Connecting Linux to TCPIP for zVSE
With zVM & zSeries

-2006-

Stephen Gracin
gracinsp@us.ibm.com

Ground Rules

For this discussion unless otherwise noted

- VSE or VSE/ESA is zVSE

- zVM is zVM 5.2

- zSeries Processor refers to z890, z990, Z9

- TCPIP for VSE is the CSI TCPIP

Moving Forward

When moving from an old VSE to the most current VSE/ESA and a ZSeries processor. **Are you:**

Going to upgrade to latest zVM or adding zVM ?

Going to bring Linux function to your zVSE environment?

Considering IP networking with/without zVM ?

Considering IP networking options with zSeries hardware ?

- With the latest software and zSeries hardware there are many connectivity options available. Which ones do you need?

zSeries Hardware
Connectivity Considerations

zSeries Hardware Connectivity

- OSA Express/2 cards – <u>Some</u> can operate in two modes, one of the modes will support SNA (VBUILD TYPE=XCA).

- HIPERSockets adapter – High speed zero latency 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/2 Cards

- May not require OSA/SF for setup, but OSA/SF can still 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. For LCS Mode OSA/SF is required to configure the card.

- 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* . Until recent releases 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 – HIPERsockets Adapter

LPAR	LPAR	LPAR	LPAR	LPAR	LPAR
LINUX	zVM	zVSE	zVSE	zVSE	LINUX

HIPERsockets   Adapter – An Ethernet LAN between LPARs
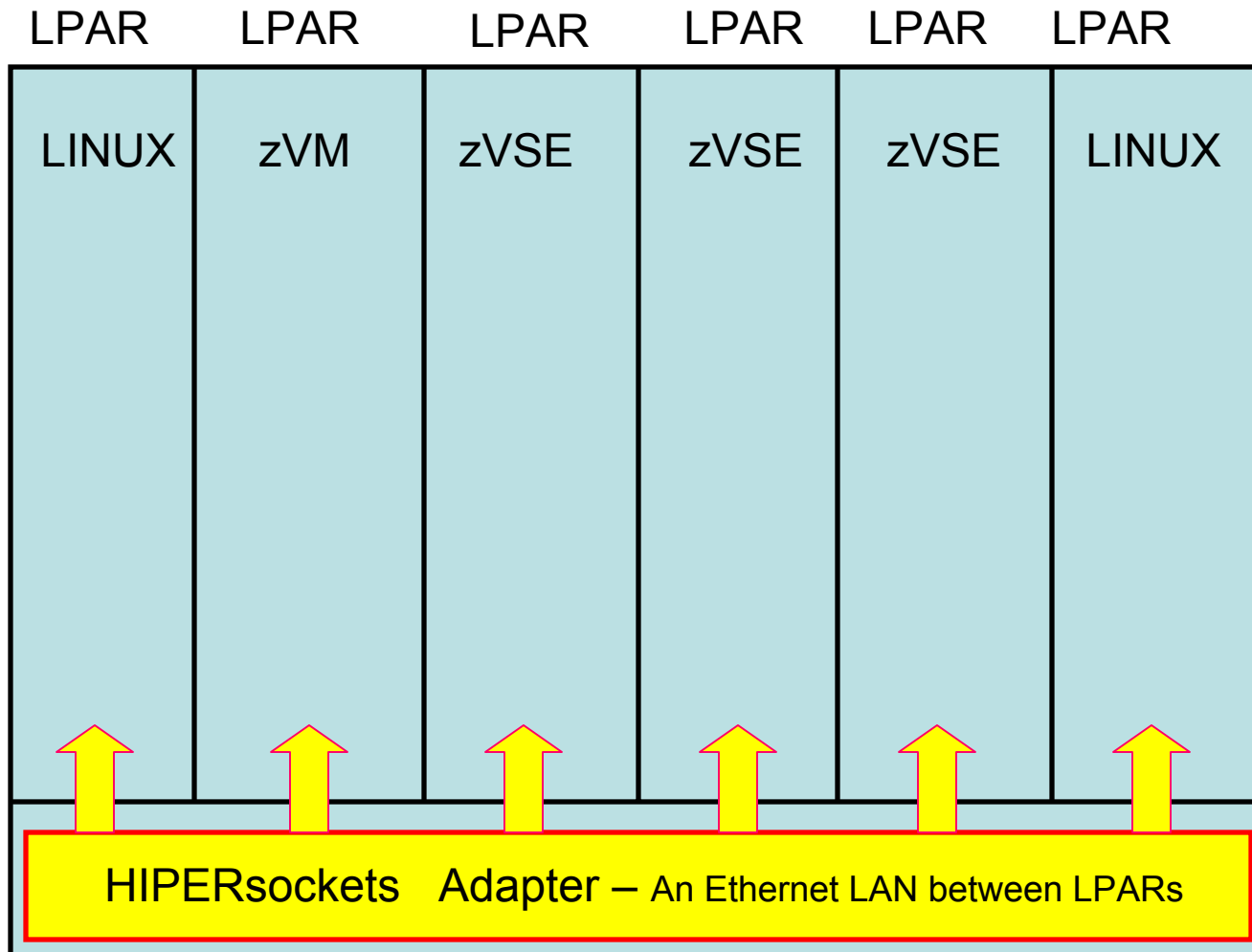
zSeries Hardware - IFL

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 will 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

LPAR	LPAR
VSE, zOS, zVM or Linux as a zVM Guest zOS VSE / zVSE LINUX zVM	zVM or Linux as a zVM Guest LINUX zVM
Processor Type **Standard CP**	Processor Type **IFL**

zVM Connectivity Options

zVM5.2

Provides Two Guest LAN types:

• QDIO – Each guest machine has a virtual NIC defined that emulates a
　　　　Gigabit OSA/E. They become hosts on a virtual LAN.

• HIPERSockets – Each guest machine has a virtual NIC that emulates the
　　　　　　　　HIPERsockets adapter. They become hosts on a virtual LAN
　　　　　　　　usually 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
　　　　this is old technology if on zSeries hardware and latest software use
　　　　QDIO/HIPER.

zVM5.1

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 the SET VSWITCH does not set VLAN tagging (VLAN ANY) VSWITCH provides LAN "bridge like" function

So what's a switch ?

- Like a Hub, 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 can be a problem.
  A Switch receives an Ethernet Frame on one of its ports and has the electronics to hold it. This allows Full Duplex communications and the ability to act upon the Frame fields like VLAN ID.

   Put the 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
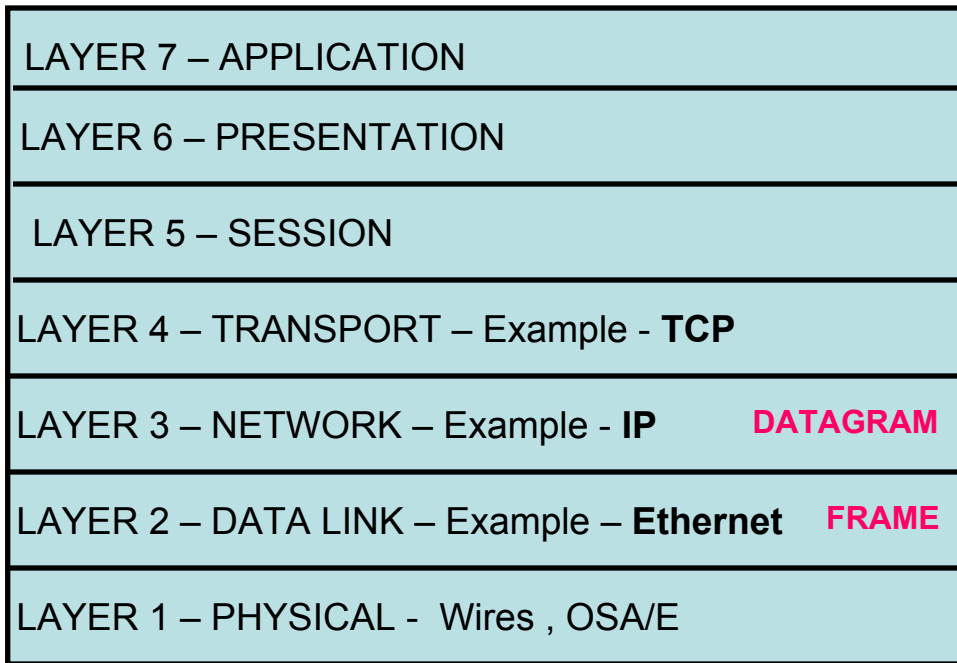
You are Here !

