# Promoting Routability

## APRICOT'98 Tutorial

17 February 1998, Manila


# Introduction

**Presenters**

Philip Smith - Cisco Systems

Anne Lord - APNIC

- **Please ask questions**
- **Some material won't be covered**




# Today's schedule

- Introduction
- Routing Terms and Concepts
- Overview of the Internet Routing Registry

  *10:30-11:00 coffee break*
- Overview of the Internet Routing Registry (continued)

  *12:30-14:00 lunch*
- Routing Registry Tools for Network Operators

  *15:30-16:00 coffee break*
- Routing Etiquette
- Future of the IRR
- End


# Promoting Routability Goals

- The Internet Routing Registry
- Routing Behaviour - “healthy” Internet
- Awareness
- Understanding
- Participation

# Acknowledgements


- PRIDE (RIPE NCC)
- MERIT
- ISI

- Cisco Systems
- APNIC





# Routing Terms and Concepts




CISCO SYSTEMS

# Overview

- IP packet forwarding
- IP packet routing
- IP routing definitions
- IP routing policies
- Why an Internet Routing Registry?

# IP Packet Forwarding

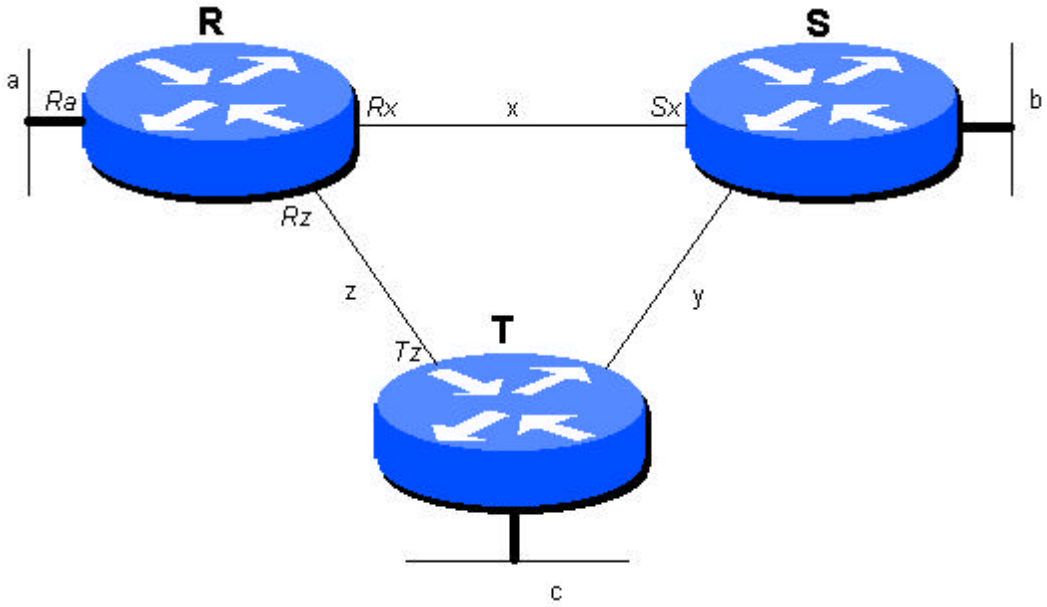
Definition

- Router makes decision on which interface a packet is sent to

Features

- Uses forwarding table
- Forwarding decisions
  - Destination address
  - Class of Service (fair queuing, committed access rate)
  - Local requirements (filtering)
- Aided by special hardware (e.g. Cisco Express Forwarding)

# IP Packet Forwarding




**Forwarding Table on Router R**

| Destination Network | Interface | Next-Hop |
|---|---|---|
| a | Ra | - |
| x | Rx | - |
| z | Rz | - |
| b | Rx | Sx |
| c | Rz | Tz |
| z | Tz | Tz |

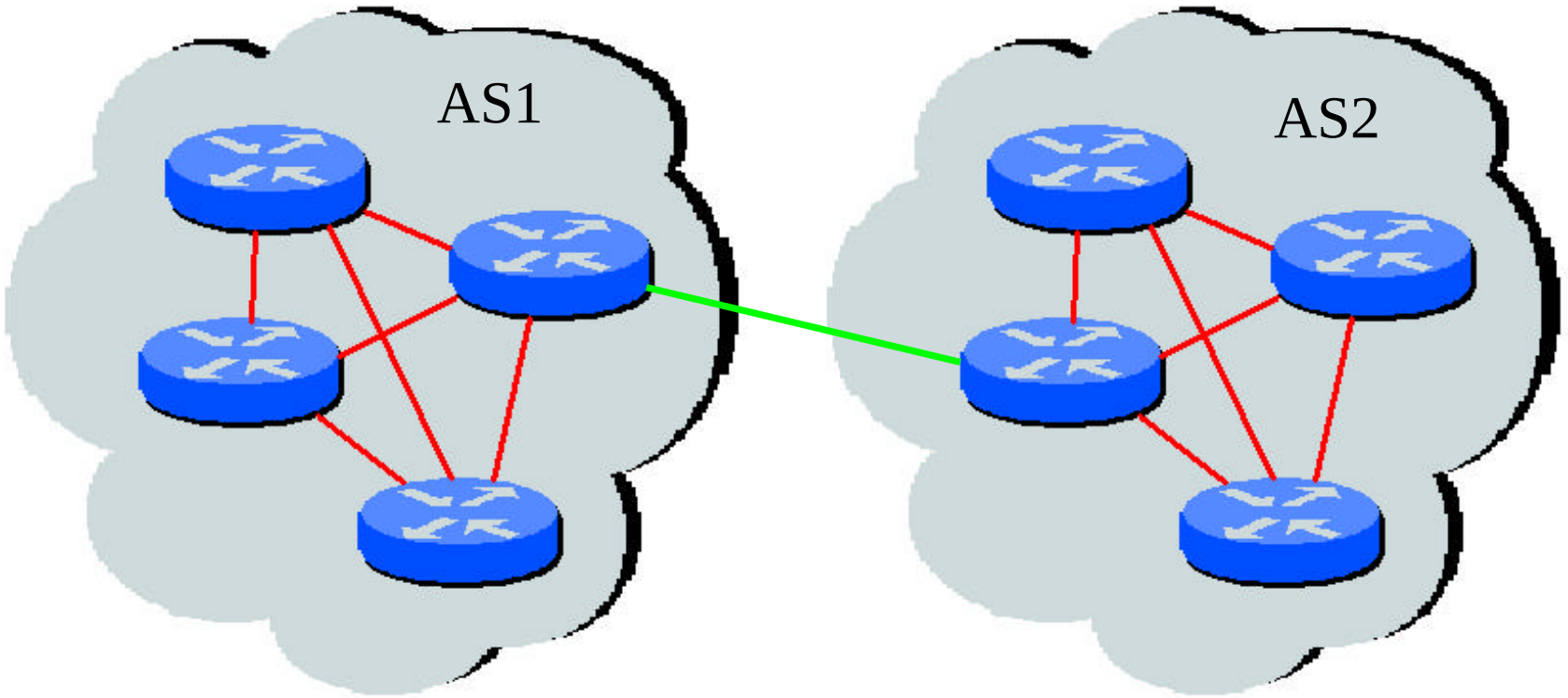
# IP Packet Routing

- Router makes decision on how to populate the forwarding table
- Copies of multiple paths to destination kept in routing table
- Only best next hop route is stored in forwarding table
- Routing decisions
  - Metrics (hopcount, delay, bandwidth, load)
  - Policies (network filtering, neighbour relationships)
  - Topology (internal, external)

# Exterior and Interior Routing

- Autonomous System (AS)
  - Collection of networks sharing the same routing policy
  - Internally connected (no islands)
- Interior Routing takes place within an AS
- Exterior Routing takes place between AS's

