# Architecting the Network

Geoff Huston
Network Technical Manager
Telstra
gih@telstra.net

# Architecture and Design

- Definition of Architectural Principles
- Translating Architecture into a Design
- Generating an Engineering Plan
- Implementing the Network
- Operational Considerations
- Policy Considerations

# Personal Experience

The Australian Internet - AARNet

- Constructed in May 1990
- Initially 45 client sites (now 450)
- Modest implementation budget initially $US 1.2M)
- Modest initial staff resources (2)
- June 95: $10M p.a. with 5 staff

Telstra

- commenced July 95
- telco Internet provider

# Architectural Principles

Assumption:

- Implementation of Public Infrastructure on a National Scope

Design issues will vary for commercial and/or corporate networks

# Architectural Principles

- Simplicity
- Functional Adequacy
- Affordability
- Implementable today
- Designed to meet actual end client requirements
- Uses (and develops) local expertise
- Where feasible uses locally available components

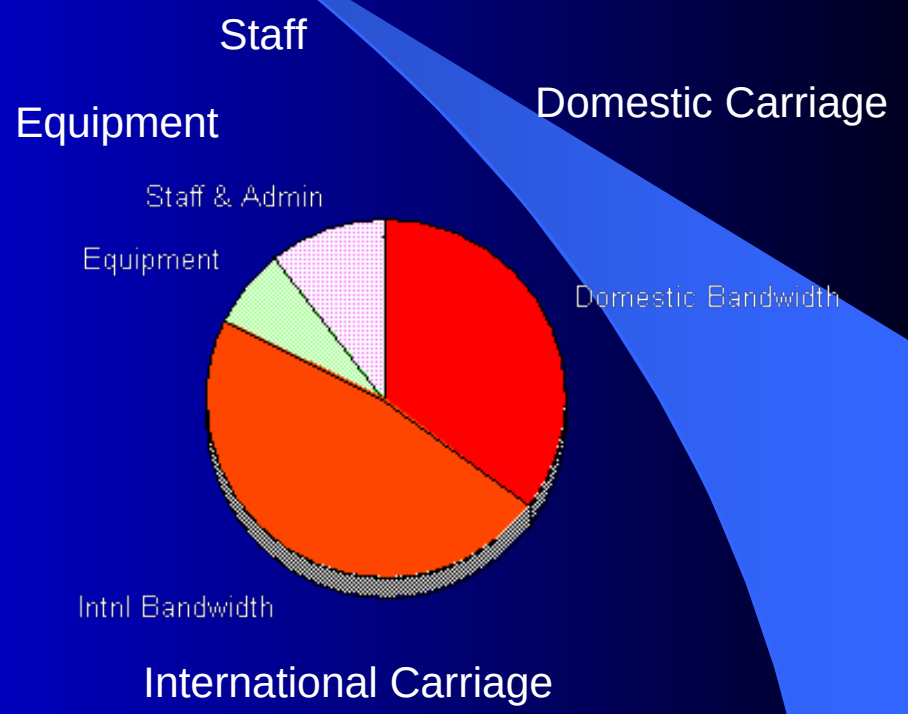
# Architectural Principles

Simplicity is the key attribute of any network architecture

Diverse, complex and uncoordinated architectures result in very high implementation and operational costs, and are resistant to subsequent incremental engineering.

# Design Considerations

- Design objective is to minimise costs and maximise capability
- Unless you are a telco bandwidth lease will dominate all other cost elements
  - even then it will probably dominate all other costs!
- The unit cost of bandwidth is the major design parameter



# Design Considerations

- Implementation and operational cost
- Network performance
- Operational reliability
- Manageability
- Extensibility

# Design Strategy

- Affordable capacity defines delivered service quality
- Solve today's problems first
- Define a service which matches current needs before matching future expectations

# Design Components

- Internet Transport Service Core
  - Leased circuits
  - Routers

# Design Components

- Access Services
  - Routers
  - Modems

# Design Components

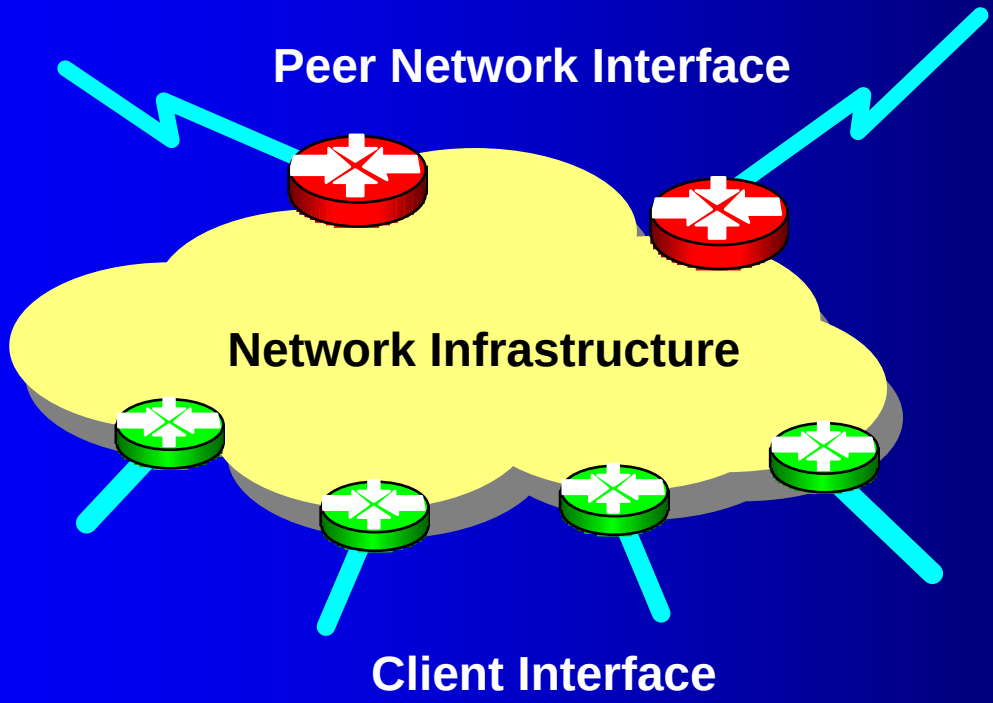
- Application Service Elements

  Servers:

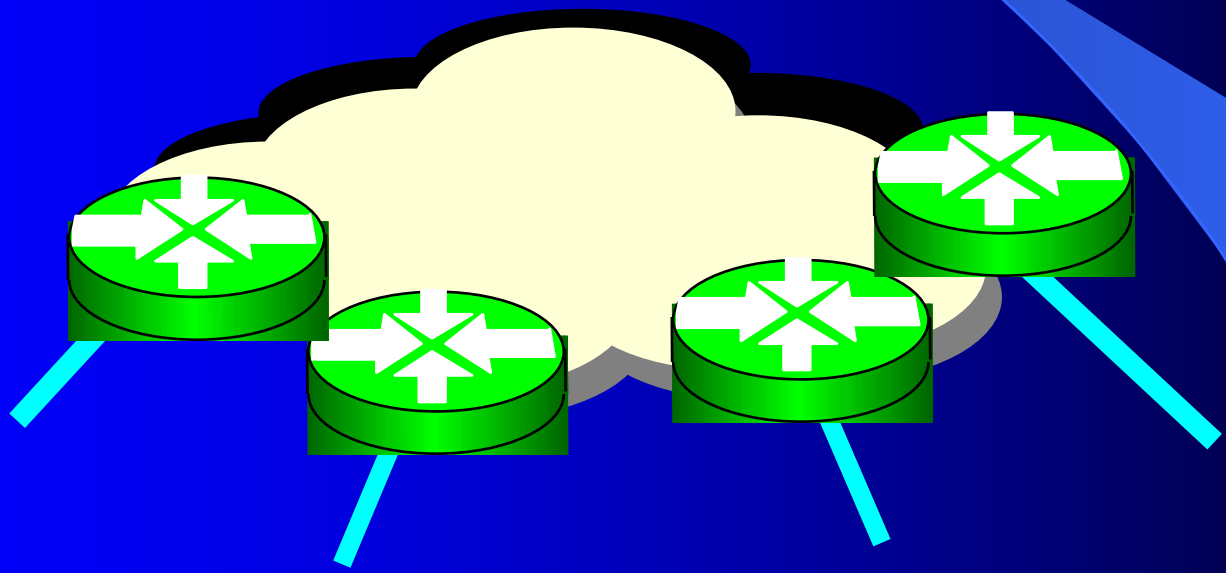
  - DNS
  - USENET
  - EMAIL
  - WWW
  - FTP
  - ARCHIE

# Abstract Design

Router Interface design model



# The Client Interface



# The Client Interface

- Single Homed Clients
  - Client uses single service provider offering "default" service
  - Client's networks are advertised via provider

# The Client Interface

- Use of RIP as Network / client boundary routing protocol?
  - ✓ simple
  - ✓ widely implemented
  - X **NOT** applicable in all cases
  - X no support for classless address exchange