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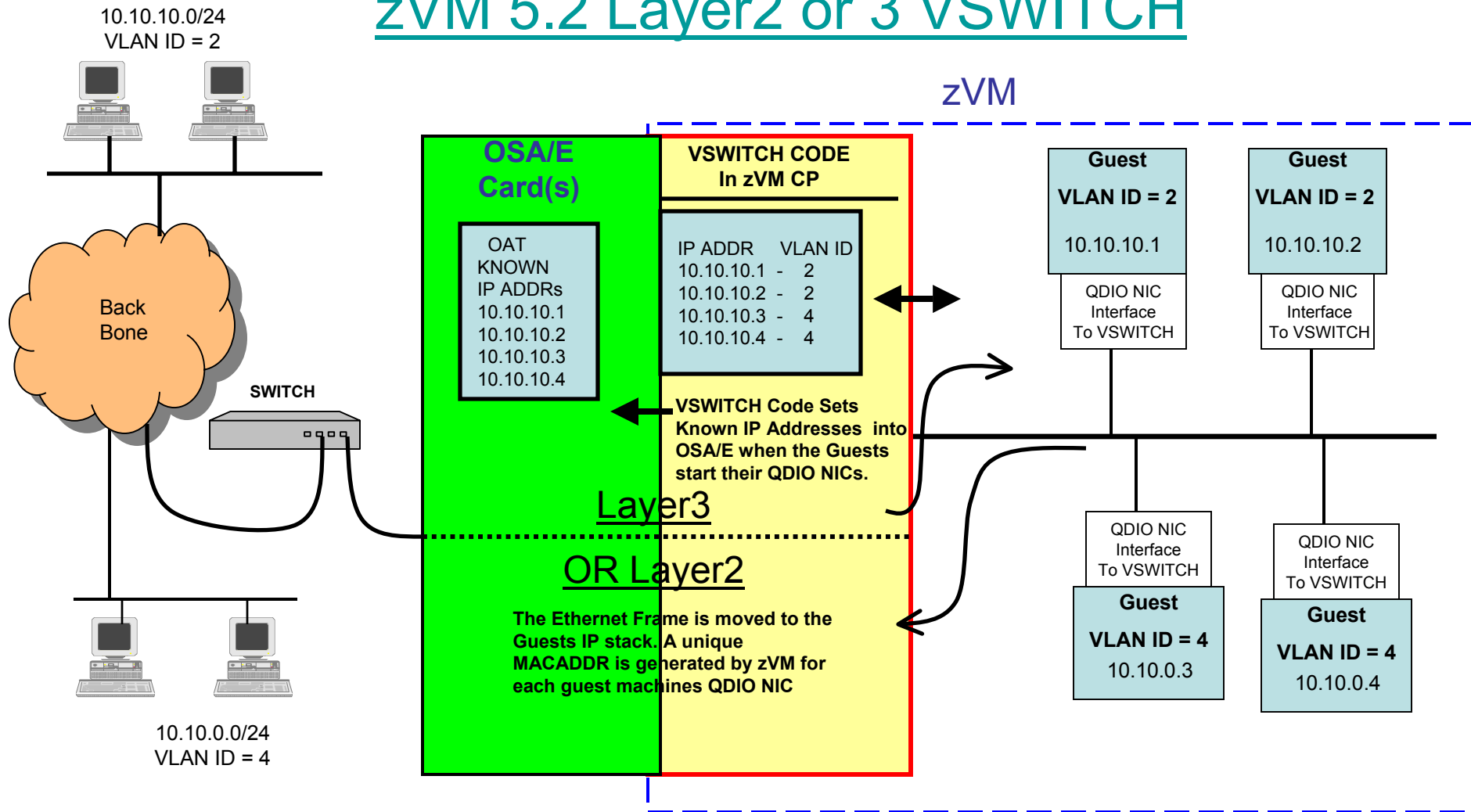
IP DATAGRAM LAYER 3

		SRC ADDR	DST ADDR		D A T A

ETHERNET FRAME LAYER 2

	DST MAC	SRC MAC	VLAN	LEN	D A T A	

VSWITCH / VLAN

- zVM5.2 Can Provide a "Bridge Like" function for Layer 3 or in some circumstances can be a true Layer 2 switch.

- Is a specialized ODIO Guest LAN

- Requires an OSA/E card in QDIO mode

- With two OSA/E cards provides hardware backup.

- Can use/manage VLAN IDs, Does Not Have too.

- zVSE environment is not VLAN aware. VSE can not manipulate an Ethernet frame to set or interrogate a VLAN ID . zVM can control a VLAN ID for the VSE guest.

- Without VLAN IDs you get an easy to use "Flat Network" layer 3 switch with OSA hardware redundancy.

zVM 5.2 Layer2 or 3 VSWITCH

zVM

10.10.10.0/24
VLAN ID = 2

Back Bone

SWITCH

10.10.0.0/24
VLAN ID = 4

OSA/E Card(s)

OAT
KNOWN
IP ADDRs
10.10.10.1
10.10.10.2
10.10.10.3
10.10.10.4

VSWITCH CODE In zVM CP

IP ADDR	VLAN ID
10.10.10.1 -	2
10.10.10.2 -	2
10.10.10.3 -	4
10.10.10.4 -	4

VSWITCH Code Sets Known IP Addresses into OSA/E when the Guests start their QDIO NICs.

Layer3

OR Layer2

The Ethernet Frame is moved to the Guests IP stack. A unique MACADDR is generated by zVM for each guest machines QDIO NIC

Guest
VLAN ID = 2
10.10.10.1
QDIO NIC Interface To VSWITCH

Guest
VLAN ID = 2
10.10.10.2
QDIO NIC Interface To VSWITCH

QDIO NIC Interface To VSWITCH
Guest
VLAN ID = 4
10.10.0.3

QDIO NIC Interface To VSWITCH
Guest
VLAN ID = 4
10.10.0.4

With a z890/990/z9 and zVM5.2the DEFINE VSWITCH command defines whether LAYER2 or LAYER3 is used. True Layer 2 requires an OSA/Express2 card.

OSA Info

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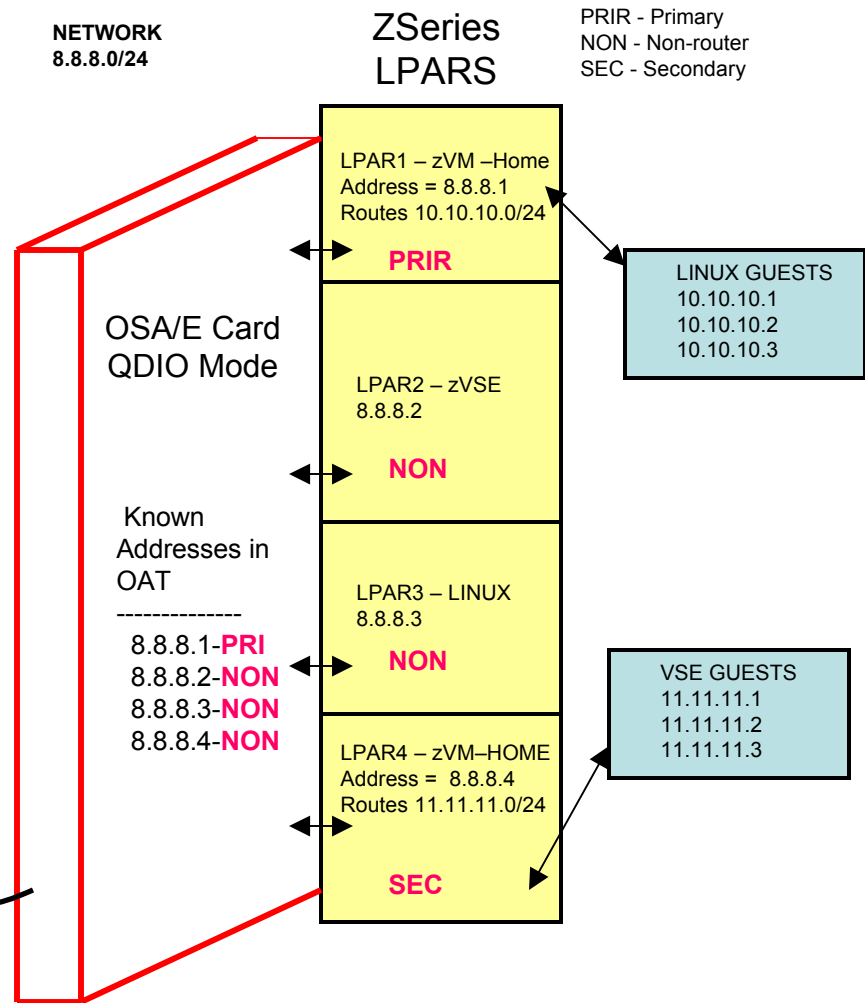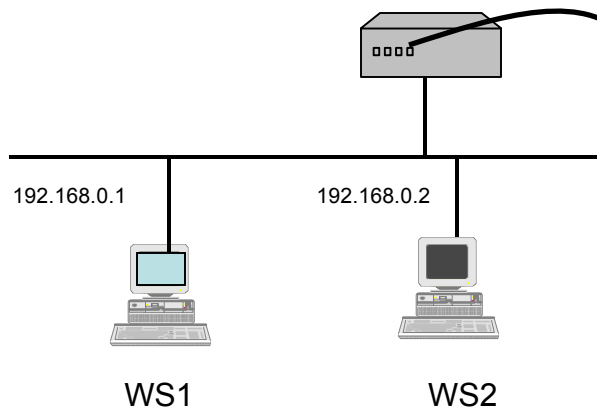
Non-router – This stack will only be presented with *Known Addresses*. This would be an IP address loaded into the OSA card when the stack initializes, such as its HOME address. Only packets whose destination address is the HOME address would be presented to the stack by the OSA card. You do not want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

NETWORK
8.8.8.0/24

ZSeries
LPARS

PRIR - Primary
NON - Non-router
SEC - Secondary

The Workstations can get to Linux Guests
10.10.10.1-3 but not to VSE Guests 11.11.11.1-3

Unless the stack is known to the OSA Card as the PRIR
Stack, only Known Addresses are presented to the stack.
Any others are not routed out of the OSA card.

