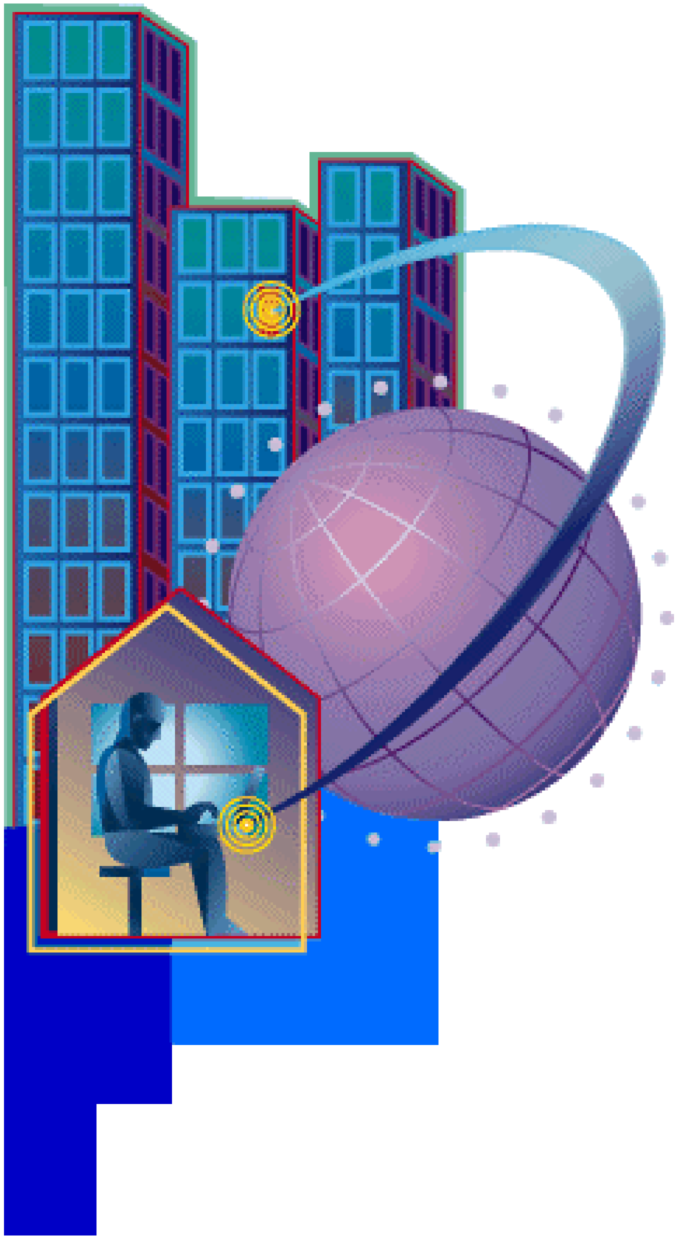
Ascend



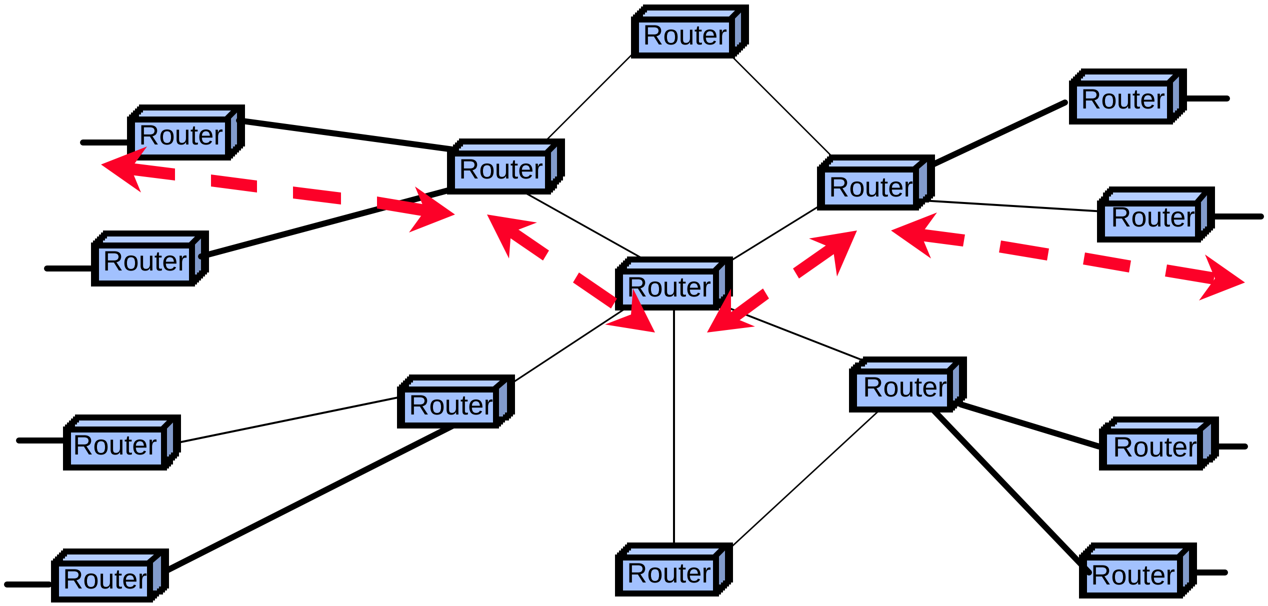
# IP Switching in the Wide Area Network

ASCEND

# *Evolution of Internet Architecture*

- **Then: Routers (Datagram Model)**
- **Now: Routers plus switched core (Overlay Model)**
  - **Frame Relay and/or ATM in the core**
  - **Higher bandwidth, lower latency and cost in the core**
  - **Traffic Engineering / Bandwidth Management**
- **Future: IP switching**
  - **Same switched core, but…**
  - **Integrate IP routing into the core**
  - **Scaling, QoS, VPNs, etc…….**

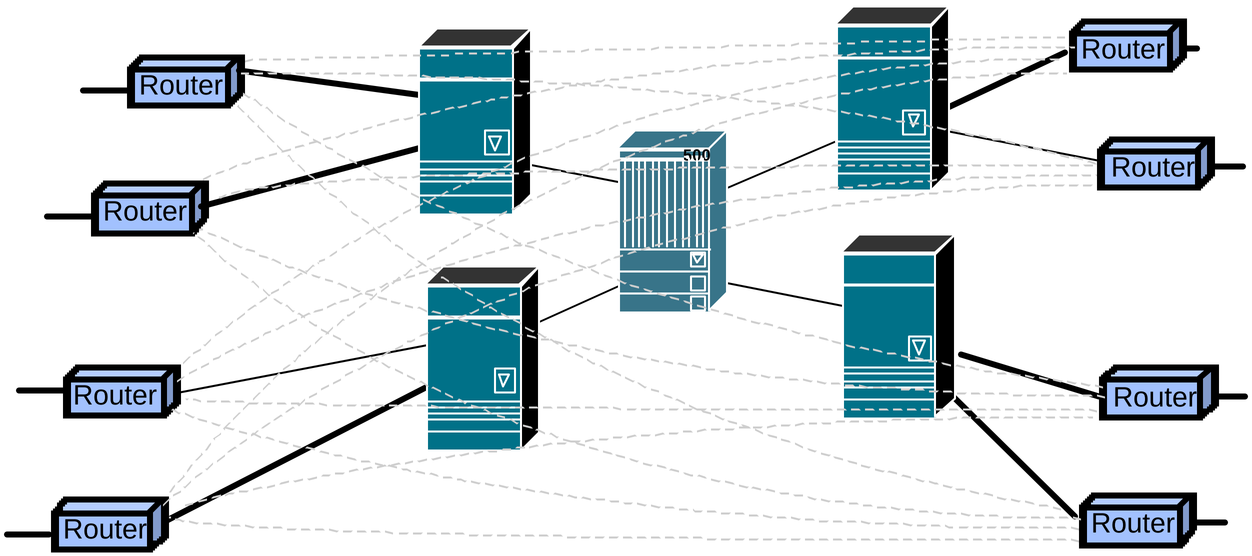
# *Before: Datagram Model*



# *Problems with Datagram Model*

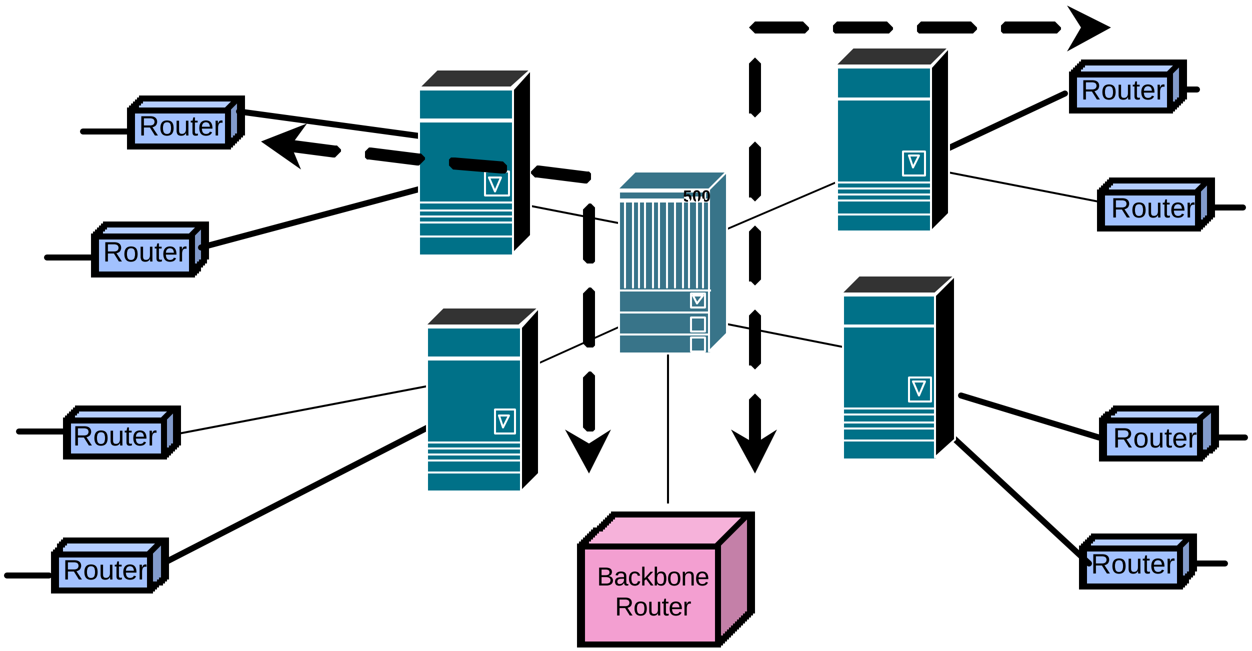
- IP only
  - Doesn't support multiservice (ATM, Frame Relay)
- Only "Best Efforts"
  - Shared QoS = no QoS
  - Router-based RSVP not scalable
- Too much latency for real time data delivery
- Traditionally, router bandwidth is limited
- Doesn't support traffic engineering or bandwidth management
  - Datagram forwarding only

