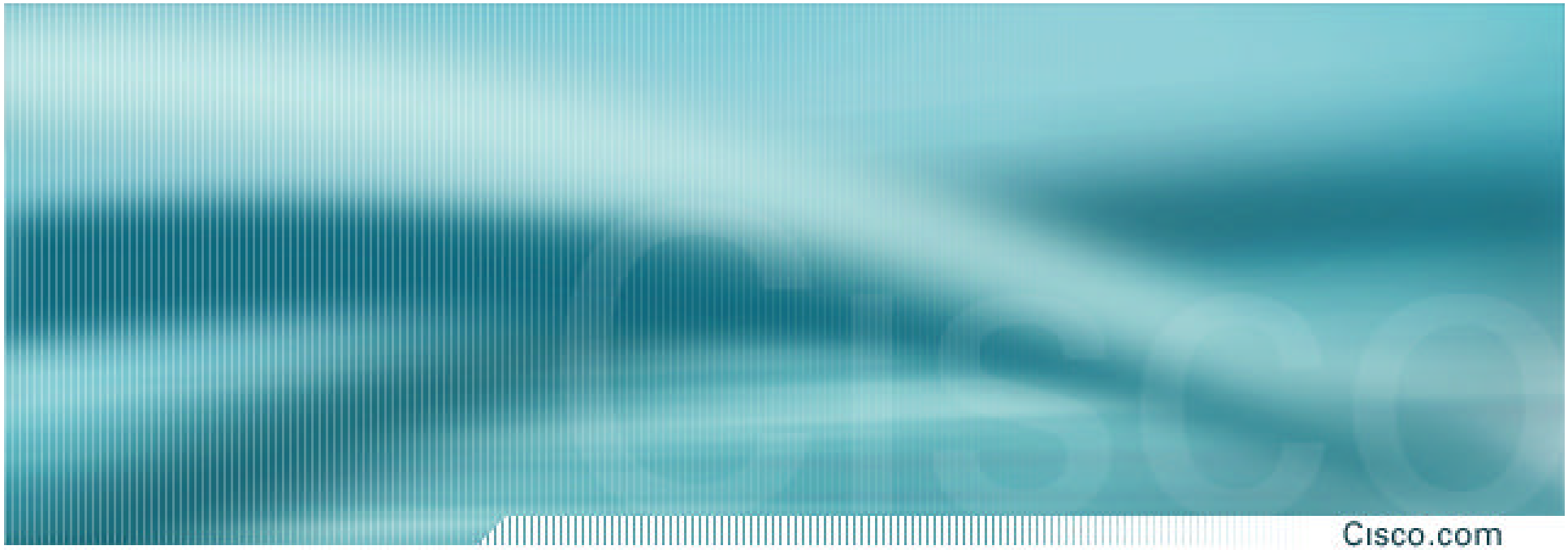


Introduction to Routing

How traffic flows on the Internet

Philip Smith pfs@cisco.com

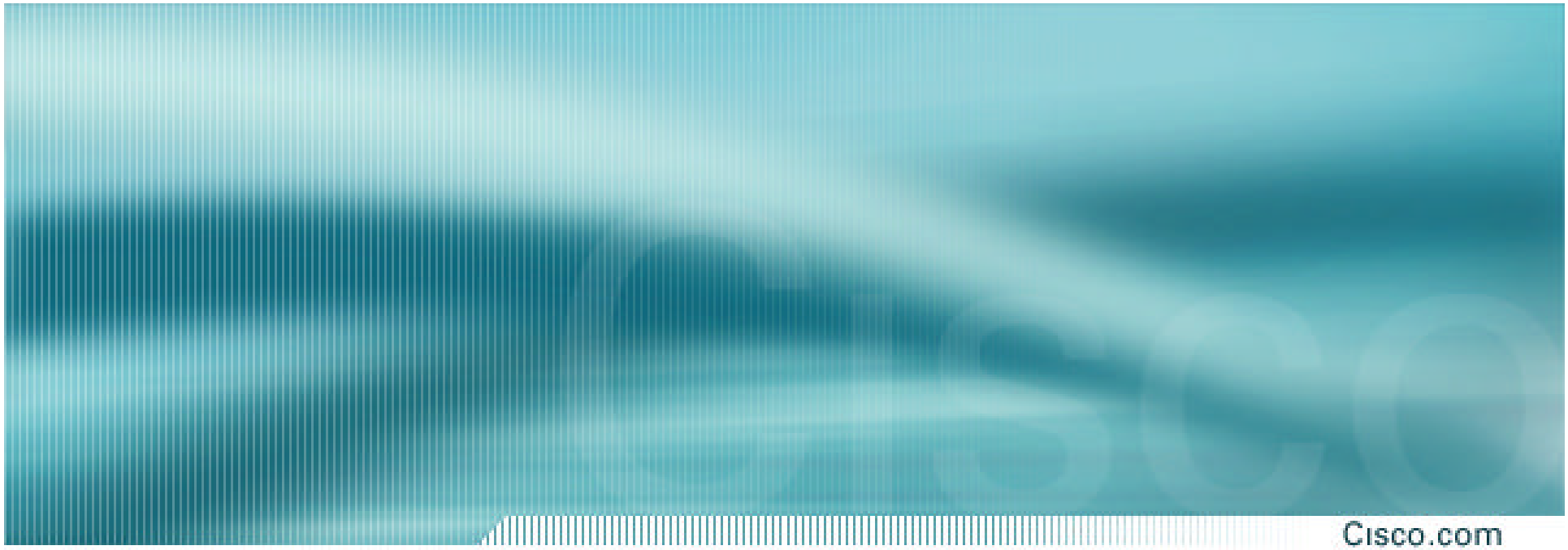
RIPE NCC Regional Meeting,

Dubai, December 2003

Agenda

Cisco.com

- Topologies and Definitions
- Routing and How it Works
- BGP
- Aggregation
- Summary



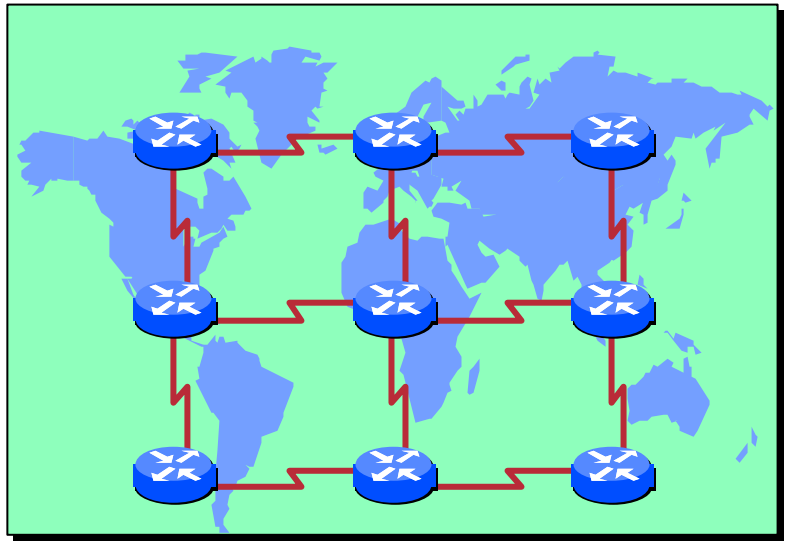
Topologies and Definitions

What does all the jargon mean?

