



VM TCP/IP Routing - Part 2

Session V23

Alan Altmark
IBM Corporation



Note

This presentation provides in-depth information on configuration of the routing components of VM TCP/IP.

References to IBM products, programs, or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe on any of the intellectual property rights of IBM may be used instead. The evaluation and verification of operation in conjunction with other products, except those expressly designed by IBM, are the responsibility of the user.

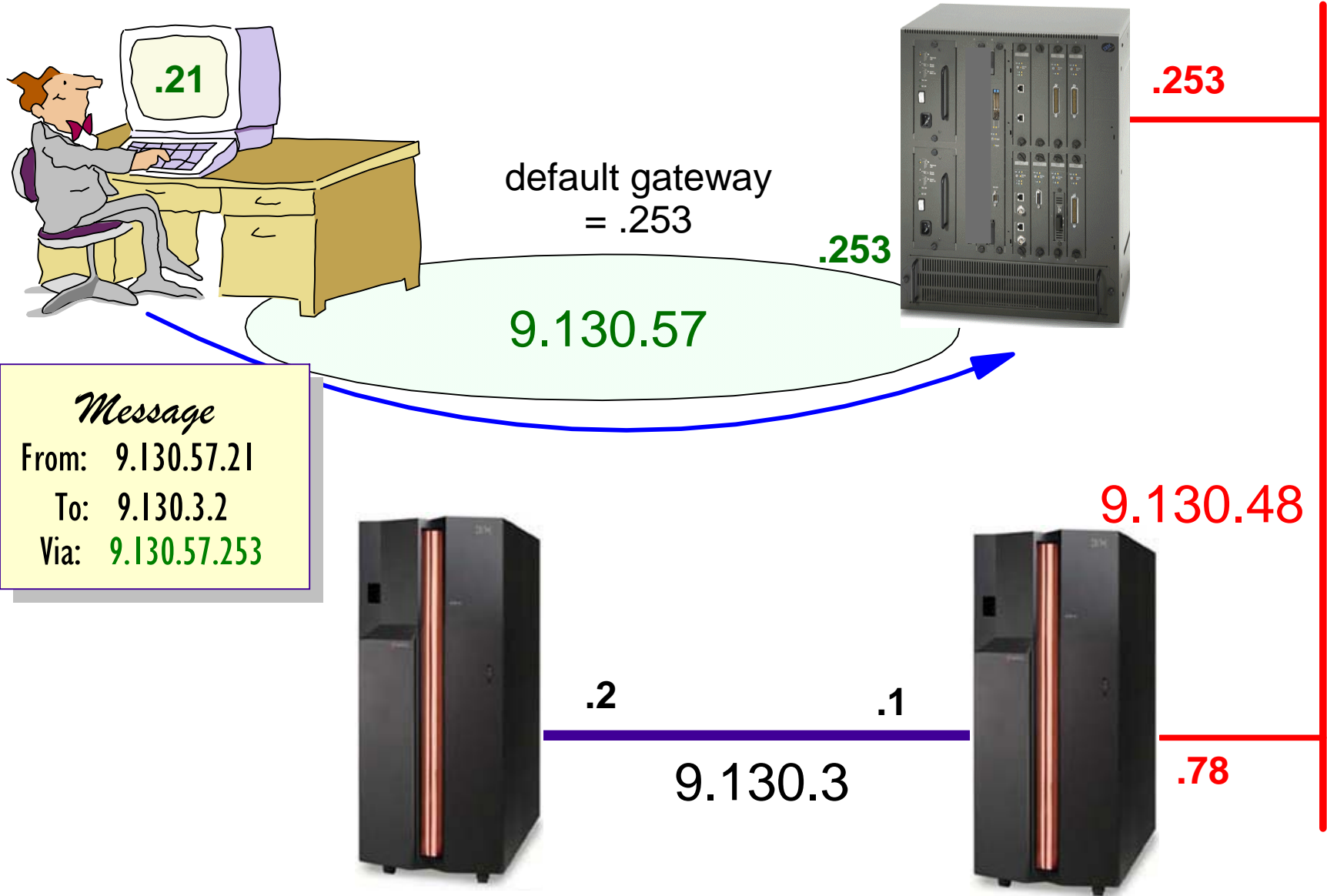
The following terms are trademarks of the International Business Machines Corporation in the United States or other countries or both: IBM IBM logo z/VM

Other company, product, and service names may be trademarks or service marks of others.

