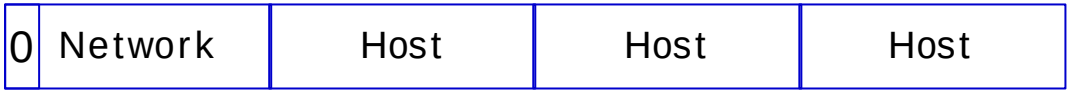
IP Routing Primer

Paul C. Huang, Ph.D.
ITRI / CCL / N300

ITRI
CCL

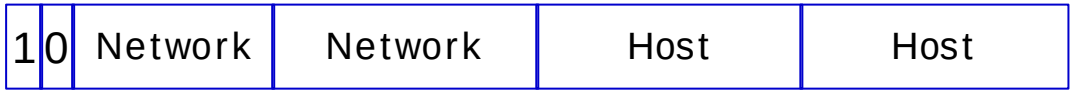
Basic Addressing / Subnetting

Class A



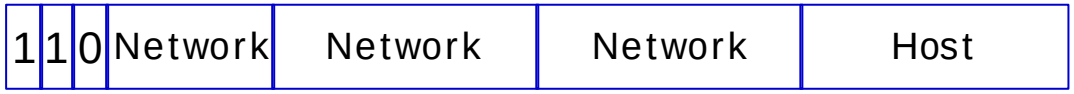
127 networks / 16,777,216 hosts
Class A natural mask 255.0.0.0
Mask of 255.255.0.0 creates 127 networks, 256 subnets/network, and 65,636 hosts/subnet

Class B



16,384 networks / 65,5366 hosts
Class B natural mask 255.255.0.0

Class C



2,097,152 networks / 256 hosts
Class C natural mask 255.255.255.0

```
IP Address   10.   0.0.1      00001010 00000000 00000000 00000001
Mask        255.255.0.0      11111111 11111111 00000000 00000000
                             Network  Subnet   Host
```

ITRI
CCL

Variable Length Subnet Mask

- Ability to configure one network with different masks
  - » more flexibility in dividing a network into multiple subnets while still maintaining an adequate number of hosts in each subnet.
  - » Example: Class C address with mask 255.255.255.X
    - – 255.255.255.252 (1111 1100) 64 subnets w/ 4 hosts each
    - – 255.255.255.248 (1111 1000) 32 subnets w/ 8 hosts each
    - – 255.255.255.240 (1111 0000) 16 subnets w/ 16 hosts each
    - – 255.255.255.224 (1110 0000) 8 subnets w/ 32 hosts each
    - – 255.255.255.192 (1100 0000) 4 subnets w/ 64 hosts each
    - – 255.255.255.128 (1000 0000) 2 subnets w/ 128 hosts each
  - » If you need to divide a Class C network into 3 subnets with 100 hosts in one subnet and 50 hosts in the remaining two subnets, VLSM is required.
  - » RIPv1 and IGRP do not handle VLSM
  - » OSPF, EIGRP, ISIS, RIPv2 handle VLSM

ITRI
CCL

Handling IP depletion

- Creative IP address space allocation
  - » VLSM
  - » Non-permanent allocation of address blocks
  - » Contiguous class C network
- Classless Interdomain Routing (CIDR)
  - » RFC 1519
- Private addressing and Network Address Translation (NAT)
  - » RFC 1918 Address allocation for private internets
  - » RFC 1631 The IP Network Address Translator (NAT)
- IP v6
  - » New Address Space

ITRI
CCL

Classless Interdomain Routing (CIDR)

- IP network represented by a <prefix, length>
  - » No more traditional IP classes A/B/C
  - » Indicates a more hierarchical Internet architecture, where each domain takes its IP addresses from a higher hierarchical level
  - » Allows multiple classes of blocks to be lumped together as an *Aggregate, CIDR block, or a Supernet*
    - – identified when the prefix length is shorter the network's natural mask.
    - – allows a list of contiguous IP networks to be summarized into one announcement.
  - » Example: <198.32.0.0/16>
    - – illegal class C network
    - – supernet because 16 < 24 (less than natural mask length)
    - – 198.32.1.0/20 is identified as a more *specific prefix* of length 20

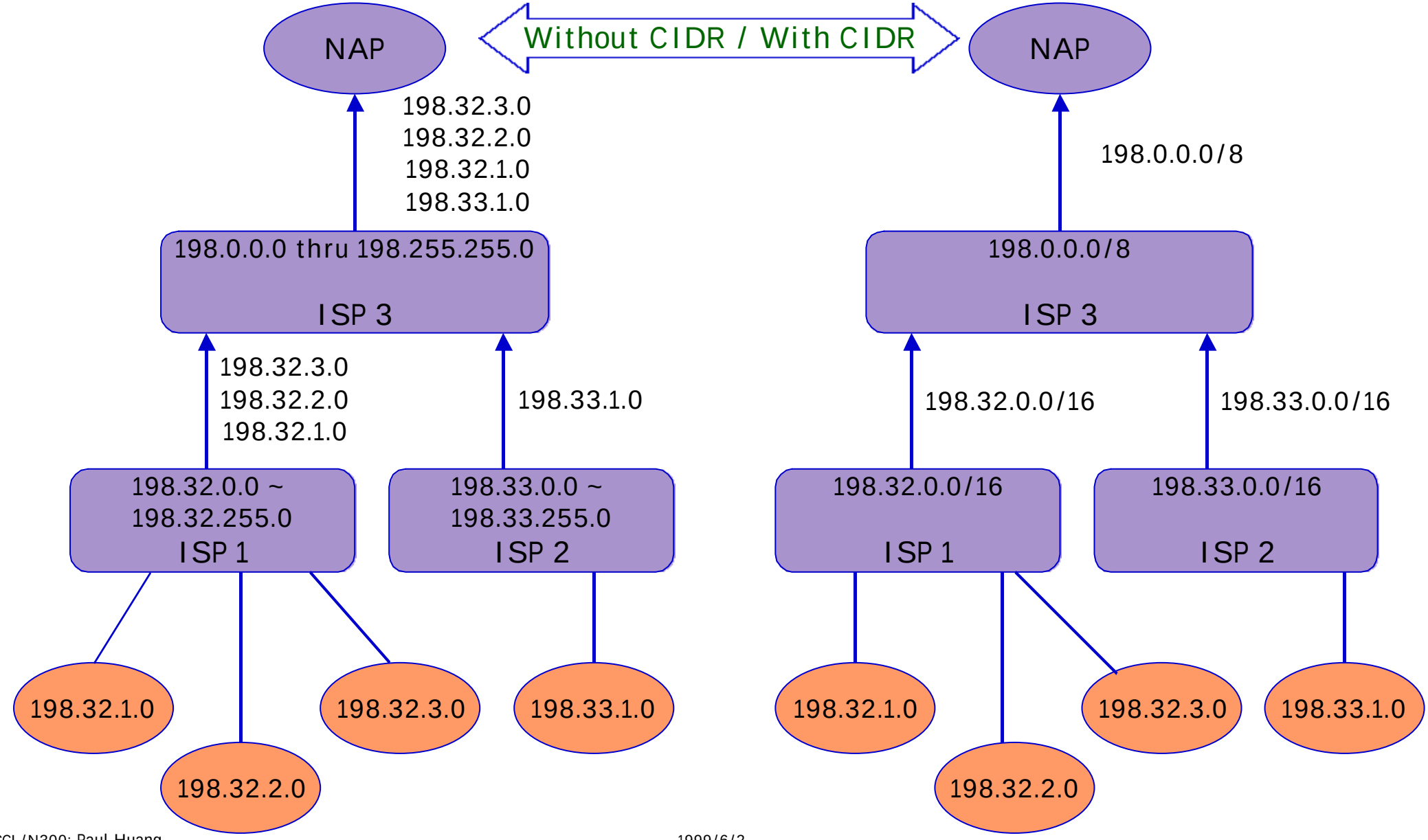
ITRI
CCL

Classless Interdomain Routing (CIDR)

- Advantages
  - » Move towards a more hierarchical, geographically-oriented network
  - » More efficient routing strategies and propagation
    - – Tremendous savings in route propagation when summarization is done close to the *leaf networks.*
    - – Aggregation permits the ISP to advertise the addresses in a single notation rather than many.
- Potential problems
  - » CIDR works only if every customer connects to his provider
    - – via one connection only (a scenario called *single-homing*)
    - – take its IP addresses from its provider's prefixes
  - » Multi-homing – connecting to multiple providers
  - » Already have previously applied IP address