# The Client Interface

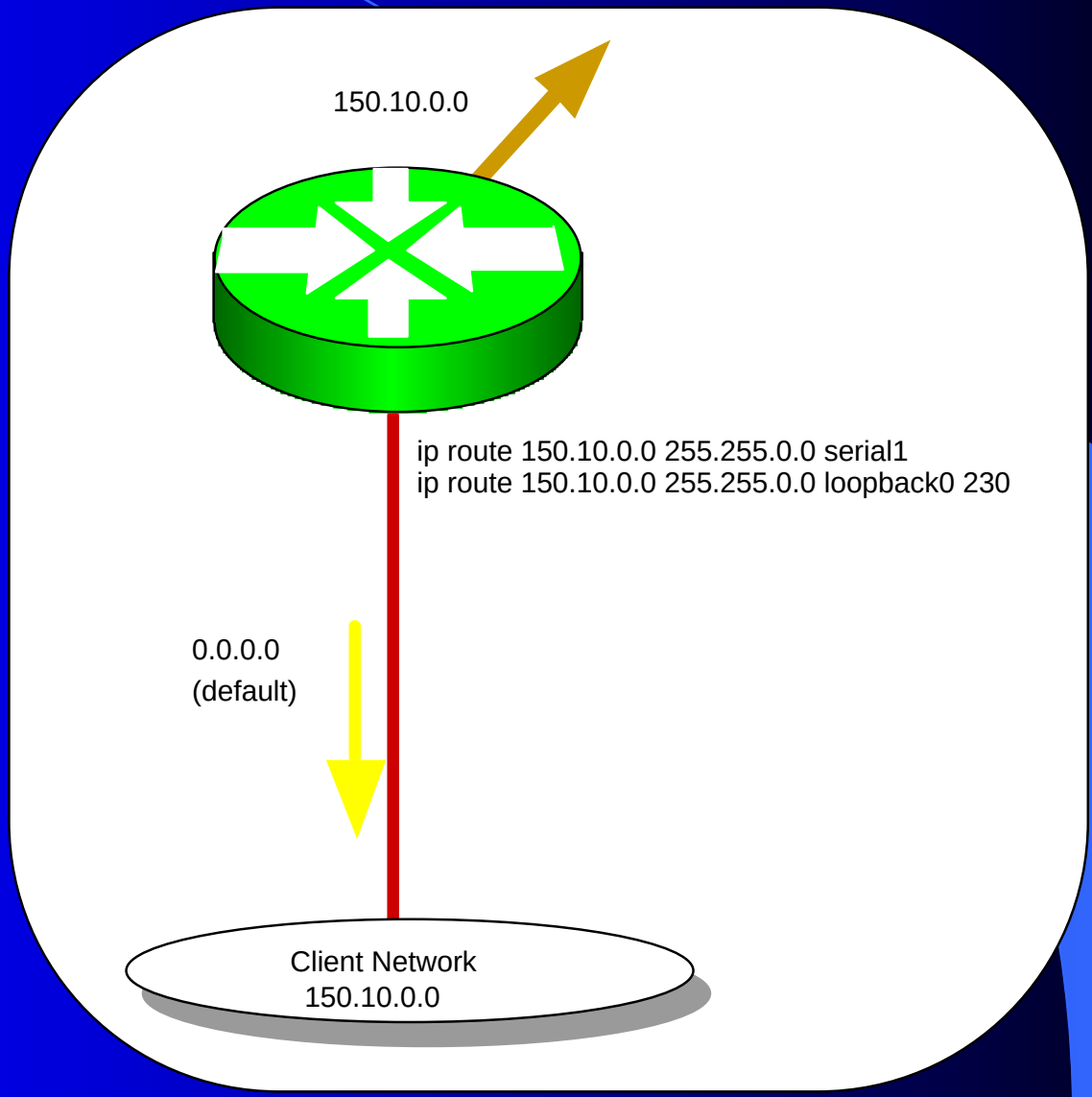
- Use of STATIC ROUTES as Network / client boundary routing protocol?
  - ✓simple
  - ✓widely implemented
  - ✓can support classless address advertisements
  - O requires careful design to scale
  - X cannot support dynamic multi-homed connections

# The Client Interface

- Use of Classless Client boundary routing protocol?
  - EIGRP - propriatary B-F Distance Vector
  - OSPF - IETF Std Link State
  - RIPV2 - IETF Std B-F Distance Vector
  - BGP4 - IETF Std Inter Domain Routing Protocol

# The Client Interface

Static routing



# The Client Interface

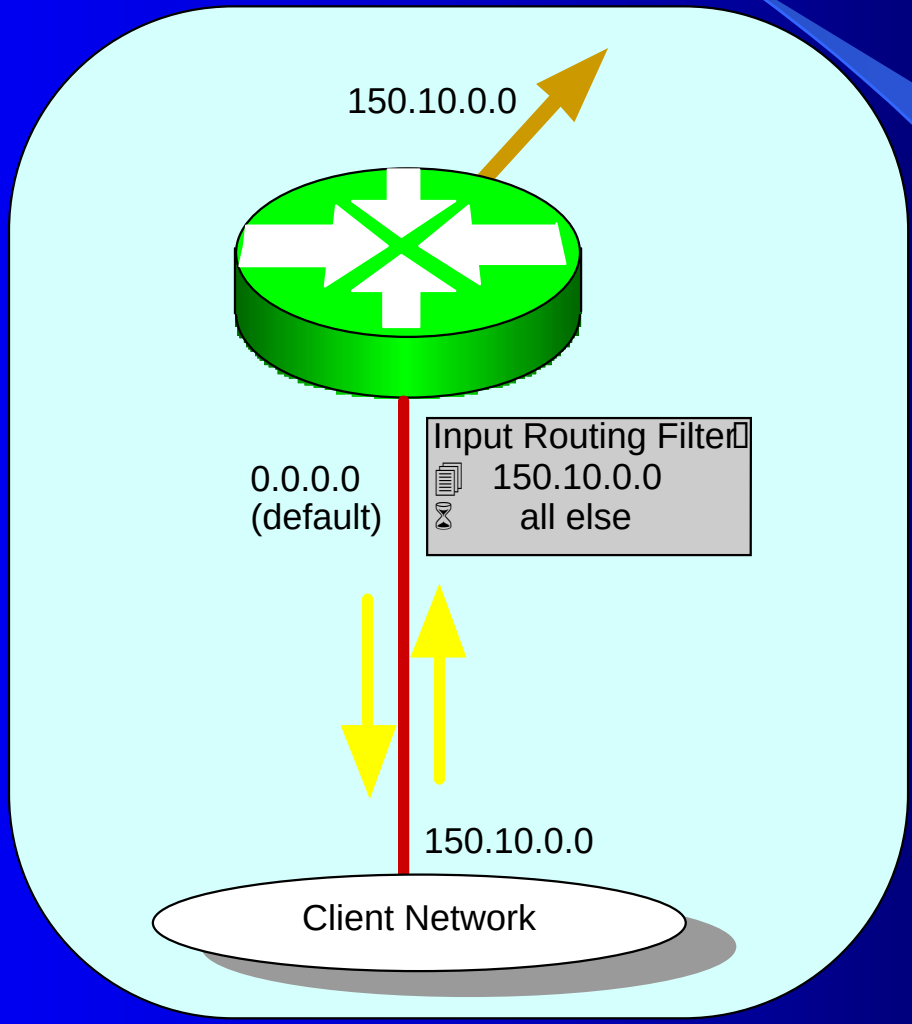
Dynamic Routing Guidelines

- Use of inbound routing filters to preserve network integrity
  - prevent client advertising bogus routes
  - preserve integrity of client network

# The Client Interface

- Dynamic Routing Guidelines
  - Use of outbound static default route to simplify client routing
    - stability of presented service
    - simplicity of presented service
    - client sees only an external default path