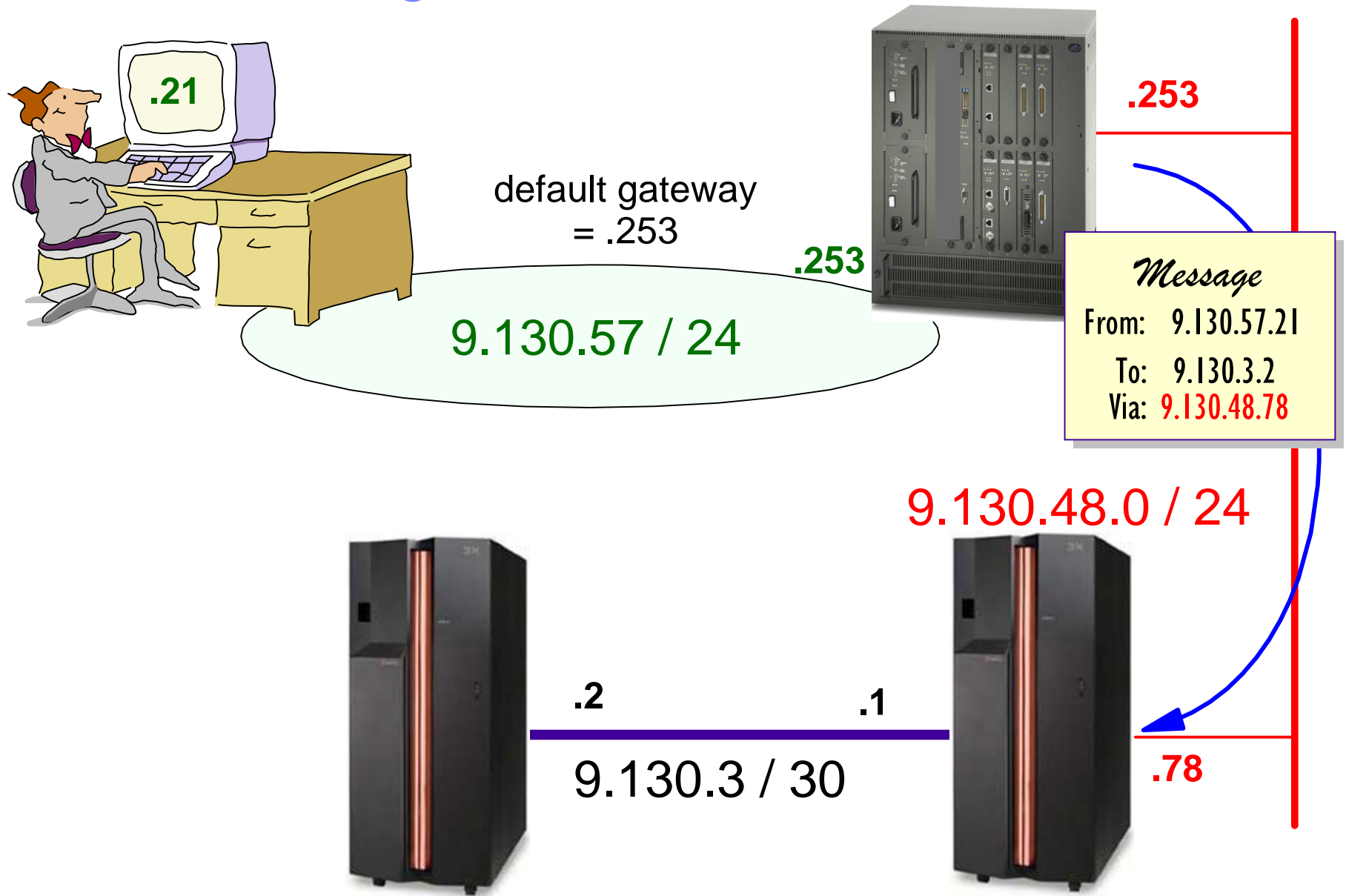
Topics

- Static Routing
- Dynamic Routing
 - ▶ RIPv2
 - ▶ OSPF
- Virtual IP addressing

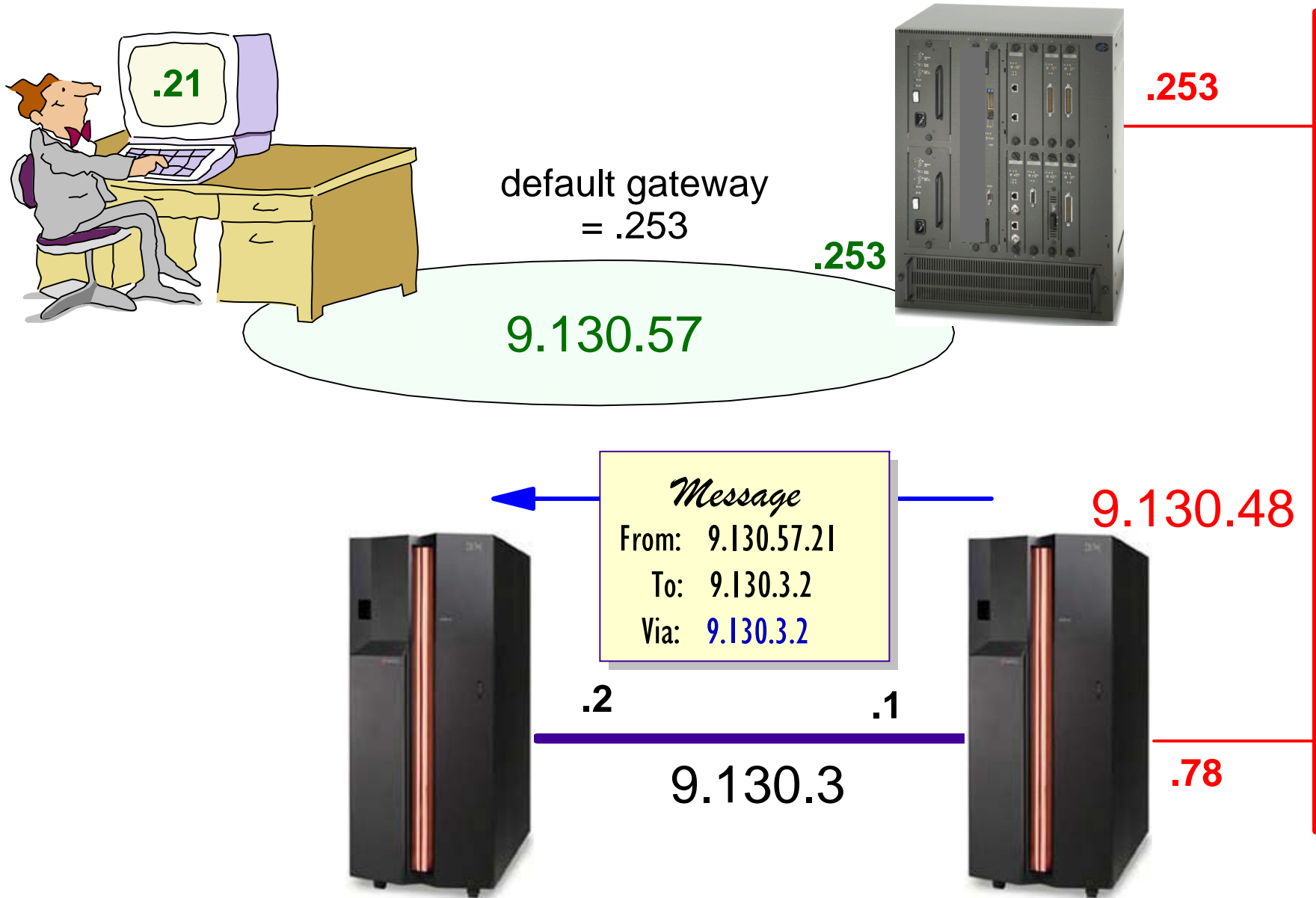
IP Packet Routing



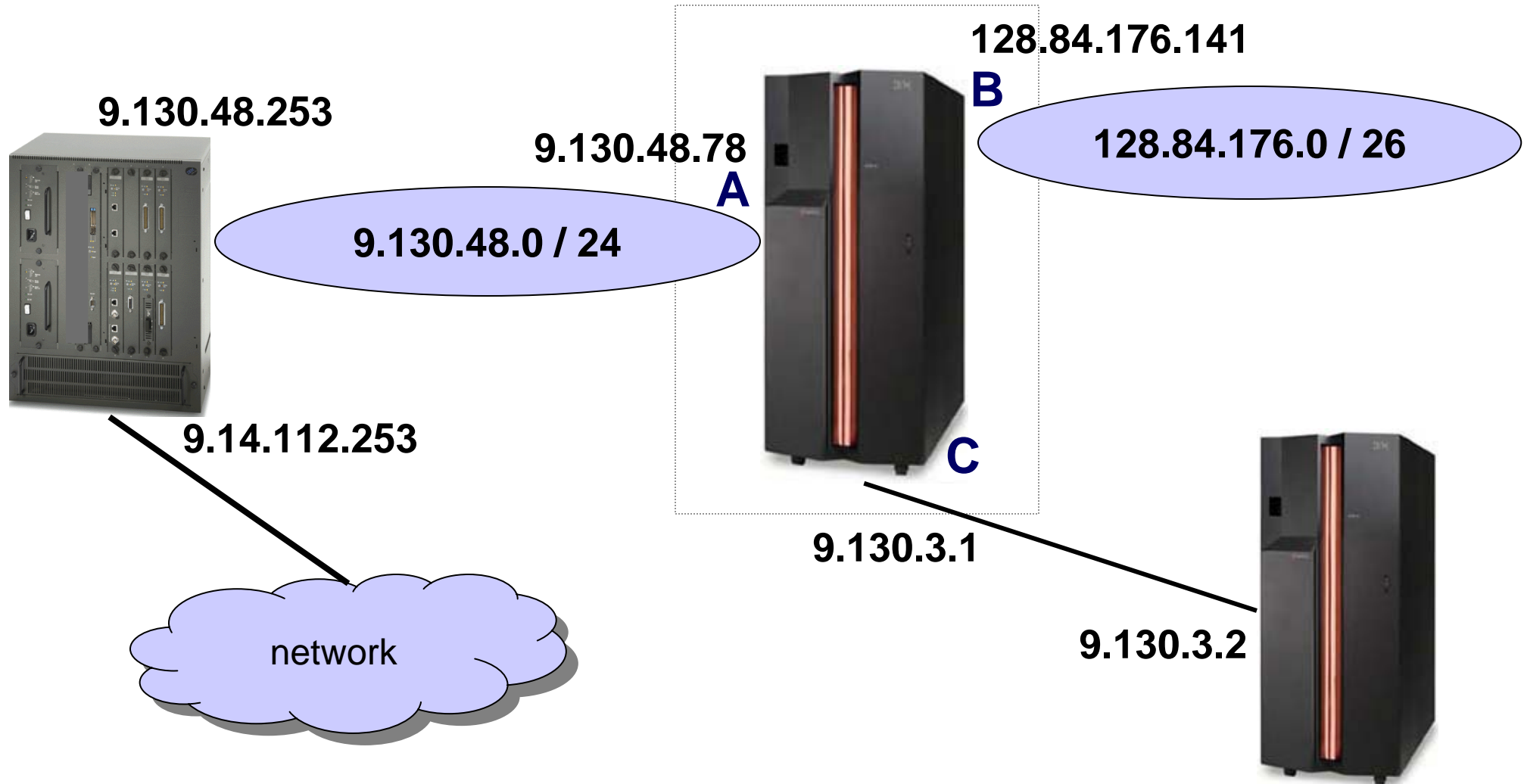
IP Packet Routing



IP Packet Routing



Routing Configuration



Static Routing Definition (pre-z/VM 5.2)

| Interface | IP Address | Subnet mask | Subnet value |
|-----------|----------------|-----------------|----------------|
| A | 9.130.48.78 | 255.255.255.0 | 9.130.48.0 |
| B | 128.84.176.141 | 255.255.255.192 | 128.84.176.128 |
| C | 9.130.3.1 | n/a | n/a |

HOME

```
9.130.48.78      A
128.84.176.141  B
9.130.3.1       C
```

```
* Network      Next hop      Ifc  MTU  subnet mask  subnet value
```

GATEWAY

```
9              =          A  1500  0.255.255.0  0.130.48.0
128.84         =          B  1500  0.0.255.192  0.0.176.128
9.130.3.2     =          C  4000  HOST
defaultnet    9.130.48.253  A  1500  0
```

Do not use BsdRoutingParms!

Static Routing Definition (z/VM 5.2)

| Interface | IP Address | Subnet mask | Subnet value |
|-----------|----------------|-----------------|----------------|
| A | 9.130.48.78 | 255.255.255.0 | 9.130.48.0 |
| B | 128.84.176.141 | 255.255.255.192 | 128.84.176.128 |
| C | 9.130.3.1 | n/a | n/a |

HOME

```

9.130.48.78/24      A
128.84.176.141/26 B
9.130.3.1/30      C

```

```

* Subnet      Mask      Next hop      Intfc MTU

```

GATEWAY

```

9.130.3.2      HOST      =      C      0
defaultnet     9.130.48.253 A      0