LPAR1 – zVM –Home
Address = 8.8.8.1
Routes 10.10.10.0/24

PRIR

LINUX GUESTS
10.10.10.1
10.10.10.2
10.10.10.3

OSA/E Card
QDIO Mode

LPAR2 – zVSE
8.8.8.2

NON

Known
Addresses in
OAT

8.8.8.1-**PRI**
8.8.8.2-**NON**
8.8.8.3-**NON**
8.8.8.4-**NON**

LPAR3 – LINUX
8.8.8.3

NON

LPAR4 – zVM–HOME
Address = 8.8.8.4
Routes 11.11.11.0/24

SEC

VSE GUESTS
11.11.11.1
11.11.11.2
11.11.11.3

192.168.0.1

192.168.0.2

WS1

WS2

VSE Connectivity
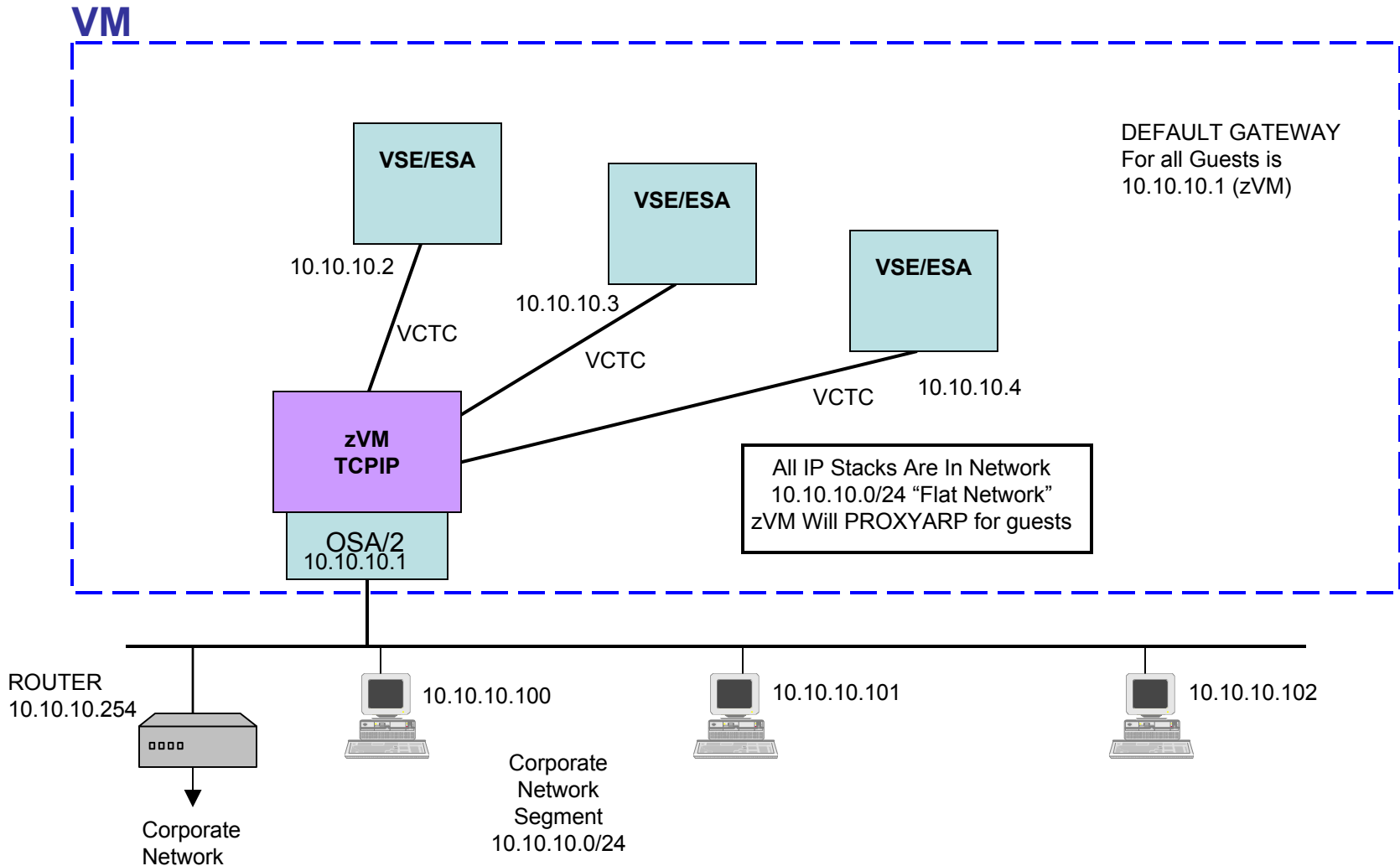# In The Past…

Connectivity History

A hub-and-spoke topology using CTC connections is most often seen with old VSE/ESA guest systems for IP and SNA. With IP it is the one topology with two IP addressing schemes.

1. The guests are in the same network as the OSA 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.

VSE

VM

VSE

VSE
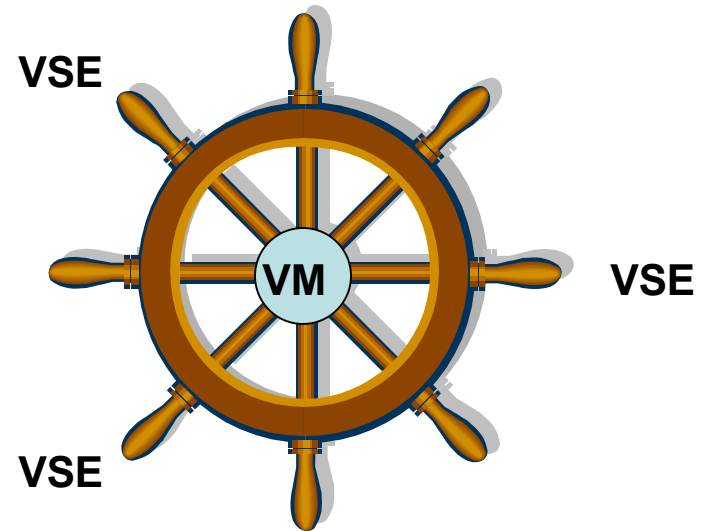
Flat Network

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

VM

VSE/ESA

10.10.10.2

VCTC

VSE/ESA

10.10.10.3

VCTC

VSE/ESA

VCTC

10.10.10.4

DEFAULT GATEWAY
For all Guests is
10.10.10.1 (zVM)

**zVM
TCPIP**

OSA/2
10.10.10.1

All IP Stacks Are In Network
10.10.10.0/24 "Flat Network"
zVM Will PROXYARP for guests

ROUTER
10.10.10.254

10.10.10.100

10.10.10.101

10.10.10.102

Corporate
Network

Corporate
Network
Segment
10.10.10.0/24

Different Network

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

VM

DEFAULT GATEWAY
For all Guests is
10.10.10.1 (zVM)

VSE/ESA

VSE/ESA

VSE/ESA

10.10.10.2

10.10.10.3

10.10.10.4

VCTC

VCTC

VCTC

10.10.10.1
**zVM
TCPIP**

OSA/2
9.9.9.1

All VSE IP Stacks Are In Network
10.10.10.0/24 , The OSA is in network
9.9.9.0/24 , zVM/TCPIP is a Router

