in the linux machine via yast or by hand

- Via yast: must have working network in order to use ssh client (such as putty from windows).
 - This is for SUSE SLESx
- Via 3270 (no network access to linux) can use line editor such as sed
 - Useful when working with cloned machine

1. In yast select network devices/network card

YaST @ lforxx93
Press F1 for Help

YaST Control Center

Software Hardware System Network Devices Network Services Security and Users Misc	Network Card
--	---------------------

[Help] [Quit]

SHARE Session 9124 All Rights Reserved VM RESOURCES LTD. 29

2. Choose the card you wish to configure; configure

YaST @ lforxx93 Press F1 for Help

Network card setup

Configure your network card here.
Adding a network card:
Choose a network card from the list of detected network cards. If your network card was not autodetected, select Other (not detected) then press Configure

Editing or Deleting:
If you press Change, an additional dialog

Network cards configuration
Network cards to configure are:
Available

- IBM OSA Express Ethernet card (0.0.e706)
- IBM OSA Express Ethernet card (0.0.eb00)
- IBM OSA Express Ethernet card (0.0.fff0)**
- IBM IUCV
- Other (not detected)

[Configure...]

Already configured devices:

- * Hipersockets Interface (HSI)
Configured with Address 10.1.2.100
- * IBM OSA Express Ethernet card (0.0.88f0)
Configured with Address 0.0.0.0

[Change...]

[Back] [Abort] [Finish]

SHARE Session 9124 All Rights Reserved VM RESOURCES LTD. 30

3. Configure the card; choose next (then in the next screens click finish then quit

YaST @ lforxx93 Press F1 for Help

Configure your IP address.
Enter the IP address (e.g., 192.168.100.99) for your computer, the (usually 255.255.255.0), and, optionally, the default gateway IP address.
Contact your network administrator for more information about the network configuration.
Clicking Next

Network address setup

Configuration Name
qeth-bus-ccw-0.0.fff0

Static Address Setup
network mask |
IP Address Subnet mask
<172.27.120.159 <5.255.255.0

Detailed settings
Host name and name server
Routing
Advanced...

[Back] [Abort] [Next]

4. Choose finish; then quit yast

YaST @ lforxx93 Press F1 for Help

Network card setup
Configure your network card here.
(0.0.e705)
Adding a network card:
(0.0.e706)
Choose a network card from the (0.0.eb00)
list of detected network cards.
If your network card was not autodetected, select Other (not detected) then press Configure.
Editing or Deleting:
If you press Change, an additional dialog in which to change the configuration opens.

Network cards configuration

Network cards to configure
Available are:
IBM OSA Express Ethernet card

IBM OSA Express Ethernet card

IBM OSA Express Ethernet card

IBM IUCV
Other (not detected)

[Configure...]

* Already configured devices:
* Hipersockets Interface (HSI)
 Configured with Address 10.1.2.100
* IBM OSA Express Ethernet card (0.0.88f0)
 Configured with Address 0.0.0.0
* IBM OSA Express Ethernet card (0.0.e704)
 Configured with Address 172.27.120.155
* IBM OSA Express Ethernet card (0.0.fff0)
 Configured with Address 172.27.120.159
* IBM OSA Express Ethernet card (0.0.ffff)
 Configured with Address 172.27.120.155

[Change...]

[Abort] [Finish] [Back]

Configuring by hand

- Configuration files for network interfaces stored in /etc/sysconfig/network in suse sles9.
- Use sed or other line editor to change files.
- IBM device configurations stored in “online control block” file system /sys
- In the example, commands are done from the /etc/sysconfig/network directory.

Cloned machine has same IP as the master ... (just after cloning):

```
# cat ifcfg-qeth-bus-ccw-0.0.fff0
BOOTPROTO='static'
BROADCAST='172.27.120.255'
IPADDR='172.27.120.155'
MTU=''
NETMASK='255.255.255.0'
NETWORK='172.27.120.0'
REMOTE_IPADDR=''
STARTMODE='onboot'
UNIQUE='3IPn.FOqOuhDmSR4'
_nm_name='qeth-bus-ccw-0.0.fff0'
```

A cautionary tale: take a copy!!

```
cp ifcfg-qeth-bus-ccw-0.0.fff0
original.ifcfg-qeth-bus-ccw-0.0.fff0
```