# Exterior and Interior Routing Diagram



# Routing Definitions

***Neighbour*** Routers exchanging routing information directly

***Peering*** The exchange of routing information between neighbours

***Announce*** Sending routing information to a neighbour

***Accept*** Receiving routing information from a neighbour

***Originate*** Insert routing information into external routing

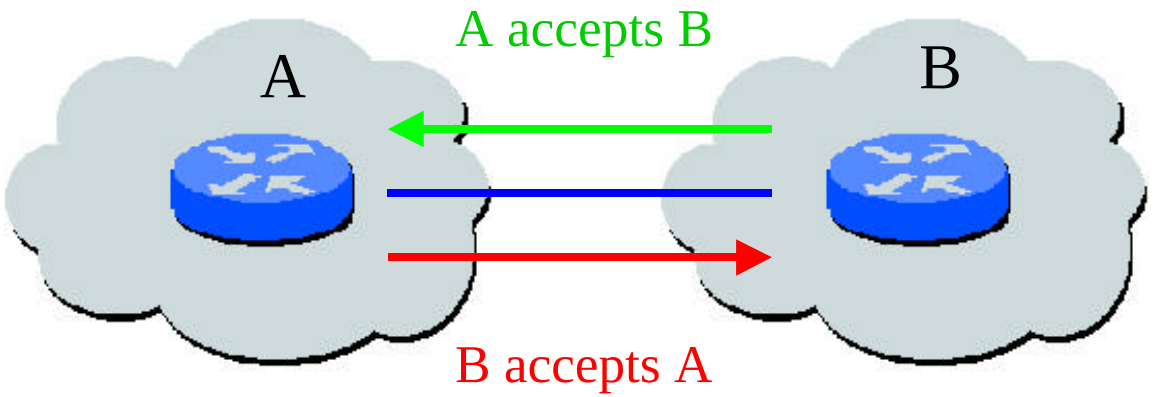
***Policy*** Schema for traffic flow

- **Applies to both internal & external routing**

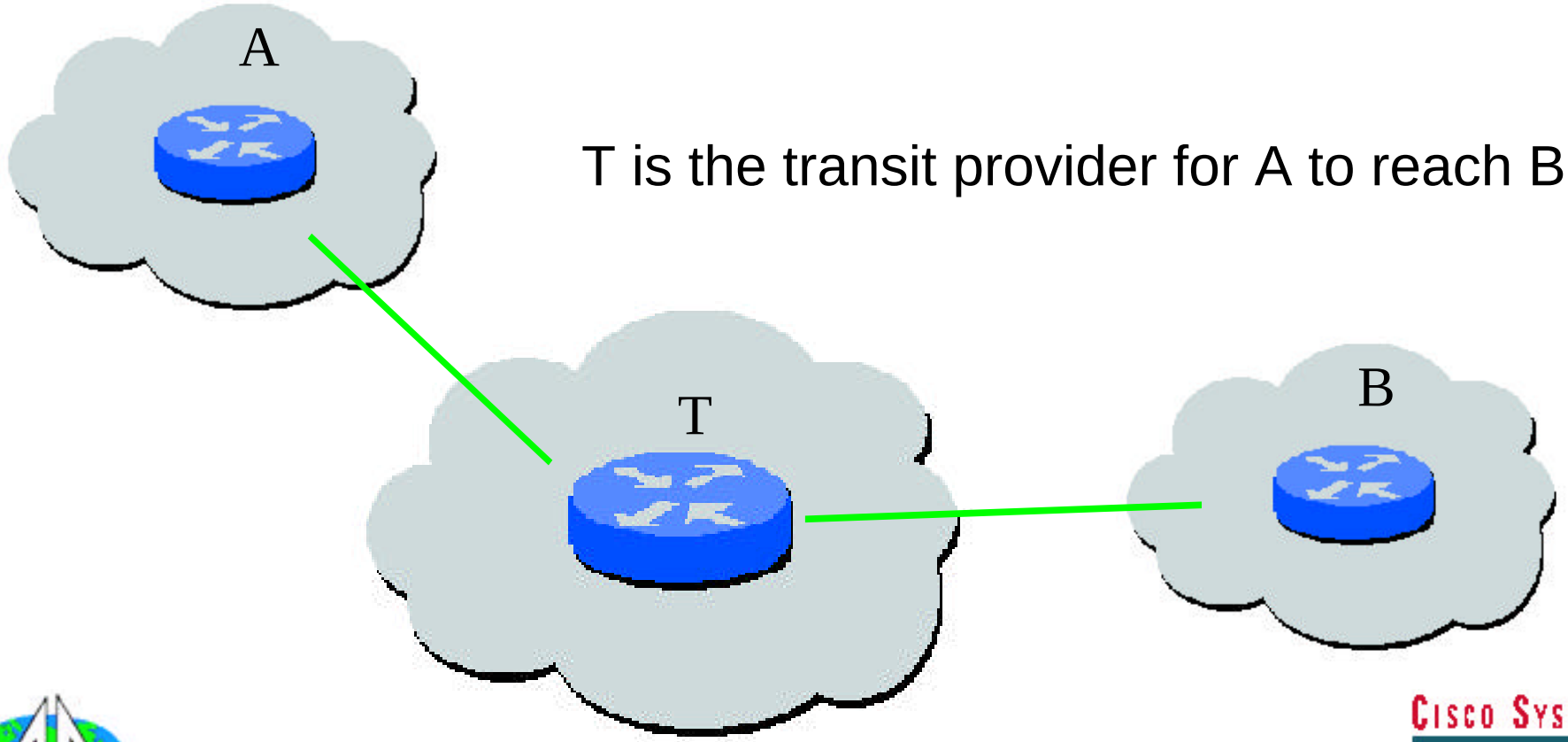
# Routing Relationships

- Neighbour
- Transit
- Multihomed
- Exchange Point

# Neighbour Relationship

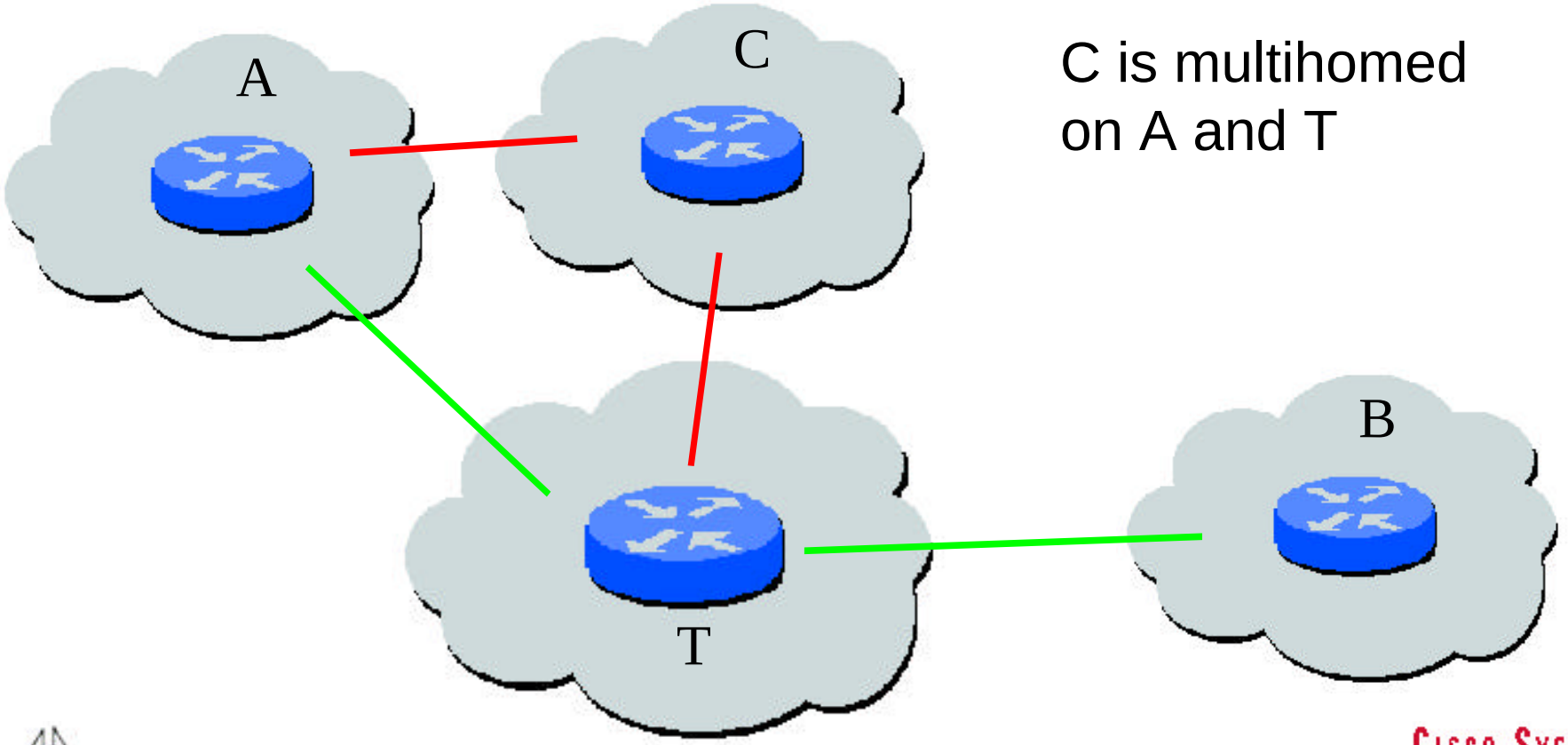


# Transit Example



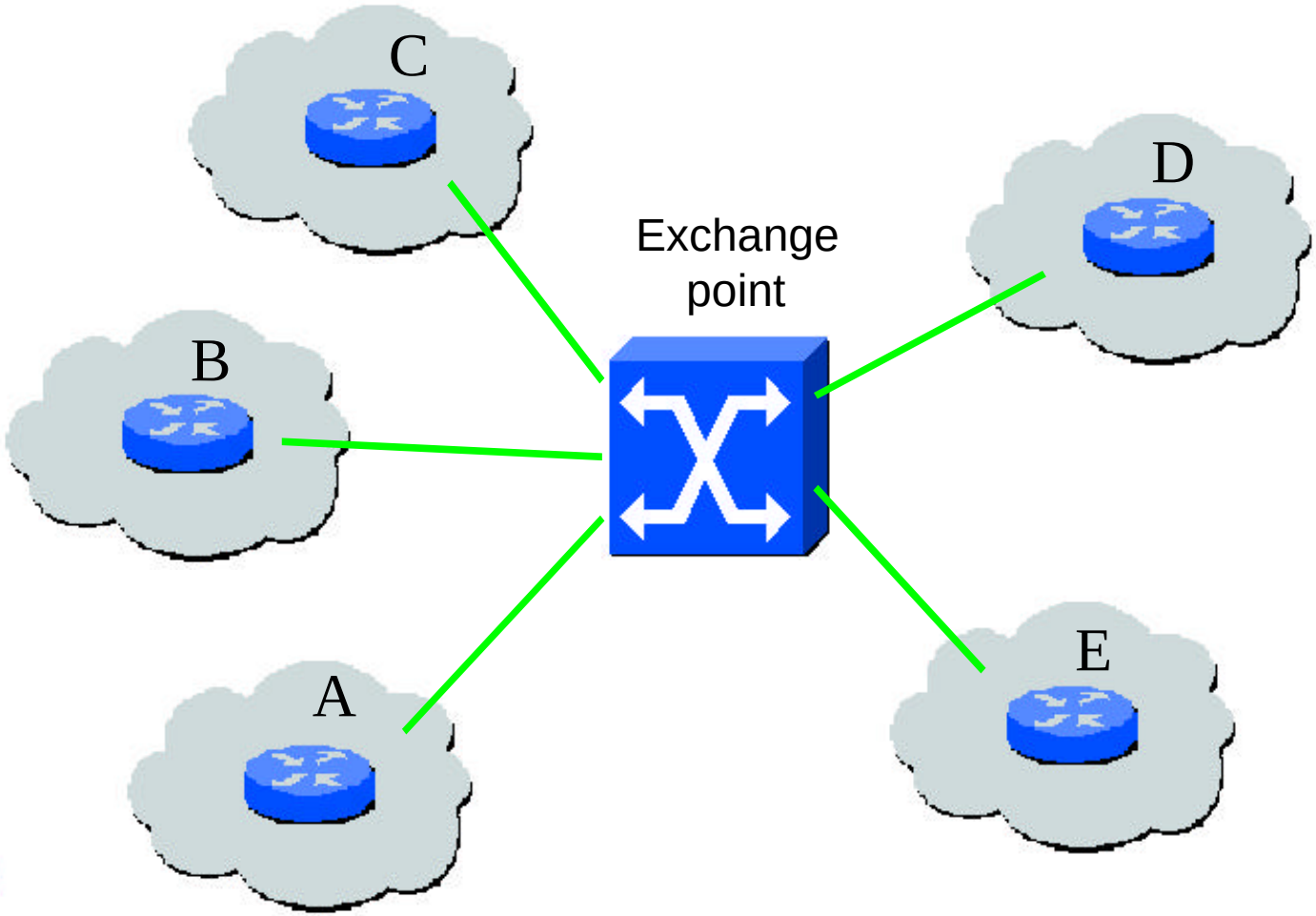

T is the transit provider for A to reach B

# Multihomed Example




C is multihomed
on A and T

# Exchange Point Example



# Complex Internetworking

- Large number of providers
- Each provider has their own rules
  - Networks advertised
  - Networks accepted
  - Special situations
- Numerous exchange points
- Numerous provider interconnects

# Routing Problems

- Have reliable knowledge about local configuration only
- Global topology not known
- Connectivity unknown
- Route groupings unknown (AS membership and groupings)

**An Internet Routing Registry is a place for documenting routing policies**




# Why an IRR?

- Register all local configurations and policies
  - Global configuration is known
  - Global topology is known
  - Connectivity can be derived
  - Paths allowed by global policy can be derived
- Register route groupings
  - AS membership
  - AS groupings
  - Other groupings (communities)
- All Internet routes known!