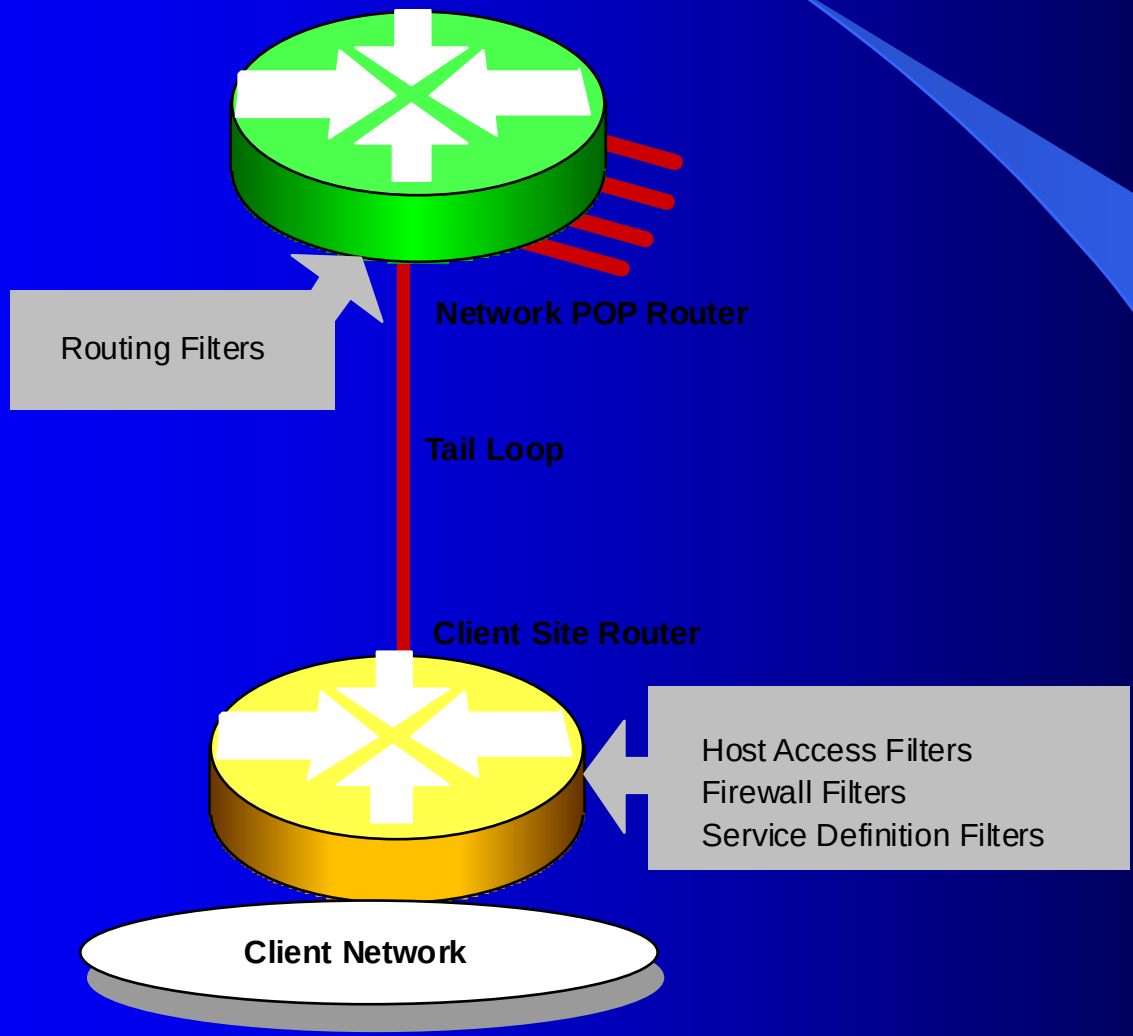
# The Client Interface



# The Client Interface

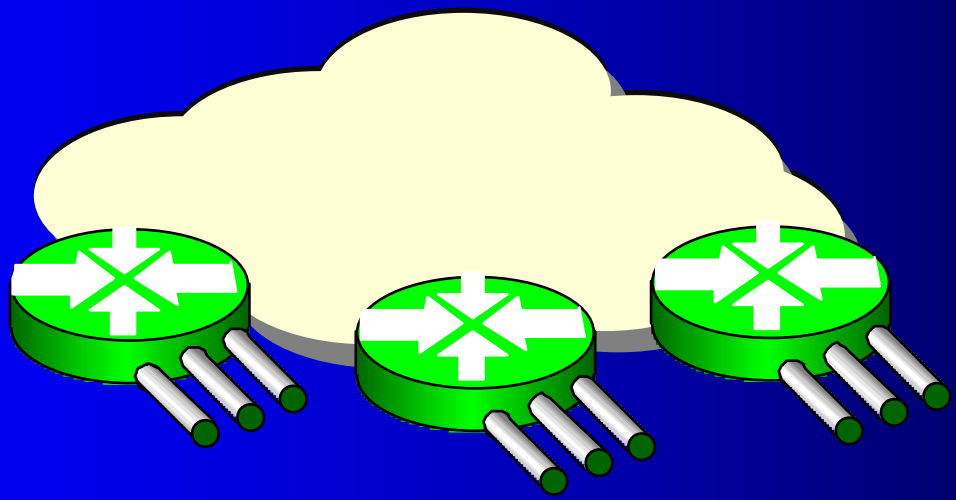
- Clear demarcation of boundary between client and network is required for consistency of service
- Single demarcation model is required for the network to ensure manageability and operability.
- The network service should never transit a client network

# The Client Interface



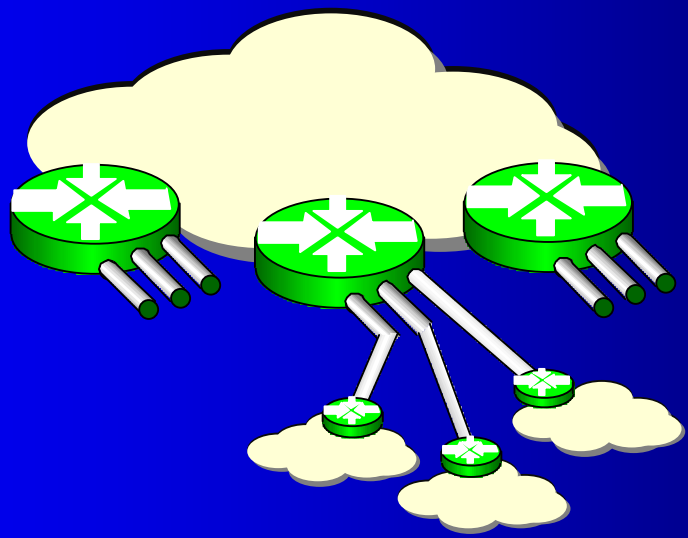
# The Client Interface

- The POP Access Model
  - Client is responsible for CPE router and tail loop
  - Network Provider provides router attachment points at a number of locations
  - Network Boundary located at POP interface



# The Client Interface

- The Comprehensive Service Model
  - Network provider installs and operates CPE router and tail loop
  - Network provider attaches to client LAN
  - Network Boundary located at LAN attachment point



# The Client Interface

- The Confused Model
  - Network Provider installs tail loop
  - Network Provider installs router interface card in client router
  - Client and network provider operate client router simultaneously

# The Client Interface

- POP or end-to-end model depends on:
  - telco bulk purchase tariff discounting
  - router vendor bulk purchase discounting
  - staff availability
  - client expertise levels
  - defined service level
- Client Site service model is preferable from a commercial perspective

# The Client Interface

- You can do both POP and end-to-end
  - as long as all routing integrity is maintained within the POP locations for all clients
    - The integrity of the system is maintained within the set of "core" routers

# The Client Connection

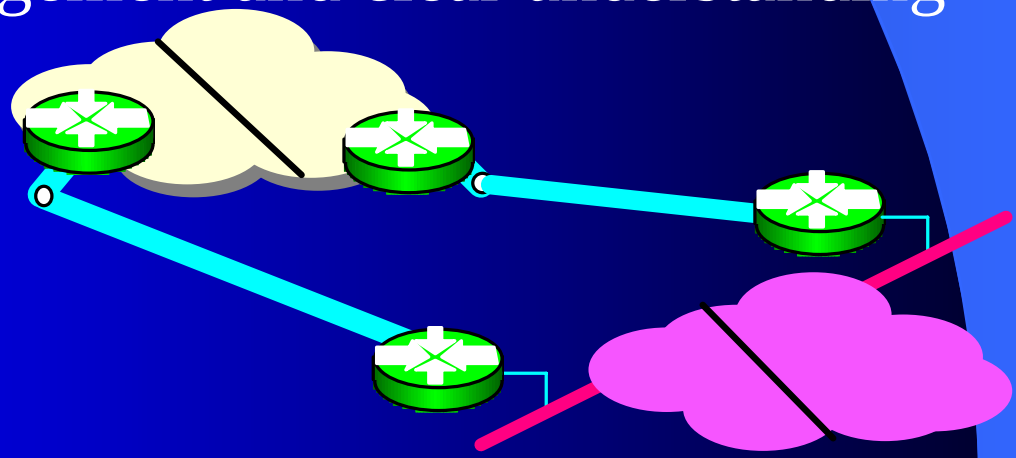
- Routers provide:
  - security capability
  - management capability
    - routing management
    - traffic management
    - service management
  - efficiency
  - integration

# The Client Connection

- SLIP / PPP implementations in hosts
  - cheap!
  - Capital price differential between hosts and router is small
  - Operating cost is higher using hosts as routers
  - use as single end host access system

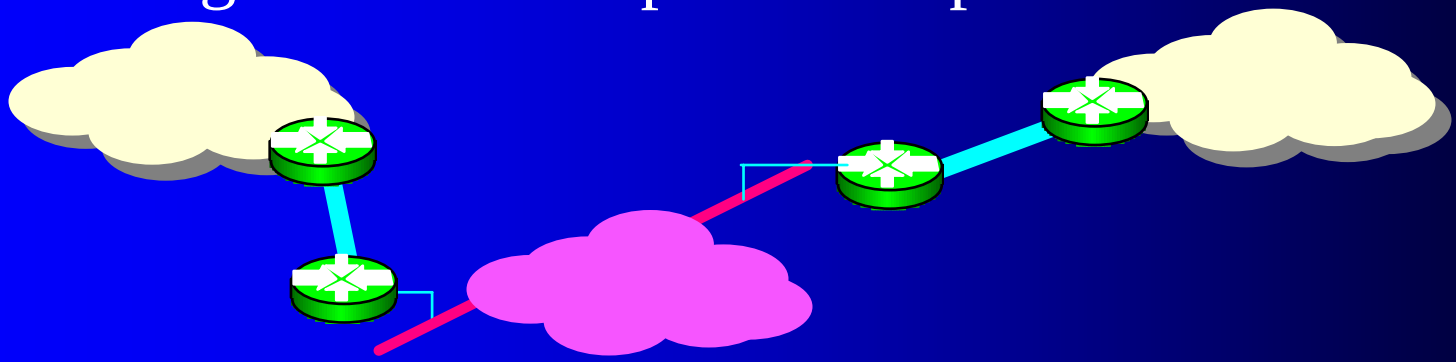
# Routing to the Client

- Multiple client interfaces
  - bifurcation of client and provider network - multiple default paths
  - asymmetric routes can be generated
  - client network internal breakage causes black hole routing
- requires careful management and clear understanding of the routing issues



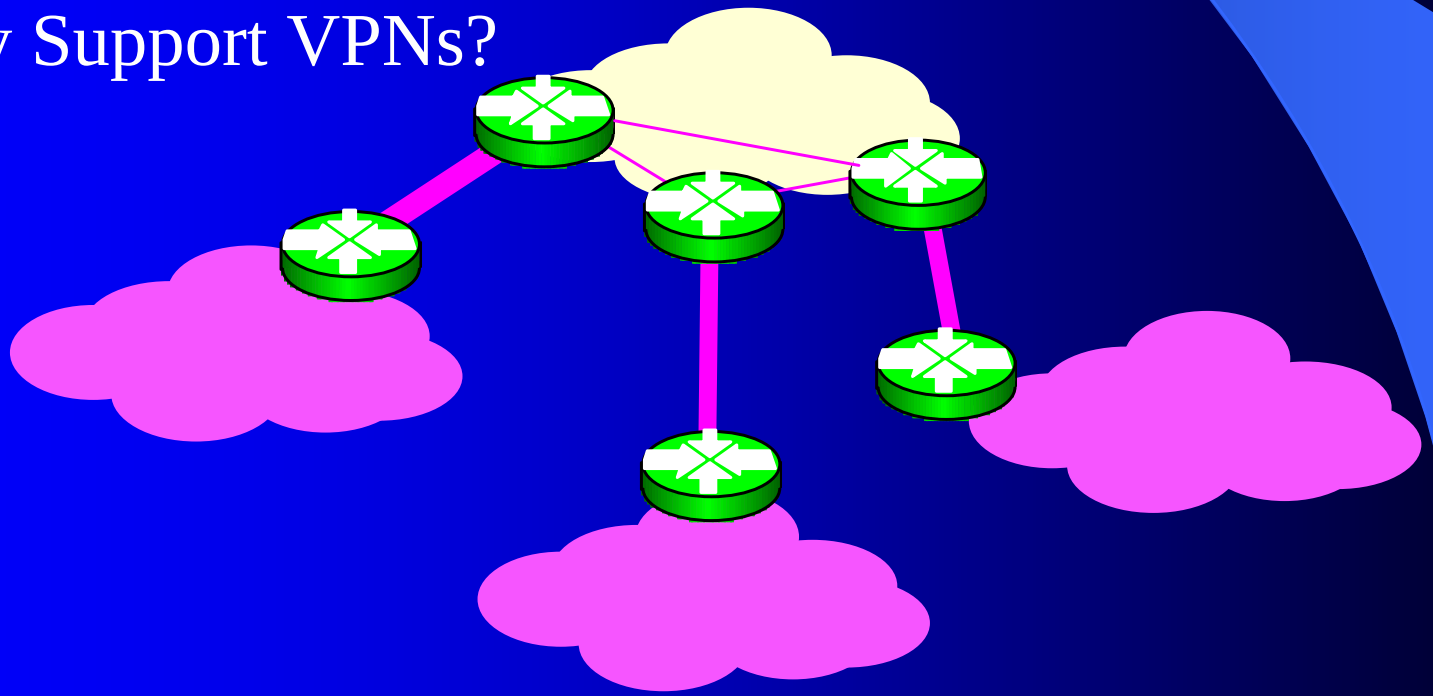
# Routing to the Client

- Multiple providers
  - Only one provider can provide "default"
  - other connected providers must resort to explicit provision of routes to enumerated networks
  - All providers must ensure that the client is not used as a transit facility through explicit route management on the part of all providers