Using sed "select lines with 155 and change to 159" in all lines and redirect output to new file temp:

```
sed s/155/159/g ifcfg-qeth-bus-ccw-0.0.fff0 > temp
sed s/155/159/g ifcfg-qeth-bus-ccw-0.0.fff0 <work # sed s/155/159/g
ifcfg-qeth-b
us-ccw-0.0.fff0 > temp
```

Display the file just created by output redirection:

```
# cat temp
cat temp
BOOTPROTO='static'
BROADCAST='172.27.120.255'
IPADDR='172.27.120.159'
MTU=''
NETMASK='255.255.255.0'
NETWORK='172.27.120.0'
REMOTE_IPADDR=''
STARTMODE='onboot'
UNIQUE='3IPn.FOqOuhDmSR4'
_nm_name='qeth-bus-ccw-0.0.fff0'
```

Rename the file:

```
# mv temp ifcfg-qeth-bus-ccw-0.0.fff0
mv temp ifcfg-qeth-bus-ccw-0.0.fff0
```

Display the configuration file:

```
# cat ifcfg-qeth-bus-ccw-0.0.fff0
cat ifcfg-qeth-bus-ccw-0.0.fff0
BOOTPROTO='static'
BROADCAST='172.27.120.255'
IPADDR='172.27.120.159'
MTU=''
NETMASK='255.255.255.0'
NETWORK='172.27.120.0'
REMOTE_IPADDR=''
STARTMODE='onboot'
UNIQUE='3IPn.FOqOuhDmSR4'
_nm_name='qeth-bus-ccw-0.0.fff0'
```

Still had the old configuration; needs to be changed

```
# ifconfig eth0
ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 02:00:00:01:FF:01
          inet addr:172.27.120.155  Bcast:172.27.120.255
          Mask:255.255.255.0
          inet6 addr: fe80::200:0:100:5/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1492  Metric:1
          errors:0 dropped:0 overruns:0 frame:0
          TX packets:6 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:2632 (2.5 Kb)  TX bytes:652 (652.0 b)
```

Take the link down

```
# ifdown eth0
ifdown eth0
eth0
eth0      configuration: qeth-bus-ccw-0.0.fff0
```

bring the link up

```
# ifup eth0
ifup eth0
eth0
eth0      configuration: qeth-bus-ccw-0.0.fff0
```

Interface is now up

```
# ifconfig eth0
ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 02:00:00:01:FF:01
          inet addr:172.27.120.159  Bcast:172.27.120.255
          Mask:255.255.255.0
          inet6 addr: fe80::200:0:100:5/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1492  Metric:1
          RX packets:24 errors:0 dropped:0 overruns:0 frame:0
          TX packets:13 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:3402 (3.3 Kb)  TX bytes:1422 (1.3 Kb)
```

Startup Messages

```
linux version 2.6.5-7.97-s390x (geeko@buildhost) (gcc version
3.3.3 (S Use linux)
) #1 SMP Fri Jul 2 14:21:59 UTC 2004
We are running under VM (64 bit mode)
:
qeth: loading qeth S/390 OSA-Express driver ($Revision:
1.77.2.20 $/$Revision: 1
.98.2.11 $/$Revision: 1.27.2.5 $/$Revision: 1.8.2.2
$/$Revision: 1.7.2.1 $/$Revi
sion: 1.5.2.4 $/$Revision: 1.19.2.7 $ :IPV6 :VLAN)
qeth: Device 0.0.ffff/0.0.ffff/0.0.ffff is a Guest LAN QDIO
card (level: v511)
with link type GuestLAN QDIO (portname:)
qeth: IP fragmentation not supported on eth0
qeth: VLAN enabled
qeth: Multicast enabled
qeth: IPV6 enabled
qeth: Broadcast enabled
SHARE Session 9124 All Rights Reserved VM RESOURCES LTD. 39
```

Definitions for TCPIPLY

Directory statement for TCPIPLY:

```
NICDEF 0800 TYPE QDIO DEVICES 3 LAN SYSTEM VMRTSW
```

PROFILE TCPIP

```
DEVICE DEV@0800 OSD 0800 NONROUTER
LINK OSASERV QDIOETHERNET DEV@0800 MTU 1500
HOME
172.27.120.158 OSASERV
GATEWAY
172.27.0.0 = OSASERV 1500 0.0.255.0 0.0.120.0
DEFAULTNET 172.27.120.254 OSASERV 1500 0
START DEV@0800
```