# *Now: Overlay Model with Full Mesh of VCs*




- **Hybrid switching + routing**

# *Now: Overlay Model with “One-Armed” Routers*

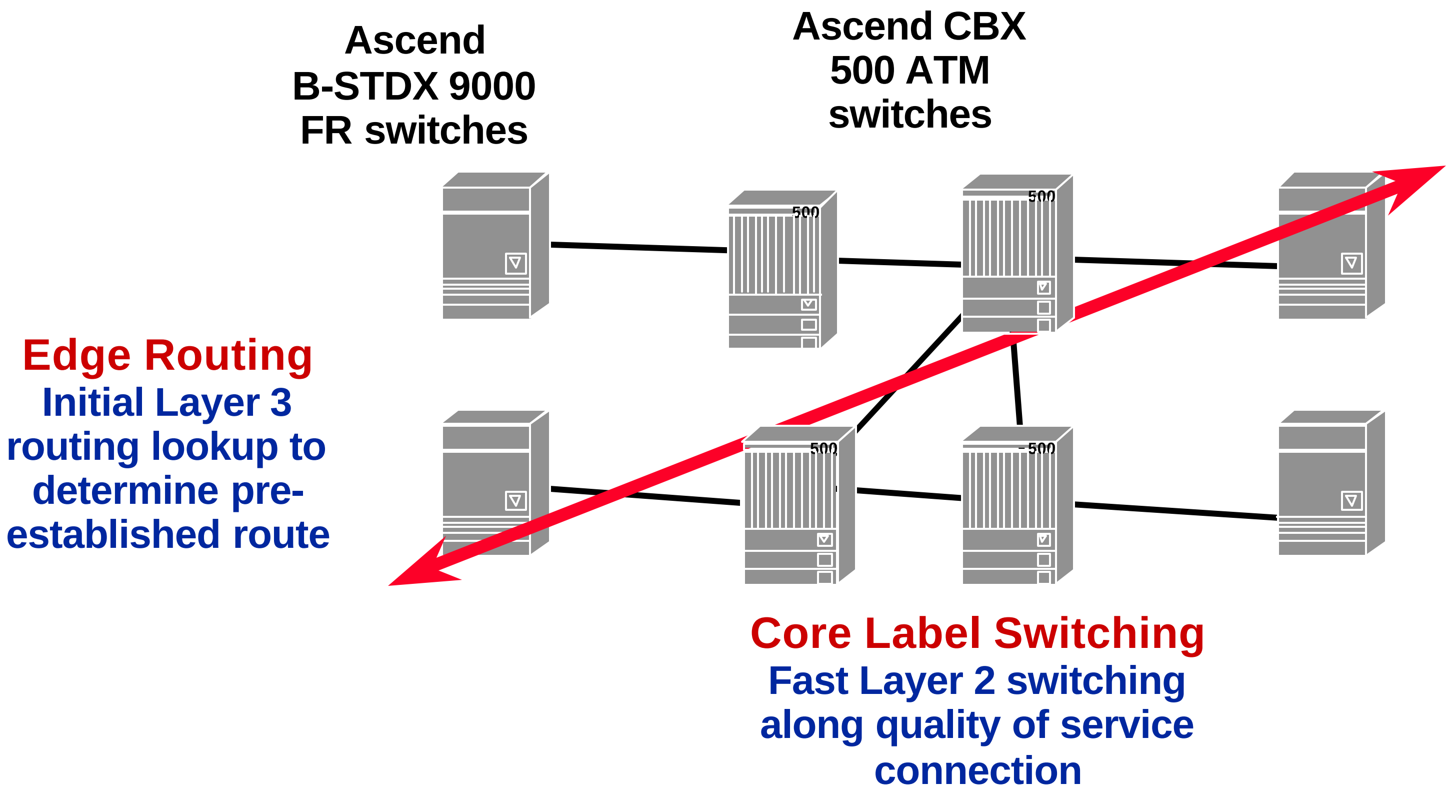



- **Hybrid switching + routing**

# *Overlay Model*

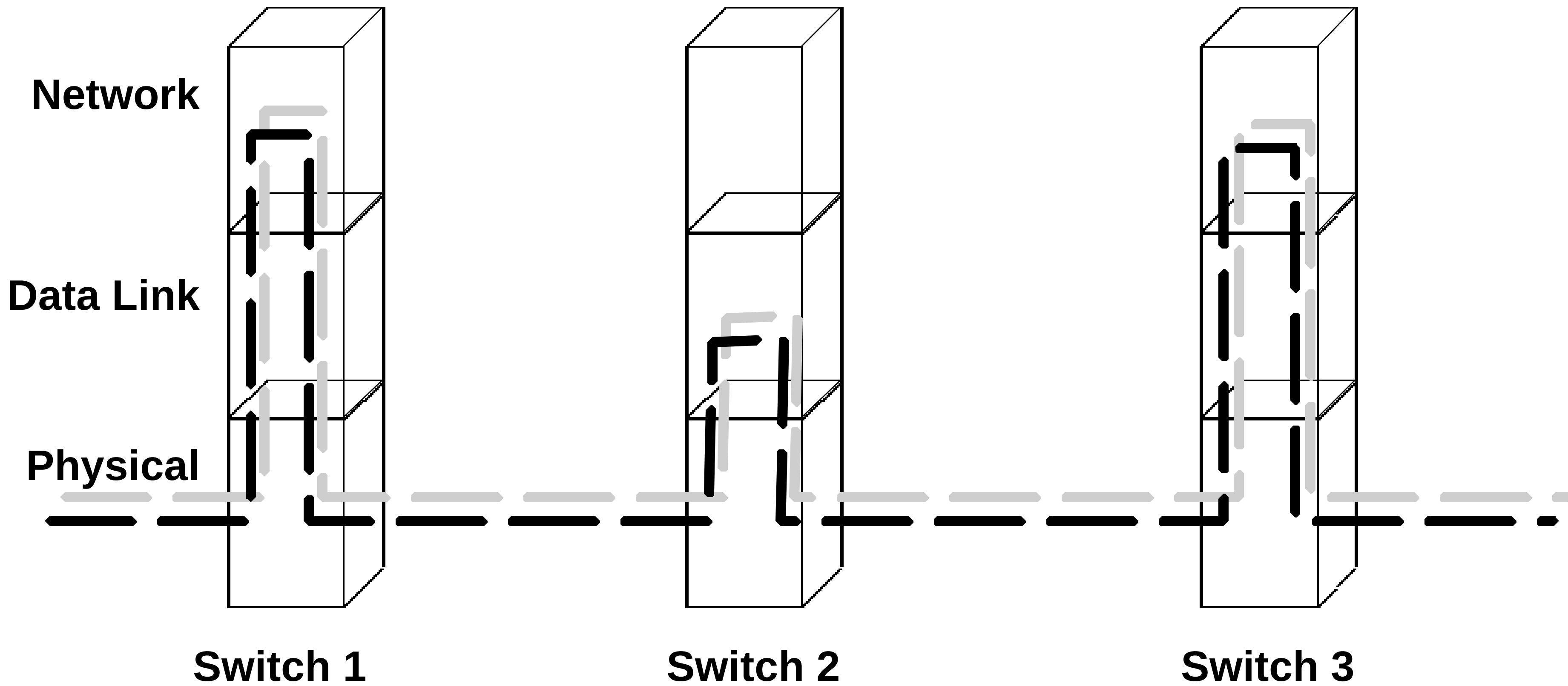
- **Switched Backbone provides**
  - **High speed interconnect, lower cost**
  - **Multi-Service**
  - **Low latency, bandwidth management**
- **But**
  - **If n-squared connectivity, scalability limit (in size)**
  - **If single “one-armed” backbone router, this is bottleneck (if backbone router is only route server, in ISP you will end up with n-squared VCs)**
  - **If partial connectivity, multiple hops may be needed across backbone**

# IP Navigator




- Edge routing determines best path to destination
- Multi-service core provides efficient label switching

# *IP Switching, "Protocol Stack" View*




- **Efficiency of switching + intelligence of routing**

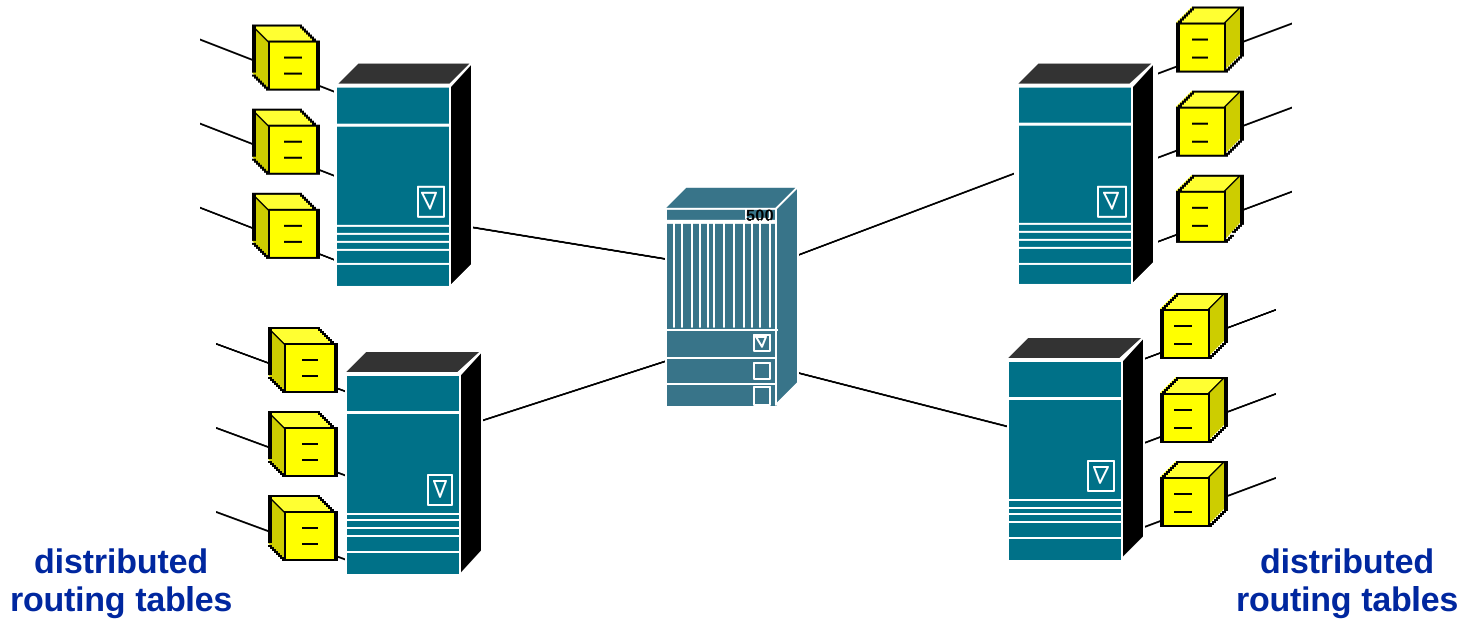
# The IP Navigator Architecture

- Based on Ascend's Virtual Network Navigator (VNN)
  - OSPF-based networking architecture
  - Proven carrier-class performance, reliability and scalability
  - Common architecture for multiservice ATM, Frame Relay, and Remote Access networks
- End-to-End QoS leverages VNN's intelligent use of available network resources
- Fully distributed routing decisions, no single point of failure

# Key Elements of IP Navigator

- **Key: Switching and Routing**
  - **Route on edge of cloud, switch inside**
- **Key: Integrated Routing**
  - **Routing for IP is aware of overall topology**
  - **Routing scaling (each switch has small number of neighbors)**
  - **Based on standard IP routing protocols (OSPF, BGP)**
- **Key: Multipoint to Point Trees**
  - **Essential to data path scaling**

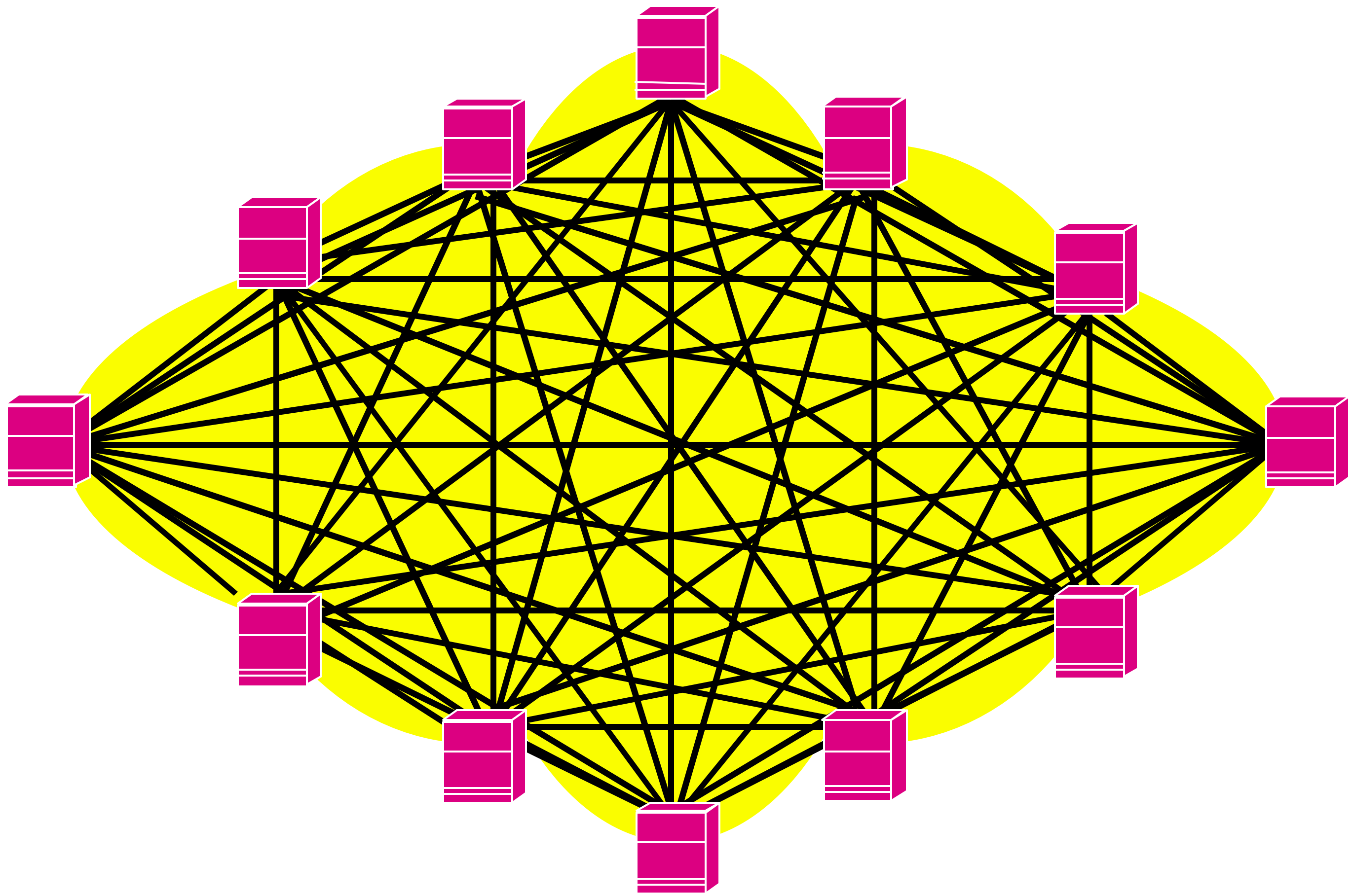
# *Distributed Routing*




- **IP Navigator distributes routing tables to each switch line card**
- **One route lookup, then cut-thru switching**