Network Topologies

Routed backbone

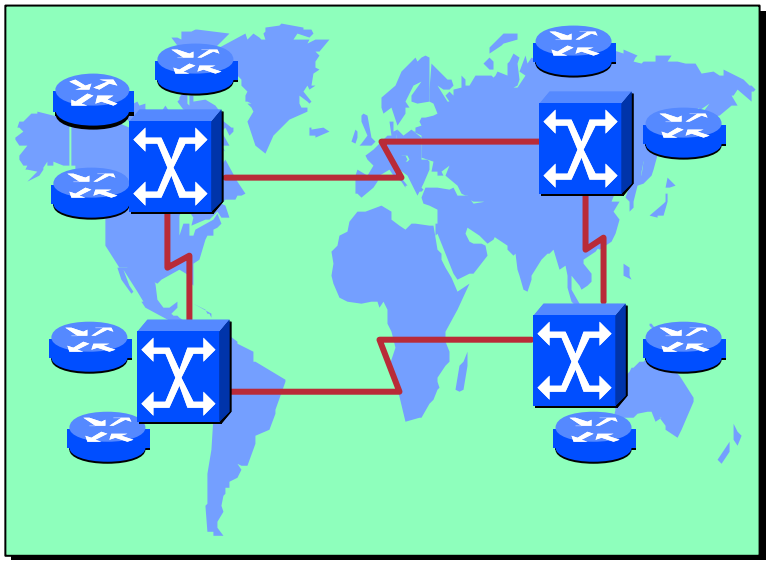
- Routers are the infrastructure
- Physical circuits run between routers
- Easy routing configuration, operation and troubleshooting



Network Topologies

Switched backbone

- frame relay or ATM
  - switches in the core
  - surrounded by routers
- Physical circuits run between switches
  - Virtual circuits run between routers
- more complex routing and debugging
- “traffic management”



Definitions

- **PoP – Point of Presence**
  - **Physical location of ISP's equipment**
  - **Sometimes called a "node"**
- **vPoP – virtual PoP**
  - **To the end user, it looks like an ISP location**
  - **In reality a back hauled access point**
  - **Used mainly for consumer access networks**
- **Hub/SuperPoP – large central PoP**
  - **Links to many PoPs**