VSWITCH Presentation Checkpoint

At this point:

- VSWITCH VMRTSW defined
- 3 virtual machines permitted to use it
- Stacks connected to VSWITCH on virtual nics:
 - LFOR0001: 2nd level VM system with TCPIP machine at 172.27.120.156
 - LFORXX93 linux machine at 172.27.120.159
 - TCPIPLY VM TCPIP stack machine at 172.27.120.158
- Additional stack machine sharing OSA port at IP address 172.27.120.155
- Two controller machines, TCPIPLZ and TCPIPLX

Will Now Show ...

- Network management commands
 - netstat
 - ping
 - Failover:
 - Device removal
 - Controller failure
 - During recovery two applications active: FTP (large transfer) and TELNET. Both applications remained available during and after recovery processing.

Before tcpip in lfor0001 joins

```
netstat arp all tcp tcpiplx ← Query the arp
VM TCP/IP Netstat Level 510      cache of the
Querying ARP cache for address * controller machine
Adapter-maintained data as of: 07/07/05 14:24:41
OSA mac
Link VMRTSWEC00LINK : QDIOETHERNET: 00025509E705 IP: 172.27.120.155
Link VMRTSWEC00LINK : QDIOETHERNET: 00025509E705 IP: 172.27.120.158
Link VMRTSWEC00LINK : QDIOETHERNET: 00025509E705 IP: 172.27.120.159
Link VMRTSWEC00LINK : QDIOETHERNET: 080020E46479 IP: 172.27.120.254
```

Physical switch mac

After LFOR0001 joins

```
netstat arp all tcp tcpiplx
VM TCP/IP Netstat Level 510
Querying ARP cache for address * Joins the arp
Adapter-maintained data as of: 07/07/05 14:35:01 table
Link VMRTSWEC00LINK : QDIOETHERNET: 00025509E705 IP: 172.27.120.155
Link VMRTSWEC00LINK : QDIOETHERNET: 00025509E705 IP: 172.27.120.156
Link VMRTSWEC00LINK : QDIOETHERNET: 00025509E705 IP: 172.27.120.158
Link VMRTSWEC00LINK : QDIOETHERNET: 00025509E705 IP: 172.27.120.159
Link VMRTSWEC00LINK : QDIOETHERNET: 080020E46479 IP: 172.27.120.254
```

First level pings from TCPIPLY

ping 172.27.120.156

Ping Level 510: Pinging host 172.27.120.156.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.002 seconds. Successes so far 1.

ping 172.27.120.158

Ping Level 510: Pinging host 172.27.120.158.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.001 seconds. Successes so far 1.

ping 172.27.120.159

Ping Level 510: Pinging host 172.27.120.159.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.001 seconds. Successes so far 1.

ping 172.27.120.155

Ping Level 510: Pinging host 172.27.120.155.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.001 seconds. Successes so far 1.

Second level pings from TCPIP in LFOR0001

ping 172.27.120.156

Ping Level 510: Pinging host 172.27.120.156.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.001 seconds. Successes so far 1.

ping 172.27.120.158

Ping Level 510: Pinging host 172.27.120.158.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.001 seconds. Successes so far 1.

ping 172.27.120.159

Ping Level 510: Pinging host 172.27.120.159.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.001 seconds. Successes so far 1.

ping 172.27.120.155

Ping Level 510: Pinging host 172.27.120.155.
Enter 'HX' followed by 'BEGIN' to interrupt.
PING: Ping #1 response took 0.001 seconds. Successes so far 1.

linux pings 1 of 2

```
lforxx93:~ # ping -c 1 172.27.120.156
PING 172.27.120.156 (172.27.120.156) 56(84) bytes of data.
64 bytes from 172.27.120.156: icmp_seq=1 ttl=60 time=0.588 ms

--- 172.27.120.156 ping statistics ---
1 packets transmitted, 1 received, 0% packet loss, time 0ms
rtt min/avg/max/mdev = 0.588/0.588/0.588/0.000 ms
lforxx93:~ # ping -c 1 172.27.120.158
PING 172.27.120.158 (172.27.120.158) 56(84) bytes of data.
64 bytes from 172.27.120.158: icmp_seq=1 ttl=60 time=0.225 ms

--- 172.27.120.158 ping statistics ---
1 packets transmitted, 1 received, 0% packet loss, time 0ms
rtt min/avg/max/mdev = 0.225/0.225/0.225/0.000 ms
```

