IBM GLOBAL SERVICES

E18

VSE Gets Connected With zSeries

Stephen Gracin

zSeries Expo | Nov. 1 - 5, 2004

Miami, FL

RETURN TO INDEX

Connecting TCPIP for VSE With Linux / zVM

And one foil on SNA

- 09/08/2004 -

Stephen Gracin
gracinsp@us.ibm.com

Ground Rules

For this discussion unless otherwise noted

- VSE or VSE/ESA is VSE/ESA2.7
- zVM is zVM 4.4
- zSeries Processor refers to z800, z900, z890, z990
- TCPIP for VSE is the CSI TCPIP

Moving Forward

Moving from VSE 2.2OLD to the most current VSE/ESA on a ZSeries processor?

Are you:

- Going to the latest zVM or add zVM ?
- Going to add Linux ?
- IP networking options with/without zVM ?
- IP networking options with zSeries hardware ?

- There is a combination for your environment. Which one is best?

zSeries Hardware
Connectivity Considerations

zSeries Hardware Connectivity

- OSA/Express cards – Some can operate in two modes, one of which supports SNA.

- HIPERSockets adapter – High speed zero latency Ethernet LAN device, can connect IP stacks within and across LPARs

- IFL's – ***I***ntegrated ***F***acility for ***L***inux, this is a zSeries engine specifically to support Linux. Not directly a networking option but can have an effect on topology.

zSeries Hardware – OSA/Express Cards

- May not require OSA/SF for setup, but OSA/SF can always be used to QUERY a card.
- Some can be set to one of two modes. QDIO or Express. This is set in the IOCP, Express mode is LCS mode.
- In QDIO or LCS mode, the cards can be shared among LPARs. They can provide IP connectivity between stacks in different LPARs but the HIPERSockets adapter is the preferred connection between LPARs.
- QDIO mode is ***Q***ueued ***D***ata ***IO*** different than TIO/SIO, SSCH. Until recently not supported by TCPIP for VSE, and there is no plan for support by VTAM.

zSeries Hardware – HIPERSockets adapter

- Is a hardware implementation that is defined to a zSeries processor via the IOCP .
- Provides high speed connection between TCPIP stacks in different LPARs or in the same LPAR.
- In some Redbooks is referred too as an iQDIO device, ***i***nternal ***QDIO*** the CHPID type is IQD.
- Not to be confused with a zVM Guest Lan TYPE=HIPER, which is an emulation of this device by zVM.

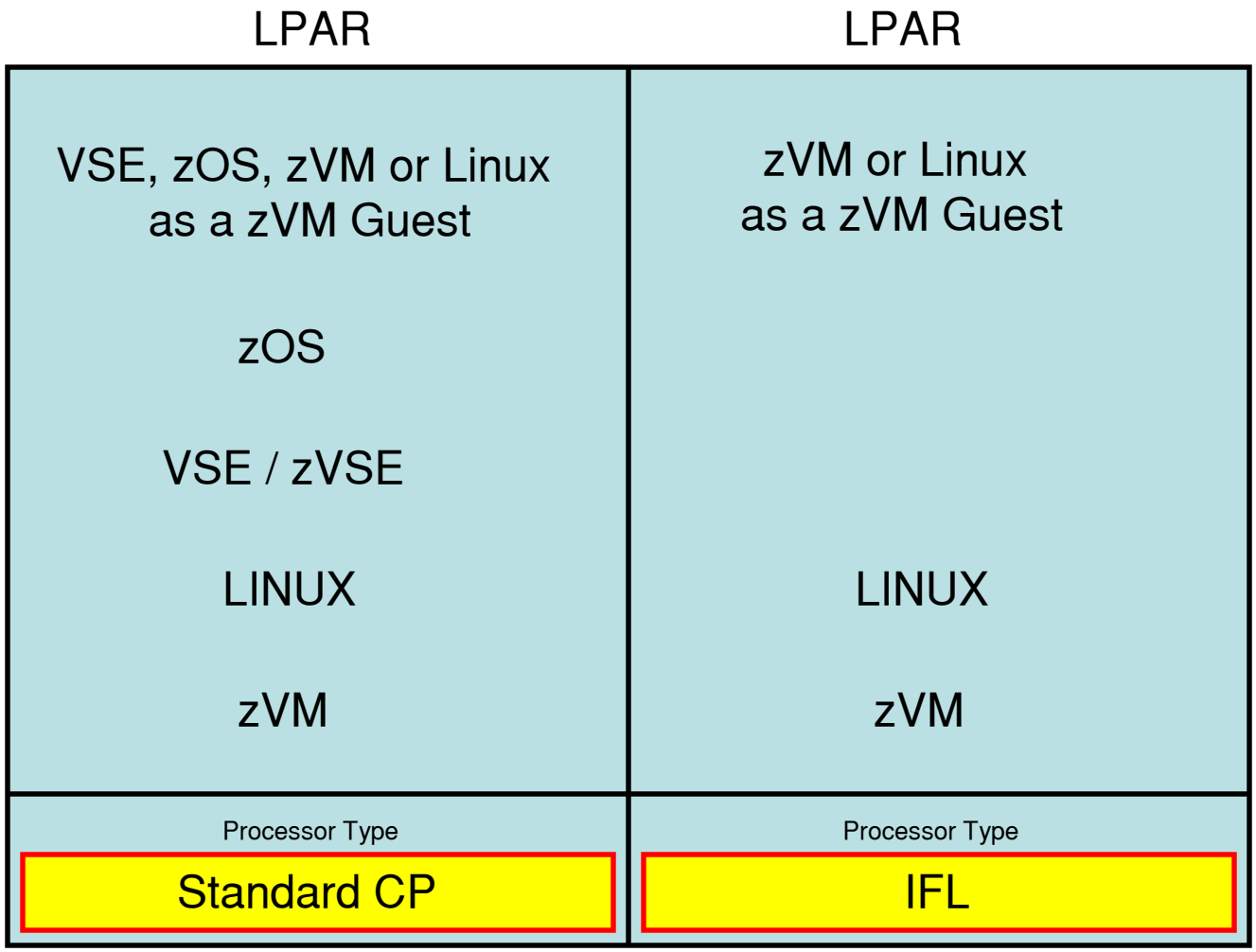
zSeries Hardware - IFL

An ***I***ntegrated ***F***acility for ***L***inux engine …

- Supports zVM or Linux. Or zVM with Linux guests.
- Does Not Support VSE/ESA or zOS, they report a processor malfunction.
- Does Not Support VSE/ESA, zOS running as a guest on zVM. The guests report a processor malfunction.

This means that a VSE connecting to a Linux would most likely be across LPARs, the HIPERSockets adapter would be the best choice.

zSeries Hardware - IFL



Software Connectivity Options

zVM4.4

Provides Two Guest LAN types:

- QDIO – Each guest machine (VM USERID) has a virtual NIC defined that emulates a Gigabit OSA/E. They become hosts on a LAN defined in zVM's SYSTEM CONFIG file.

- HIPERSockets – Each guest machine (VM USERID) has a virtual NIC that emulates a whatever on a HIPERSockets adapter. They become hosts on a LAN defined in the SYSTEM CONFIG file.

Provides Point-to-Point:

- IUCV – Inter-user communications vehicle a memory transfer p2p network device that can be used zVM to zVM , Linux to Linux or zVM to Linux. VSE does not support this connection.

- VCTC – Virtual Channel to Channel P2P connection supported by VSE, zVM and Linux.

zVM4.4

VSWITCH – *V*irtual *S*witch

- A special QDIO Guest LAN. The added VSWITCH function can provide VLAN capability. This can employ 802.1Q VLAN tagging where a VLAN ID is added to the Ethernet frame. If with SET VSWITCH the VLAN tagging is set not to filter (VLAN ANY) VSWITCH provides LAN “bridge like” function

So what’s a switch ?

- Kind of like a Hub in that it connects Ethernet cards and therefore IP stacks controlling those cards together. With a Hub a packet on one port is seen on all ports simultaneously, this leads to collisions.
  A Switch receives an Ethernet frame on one of its ports and has the electronics to read in the Frame and hold it. This allows Full Duplex communications and the ability to act upon the information like adding a VLAN ID to a Ethernet Frame or interrogating the IP addresses in a Packet.
  Put similar function in zVM/CP and it is a VSWITCH