# Old Point-to-Point Model

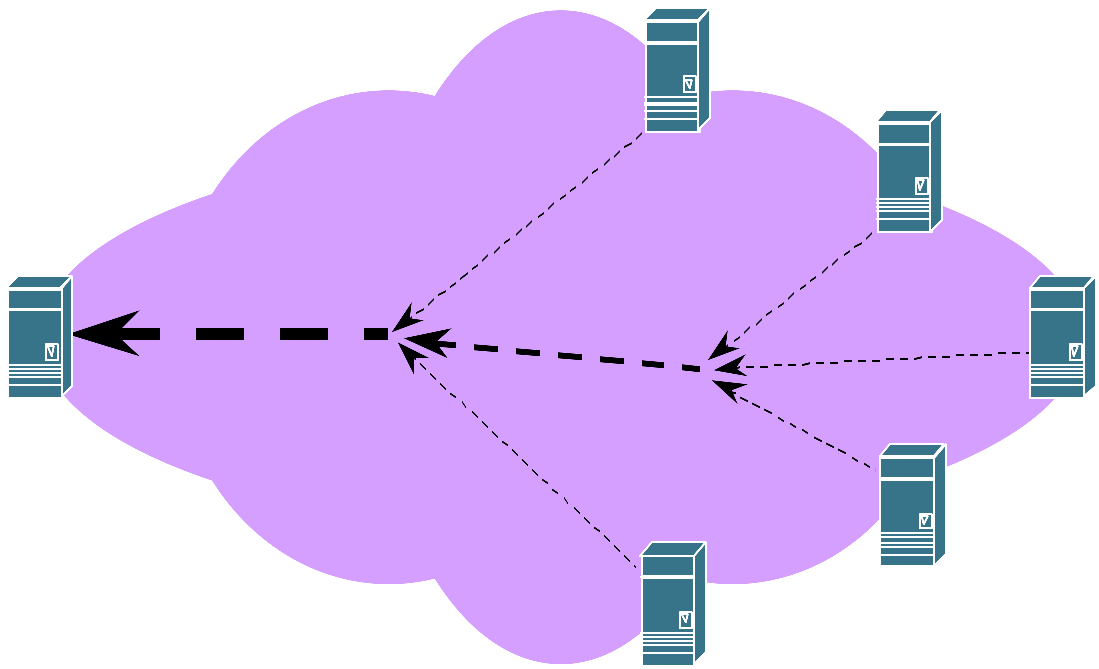
## … exponential $N^2$ Scaling Problem



ASCEND

# *Multipoint to Point Trees (MPTs)*

- **MPTs are essential for scalability**
  - **Allow O(n) Scaling**
  - **Simplify packet forwarding by merging streams**

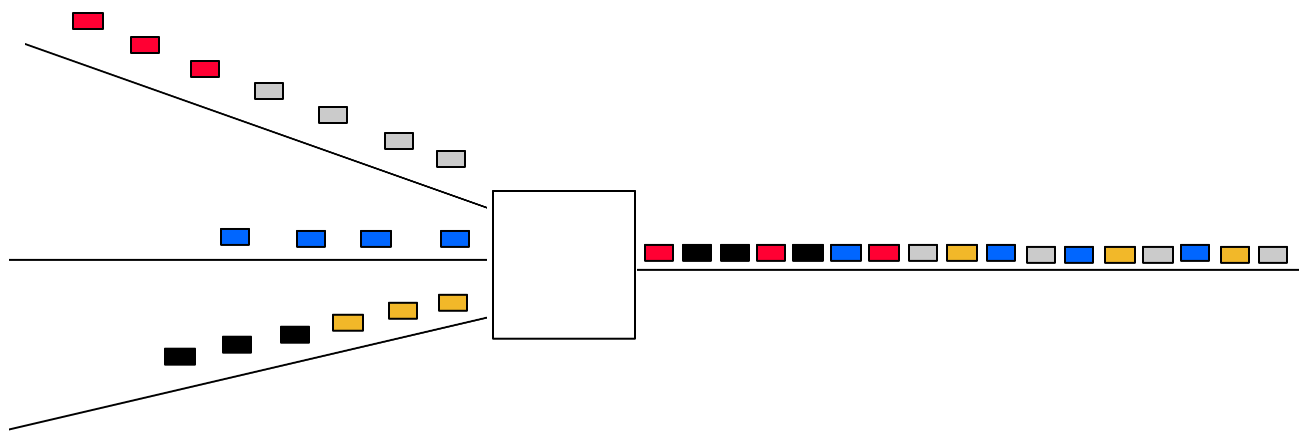


# *Stream Merge for MPTs*

- **Over frame media, merging is relatively straightforward**
  - **Two or more incoming labels can be mapped to one outgoing label**
  - **Packets from different sources can be interleaved as desired**
  - **Streams are merged only when the downstream treatment is the same (including the path to the egress node, and the QoS)**
- **Over ATM: There is another Issue, cell interleave**

# *Cell Interleave Issue*

- **AAL5 requires that within each VC, the cells associated with a packet are contiguous**
- **This is an issue when streams are merged**



- **IP Navigator solves this problem by merging VPs, using the VCI to identify each source within the VP**

# VP Merge for MPTs

- **Virtual paths automatically set up between all switches in the network**
- **VP Switching used at intermediate switches**
- **N virtual paths interconnect N switches**
  - **VPI specifies the root of the tree (destination)**
  - **VCI within each path specifies the leaf (source) (VCIs can be ignored at intermediate points)**
  - **VCIs allows cells from different packets to be interleaved, differentiated by standard SAR at root.**

# *VNN Routing*

- **Based on OSPF**
  - **The preferred standard for IP interior routing (within a domain)**
  - **Robust Link State Routing Protocol**
  - **Stable, fully debugged, mature technology**
  - **OSPF was authored by John Moy of Ascend**
- **VNN Extends OSPF to support VC /QoS routing**
  - **Network resources announced in routing protocol (bandwidth, delay, ...)**
  - **Source (for VCs) or Root (for MPTs) picks path for call based on requirements of VC and capabilities of the network**
  - **Explicit routing used to set up the VC / MPT**

# *Why Explicit Routing?*

- **IP Navigator uses explicit routing for setting up VCs and MPTs**
  - **The path to be followed is included in setup message**
  - **This would be too expensive for pure datagram networking**
  - **Is minimal overhead for VCs (only call setup needs explicit route)**
- **Explicit Routing has multiple advantages**
  - **Ensures consistent path will be followed**
  - **Allows “clever” route computations**
  - **Prevents loops, permits crankback and retry**
  - **Important for QoS routing, traffic engineering**
- **Pioneered by the Core Switching Division of Ascend**
  - **Approach has been adopted by ATM PNNI routing**

# VC Granularity and Latency

- **IP Navigator sets up MPTs (VCs) based on topology**
  - **MPTs set up to egress nodes based on routing protocol exchange**
- **VCs are set up prior to arrival of user data**
  - **Minimizes latency in forwarding user data traffic**
- **Some alternatives (MPOA, Ipsilon, Toshiba's CSR) set up VCs on host to host granularity**
  - **This precludes setting up all possible VCs a priori**
  - **This in turn requires VC setup to be based on data traffic**
  - **This adds a serious latency problem: what do you do with a packet while waiting for the VC to be set up? (Dump it? Buffer it?)**
  - **Also, what do you do if there are too many host to host flows?**
  - **This is a serious scaling issue**

# Hierarchy

- One OSPF/VNN area can extend to approximately 400 nodes, 1000 links
- For larger networks, IP Navigator allows use of hierarchy
  - Follows OSPF two-level hierarchy design
  - Nodes know the full topology in their area only
  - Area border routers/switches participate in multiple areas, including backbone area
  - Nodes know the router IDs and IP reachability of nodes in other areas (this can scale to tens of thousands of reachability advertisements)
  - MPTs can terminate at area border routers/switches, or extend between areas

# Interconnecting Routing Domains

- **The advanced features of IP Navigator (MPTs, QoS, Traffic engineering, etc..) operates within a single IP Routing Domain**
- **For interconnecting multiple domains: standard IP "Datagram" forwarding is used**
- **BGP is the standard routing protocol used between Routing Domains**
  - **External routes are fed into the Routing Domain using either I-BGP or by injecting routes into OSPF**
  - **EGP and static routes are available as options**
- **Also note VPN support, discussed later in this talk**
  - **Customer network may attach as a VPN to an IP Navigator carrier**

# *Multicast*

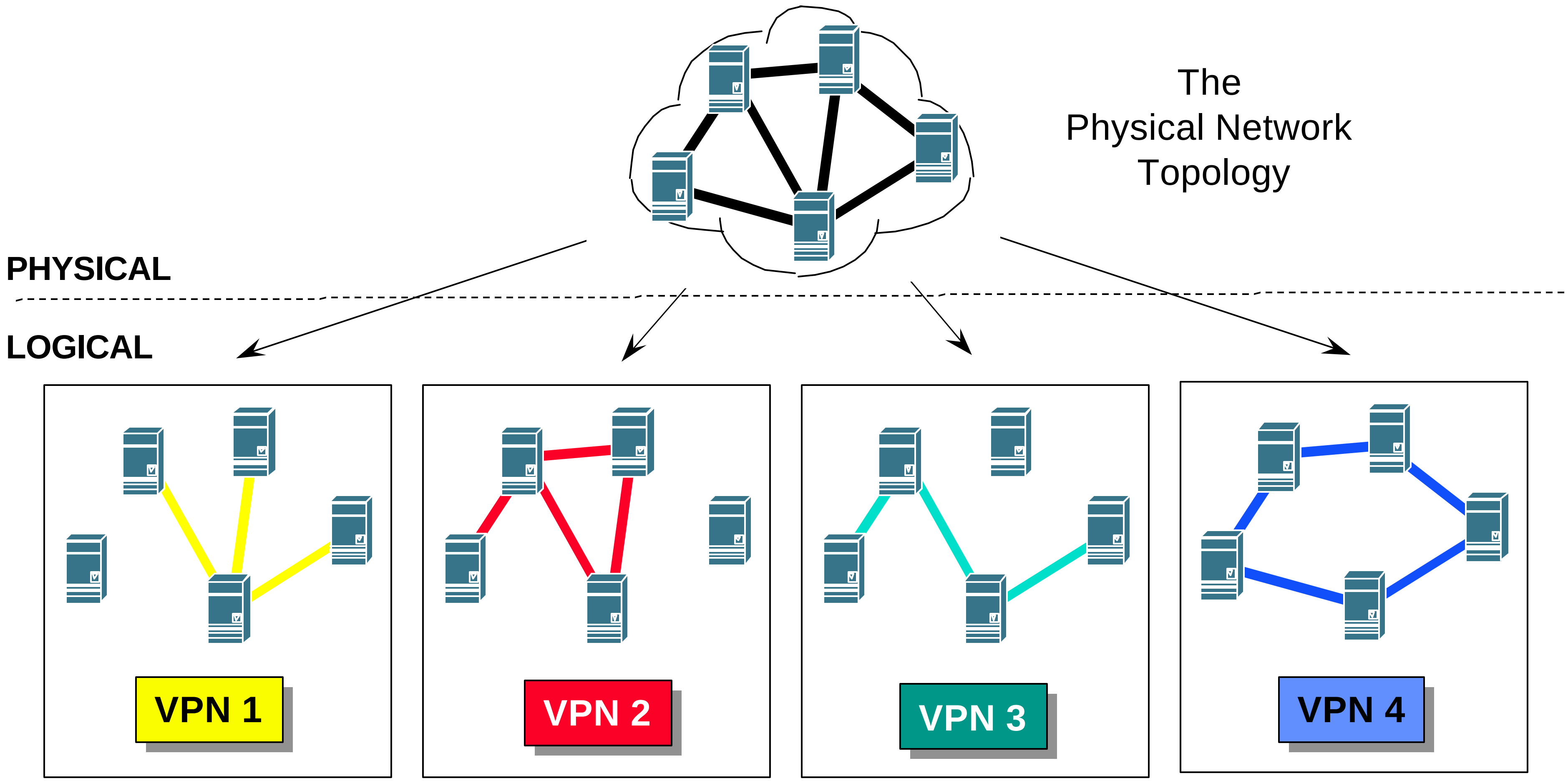
- **IP Navigator supports standard IETF multicast routing protocols**
  - **MOSPF is used in the core of the IP Navigator cloud**
  - **DVMRP, PIM (dense and sparse modes) and MOSPF supported at edge**
  - **IGMP supported for host registration to join multicast groups**
- **IP Multicast is mapped to ATM point-to-multipoint connections**
  - **Point to multipoint trees can be set up on demand for efficient multicast**
  - **Default hop by hop forwarding used until the PMT is set up**