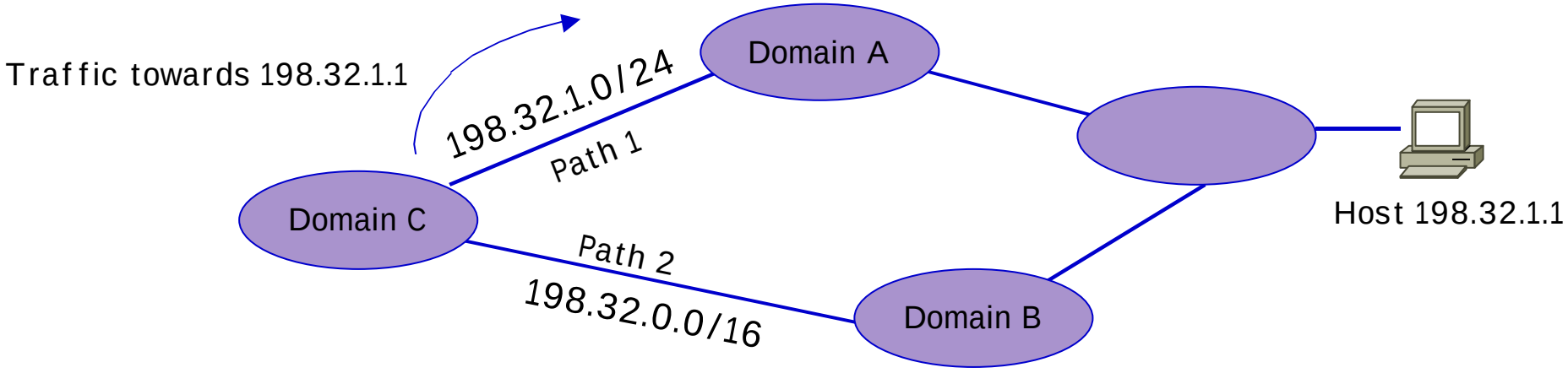
ITRI
CCL

CIDR Example

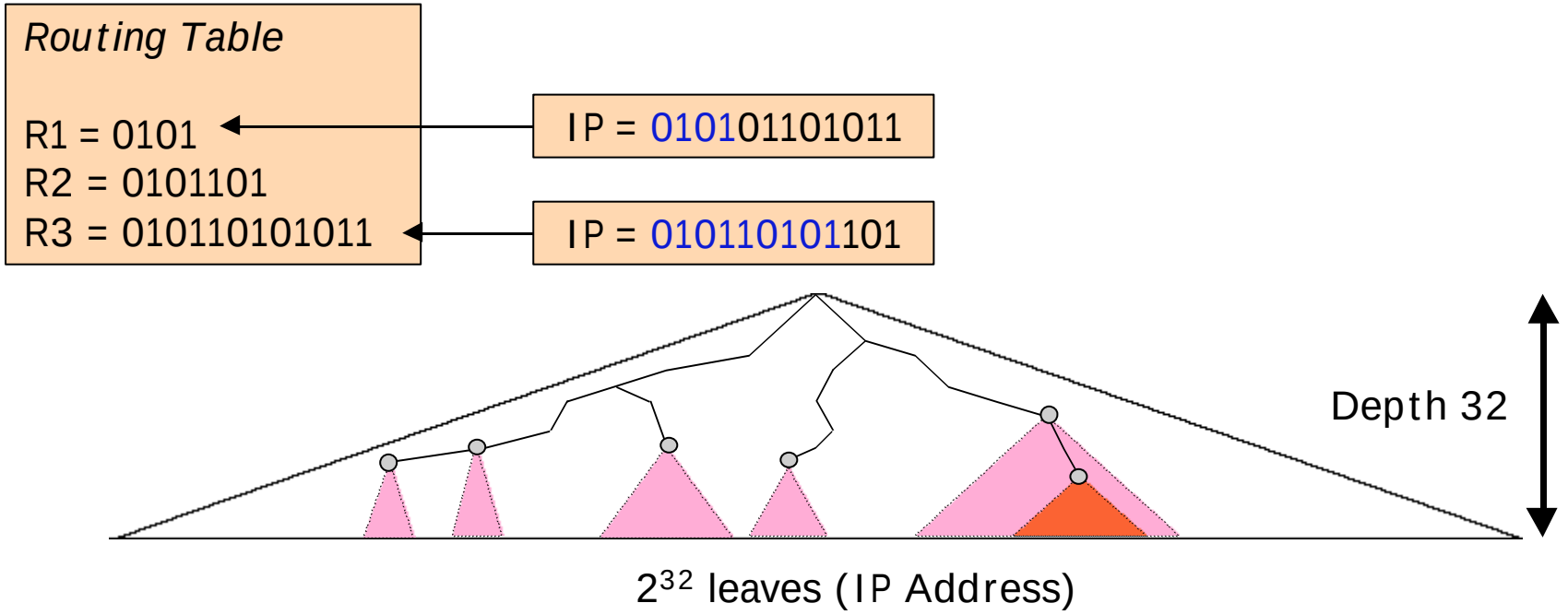


Longest Match Routing Rule

- Routing is always done on a longest match basis
  - » will always follow the longer mask for two different length prefix of the same network
  - » destinations connected to multiple domains must always be explicitly announced in their most specific, not aggregate forms.
  - » Example:
    - 198.32.1.0/24 via path 1
    - 198.32.0.0/16 via path 2