```

MTUs are on the LINK statement!

Static Routing - New for z/VM 5.1

The MTU can be specified on LINK

- Specify MTU of 0
 - ▶ GATEWAY
 - ▶ MPROUTE CONFIG

- Avoids duplicate information and conflicts

```
DEVICE ETH0 OSD 1F00
```

```
LINK ETH0 QDIOETHERNET ETH0 MTU 1500
```

Static Routing - New for z/VM 5.2

- The interface subnet can be specified on HOME
- Normal BSD mask or length (/nn)

HOME

```
9.130.48.78/24          ETH0
9.130.15.128 255.255.255.0 ETH1
```

| * | Subnet | Mask | Next hop | Intfc | MTU |
|---|--------|------|----------|-------|-----|
|---|--------|------|----------|-------|-----|

GATEWAY

| | | | | | |
|--|---------------|---------------|-------------|------|---|
| | 9.150.20.0/24 | | 9.130.48.5 | ETH0 | 0 |
| | 9.150.30.0 | 255.255.255.0 | 9.130.15.16 | ETH1 | 0 |
| | defaultnet | | 9.130.48.1 | ETH0 | 0 |

GATEWAY Arcana (pre-z/VM 5.2)

■ Network

- ▶ Value is class A, B, or C network only, not subnet
- ▶ Network value depends on class
- ▶ Trailing zeros may be omitted
- ▶ Must provide a default, defaultnet
 - e.g. 9, 148.12, 200.1.59

■ First hop

- ▶ "=" indicates direct link
- ▶ host IP address indicates indirect route (router)

GATEWAY Arcana (pre-z/VM 5.2)

■ Subnet Mask

- ▶ network class octets must be zero
 - e.g. A: 0.x.y.z, B: 0.0.y.z, C: 0.0.0.z
- ▶ "HOST" indicates point-to-point link
- ▶ zero indicates no subnetting

■ Subnet Value

- ▶ network class octets must be zero
- ▶ host bits not defined by mask must be zero
 - e.g. 0.0.255.240