PoP Topologies

- **Core routers**

  high speed trunk connections

- **Distribution routers**

  higher port density, aggregating network edge to the network core

- **Access routers**

  high port density, connecting the end users to the network

- **Border routers**

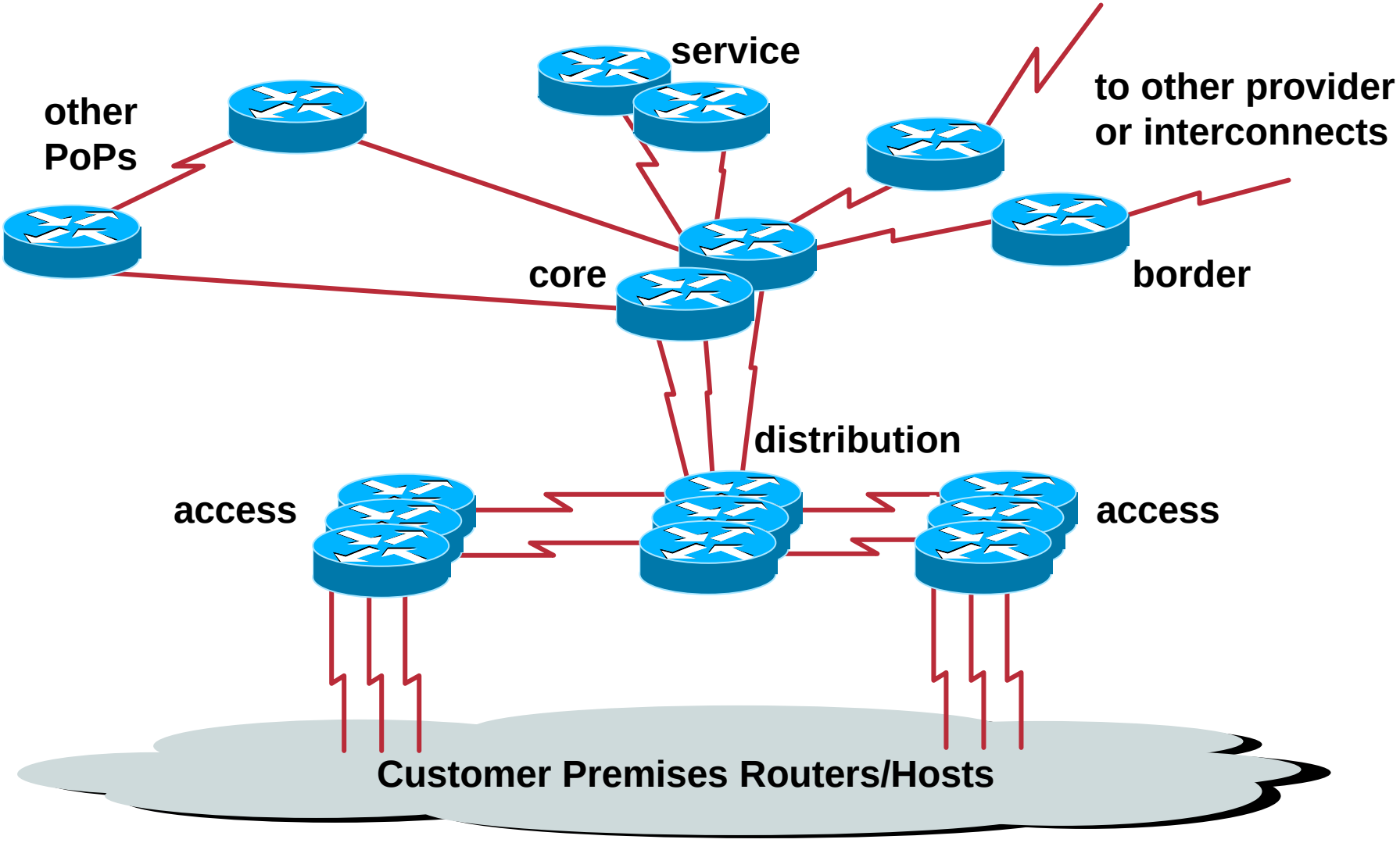
  connections to other providers

- **Service routers**

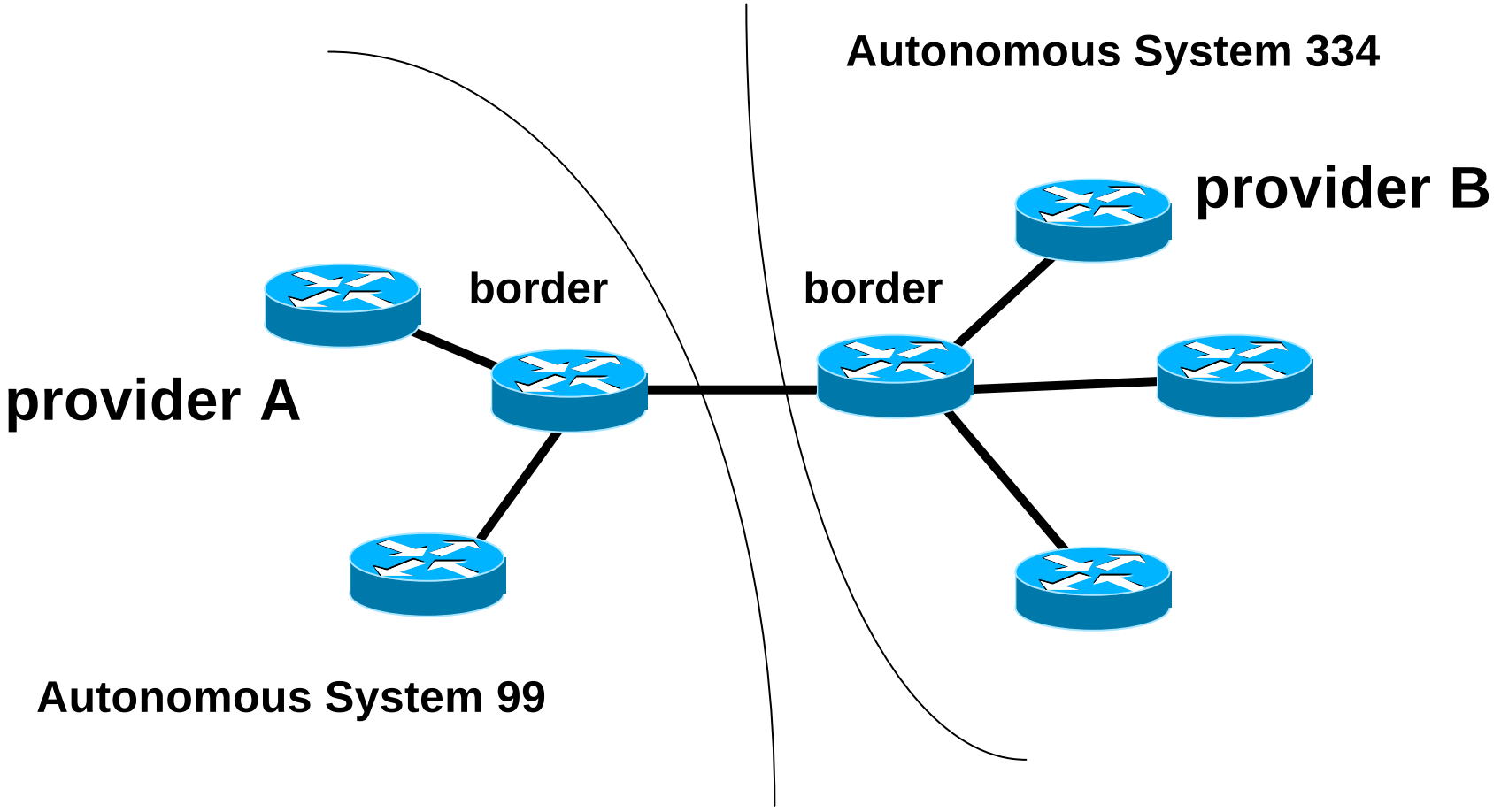
  hosting and servers

- **Some functions might be handled by a single router**

PoP Topologies

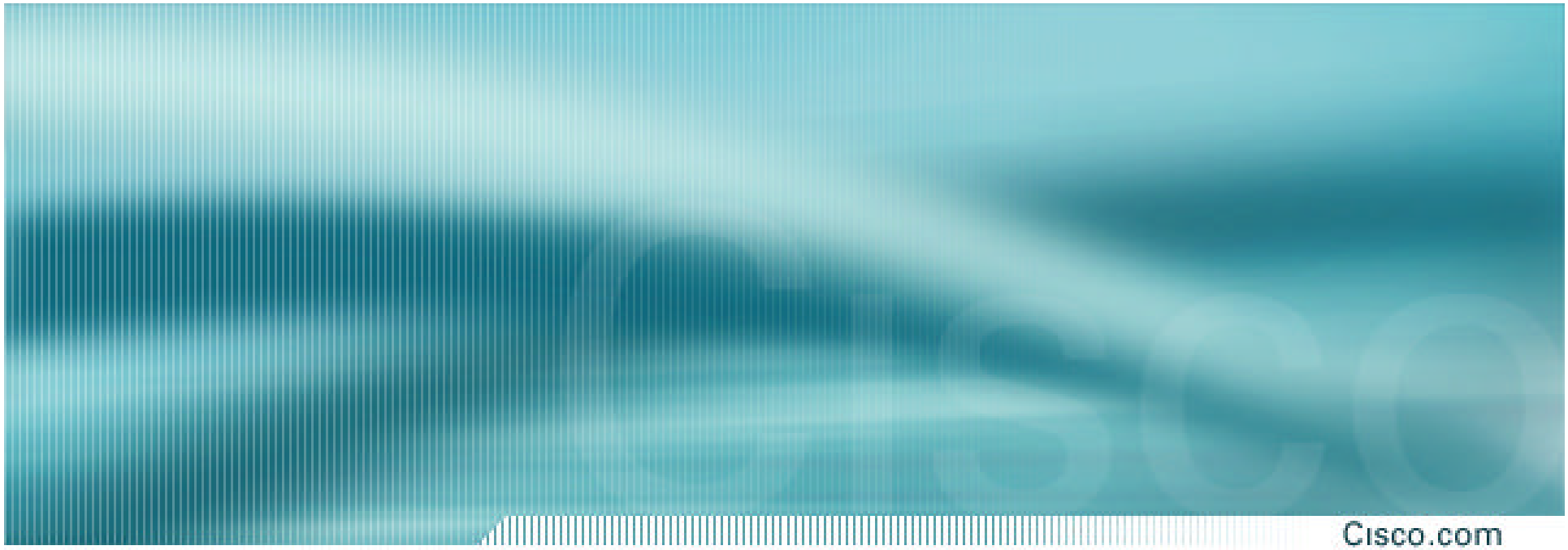


Private Interconnect



Public Interconnect Point

- **IXP – Internet eXchange Point**
- **NAP – Network Access Point**
- **Location or facility where several ISPs are present and connect to each other over a common shared media**



Routing Basics

Now we understand the high level, how does it all work…?

How Does Routing Work?

- Internet is made up of the ISPs who connect to each other's networks
- How does an ISP in Kenya tell an ISP in Japan what customers they have?
- And how does that ISP send data packets to the customers of the ISP in Japan, and get responses back
  - After all, as on a local ethernet, two way packet flow is needed for communication between two devices

How Does Routing Work?

- ISP in Kenya could buy a direct connection to the ISP in Japan
  - But this doesn't scale – thousands of ISPs, would need thousands of connections, and cost would be astronomical
- Instead, ISP in Kenya tells his neighbouring ISPs what customers he has
  - And the neighbouring ISPs pass this information on to their neighbours, and so on
  - This process repeats until the information reaches the ISP in Japan

How Does Routing Work?

- This process is called “Routing”
- The mechanisms used are called “Routing Protocols”
- Routing and Routing Protocols ensures that the Internet can scale, that thousands of ISPs can provide connectivity to each other, giving us the Internet we see today

How Does Routing Work?

- **ISP in Kenya doesn't actually tell his neighbouring ISPs the names of the customers**

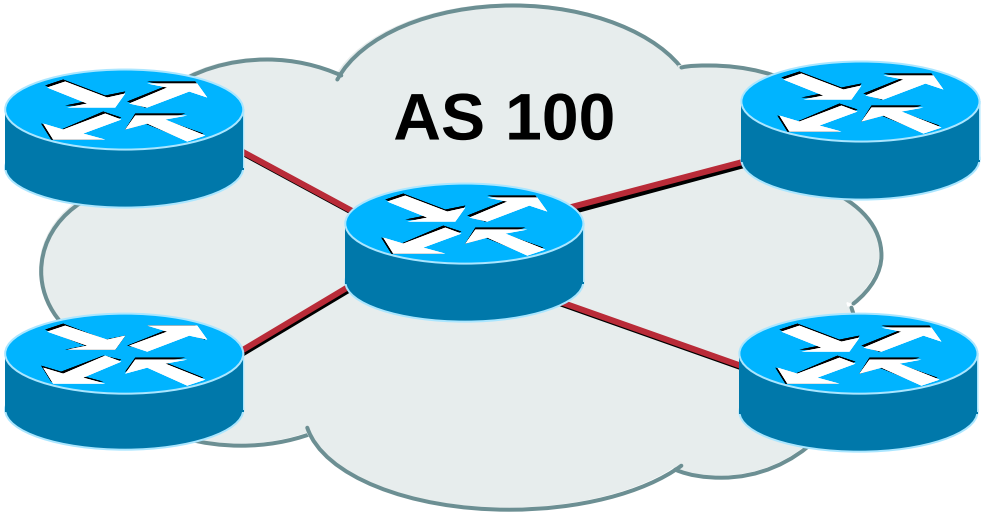
  (network equipment does not understand names)

- **Instead, he has received an IP address block as a member of the Regional Internet Registry serving Kenya**

  He announces this address block to his neighbouring ISPs – this is called "announcing a route"

  His customers have received address space from this address block as part of their "Internet service"

Autonomous System (AS)



- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control
- Identified by 16-bit integer, of which 1-64511 are available for public use

More Definitions

- **Neighbours**

  AS's which directly exchange routing information

  Routers which exchange routing information

- **Announce**

  send routing information to a neighbour

- **Accept**

  receive and use routing information sent by a neighbour

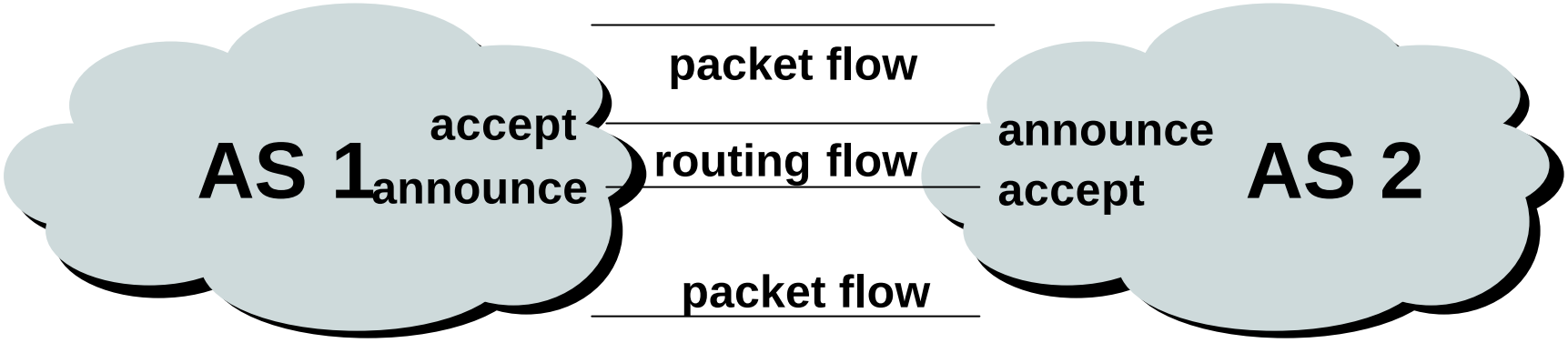
- **Originate**

  insert routing information into external announcements (usually as a result of the IGP)