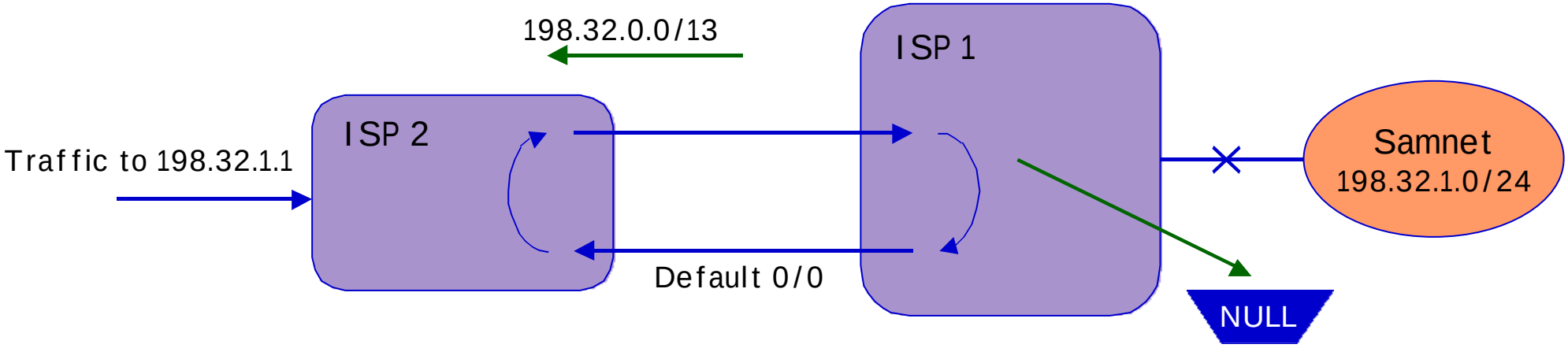


Longest Prefix Match



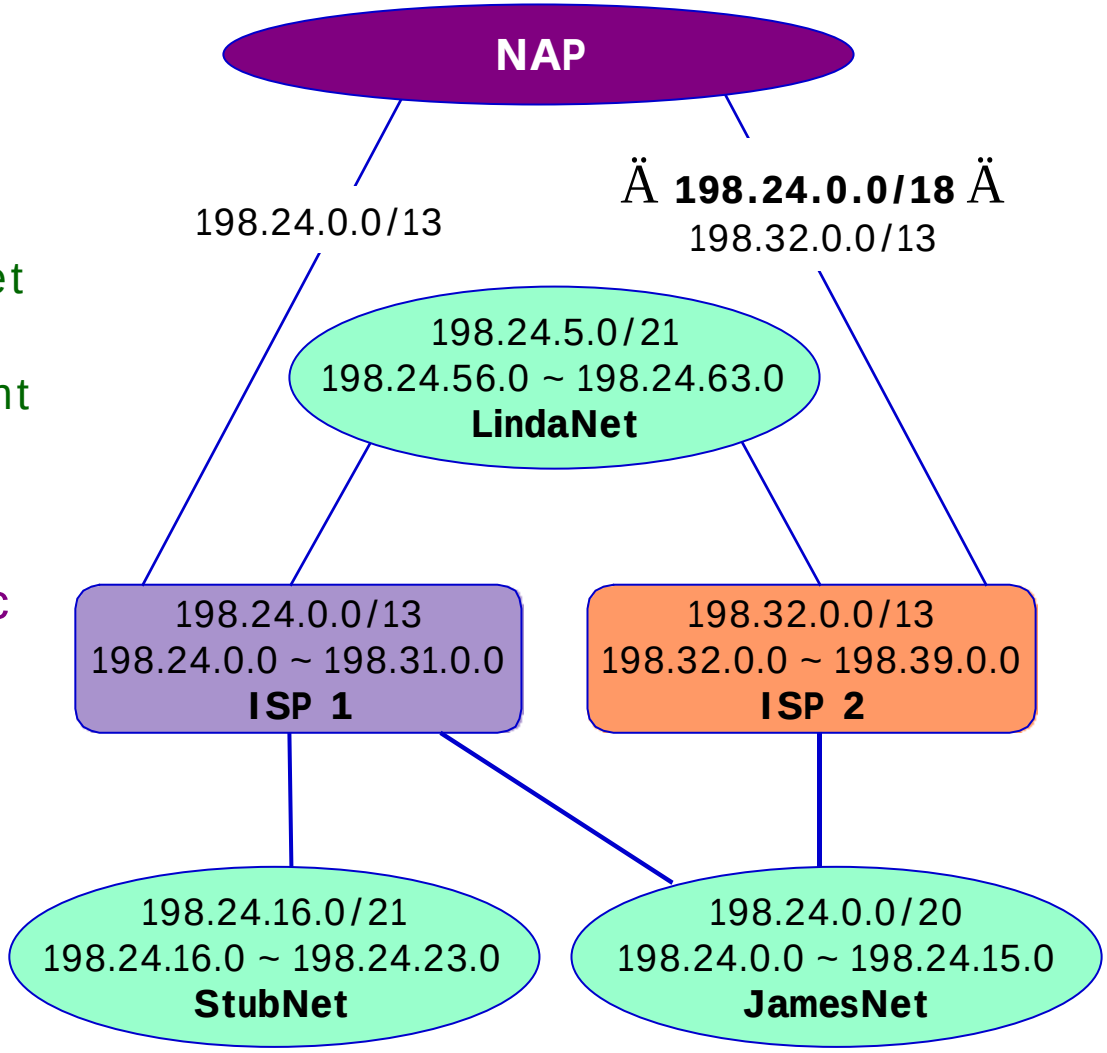
Routing Loop

- Occurs when traffic circles back and forth between domains, never reaching its final destination.
  - » Prevention
    - – Must not follow a less specific route for a destination that matches one of its own aggregate routes.
    - – Drop all packets destined to an unreachable destination less specific than the aggregate route
    - – Beware of *"Black holes"* where traffic reaches and stops at a destination that is not its intended destination, but from which it cannot be forwarded.



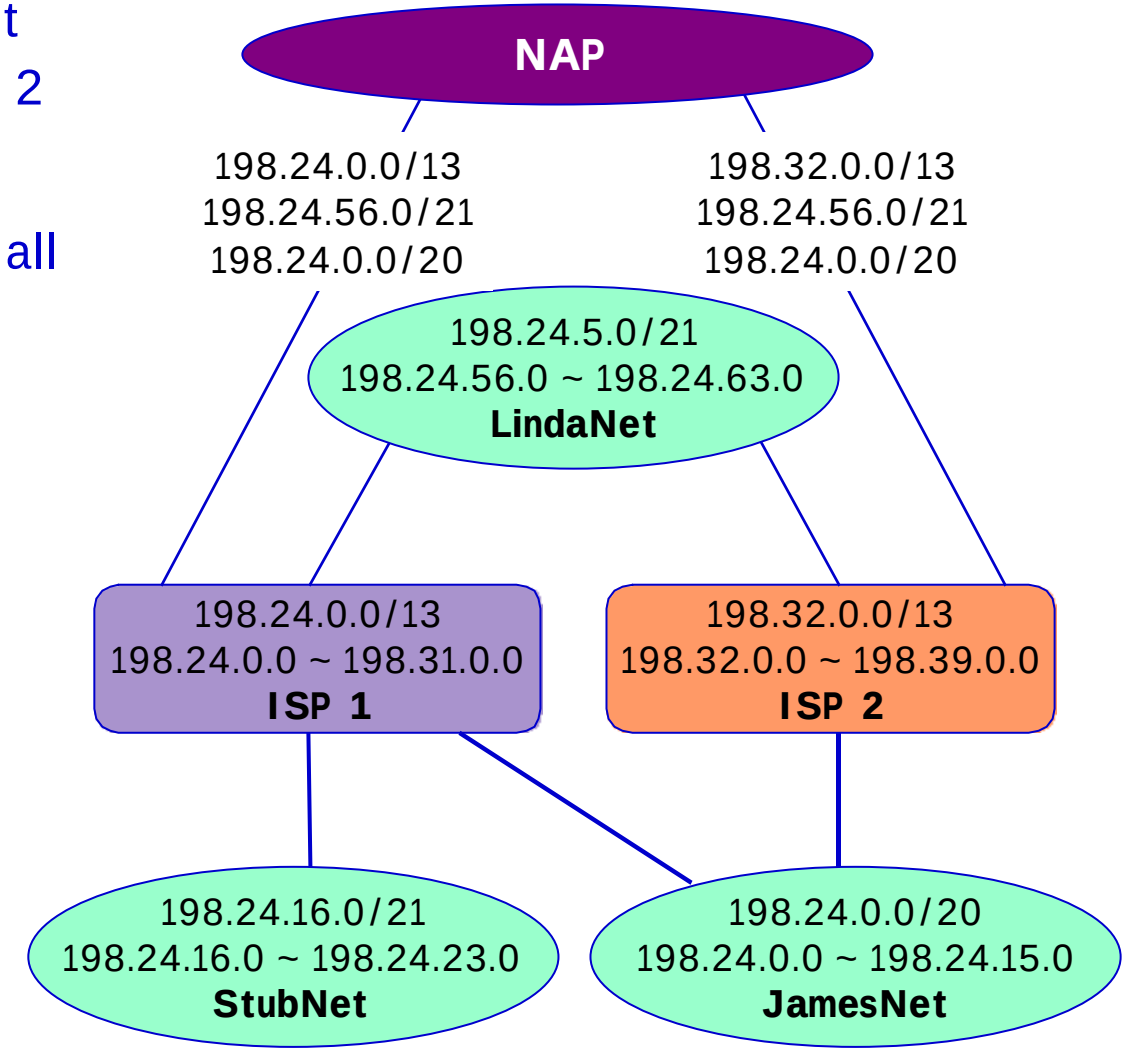
Multihoming Scenario: Address taken from one provider

- Advertising aggregates is tricky
  - » no one is allowed to aggregate someone else's routes (proxy aggregation) unless
    - – the aggregating party is a superset of the other party
    - – both parties are in total agreement
- Example
  - » 198.24.0.0/18 will create a black hole for StubNet, because traffic will follow the longest match and end up in ISP 2 with no link to StubNet.



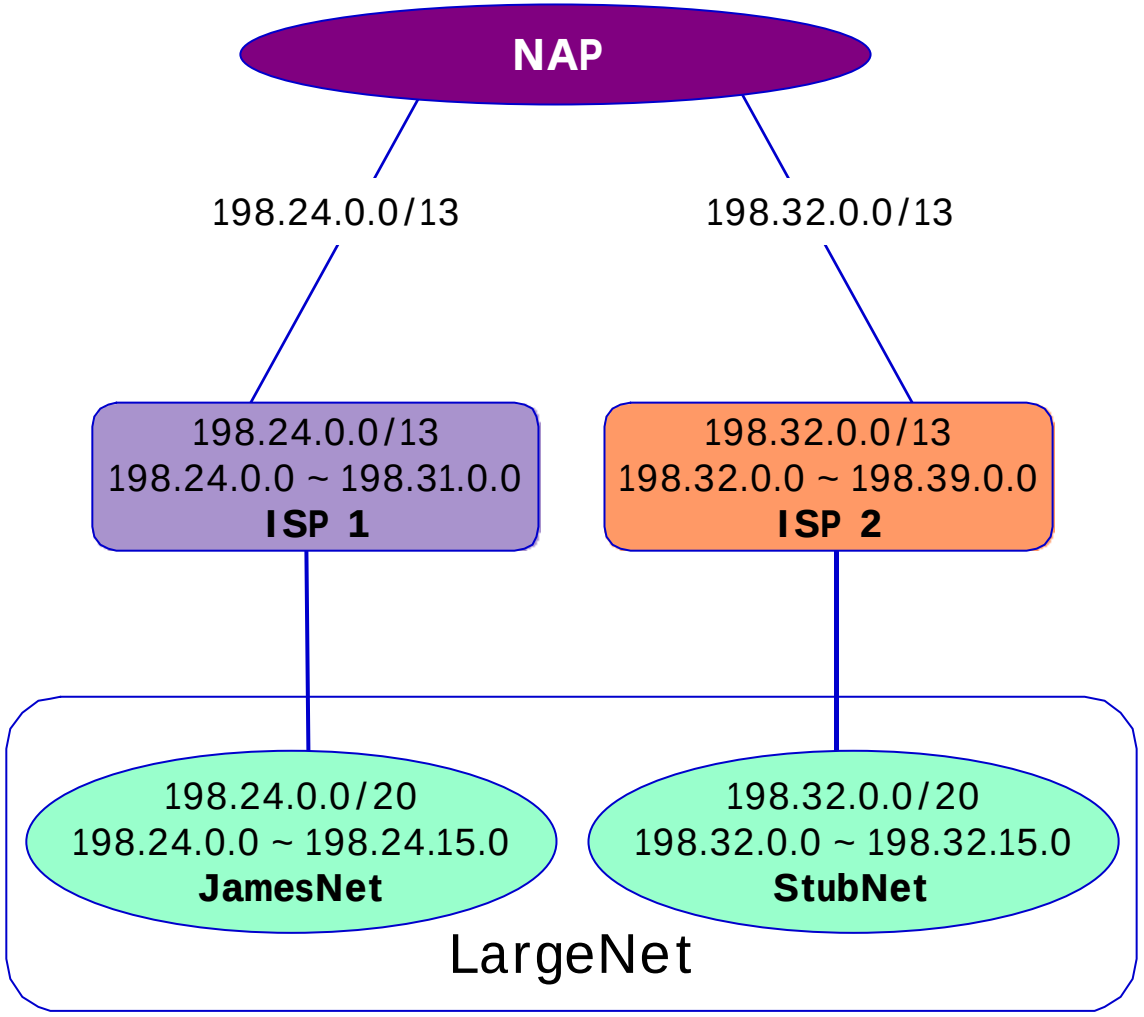
ITRI CCL Multihoming Scenario: Address taken from one provider

- In this case, explicit aggregates must be advertised by both ISP 1 and ISP 2
- If ISP 1 only advertise the less-specific aggregate of 198.24.0.0/13, all traffic towards JamesNet and LindaNet would always follow the longest match via ISP2.