linux pings 2 of 2

```
lforxx93:~ # ping -c 1 172.27.120.159
PING 172.27.120.159 (172.27.120.159) 56(84) bytes of data.
64 bytes from 172.27.120.159: icmp_seq=1 ttl=64 time=0.064 ms

--- 172.27.120.159 ping statistics ---
1 packets transmitted, 1 received, 0% packet loss, time 0ms
rtt min/avg/max/mdev = 0.064/0.064/0.064/0.000 ms
lforxx93:~ # ping -c 1 172.27.120.155
PING 172.27.120.155 (172.27.120.155) 56(84) bytes of data.
64 bytes from 172.27.120.155: icmp_seq=1 ttl=60 time=0.664 ms

--- 172.27.120.155 ping statistics ---
1 packets transmitted, 1 received, 0% packet loss, time 0ms
rtt min/avg/max/mdev = 0.664/0.664/0.664/0.000 ms
```


QUERY VSWITCH VMRTSW DETAILS

```
VSWITCH SYSTEM VMRTSW Type: VSWITCH Connected: 3 Maxconn: INFINITE
PERSISTENT RESTRICTED NONROUTER Accounting: OFF
VLAN Unaware
State: Ready
IPTimeout: 5 QueueStorage: 8
Portname: UNASSIGNED RDEV: EC00 Controller: TCPIPLZ VDEV: EC00
Portname: UNASSIGNED RDEV: EB00 Controller: TCPIPLZ VDEV: EB00 BACKUP
VSWITCH Connection:
RX Packets: 8878 Discarded: 4 Errors: 0
TX Packets: 9215 Discarded: 0 Errors: 0
RX Bytes: 800654 TX Bytes: 1911124
239.255.255.253 MAC: 01-00-5E-7F-FF-FD
FFFE::1 MAC: 33-33-00-00-00-01 Local
FFFE::1:FFFD:FFFE MAC: 33-33-FF-01-FF-02 Local
:
```

1 of 3 ...

QUERY VSWITCH VMRTSW DETAILS

```
:
Adapter Owner: LFORXX93 NIC: FFFC Name: UNASSIGNED
RX Packets: 568 Discarded: 0 Errors: 0
TX Packets: 276 Discarded: 0 Errors: 0
RX Bytes: 74526 TX Bytes: 41076
Device: FFFE Unit: 002 Role: DATA
Options: Broadcast Multicast IPv6 IPv4 VLAN
Unicast IP Addresses:
172.27.120.159 MAC: 02-00-00-01-FF-02
FE80::200:0:201:FF02 MAC: 02-00-00-01-FF-02 Local
Multicast IP Addresses:
224.0.0.1 MAC: 01-00-5E-00-00-01
224.0.0.251 MAC: 01-00-5E-00-00-FB
:
```

2 of 3 ...

QUERY VSWITCH VMRTSW DETAILS

3 of 3 ...

```
:
Adapter Owner: LFOR0001 NIC: FFFC Name: UNASSIGNED
  RX Packets: 135      Discarded: 0      Errors: 0
  TX Packets: 49      Discarded: 0      Errors: 0
  RX Bytes: 33273    TX Bytes: 6902
  Device: FFFE Unit: 002 Role: DATA
  Options: Broadcast Multicast IPv4 VLAN
  Multicast IP Addresses:
    172.27.120.156      MAC: 02-00-00-00-00-04
  Multicast IP Addresses:
    224.0.0.1          MAC: 01-00-5E-00-00-01
Adapter Owner: TCPIPLY NIC: 0800 Name: UNASSIGNED
  RX Packets: 126      Discarded: 0      Errors: 0
  TX Packets: 31      Discarded: 0      Errors: 0
  RX Bytes: 31768    TX Bytes: 5210
  Device: 0802 Unit: 002 Role: DATA
  Options: Broadcast Multicast IPv4 VLAN
  Unicast IP Addresses:
    172.27.120.158      MAC: 02-00-00-00-00-02
    224.0.0.1          MAC: 01-00-5E-00-00-01
```

SHARE Session 9124

All Rights Reserved VM RESOURCES LTD.