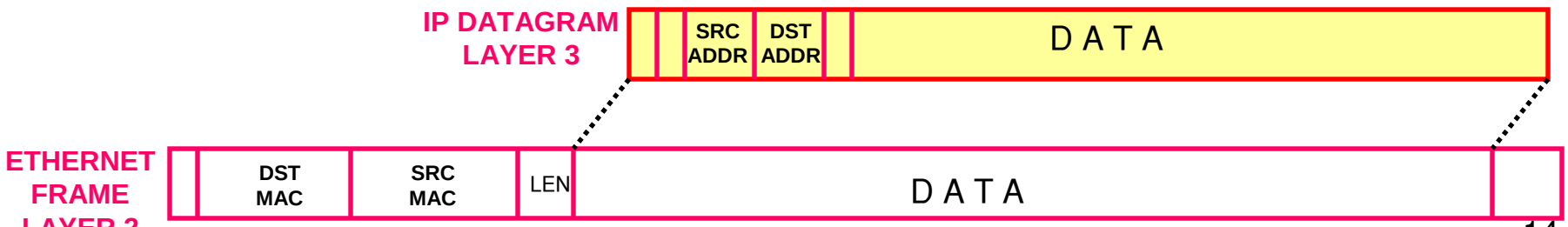
What a switch Looks at

OSI Stack Model

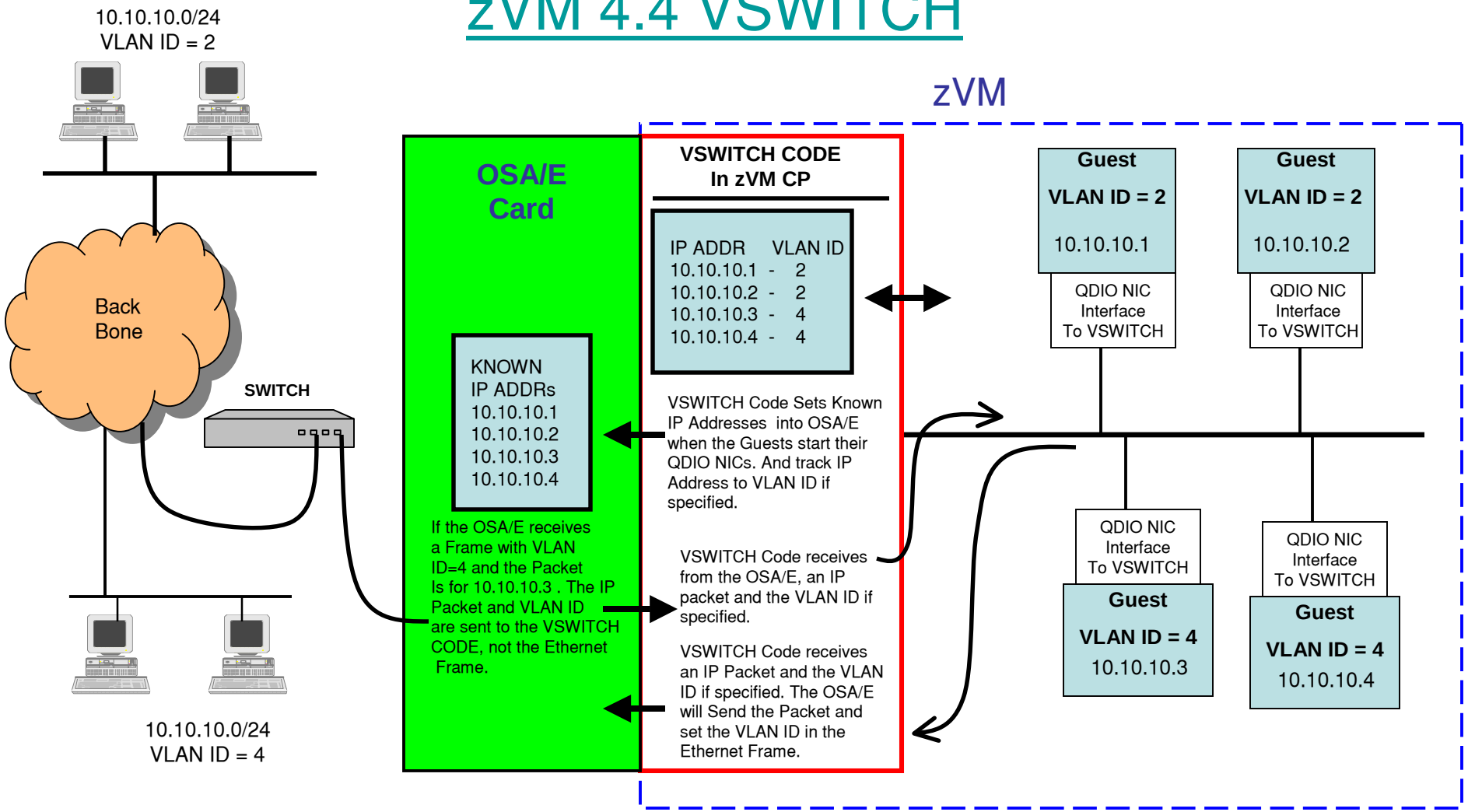
LAYER 7 – APPLICATION
LAYER 6 – PRESENTATION
LAYER 5 – SESSION
LAYER 4 – TRANSPORT – Example - **TCP**
LAYER 3 – NETWORK – Example - **IP** **DATAGRAM**
LAYER 2 – DATA LINK – Example – **Ethernet** **FRAME**
LAYER 1 – PHYSICAL - Wires , OSA/E

You are Here !

A layer 2 SWITCH is DATA LINK layer based upon MAC addresses in IP DATAGRAM. Layer 3 switch is based upon IP addresses. 802.1Q VLAN ID tag is in the Ethernet FRAME.



zVM 4.4 VSWITCH



VSWITCH changes with zVM5.1 VSWITCH is a true Layer 2 switch

VSWITCH / VLAN

- Can Provide a “Bridge Like” function for IP packets
- Not really a layer 2 switch, Ethernet frames are not exchanged in zVM4.4 they will be in zVM5.1 and later.
- Is a specialized ODIO Guest LAN
- Requires an OSA/E card in QDIO mode
- Can provide Hardware redundancy with two OSA/E cards.
- Can use/manage VLAN IDs, Does Not Have too.
- VSE does not have a 801.2Q driver. VSE can not manipulate a Ethernet frame to set a VLAN ID itself, but zVM can control a VLAN ID for the VSE guest. Or without any VLAN IDs you still get an easy to manage “Flat Network” with OSA hardware redundancy.

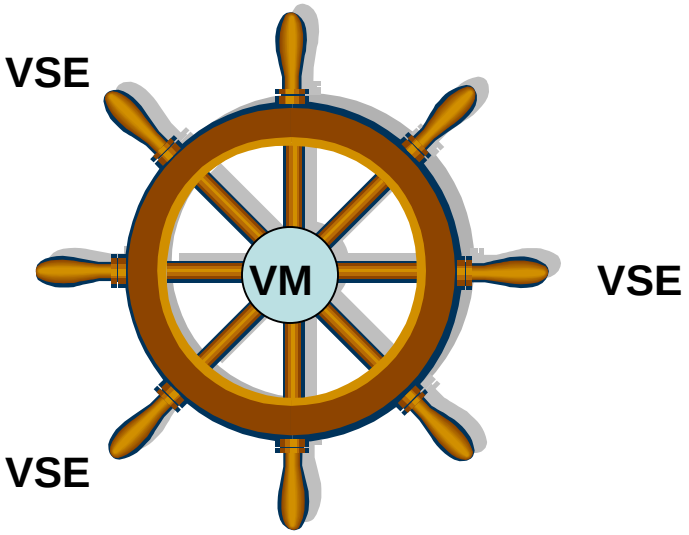
VSE Connectivity History

Connectivity History

A hub-and-spoke topology using point-to-point connections is most often seen with older releases of VSE/ESA running on VM/ESA. One topology with two IP addressing schemes.

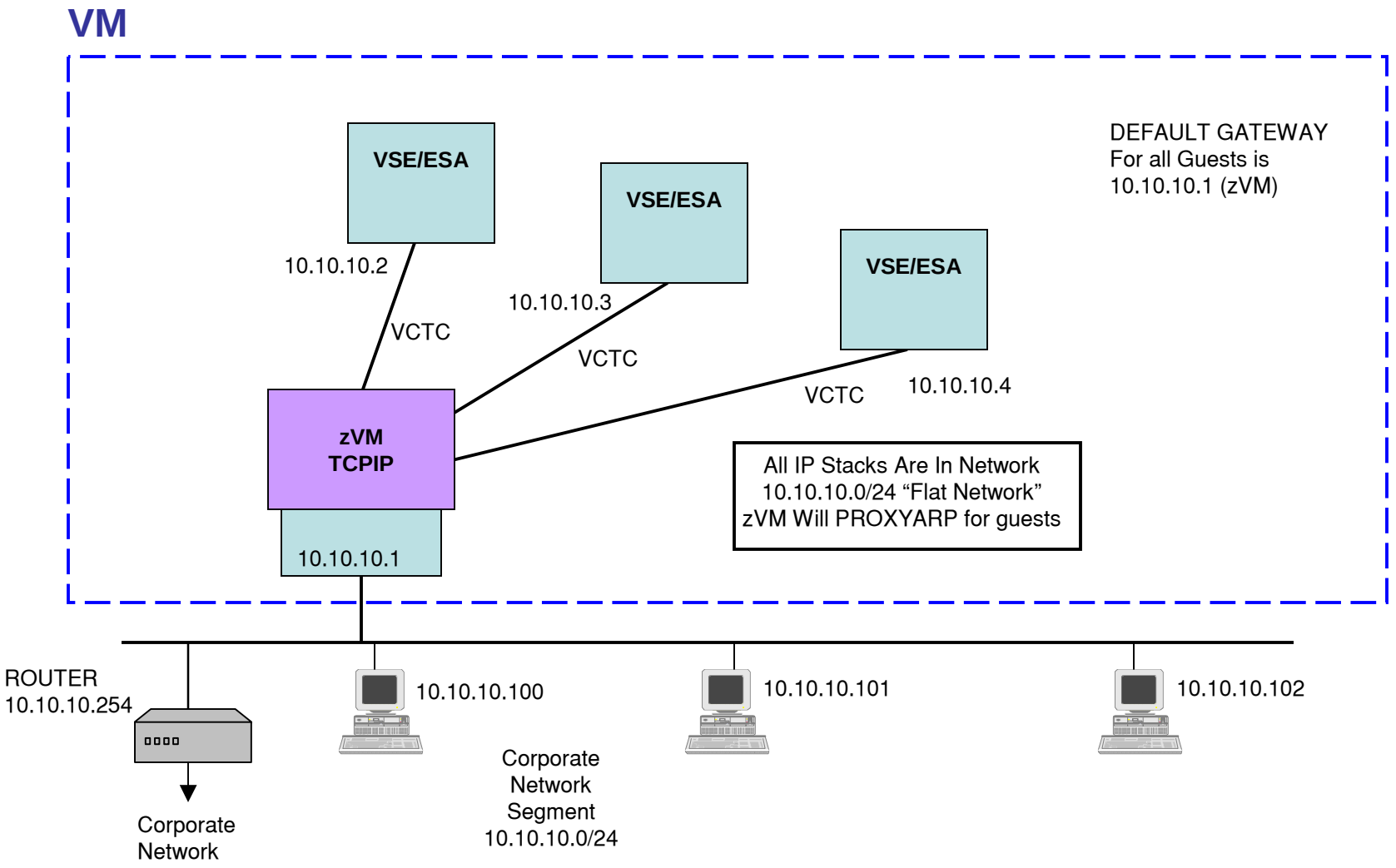
1. The guests are in the same network as the VM stack and its IP controller. This scheme, the “Flat Network”, is usually the easiest to integrate into an existing network.