ROUTER
9.9.9.254

9.9.9.100

9.9.9.101

9.9.9.102

Corporate
Network
Segment
9.9.9.0/24

To Everybody
else
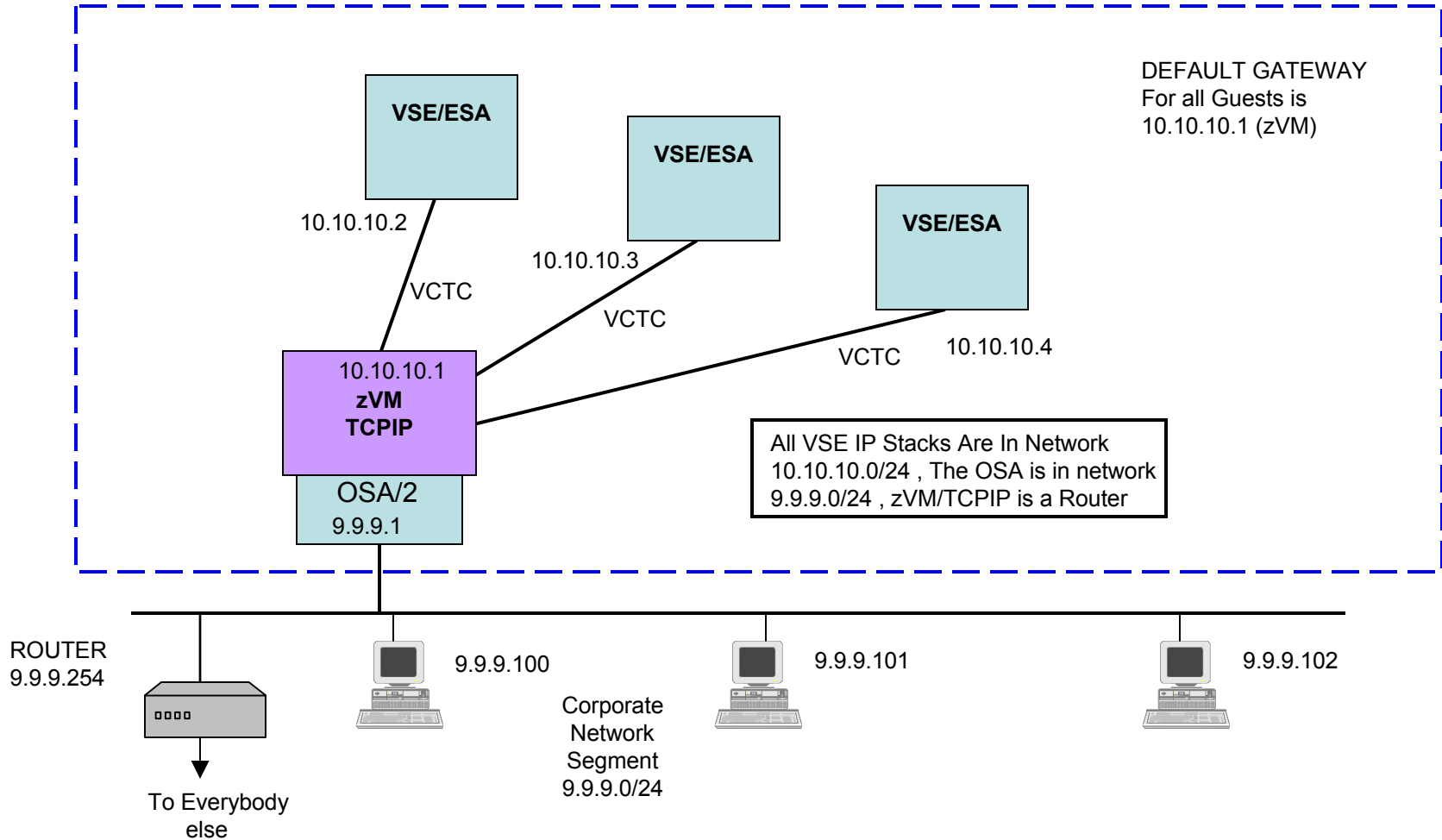
VSE/ESA

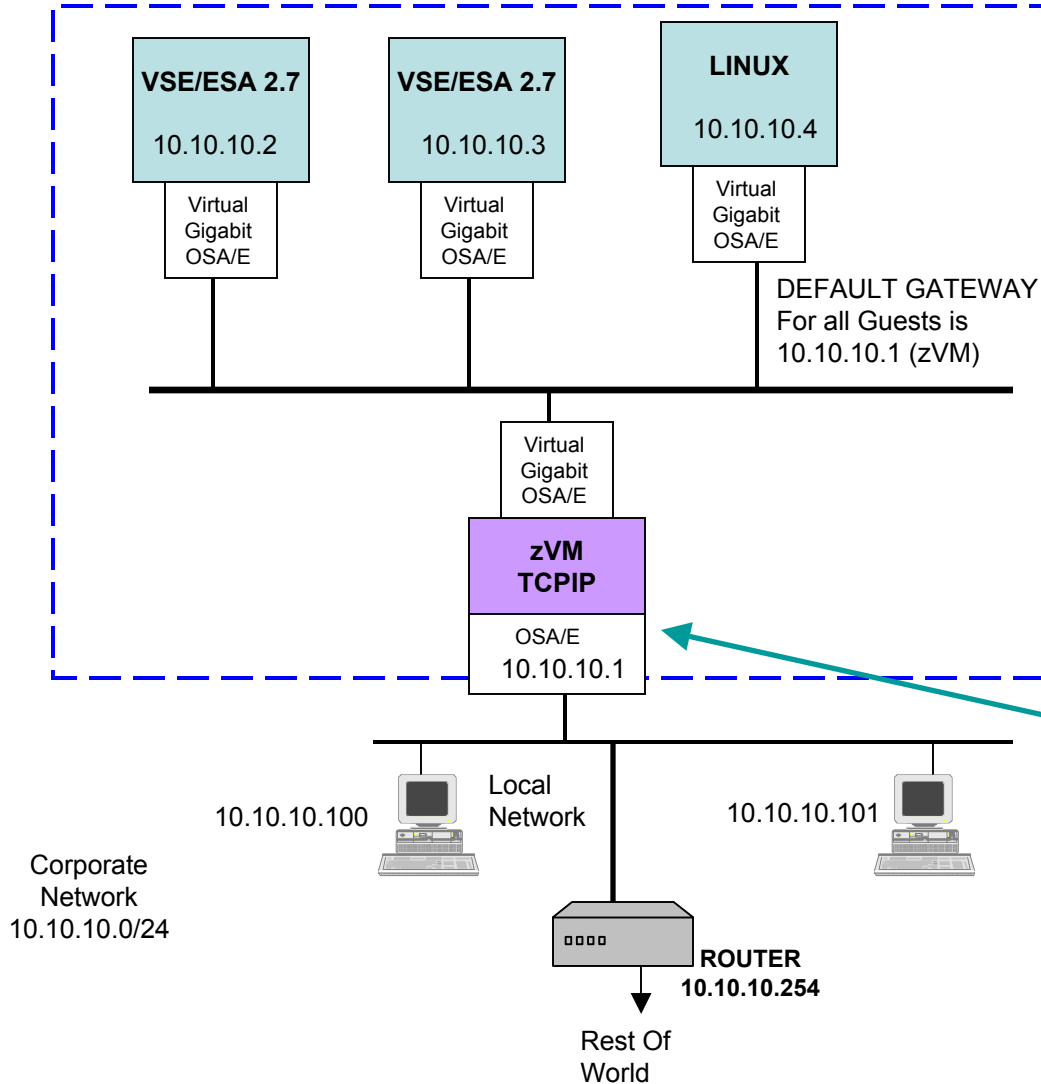Moving forward assumes that we will upgrade to zVSE and zVM 5.2

- zVSE brings support for QDIO and HIPERSocket adapters

- zVM5.2 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.

- zVM5.2 also bring support for VSWITCH which allows the guest machines
  To bridge to the corporate network.

Migration Observations

- With the latest zVSE , 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

zVM with zSeries Hardware

VSE/ESA 2.7

10.10.10.2

Virtual Gigabit OSA/E

VSE/ESA 2.7

10.10.10.3

Virtual Gigabit OSA/E

LINUX

10.10.10.4

Virtual Gigabit OSA/E

DEFAULT GATEWAY
For all Guests is
10.10.10.1 (zVM)

Virtual Gigabit OSA/E

zVM TCPIP

OSA/E
10.10.10.1

Local Network

10.10.10.100

10.10.10.101

Corporate Network
10.10.10.0/24

ROUTER
10.10.10.254

Rest Of World

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.