# Overview of the Internet Routing Registry (IRR) (Session One)




# Overview

- Definition
- Entities and Relationships
- History and Development
- Benefits
- Key Objects and Syntax of RIPE-181
- Key Objects and Syntax of RPSL
- Registering and Updating Information

# Definition

"A public and authoritative distributed repository of information"

- Public databases
- Distributed repository of information
- Have authoritative data
- Vendor independent

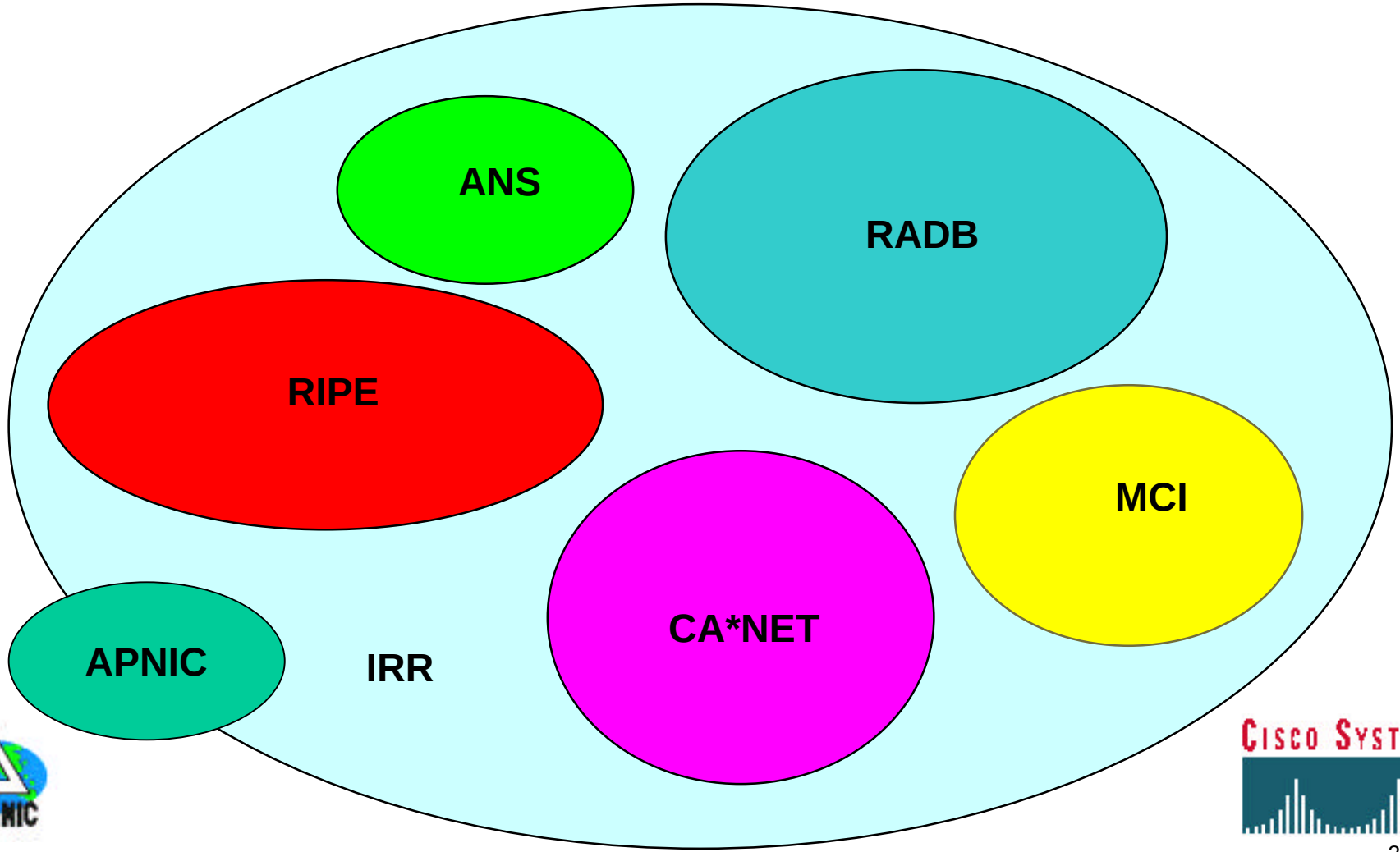
# Composition

- Routing policy details
- Routes and their aggregates
- Topology linking AS's
- Network components such as routers

- Is separate from other information such as domains and networks




# Entities



# Relationship Table

| Registry | Routing Policy | Routes | Networks | Domains |
|---|---|---|---|---|
| APNIC | Yes | No | Yes | No |
| RIPE | Yes | Yes | Yes | Yes |
| RADB | Yes | Yes | No | No |
| MCI | Yes | Yes | No | No |
| ANS | Yes | Yes | No | No |
| CA*NET | Yes | Yes | No | No |
| InterNIC | No | No | Yes | Yes |
| ARIN | ? | ? | Yes | ? |

*Shaded area = IRR*

# Relationships

- MCI, ANS and CA*NET – provider run RR's
- RIPE RR – European providers
- RADB – Default RR for rest of world
- APNIC
  - Will be full member of IRR in the future
- All RRs mirror each other

# History and Development

- Network information stored by MILNET and NSFNET/Merit
- 1989 European network management database
  - Network numbers
  - Contact information
- 1992 RIPE database authoritative
  - Network numbers
  - Contact information
  - AS information

# History and Development

- 1993 Major development
  - MERIT PRDB tools
  - PRIDE project by RIPE NCC
  - RIPE-81 MERIT and PRIDE efforts
- 1993 APNIC and InterNIC formed
- 1994 RIPE-81 → RIPE-181
- 1995 RIPE-181 → RFC1786
  - Defacto standard
  - Accepted by all IRR's at time

# History and Development

- 1995 IETF RPS Working Group
  - Continues development of RFC1786
  - Ripe-181 now expanded to RPSL
  - RFC2280 published on 31 Jan 1998
- Near future
  - Acceptance by all IRRs to convert to RPSL
  - Current draft is *draft-ietf-rps-transition-02.txt*

# Benefits of an IRR

- Operational Support
- Information
- Configuration
- Improved Quality of Service

# Operational Support

- Tools for consistency checking
- Problem Diagnosis
- Contact Information
- Details to be covered in the next sections




# Information

- Routing policy repository
- Map of global routing topology
- Routing policy between any neighbouring AS's
- Device independent description of routing policy

# Configuration

- Supports network filtering
- Configuring routers and policies
- Revision control
- Sanity checking
- Simulation?

# Improved Quality of Service

**All this adds up to improved quality of service**

**Participation is essential!**

# Overview of the Internet Routing Registry (IRR) (Session Two)




# Overview

- Definition
- Entities and Relationships
- History and Development
- Benefits
- IRR Objects and Syntax of RIPE-181
- IRR Objects and Syntax of RPSL
- Registering and Updating Information




# Key IRR Objects and Syntax of RIPE-181

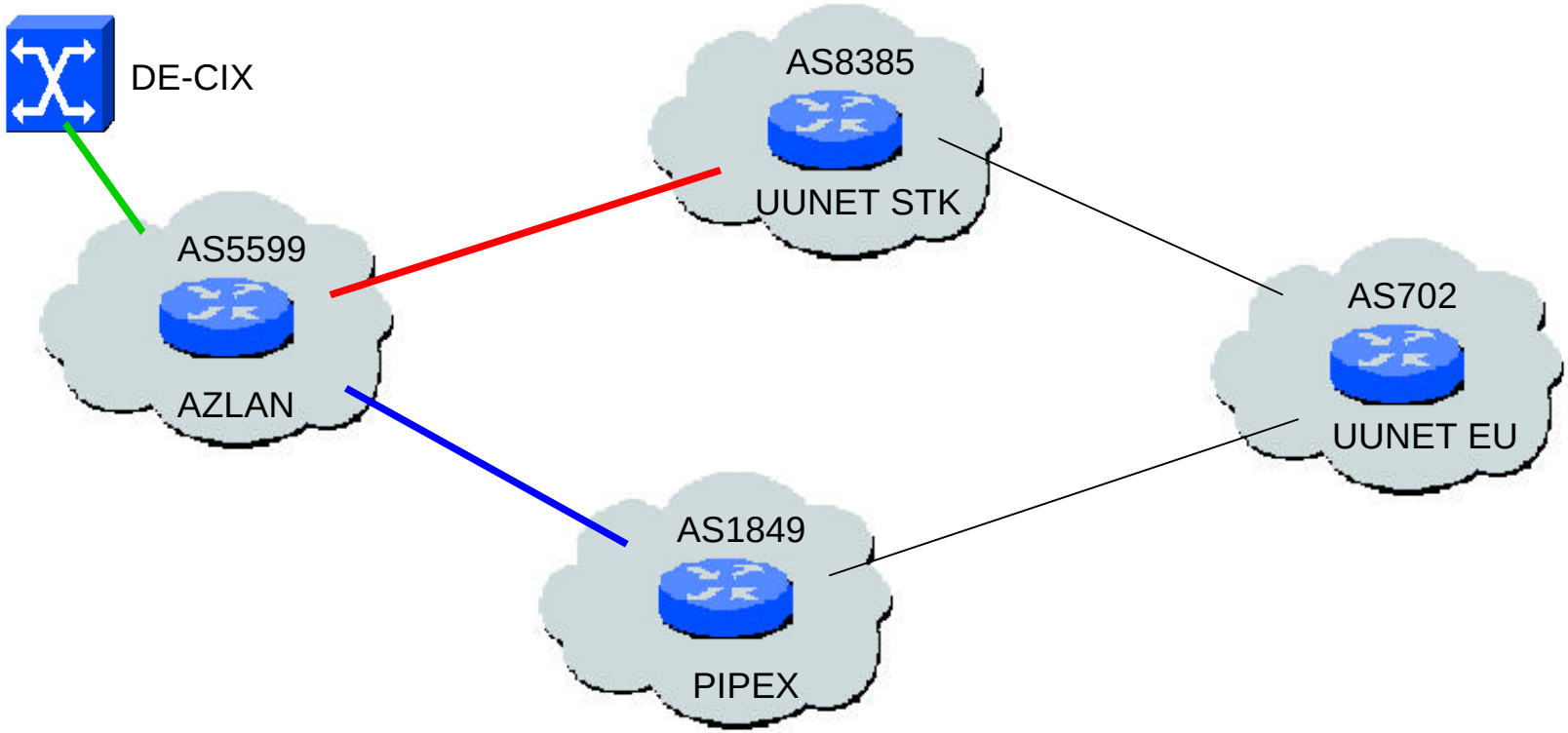
- Representation
- AS-Object
- AS-Macro
- Route Object
- Authorisation - Maintainer object

# Representation

- ASCII printable
- Attributes by “**tag: value**” lines
- Objects separated by empty lines

- RIPE-181
- RPSL

# Real World Example!



# AS-Object

| | |
|---|---|
| aut-num: | AS5599 |
| descr: | Azlan Scandinavia |
| descr: | Internet Business Unit |
| descr: | Glostrup NOC |
| as-in: | from AS1849 100 accept AS-PIPEXEURO |
| as-in: | from AS1835 100 accept AS1835 |
| as-in: | from AS2863 100 accept AS2863 |
| as-in: | from AS3292 100 accept AS-DKNET AS3292 |
| as-in: | from AS3308 100 accept AS3308 |
| as-in: | from AS5492 100 accept AS5492 |
| as-in: | from AS5509 100 accept AS5509 |
| as-in: | from AS6785 100 accept AS6785 |
| as-in: | from AS6834 100 accept AS6834 |
| as-in: | from AS8526 100 accept AS8526 |
| as-in: | from AS8385 100 accept {146.188.0.0/16} |

| | |
|---|---|
| as-out: | to AS1849 announce AS5599 |
| as-out: | to AS1835 announce AS5599 |
| as-out: | to AS2863 announce AS5599 |
| as-out: | to AS3292 announce AS5599 |
| as-out: | to AS3308 announce AS5599 |
| as-out: | to AS5492 announce AS5599 |
| as-out: | to AS5509 announce AS5599 |
| as-out: | to AS6785 announce AS5599 |
| as-out: | to AS6834 announce AS5599 |
| as-out: | to AS8526 announce AS5599 |
| as-out: | to AS8385 announce AS5599 |
| default: | AS8385 100 |
| admin-c: | MW89-RIPE |
| tech-c: | KE30-RIPE |
| mnt-by: | AS5599-MNT |
| changed: | klaus@azlan.net 970207 |
| changed: | klaus@azlan.net 971209 |
| source: | RIPE |

Connection to exchange point
Connection transit provider
Connection to backup provider

# Syntax for AS Object

Can represent policy using

- Boolean expressions (AND, OR, NOT)
- Keyword ANY - means "everything"
- Communities and AS-MACROs
- Route lists - {prefixes}
- Cost to indicate preference
- Attribute DEFAULT - accept 0.0.0.0/0

# Fields in AS-Object

- Mandatory Fields
  - aut-num, descr, admin-c, tech-c, mnt-by, changed, source, **as-in**, **as-out**
- Optional Fields
  - as-name, interas-in, interas-out, as-exclude, default, guardian, remarks, notify

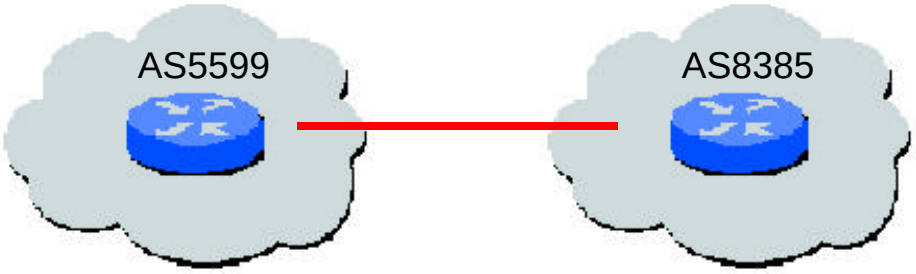
# IP Routing Policy

- Relationship between AS’s
- What to announce to each neighbour
- What to accept from each neighbour
- Selection between multiple paths
- Preferred paths
- Use default route?