2. The guests are in a different IP network.



Flat Network

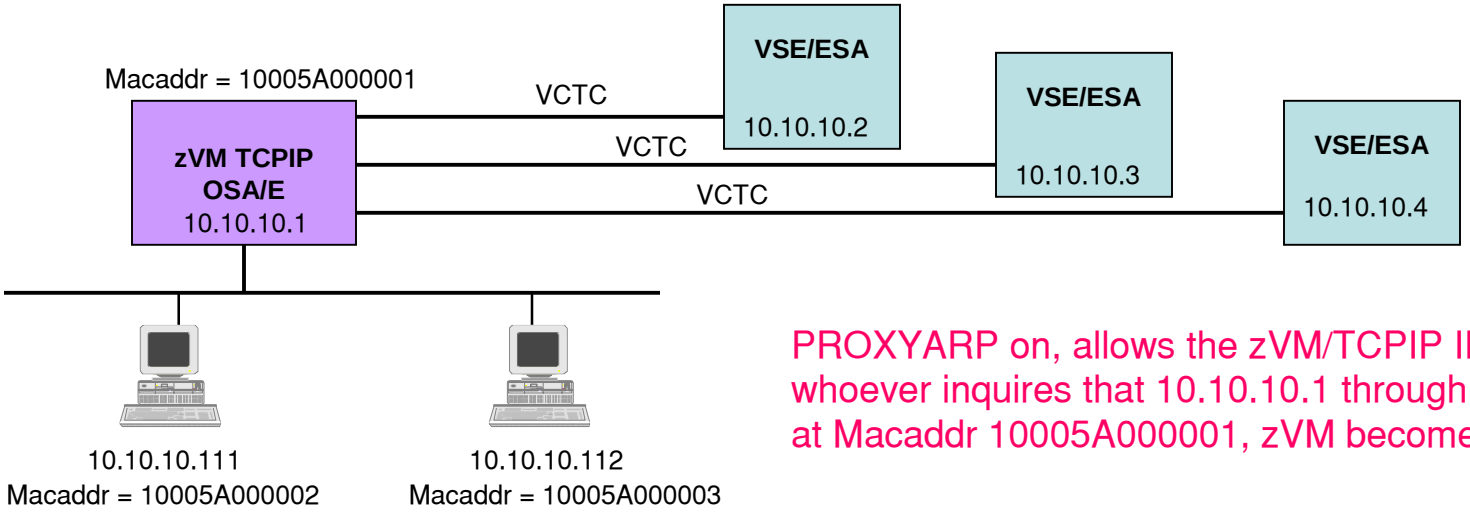
All the VSE/ESA guests and zVM are in the same IP network the ASSORTEDPARMS statement PROXYARP makes this possible.



What's PROXYARP

ARP - **A**ddress **R**esolution **P**rotocol mapping of an IP address to a machine address, an address assigned to the NIC when manufactured.

PROXYARP – Where a machine will respond to (ARP) requests for hosts other than its self.



PROXYARP on, allows the zVM/TCPIP IP Stack to tell whoever inquires that 10.10.10.1 through 10.10.10.4 are at Macaddr 10005A000001, zVM becomes their proxy.

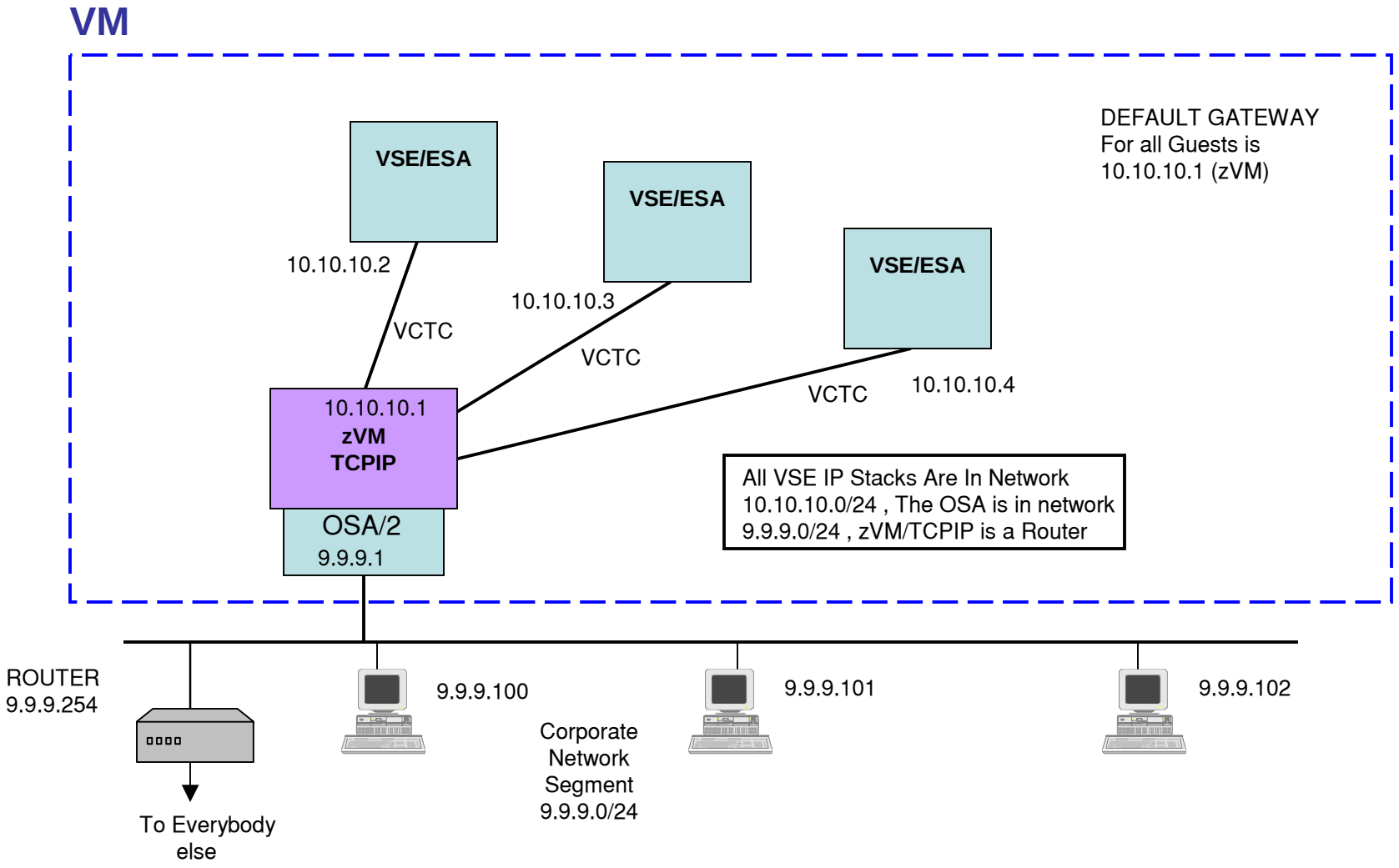
In file PROFILE TCPIP

```
; ------------------------
 ASSORTEDPARMS
   PROXYARP
```

The ASSORTEDPARMS statement PROXYARP when uncommented enables the function

Different Network

VSE/ESA connected to zVM via VCTC connections as this is what VSE supported. VM stack which is the router. This example the guests are in a different network than the OSA card.



OSA/E

OSA card settings to be aware of. Not understanding these settings has caused many problems when trying to use a stack as a router.

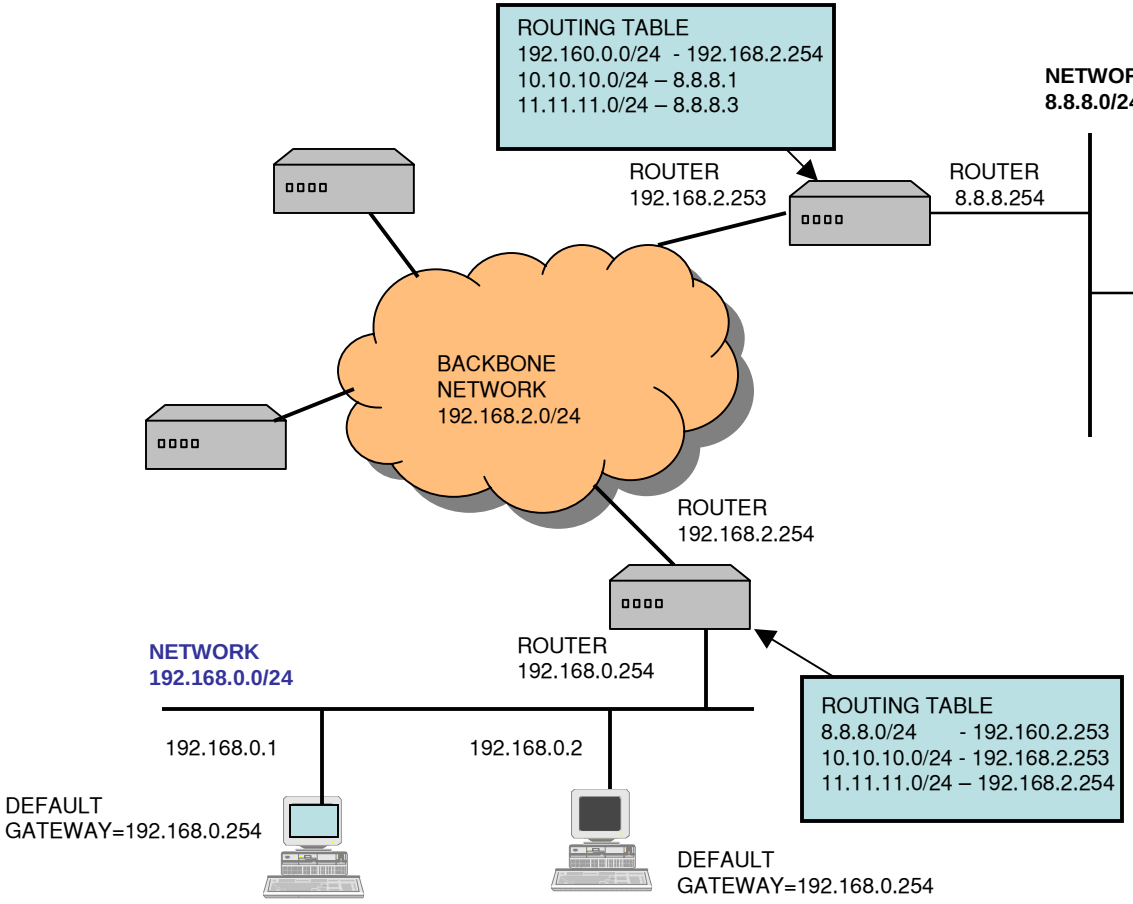
Primary Router – The stack that has this value will be presented with packets that have any destination IP address. You want this setting if your stack is a router. There can only be one Primary at a time per OSA port.