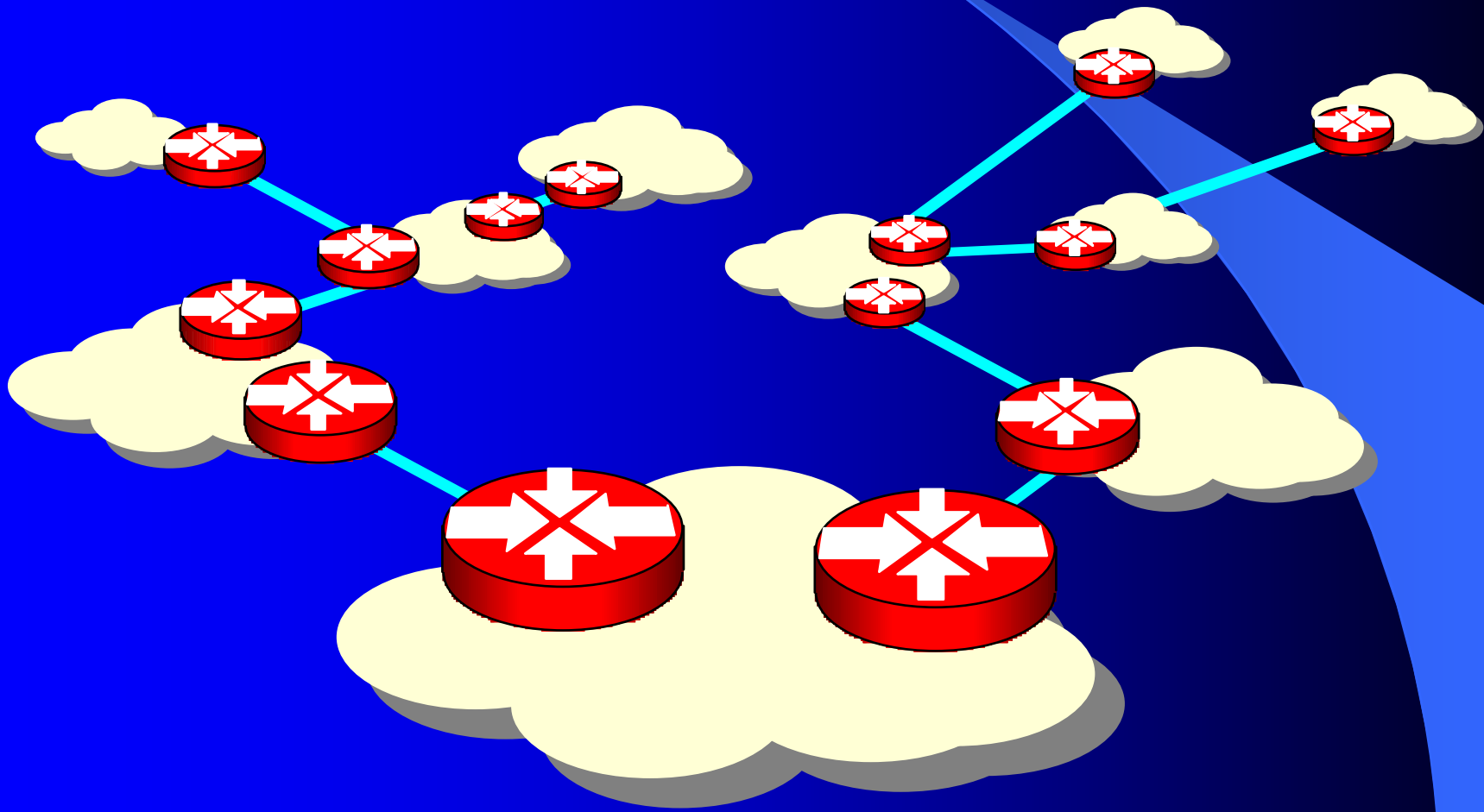


# Distributed Client support

- VPN architecture issues
  - VPNs via filtering
  - VPNs via tunnelling
- Why Support VPNs?



# Network Peer Interface



# Network Peer Interface

- Who is my peer?
  - Differentiating between:
    - client network (they pay me!)
    - service provider network (I pay them!)
    - peer network (we pay each other!)
- There are no Internet mechanisms to determine who is a peer network

# Network Peer Interface

- Where do I peer?
  - Onshore 1:1
  - Onshore at a layer 2 exchange
  - Offshore via Service Provider
  - Offshore at a layer 2 exchange

# Network Peer Interface

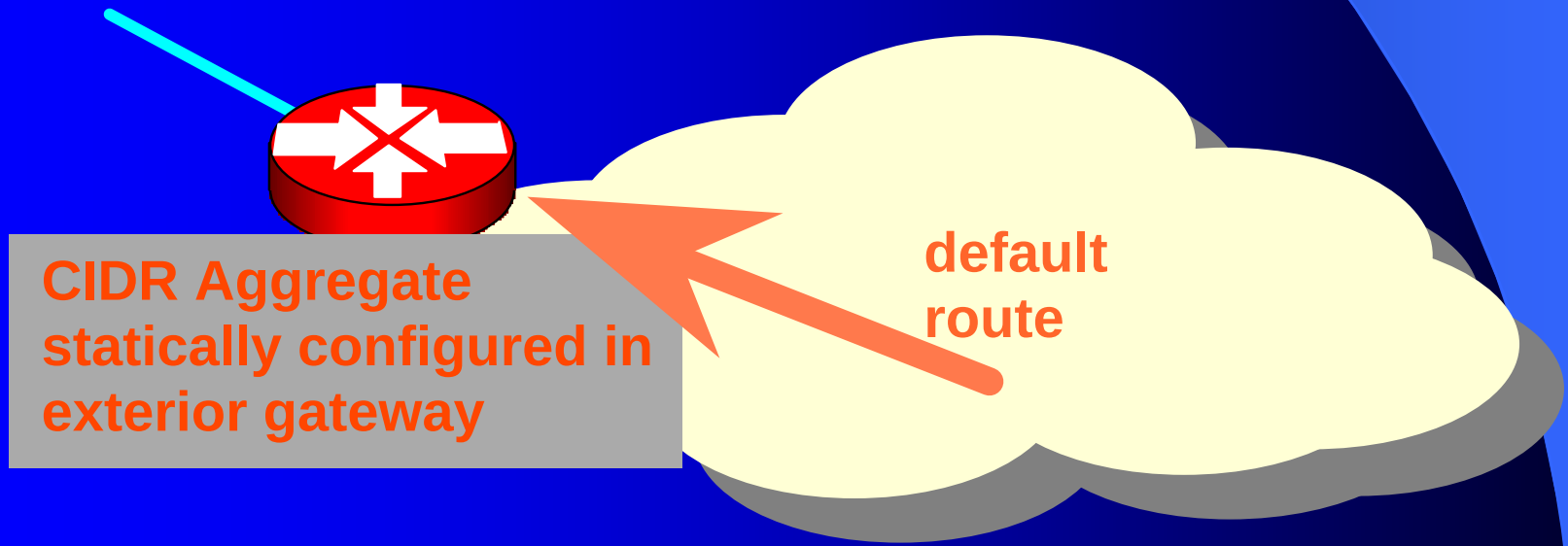
- Routing Considerations
  - Export routes via BGP4 using CIDR
  - Import routes using whatever works easily!
- Operational Considerations
  - Minimise bandwidth used by routing
  - maximise operational stability

# Network Route Management

- Obtain registered Autonomous System number (AS)
  - from IANA or your Regional Registry
- Generate aggregate mask which covers all announced networks
- Announce CIDR aggregate to peer via BGP4 session