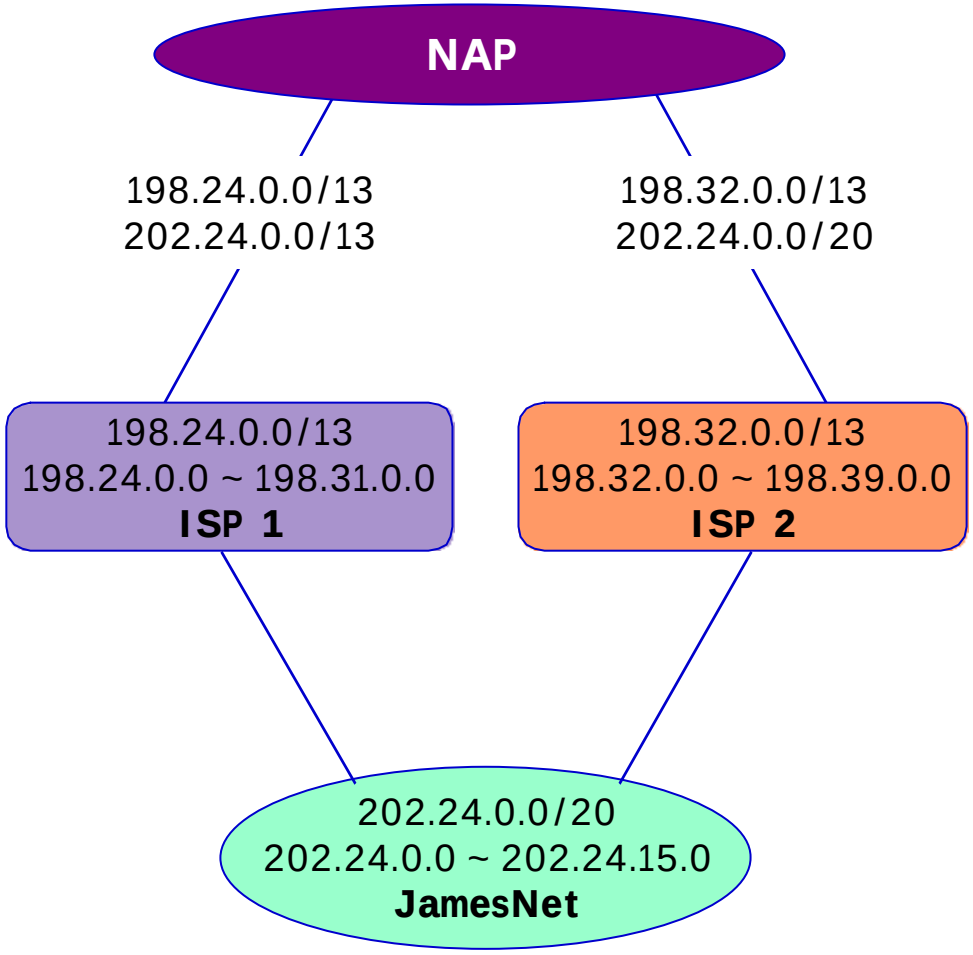
ITRI CCL Multihoming Scenario: Address taken from different provider

- In this example, each provider will be able to aggregate its own address space without having to list specific ranges from the other provider.
- Major drawback
  - » backup routes to multihomed organizations are not maintained.
  - » each only advertise its own block of addresses, not from other ISPs
  - » portions of network will be out if one of the ISPs is out.



Multihoming Scenario: Address taken from none of the provider

- In this example, the address space is not from those of ISP 1 or ISP 2.
  - » Both ISPs must advertise a specific aggregate on top of their own ranges.
- Major drawback
  - » increase the size of the routing table for all routers.

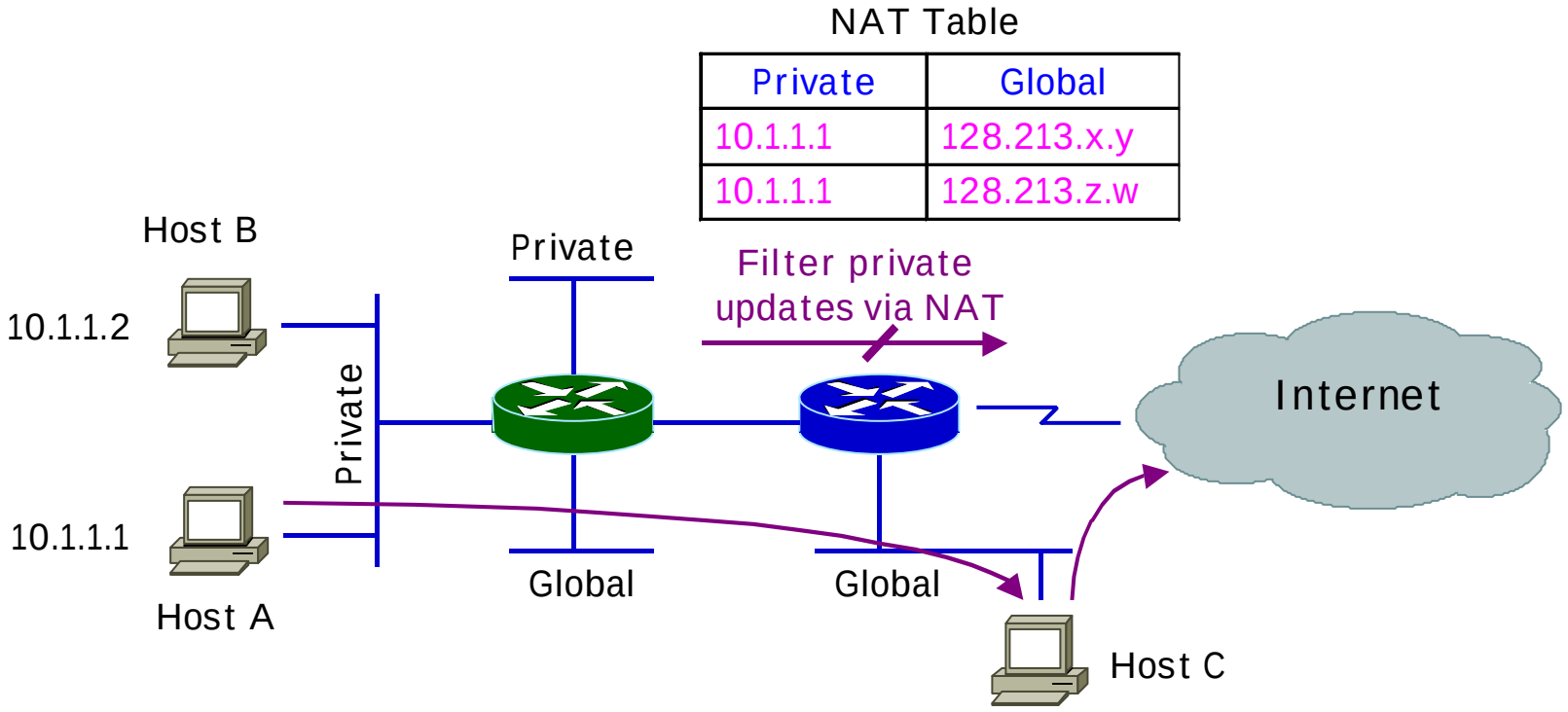


Private Addressing and NAT

- Global connectivity
  - » Globally unique IP address that are recognized inside and outside the organization.
  - » hosts inside an organization have direct access to both internal hosts and Internet hosts.
- Private connectivity
  - » hosts inside an organization have direct access to internal hosts only, not Internet hosts.
    - – Examples: ATM machines, cash registers, etc.
  - » Proxy gateway required to access Internet host.
    - – Companies have the responsibility of applying routing filters to prevent the private networks from being leaked to the Internet.
  - » Advantages:
    - – fewer unique global IP addresses required
    - – security for internal hosts
  - » Disadvantages:
    - – renumbering of all host required if later decide to join globally; problem may be mitigated by using DHCP (Dynamic Host Configuration Protocol)

Private Addressing and NAT

- Private Internet address space allocated
  - » 10.0.0.0 ~ 10.255.255.255 (a single class A network number)
  - » 172.16.0.0 ~ 172.31.255.255 (16 contiguous class B network number)
  - » 192.168.0.0 ~ 192.168.255.255 (256 contiguous class C network number)
- Telnet requires a two step procedure
- NAT has to dynamically map private IP addresses to a smaller pool of global addresses