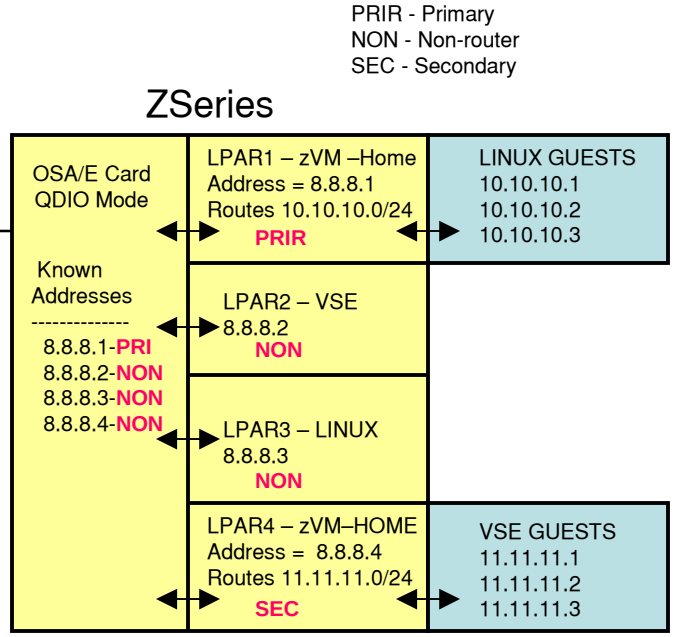
Secondary Router – The stack with this setting will be a non-router when the primary stack is up. If the primary stack goes down, this stack becomes primary until the original primary comes back up. This setting would be useful for a redundant stack/card backup scenario.

Non-router – This stack will only be presented with *Known Addresses*. This would be an IP address loaded into the OSA card when the stack initializes, such as its HOME address. Only packets whose destination address is the HOME address would be presented to the stack by the OSA card. You do not want this if your stack is a router.

Note: At this time VSE/TCPIP can not set Primary, it is always a non-router. zVM, zOS, Linux can be Primary.

OSA Cards – PRIR / SEC / NON





Host 192.168.0.1 Tries to telnet to Linux Guest 10.10.10.2 and is successful.

Host 192.168.0.2 Tries to telnet to VSE Guest 11.11.11.1 and is Not successful. But it can TN3270 to zVM 8.8.8.3

The OSA card has a table of Known Addresses that are set as the stacks become active. When the card receives a packet it checks the Known Addresses, if there is a match the packet is given to that stack. If no match , before discarding the packet the card checks if any stack is PRIR . The setting of PRIR, Primary Router means this one stack, and there can only be one PRIR, is presented all non- matched packets.

Host 192.168.0.1 sends a packet with a Destination Address of 10.10.10.2 which is not a known address. But the zVM/TCPIP stack in LPAR1 is PRIR so it is given the packet. Host 192.168.0.2 sends a packet with a Destination Address of 11.11.11.1 which is not Known and this stack is not PRIR so the packet is discarded. It can TN3270 to VSE since 8.8.8.2 is a Known Address.

VSE/ESA

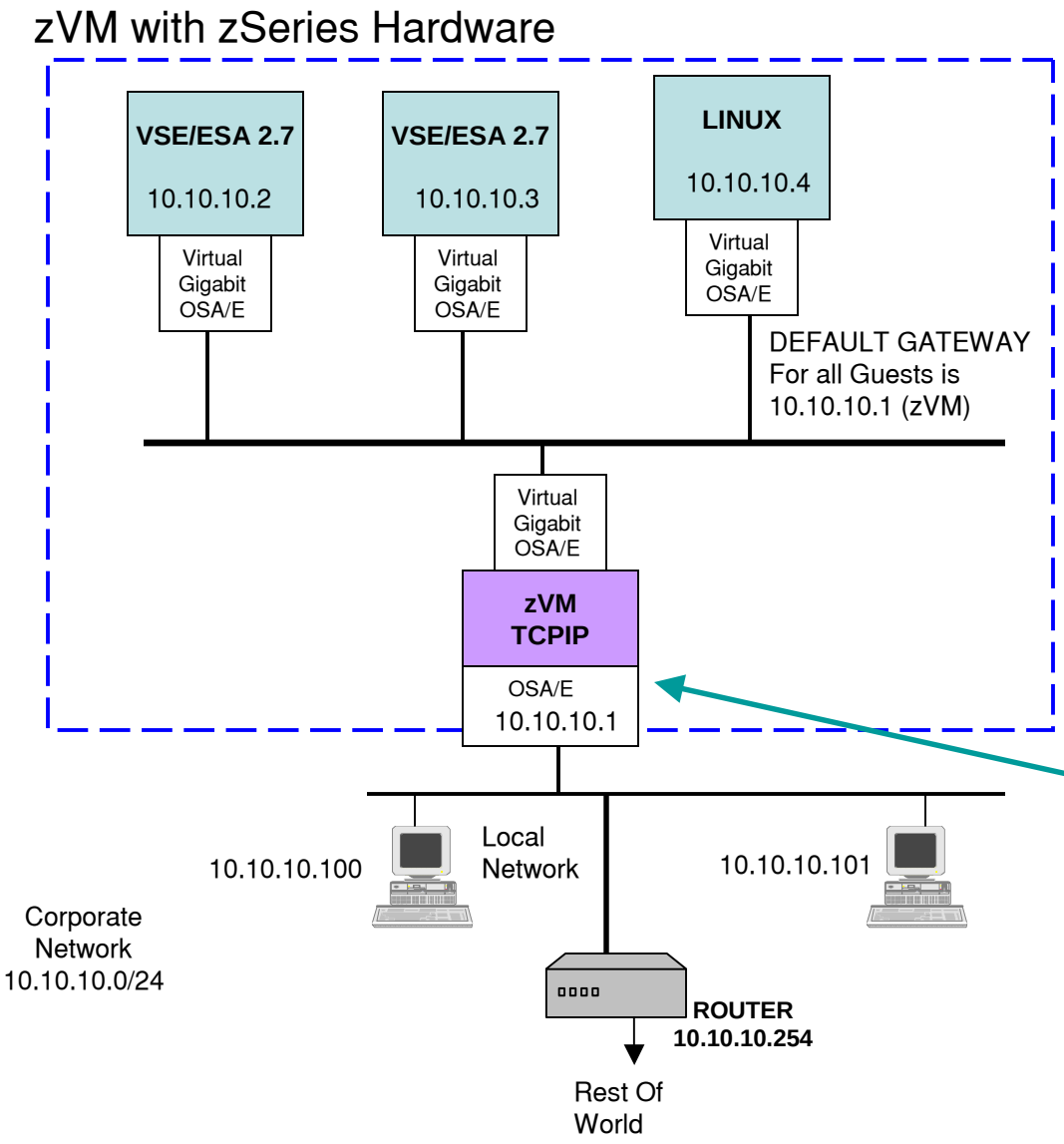
Moving forward assumes that we will upgrade to VSE/ESA 2.7 and zVM 4.4

- VSE/ESA 2.7 brings support for QDIO and HIPERSocket adapters
- zVM4.4 brings support for QDIO and HIPERSocket guest lan emulation The guest lan emulation allows each guest to operate a virtual adapter (NIC). Each guest has a virtual adapter connecting it to the guest LAN.
- zVM4.4 also bring support for VSWITCH which allows the guest machines To bridge to the corporate network. Definitions for the zVM/TCPIP stack In PROFILE TCPIP may not be necessary.

Migration Observations

- With the latest VSE , zVM and hardware the number of topology options Increases dramatically. There is usually more than one way to provide a function.
- The changes may be all inside the zSeries processor, not apparent to the external network structure.
- The apparent changes will be in performance, ease of use, system management
- Using the latest hardware/software we have for example, at least four ways to migrate our original Flat Network.

Migrate Flat Network – To Guest LAN QDIO or HIPER



Each Guest Machine and zVM TCPIP has a virtual Gigabit OSA/E defined as their interface to the Guest LAN. The Virtual NIC (Gigabit card) is Defined in the Guest machines Directory Entry.

zVM TCPIP routes packets between the guest LAN And the real OSA/E to local network.