# Policy Examples

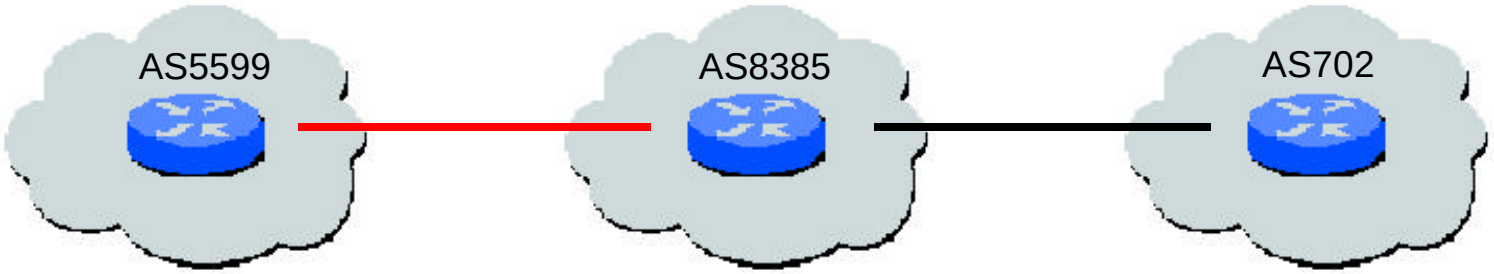
- Basic
- Transit
- Multihoming
- Exchange Point

# Basic Policy




```
aut-num:   AS5599
as-out:     to AS8385 announce AS5599
as-in:      from AS8385 100 accept {146.188.0.0/16}

aut-num:   AS8385
as-out:     to AS5599 announce {146.188.0.0/16}
as-in:      from AS5599 100 accept AS5599
```

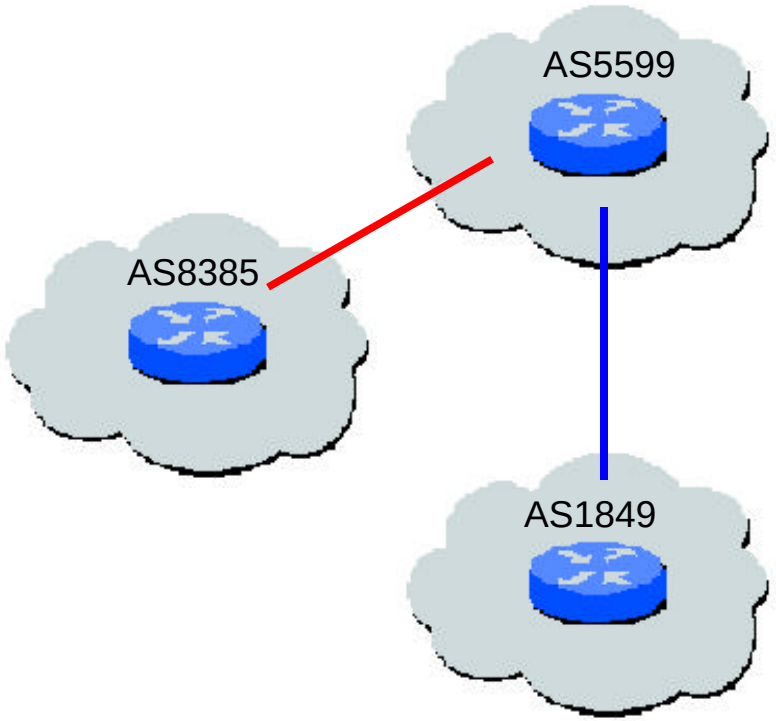
# Transit Policy




```
aut-num:  AS8385
as-in:      from AS702 100 accept ANY
as-in:      from AS5599 100 accept AS5599
as-out:     to AS702 announce AS8385 AS5599 AS8473 AND NOT {0.0.0.0/0}
as-out:     to AS5599 announce {146.188.0.0/16}
default:    AS702     50 {146.188.0.0/16}

aut-num:  AS702
as-in:      from AS8385 100 accept AS8385 AS5599 AS8473
as-out:     to AS8385 announce ANY
```

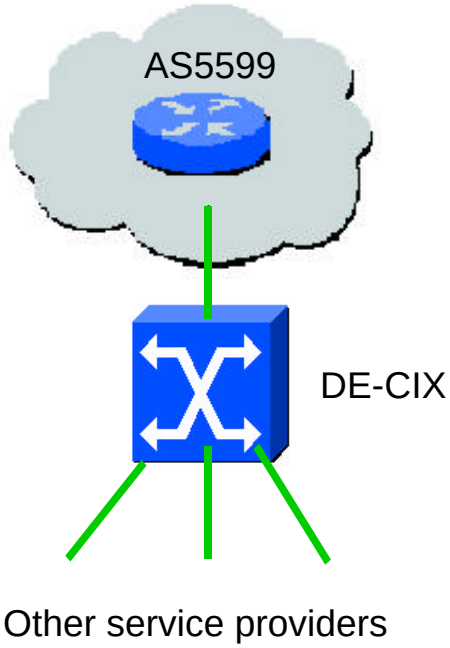
# Multihoming Policy

```
aut-num:  AS5599
as-in:    from AS1849 100 accept AS-PIPEXEURO
as-in:    from AS8385 100 accept {146.188.0.0/16}
as-out:   to AS8385 announce AS5599
as-out:   to AS1849 announce AS5599

aut-num:  AS1849
as-in:    from AS5599 100 accept AS5599
as-out:   to AS5599 announce AS-PIPEXEURO

aut-num:  AS8385
as-out:   to AS5599 announce {146.188.0.0/16}
as-in:    from AS5599 100 accept AS5599
```



# Exchange Point Policy

```
aut-num:   AS5599
as-out:      to AS1835 announce AS5599
as-out:      to AS2863 announce AS5599
as-out:      to AS3292 announce AS5599
as-out:      to AS3308 announce AS5599
as-out:      to AS5492 announce AS5599
as-out:      to AS5509 announce AS5599
as-out:      to AS6785 announce AS5599
as-out:      to AS6834 announce AS5599
as-out:      to AS8526 announce AS5599
```



# AS Macro

- Collection of AS's or other AS macros
- Describes membership of a set
- Contains no policy info
- Scales better
- Can differentiate between customer and peer routes




# Fields in AS Macro

- Mandatory Fields
  - as-macro, descr, **as-list**, tech-c, admin-c, mnt-by, changed, source
- Optional Fields
  - guardian, remarks, notify

# AS Macro

```
as-macro:   AS-UUNETSTK
descr:      UUNET customer routes in Stockholm
as-list:    AS-TAIDE
as-list:    AS-KOLUMBUS
as-list:    AS1759
as-list:    AS8385
as-list:    AS702
tech-c:     KCH251
admin-c:    ES199
remarks:    AS702 Stockholm routes are community tagged
notify:     intl-net-eng@uu.net
mnt-by:     UUNET-MNT
changed:    annel@uu.net 971113
source:     RIPE
```

Used in

```
aut-num:    AS702
as-out:     to AS1759 announce AS-UUNETSTK
```




# Route Object

- Represents a route in the Internet
- Contains all membership information
- Only one origin possible
- Classless (should be aggregated)
- Can support **hole**s and **withdrawn**

# Fields in Route Object

- Mandatory Fields
  - route, descr, **origin**, mnt-by, changed, source
- Optional Fields
  - hole, withdrawn, comm-list, remarks, notify

# Route Object

```
route:      195.129.0.0/19
descr:      UUNET-NET
origin:     AS702
remarks:    UUNET filter inbound on prefixes longer than /24
notify:     intl-net-eng@uu.net
mnt-by:     UUNET-MNT
changed:    annel@uu.net 970501
source:     RIPE
```

# Route Object

```
route:      194.216.0.0/16
descr:      PIPEX-BLOCK194216
origin:     AS1849
hole:       194.216.59.0/24
remarks:    UUNET UK filter inbound on prefixes longer than /24
mnt-by:     AS1849-MNT
changed:    philip@uk.uu.net 19980107
source:     RIPE
```

```
stk-gw1>show ip bgp 194.216.0.0 255.255.0.0 longer-prefixes
BGP table version is 53607058, local router ID is 195.242.36.254
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop        Metric LocPrf Weight Path
*> 194.216.0.0/16   146.188.30.162                 0      702 1849 i
*> 194.216.59.0     146.188.30.162                 0      702 701 3491 5557 i
```

# RPSL

- Enhanced over RIPE-181
- Detailed – can generate router configurations
- Extensible – caters for new routing protocols
- **NOT** designed to be router configuration language, but a routing **policy** specification language

RPSL described in ftp://ftp.apnic.net/ietf/rfc/rfc2280.txt

# Key IRR Objects and Syntax of RPSL

- "Object" in ripe-181 becomes "Class" in RPSL
  - aut-num class
  - as-set class
  - route class
  - route-set class *NEW!*

# Representation

- As RIPE-181, plus
- Attribute’s value can be split over multiple lines
- Comments can be added with a # at the start of a line
- Supports regular expressions
- Supports PGP encryption

# Availability

- RSPL coming real soon
- User documentation
- Tutorials
- RAToolset updated




# Conversion to RPSL

- All IRRs have agreed to transition to RPSL
- Basic conversion very simple
- Transition Plan

  *ftp://ftp.apnic.net/ietf/internet-drafts/draft-ietf-rps-transition-02.txt*

| | **Ripe-181** | **RPSL** |
|---|---|---|
| Phase1 | Read/Write | |
| Phase2 | Read/Write | Read |
| Phase3 | Write | Read/Write |
| Phase4 | | Read/Write |


# Conversion Tool

- *http://www.isi.edu/ra/rps/transition/*

**Welcome to the RPSL Transition Page**
This page presents the latest information on the transition from RIPE-181 to the new Routing Policy Specification Language. The information will be updated frequently; visit often to stay up-to-date.