# *Advanced Services Delivery*

- **IP Navigator supports multiservice networks**
  - **Simultaneous IP via PPP, ATM, and FR in same network!**
  - **FR and ATM used both for native services and IP access**
- **IP Navigator enables delivery of premium IP services:**
  - **Virtual Private Networks**
  - **End-to-end QoS**
  - **Differentiated Service Offerings**
  - **Bandwidth Management/Traffic Engineering**

# *IP Navigator Builds Multiple VPNs*

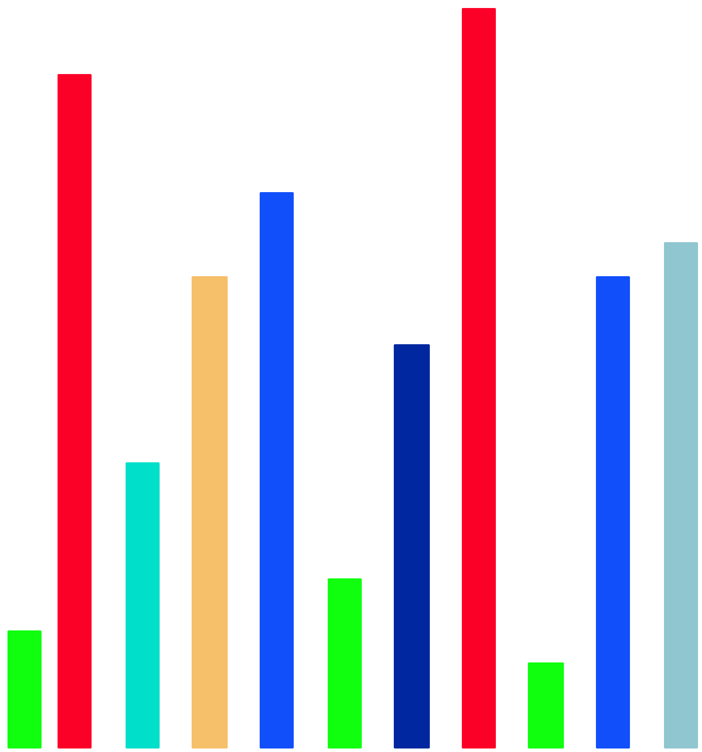
## Independent logical maps for each Virtual Private Network



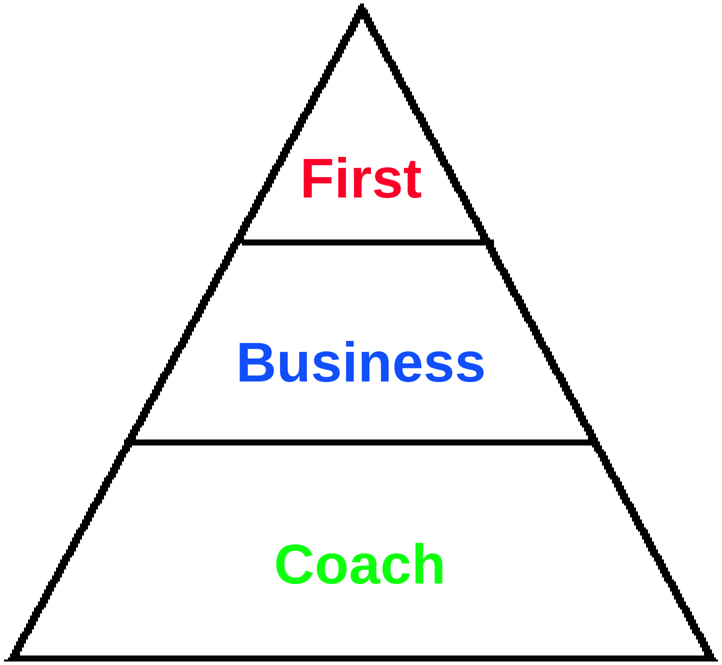
# *Quality of Service*

- **What Granularity Level to Support?**
  - **Fine Grain (host to host flows)**
  - **Coarse Grain (differentiated service classes)**
- **How to Determine what level of QoS is needed**
  - **Explicit request (RSVP, ATM UNI)**
  - **Provisioning**
  - **IP ToS in header**

# *Range of QoS Granularity*



Individual (fine grain)




Grouped

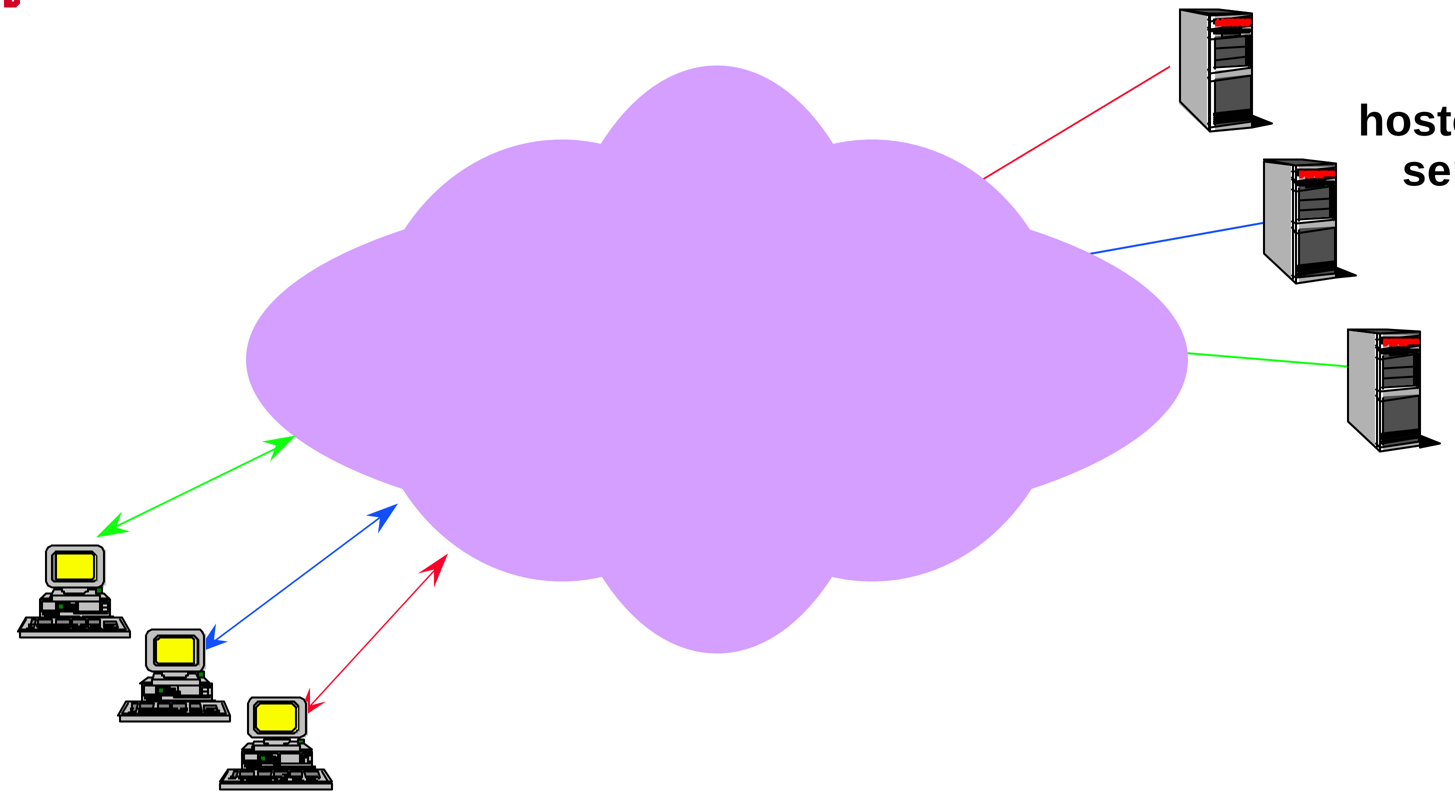
# *Ascend IP Navigator QoS Support*

- Best effort traffic uses IP Navigator Multipoint-to-Point Trees (or Point-to-Multipoint trees for multicast)
- For coarse-grained QoS, separate MPTs can be set up for each of several classes of service
  - IETF is beginning work on coarse grained QoS (referred to as Differentiated Services)
- For fine-grained QoS, IP traffic is carried over ATM VCs that provide the required characteristics
  - PVCs can be provisioned with QoS attributes
    - Useful for recurring bulk flows
  - Shortcut SVCs can be dynamically established
    - Based on signaling or flow detection

# *How to Determine Which QoS to Use*

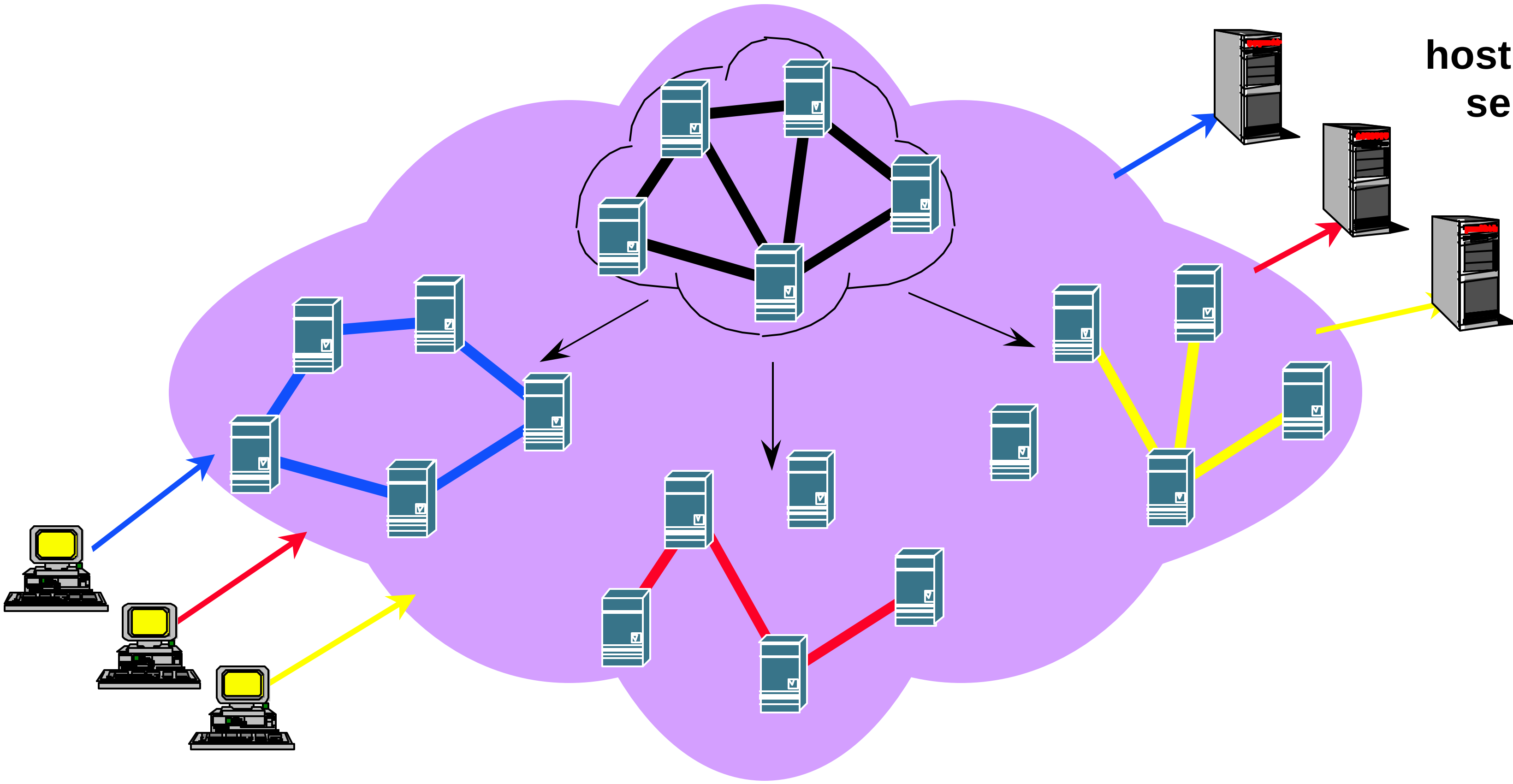
- **Traffic that requires a particular QoS can be dynamically detected or signaled for a number of ways:**
  - **RSVP**
  - **NHRP (with QoS extensions)**
  - **Provisioned flow description: any or all of source and/or destination IP address (with CIDR masks), ToS, IP protocol ID, source and/or destination protocol port numbers**
- **For fine grained QoS, SVCs are always shortcuts from source to destination**
- **For coarse grained QoS, IP packets are mapped onto appropriate MPTs**