# Single Exterior Peer

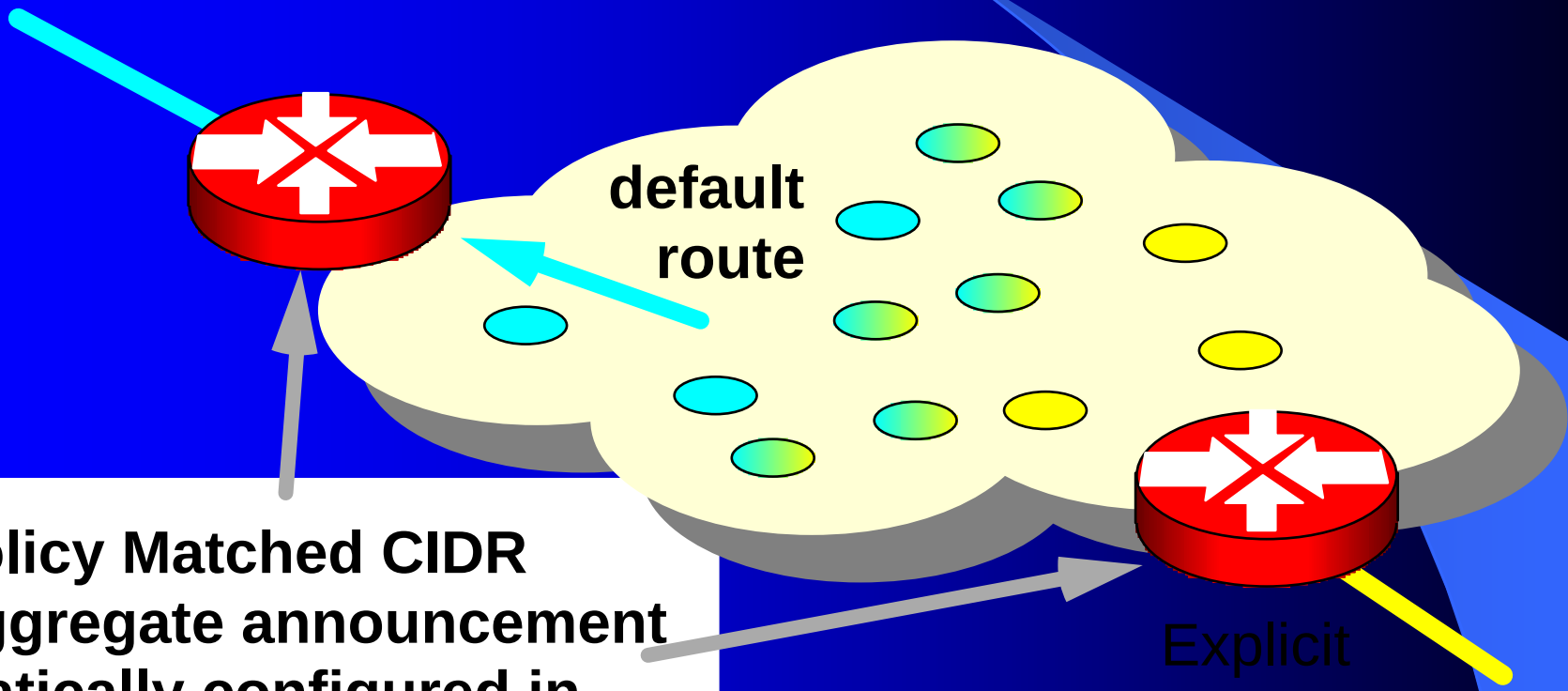
- Announce local nets via CIDR aggregate using BGP4
- Synthesise static default route directed to exterior peer gateway



# Multiple Exterior Peers

- Externally Imposed Policy differentiation
  - For example:
    - Academic & Research peer external network
    - Commercial peer external network
- Routing is Destination address-based - not source address
  - Default route based on best policy match
  - Explicit routes are imported from other network peers
  - Traffic path based on destination net - not local source policy

# Multiple Exterior Peers

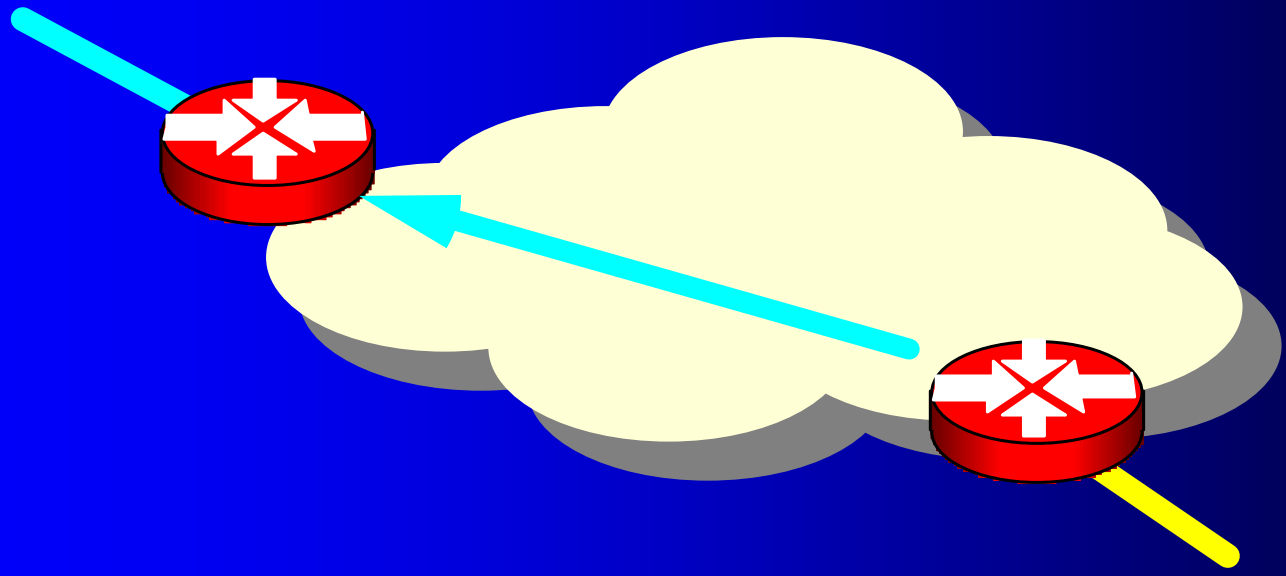



**Policy Matched CIDR Aggregate announcement statically configured in exterior gateways**

Explicit routes imported via BGP4

# Multiple Exterior Peers

- Transit Arrangement
  - Importation of transiting AS network numbers
  - Announcement of transiting networks via AS path mechanism



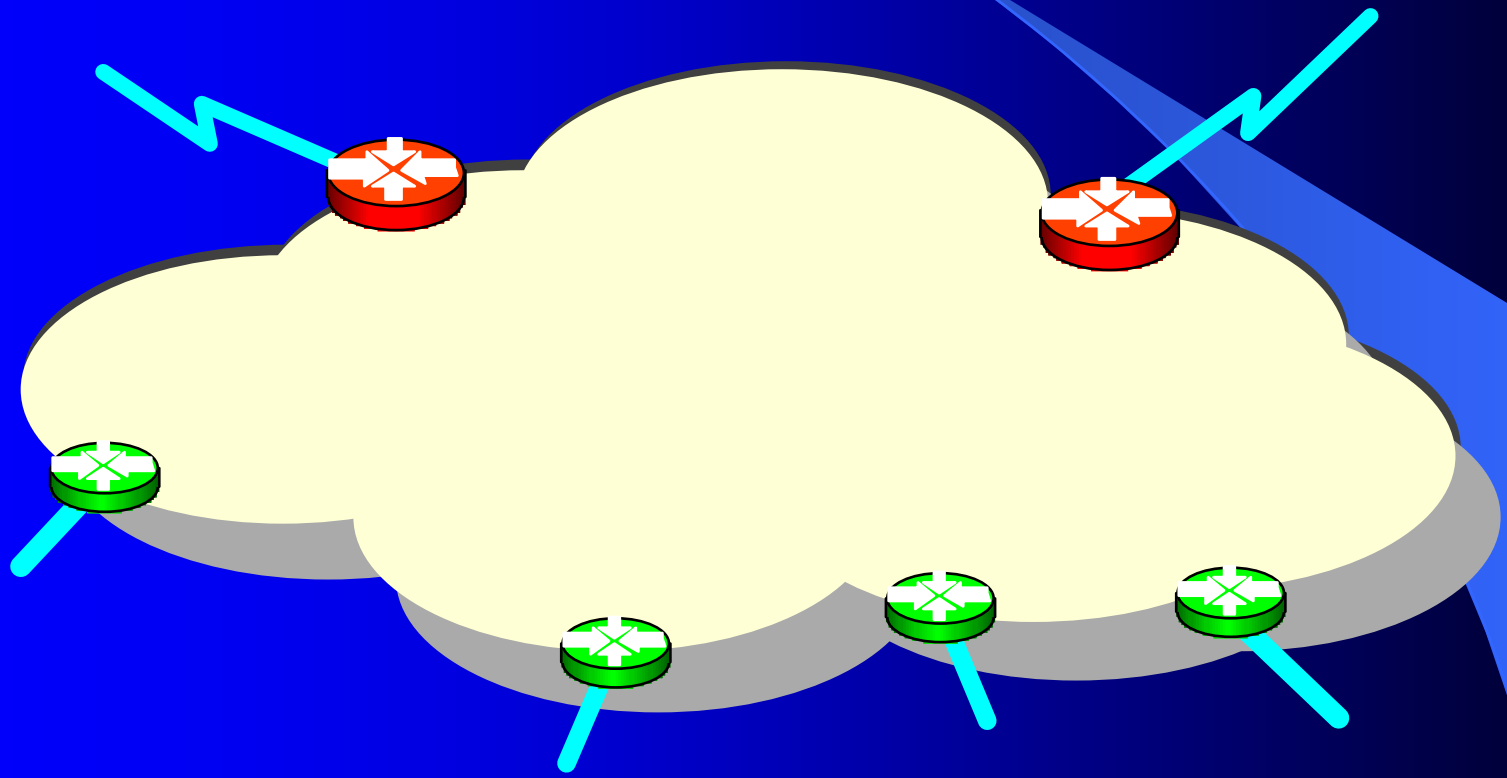
# Exterior Peering

- Importing a default route is cost effective and highly efficient as long as there is a suitable policy and capability match with the peer
- Default-less routing is expensive, time-consuming, and can be unstable
- Default-less routing allows greater levels of self-determination of policy - with an operational cost