- Query a mirrored copy of the Internet Routing Registry
- Query a mirrored copy of the IRR that has been converted to RPSL
- Convert RIPE-181 objects to RPSL / Perform RPSL syntax checks
- Access ISI's RPSL-capable database server
- Download the RIPE-to-RPSL converter tool
- Transition Plan
- RPSL Transition Presentation to NANOG, October 1997

# Aut-Num Class Example

**RPSL**

| | |
|---|---|
| aut-num: | AS5599 |
| **as-name**: | UNSPECIFIED |
| descr: | Azlan Scandinavia |
| descr: | Internet Business Unit |
| descr: | Glostrup NOC |
| **import**: | from AS1849 |
| | action pref = 100; |
| | accept AS-PIPEXEURO |
| <snip> | |
| **export**: | to AS1849 |
| | announce AS5599 |
| <snip> | |
| default: | to AS8385 |
| | action pref = 100; |
| | networks ANY |
| admin-c: | MW89-RIPE |
| tech-c: | KE30-RIPE |
| remarks: | This data is automatically converted |
| remarks: | from the RIPE181 registry (19980106) |
| notify: | as-guardian@azlan.net |
| mnt-by: | AS5599-MNT |
| changed: | klaus@azlan.net 19970207 |
| source: | RIPE |

**RIPE-181**

| | |
|---|---|
| aut-num: | AS5599 |
| descr: | Azlan Scandinavia |
| descr: | Internet Business Unit |
| descr: | Glostrup NOC |
| as-in: | from AS1849 100 accept AS-PIPEXEURO |
| <snip> | |
| as-out: | to AS1849 announce AS5599 |
| <snip> | |
| default: | AS8385 100 |
| admin-c: | MW89-RIPE |
| tech-c: | KE30-RIPE |
| notify: | as-guardian@azlan.net |
| mnt-by: | AS5599-MNT |
| changed: | klaus@azlan.net 19970207 |
| source: | RIPE |

# Route Class Example

**RIPE-181**

```
route:      194.216.0.0/16
descr:      PIPEX-BLOCK194216
origin:     AS1849
hole:       194.216.59.0/24
remarks:    UUNET UK filter inbound on prefixes longer than /24
mnt-by:     AS1849-MNT
changed:    philip@uk.uu.net 19980107
source:     RIPE
```

**RPSL**

```
route:      194.216.0.0/16
descr:      PIPEX-BLOCK194216
origin:     AS1849
hole:       194.216.59.0/24
remarks:    UUNET UK filter inbound on prefixes longer than /24
mnt-by:     AS1849-MNT
changed:    philip@uk.uu.net 19980107
source:     RIPE
```

# AS-Set Class Example

**RIPE-181**

```
as-macro:     AS-UUNETSTK
descr:        UUNET customer routes in Stockholm
as-list:      AS-TAIDE
as-list:      AS-KOLUMBUS
as-list:      AS1759
as-list:      AS8385
as-list:      AS702
tech-c:       KCH
admin-c:      ES199
remarks:      AS702 Stockholm routes are community tagged
notify:       intl-net-eng@uu.net
mnt-by:       UUNET-MNT
changed:      annel@uu.net 971113
source:       RIPE
```

**RPSL**

```
as-set:       AS-UUNETSTK
descr:        UUnet customer routes in Stockholm
members:      AS-TAIDE, AS-KOLUMBUS, AS702, AS1759, AS8385
remarks:      AS702 Stockholm routes are community tagged
remarks:      This data is automatically converted from the RIPE181 registry (19980127)
tech-c:       KCH
admin-c:      ES199
notify:       intl-net-eng@uu.net
mnt-by:       UUNET-MNT
changed:      annel@uk.uu.net 971113
changed:      davidk@ISI.EDU 19980127
source:       RIPE
```

# How to Register and Update Information in IRRs

- Frequently used objects
- Update Procedures
  - Modifying objects
  - Deleting objects
  - Submitting objects
  - Authorisation/notification
  - Warnings and errors
  - NIC handles

# Frequently Used Objects

- Person: contact person
- Maintainer: authorisation of objects
- Inetnum: address assignment
- Aut-num: autonomous systems
- As-macro: set of ASes
- Route: announced routes




# Unique Keys

- Uniquely identifies object
- Updating object overwrites old entry – need unique key
- Used in querying by **whois**
- Web based full text searches are becoming available

# Unique Keys

- Person: name plus NIC handle
- Maintainer: maintainer name
- Inetnum: network number
- Aut-num: AS number
- As-macro: AS-macro name
- Route: route value plus origin

# Modifying an Object

Before

| | |
|---|---|
| **person:** | **Philip F. Smith** |
| address: | UUNET UK |
| address: | Internet House |
| address: | 332 Science Park |
| address: | Milton Road |
| address: | Cambridge CB4 4BZ |
| address: | England, UK |
| phone: | +44 1223 250100 |
| fax-no: | +44 1223 250101 |
| e-mail: | philip@uk.uu.net |
| **nic-hdl:** | **PFS2-RIPE** |
| notify: | philip@uk.uu.net |
| changed: | philip@uk.uu.net 19971202 |
| source: | RIPE |

Submitted and After

| | |
|---|---|
| **person:** | **Philip F. Smith** |
| address: | Cisco Systems Australia |
| address: | Level 13, 80 Albert Street |
| address: | Brisbane 4000 |
| address: | QLD |
| address: | Australia |
| phone: | +61 7 3309 8602 |
| fax-no: | +61 7 3211 3889 |
| e-mail: | pfs@cisco.com |
| e-mail: | philip@dial.pipex.com |
| **nic-hdl:** | **PFS2-RIPE** |
| notify: | philip@dial.pipex.com |
| changed: | pfs@cisco.com 19980209 |
| source: | RIPE |

- Unique keys must stay the same
- Remember to use current date

# Deleting an Object

```
person:     Philip F. Smith
address:    UUNET UK
address:    332 Science Park
address:    Milton Road
address:    Cambridge
address:    England, UK
phone:      +44 1223 250100
fax-no:     +44 1223 250101
e-mail:     philip@uk.uu.net
nic-hdl:    PFS2-RIPE
notify:     philip@uk.uu.net
changed:    philip@uk.uu.net 19971202
source:     RIPE
delete:     philip@dial.pipex.com left company
```

- **delete** deletes object from database
- current object must be submitted exactly as is, only with extra **delete** line
- If there is a **mnt-by** line, need the password!

# Submitting Objects

- E-mail interface

  *auto-rr@mci.net*

  - Robot mail box
  - Send all updates to this mailbox
  - Can use LONGACK in subject line

  *apnic-dbm@rs.apnic.net*

  - Human mail box
  - Questions and comments

# Authorisation/Notification

```
route:      194.216.0.0/16
descr:      PIPEX-BLOCK194216
origin:     AS1849
hole:       194.216.59.0/24
remarks:    UUNET UK filter inbound on prefixes longer than /24
mnt-by:     AS1849-MNT
notify:     support@uk.uu.net
changed:    philip@uk.uu.net 19980107
source:     RIPE
```

- **mnt-by** the **maintainer** object
- **notify** who is notified of changes

# Maintainer Object

- Who is authorised
- Authorisation Method
  - email-from
  - crpyt-pw - encrypted password
- Mandatory Fields
  - **mntner**, descr, admin-c, tech-c, upd-to, auth, mnt-by
- Optional Fields
  - mnt-nfy, changed, notify, source

# Maintainer Object

Maintainer Object AS1849-MNT

```
mntner:     AS1849-MNT
descr:      AS 1849 Maintainer - PIPEX UK
admin-c:    PFS2-RIPE
tech-c:     PFS2-RIPE
upd-to:     philip@uk.uu.net
mnt-nfy:    netdev@uk.uu.net
auth:       CRYPT-PW fjOImdmwKsx
mnt-by:     AS1849-MNT
changed:    philip@uk.uu.net 19980109
source:     RIPE
```

Object has to be registered manually




# Authorisation/Notification

```
route:      194.216.0.0/16
descr:      PIPEX-BLOCK194216
origin:     AS1849
hole:       194.216.59.0/24
hole:       194.216.136.0/23
remarks:    UUNET UK filter inbound on prefixes longer than /24
mnt-by:     AS1849-MNT
passwd:     c4Ange5
notify:     support@uk.uu.net
changed:    philip@uk.uu.net 19980109
source:     RIPE
```

- New **hole** to be added.
- **passwd** field to allow change
- <support@uk.uu.net> will be notified of this change
- updated **changed** field

# Warnings and Errors

- Warnings
  - Object corrected and accepted
  - Notification of action taken sent in acknowledgement
- Errors
  - Object not corrected and not accepted
  - Diagnostics in acknowledgement
  - If not understandable, send e-mail to a human!
- Syntax checking is very strict, so be careful

# NIC Handles

| | |
|---|---|
| mntner: | AS1849-MNT |
| descr: | AS 1849 Maintainer - PIPEX UK |
| **admin-c:** | **PFS2-RIPE** |
| **tech-c:** | **PFS2-RIPE** |
| upd-to: | philip@uk.uu.net |
| mnt-nfy: | netdev@uk.uu.net |
| auth: | CRYPT-PW fjOImdmwKsx |
| mnt-by: | AS1849-MNT |
| changed: | philip@uk.uu.net 19980109 |
| source: | RIPE |

**PFS2-RIPE** is the NIC Handle of the person

- Only way of avoiding ambiguity in person objects
- Mandatory
- Format: <initials><number>-<regional registry>
- Local differences for obtaining NIC Handles.

# Tools and Resources




# Overview

- What tools & resources are there?
- Where can I get them?
- How do they help me?

# What Resources?

- RAToolset
- PRIDE
- RIPE Whois
- Looking glasses eg. *http://nitrous.digex.net*
- Traceroute servers eg. *http://www.boardwatch.com/isp/trace.htm*

# Sources

| Resources | Platform | Location |
|---|---|---|
| RAToolset<br>PRIDE<br>RIPEWhois<br>IPMA | Unix/tcl7.4/tk2.0/g++2.7.2<br>Unix/perl4<br>Unix<br>Unix/Win95/ie3/netscape3 | www.isi.edu/ra/RAToolSet<br>ftp.ripe.net:/pride/tools/<br>ftp.ripe.net:/tools/ripe/ripe-whois-tools-2.2.tar.gz<br>www.merit.edu/ipma |

# RAToolSet

- Current version 3.5.7
- Runs on (most) UNIX platforms
- Requires g++, tcl and tk
  - Not included in toolset
- Excellent  for housekeeping, debugging & configuration

Highly recommended!