Primary Class of Routing Protocols

- Link-state protocols
  - » ISIS, OSPF
  - » Information sent describes the state of the link attached to the router
    - – IP network or subnets assigned
    - – cost of the links (inversely proportional to bandwidth)
  - » Advantages:
    - – greater scalability
      - • no hop count
      - • supports VLSM and CIDR
      - • hierarchy allows aggregation
    - – faster convergence
      - • immediate flooding of link states to the domain
      - • less overhead from route exchanges
    - – load balancing via link costs
  - » Disadvantages:
    - – difficult to configure and troubleshoot.
- Distance Vector protocol
  - » RIP, IGRP
  - » Information sent
    - – all the networks that the router knows about, not just the networks it is connected to.
    - – Advertising router increases the metric (typical hop-count) of the router it has heard about. (maximum hop-count 15)
  - » Advantages:
    - – simple configuration
  - » Disadvantages:
    - – cannot take into account link speeds
    - – slower convergence
    - – large routing table exchange and maintenance.
    - – hold-down/flush timers slows convergence to minutes
    - – temporary routing loops
    - – less scalable, lacks hierarchy
    - – classful, lack VLSM or CIDR features

Choosing routing protocols

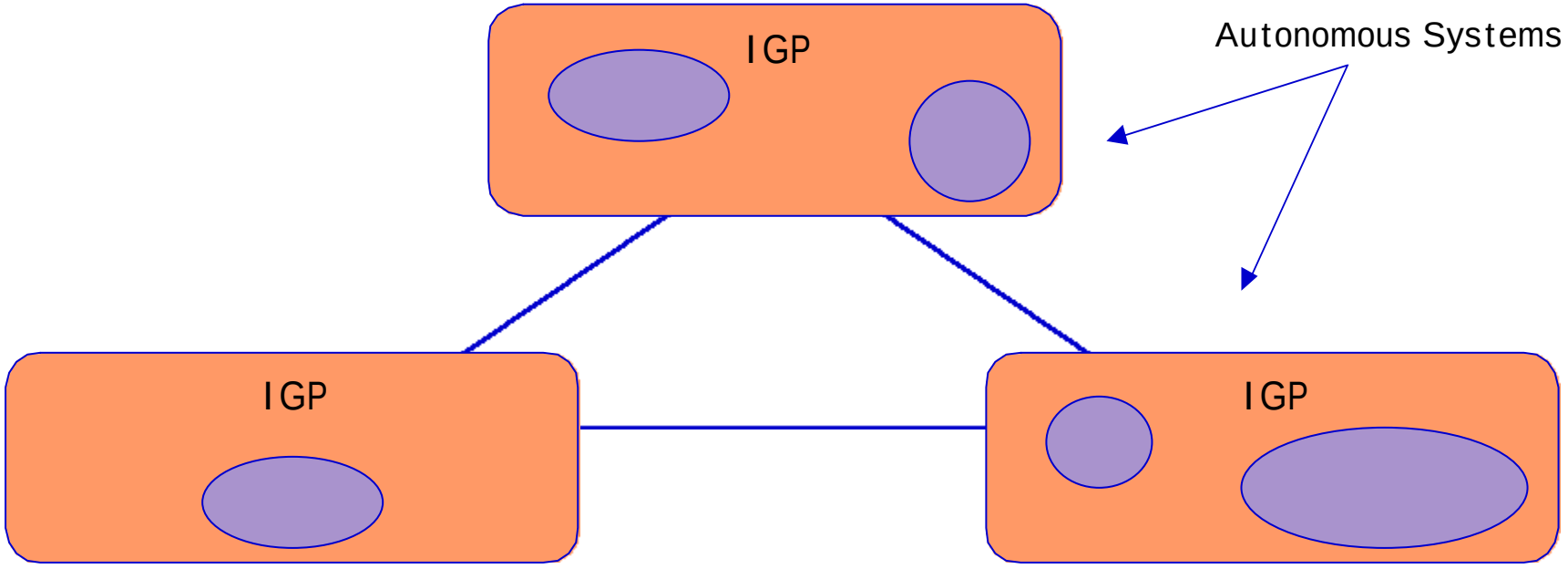
- RIPv1
  - » Networks with no redundant links or parallel paths and no requirements for VLSM (variable length subnet masking) or discontiguous major networks
- RIPv2
  - » Network that requires VLSM and/or discontiguous major networks and does not have redundant links or parallel paths
- IGRP
  - » Network that has redundant links or parallel paths and does not require VLSM or discontiguous major networks
- OSPF, EIGRP, ISIS
  - » Network with parallel paths and/or redundant links and that requires VLSM or discontiguous major networks

Routing methodology

- Static routing
  - » routes to destination are listed manually
  - » network reachability is not dependent on the state of the network;
    - – i.e. whether a destination is up or down, the static routes would remain in the routing table, and traffic will still be sent to that destination.
- Default routing
  - » “last resort” outlet for destination traffic that are unknown to the router.
  - » Easiest form of routing for a domain connected to a single exit point.
- Dynamic routing
  - » routes are learned via internal or external routing protocol
  - » network reachability is dependent on the existence and state of the network.
    - – i.e. if a destination is down, the route would disappear from the routing table and traffic would not be sent toward that destination.
- Most networks use a combination of all three routing methodology

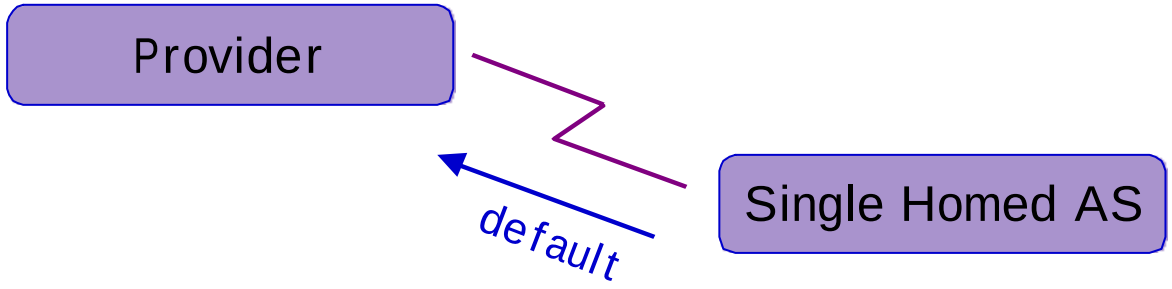
Autonomous System (AS)

- AS is a set of routers having a single routing policy, running under a single technical administration.
  - » A collection of IGPs working together to provide interior routing
  - » Each AS has an identifying number, assigned by an Internet Registry.
  - » Routing information between ASs is exchanged via an exterior gateway protocol such as BGPv4.



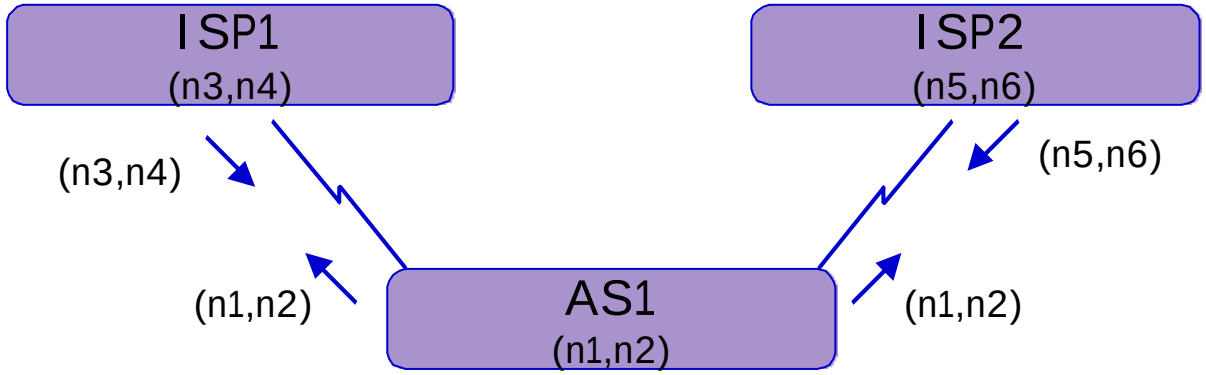
Stub AS

- Reaches the network outside its domain via a single exit point.
  - » All traffic defaults to the provider, therefore don't have to learn Internet routes from the provider
  - » Still need to advertise routes to the providers
    - – via static entries -- good for small set of aggregate routes
    - – via IGP protocol -- stub AS have difficult time getting registered AS numbers because the routing policies tend to follow the provider and therefore don't require unique AS number.
    - – via BGP protocol -- Provider will give the customer a private AS number (65412 ~ 65535)