# *Provisioned QoS*




- **QoS determined by source, destination or service class**
- **Range of granularities possible: coarse to fine**

# *Provisioned QoS and VPNs*




**Separate route table for each QoS class or VPN**

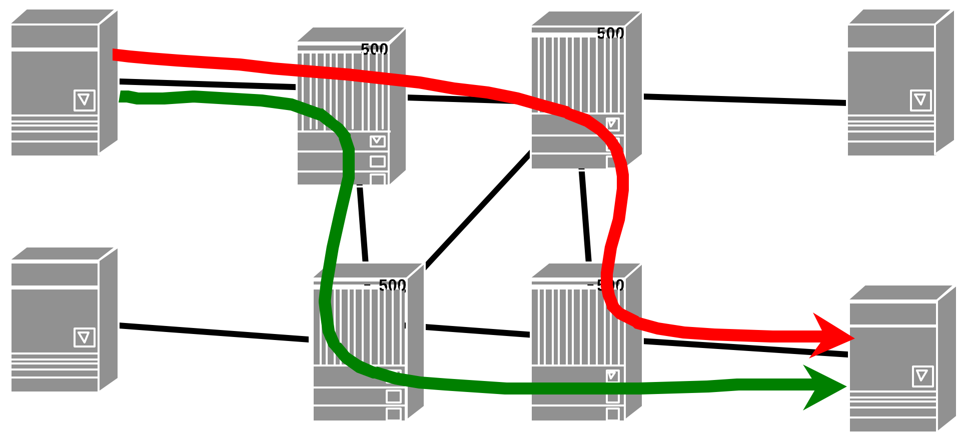
# *Why Not Just Use RSVP?*

- **RSVP cannot scale for use on the Internet, and its use is confined to enterprise or corporate networks**
- **Problems include:**
  - **Too much state information per connection**
  - **No flow aggregation, too many individual flows**
  - **Billing, security, policy, hacking RSVP protocol?**
  - **Much too complex**
- **Many Internet Service Providers will not deploy RSVP due to these problems**
- **RSVP may go the way of ATM to the desktop**
  - **Or only used for signaling on the network edge**

# Traffic Engineering

- **Traffic Engineering (aka “Bandwidth Management”)**
  - **Intelligent path selection**
  - **Equalize the loading of the network, use the network efficiently**
  - **Network operator controls the flow of user data**
- **Based on**
  - **Network capabilities (especially bandwidth, also delay, etc..)**
  - **Level of user data**
  - **Requested QoS**
- **This is related to, but separate from, QoS routing**
  - **Also applies to “best effort” traffic**

# *Traffic Engineering (con't)*

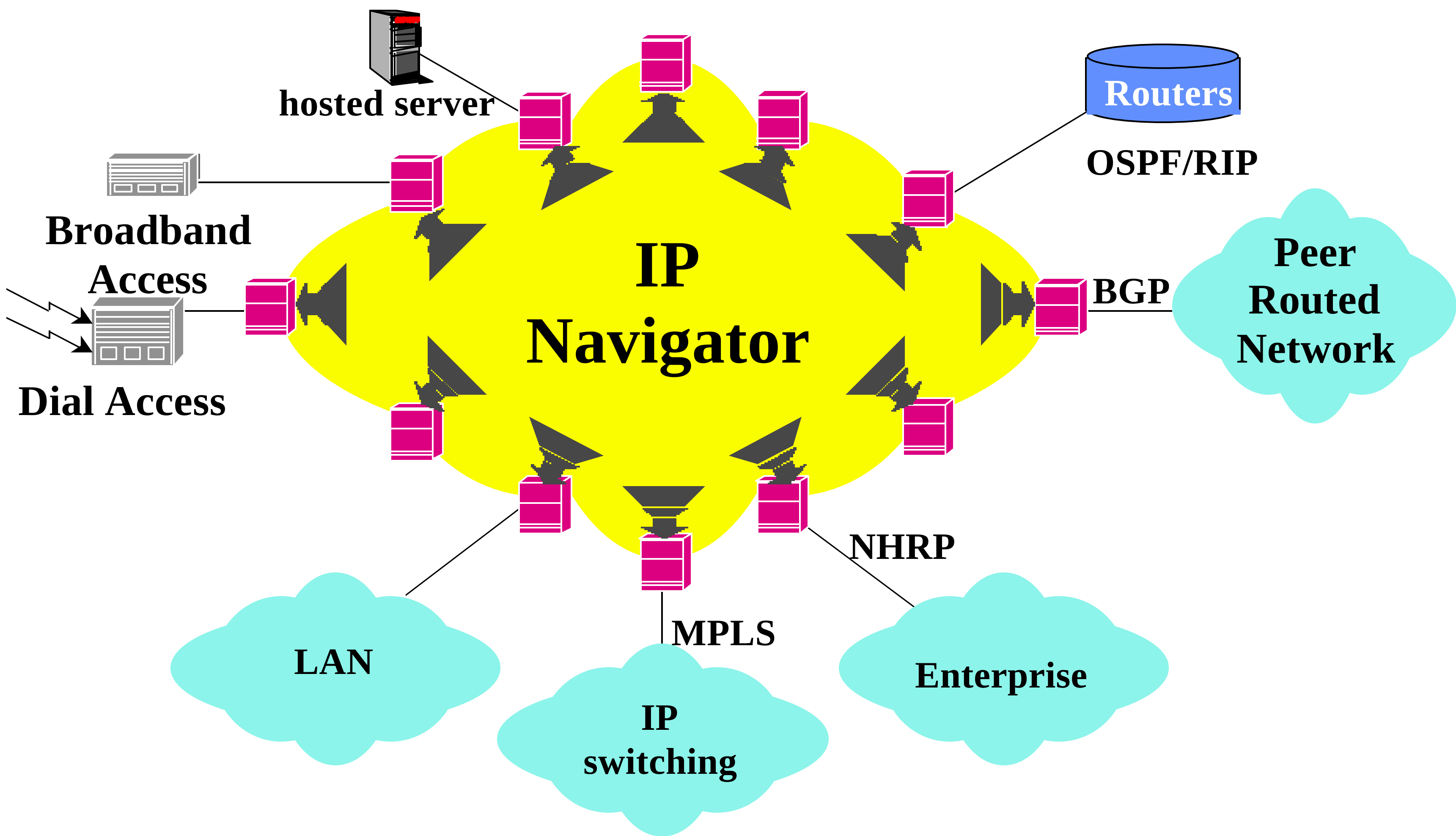



- **Provision a different route than IP routing would use to override routing and efficiently utilize bandwidth**

# *IP Navigator and Standards*

- There are two approaches
  - IP Navigator in the core of a network, standards at the edge
  - Standardize protocol functionality similar to IP Navigator, allow multivendor network core
- We are actively pursuing both approaches

# *IP Navigator in the Core*




**IP Navigator is the "agnostic" core of the network**

# *IETF Multi-Protocol Label Switching*

- **IETF is working hard on an MPLS standard**
  - **Combines features of IP Navigator, Tag Switching, and ARIS**
  - **These are relatively similar approaches**
- **Working group has two core documents jointly produced by Ascend, Cisco, and IBM:**
  - **MPLS Framework document, authored by Ross Callon of Ascend plus five co-authors**
  - **MPLS Protocol Architecture specification, authored by Ross Callon plus two co-authors**
- **Ascend will continue to contribute to the standard, and will support the standard when it is done**

# *Features of MPLS*

- **Same Basic Key Features as IP Navigator**
  - **Integrate switching and routing**
  - **IP routing on switches as well as routers**
  - **Multipoint to Point trees**
  - **Circuits set up based on control traffic / topology**
- **Explicit routing is optional**
  - **Consequence of not being able to choose a routing protocol**
- **More things are optional**
  - **Implies the need to deal with interworking of multiple options**
  - **This is a normal consequence of the standards process**

# *MPLS Terminology*

- **Some terminology is a bit different from IP Navigator**
- **MPLS uses "labels" for L2 forwarding.  ATM VPI/VCI is one example of a label**
- **A "Label Switching Router" (LSR) is a router which understands the MPLS  protocol**
- **MPLS has an explicit "label distribution protocol" (LDP) for setting up "label switched paths" (LSP)**
- **The set of packets which flow down a label switched path is referred to as a "stream"**

# *MPLS Issues*

- How to Deal with Loops
- Granularity
- Local versus Egress Control
- VC merge versus VP merge
- TTL and Traceroute

# MPLS, Dealing with Loops

- We are concerned about looping label switched paths
- Loop Prevention
  - Refers to methods that prevent loops from being set up
  - Can use explicit routing
  - Can use an explicit loop prevention protocol as part of the label distribution protocol
- Loop Detection
  - Refers to methods that allow loops to be set up, then notices and fixes it
- Loop Mitigation
  - Refers to methods which minimize the damage caused by loops
  - EG: TTL, Per-VC queuing, and weighted fair queuing
- IP Navigator doesn't have this problem due to use of explicit routing

# *MPLS Granularity*

- **MPLS allows a wide range of "granularity" with respect to what IP packets get mapped to a particular label**
  - **Packets whose destination matches a particular address prefix**
  - **Packets whose destination matches any of a list of address prefixes**
  - **Packets routed via a particular egress node (easy with link state routing, hard with distance vector routing)**
- **Multiple fine grain LSPs can be mapped into one coarse grain LSP**
- **However, a coarse grain LSP cannot be mapped into a fine grain LSP**
- **Thus there needs to be a way to ensure consistency**
  - **IP Navigator doesn't have this problem, because it uses egress node granularity with link state routing (OSPF)**

# *Local vs. Egress Control*

- **What Initializes the exchange of labels for a particular LSP?**
- **Local Control**
  - **Initially, any node can exchange labels with its neighbors**
  - **Streams are "spliced together" locally**
  - **Stream granularity needs to be assured somehow**
  - **Requires consistent management or can cause temporary thrashing at start**
- **Egress Control**
  - **Initially, egress node provides labels to neighbors**
  - **Each node, when it gets label from downstream, provides consistent label to upstream**
  - **Once you "learn the granularity", can correct locally when routing changes**
- **Egress control is necessary in some situations, local control can be useful (but is not needed) in other situations**

# VC Merge vs. VP Merge

- VP merge has already been explained in the IP Navigator discussion
- VC merge is an alternative: Each switch buffers the cells associated with a particular IP packet until the entire packet has arrived, and then forwards them all together.
- VC merge is not consistent with most existing ATM switch fabrics, requires additional buffering
- However, VC merge makes more efficient use of the VPI/VCI space (since each LSP requires only one VCI, rather than one VPI)
- It appears that the standard is going to support both options

# *TTL and Traceroute*

- **IP routers decrement the "Time to Live" at each hop**
- **This is useful to stop looping packets, and to support Traceroute (an IP application which traces the path of a packet)**
- **ATM switches don't know about TTL, and thus can't decrement it**
- **IP Navigator and MPLS solve this in the same fashion: By having the nodes which do IP forwarding decrement the TTL by the length of the label switched path**