Dynamic Routing

- **MRoute** server communicate routing information
 - ▶ Open Shortest Path First, OSPF
 - ▶ Routing Information Protocol, RIP
 - ▶ Status of local links (up / down)
 - ▶ List of directly connected networks
 - ▶ Routes to other networks or hosts learned from other servers

- Modifies IP routing table in stack

- Provides route to next hop in network

Routing Information Protocol – RIPv2

- Smaller networks – maximum 15 hops
- Best route: Fewest hops
- Multicasts routing table every 30 seconds
 - ▶ Multicasts link state change (up/down) immediately
- Listens for multicasts from other routed servers
 - ▶ Must be reminded at least once every 3 minutes
 - ▶ Uses learned routing information
- Do not use RIPv1

Open Shortest Path First (OSPF)

- Medium to Large networks – no limit on hops
- Best route: Lowest cost, as defined by IP administrators
 - ▶ Speed, usage fees, administrator preference
- Every host is a member of an *area*
 - ▶ Area topology maintained by each router in the area
 - ▶ Area *border routers* maintain routes to other areas
 - ▶ Avoids massive router tables for better performance
- Multicasts link state changes immediately
- Forwards link states learned from other routers

MPROUTE CONFIG

```
; Set default route
Default_Route Name=ETH0 Next_Hop=9.130.48.254;

; Connect to network that runs RIPv2
RIP_Interface
    IP_Address=9.130.48.78
    Name=A
    Subnet_Mask=255.255.255.0
    MTU=0          Use MTU from LINK statement
    Receive_Dynamic_Hosts=YES
    Send_Host_Routes=YES
    RIPv2=YES;
```

MPROUTE CONFIG

Area

```
Area_Number=1.1.1.1      (default is 0.0.0.0)  
Authentication_Type=None;
```

; Connect to backbone network that runs OSPF

OSPF_Interface

```
IP_Address=128.84.176.141  
Name=B  
Subnet_Mask=255.255.255.192  
Attaches_To_Area=0.0.0.0  
Cost0=2  
Router_Priority=1;
```

MPROUTE CONFIG

```
; Define point-to-point interface to host that  
; does not use dynamic routing.
```

```
Interface
```

```
  IP_Address=9.130.3.1
```

```
  Name=C
```

```
  Subnet_Mask=255.255.255.252
```

```
  Destination_Addr=9.130.3.2
```

z/VM 5.2: NO MASK OF 255.255.255.255 !!!

YOU MUST USE SUBNETS ON POINT-TO-POINT LINKS

A few words about OSPF...

- OSPF is complicated
 - ▶ Get education or work with an experienced person
 - ▶ Don't guess - know
 - ▶ If you guess, you will be caught
 - ▶ Always provide subnet masks

- Don't try to take shortcuts
 - ▶ Does not work with Proxy ARP (multicast!)
 - ▶ You have to define ALL interfaces else defaults will be used!

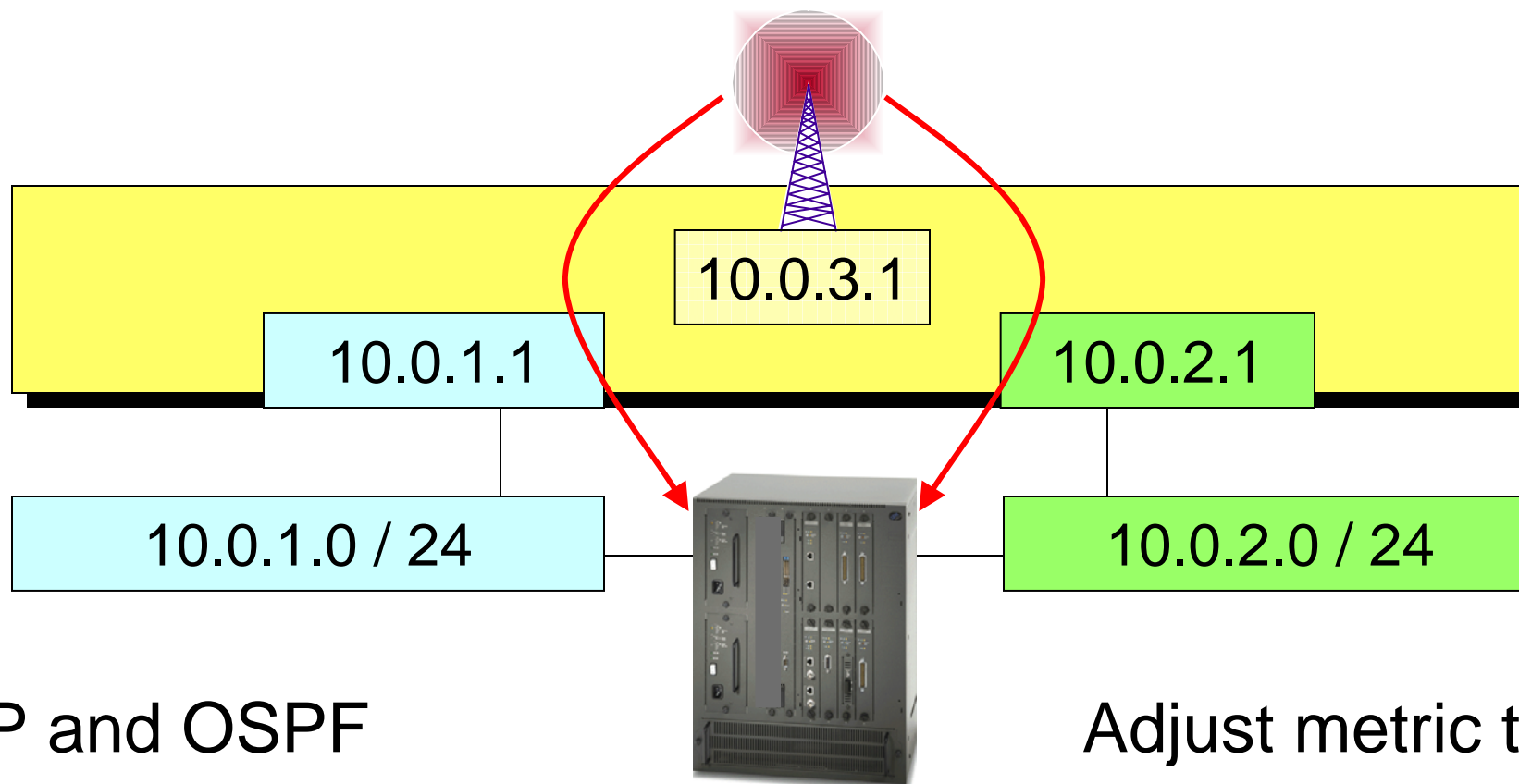
NETSTAT GATE (pre-z/VM 5.2)