# Exterior Peering

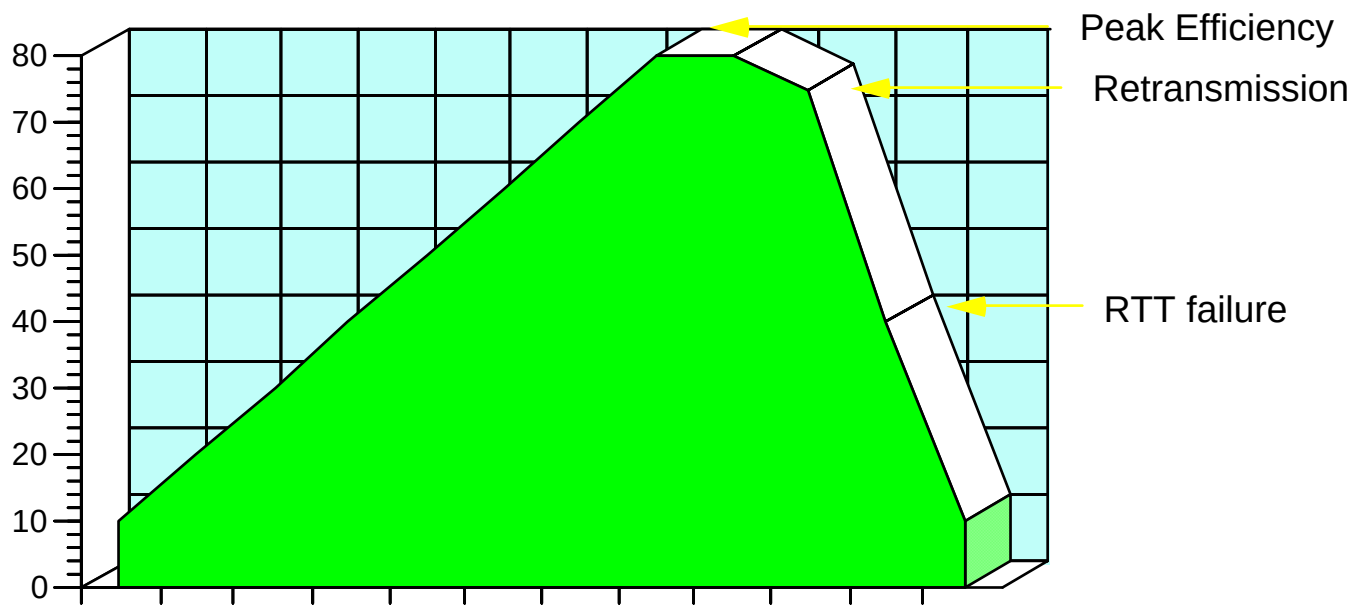
- Use a simple model initially:
  - Single exterior peer
  - Derived default route
  - Announce CIDR aggregate to peer

# Network Infrastructure



# Network Infrastructure

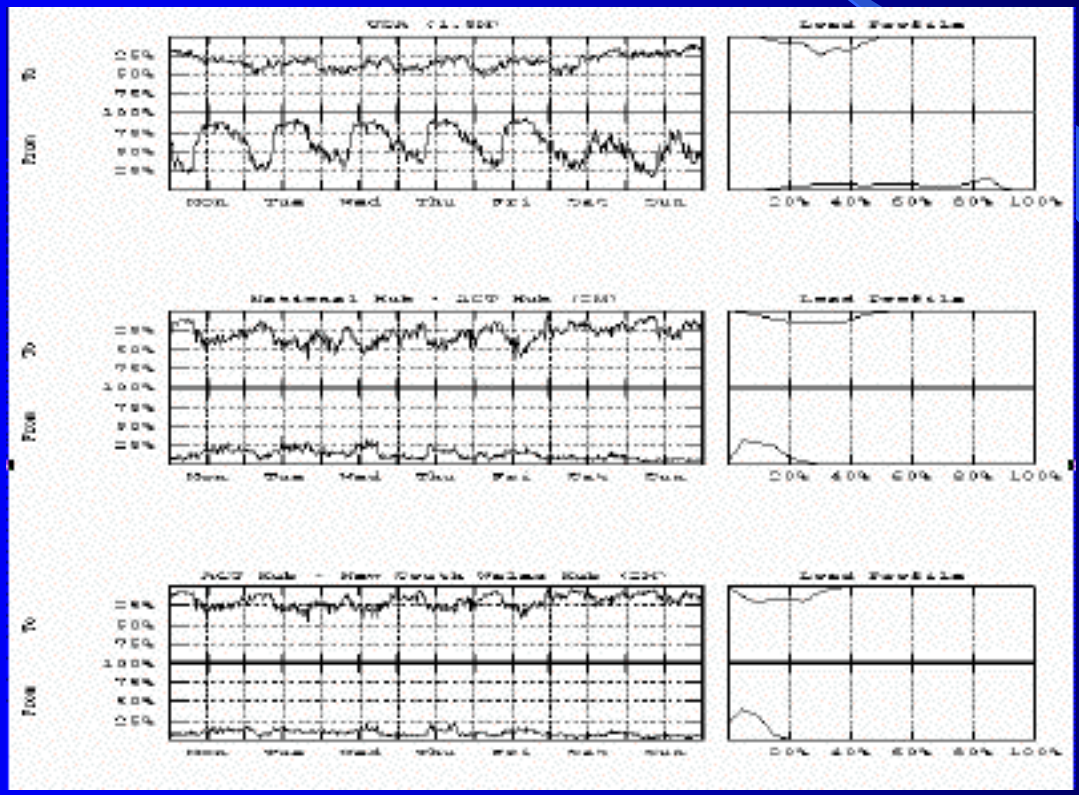
- Bandwidth is a coarse control tool



# Network Infrastructure

- Engineer capacity for peak demand periods
- Understand end-to-end flow patterns
- Attempt to avoid sustained (> 15 minutes) acute congestion on any link
- Constantly monitor bandwidth utilisation and flow patterns
- Generate trend patterns and plan accordingly

# Network Infrastructure



# Network Infrastructure

- Communications techology choices:
  - Dedicated Facilities
    - point to point leased circuit
    - point to point radio
  - Common Switched Facilities
    - X.25
    - Frame Relay
    - SMDS access
    - ATM

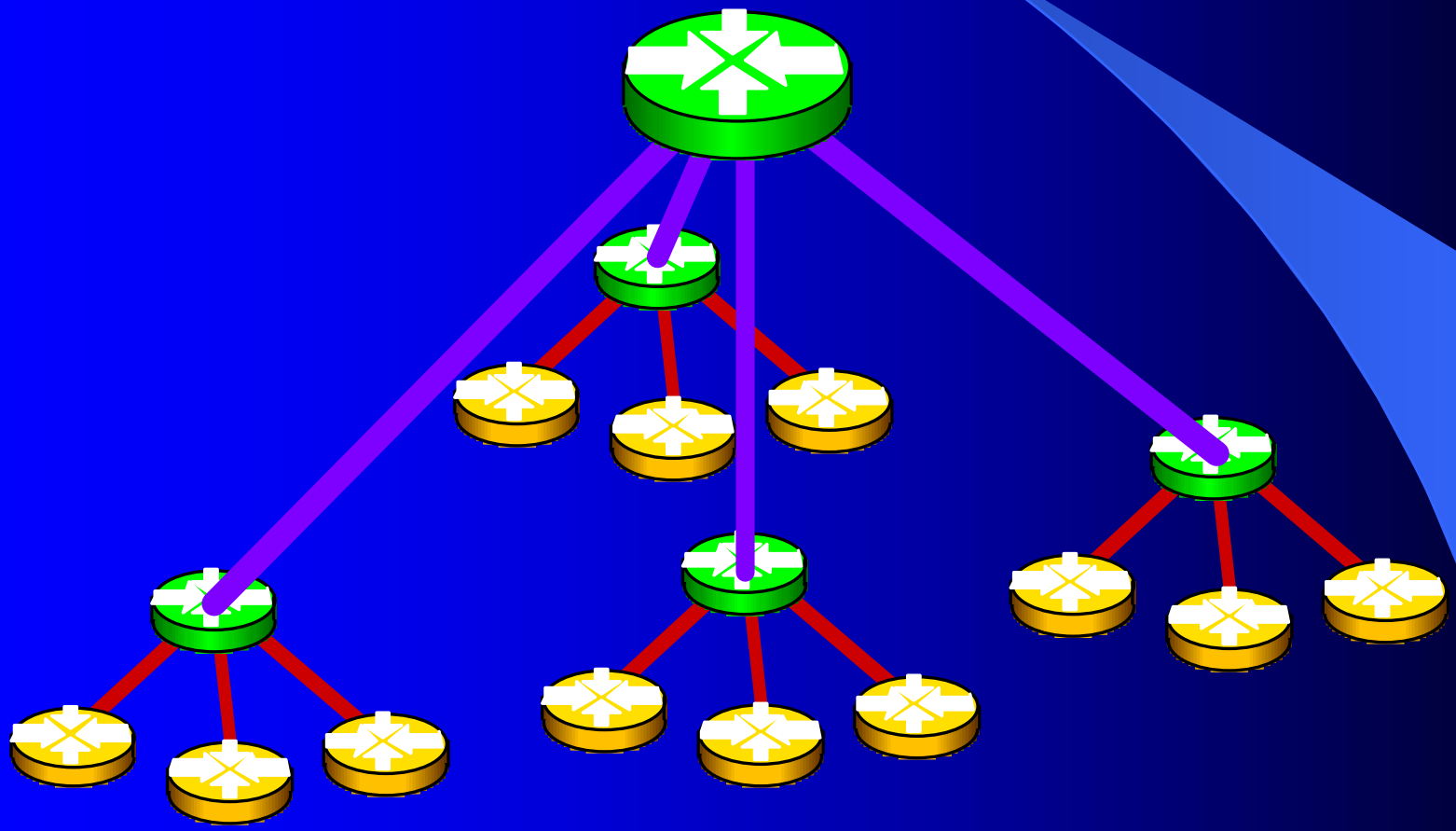
# Network Infrastructure

- Leased circuit design
  - Performance
  - Reliability
  - (In)Flexibility
  - Cost