# RAToolSet Tools

- Some useful tools can help to
  - Generate router configurations

    **RTconfig**
  - Update aut-num, as-macro objects

    **AOE - Aut-num Object Editor**
  - Update route-object

    **ROE - Route Object Editor**
  - Give advice on CIDRisation

    **CIDRadvisor**

# RAToolSet

- Other useful tools are
  - A traceroute which compares RR and real world

    **PRtraceroute**
  - Checks syntax & semantics of your AS objects

    **PRcheck**
  - Expand AS macro to see content

    **Peval**

- PRTools *very* similar to PRIDE tools but supported

# RAToolset Uses

| RAToolset | Debugging | Configuration | Housekeeping |
|---|---|---|---|
| AOE | | *yes* | *yes* |
| ROE | | *yes* | *yes* |
| CidrAdvisor | | *yes* | *yes* |
| RTConfig | | *yes* | |
| PRCheck | | | *yes* |
| Peval | *yes* | | |
| Prtraceroute | *yes* | | |

# RtConfig Example

From RAToolSet

- Automating router configuration builds
- Eases the management of access lists
- Reduces possibility of human error
- Interacts with the routing registry

# RtConfig Example

**RtConfig –supress_martian –s ripe < ip-lists.template**

- Suppress known "martian" networks
- Start with RIPE database for search
- Use "ip-lists.template" file

# RtConfig Example

```
!
interface serial 5/1
  ip address 195.242.32.5 255.255.255.252
!
router bgp 8385
  neighbour 195.242.32.6 remote-as 5599
@RtConfig set cisco_access_list_no = 100
@RtConfig set cisco_max_preference = 100
@RtConfig set cisco_map_name = "azlan-in"
@RtConfig import AS8385 195.242.32.5/32 AS5599 195.242.32.6/32
!
```

ip-lists.template file – can be normal configuration file

# RtConfig Example

```
no access-list 100
access-list 100 deny   ip host 0.0.0.0 any
access-list 100 deny   ip 127.0.0.0 0.255.255.255 255.0.0.0 0.255.255.255
access-list 100 deny   ip 10.0.0.0 0.255.255.255 255.0.0.0 0.255.255.255
<snip>
access-list 100 deny   ip 224.0.0.0 31.255.255.255 224.0.0.0 31.255.255.255
access-list 100 deny   ip any 255.255.255.128 0.0.0.127
access-list 100 permit ip 195.24.0.0   0.0.0.0   255.255.224.0   0.0.0.0
access-list 100 deny ip 0.0.0.0 255.255.255.255 0.0.0.0 255.255.255.255
!
no route-map azlan-in
route-map azlan-in permit 1
 set local-preference 100
!
router bgp 8385
neighbor 195.242.32.6 route-map azlan-in in
neighbor 195.242.32.6 distribute-list 100 in
```

Output from RtConfig

# RtConfig Example

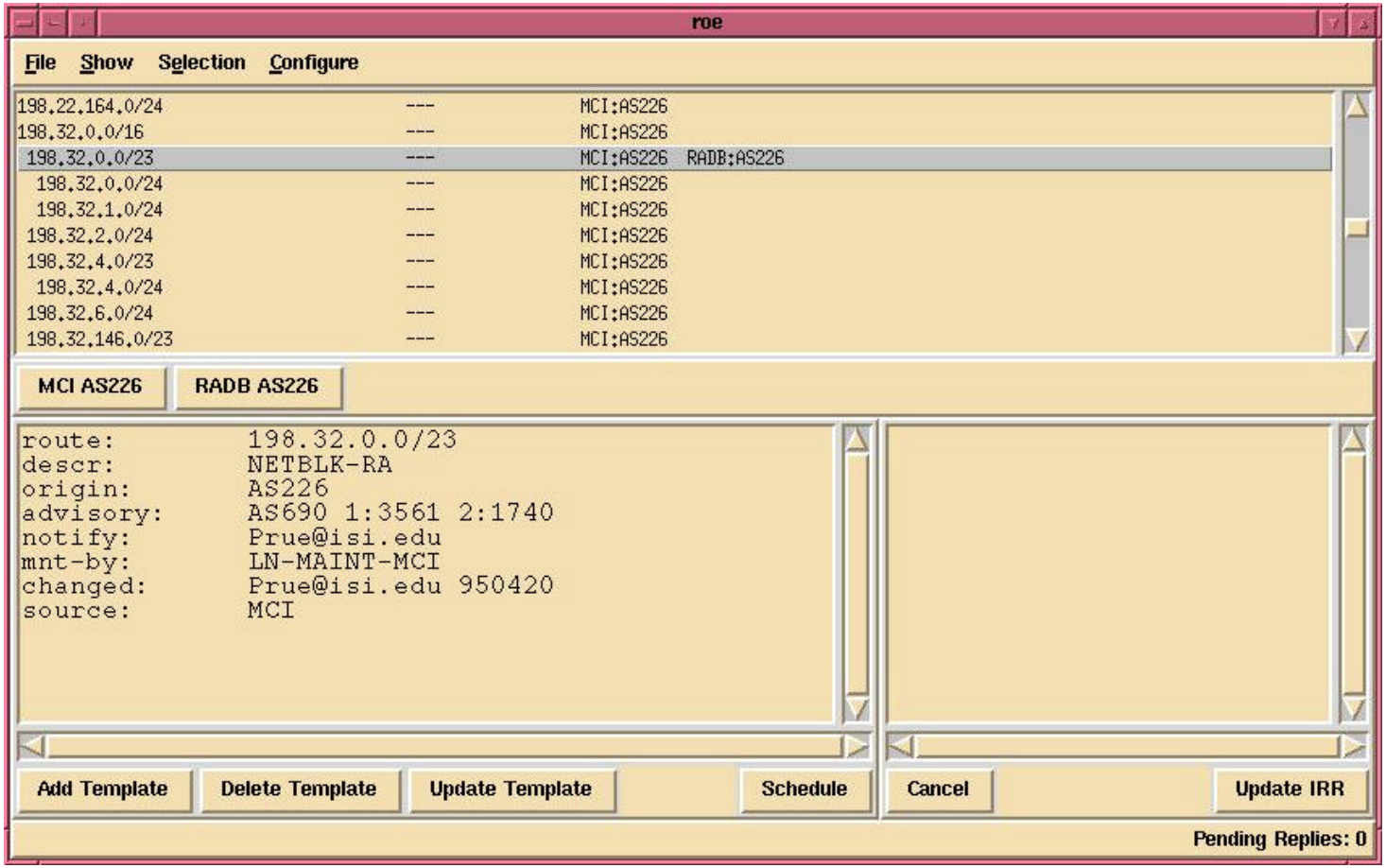
- RtConfig only supported by RADB.
  - RADB mirrors all other RRs
- Martian networks are defined in RtConfig
- Default produces “route-map foo"
- Default sets local pref 1000

# ROE uses

Route object editor used to

- Check for consistency of route objects in IRRs
- Synchronise route object entries in different IRRs
- Detect missing or unwanted route objects

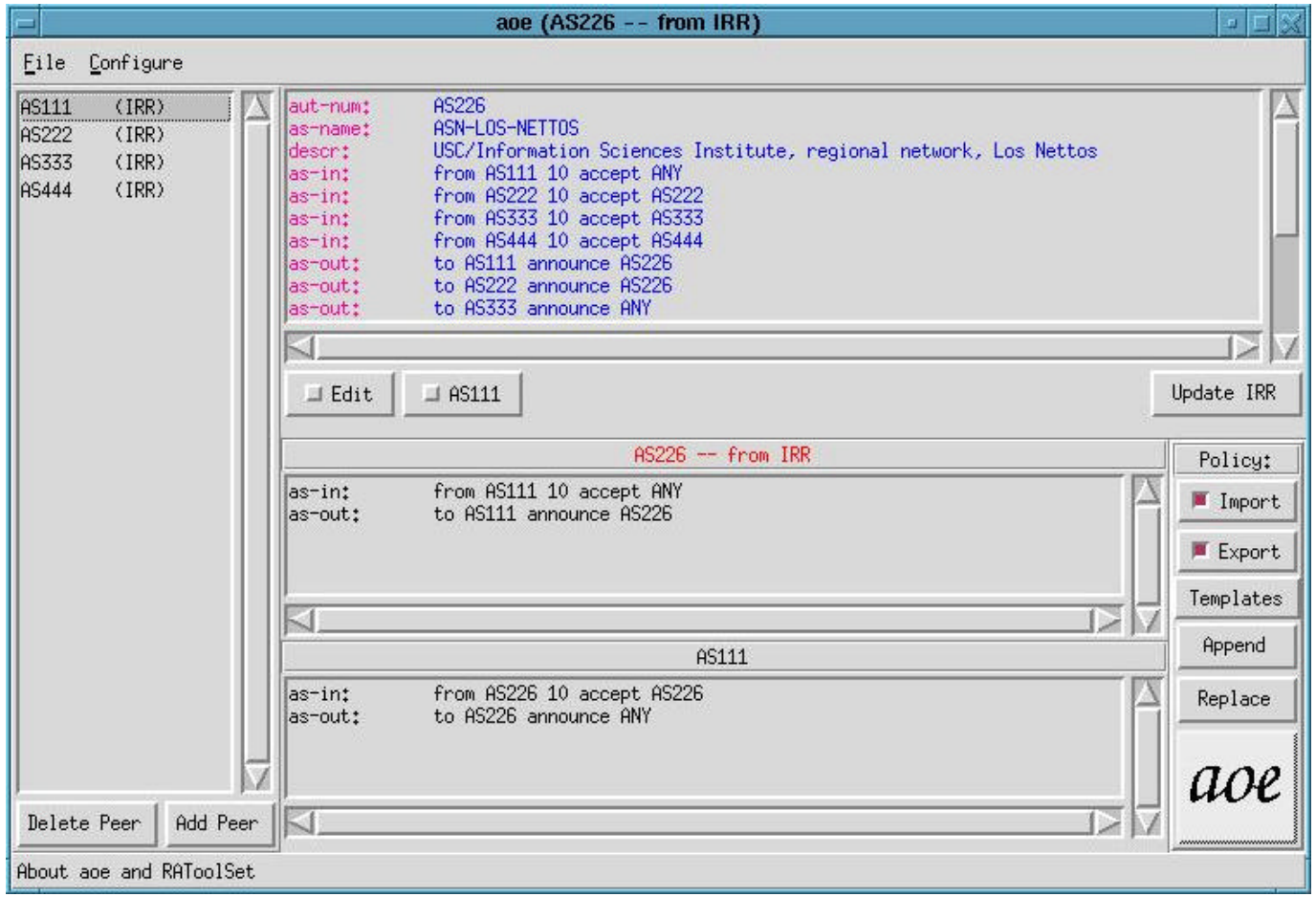
# ROE example


roe
File Show Selection Configure
198.22.164.0/24 --- MCI:AS226
198.32.0.0/16 --- MCI:AS226
198.32.0.0/23 --- MCI:AS226 RADB:AS226
198.32.0.0/24 --- MCI:AS226
198.32.1.0/24 --- MCI:AS226
198.32.2.0/24 --- MCI:AS226
198.32.4.0/23 --- MCI:AS226
198.32.4.0/24 --- MCI:AS226
198.32.6.0/24 --- MCI:AS226
198.32.146.0/23 --- MCI:AS226
MCI AS226
RADB AS226
route: 198.32.0.0/23
descr: NETBLK-RA
origin: AS226
advisory: AS690 1:3561 2:1740
notify: Prue@isi.edu
mnt-by: LN-MAINT-MCI
changed: Prue@isi.edu 950420
source: MCI
Add Template
Delete Template
Update Template
Schedule
Cancel
Update IRR
Pending Replies: 0