| NetAddress | FirstHop | Flgs | PktSz | Subnet | Mask | Subnet Value | Link |
|---------------|---------------|------|-------|---------------|-------|---------------|------------|
| ----- | ----- | ---- | ----- | ----- | ----- | ----- | ----- |
| Default | 9.130.48.254 | UG | 2000 | <none> | | | ISRING |
| 9.0.0.0 | 9.130.235.1 | UG | 1500 | <none> | | | ETRING |
| 9.130.3.17 | 9.130.248.99 | UGH | 2000 | HOST | | | FDRING |
| 9.130.3.35 | <direct> | UH | 576 | HOST | | | ALANCTC |
| 9.130.3.36 | <direct> | HS | 9216 | HOST | | | LINUX2 |
| 9.130.3.38 | <direct> | HS | 1500 | HOST | | | LINUX2_CTC |
| 9.0.0.0 | <direct> | U | 2000 | 0.255.255.240 | | 0.130.3.48 | ISRING |
| 9.130.3.129 | 9.130.248.99 | UGH | 2000 | HOST | | | FDRING |
| 9.130.3.162 | 9.130.248.99 | UGH | 2000 | HOST | | | FDRING |
| 9.0.0.0 | <direct> | U | 2000 | 0.255.255.0 | | 0.130.48.0 | ISRING |
| 9.130.48.134 | 9.130.248.99 | UGH | 2000 | HOST | | | FDRING |
| 9.0.0.0 | <direct> | U | 1500 | 0.255.255.0 | | 0.130.176.0 | ETRING1 |
| 9.130.176.111 | 9.130.176.100 | UGH | 1500 | HOST | | | ETRING1 |
| 9.0.0.0 | <direct> | U | 1500 | 0.255.255.192 | | 0.130.235.0 | ETRING |
| 9.0.0.0 | <direct> | U | 2000 | 0.255.255.240 | | 0.130.248.96 | FDRING |
| 9.0.0.0 | <direct> | U | 9180 | 0.255.255.240 | | 0.130.248.112 | ATMTEST |
| 9.130.249.32 | 9.130.248.99 | UGH | 2000 | HOST | | | FDRING |

NETSTAT GATE (z/VM 5.2)

| Subnet Address | Subnet Mask | FirstHop | Flgs | PktSz | Metric | Link |
|----------------|-----------------|------------|------|-------|--------|------|
| Default | <none> | 9.56.63.3 | UG | 1492 | 34 | ETH2 |
| Default | <none> | 9.56.63.4 | UG | 1492 | 34 | ETH2 |
| Default | <none> | 9.56.212.3 | UG | 1492 | 34 | ETH1 |
| Default | <none> | 9.56.212.4 | UG | 1492 | 34 | ETH1 |
| 9.56.63.0 | 255.255.255.0 | <direct> | U | 1492 | 1 | ETH2 |
| 9.56.208.0 | 255.255.255.248 | 9.56.63.3 | UG | 1492 | 101 | ETH2 |
| 9.56.208.0 | 255.255.255.248 | 9.56.63.4 | UG | 1492 | 101 | ETH2 |
| 9.56.208.0 | 255.255.255.248 | 9.56.212.3 | UG | 1492 | 101 | ETH1 |
| 9.56.208.0 | 255.255.255.248 | 9.56.212.4 | UG | 1492 | 101 | ETH1 |
| 9.56.208.8 | 255.255.255.248 | 9.56.63.3 | UG | 1492 | 101 | ETH2 |
| 9.56.208.8 | 255.255.255.248 | 9.56.63.4 | UG | 1492 | 101 | ETH2 |
| 9.56.208.8 | 255.255.255.248 | 9.56.212.3 | UG | 1492 | 101 | ETH1 |
| 9.56.208.8 | 255.255.255.248 | 9.56.212.4 | UG | 1492 | 101 | ETH1 |
| 9.56.208.16 | 255.255.255.252 | 9.56.63.3 | UG | 1492 | 101 | ETH2 |
| 9.56.208.16 | 255.255.255.252 | 9.56.63.4 | UG | 1492 | 101 | ETH2 |
| 9.56.208.16 | 255.255.255.252 | 9.56.212.3 | UG | 1492 | 101 | ETH1 |
| 9.56.208.16 | 255.255.255.252 | 9.56.212.4 | UG | 1492 | 101 | ETH1 |
| 9.56.208.40 | HOST | 9.56.63.3 | UGH | 1492 | 2 | ETH2 |
| 9.56.208.40 | HOST | 9.56.212.3 | UGH | 1492 | 2 | ETH1 |