# Network Infrastructure

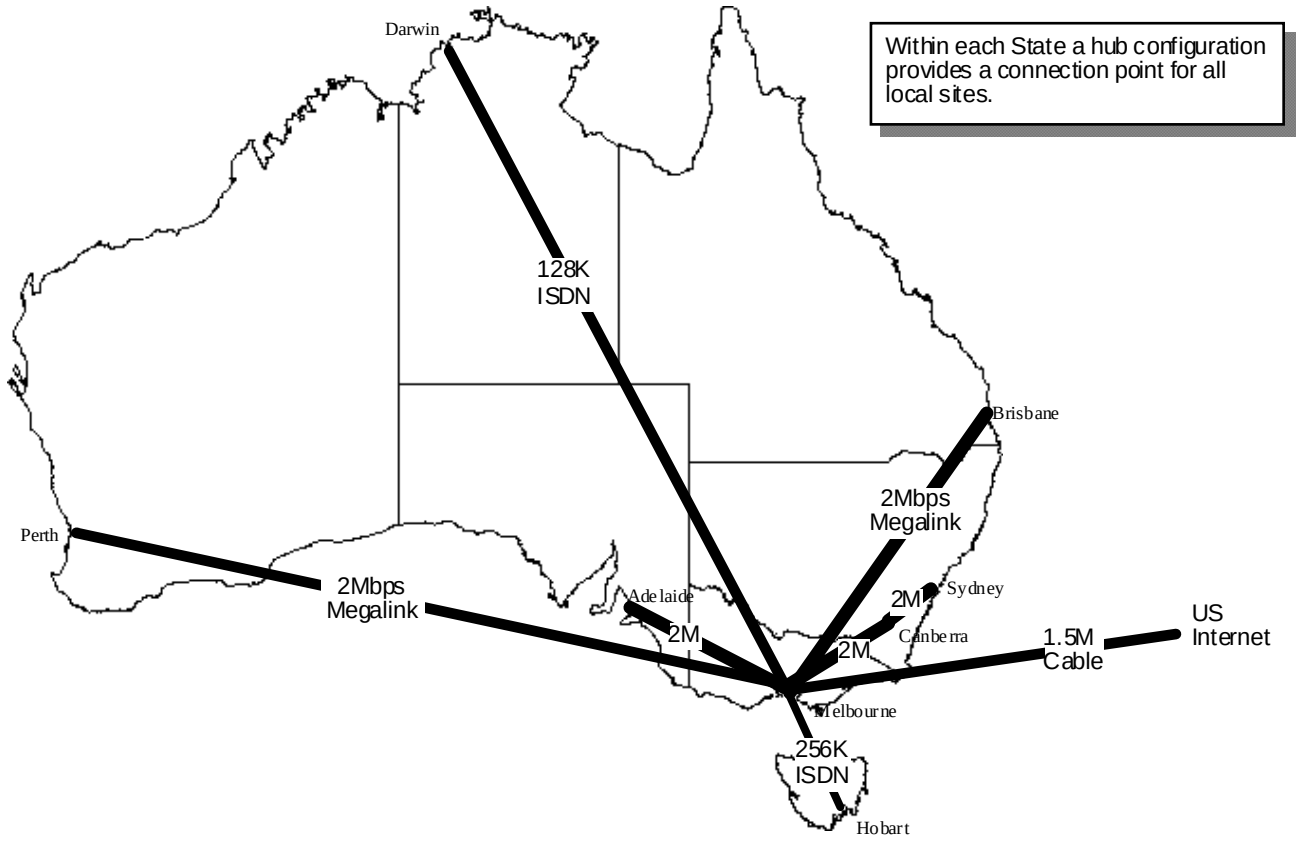
- Hierarchy (Star) Topology
  + Minimal Cost
  + Simple Topology
  + Maximal efficiency
  - Critical points of failure

# Network Design



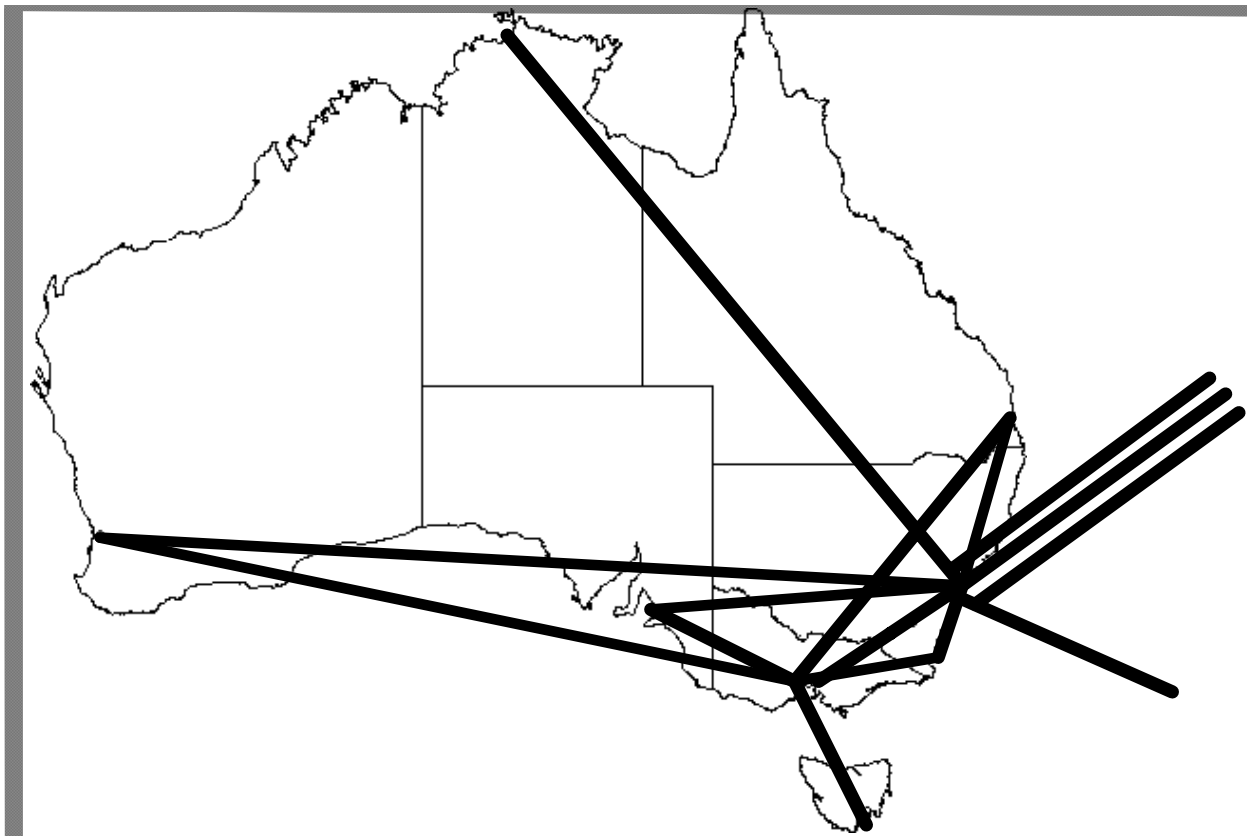
# Network Design

**The Australian Academic and Research Network**



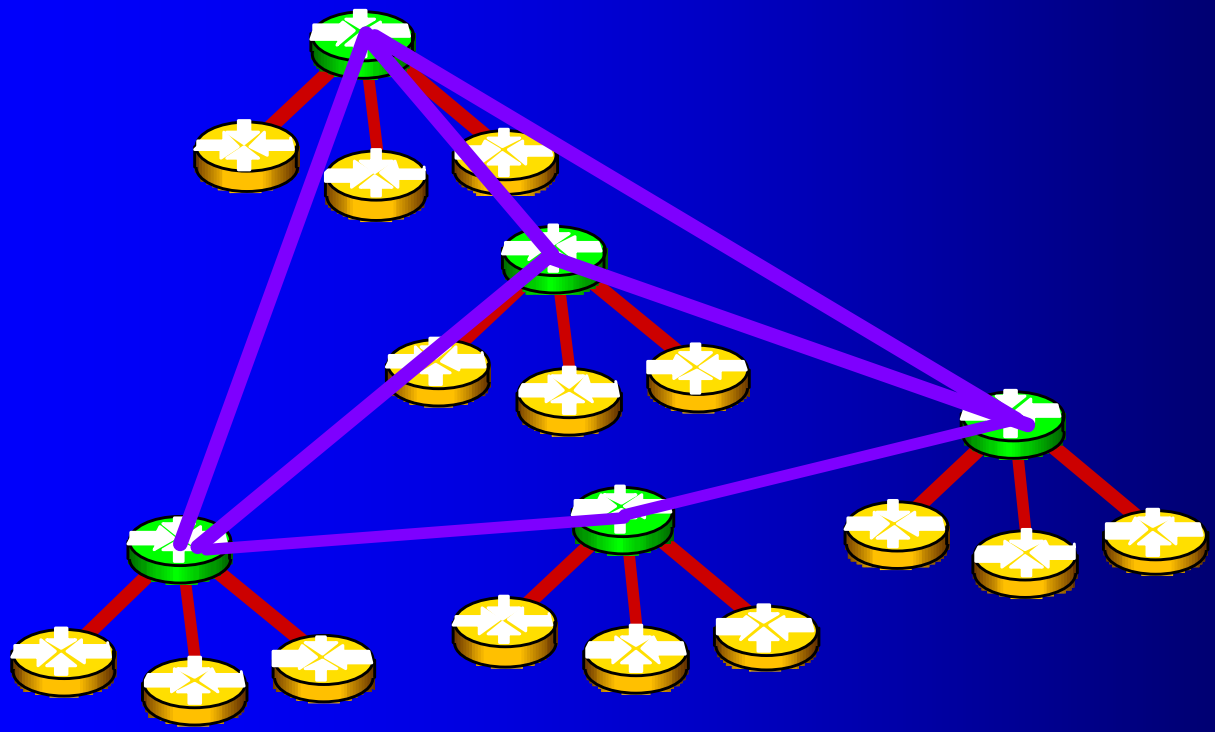
# Network Design

Telstra Internet Network - June 1996



# Network Infrastructure

- Mesh Topology
  + Resiliancy against link or site failure
  - Higher communications lease cost



# Network Infrastructure

- Hybrid - Resiliancy via Dial-on-Demand
  - Establish backup circuits using ISDN, X.25 or modems
  - Issue of matching backup capacity against primary capacity

# Network Infrastructure

- Access to common switched services
  - X.25
  - Frame Relay
  - SMDS
  - ATM

# Network Infrastructure

- Switched Network Design Issues
  - Delivered Service contract (and enforceability)
  - Tariff issues
  - Dynamic vs static virtual channels
  - Efficiency
  - Congestion behaviour