51

Before removing the rdevs

```
q ec00-ec02 eb00-eb02
OSA EC00 ATTACHED TO TCPIPLX EC00
OSA EC01 ATTACHED TO TCPIPLX EC01
OSA EC02 ATTACHED TO TCPIPLX EC02
OSA EB00 ATTACHED TO TCPIPLX EB00
OSA EB01 ATTACHED TO TCPIPLX EB01
OSA EB02 ATTACHED TO TCPIPLX EB02
q vswitch vmrtsw
VSWITCH SYSTEM VMRTSW Type: VSWITCH Connected: 4 Maxconn:
INFINITE
PERSISTENT RESTRICTED NONROUTER Accounting:
OFF
VLAN Unaware
State: Ready
IPTimeout: 5 QueueStorage: 8
Portname: UNASSIGNED RDEV: EC00 Controller: TCPIPLX VDEV: EC00
Portname: UNASSIGNED RDEV: EB00 Controller: TCPIPLX VDEV: EB00
BACKUP
```

SHARE Session 9124

All Rights Reserved VM RESOURCES LTD.

52

Remove the Rdevs

```
det ec00-ec02 tcpiplx
```

```
TCPIPLX : EC00-EC02 DETACHED BY TCPMAINT
EC00-EC02 DETACHED TCPIPLX
TCPIPLX : 17:19:22 DTCOSD082E VSWITCH-OSD shutting down:
HCPSWU2830I VSWITCH SYSTEM VMRTSW status is devices attached.
HCPSWU2830I TCPIPLX is VSWITCH controller.
HCPSWU2830I VSWITCH SYSTEM VMRTSW status is in error recovery.
HCPSWU2830I TCPIPLX is new VSWITCH controller.
```

Also have performed a cable pull. Recovery proceeds similar to detaching the real devices

TCPIPLX Recovery Messages 1 of 2

```
TCPIPLX : 17:19:22 DTCPRI385I Device VMRTSWEC00DEV:
TCPIPLX : 17:19:22 DTCPRI386I Type: VSWITCH-OSD, Status: Ready
TCPIPLX : 17:19:22 DTCPRI387I Envelope queue size: 0
TCPIPLX : 17:19:22 DTCPRI388I Address: EC00
TCPIPLX : 17:19:22 DTCQDI001I QDIO device VMRTSWEC00DEV device number
EC02:
TCPIPLX : 17:19:22 DTCQDI007I Disable for QDIO data transfers
TCPIPLX : 17:19:22 DTCOSD361I VSWITCH-OSD link removed for VMRTSWEC00DEV
TCPIPLX : 17:19:22 DTCOSD080I VSWITCH-OSD initializing:
TCPIPLX : 17:19:22 DTCPRI385I Device VMRTSWEB00DEV:
TCPIPLX : 17:19:22 DTCPRI386I Type: VSWITCH-OSD, Status: Not started
TCPIPLX : 17:19:22 DTCPRI387I Envelope queue size: 0
TCPIPLX : 17:19:22 DTCPRI388I Address: EB00
TCPIPLX : 17:19:22 DTCQDI001I QDIO device VMRTSWEB00DEV dev number EB02:
TCPIPLX : 17:19:22 DTCQDI007I Enabled for QDIO data transfers
```

TCPIPLX Recovery Messages 2 of 2

```
TCPIPLX : 17:19:22 DTCOSD238I ToOsd: IPv4 multicast support enabled for VMRTSWEB00DEV
TCPIPLX : 17:19:22 DTCOSD319I ProcessSetArpCache: Supported for device VMRTSWEB00DEV
TCPIPLX : 17:19:22 DTCOSD341I Obtained MAC address 000255899D45 for device VMRTSWEB00DEV
TCPIPLX : 17:19:22 DTCOSD238I ToOsd: IPv6 multicast support enabled for VMRTSWEB00DEV
TCPIPLX : 17:19:22 DTCOSD360I VSWITCH-OSD link added for VMRTSWEB00DEV
HCPSWU2830I VSWITCH SYSTEM VMRTSW status is ready.
HCPSWU2830I TCPIPLX is VSWITCH controller.
TCPIPLX : 17:19:26 DTCOSD246I VSWITCH-OSD device VMRTSWEB00DEV: Assigned IPv4 address 172.27.120.159
TCPIPLX : 17:19:26 DTCOSD246I VSWITCH-OSD device VMRTSWEB00DEV: Assigned IPv4 address 172.27.120.156
TCPIPLX : 17:19:26 DTCOSD246I VSWITCH-OSD device VMRTSWEB00DEV: Assigned IPv4 address 172.27.120.158
```