Virtual IP Addressing - the soul of the machine

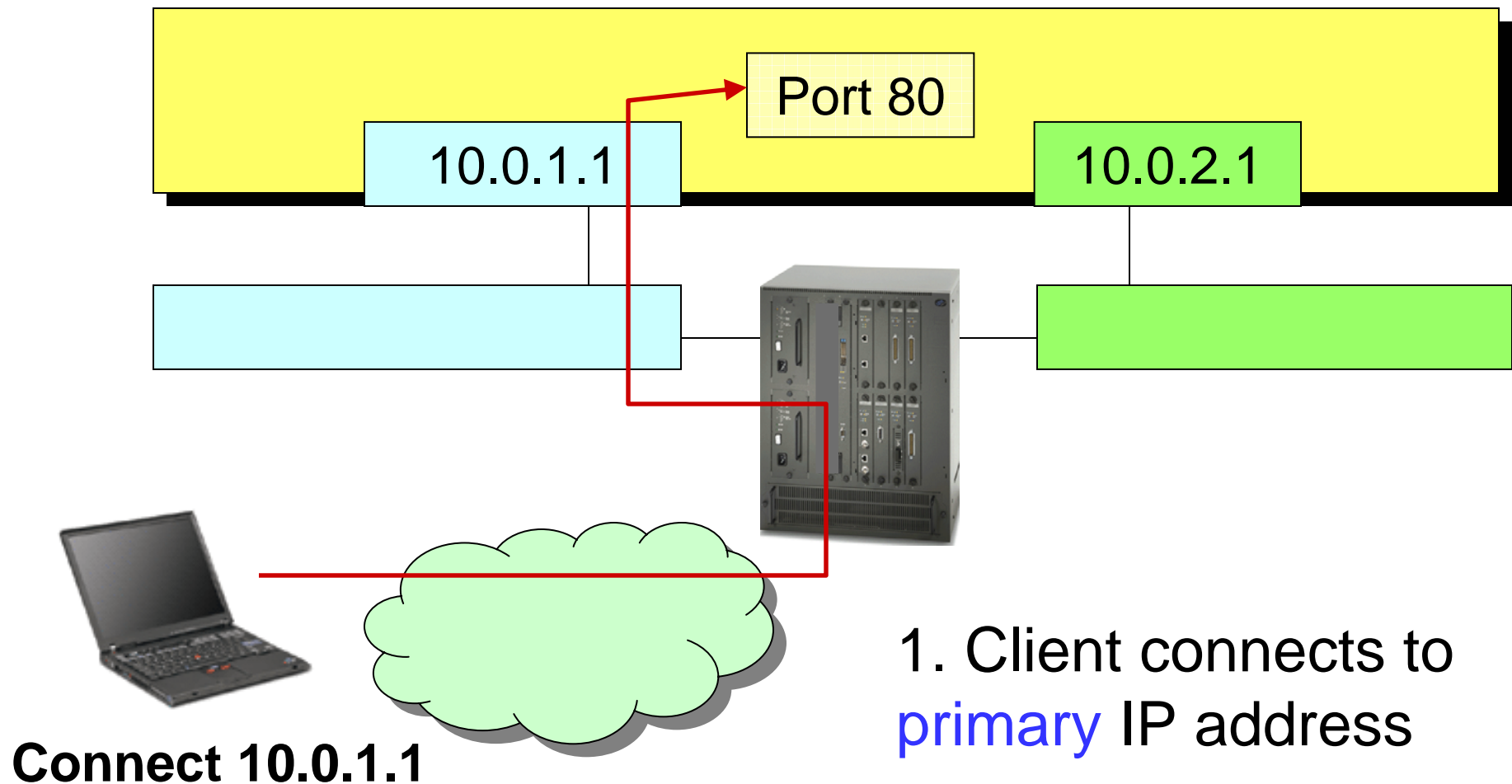


RIP and OSPF
broadcast virtual
route via physical
adapters

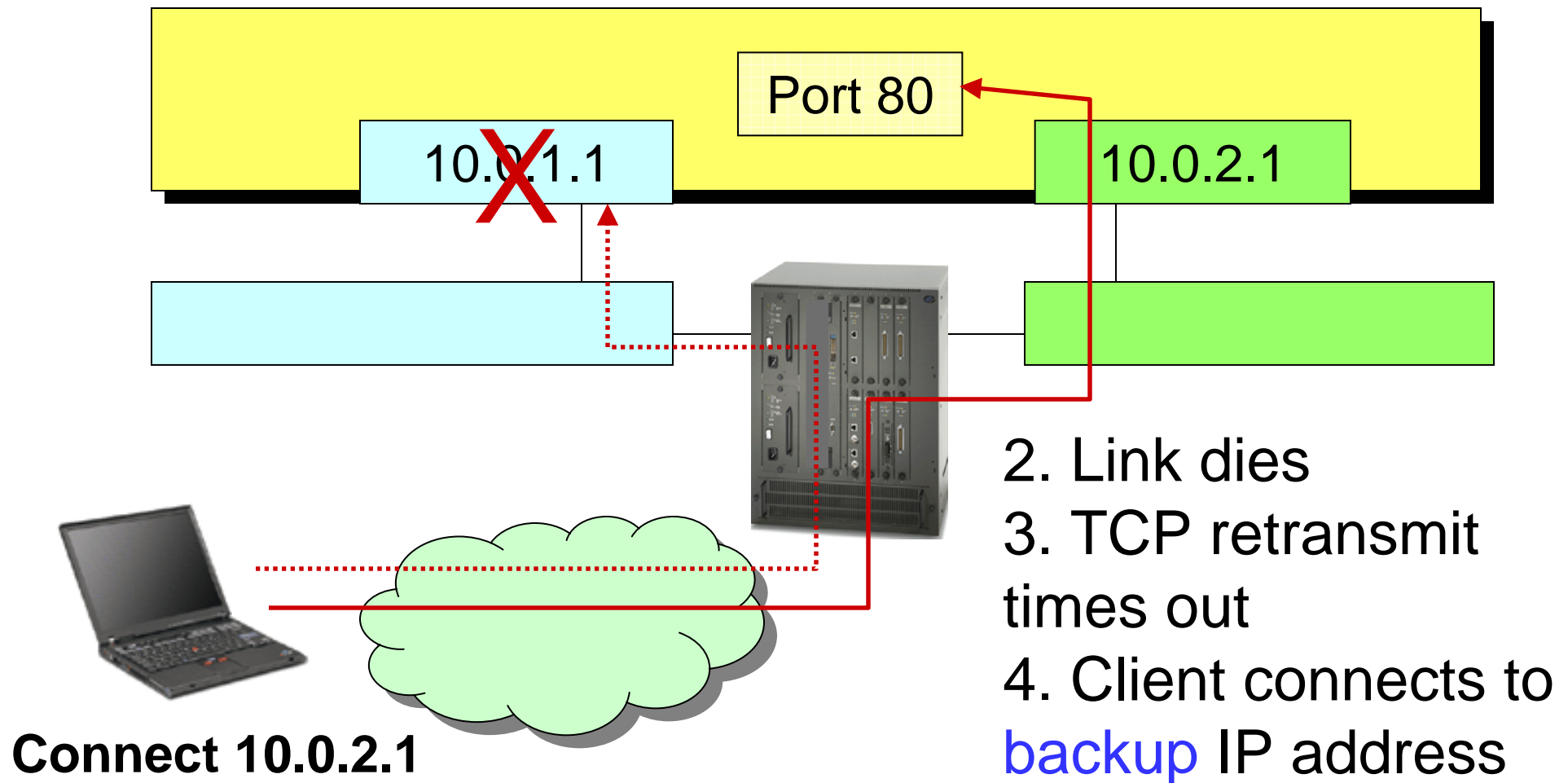
Destination 10.0.3.1
Via 10.0.1.1 metric 1
Via 10.0.2.1 metric 1