- **Peers**

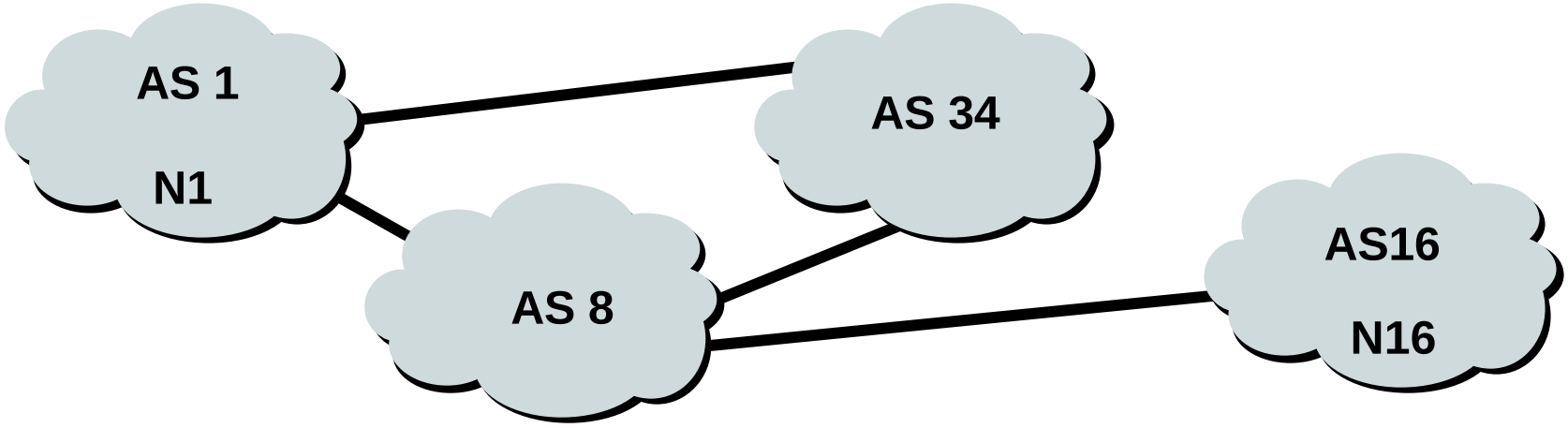
  routers in neighbouring AS's or within one AS which exchange routing and policy information

Routing flow and Packet flow



- For networks in AS1 and AS2 to communicate:
  - AS1 must announce to AS2
  - AS2 must accept from AS1
  - AS2 must announce to AS1
  - AS1 must accept from AS2
- Direction of Traffic flow is always opposite to the direction of the flow of Routing information

Routing Flow/Packet Flow: With multiple ASes

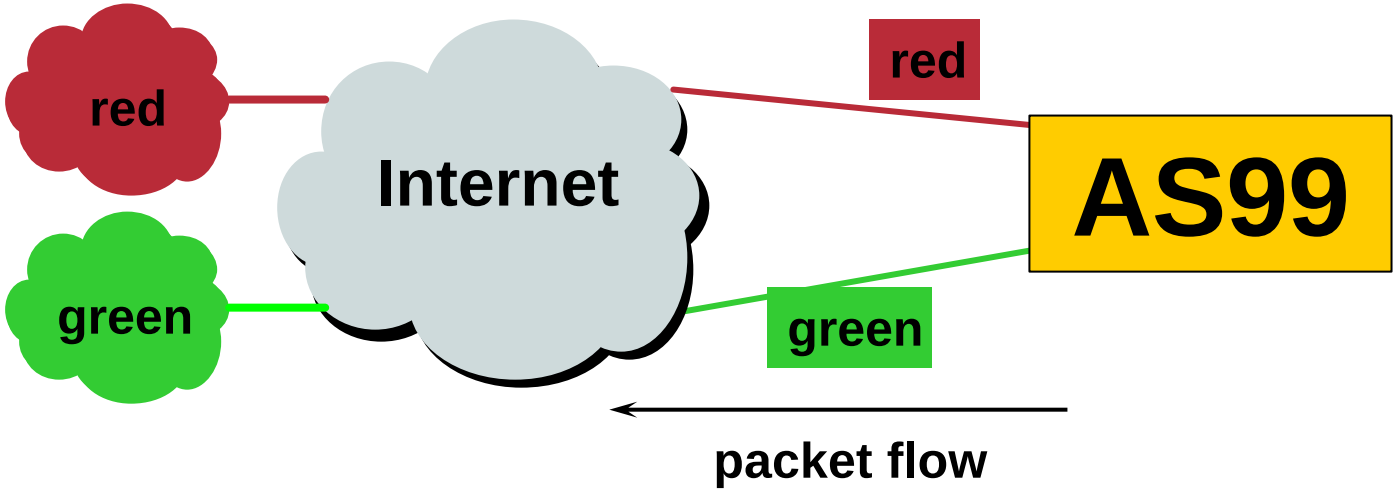


- For net N1 in AS1 to send traffic to net N16 in AS16:
  - AS16 must originate and announce N16 to AS8
  - AS8 must accept N16 from AS16
  - AS8 must announce N16 to AS34 or AS1
  - AS1 must accept N16 from AS8 or AS34
- For two way packet flow, similar policies must exist for N1

Routing Policy

- Used to control traffic flow in and out of an ISP network
- ISP makes decisions on what routing information to accept and discard from its neighbours
  - Individual routes
  - Routes originated by specific ASes
  - Routes traversing specific ASes
  - Routes belonging to other groupings
    - Groupings which you define as you see fit

Routing Policy Limitations

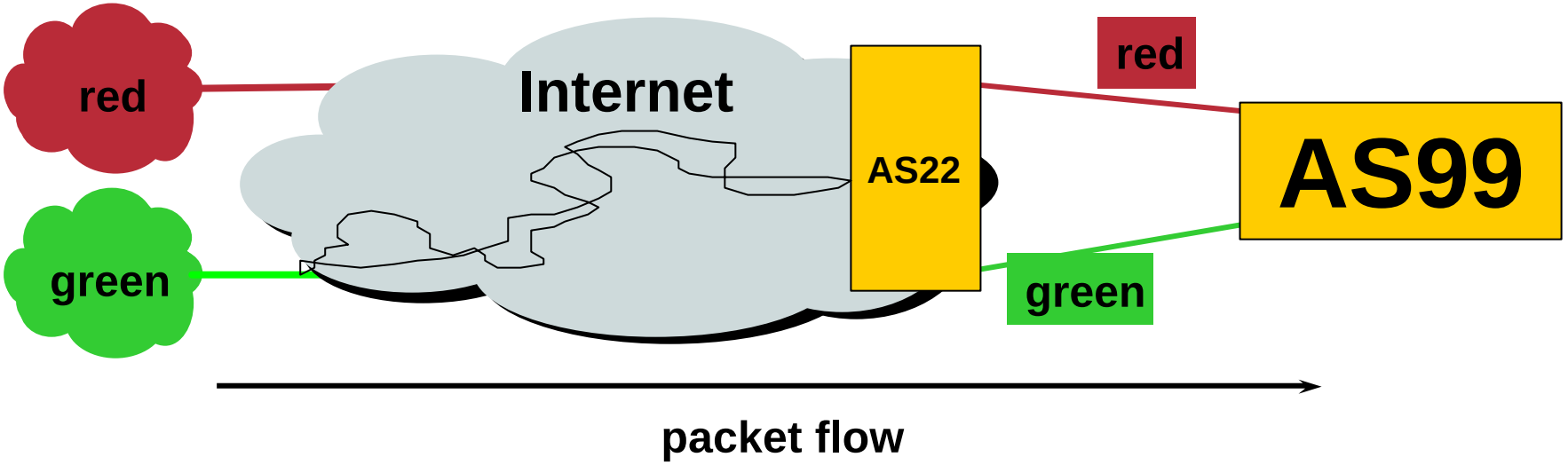


- **AS99 uses red link for traffic to the red AS and the green link for remaining traffic**
- **To implement this policy, AS99 has to:**