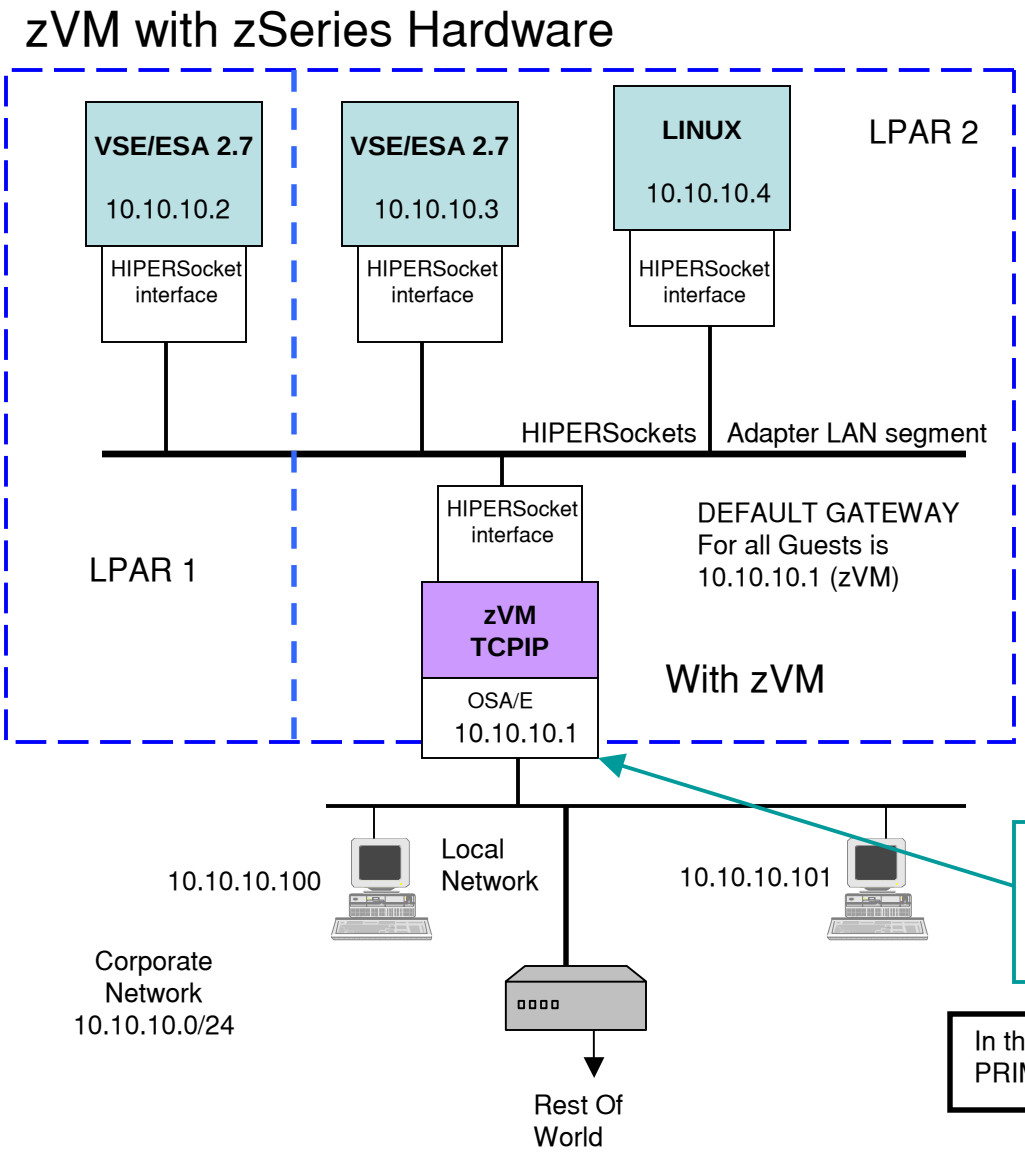
zVM4.3 and 4.4 allow PROXYARP for a guest LAN (Broadcast Network) once APAR PQ80577 is Applied. You must have a HOST routes in zVM/TCPIP For the Guests for PROXYARP to work.
A NETWORK route will not work. If PROXYARP and HOST routes are not defined in PROFILE TCPIP the Guests can not be in the 10.10.10.0/24 corporate network

The guest LAN (VM/CP) carries interguest traffic the zVM TCPIP stack is not required for this function.

FYI - The zVM stack routes for guest machines On the Guest LAN and provides TN3270 access For zVM access.

In this case zVM needs the OSA card as PRIMARY ROUTER.

Migrate Flat Network - To HIPERSockets Adapter



The HIPERSockets adapter is part of the zSeries Hardware. The devices the Guest machines and zVM TCPIP use are defined in the IOCP.

zVM TCPIP routes packets between the HIPERSockets Adapter LAN and the OSA/E to local network.

zVM4.3 and 4.4 allow PROXYARP for a guest LAN (Broadcast Network) once APAR PQ80577 is Applied. You must have a HOST routes in zVM/TCPIP For the Guests for PROXYARP to work.
A NETWORK route will not work. If PROXYARP And HOST routes are not used the VSE/ESA guests can not be in the 10.10.10.0/24 corporate network

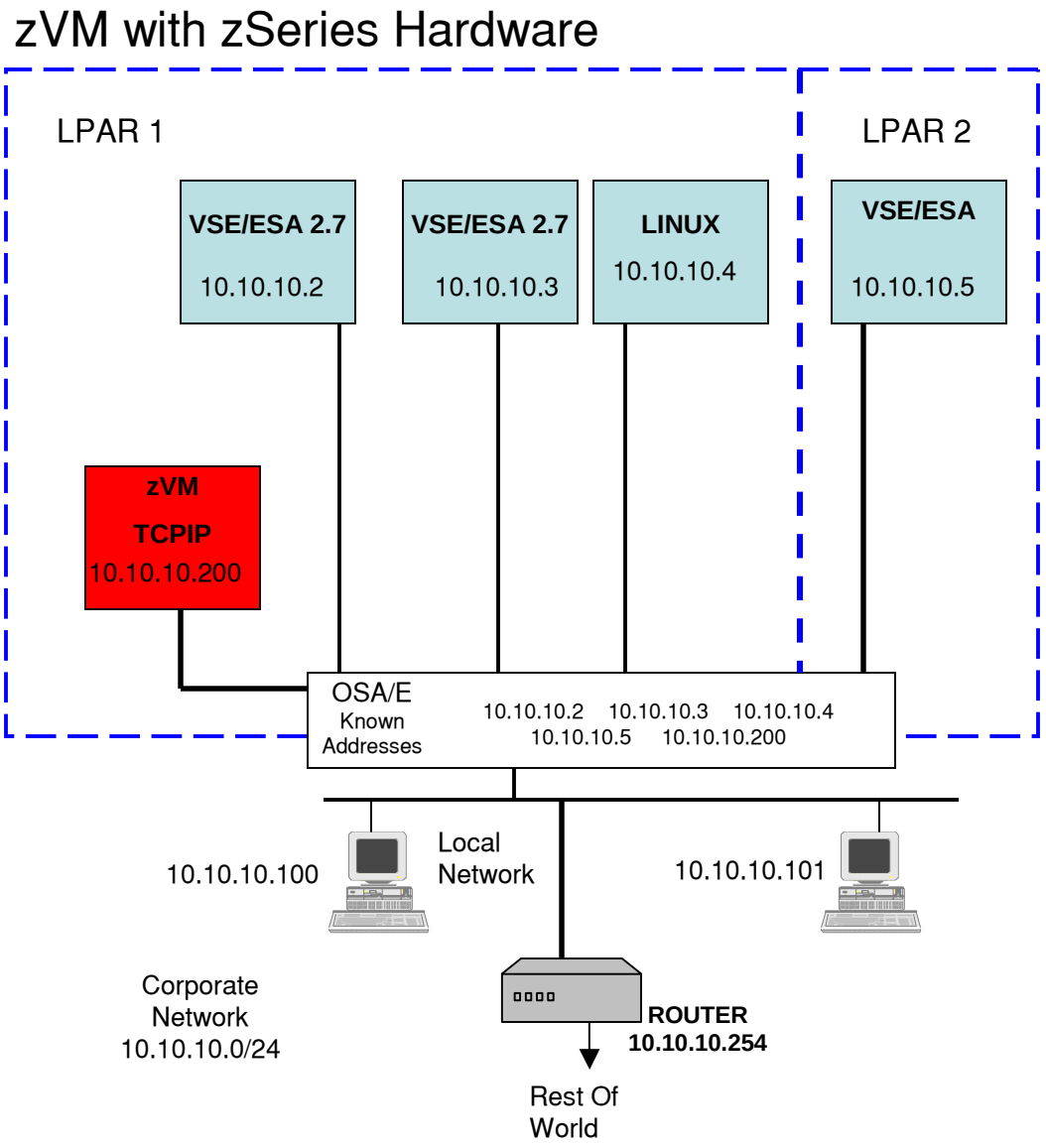
The HIPERSockets Adapter LAN carries interguest traffic, the zVM TCPIP stack is not needed for this function. Since it is a hardware solution it can connect guest machines in different LPARs.

The HIPERSockets adapter is very fast.

FYI – The zVM/TCPIP stack routes for the Hosts on the HIPERSocket Adapter LAN Segment and also provides 3270 connectivity.

In this case zVM needs the OSA card as PRIMARY ROUTER.

Migrate Flat Network - To Direct OSA/E Attach



Each VSE/ESA guest is given a set of Sub channel Addresses with which to access the OSA card.