# *IP Navigator and MPLS*

- **Very similar approaches**
  - **Use IP routing on routers and switches**
  - **IP forwarding at the edge, label (VPI/VCI) forwarding in core**
  - **Allows High Speed Forwarding, Traffic Engineering**
- **MPLS is work in progress**
  - **Completion likely in late 1998 or early 1999 (wild guess!)**
  - **No QoS, VPN, other added value support in initial version**
- **IP Navigator makes simplifying assumptions**
  - **OSPF as the Interior Routing Protocol**
  - **Does not have to be "all things to all people"**
  - **This allows IP Navigator to be ready much sooner, and to be more easily extended to value added features (VPNs, QoS, Traffic Engineering)**

# *Other Proprietary Approaches*

- Tag Switching, Cisco Systems
- Aggregate Route-based IP Switching (ARIS), IBM
- "IP Switching", Ipsilon
- Cell Switching Router (CSR), Toshiba, Ennovate Networks

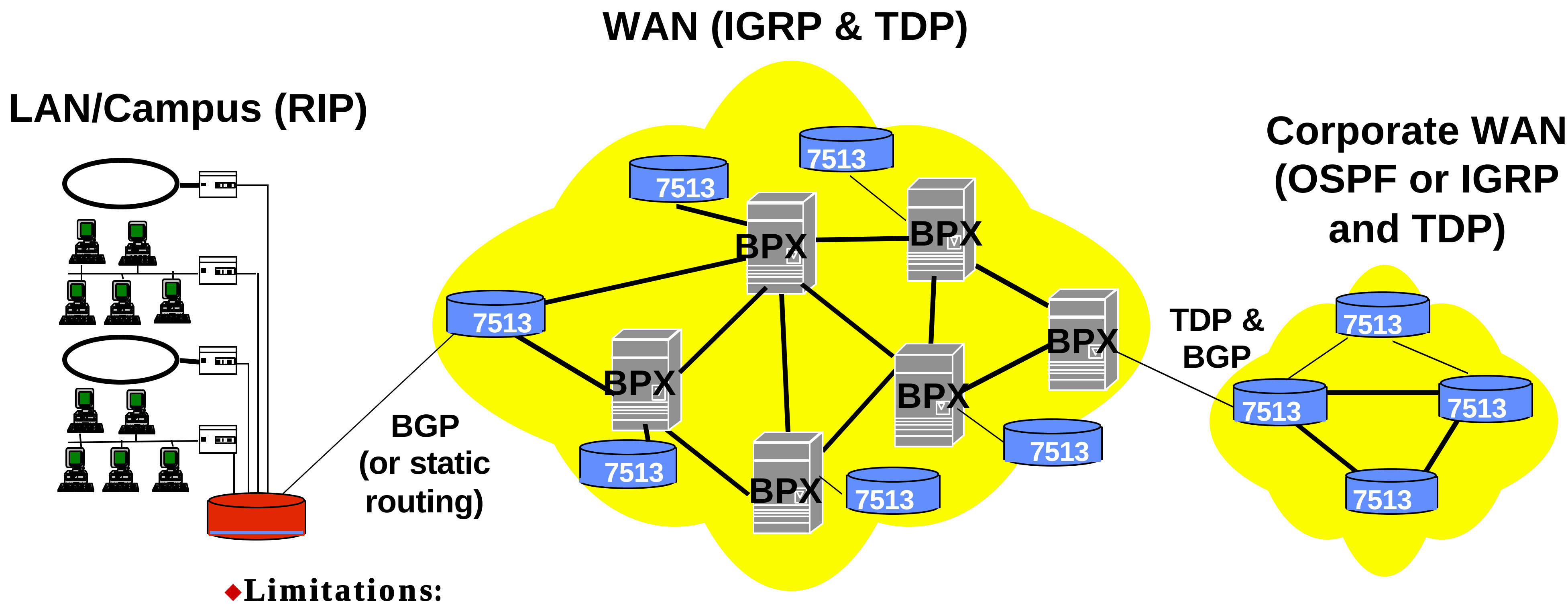
# Tag Switching

- **Cisco's Tag Switching is similar to IP Navigator and MPLS**
  - **IP routing on switches, topology-driven label assignment, routing at the edge and switching in the core, ...**
- **Cisco is also participating heavily in the MPLS effort**
  - **We are working with them (and IBM) closely**
  - **Tag Switching keeps changing in response to MPLS work**
- **Tag Switching's architecture**
  - **Is very complex, with a lot of features (much legacy to support)**
  - **Lots of documents, lots of authors, not completely consistent (but, this is work in progress, much like MPLS)**

# Tag Switching (con't)

- Tag switching is largely a subset of MPLS
  - Or perhaps more fairly, essentially all of the important features of tag switching are included in the MPLS standard
  - This is sensible, since Cisco has the same issues (compatibility with everything) and has been very active in the MPLS effort (along with Ascend, IBM and other companies)
- There is a great deal of documentation on Tag Switching, not all of which is consistent
  - This again is a reasonable thing: For a very large company, it is sensible to get an overview of the possibilities, ship a much simpler subset, participate in standards, and then straighten out the details
  - This makes it a bit hard to know what is planned for any particular release

# *Tag Switching, as envisioned by Cisco*




- **Limitations:**
  - **Complexity of new, unproven TDP protocol in network**
  - **No end-to-end QoS for IP**
  - **No value added services**
  - **Cisco router architecture limits performance**
- **Benefits: Extends life of poorly-performing Cisco Routers!**

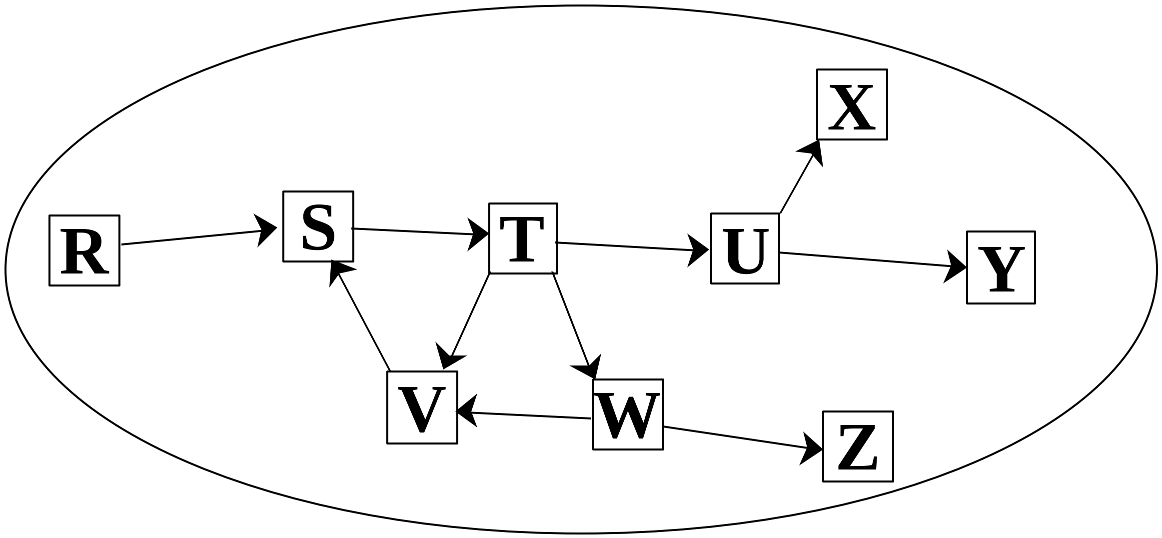
# *Key Aspects of Tag Switching*

- **Compatible with wide range of Routing Protocols**
  - **This is necessary given Cisco's installed base**
- **Local control of LSP setup**
  - **No loop prevention**
- **Allows hierarchical "stack" of tags**
  - **Hierarchy is the same as  MPLS**
- **Multicast label exchange is piggybacked on PIM**

# *Tag Switching Issues*

- **Tag Switching is independent of the routing protocol**
  - **They have to do this to be consistent with IGRP, IS-IS, etc..**
  - **Means explicit routing is hard (in general), therefore is option only**
  - **Makes the result more complex than IP Navigator**
- **They don’t prevent loops**
  - **Over ATM and Frame Relay, just hope they don’t happen (okay for unicast with VC queuing, issue regarding multicast)**
  - **Over special Tag / MPLS switches, use TTL**

# *Multicast Loop Grenades*

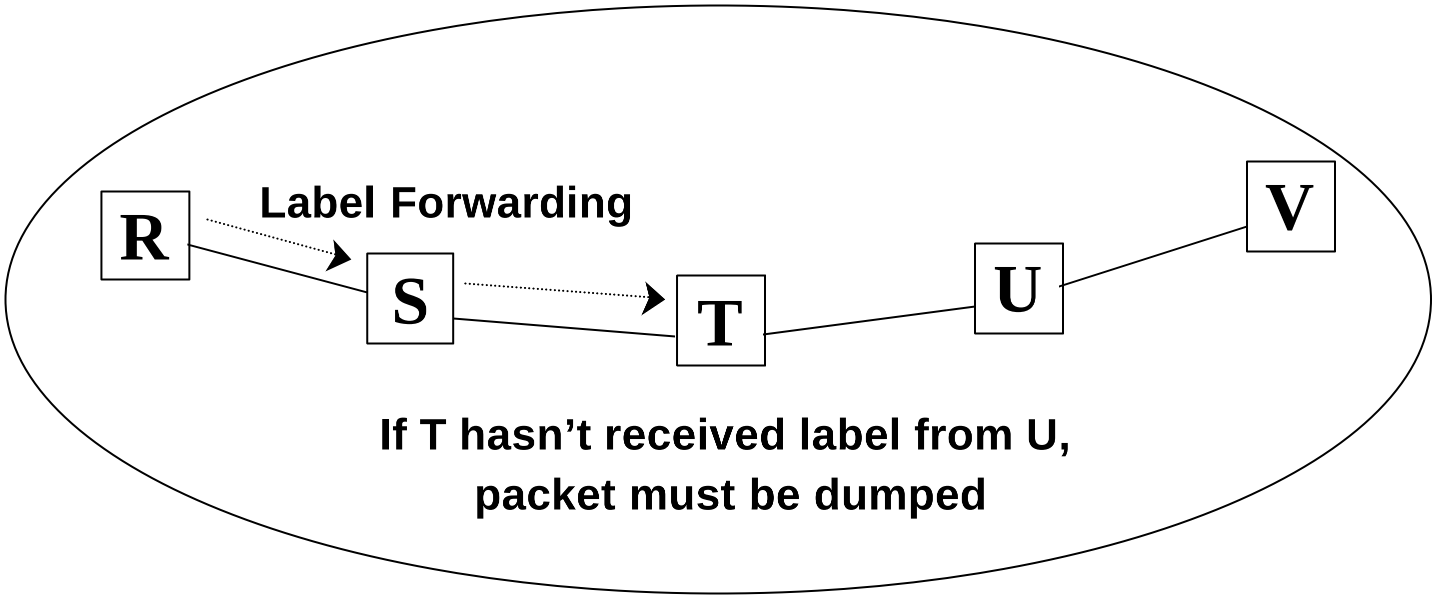



- **Consider multicast packet from ingress R, over ATM media**
- **Assume some but not all branches loop**
- **HUGE numbers of duplicate packets get delivered on other branches**
- **This is not a good thing**

# *Tag Switching Local Control*

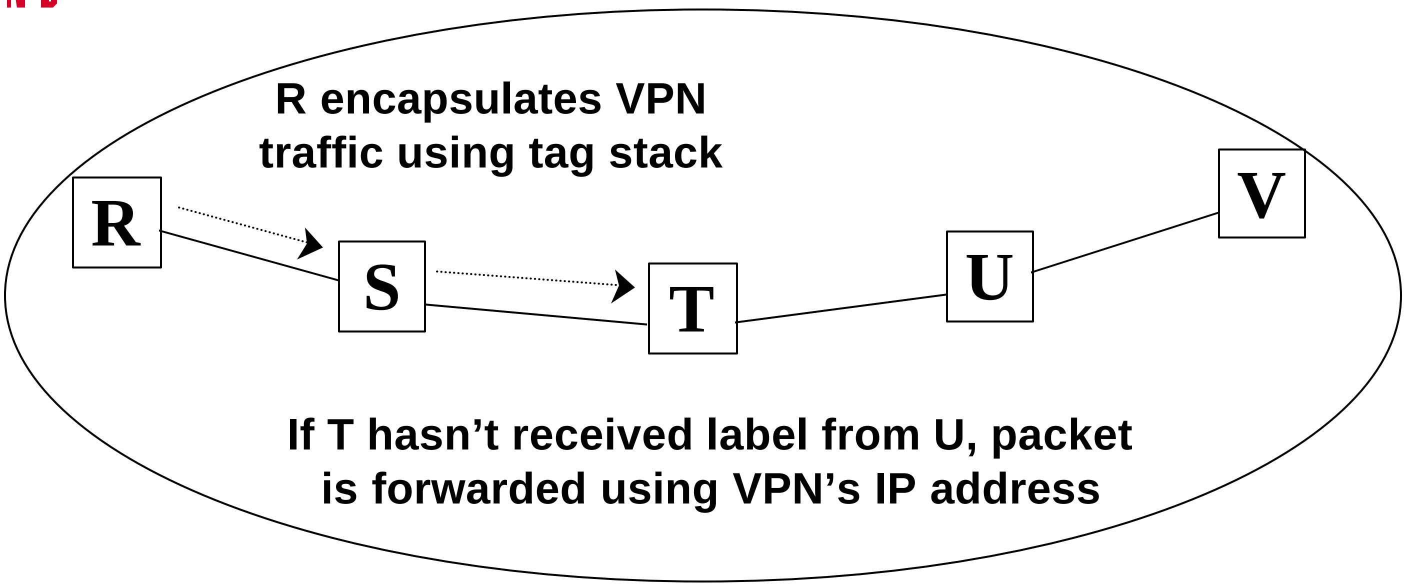
- **Consistent granularity ensured by network management, or by thrashing**
- **Makes egress granularity hard (or impossible) to accomplish**
- **There seem to be a bug with regard to hierarchical operation with BGP and with VPN support**

# *Local Control / BGP Bug*




- Suppose border nodes run BGP, only interior routes used inside ISP
- R puts local tag onto packet, thinks LSP gets to V
- But: LSP isn't set up yet past T.  T can't forward the packet.