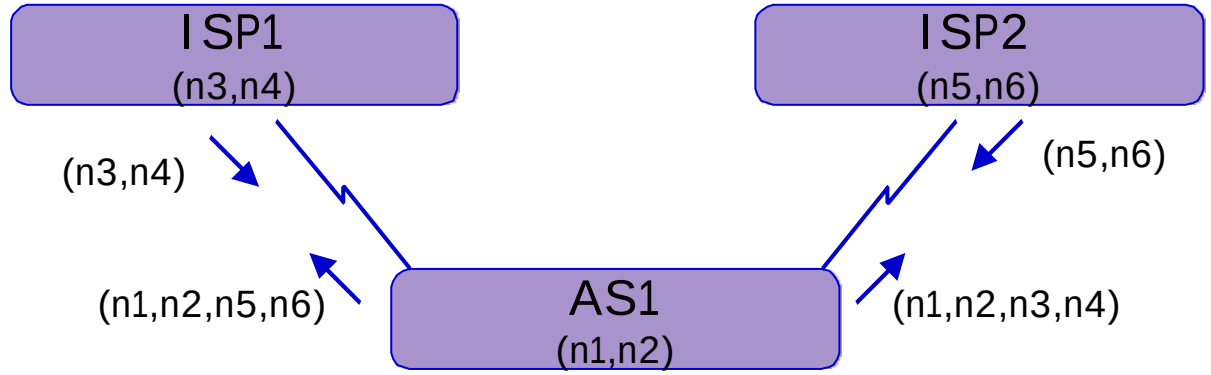
Multihomed Nontransit AS

- Multihomed
  - » more than one exit point to the outside world
- Nontransit
  - » does not allow transit traffic to go through it (IP source /destination address outside of AS)
  - » will not advertise routes it learned from other AS
  - » transit traffic can still be forced via default or static routing
  - » does not require BGP4



Multihomed Transit AS

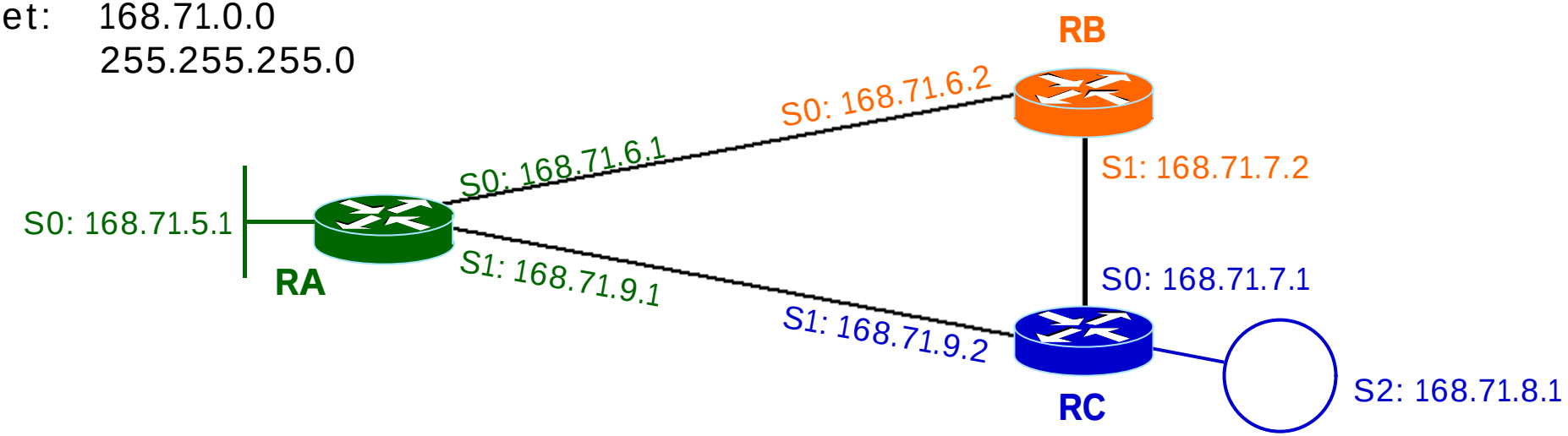
- Multihomed
  - » more than one exit point to the outside world
- Transit
  - » Uses BGP inside an AS as a pipe to exchange BGP updates (IBGP)
  - » Routers running IBGP are called transit routers
  - » Routers running EBGP are called border routers
  - » Will advertise routes it learned from other AS



ITRI
CCL

Router Configuration

Major Net: 168.71.0.0
Mask: 255.255.255.0



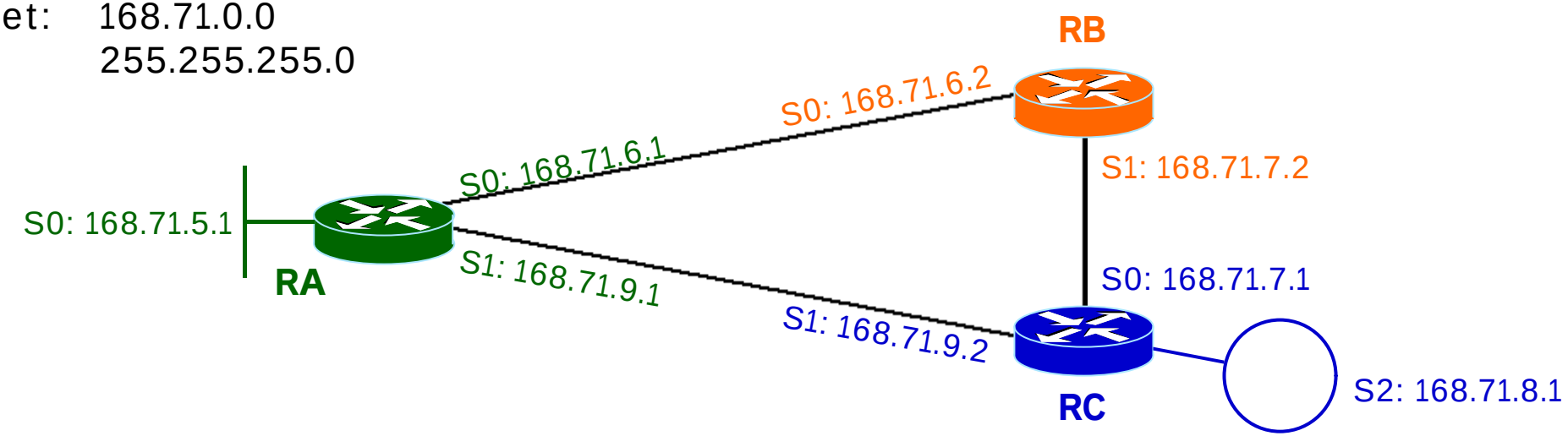
- Router A's Configuration
  - » interface ethernet0
  - » ip address 168.71.5.1/255.255.255.0
  - » ip address 171.68.207.164/255.255.255.128 secondary
  - » interface serial0
  - » ip address 168.71.6.1/255.255.255.0
  - » bandwidth 128
  - » interface serial1
  - » ip address 168.71.9.1/255.255.255.0
  - » bandwidth 128

- Router A's Configuration
  - » router rip
  - » network 168.71.0
  - » passive-interface Ethernet0
    (prevents RIP from advertising routing updates out Ethernet0)
  - » ip route 0.0.0.0 0.0.0.0 Ethernet0

ITRI
CCL

Router Configuration

Major Net: 168.71.0.0
Mask: 255.255.255.0

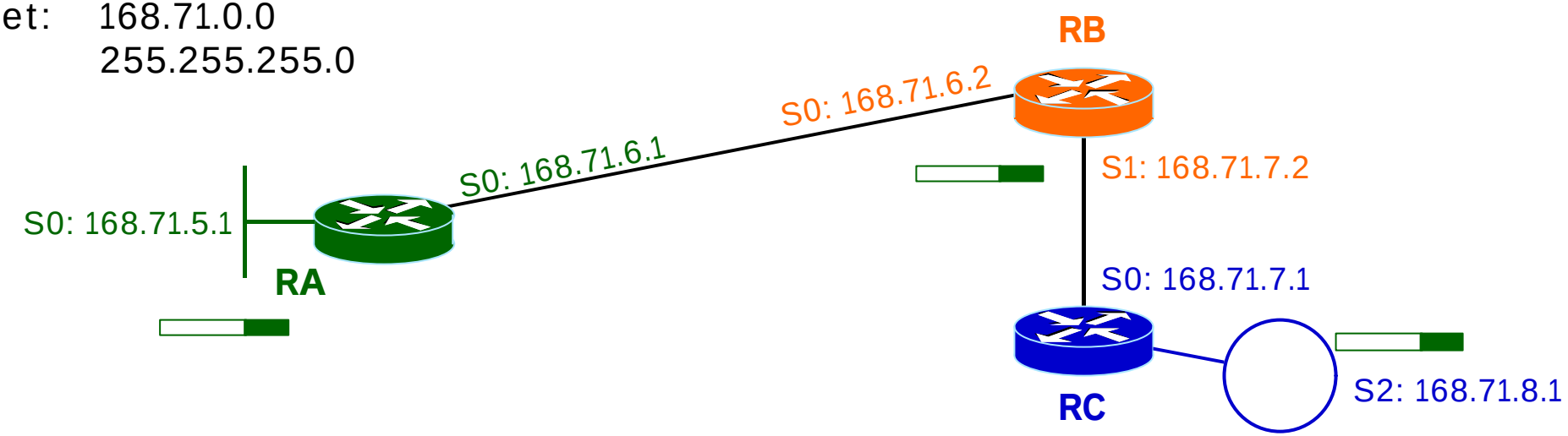


- Router B's Configuration
  - interface serial0
  - ip address 168.71.6.2/255.255.255.0
  - bandwidth 128
  - interface serial1
  - ip address 168.71.7.2/255.255.255.0
  - bandwidth 64
  - router rip
  - network 168.71.0.0