Kill Controller Machine

```
q controller
Controller TCPIPLX Available: YES VDEV Range: * Level 510
Capability: IP ETHERNET VLAN_ARP
SYSTEM VMRTSW Primary Controller: * VDEV: EC00
SYSTEM VMRTSW Backup Controller: * VDEV: EB00
```

force tcpipx

```
USER DSC LOGOFF AS TCPIPLX USERS = 16 FORCED BY TCPMNLAB
HCPSWU2843E The path was severed for TCP/IP Controller TCPIPLX.
HCPSWU2843E It was managing device EC00 for VSWITCH SYSTEM VMRTSW.
HCPSWU2843E The path was severed for TCP/IP Controller TCPIPLX.
HCPSWU2843E It was managing device EB00 for VSWITCH SYSTEM VMRTSW.
```

Recovery controller messages 1 of 2

```
TCPIPLZ : 17:22:14 DTCOSD360I VSWITCH-OSD link added for VMRTSWEC00DEV
TCPIPLZ : 17:22:14 DTCOSD080I VSWITCH-OSD initializing:
TCPIPLZ : 17:22:14 DTCPRI385I Device VMRTSWEC00DEV:
TCPIPLZ : 17:22:14 DTCPRI386I Type: VSWITCH-OSD, Status: Not started
TCPIPLZ : 17:22:14 DTCPRI387I Envelope queue size: 0
TCPIPLZ : 17:22:14 DTCPRI388I Address: EC00
TCPIPLZ : 17:22:14 DTCQDI001I QDIO device VMRTSWEC00DEV device number EC02:
TCPIPLZ : 17:22:14 DTCQDI007I Enabled for QDIO data transfers
TCPIPLZ : 17:22:14 DTCOSD238I ToOsd: IPv4 multicast support enabled for
VMRTSWEC00DEV
TCPIPLZ : 17:22:14 DTCOSD319I ProcessSetArpCache: supported for device
VMRTSWEC00DEV
TCPIPLZ : 17:22:14 DTCOSD341I obtained MAC address 00025509E705 for device
VMRTSWEC00DEV
TCPIPLZ : 17:22:14 DTCOSD238I ToOsd: IPv6 multicast support enabled for
VMRTSWEC00DEV
```

Recovery controller messages 2 of 2

```
HCPSWU2830I VSWITCH SYSTEM VMRTSW status is ready.
HCPSWU2830I TCPIPLZ is VSWITCH controller.
TCPIPLZ : 17:22:14 DTCOSD360I VSWITCH-OSD link added for
VMRTSWEC00DEV
TCPIPLZ : 17:22:18 DTCOSD246I VSWITCH-OSD device VMRTSWEC00DEV:
Assigned IPv4 address 172.27.120.159
TCPIPLZ : 17:22:18 DTCOSD246I VSWITCH-OSD device VMRTSWEC00DEV:
Assigned IPv4 address 172.27.120.156
TCPIPLZ : 17:22:18 DTCOSD246I VSWITCH-OSD device VMRTSWEC00DEV:
Assigned IPv4 address 172.27.120.158
```

Additional Documentation

- REDP-3719-00 linux on IBM zSeries and S/390: VSWITCH and VLAN Features of z/VM 4.4
- SC24-6080-00 z/VM V5R1.0 Connectivity Guide chapter 2 and more
- SC24-6125-00 z/VM V5R1.0 TCP/IP Planning and Customization
- GC24-6102 z/VM 5.1 Getting Started with Linux on zSeries
- SC33-8289-01 linux on system z/9 and z/series Device Drivers, Features, and Command

Penultimate thoughts

- Recovery based on CP artifacts as opposed to, say, VIPA methods.
- Extends existing network topologies horizontally.
- No need for additional subnets once you transcend cultural barriers with network administrator.
- Ideally suited to linux virtual machine environments.

Final Thoughts

- Wow!
- Recovery of both failures took just a few seconds.
- VSWITCHes can also support VLANs – not discussed today.
- Recommended approach to linux on z/VM networks.
- Remember: CP manages the devices and the switch table.