# Network Infrastructure Design

- “Core” routers driving major internal trunk lines
- “Boundary Routers” providing client connection point
- “Access Routers” used on client site

# Routing within the Network

- Choosing an Interior Routing Protocol
  - RIP (V2)
  - OSPF
  - (E)IGRP
- Classless routing protocols are now essential for this domain

# Routing within the Network

- Integrity and stability of the routing domain is essential
- The routing protocol is not used to learn new routes
  - authenticity of the route
  - security and integrity of the routing domain
- The routing protocol is only used to promulgate the route within the network

# Routing within the Network

- Use of static routes to dampen down route flaps
  - A  transition of a route (up / down) causes all routers to undertake a cache dump and rebuild the route table from received routing information
  - Route Flapping can destroy network performance
- default is synthesised to all network clients through presentation to the client of a static default route

# Service Management

- Use of router facilities to define service levels
  - form of bandwidth management:
    - transmission priority lists
    - bandwidth class scheduling
  - Can improve performance of defined services under load
- Effectively such measures are within the area of "congestion management"
  - The intent is to provide resources to some services when the bandwidth resource is under load

# Service Management

- Priority Example:
  - High priority on packets to and from port 23 (telnet) and 513 (rlogin)
  - Low priority on packets to/from port 119 (net news)
- Class Scheduling
  - Allow telnet and rlogin up to 50% of available bandwidth when under contention
  - Allow nntp up to 2% of bandwidth when under contention
- Class Scheduling is a more stable approach to congestion management

# Network Operation

- Management of IP numbers is critically important:
  - Ensure network number registration information is accurate
  - Publish correct IP numbers to external network peers
  - Ensure that correct IP numbers are routed
  - Ensure that end clients are using correctly allocated numbers

# Operation of a Service

- Service Quality is achieved by a match of capability to demand:
  - technical capability to carry user load
  - financial capability to provide adequate resource

# Stitching it all Together

- roll out
- shipping
- end site training / interaction

# Operational Management

- All active elements of the network centrally managed
- SNMP used as platform for management
- routers are the central component of operations

# Operational Management

- snmp traps used for exception reporting
- never underestimate the power of ping !
- traceroute - the route reporter
- dig - DNS diagnosis

# Operational Management

- Each management environment has particular requirements
- Routers are the most reliable network element
- carrier services are the greatest point of vulnerability
- careful router configuration will isolate LAN faults

# Operational Management

- Internet issues - working within a larger muti-provider environment:
  - NOC obligations
  - trouble ticket management

# Reporting

- Goals of data collections and reporting:
  - operational management
  - trend analysis of traffic volumes
  - monitor levels of delivered service
  - monitor usage patterns
  - marketing material!

# Reporting

- Balance of cost of data collection and analysis against benefit of resultant data sets
- Data collection points affect ability to gather data

# Reporting

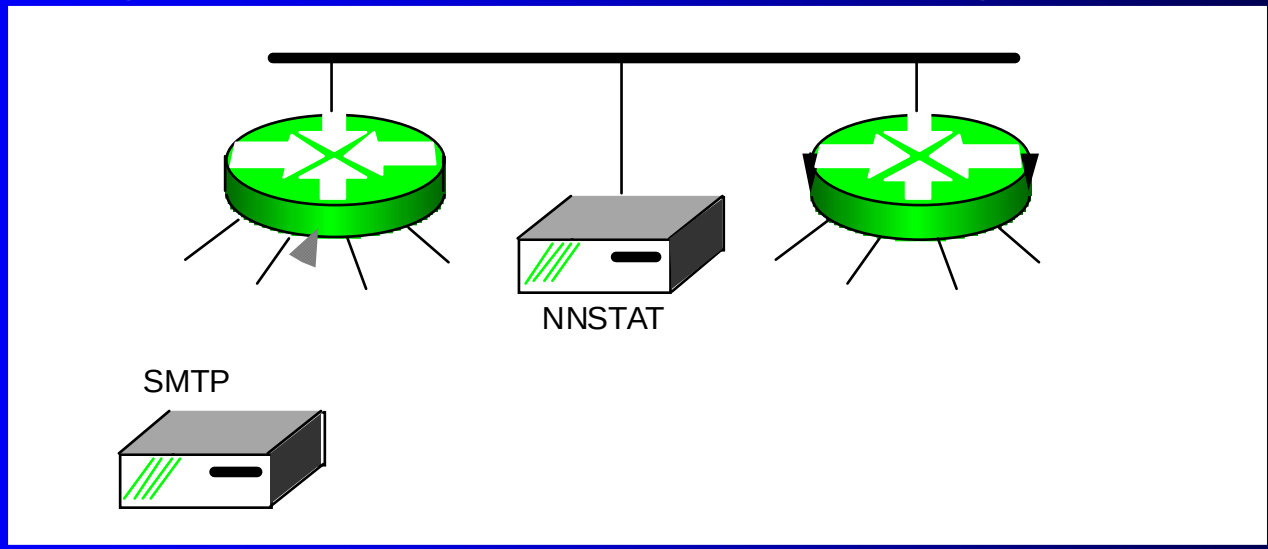
- Routers:
  - Interface volumes
  - Line errors
  - routing tables
  - router resource use

# Reporting

- nnstat - ethernet monitoring with a host
  - gather packet header information
  - source - destination volumes
  - application generated volumes
  - highly flexible data gathering ability
  - expensive to deploy!

# Example Data Collection Architecture

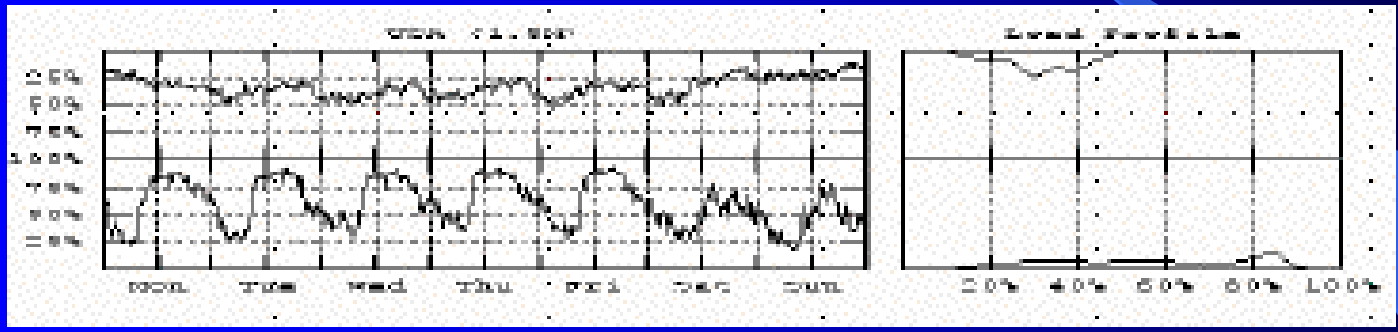
- Routers
  - 15 minute interface volumes and error count
- nnstat
  - Deployed at network peer boundary



# Network Reports

- weekly report of 15 minute link load levels

# Weekly Link Report

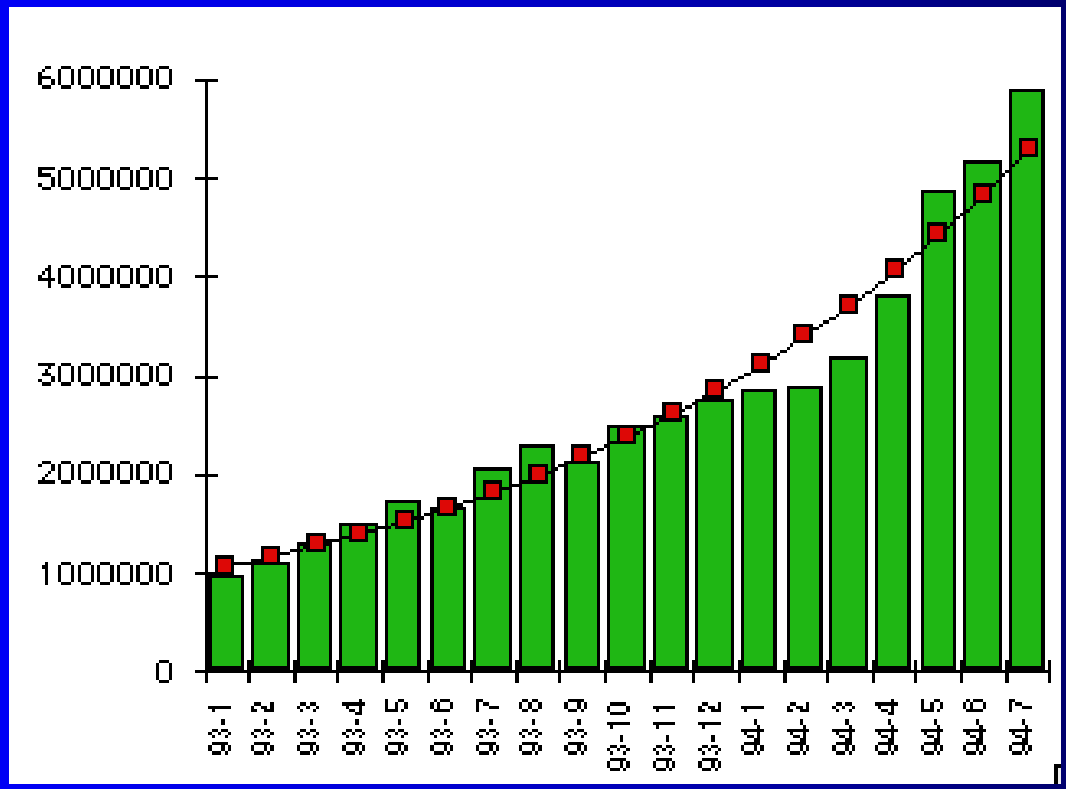



15 minute line sample normalized to % utilization

Histogram of samples as load signature

# Network Reports

- monthly reports
- quarterly trend reports and projections



# Policy Considerations

- This space intentionally left blank!

# Summary

- Network Design defined by router interaction
  - Client Service interface
  - Network Peer interface
  - Internal network design
- Operational Considerations
- Policy Considerations