zVM allows PROXYARP for a guest LAN (Broadcast Network) . You must have HOST routes in zVM/TCPIP For each Guests for PROXYARP to work. A NETWORK route will not do. 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
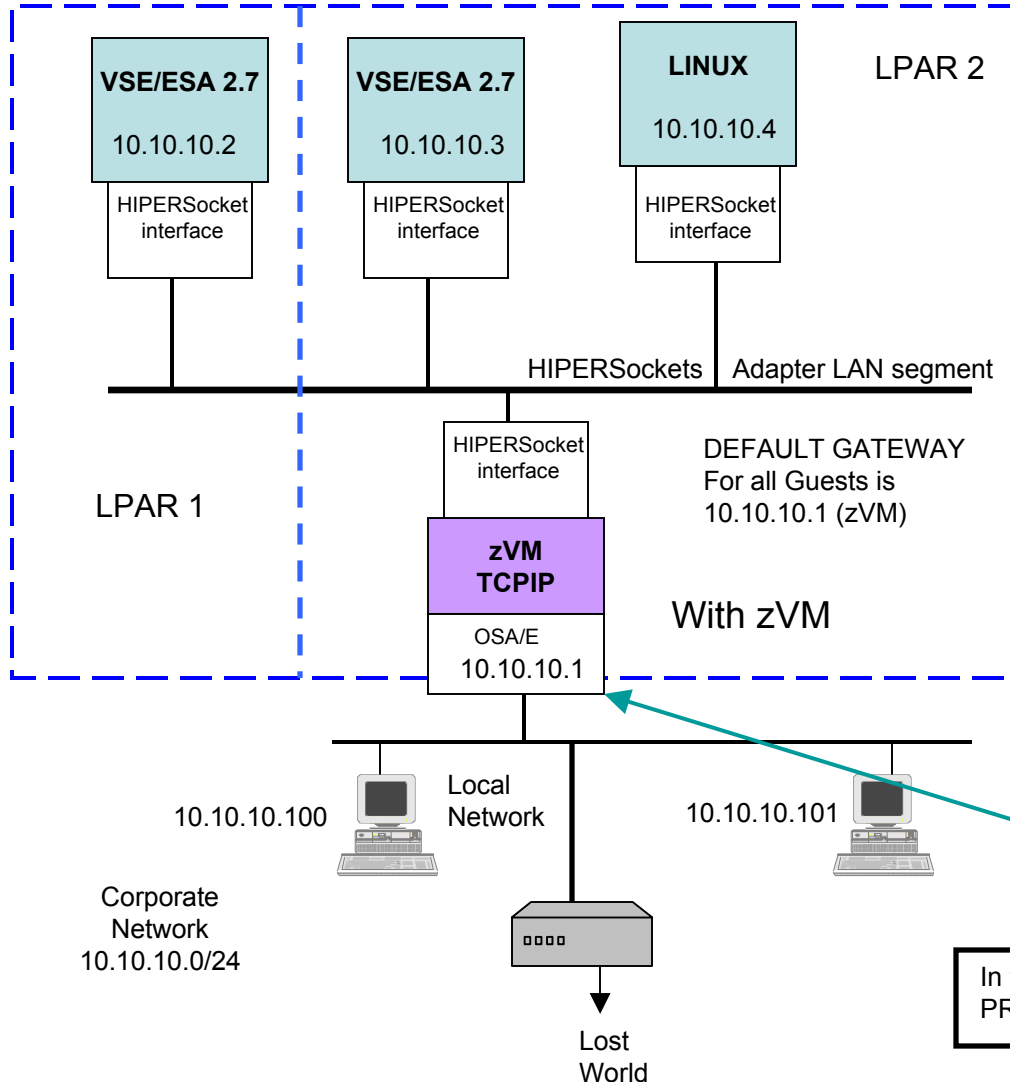
The guest LAN (VM/CP) carries interguest traffic the zVM TCPIP stack is not required for this function.

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

zVM with zSeries Hardware

VSE/ESA 2.7	VSE/ESA 2.7	LINUX	LPAR 2
10.10.10.2	10.10.10.3	10.10.10.4	

HIPERSocket interface | HIPERSocket interface | HIPERSocket interface

HIPERSockets | Adapter LAN segment

HIPERSocket interface

DEFAULT GATEWAY
For all Guests is
10.10.10.1 (zVM)

zVM TCPIP

OSA/E
10.10.10.1

With zVM

LPAR 1

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.

zVM allows PROXYARP for a guest LAN (Broadcast Network) . You must have a HOST route in zVM/TCPIP For each Guests for PROXYARP to work. A NETWORK route will not do. 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.
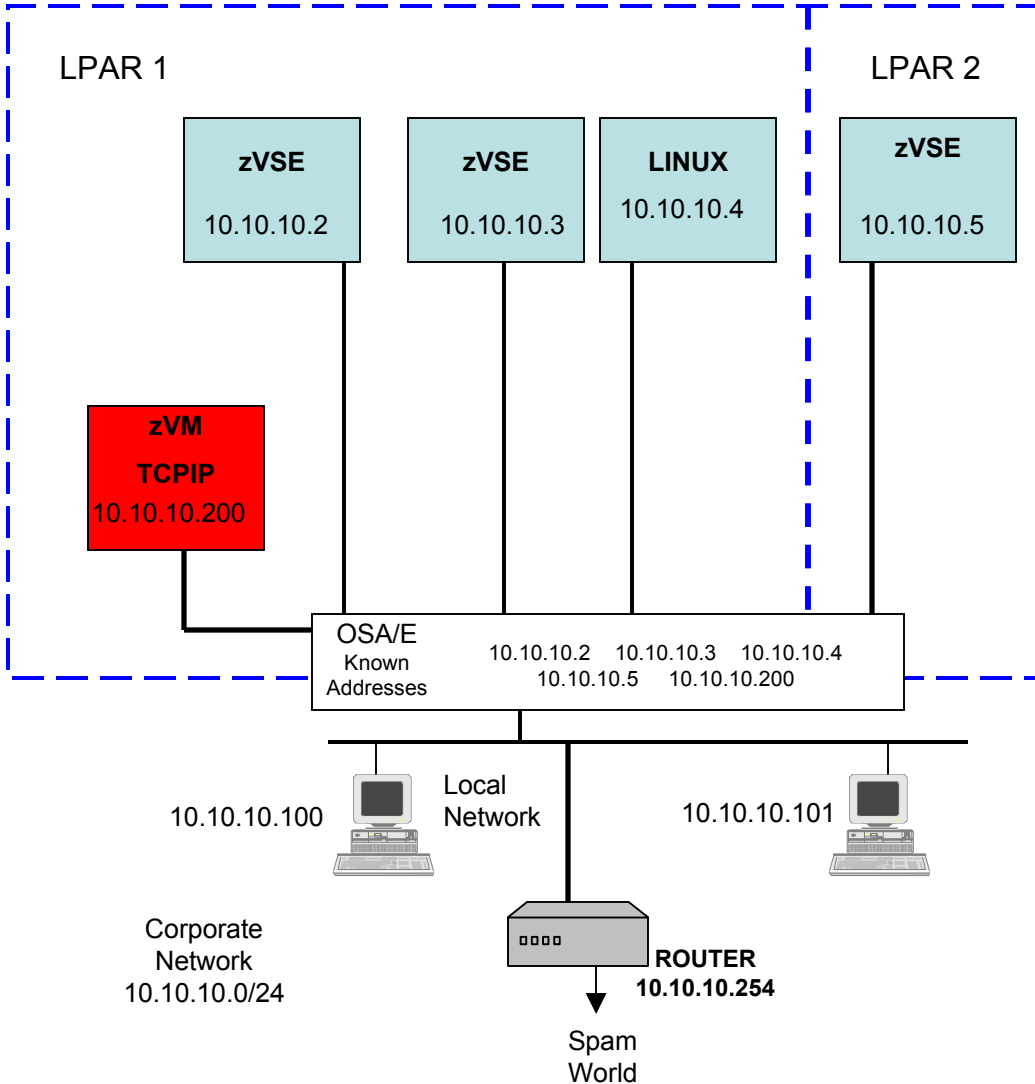
The HIPERSockets adapter is very fast.