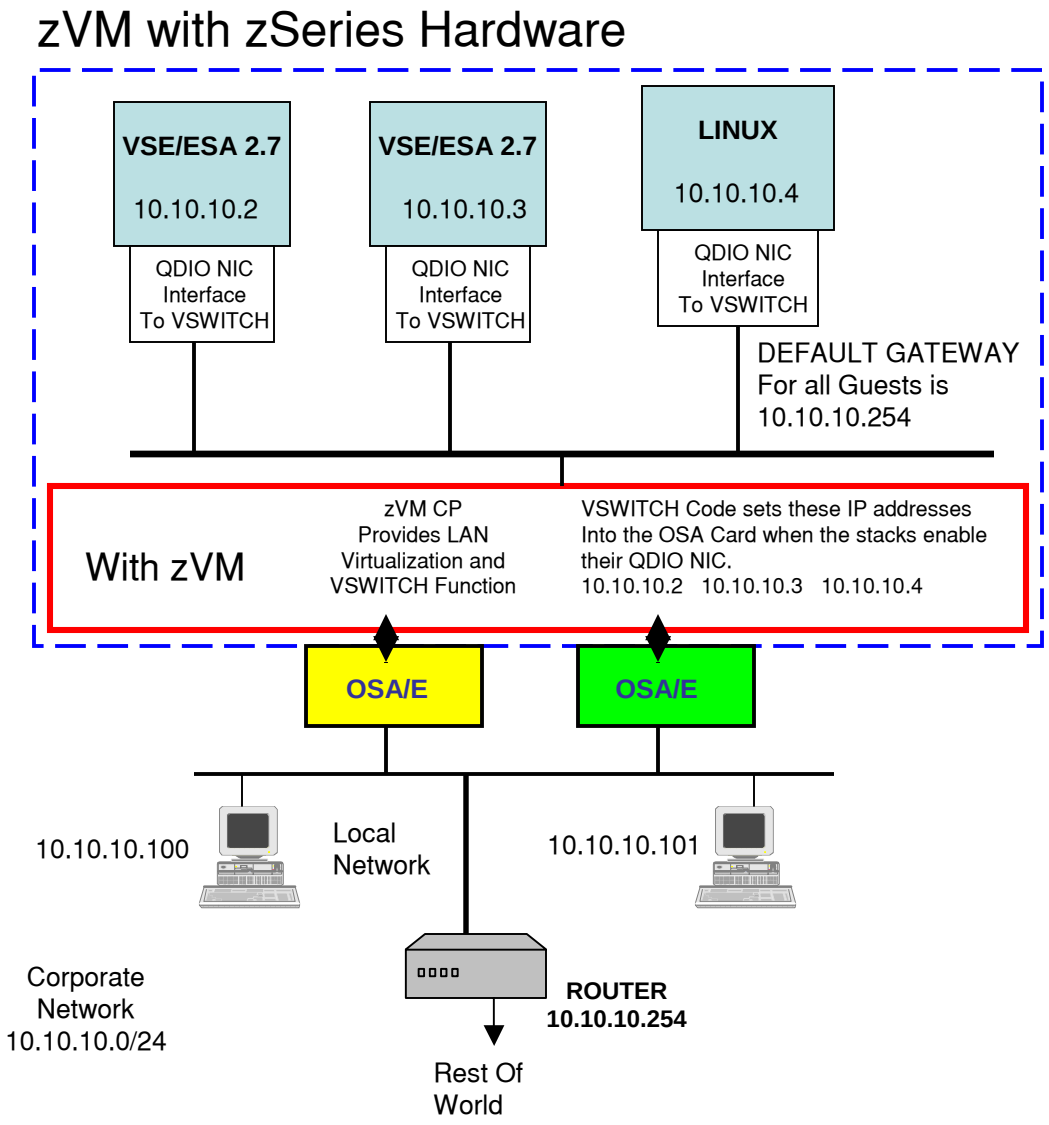
The OSA/E card builds an internal ARP table to track the IP address of stacks connected to the card. If a packet is destined to go to 10.10.10.4 from 10.10.10.2 then the packet does not leave the card to be routed.

Since the OSA/E sub channel addresses can be Shared across LPARs this is a interguest and a LPAR-to-LPAR IP connectivity option. Not as fast as the HIPERSockets Adapter.

You can define some sub channel addresses of the OSA/E card for use by zVM/TCPIP for TN3270 access. Otherwise a 2074 or OSA3270 is required for non-SNA 3270 device access.

In this case all stacks can be NON-ROUTER. None have networks behind them they are routing for.

MOVE - To VSWITCH



The VSWITCH by its nature provides a "flat network" topology, it acts like an IP bridge between the virtual Network and the corporate network. At zVM4.4 VSWITCH can provide VLAN support and "bridge like" function but it is not a layer 2 switch.

With VSWITCH zVM may not have a HOME IP address, a PROFILE TCPIP for the zVM IP stacks is not required For VSWITCH to function.

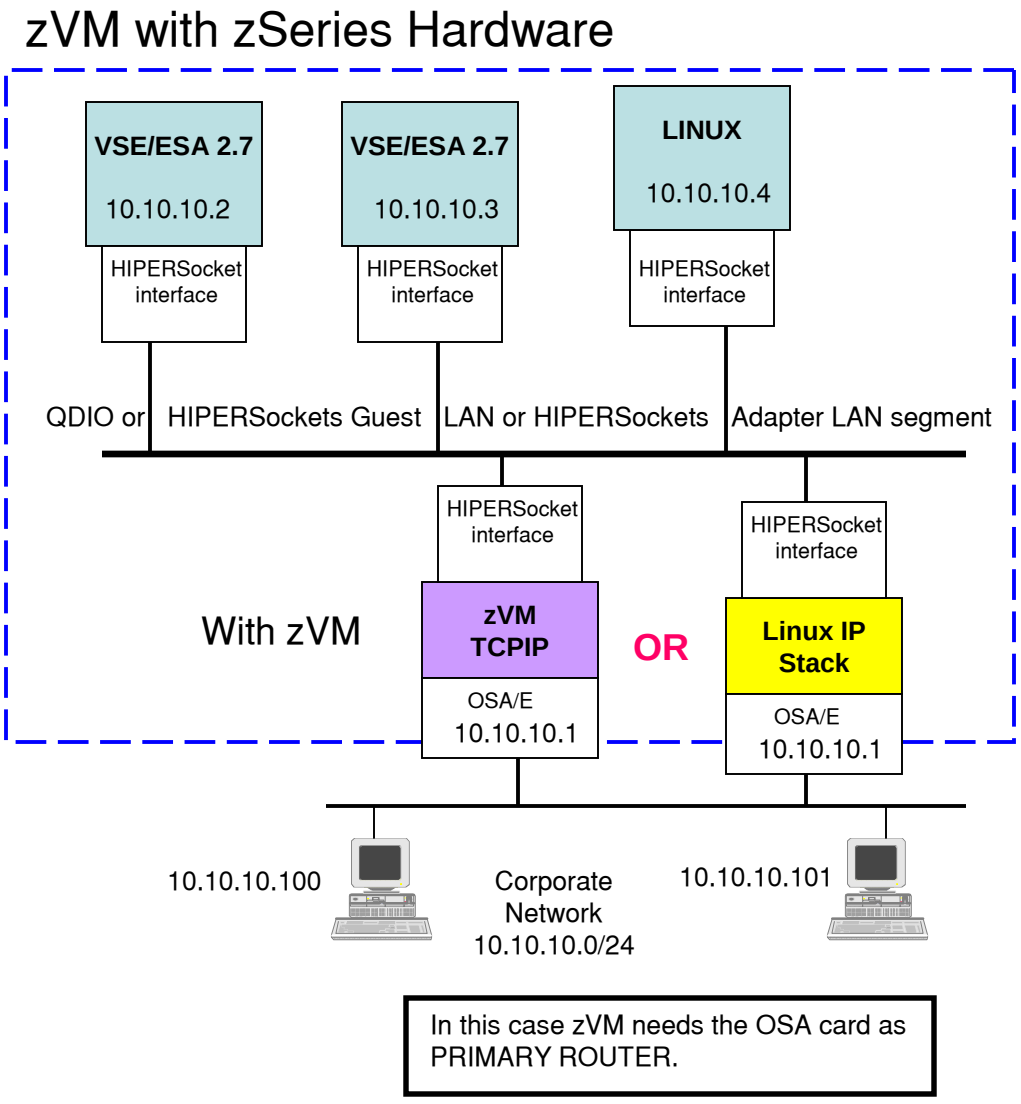
The VSWITCH can function with one OSA/E. Two OSA/E cards can be defined to the VSWITCH. One active the other provides hardware backup.

You can define some sub channel addresses of one of the OSA/E cards for use by zVM/TCPIP for TN3270 Access. Otherwise a 3174 / 2074 device is required for non-SNA 3270 device access.

The nature of VSWITCH makes PRI-ROUTER vs. NON-ROUTER not an issue. As the Guest Machines initialize their virtual NICs the information is also added to the OSA/E by the VSWITCH function. Their IP addresses become *Known Addresses* to the OSA/E

Interface To The World !
Which Stack?

Which Stack?