# AOE uses

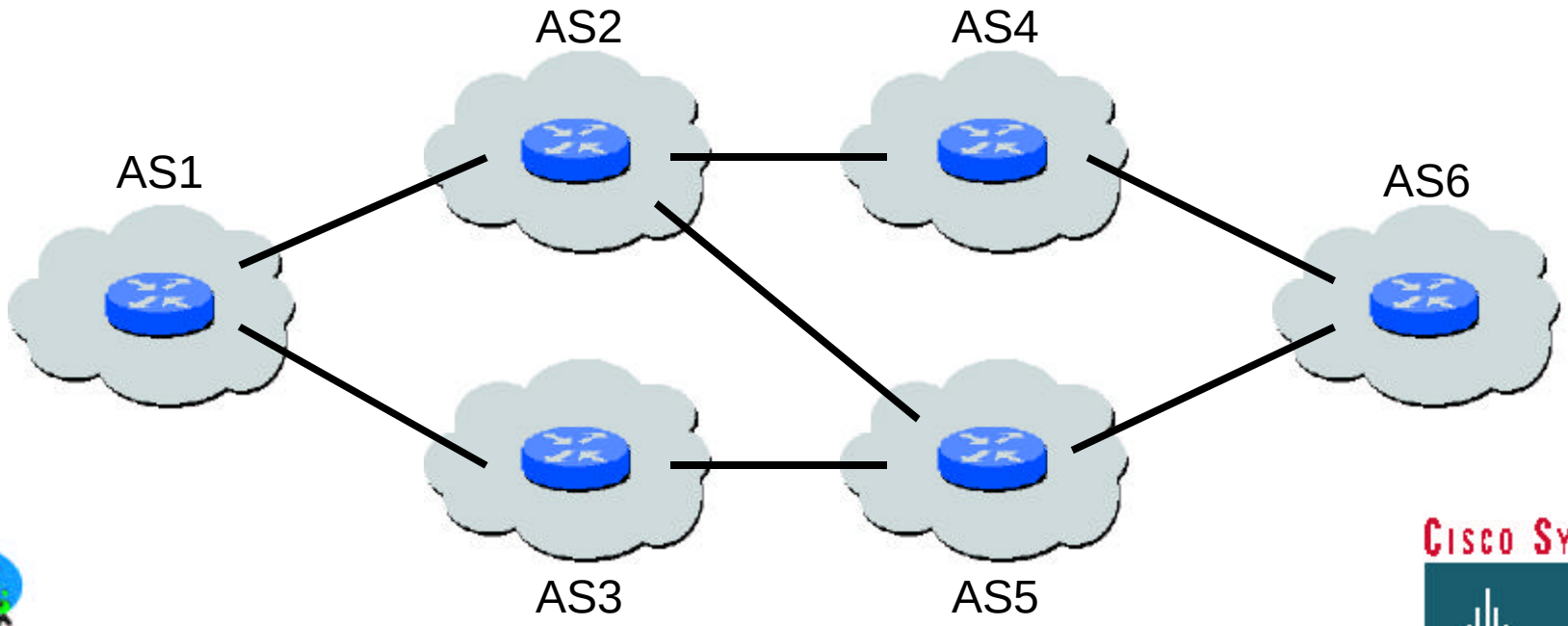
- AS object editor used to
  - Generate AS objects and policies as-in and as-out
    - *from BGP dump*
    - *from peers AS objects*
    - *manually*
  - Check policies listed in AS Object on the IRRs
  - Check policies according to BGP dump

# AOE example


aoe (AS226 -- from IRR)
File Configure
AS111 (IRR)
AS222 (IRR)
AS333 (IRR)
AS444 (IRR)
aut-num: AS226
as-name: ASN-LOS-NETTOS
descr: USC/Information Sciences Institute, regional network, Los Nettos
as-in: from AS111 10 accept ANY
as-in: from AS222 10 accept AS222
as-in: from AS333 10 accept AS333
as-in: from AS444 10 accept AS444
as-out: to AS111 announce AS226
as-out: to AS222 announce AS226
as-out: to AS333 announce ANY
Edit
AS111
Update IRR
AS226 -- from IRR
as-in: from AS111 10 accept ANY
as-out: to AS111 announce AS226
AS111
as-in: from AS226 10 accept AS226
as-out: to AS226 announce ANY
Policy:
Import
Export
Templates
Append
Replace
aoe
Delete Peer
Add Peer
About aoe and RAToolSet

# PRtraceroute

- Modified traceroute which includes AS information and a comparison between the real route and the route according to the IRR.



# PRtraceroute Example

```
% prtraceroute -lv collegepk-cr9.bbnplanet.net
traceroute with AS and policy additions [Jan 13 20:21:19 UTC]

                from  AS109 lovefm.cisco.com (171.68.228.35)
                to     AS86 collegepk-cr9.bbnplanet.net (192.239.103.9)

 1  AS109  al.cisco.com                    171.68.228.3     [I] 4 1 1 ms
 2  AS109  acorn.cisco.com                 171.68.0.134     [I] 2 1 1 ms
 3  AS109  gaza-gw2.cisco.com              171.68.0.91      [I] 2 1 1 ms
 4  AS109  sj-wall-2.cisco.com             198.92.1.138     [I] 3 3 2 ms
 5  AS109  barrnet-gw.cisco.com            192.31.7.37      [I] 4 3 2 ms
 6  AS200  paloalto-cisco.bbnplanet.net    131.119.26.9     [?] 4 4 3 ms
 7  AS200  paloalto-br1.bbnplanet.net      131.119.0.193    [I] 7 8 7 ms
 8    AS1  chicago2-br1.bbnplanet.net      4.0.1.2          [E1] 58 59 58 ms
 9    AS1  collegepk-br1.bbnplanet.net     4.0.1.6          [I] 82 73 75 ms
10   AS86  collegepk-cr9.bbnplanet.net     128.167.252.9    [E1] 86 81 ms

AS Path followed: AS109 AS200 AS1 AS86

 AS109 = Cisco Systems
 AS200 = BBN Planet Western Region
   AS1 = BBN Planet backbone
  AS86 = SURAnet Northern AS
```

| | |
|---|---|
| **ERROR** | **hop should not have been taken** |
| **NH ASx** | **possible NEXT_HOP followed** |
| **I** | **intra AS hop** |
| **En** | **nth choice inter AS hop** |
| **Dn** | **nth choice default hop** |
| **C** | **connected hop** |
| **?** | **No information in IRR** |

# RIPE whois

- Runs on most (UNIX) platforms
- Easy to install
- Can use to query all other IRR's
- Expanded whois functionality
- Good for housekeeping, debugging, operations

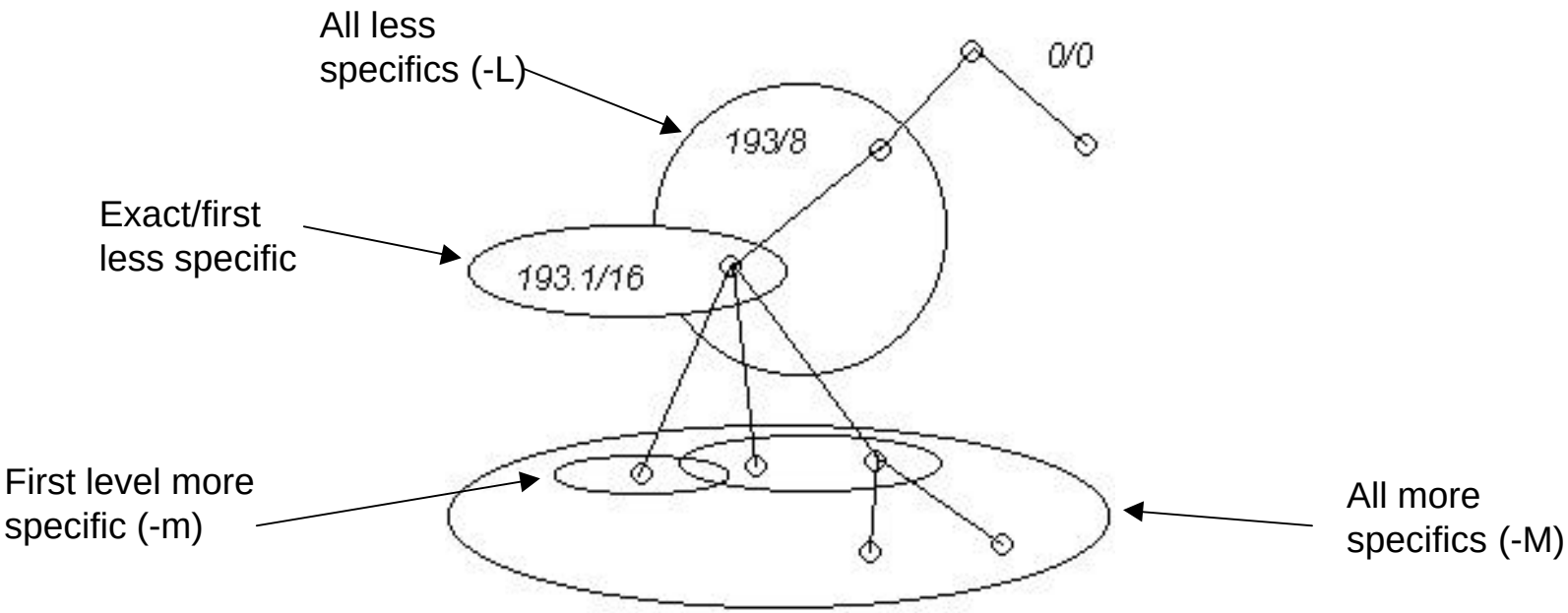
Highly recommended**!**

# RIPE whois flags

**Whois help**

- **i** inverse lookup for specified attributes
- **L** find all less specific matches
- **r** turn off recursive lookups
- T type only - look only for object of certain type
- a search all databases
- t show template for object of certain type
- v verbose information for object type

# Example whois DB query




Example query: 193.1.0.0/16

# Whois -i flag

- **Inverse lookup for special arguments**

Examples

- whois -i tech-c,admin-c,zone-c DR222 [NIC handle]
- whois -i notify anne@apnic.net [Email]
- whois -i origin AS702 [AS number]
- whois -i mnt-by AS1849-MNT [Maintainer object]

**Very useful!**

# Routing Etiquette




# Overview

- Motivation
- CIDR and aggregation
- “The Swamp”
- Renumbering
- Dampening
- NAT
- Filtering Policies

# Motivation – Problems on the Internet

- Concern about rapidity of Internet growth
  - Size of Internet: Network Wizards survey
    *http://www.nw.com/zone/WWW/report.html*
- Large number of routes
  - History chart
    *http://www.employees.org/~tbates/cidr.hist.plot.html*
- Routing churn
  - MERIT network report
    *http://www.merit.edu/ipma/reports*
- Difficulties diagnosing problems
- Quality of Service


# CIDR and aggregation

- Meaning?
- Examples
- Effects
- Positive efforts




# CIDR – Meaning

- “Classless InterDomain Routing”
  *ftp://ftp.apnic.net/ietf/rfc/rfc1519.txt*
- Concept of class A, class B and class C is no more
- New terminology – Address Prefixes
- Supported by BGP4, OSPF, ISIS and EIGRP

# CIDR – Effects

If Internet were unaggregated

- Would be over 200000 networks
- What size of routers required then?
- How stable would Internet be?

# CIDR – Examples

- Announce network allocation block, not the individual networks within that
- On Cisco routers
  - Redistribute static
  - Aggregate address
  - Network/mask pair




# CIDR – Examples

Redistribute static

```
router bgp 1849
network 194.216.0.0
redistribute static
! Must have a matching IGP route
ip route 194.216.0.0 255.255.0.0 null0
```

Aggregate address

```
router bgp 1849
network 194.216.0.0
aggregate-address 194.216.0.0 255.255.0.0
! More specific route must exist in BGP table
```

Network/mask pair

```
router bgp 1849
network 194.216.0.0 mask 255.255.0.0
! Must have a matching IGP route
ip route 194.216.0.0 255.255.0.0 null0
```


# CIDR – Positive efforts

- Many ISP's filter out networks longer than /24
- ISP's filter according to policy registered in the IRR
  - eg. Sprint
- No aggregation/bad aggregation could result in no connectivity!

# Aggregation

- Announce aggregate to rest of Internet
- Put it in routing registry (route object)
- Keep specifics internal to network
  - Aggregate internally where possible
  - Use iBGP for carrying customer networks
  - Use IGP for carrying backbone links
  - Reduces load on IGP
  - Improves customer's connectivity




# Aggregation – Good Example

- Customer link goes down
  - Their /24 network becomes unreachable
- /19 aggregate is still announced
  - No BGP holddown problems
  - No BGP propagation delays
  - No dampening by other ISP's

# Aggregation – Good Example

- Customer link returns
- Their /24 network is now reachable again
- The whole Internet becomes visible immediately
- Quality of Service perception!