Local
Network

10.10.10.100

10.10.10.101

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

Corporate
Network
10.10.10.0/24

Lost
World

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

Migrate Flat Network - To Direct OSA/E Attach

zVM with zSeries Hardware

LPAR 1

LPAR 2

zVSE

10.10.10.2

zVSE

10.10.10.3

LINUX

10.10.10.4

zVSE

10.10.10.5

zVM

TCPIP

10.10.10.200

OSA/E
Known
Addresses

10.10.10.2   10.10.10.3   10.10.10.4
10.10.10.5   10.10.10.200

10.10.10.100

10.10.10.101

Local
Network

Corporate
Network
10.10.10.0/24

ROUTER
10.10.10.254

Spam
World

**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
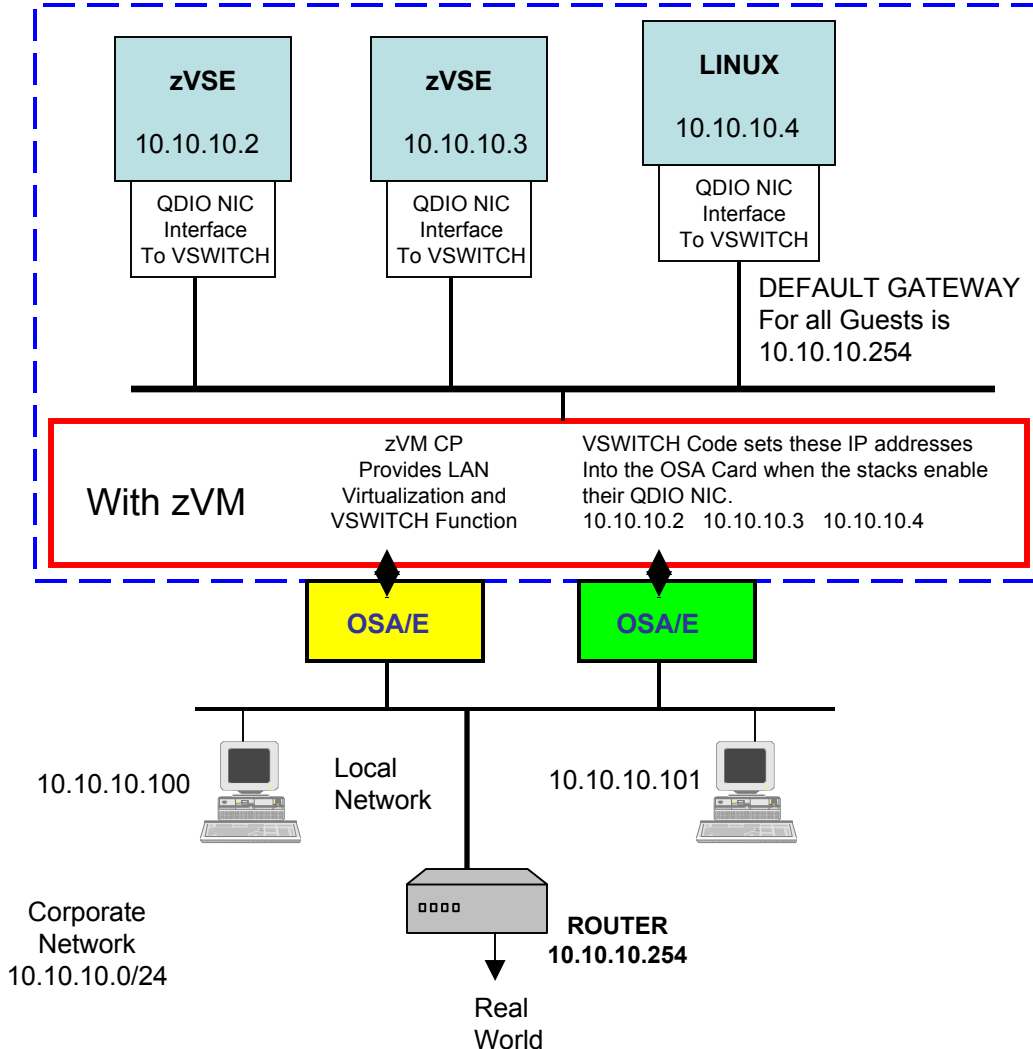to be routed.**

**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

zVM with zSeries Hardware

zVSE	zVSE	LINUX
10.10.10.2	10.10.10.3	10.10.10.4

QDIO NIC Interface To VSWITCH

QDIO NIC Interface To VSWITCH

QDIO NIC Interface To VSWITCH

DEFAULT GATEWAY
For all Guests is
10.10.10.254

With zVM

zVM CP
Provides LAN
Virtualization and
VSWITCH Function

VSWITCH Code sets these IP addresses
Into the OSA Card when the stacks enable
their QDIO NIC.
10.10.10.2   10.10.10.3   10.10.10.4

OSA/E

OSA/E

10.10.10.100

Local
Network

10.10.10.101

Corporate
Network
10.10.10.0/24

**ROUTER
10.10.10.254**

Real
World

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 zVM5.1 VSWITCH can be a layer 2 or Layer 3 switch.

With just VSWITCH enabled zVM will not have a HOME IP Address in a PROFILE TCPIP. You can have A functioning VSWITCH providing connectivity to Your Guests but have no connectivity to zVM its self.

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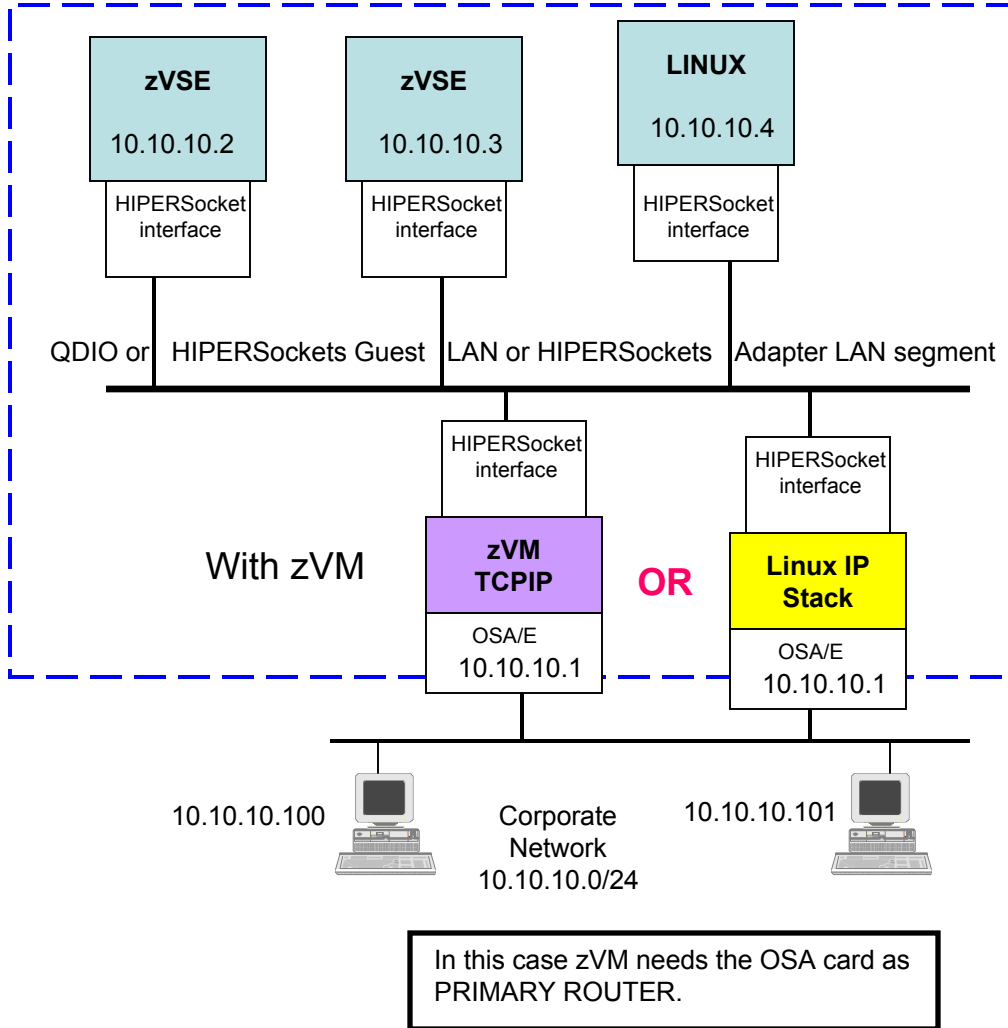
You can define some sub channel addresses of one of the OSA/E cards for use by zVM/TCPIP for TN3270 Access. Otherwise a 3174 / 2074 device is required for non-SNA 3270 device access.

The nature of VSWITCH makes PRI-ROUTER vs. NON-ROUTER not an issue. As the Guest Machines initialize their virtual NICs in a Layer 3 environment the information is also added to the OSA/E by the VSWITCH function.

Interface To The World !
Which Stack?

Which Stack?

zVM with zSeries Hardware

zVSE	zVSE	LINUX
10.10.10.2	10.10.10.3	10.10.10.4
HIPERSocket interface	HIPERSocket interface	HIPERSocket interface