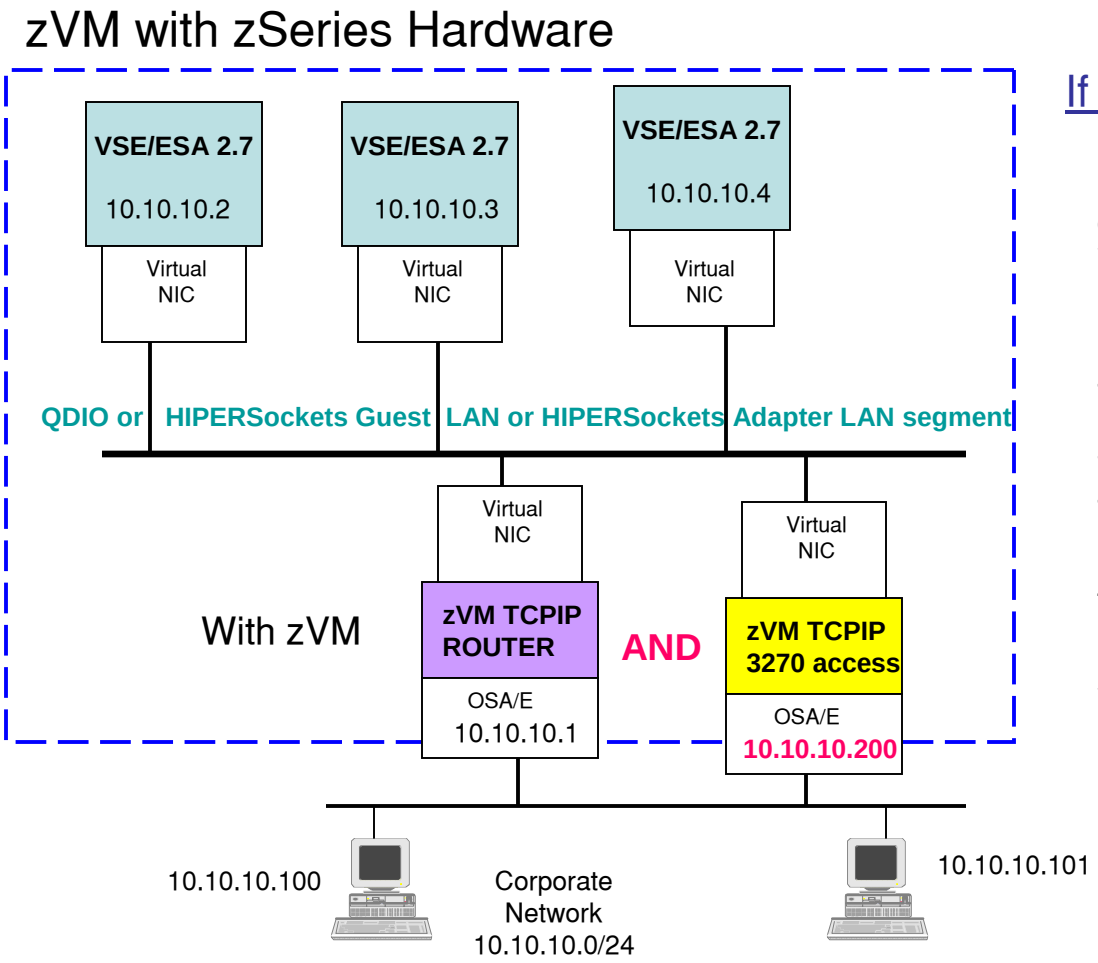
zVM or a Linux stack can be a Router for a QDIO or HIPERSockets Guest LAN or a HIPERSockets adapter LAN.

Some differences are:

The zVM/TCPIP stack can provide 3270 Connectivity to zVM for system programming activities, where the Linux stack can not. MPROUTE is built into zVM for dynamic routing protocols another package has to be added to Linux.

The Linux stack can provide a firewall like IPTABLES , functions like NAT, IP Aliasing that zVM does not have.

Two (Or more) zVM Stacks!



If using one stack as a ROUTER

In this case zVM stack with HOME address of 10.10.10.1 is the ROUTER for the VSE/ESA machines.

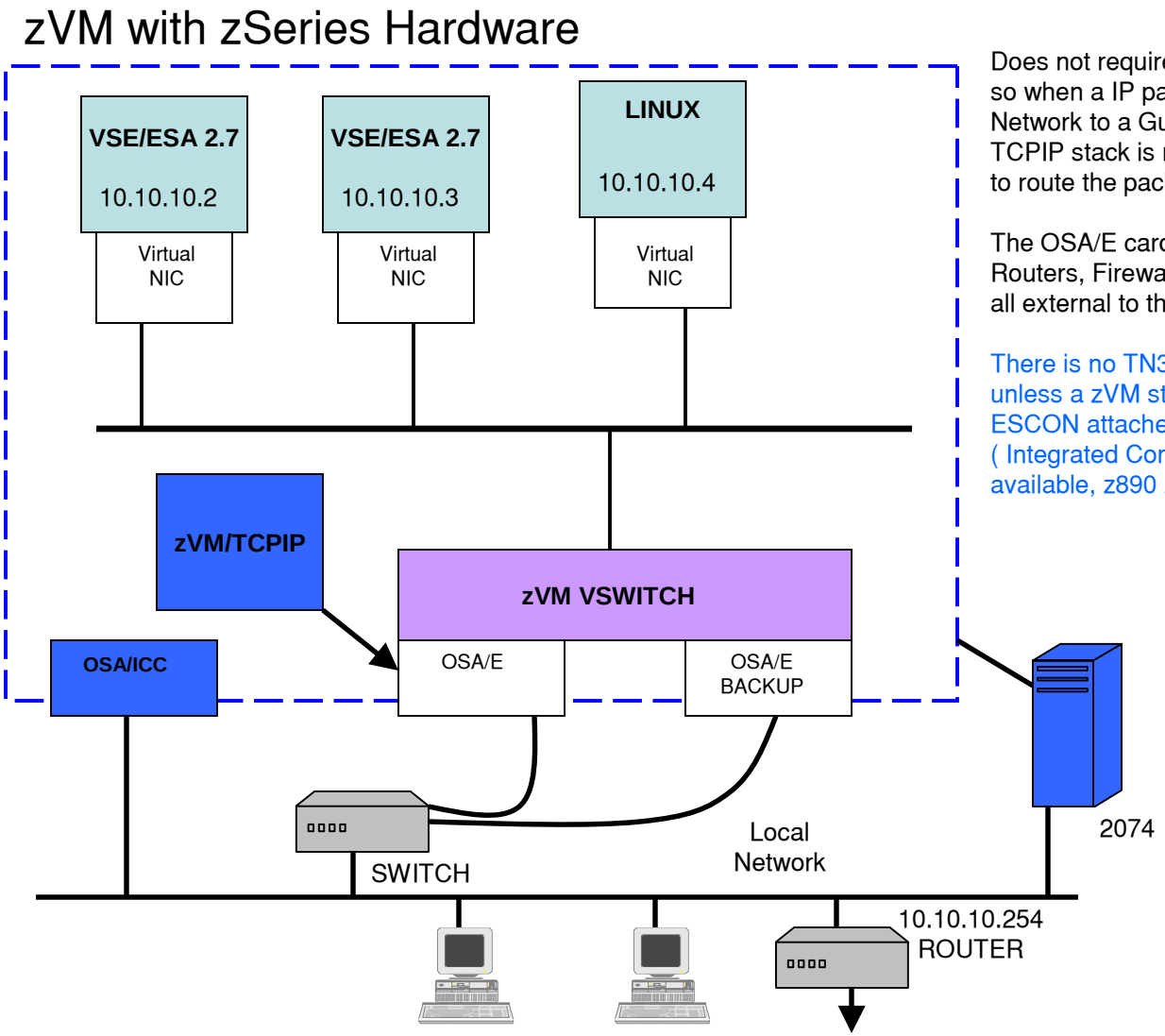
Either stack can provide 3270 access to zVM

3270 access is the primary mission of the zVM stack with the HOME address of 10.10.10.200.

Then two zVM stacks allow a system programmer to take down the ROUTER stack and still have 3270 access from a remote site.

In this case zVM 10.10.10.1needs the OSA card as PRIMARY ROUTER, 10.10.10.200 does not. So 10.10.10.1 and 10.10.10.200 can each be given a set of sub channel addresses On the same OSA card. No need of multiple cards.

No Stack - VSWITCH



Does not require a TCPIP stack for routing so when a IP packet goes from the Local Network to a Guest Machine. A machine running a TCPIP stack is not dispatched to route the packet. Save some cycles.

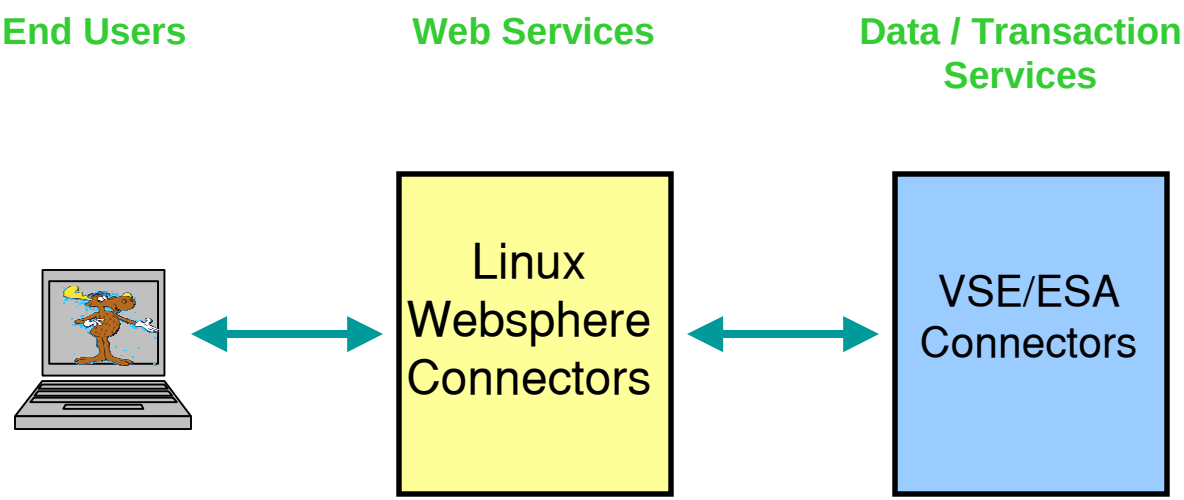
The OSA/E cards do not have an IP address, the Routers, Firewalls are usually all external to the zSeries and zVM.

There is no TN3270 access to zVM with VSWITCH unless a zVM stack is added, or there is a 2074 ESCON attached to the processor. Or a OSA/ICC (Integrated Console Controller / OSA3270) is available, z890 z990 only.