  Accept routes originating from the red AS on the red link

  Accept all other routes on the green link

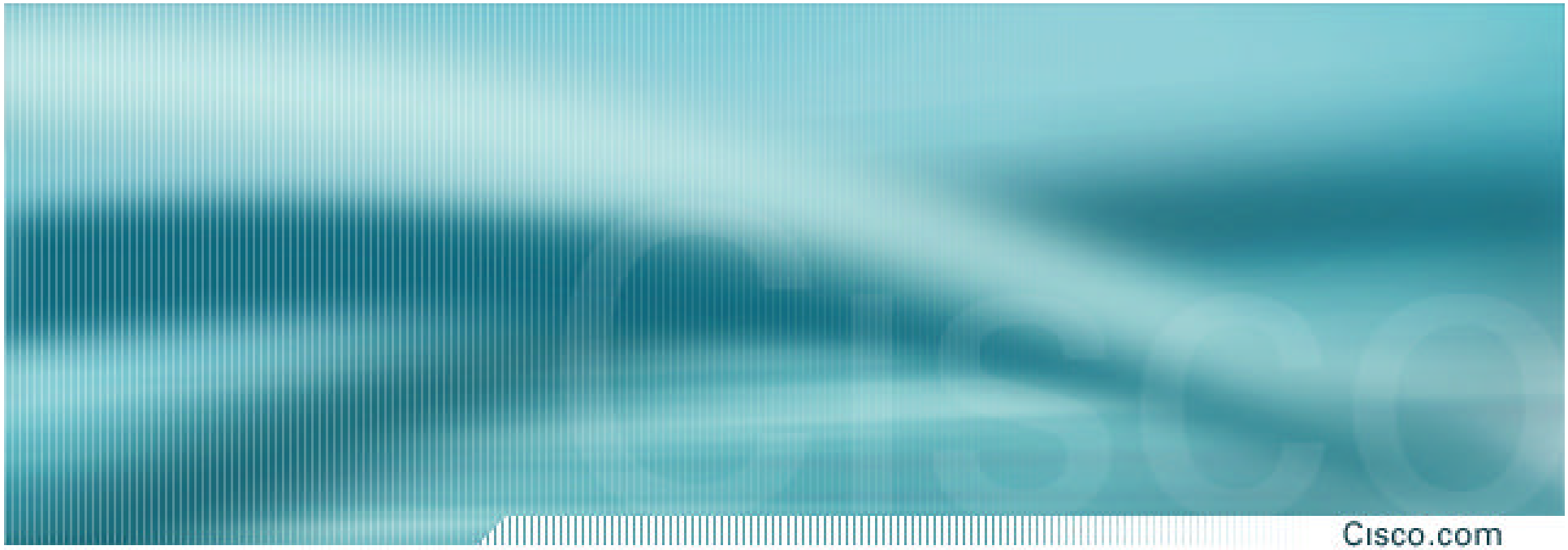
Routing Policy Limitations



- AS99 would like packets coming from the green AS to use the green link.
- But unless AS22 cooperates in pushing traffic from the green AS down the green link, there is very little that AS99 can do to achieve this aim

Routing Policy Limitations

- **In the Internet today:**
  - **130000 prefixes (not realistic to set policy on all of them individually)**
  - **16500 origin AS's (too many)**
  - **Routes tied to a specific AS or path may be unstable regardless of connectivity**
- **Groups of ASes are a natural abstraction for filtering purposes**



Routing Protocols

We now know what routing means…

…but what do the routers get up to?

Routing Protocols

- **Routers use “routing protocols” to exchange routing information with each other**
  - **IGP is used to refer to this process on routers running inside an ISP’s network**
  - **EGP is used to refer to the process running between routers bordering directly connected ISP networks**

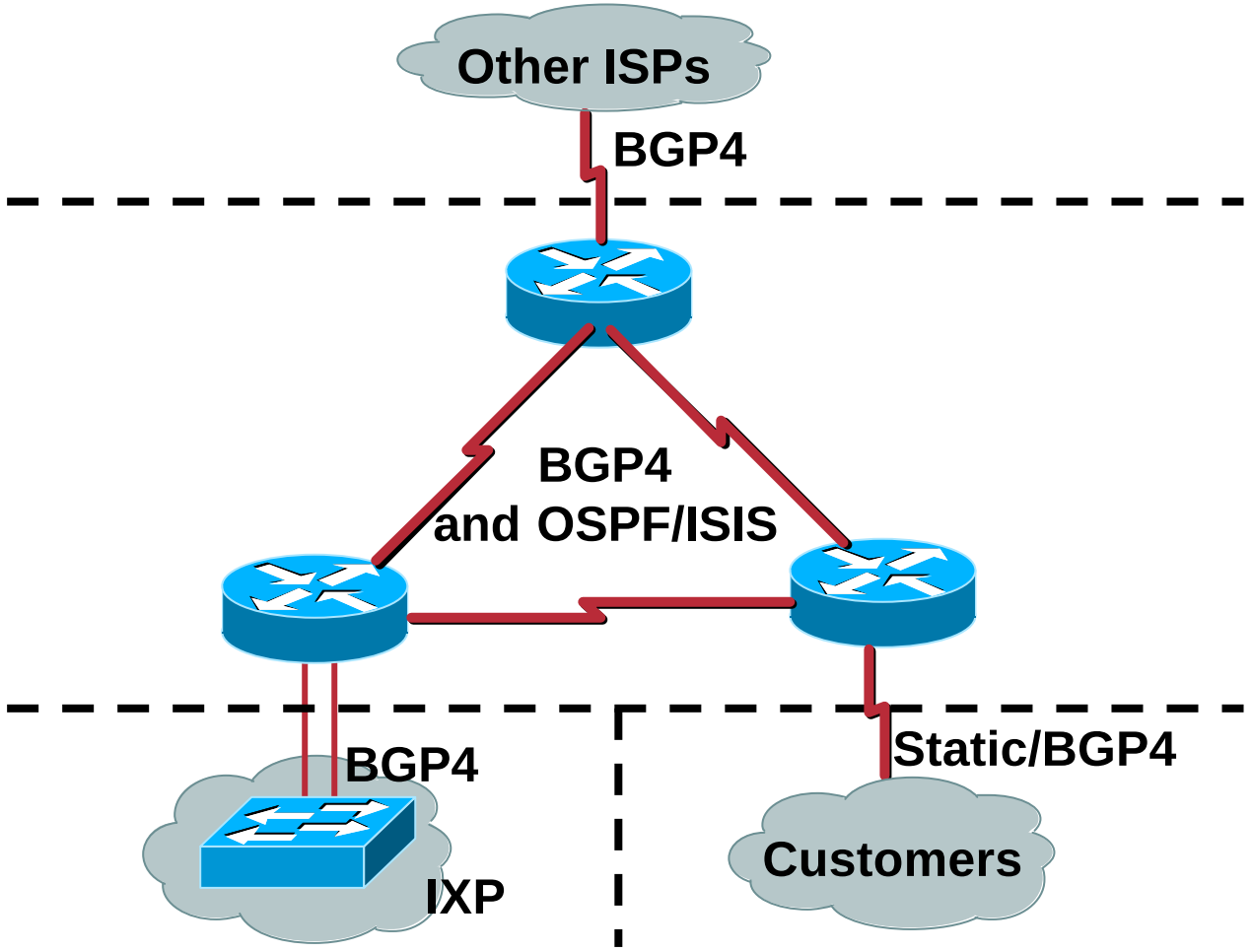
What Is an IGP?