Adjust metric to
give preference
to a particular
adapter

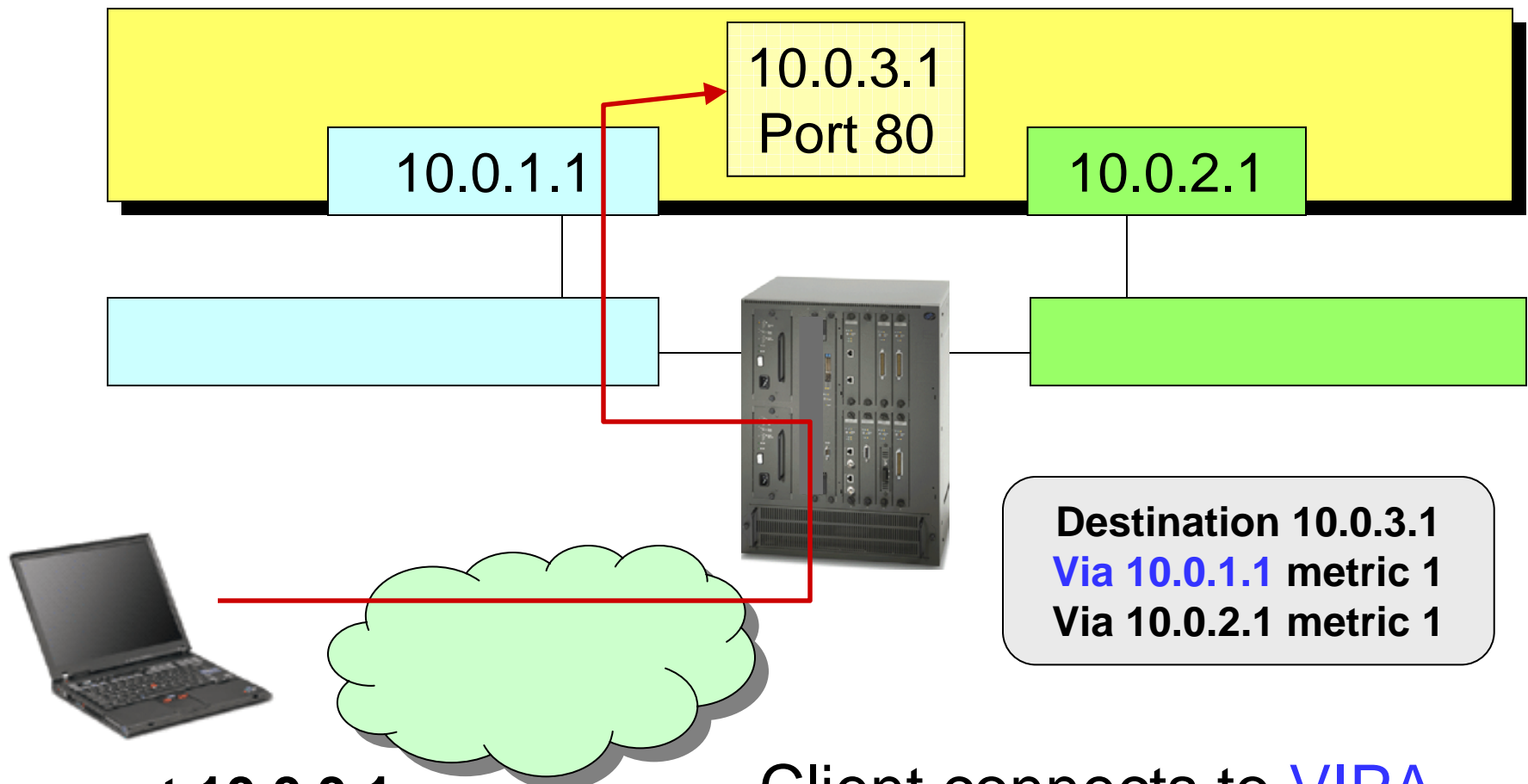
Without VIPA: Session establishment



Without VIPA: User-Initiated Failover



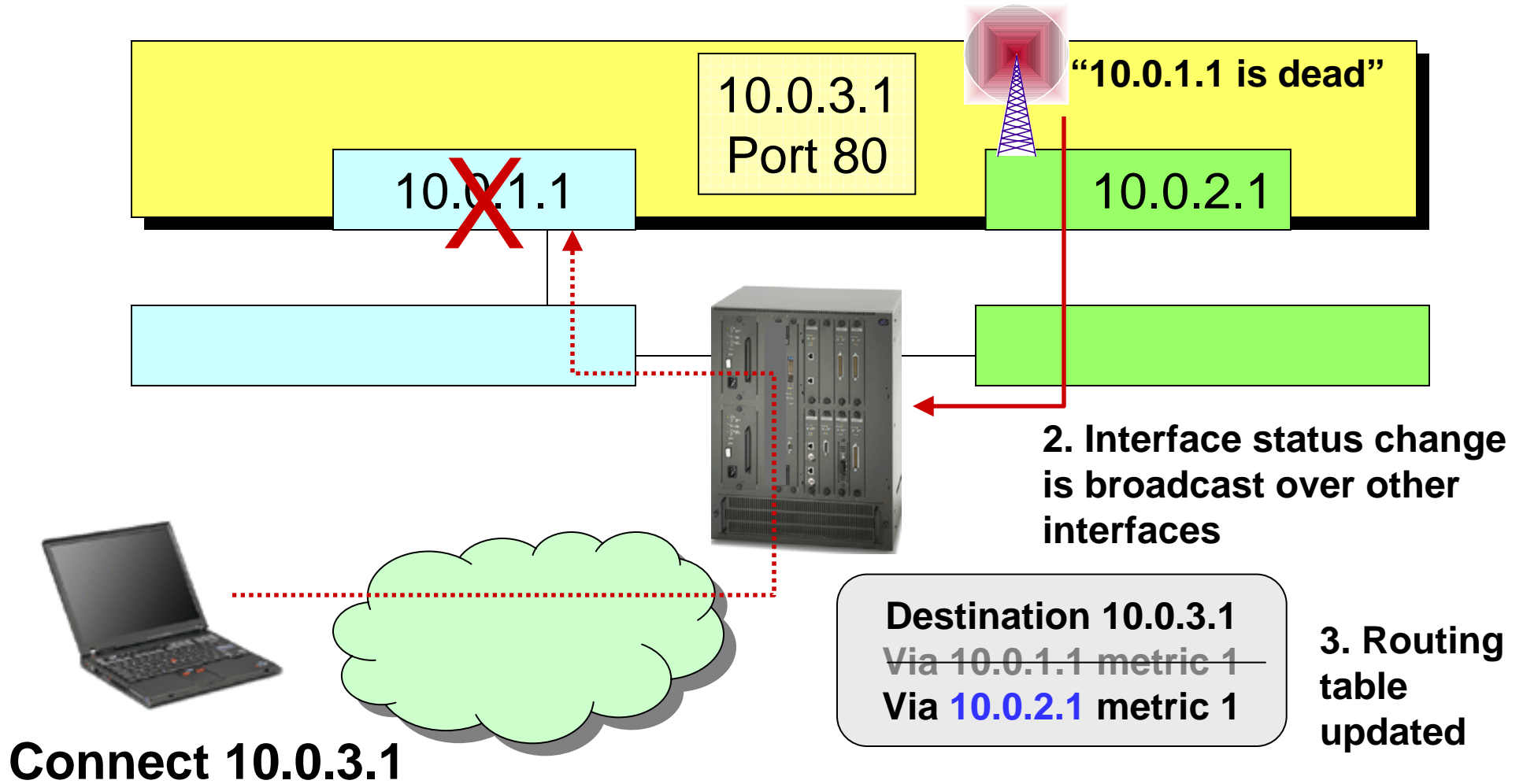
With VIPA: Session establishment



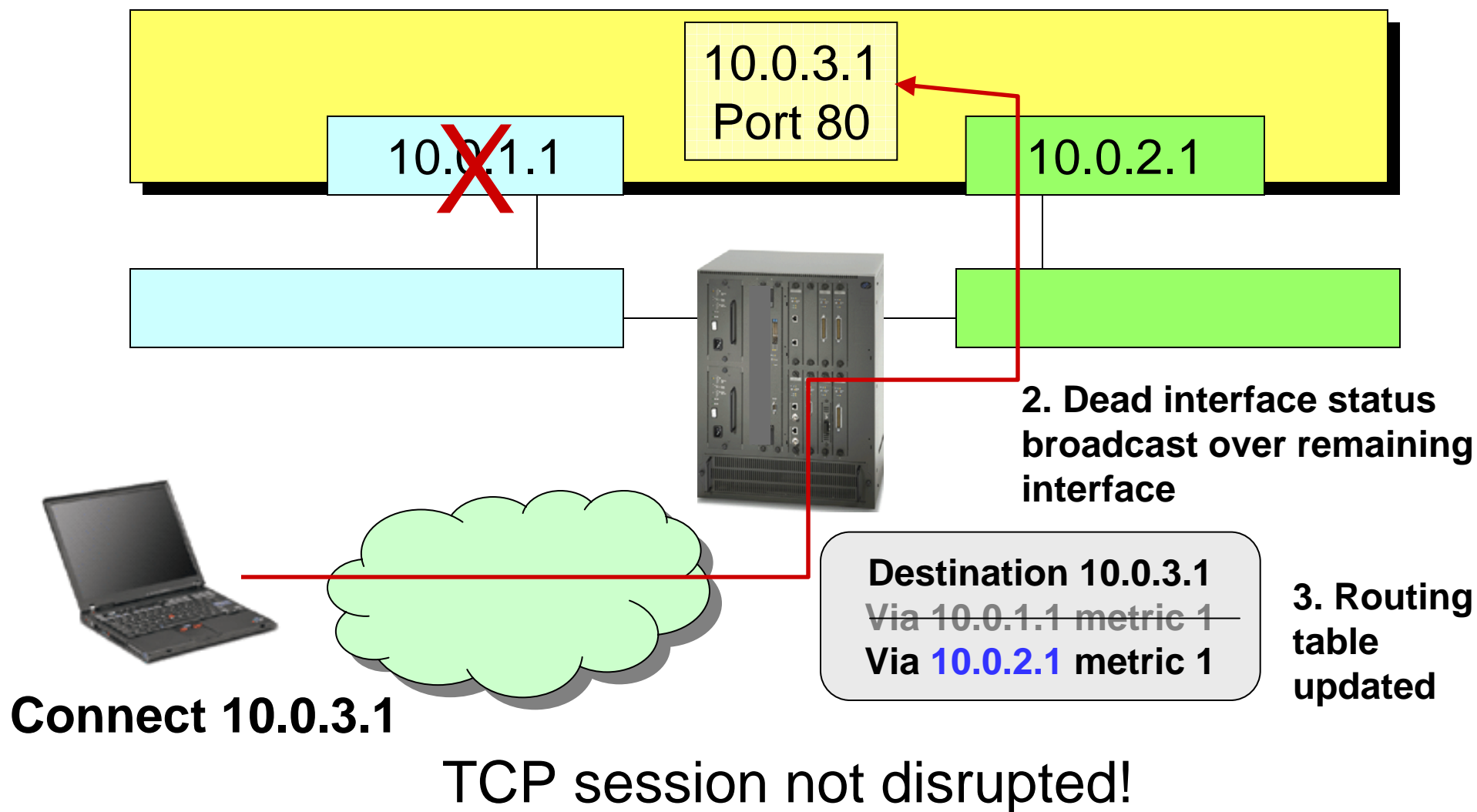
Connect 10.0.3.1

Client connects to **VIPA**

With VIPA: Network-Initiated Failover



With VIPA: Network-Initiated Failover



VIPA Benefits

- Insulates clients from IP address changes
- Protects clients from interface outages
- Provides increased application availability
- Designed for use with dynamic routing
 - ▶ Can use with static routing, but no protection from interface outage unless using routers with black hole detection

IPv4 Addressing Quick Reference

| Class | First octet | Network |
|-------|-------------|---------|
| A | 0-127 | a.0.0.0 |
| B | 128-191 | a.b.0.0 |
| C | 192-223 | a.b.c.0 |
| D | 224-239 | n/a |

| Mask size | Last octet | binary | subnetwork | # hosts |