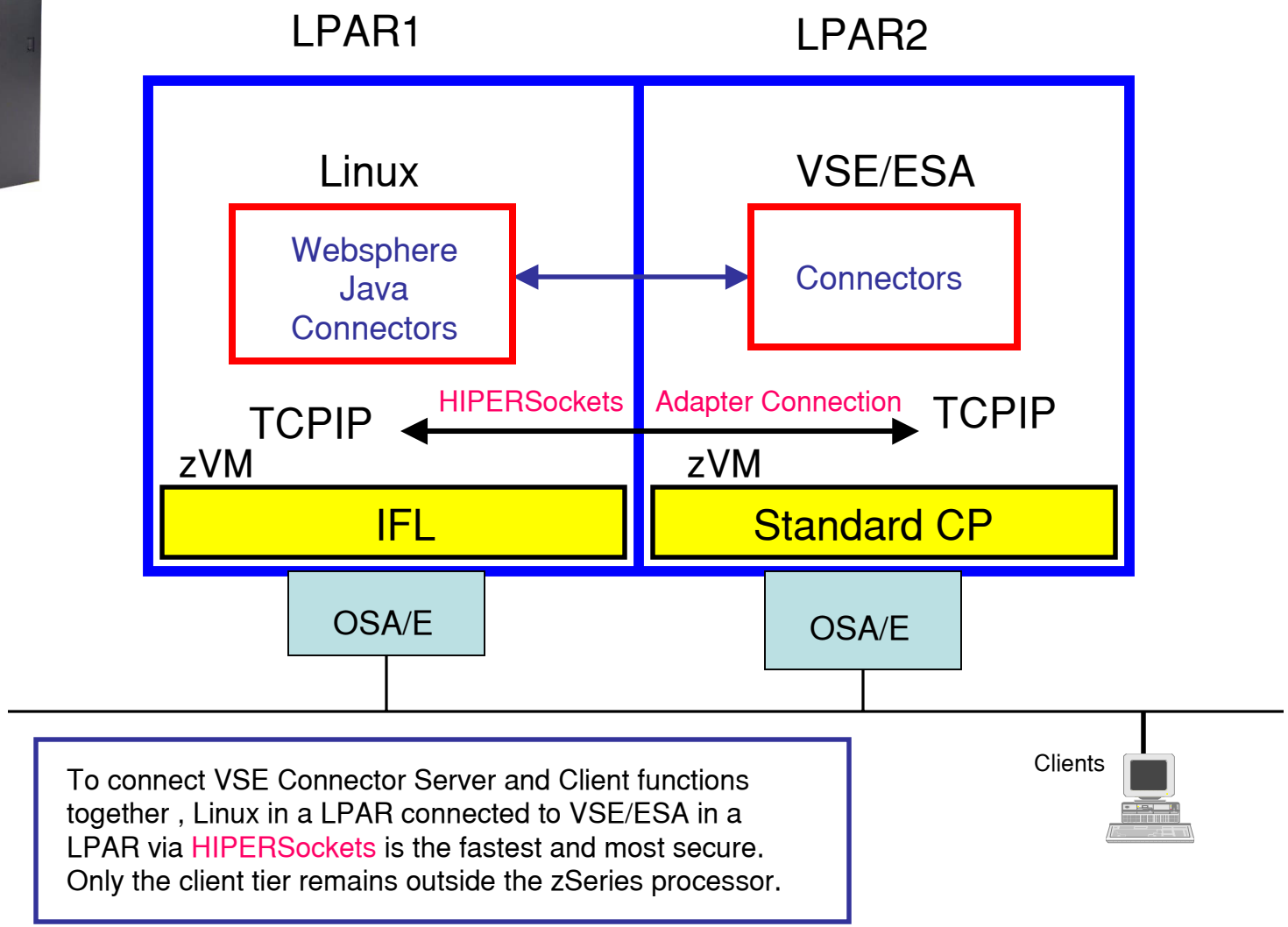
Connecting
VSE To Linux

VSE/ESA

- The three tier e-Busniess structure to bring your VSE data to the web.

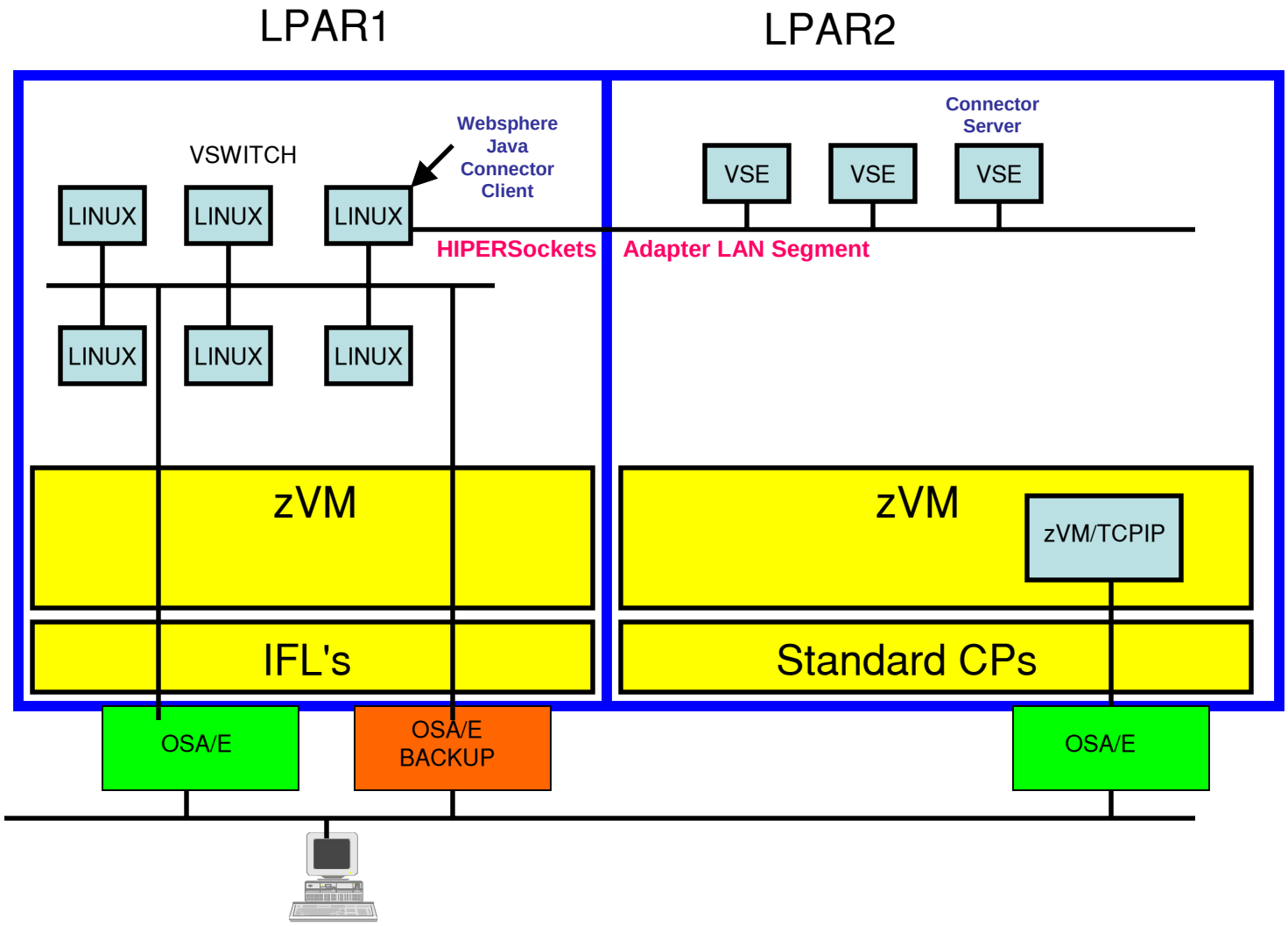


VSE, zSeries, VSE Connectors, zLinux
The Environment In a zSeries



To connect VSE Connector Server and Client functions together , Linux in a LPAR connected to VSE/ESA in a LPAR via HIPERSockets is the fastest and most secure. Only the client tier remains outside the zSeries processor.

VSE, zSeries, VSE Connectors, zLinux
One Example Of Multiple Guests



SNA ?

- The HIPERSockets adapter of the zSeries processors is an IP device only, no SNA support.

- An OSA/E card running in QDIO mode is an IP device only, no SNA support.

- Depending upon which zSeries processor is in use, a SNA capable OSA/E can be ordered. It will be an Ethernet 10/100T, 10/100/1000TX or Token Ring card that can be set to Express mode (**CHPID TYPE OSE**). It functions as LAN Channel Station (**LCS mode , 3172**). In this mode the card is SNA capable and requires you use OSA/SF to tailor the card. It can then operate as a XCA major node.

- All other OSA/E cards can only be QDIO mode (**CHPID TYPE OSD**) which is TCPIP only.

Summary

- The HIPERSockets adapter of the zSeries is the fastest inter-LPAR connection.
- zVM provides two Guest LANs QDIO and HIPER for intra-LPAR connections
- The VSWITCH provides 802.1Q support and removes the need to have an IP stack as a router.
- Latest Hardware/Software provide a multitude of IP connectivity options

There will be more than one way to set up your network, spend time planning.

- VSE/ESA and LINUX on a zSeries connected via HIPERSockets adapter using VSE e-Busniess Connectors, is the way to bring your VSE data to the WWW.

Glossary

Flat Network – All Hosts are in the same IP network, do not find this as an official definition but its use is ubiquitous

IFL – Integrated Facility for Linux. A Central Processor that supports only zVM and/or Linux.

NIC – Network Interface Card, like the Ethernet adapter in a personal computer.

Penguin – Bird that looks like a Headwaiter

P2P – Point-to-point connection

Standard CP – Standard Central Processor that supports all operating systems

The End