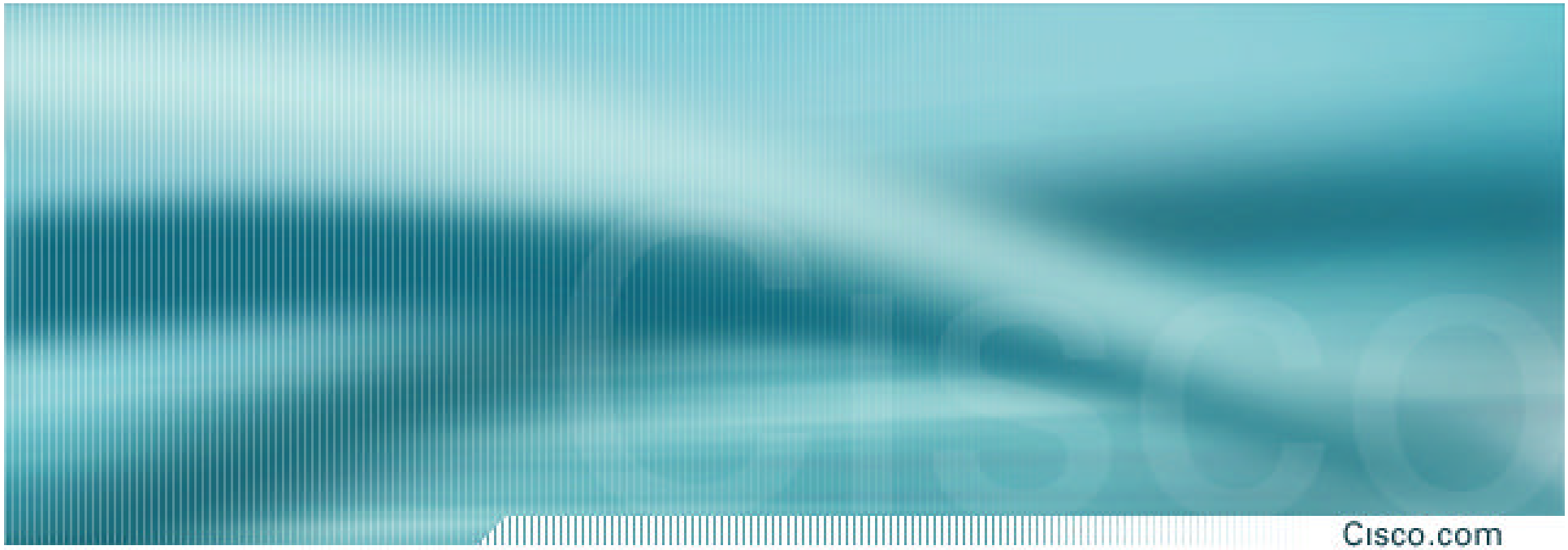
- **Interior Gateway Protocol**
  - **Used within an Autonomous System**
  - **Carries internal infrastructure prefixes**
  - **Examples – OSPF, ISIS, EIGRP…**
- **Needed for scaling the ISP's backbone**
  - **Only used for ISP's infrastructure addresses, not customers**
  - **Design goal is to minimise number of prefixes in IGP to aid scalability and rapid convergence**

What Is an EGP?

- **Exterior Gateway Protocol**
  - **Used to convey routing information between Autonomous Systems**
  - **De-coupled from the IGP**
  - **Current EGP is BGP4**
- **Allows scaling to a large network**
- **Defines administrative boundaries**
- **Used to apply Routing Policy**

Hierarchy of Routing Protocols





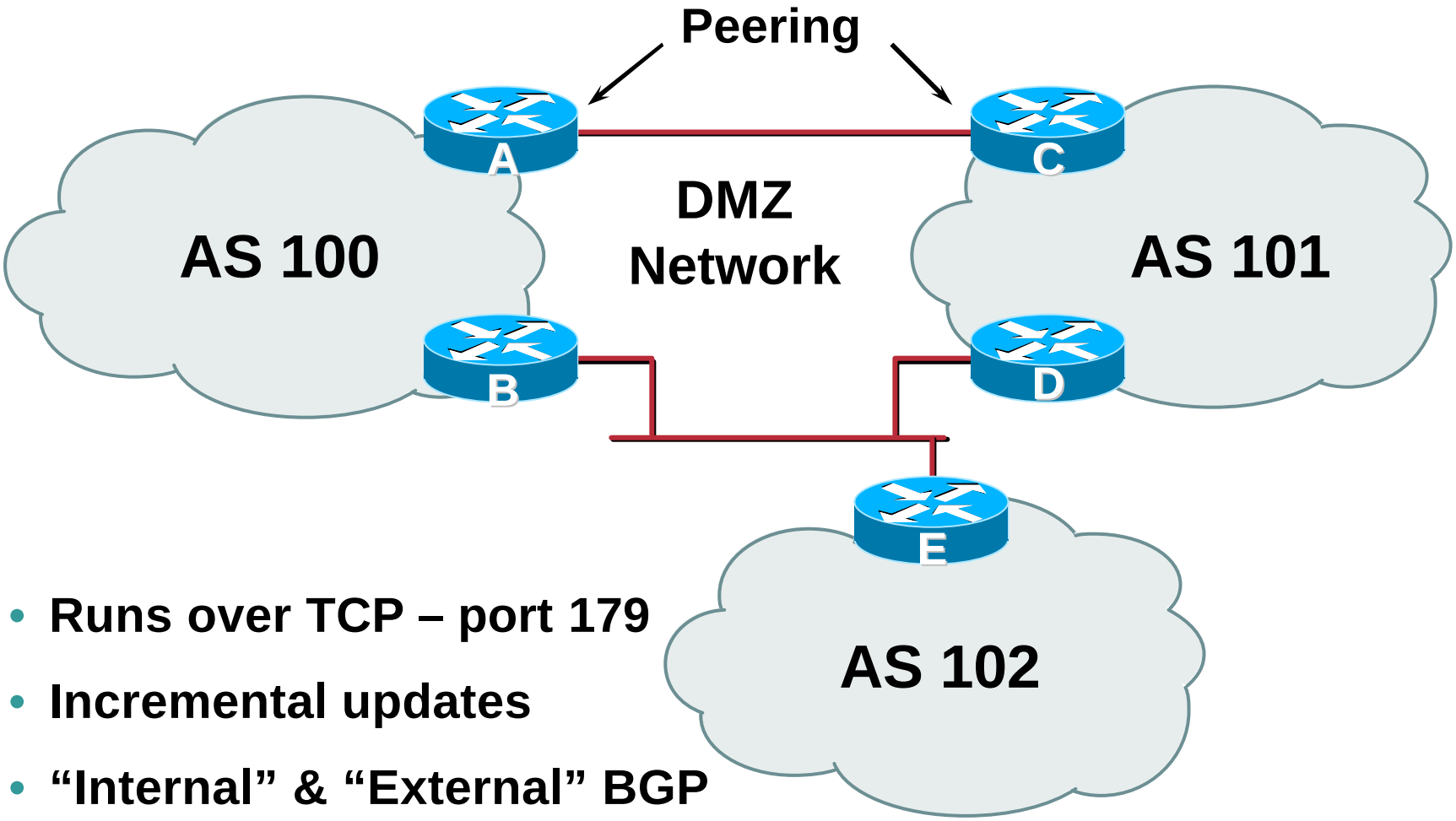
BGP Basics

People (and routers) talk about BGP – what is it?

Border Gateway Protocol

- **Routing Protocol used to exchange routing information between networks**
  - **exterior gateway protocol**
- **Described in RFC1771**
  - **work in progress to update**
  - **www.ietf.org/internet-drafts/draft-ietf-idr-bgp4-23.txt**
- **The Autonomous System is BGP's fundamental operating unit**
  - **It is used to uniquely identify networks with common routing policy**

BGP Basics



- Runs over TCP – port 179
- Incremental updates
- “Internal” & “External” BGP
- DMZ is shared network between ASes