|-----------|------------|-----------|---|---------|
| /25 | 128 | 1000 0000 | 2: 0 128 | 126 |
| /26 | 192 | 1100 0000 | 4: 0 64 128 192 | 62 |
| /27 | 224 | 1110 0000 | 8: 0 32 64 96 128 160 192 224 | 30 |
| /28 | 240 | 1111 0000 | 16: 0 16 32 48 64 80 96 112 128 144 160 176 192 208 224 240 | 14 |
| /29 | 248 | 1111 1000 | 32: 0 8 16 24 32 40 48 56 64 72 80 88 96 104 112 120 128 136 144 152 160 168 176 184 192 200 208 216 224 232 240 248 | 6 |
| /30 | 252 | 1111 1100 | 64: 0 4 8 16 20 24 28 32 36 ... | 2 |

Read More About It...

- z/VM TCP/IP Planning and Customization, SC24-6019
- TCP/IP Illustrated, Vol. 1
Addison Wesley
W. Richard Stevens
ISBN 0-201-63346-9
- Internetworking with TCP/IP
Prentice Hall
Douglas P. Comer
ISBN 0-13-216987-8

Contact Information

- By e-mail: `Alan_Altmark@us.ibm.com`
- In person: USA 607.429.3323
- On the Web: <http://ibm.com/vm/devpages/altmarka>
- Mailing lists:
`IBMTCP-L@vm.marist.edu`
`VMESA-L@listserv.uark.edu`
`LINUX-390@vm.marist.edu`

<http://ibm.com/vm/techinfo/listserv.html>