QDIO or   HIPERSockets Guest   LAN or HIPERSockets   Adapter LAN segment

HIPERSocket interface

HIPERSocket interface

With zVM

zVM TCPIP   **OR**   **Linux IP Stack**

OSA/E
10.10.10.1

OSA/E
10.10.10.1

10.10.10.100   Corporate Network   10.10.10.101
10.10.10.0/24

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

zVM or a Linux stack can be a Router for a QDIO or HIPERSockets Guest LAN or a HIPERSockets adapter LAN.
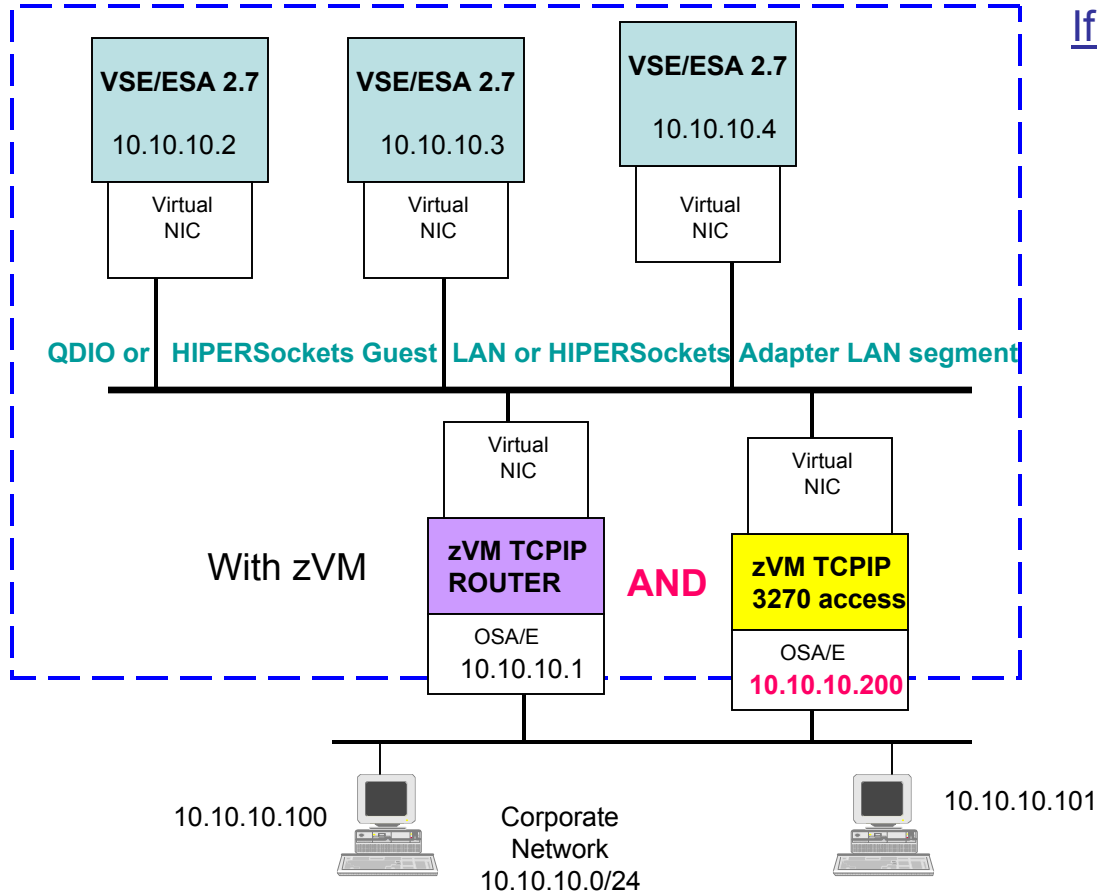
Some differences are:

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.

The trend now is use VSWITCH, have a zVM stack for TN3270 connectivity and have ROUTERS/Firewalls external to the zVM virtual network.

Two (Or more) zVM Stacks!

zVM with zSeries Hardware

VSE/ESA 2.7

10.10.10.2

Virtual NIC

VSE/ESA 2.7

10.10.10.3

Virtual NIC

VSE/ESA 2.7

10.10.10.4

Virtual NIC

QDIO or HIPERSockets Guest LAN or HIPERSockets Adapter LAN segment

Virtual NIC

Virtual NIC

With zVM

zVM TCPIP ROUTER

AND

zVM TCPIP 3270 access

OSA/E
10.10.10.1

OSA/E
10.10.10.200

10.10.10.100

Corporate Network
10.10.10.0/24

10.10.10.101

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.

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.
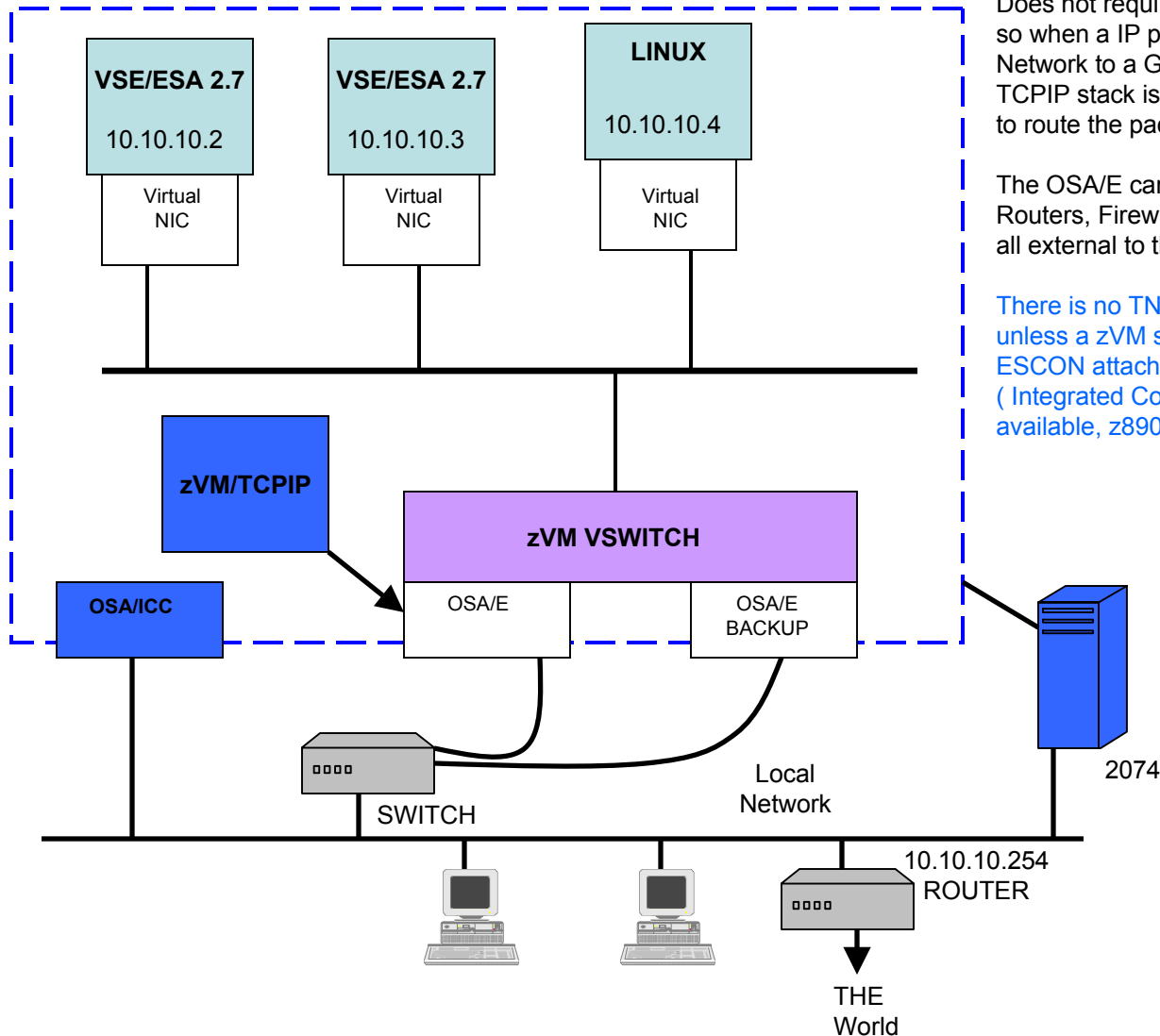
Either stack can provide 3270 access to zVM

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.

No Stack Routes For Guests - VSWITCH

zVM with zSeries Hardware

VSE/ESA 2.7

10.10.10.2

VSE/ESA 2.7

10.10.10.3

LINUX

10.10.10.4

Virtual NIC

Virtual NIC

Virtual NIC

zVM/TCPIP

zVM VSWITCH

OSA/E

OSA/E BACKUP

OSA/ICC

SWITCH

Local Network

2074

10.10.10.254
ROUTER

THE World

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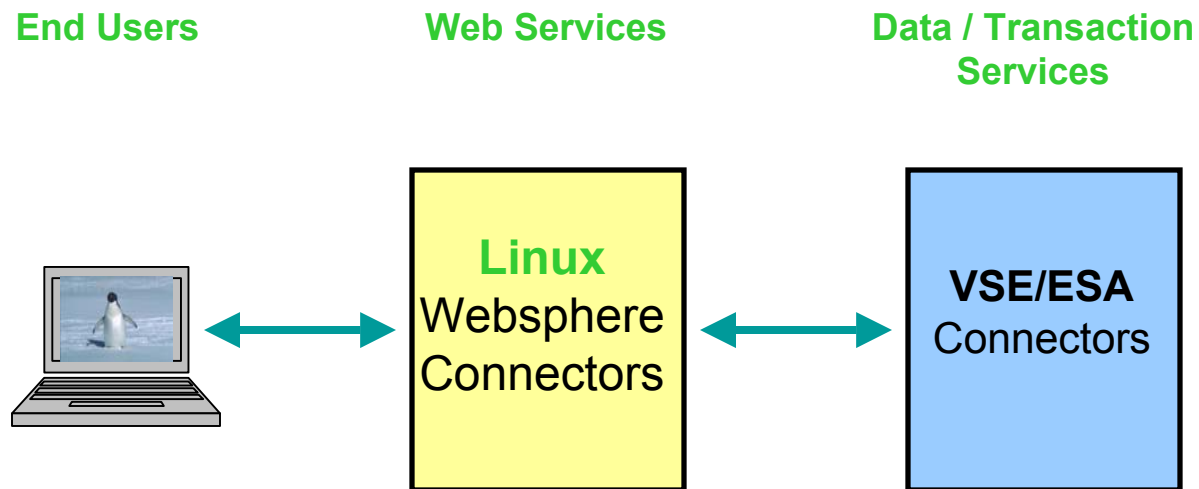
The OSA/E cards do not have an IP address, the Routers, Firewalls are usually all external to the zSeries and zVM.