BGP General Operation

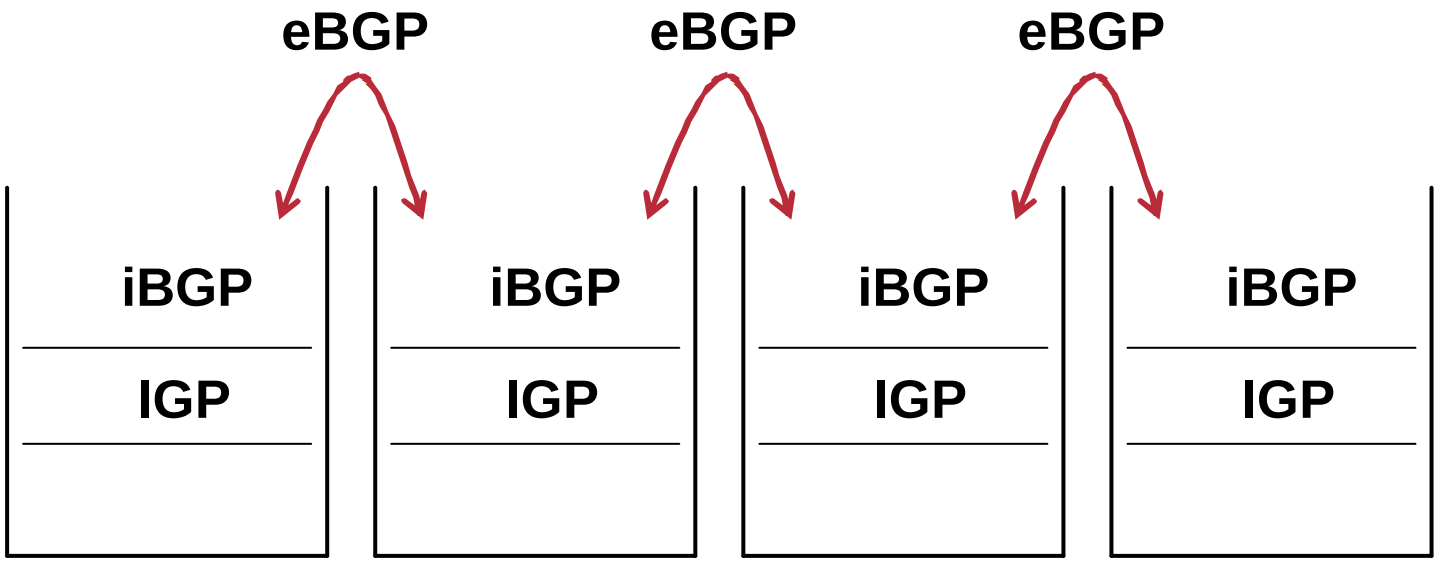
- Learns multiple paths via internal and external BGP speakers
- Picks the best path and installs in the forwarding table
- Best path is sent to external BGP neighbours
- Policies applied by influencing the best path selection

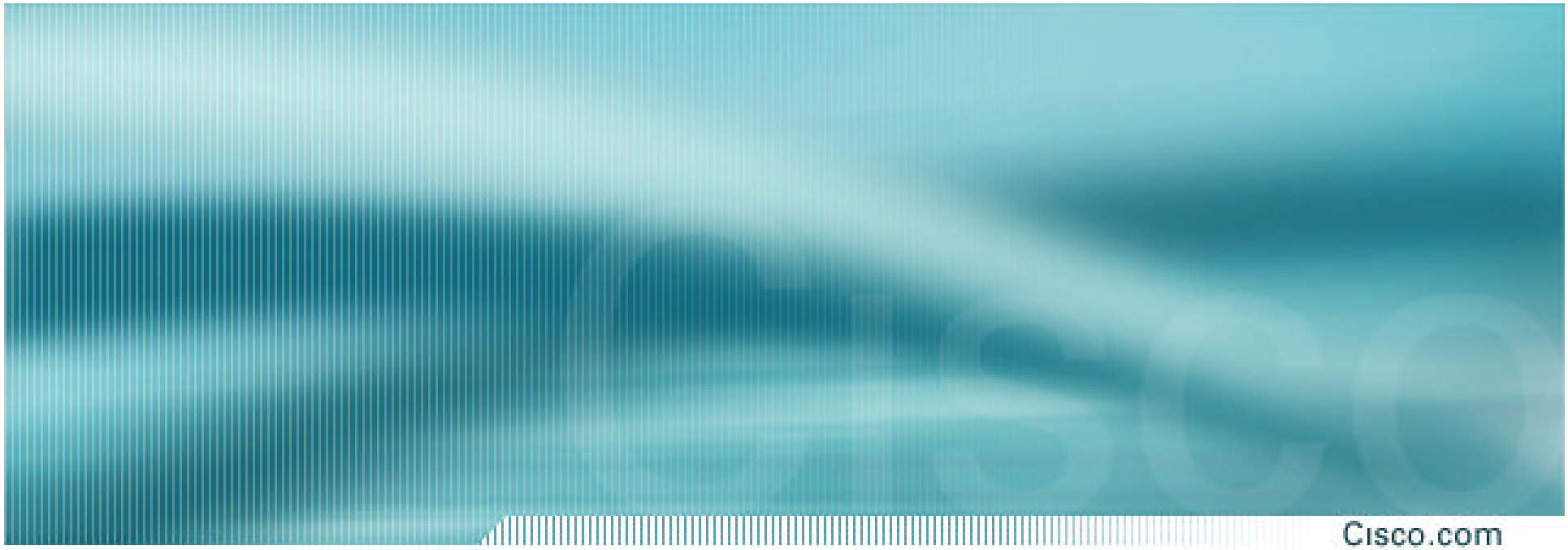
eBGP & iBGP

- BGP used internally (iBGP) and externally (eBGP)
- iBGP used to carry
  - some/all Internet prefixes across ISP backbone
  - ISP's customer prefixes
- eBGP used to
  - exchange prefixes with other ASes
  - implement routing policy

BGP/IGP model used in ISP networks

- **Model representation**





Aggregation

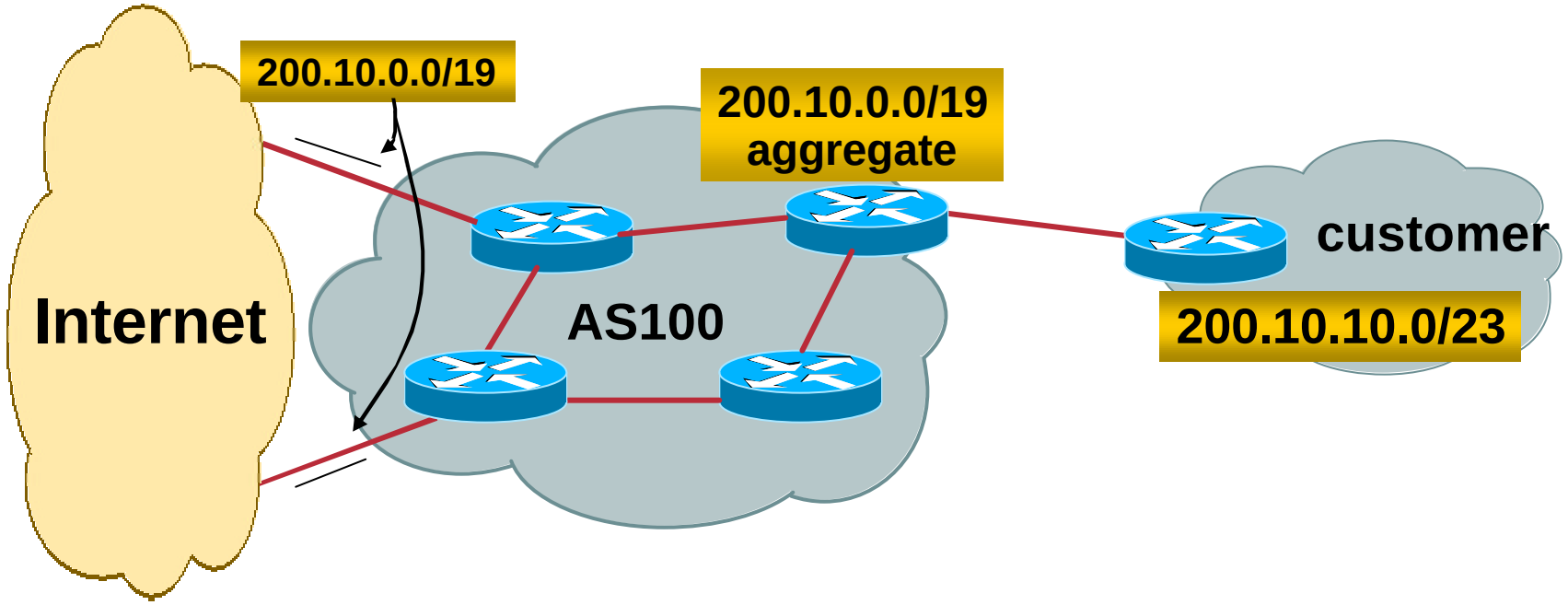
How to announce reachability information to the Internet

"Quality or Quantity?"