# Aggregation – Bad Example

- Customer link goes down
  - Their /24 network becomes unreachable
- Their ISP doesn't aggregate the /19
  - /24 network withdrawal announced to ISP peers
  - Starts rippling through the Internet
  - Added load on Internet peering routers as network is removed from routing table

# Aggregation – Bad Example

- Customer link returns
  - Their /24 network is now reachable again from their ISP
  - The /24 network is announced again to ISP peers
  - Starts rippling through Internet
  - Load on Internet peering routers as network added to routing table
  - Some ISP's dampen, so they may decide not to add
  - The whole Internet may take 10 to 20 minutes to be visible again
- Quality of Service Perception?




# Aggregation – Summary

- First example is how everyone should do it
  - Adds to Internet stability
  - Reduces size of routing table
  - Reduces routing churn
  - Improves quality of service to customers.
- Second example is how not to do it
  - Too many do!

# “The Swamp”

- What is it?
- Cause
- Typical Problems
- Solutions
- Efforts

# Swamp?

- So called areas of poor or no aggregation
- 192/8 worst offender for routed networks
- 192/3 space uses 40844 networks, the rest is another 7000 routes only

| Block | Networks | Block | Networks | Block | Networks | Block | Networks |
|---|---|---|---|---|---|---|---|
| 192/8 | 6369 | 197/8 | 3 | 202/8 | 2024 | 207/8 | 2401 |
| 193/8 | 2185 | 198/8 | 4001 | 203/8 | 2761 | 208/8 | 1570 |
| 194/8 | 2646 | 199/8 | 3750 | 204/8 | 3917 | 209/8 | 1151 |
| 195/8 | 973 | 200/8 | 1011 | 205/8 | 2463 | 210/8 | 402 |
| 196/8 | 359 | 201/8 | 0 | 206/8 | 2858 | 211/8 | 0 |

# Swamp Cause?

- Early growth of Internet
- It was never a problem
- Small size of routing tables
- Few thousand networks only
- Lack of foresight by all

# Swamp Persists

- Lazy or technically unaware service providers
- Unannounced allocated networks
- Perceived market impact
- Technical solutions keep ahead of problem (faster routers, bigger memory, CIDR) so far!




# Solutions

- Don’t route 192/8 or other PA space
  - Encourage customers to renumber into your assigned space
  - Point out problems of using 192/8 space – quality of service!
- Aggregate!
- Don’t route anything smaller than your assigned block
- Be prudent when routing prefixes longer than /16 out of the former B space
- Encourage other ISP's to do the same
  - Don’t route/accept their “poor citizenship”
  - Consensus needs to be built among all ISPs

# Solutions – Special Cases

- Multihoming
  - Fragments address space
  - Think carefully about set up and requirements
  - Load balancing versus resiliency versus routing announcements.

Refer to MCI documentation

*http://infopage.mci.net/Routing*

# Efforts

- Bill Manning’s “drainage scheme”
- Registries
- ISPs
- The Internet needs YOU!

# Renumbering - motivation

- Same as motivation for aggregation
  - holes are bad
  - routing may not be optimal using “swamp” address space
- First time connection to the Internet requires
  - Legal address space
  - Sensible/practical addressing scheme
- New provider
  - Renumber from old provider’s address space to new
  - Helps reducing address space fragmentation
  - Improves routability

# Renumbering – how to?

- PIER – “Procedures for Internet Enterprise Renumbering” paper
  - *http://www.isi.edu/div7/pier/papers.html*
- Be aware of effect on essential services
  - e.g. DNS ttl lowering
- DHCP (dynamic host configuration protocol)
- Secondary IP addresses
- Not difficult but needs planning!

# Dampening

- Why do it?
- How to do it
- Current situation

# Dampening – why?

- Route flaps
  - Rapid withdrawal and reannouncement of routes
  - Requires significant router CPU
  - Leads to network instability
- Many ISP's now dampen at their network borders

# Dampening – how to

- Cisco has “bgp dampening” command in IOS
- Parameters
  - Defaults not so good
  - Recommendation of RIPE routing-wg *ftp://ftp.ripe.net/ripe/archives/routing-wg/**
- State dampening policy in AS object in routing registry - see AS1849 for example
- Dampening policy can be expressed in RPSL
- Case studies *http://www.cisco.com/warp/public/459/16.html*


# Dampening – Caution

- Be aware of potential problems
- Unreachability could be due to dampening, not disconnection
- Border routers need more significant CPU and memory
- Train your staff!

# NAT

- Network Address Translation
  *ftp://ftp.apnic.net/ietf/rfc/RFC1631*
- Used by firewalls or simpler “gateway” systems.
- Avoids the need for renumbering
- Helps conserve address space
- Concern about “concept” - ongoing discussion at IETF and elsewhere
- Mailing list *<nat@livingston.com>*

# Filtering Policies

- Filter announcements by peers
  - AS list
  - prefixes
- Only accept what is listed in routing registry
  - Avoids configuration errors
  - Avoids routing problems
  - Authorisation?
- Only announce what you list in routing registry!
- Keep routing registry and filters up to date!

# Filtering Policies - Prefix

- Don’t announce or accept *RFC1918 networks*
- Don’t announce or accept Martian networks

```
access-list 110 deny ip 10.0.0.0 0.255.255.255 255.0.0.0 0.0.0.255
access-list 110 deny ip 19.255.0.0 0.0.255.255 255.255.0.0 0.0.255.255
access-list 110 deny ip 127.0.0.0 0.0.0.255 255.255.255.0 0.0.0.255
access-list 110 deny ip 128.0.0.0 0.0.255.255 255.255.0.0 0.0.255.255
access-list 110 deny ip 172.16.0.0 0.15.255.255 255.240.0.0 0.15.255.255
access-list 110 deny ip 192.0.0.0 0.0.0.255 255.255.255.0 0.0.0.255
access-list 110 deny ip 192.0.2.0 0.0.0.255 255.255.255.0 0.0.0.255
access-list 110 deny ip 192.168.0.0 0.0.255.255 255.255.0.0 0.0.255.255
access-list 110 deny ip 223.255.255.0 0.0.0.255 255.255.255.0 0.0.0.255
access-list 110 deny ip 224.0.0.0 31.255.255.255 224.0.0.0 31.255.255.255
access-list 110 deny ip any 255.255.255.128 0.0.0.127
```

# Prefix Length Filtering

- Minimum prefix length filtering
  - e.g. Sprint /19, UUNET UK /24
- Reduces size of routing table
- Smaller networks more likely to flap

# Summary

- Looked at some of the methods of improving routability
- Business will benefit
- Quality of Service improvements
  - Happier customers
  - Less congested links
- Use the routing registries to document your policy
  - Poor or no documentation causes many of the problems today




# Future of the IRR




# Overview

- Open Issues
- Ways Forward
- Blue Sky

# Open Issues

- Why aren’t the tools used more today?
- What other tools should be available?
- Ignorance of the purpose of routing registries?
- No local routing registry?
- Security fears?

# Tool Usage

- Are they too complicated?
- Are there enough/too many?
- Are they too complicated/not sophisticated enough?

# Tool Availability

- Are there other tools which should be available? Which?
- Router able to “automatically” build configuration from routing registries?
- Router configuration seamlessly changeable when information in routing registry is changed?

# Tool Availability

- Should software available be offered as commercial package
  - Better bundled
  - Better debugged
  - Better supported
  - Better integrated
  - Training
- Most tools are freely available public efforts “for the good of the community”

# Routing Registries

***Belief that Internet will work without Routing Registries***

- It will, but for how long? Hours, days, weeks?
- Many ISP's rely heavily on data kept in registry
- Subset of tools available are being used on a daily basis

# Routing Registries

- Should each ISP run their own local routing registry
  - As mirror of their regional routing registry?
  - As a part of some global distributed routing registry?
- Software availability, scalability, data integrity, security, etc…?

# Training

- Is there enough training on “promoting routability”?
- Headcount requirement
  - Depends on organisation size
  - The will to be “good citizen” – too easy to be irresponsible
- Organisational awareness of the issues ➔ overall improved efficiency, quality of service and support

# Ways Forward

- APNIC routing registry
- AP region ISP's use it to
  - Register networks and routing policy
  - Configure border routers
  - Debug network problems

# Ways Forward

- Routing registries enhancements
- Feedback on tool enhancements
- Feedback to vendors on equipment configuration enhancements
- More training, more spreading the word

# Blue Sky

- Everyone uses routing registries
- Routers automagically configure from routing registry
- Everyone aggregates
- Prefixes longer than /19 aren’t routed apart from in special cases

# Summary

- Looked at the usefulness of routing registries
- Shown tools that help diagnose and solve routing problems more easily
- Made you aware of some of the issues affecting the Internet today
- **Shown how you can make a useful contribution to the seamless functioning of the Internet**

# Questions and Answers?

# IRR Reading List

1. RFC1786 “Representation of IP Routing Policies in a Routing Registry”
   ftp://ftp.apnic.net/ietf/rfc/rfc1786.txt
2. RPSL
   ftp://ftp.apnic.net/ietf/rfc/rfc2280.txt
   ftp://ftp.apnic.net/ietf/internet-drafts/draft-ietf-rps-transition-02.txt
   http://www.isi.edu/ra/rps/transition
   ftp://ftp.apnic.net/ietf/internet-drafts/draft-ietf-rps-appl-rpsl-00.ps
   Application of RPSL on the Internet
3. RATools
   Tools http://www.isi.edu/ra/*
   Mailing List <ratoolset@isi.edu>
   Subscribe ratoolset <majordomo@isi.edu>

# IRR Reading List

4. PRIDE
   - Slides ftp://ftp.ripe.net/pride/docs/course-slides
   - Guide ftp://ftp.ripe.net/pride/docs/guide-2.0txt.{ps}.tar.gz
   - Tools ftp://ftp.ripe.net/pride/tools/*
5. IRR authorization/notification
   - ftp://ftp.ripe.net/ripe/docs/ripe-120.txt
6. RADB pointers
   - http://www.ra.net
   - http://www.ra.net/.faq.htm
7. ISP run RR User documents
   - http://infopage.mci.net/Routing

# Useful URL's & Reading

1. BGP Dampening
   - http://www.cisco.com/warp/public/459/16.html
   - http://www.ripe.net/mail-archives/routing-wg/current/msg00010.html
     European recommendations for route flap dampening
   - ftp://engr.ans.net/pub/slides/nanog-feb-1995-route-dampen.ps
2. Routing Discussion
   - http://www.ripe.net/wg/routing/index.html
3. AS numbers
   - ftp://ftp.apnic.net/ietf/rfc1930.txt
     Guidelines for creation, selection, and registration of an AS
4. NANOG archive
   - http://www.merit.edu/mail.archives/html/nanog/maillist.html
5. CIDR
   - ftp://ftp.apnic.net/ietf/rfc/rfc{1517,1518,1519}.txt
   - http://www.ibm.net.il/~hank/cidr.html
   - ftp://ftp.uninett.no/pub/misc/eidnes-cidr.ps.Z
     Network addressing when using CIDR


# Useful URL's & Reading

6. Address Allocation and Private Internets
   ftp://ftp.apnic.net/ietf/rfc/rfc1918.txt
7. Traceroute server repository
   http://www.boardwatch.com/isp/trace.htm
8. ISP Tips
   http://www.amazing.com/internet/faq.html
9. BGP Table
   http://www.telstra.net/ops/bgptable.html
   http://www.employees.org/~tbates/cidr.hist.plot.html
   http://www.merit.edu/ipma/reports
10. Route server views
   http://www.caida.org