# *Local Control / VPN Bug*




- Suppose border nodes are using tags to encapsulate packets from a VPN
- The VPN is using local IP addresses (not valid inside the core network)
- But: LSP isn't set up yet past T.  T thinks that it can forward the packet.
- Packet is forwarded using an invalid IP address. This is really bad.

# *Local Control, Revisited*

- **These bugs occur because LSPs are set up via local control, but are assumed to have wider semantics (global within the area) even though there is no way to know whether the wider semantics are actually correct**
  - **Packets may be lost (if encapsulated label or address is unknown)**
  - **or, may be delivered to the wrong place based on the wrong local IP address**
- **Local control is a dangerous optimization**
  - **This is an example of being “optimistic”**
  - **Optimism isn’t always the right protocol design**

# *Tag Switching Summary*

- Tag Switching is independent of the routing protocol
  - Complexity of new, unproven TDP protocol in network
- No loop prevention
- Local control
- IP Navigator is much simpler
  - Easier to extend to traffic engineering and QoS
  - Prevents loops
  - Addition to existing mature OSPF routing

# ARIS

- **ARIS is also very similar to MPLS, and is basically a subset of MPLS**
  - **IP Forwarding on the edge, label forwarding in the core**
  - **Extend IP routing into Switches**
  - **Topology based Label Assignment**
  - **Support for multipoint to point trees**
  - **Compatible with “standard vanilla” IP routers**
  - **Uses same hierarchical stack of tags**
- **As a senior engineer said at a recent IETF: “It is obvious that we were all trying to do the same thing”**

# *Key Features of ARIS*

- **Works with a range of routing protocols**
- **Provides explicit loop prevention mechanism**
  - **Important for operation over ATM, especially for multicast**
  - **Useful in other cases**
- **TTL to egress is known**
  - **This also ensures Traceroute support**
- **Egress / Root Control**
  - **Ingress control for multicast**
  - **Ensures consistent granularity, semantics**
  - **Controls whether to use loop prevention in any particular case**
  - **Allows policy control**

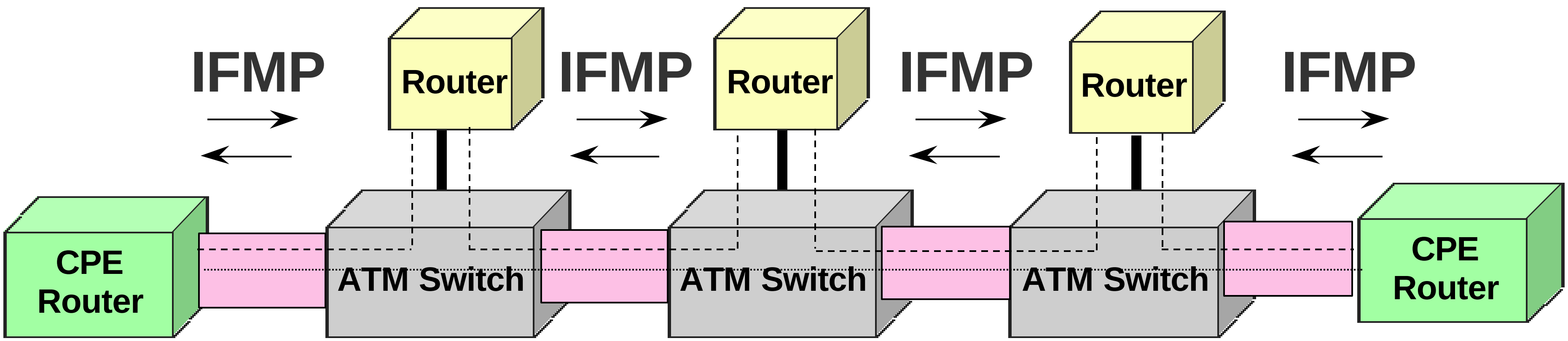
# Key Features of ARIS (con't)

- Multipoint to Point Trees
  - VP Merge in the short term
  - As in IP Navigator, this allows compatibility with existing ATM switches, and minimizes latency
  - VC merge as the long term approach to scaling / MPTs
- Multicast as integral part of LDP
  - ARIS doesn't require any particular multicast routing protocol
- Tunneling / Hierarchy
- Explicit routing available as an option
- ARIS solves the semantic contradiction and looping bugs of Tag, otherwise is relatively similar

# *Ipsilon's IP Switching*

- **Combine IP routing and ATM switches (also IPX and Bridging)**
- **Ipsilon Flow Management Protocol (IFMP, RFC 1953)**
  - **IP Switching nodes cooperate to cut through long-lasting flows**
  - **Key is the identification of flows of IP data -- Traffic Driven Model**
- **General Switch Management Protocol (GSMP, RFC 1987)**
  - **Connect "router brain" to ATM switch**
  - **Directs ATM switch where to direct each IP flow**
- **Note: An RFC is NOT NECESSARILY a standard**

# *Ipsilon's IP Switching (con't)*




## *Issues*

- VC required for every TCP connection - this will not scale
- Proprietary IFMP protocol must be implemented on CPE Routers (to avoid using Gateway Switch)

# *Ipsilon Flow Management Protocol*

- **IFMP is a protocol between two adjacent nodes**
- **A node tells its neighbor: "In the future, use this label for this flow"**
  - **Downstream node gives label to upstream neighbor**
  - **Flow may be host to host, or host+port to host+port**
  - **Flows follow path provided by normal IP routing**
- **If my downstream neighbor gives me a label, and I have given a corresponding label upstream, then I can "Label Swap"**
  - **This is similar to MPLS local control**
  - **Like MPLS, labels are unidirectional**
- **Very fine grain labels require traffic-driven label assignment**

# General Switch Management Protocol

- **Protocol between the "IP Switch Controller" (IP router brain) and the ATM switch**
- **Implies that only a very simple protocol (the ATM switch side of GSMP) needs to be implemented on an ATM switch. In conjunction with the switch controller, this allows the ATM switch to operate like a router**

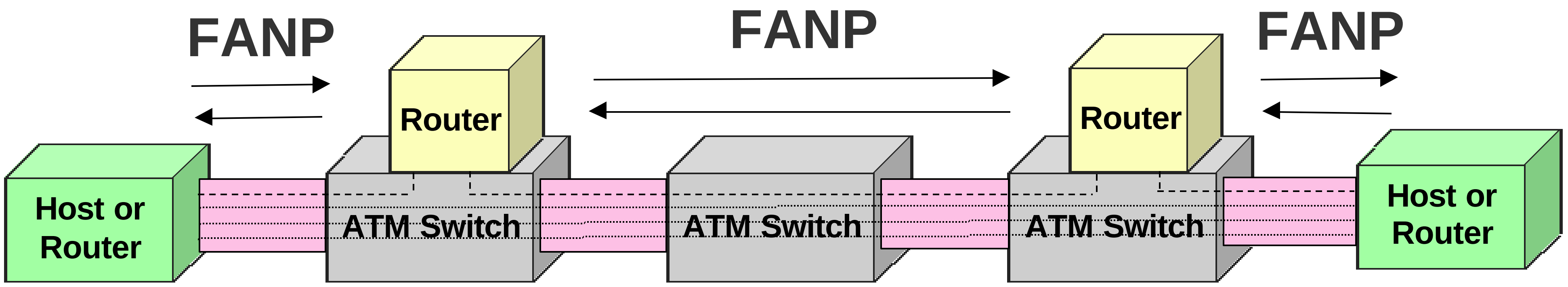
# *IP Navigator vs. Ipsilon*

- **Ipsilon's approach is based on fine grain host to host flows, or even application to application flows**
  - **This might be okay in the campus environment**
  - **This implies a severe scaling problem in ISPs and major backbones (latency of call setup, amount of state information)**
- **Ipsilon might allow traffic engineering, but no work along these lines**
- **Ipsilon is pretty much ignoring standards**
  - **In particular, they are not involved in the MPLS effort**
  - **Yet, approach requires customer premise equipment to support IFMP**
- **IP Navigator therefore has huge advantages over Ipsilon's IP switching in major backbones and Internet Service Providers**
  - **Scaling, latency, standardization, traffic engineering**

# Cell Switch Router

- CSRs are combined ATM Switches and Routers
  - CSRs can be connected either directly, or via ATM switches
  - Similar to Ipsilon approach, but ATM switches can be mixed in
  - For supporting ATM service, CSRs can act as normal ATM switches
- Default path via routers is available
- VCs are allocated to individual flows dynamically
  - Operation of the CSR is flow based: ID Allocations are always triggered by the arrival of data packets ("trigger packets")
  - Creates cut-through, but along same path of CSRs followed by IP routing
- Uses the Flow Attribute Notification Protocol to exchange bindings

# Cell Switch Router (con't)




## *Note*

- Can run through ordinary ATM switch
- Requires FANP in CPE host or router
- Like Ipsilon, uses fine-grained flows

# *Flow Attribute Notification Protocol (FANP)*

- **Runs between neighboring CSRs**
- **Exchanges bindings: IP source/destination address to VCID**
- **Bindings are allocated by the upstream CSR, told to the downstream CSR (opposite direction to Ipsilon label bindings)**
- **Soft State: Downstream CSR periodically refreshes the state, state is reset by timeout**
- **If state is lost, packets can be forwarded as datagrams**
- **Route computation is based on normal IP routing protocols (and PNNI for ATM SVCs)**

# *Ipsilon and CSR: I Don't Get It!*

- **For ISPs, carriers, major corporate backbones**
  - **Need topology-driven VCs / LSPs**
- **For Campus Networks**
  - **Need rich routing support**
  - **Need LANE support, virtual networks**
  - **Why bother with ATM in the campus if you don't want virtual networks?**
- **Thus, I don't understand where Ipsilon's approach or the CSR are going to be used**
  - **Perhaps there's a use to support IP over ATM to the home??**
  - **At best, it is one way to build a fast router, except…**
  - **To actually go fast, Ipsilon / CSR routers need to be fed by a "participating" router (or Gateway node for Ipsilon)**

# *IP Over ATM: Other Alternatives*

- **There are many other approaches available for running IP over ATM**
  - **RFC 1483 Encapsulation**
  - **LAN Emulation**
  - **RFC 1577 Classical IP Over ATM**
  - **Next Hop Resolution Protocol (NHRP)**
  - **Multiprotocol over ATM (MPOA)**
  - **Integrated PNNI**

# RFC 1483 Encapsulation

- Provides format for Multiprotocol encapsulation over ATM
- Allows ATM VCs to be used as point to point links (e.g., between routers)
- This is a limited approach, but very useful
  - Point to point links can be used to interconnect routers, to join routers to hosts, etc..
  - Reconfiguration becomes easy
- Same encapsulation can be used with other techniques

# ATM Forum LAN Emulation

- Allows set of ATM attached devices to operate as on common LAN (Ethernet or Token Ring)
- Devices which are attached to the emulated LAN can be physically distributed on an ATM subnet
- Flexible approach
  - "Nearly every protocol" can run over LAN Emulation
  - Emulated LANs can be bridged to "real" LANs
  - Must interconnect multiple ELANs with routers
- Supports Virtual Networks, easy reconfiguration, ATM bandwidths
- Scales like Bridging, includes MAC header
- Inefficient, uses "single point of failure" servers (they are working on fixing this)