Aggregation

- **Aggregation means announcing the address block received from the RIR to the other ASes connected to your network**
- **Subprefixes of this aggregate *may* be:**
  - **Used internally in the ISP network**
  - **Announced to other ASes to aid with multihoming**
- **Unfortunately too many people are still thinking about class Cs, resulting in a proliferation of /24s in the Internet routing table**

Aggregation – Example



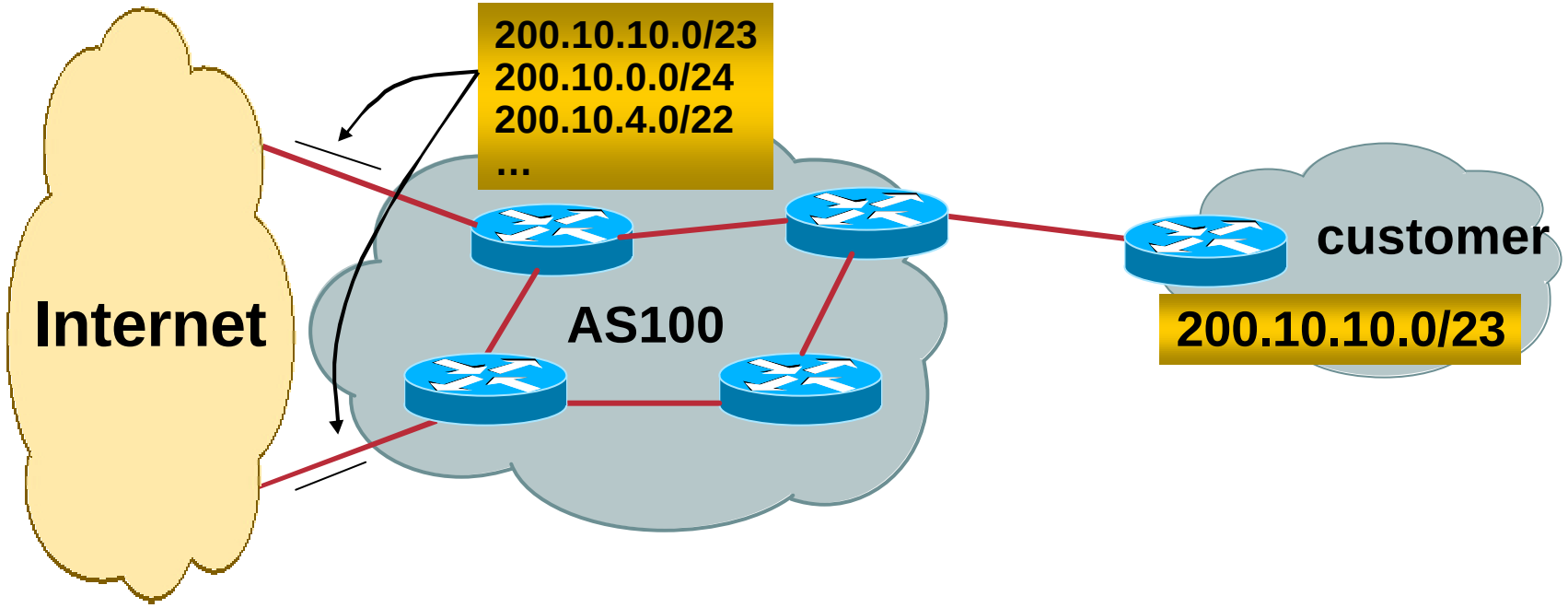
- **Customer has /23 network assigned from AS100's /19 address block**
- **AS100 announced /19 aggregate to the Internet**

Aggregation – Good Example

- Customer link goes down
  - their /23 network becomes unreachable
  - /23 is withdrawn from AS100's iBGP
- /19 aggregate is still being announced
  - no BGP hold down problems
  - no BGP propagation delays
  - no damping by other ISPs

- Customer link returns
- Their /23 network is visible again
  - The /23 is re-injected into AS100's iBGP
- The whole Internet becomes visible immediately
- Customer has Quality of Service perception

Aggregation – Example



- Customer has /23 network assigned from AS100's /19 address block
- AS100 announces customers' individual networks to the Internet

Aggregation – Bad Example

- Customer link goes down
  - Their /23 network becomes unreachable
  - /23 is withdrawn from AS100's iBGP
- Their ISP doesn't aggregate its /19 network block
  - /23 network withdrawal announced to peers
  - starts rippling through the Internet
  - added load on all Internet backbone routers as network is removed from routing table

- Customer link returns
  - Their /23 network is now visible to their ISP
  - Their /23 network is re-advertised to peers
  - Starts rippling through Internet
  - Load on Internet backbone routers as network is reinserted into routing table
  - Some ISP's suppress the flaps
  - Internet may take 10-20 min or longer to be visible
  - Where is the Quality of Service???

Aggregation – Summary

- **Good example is what everyone should do!**
  - **Adds to Internet stability**
  - **Reduces size of routing table**
  - **Reduces routing churn**
  - **Improves Internet QoS for everyone**
- **Bad example is what too many still do!**
  - **Why? Lack of knowledge?**

The Internet Today (December 2003)

- **Current Internet Routing Table Statistics**

BGP Routing Table Entries	130957
Prefixes after maximum aggregation	80657
Unique prefixes in Internet	62723
Prefixes smaller than registry alloc	57878
/24s announced	71544
only 5492 /24s are from 192.0.0.0/8	
ASes in use	16265

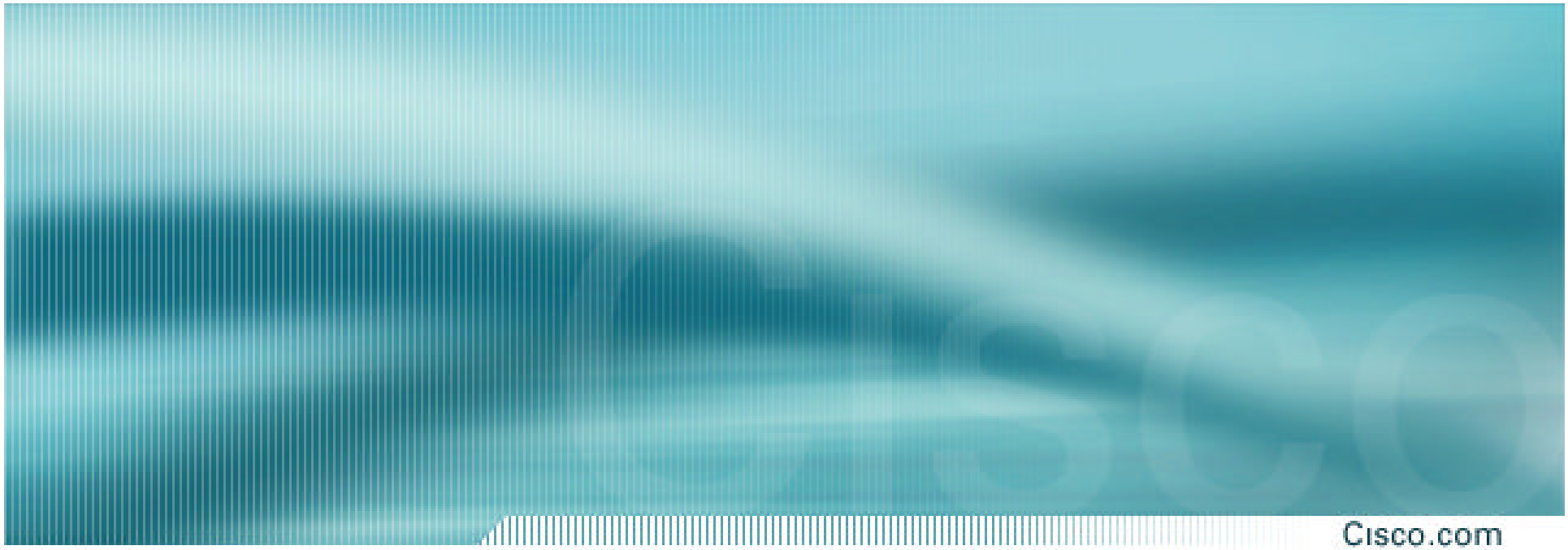
Efforts to improve aggregation

- **The CIDR Report**
  - **Initiated and operated for many years by Tony Bates**
  - **Now combined with Geoff Huston's routing analysis**
  - **www.cidr-report.org**
  - **Results e-mailed on a weekly basis to most operations lists around the world**
  - **Lists the top 30 service providers who could do better at aggregating**

Agenda

Cisco.com

- Topologies and Definitions
- Routing and How it Works
- BGP
- Aggregation
- Summary



Introduction to Routing

How traffic flows on the Internet

Philip Smith pfs@cisco.com

RIPE NCC Regional Meeting,

Dubai, December 2003