- Router C's Configuration
  - interface tokenring0
  - ip address 168.71.8.1/255.255.255.0
  - interface serial0
  - ip address 168.71.7.1/255.255.255.0
  - bandwidth 64
  - interface serial1
  - ip address 168.71.9.2/255.255.255.0
  - router rip
  - network 168.71.0.0

What a Router does

Major Net: 168.71.0.0
Mask: 255.255.255.0

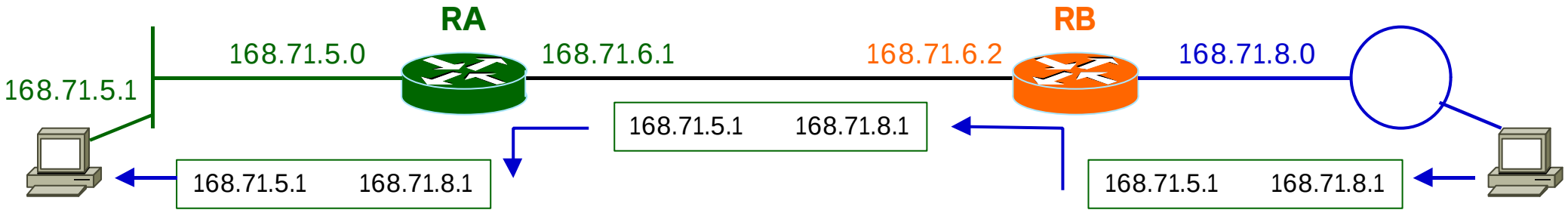


- Router A forwards a packet with destination address 168.71.8.2/255.255.255.0 to Router B
  - » does not know nor care if 168.71.8.2 is directly attached to Router B
  - » router terminology: network prefix 168.71.8.0 is *known via* 168.71.6.2, the advertising router.

ITRI
CCL

How forwarding decisions are made

Major Net: 168.71.0.0
Mask: 255.255.255.0



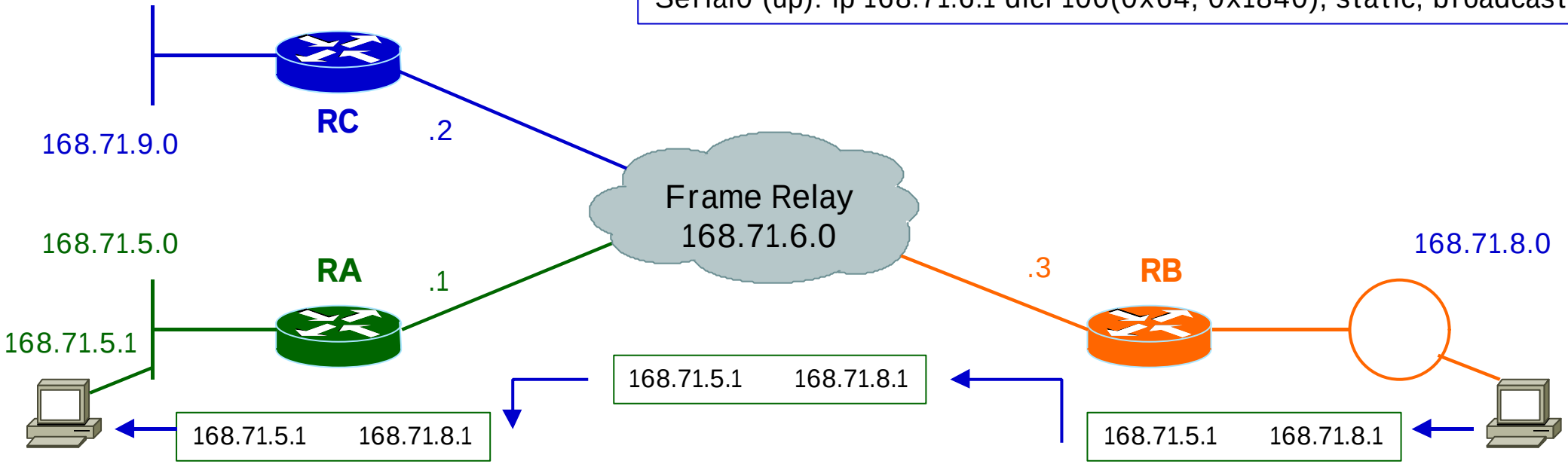
Protocol	Network	Interface	Next Hop
IGRP	168.71.5.0	Serial0	168.71.6.1
Connected	168.71.6.0	Serial0	
Connected	168.71.8.0	Token0	

Forwarding decisions for Multi-point Interfaces

Major Net: 168.71.0.0
Mask: 255.255.255.0

ARP entry required

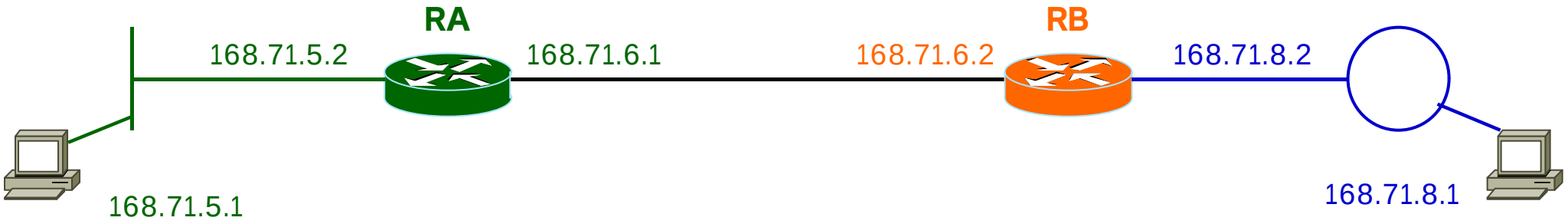
RouterB# show frame-relay map
Serial0 (up): ip 168.71.6.1 dlci 100(0x64, 0x1840), static, broadcast



Protocol	Network	Interface	Next Hop
IGRP	168.71.5.0	Serial0	168.71.6.1
Connected	168.71.6.0	Serial0	
Connected	168.71.8.0	Token0	
IGRP	168.71.9.0	Serial0	168.71.6.2

ITRI CCL End systems sending Packets to other subnets

Major Net: 168.71.0.0
Mask: 255.255.255.0



0000.0C00.1234 ↔ 0000.0C01.8793

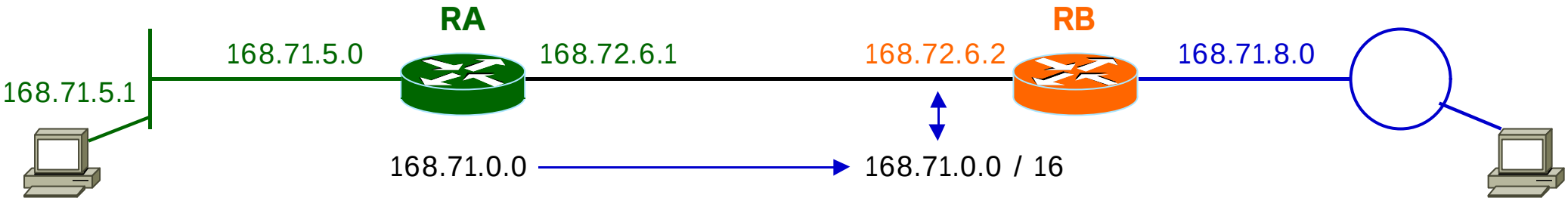
Layer 2

0000.0F00.8674 ↔ 0000.0D00.5463

168.71.5.1 ↔ Layer 3 ↔ 168.71.8.1

ITRI
CCL

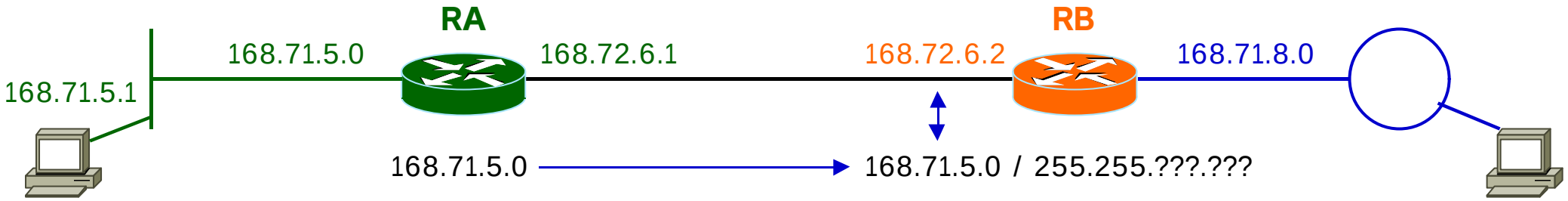
More Routing Terminology



- Major Net
  - » refers to the network-only portion of an Internet address
- Discontiguous network
  - » IP subnet address from a major network are applied to physical networks that are separated by a network with a different major net.
- Summarize
  - » To report only the major net portion of an address, even though there are subnets of the address in the routing table.
- Aggregation
  - » To reduce the number of subnets in a routing advertisement by advertising only the common portion of the subnet address that is required to provide full IP connectivity.