# *RFC 1577 "Classical IP"*

- **Defines Logical IP Subnets**
  - **Conceptually similar to LANE, but:**
  - **Without the MAC headers**
  - **Optimized for IP (more efficient than LANE)**
  - **Defines operation of IP subnets over ATM**
- **As with LANE, defines operation within one subnet, interconnection is possible via routers**
- **Uses non-redundant ATMARP server (this is being fixed in IETF)**

# *Next Hop Resolution Protocol (NHRP)*

- **Classical IP allows operation within a Logical IP Subnet**
  - **LIS interconnection requires routers**
- **Implies “extra hops” across an ATM network**
  - **Packets traverse routers needlessly**
- **NHRP can be used to “shortcut” these extra router hops**
  - **Based on Query / Response model, uses redundant servers to translate dest. IP address to ATM address**
  - **Shortcuts are created by opening SVCs to carry observed data traffic**

# *Multi-Protocols over ATM (MPOA)*

- **MPOA is heavily oriented towards enterprise networks**
  - **Combines ATM Forum LANE with NHRP, supports virtual LANs and separated forwarding and routing**
  - **IP-address-to-ATM-address mapping requires "per host" granularity**
  - **This in turn requires query-based model, uses non-redundant servers (including route servers!)**
- **MPOA requires traffic-generated VC establishment**
  - **Significant latency issue**
  - **Serious scaling issue for ISPs, large backbone networks**

# *Classical IP, LANE or MPOA in the WAN?*

- **Classical IP, LANE and MPOA are heavily oriented towards virtual networks**
- **Virtual networks requires per-host granularity, which requires traffic-generated VC establishment**
  - **Significant latency issue**
  - **Probably makes sense in a campus with good data locality**
  - **Serious scaling issue if adopted for ISPs, large backbone networks**

# *Classical IP, LANE or MPOA in the WAN? (con't)*

- **A node's IP address (and MAC address) is independent from its ATM address**
  - **A host can be moved on the ATM network without any change in its IP and MAC addresses**
  - **MAC address and IP address to ATM address mapping requires "per host" granularity**
- **This requires servers, plus a query-based model**
  - **LE-ARP, ATMARP, NHRP, MPOA router servers**
  - **Serious reliability concern**
- **Classical IP, LANE, and MPOA are primarily appropriate for campus nets, and inappropriate for the WAN**

# Integrated PNNI

- Single Routing Protocol for support of ATM and IP
  - Similar motivation to the routing aspect of MPLS
  - Extendible to other protocols
- Based on extension of ATM Forum PNNI routing, provide QoS routing for IP
- Compatible with ATM switches running standard PNNI
- Requires SVCs to be opened to carry observed user traffic
- Not much ongoing work, overtaken by events

# *Summary: IP Navigator Benefits*

- **Designed for carriers, ISPs, major backbones**
  - **Carrier-class reliability, scalability, and availability**
- **Optimizes Infrastructure**
  - **Scalability to meet exponential IP growth**
  - **Significant performance increase**
  - **Premium services: Traffic Engineering, End-to-end QoS, VPNs**
- **Financial Benefits**
  - **Investment protection (software upgrade to existing network)**
  - **Compatibility with existing and emerging standards (MPLS)**
  - **Flexible multiservice platform (ATM and FR in addition to IP)**
  - **Incremental revenue from premium services**

# *IP Navigator Availability*

- **Software upgrade**
  - **Support for OSPF, RIP-2, and BGP-4 protocols**
  - **Greater than 1/4 million IP routes supported**
  - **Natural evolution of Ascend's WAN switching architecture**
- **IP Navigator Upgrade has been in Beta Sites since August 1997**
- **First Customer Ship on October 31, 1997**

# *Emerald Release Features*

## B-STDX 9000 - Release 5.0

- Support for native IP
- OSPF, RIP-2, BGP-4, TCP/IP, Static Routing
- OSPF Areas
- Traffic Filters
- Provisioned QoS for IP Virtual Paths

## CBX 500 Release 2.5

- Support IP virtual path transport for 9000

## Network Management

- All new IP Navigator features supported, standard MIBs

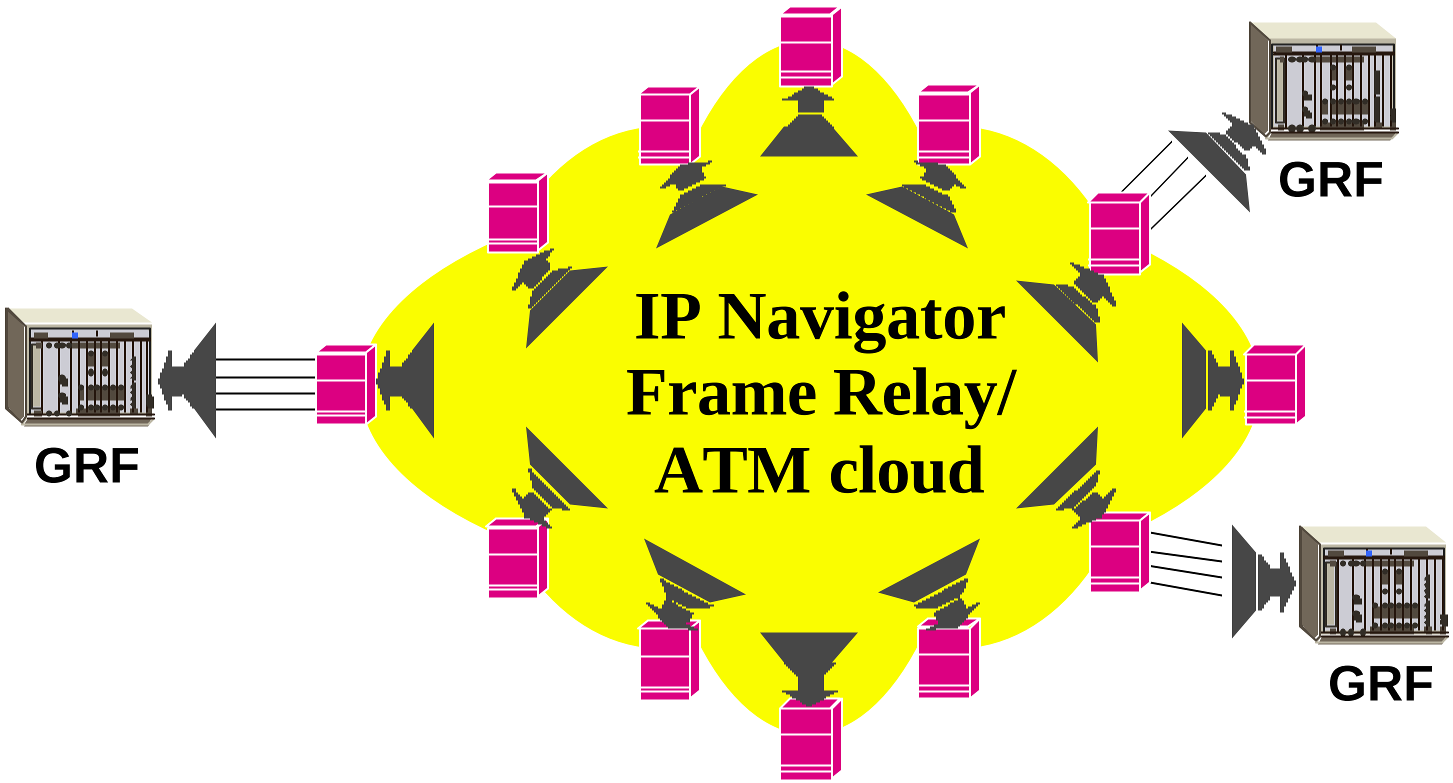
# Jade Release Features

- **B-STDX Release 6.0**
  - **Multicast Protocols: DVMRP, MOSPF, IGMP**
  - **QoS for IP MPTs**
  - **100MB Fast Ethernet Interface Card**
- **CBX 500 - Release 3.0**
  - **Native IP support on ATM access**
  - **10/100MB Ethernet, 6 Port DS3 Frame Relay Cards**
  - **NHRP/Server Card**
- **GX 550 (Garnet) - Release 1.0**
  - **IP Navigator across all ATM trunks**
- **Network Management**
  - **Simple Command Line Interface**

# *Subsequent Release Features*

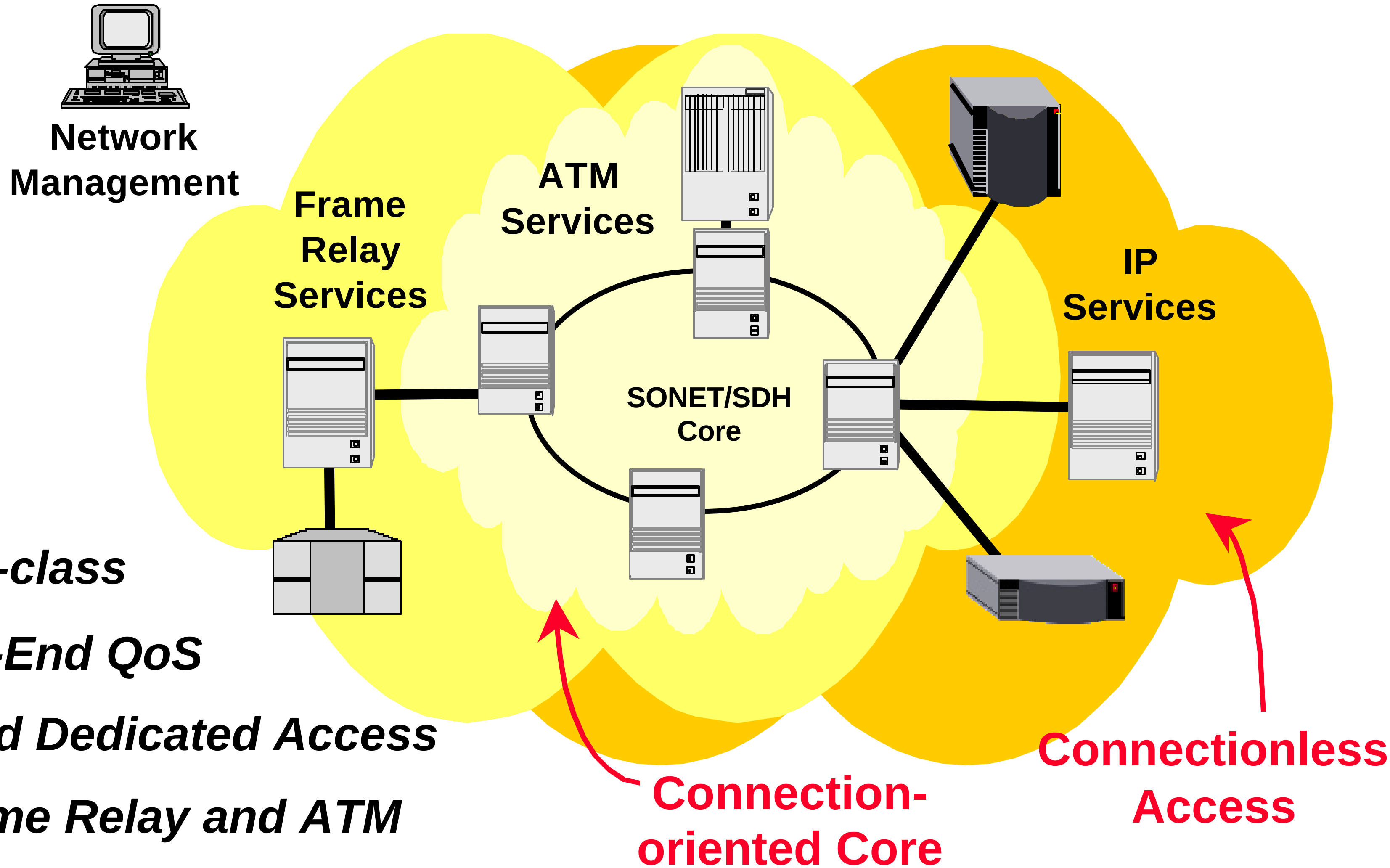
- **B-STDX 9000**
  - **Dial-in Access, Radius integration**
  - **Virtual Private Networks**
  - **Larger Routing tables supported**
- **CBX 500**
  - **6 Port DS3 ATM Interface Card**
  - **Virtual Private Networks**
  - **Larger Routing tables supported**
- **GX 550**
  - **Frame Interfaces**

# *GRF/IP Navigator Integration Plans*




- GRF implements MPT stub via OSPF, ARP
- No IP processing on 9000/500, switching only
- GRF will implement MPLS when standard is available

# The New Public Multiservice Network




- *Carrier-class*
- *End-to-End QoS*
- *Dial and Dedicated Access*
- *IP, Frame Relay and ATM*
- *Integrated Network Management*

ASCEND

# *Thank you!*