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.
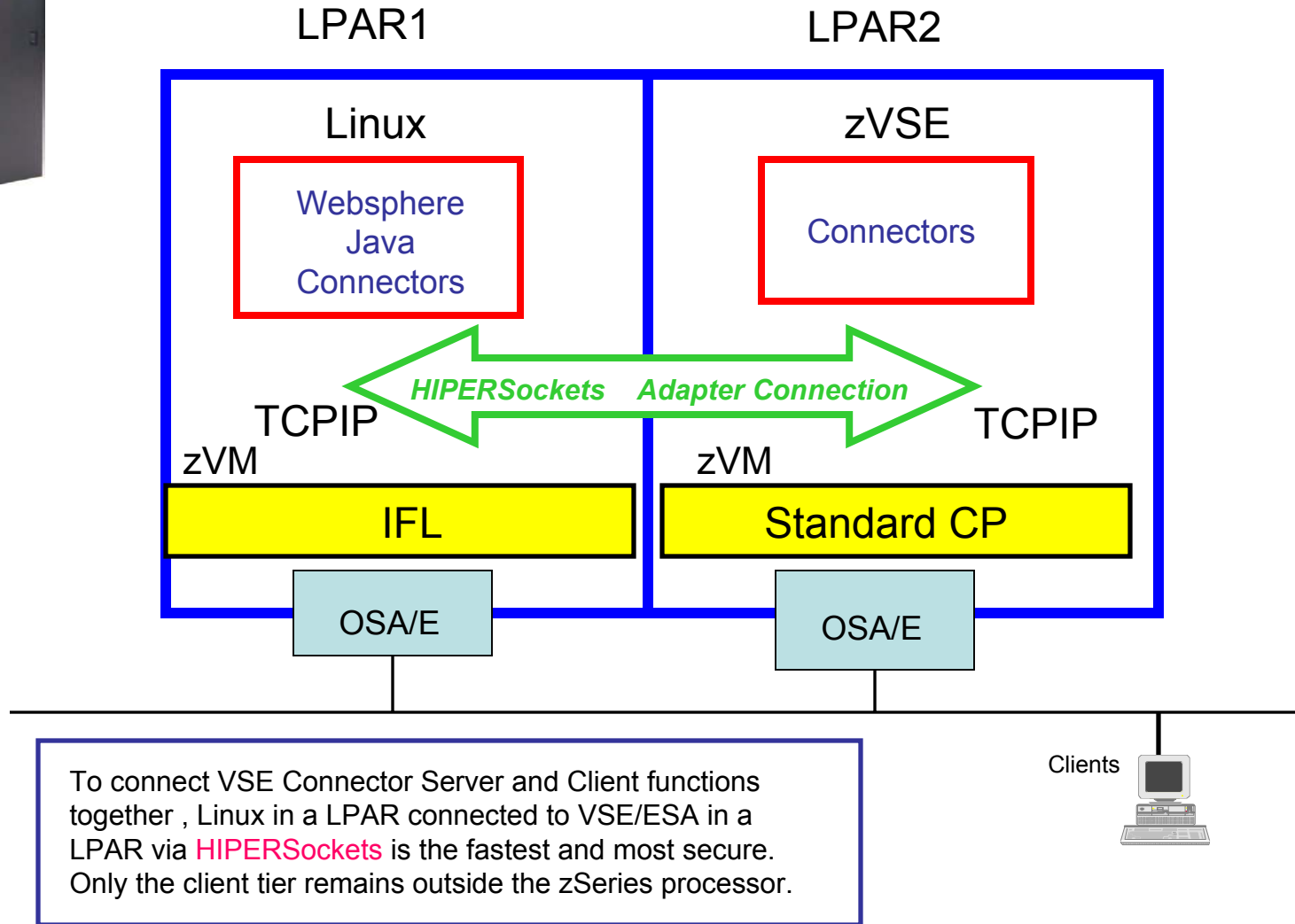
Connecting
zVSE To Linux

VSE/ESA

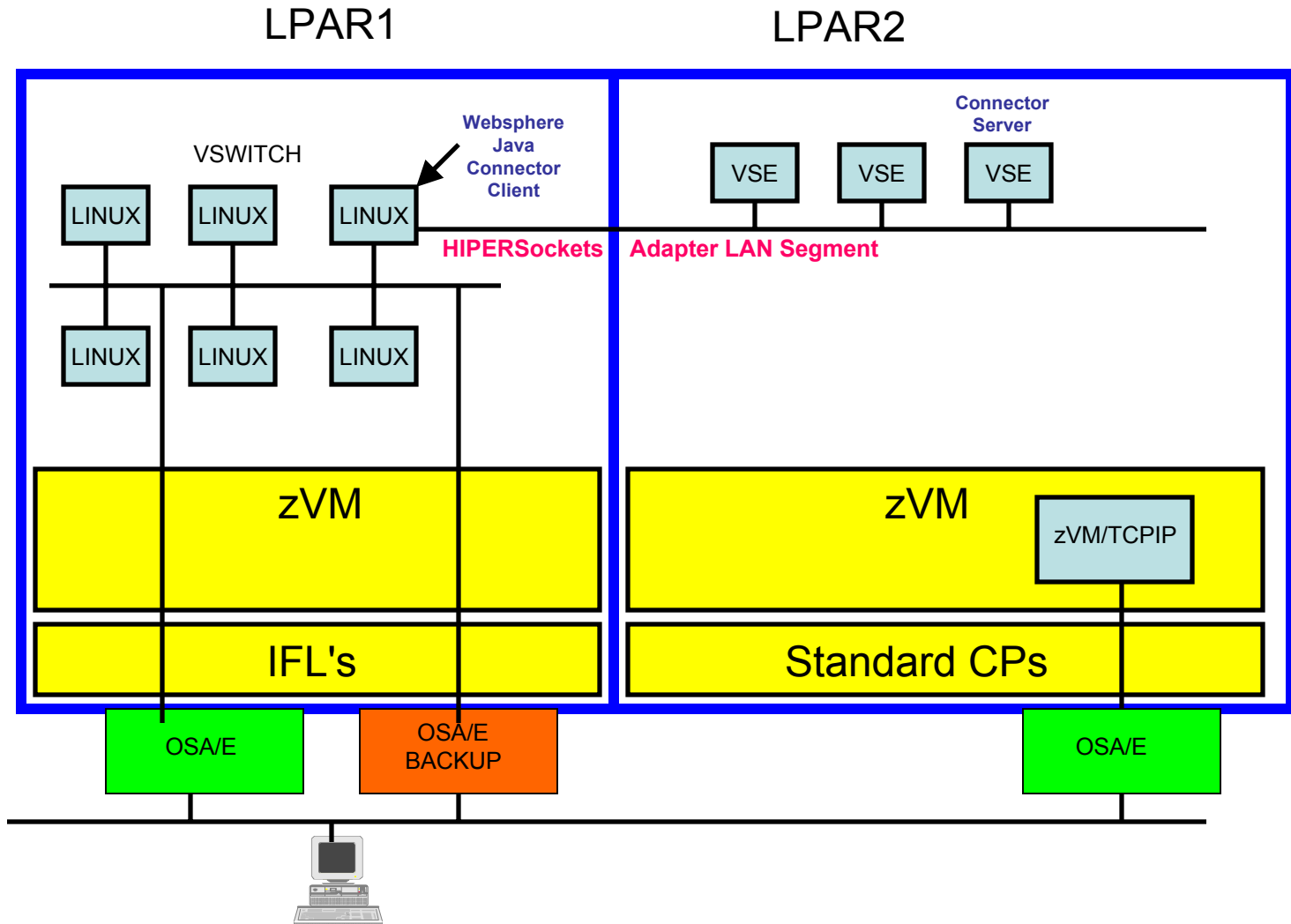- The three tier e-Busniess structure to bring your VSE data to the web.

End Users   **Web Services**   **Data / Transaction Services**

Linux
Websphere
Connectors

VSE/ESA
Connectors

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

LPAR1

LPAR2

Linux

zVSE

Websphere
Java
Connectors

Connectors

HIPERSockets   Adapter Connection

TCPIP

TCPIP

zVM

zVM

IFL

Standard CP

OSA/E

OSA/E

Clients

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

LPAR1

LPAR2

VSWITCH

Websphere Java Connector Client

Connector Server

| LINUX | LINUX | LINUX |

| VSE | VSE | VSE |

HIPERSockets **Adapter LAN Segment**

| LINUX | LINUX | LINUX |

zVM

zVM

zVM/TCPIP

IFL's

Standard CPs

OSA/E

OSA/E BACKUP

OSA/E

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.

The End

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 also 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