ITRI
CCL

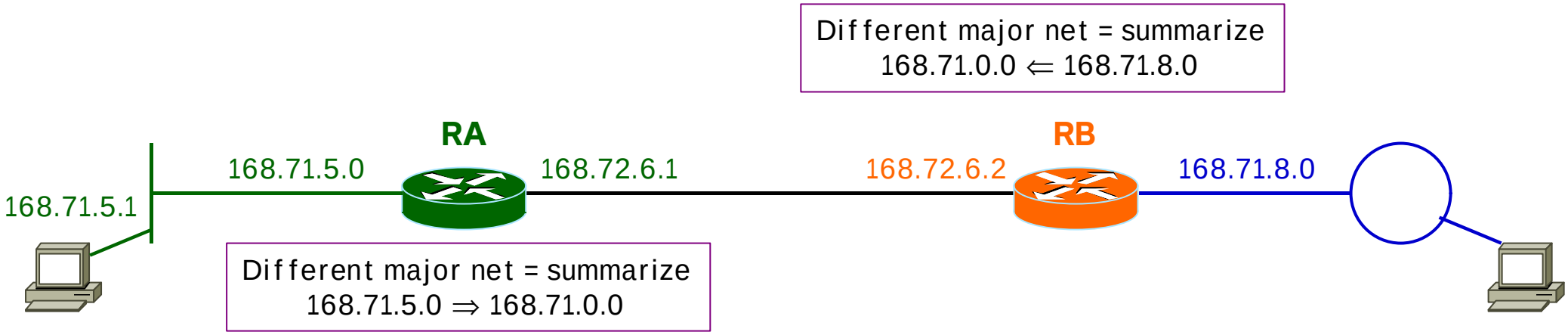
Deriving network masks



- Two methods / rules
  - » Advertisements for subnets that are part of the same major networks as the IP address of the interface they are receiving over will have the subnet mask of the receiving interface applied to them
  - » All other networks in the advertisement should be for major networks, not subnets (thru summarization)
- Example: incorrect advertising subnet over a major net boundary
  - » 1. Ignore route advertisement
  - » 2. Install it with a mask of 255.255.255.255, indicating that it is a host route
  - » 3. Guess the mask size.

ITRI
CCL

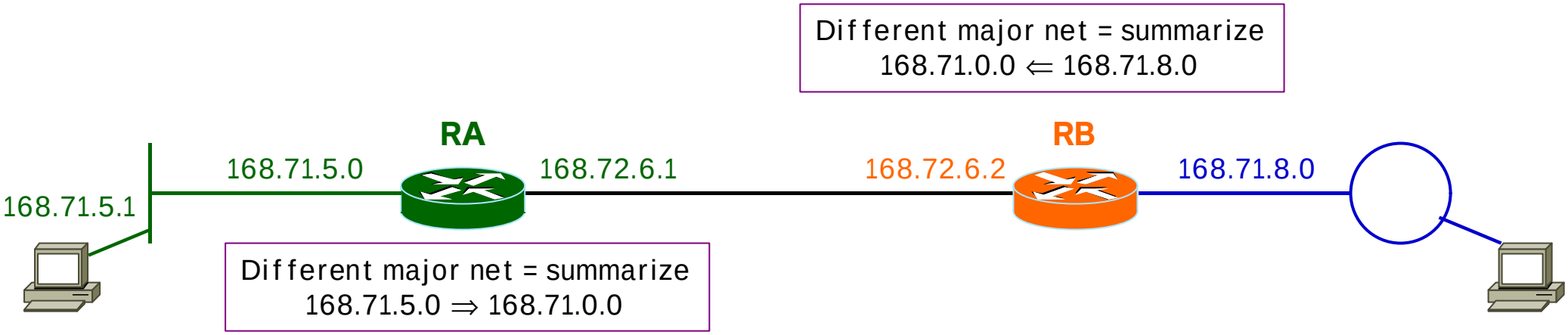
Understanding Summarization



- Summarization rules
  - » If a network or subnet in the table is part of the same network address space (major net) and has the same mask as the outbound interface, it is advertised over the interface.
  - » If a network or subnet is part of the same major net as the outbound interface but has a different mask, it is not advertised.
  - » If a network or subnet is part of a different address space, the outbound interface mask is ignored and only the summarized major network portion of the route is advertised.

ITRI
CCL

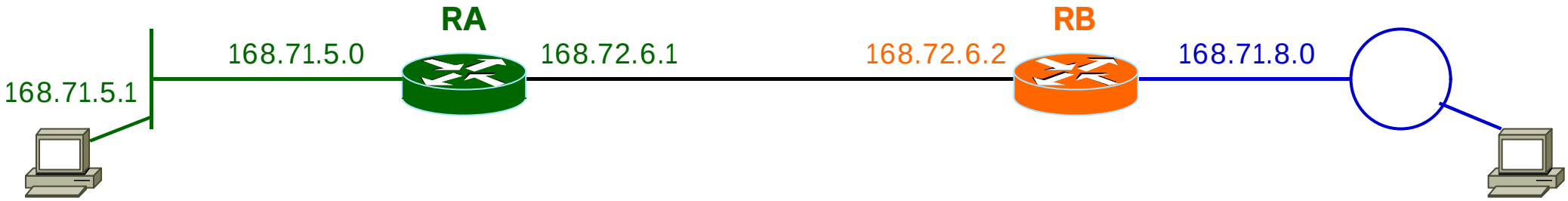
Understanding Subnet 0



- Subnet 0 refers to the use of 0s in the subnet portion of an address covered by the subnet mask.

	Subnet 0	
Network Address	168.71.0.0	**10101000.01000111.00000000.00000000**
Subnet 0 mask	255.255.255.0	**11111111.11111111.11111111.00000000**
	Summarized Route	
Network Address	168.71.0.0	**10101000.01000111.00000000.00000000**
Summarized mask	255.255.0.0	**11111111.11111111.00000000.00000000**

Summarization caused by discontiguous network



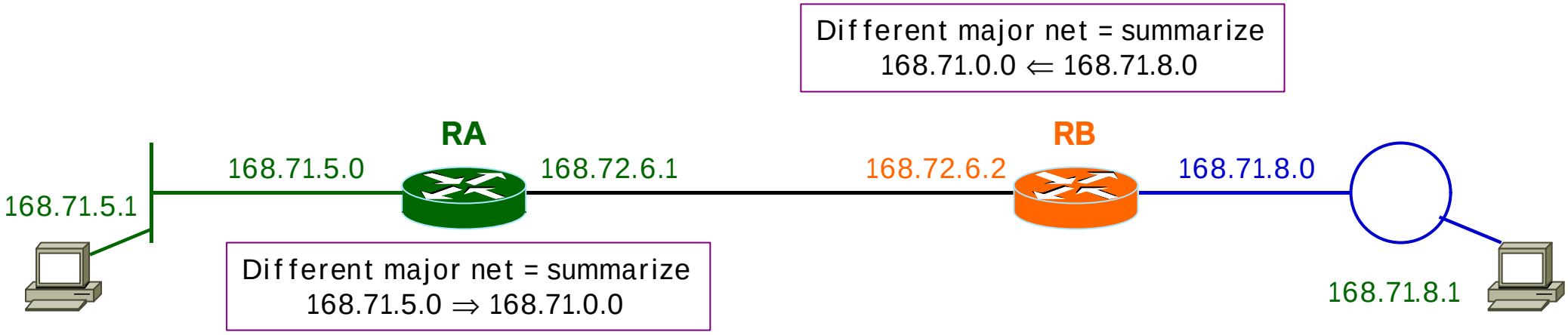

```
RouterA# conf t
RouterA#-config# interface s0
RouterA #-config-if# ip address 168.72.6.1
                                255.255.255.0
RouterA# show interface serial 0
Serial0 is up, line protocol is up
Internet address is 168.72.6.1/24
MTU 1500 bytes, BW 128 kbit, DLY 20000
    usec, rely 255/255, load 1/255
```

```
RouterB# conf t
RouterB#-config# interface s0
RouterB #-config-if# ip address 168.72.6.2
                                255.255.255.0
RouterB# show interface serial 0
Serial0 is up, line protocol is up
Internet address is 168.72.6.2/24
MTU 1500 bytes, BW 128 kbit, DLY 20000
    usec, rely 255/255, load 1/255
```

```
Router rip
network 168.71.0.0
network 168.72.0.0
```

ITRI
CCL

RIP cannot reach discontiguous subnets



- Routers running RIP ignores summarized routes for networks to which they have connections; therefore RIP doesn't work with discontiguous networks.
  - » Checks network number and mask with its interface before sending out a packet
  - » RA does not have a route installed to 168.71.0.0 even though RB is advertising it; vice versa with RB
  - » therefore RA can't ping 168.71.0.1 and RB can't ping 168.71.5.1