Practical Layer 2 and Layer 3 Alternatives for Open Access Connectivity

Wes Berkey
Manager, Digital Transport Network Engineering
Transmission Network Systems
Scientific-Atlanta, Inc.

5030 Sugarloaf Parkway
Lawrenceville, GA 30042
Telephone: 770-236-7767
E-mail: **wes.berkey@sciatl.com**

Executive Summary

This paper describes the fundamental issues associated with providing open access connectivity in Data Over Cable Service Interface Specification (DOCSIS) 1.1 networks, examines and compares some practical transport alternatives available to Multi-System Operators (MSOs) planning to enable open access, offers a roadmap for cost-effective initial implementation and migration for open access connectivity, and provides some suggestions for selecting the appropriate initial architecture according to individual business needs.

Fundamentals of Open Access Transport Networks

Leading cable operators have been experimenting with a number of existing and emerging technologies that can be used to deliver open access connectivity from the CMTS to one or more open access peering points. The hype surrounding open access and the availability of multiple competing technologies and products that are being positioned as open access enablers can make it easy to lose sight of the fundamental issues and requirements for providing access to multiple ISPs over MSO networks.

Although there is no standard approach to open access network design, it is generally accepted that the transport network should have the following qualities:

1) Ability to provision multiple providers
2) IP address management
3) QoS
4) Traffic separation
5) Latency management
6) Rate limiting and filtering
7) Interoperability

Several technologies are available to enable open access, although a few have significant drawbacks that will limit their effectiveness. All of the approaches listed in Table 1 provide the ability to connect multiple Internet Service Providers (ISPs) to the MSO packet transport network.

Table 1 Open access enabling technologies

Technology	Advantages	Disadvantages
Parallel Network	Simple provisioning	Lack of available RF bandwidth High cost (duplicate components) Doesn't scale well
Network Address Translation (NAT)	Traffic security	No path awareness Doesn't scale well

IP tunneling	Scalable Differentiates traffic through encapsulation	Multiple applications share same tunnel, QoS Requires routing at the edge
Policy based routing or switching	Scalable Packet forwarding based on a wide range of attributes	Requires intelligent devices at the edge
Multiprotocol Label Switching (MPLS) with policy based routing or switching	Scalable Low latency Enables service level agreements Packet forwarding based on a wide range of attributes	New technology Requires intelligent devices at the edge

Parallel Network

In order to implement a parallel network, the MSO has to allocate additional forward and reverse bandwidth in the Hybrid Fiber-Coax (HFC) plant in order to provision multiple ISPs. This solution is impractical in today's HFC network where bandwidth is at a premium, and because a parallel network would have to be installed for each ISP, it is not scalable.

Network Address Translation (NAT)

NAT provides address translation from a local IP address to a globally unique IP address. The NAT device is aware of the incoming IP address, however it has no awareness of the overall inbound or outbound path. This technology is simple, and may be suitable for some deployments, however it is limited in scalability and service differentiation.

IP tunneling

Tunneling operates by encapsulating a network protocol within packets carried by a second network. In this case a customer's packets would be encapsulated in IP for transport to the ISP. Since multiple applications in the same packet flow to a user must share the same tunnel, it is difficult to enable multicast services (enable other users to access the packets) and provide Quality of Service (QoS) where different applications could be given priority.

Policy Based Routing or Switching

Routing or switching decisions are based on a Service Level Agreement (SLA) that will specify destination and may specify guaranteed data rate, QoS, and primary and secondary path. The identification and classification of the packet at the edge can be based upon a variety of attributes such as source address, destination address, service flow identification, Type of Service (TOS), and Virtual Local Area Network (VLAN) tag. Although having the ability to perform packet classification and policing can add cost to a network device, the operator can use these devices to optimize a given network's performance. Devices that can perform policy based routing or switching are Layer 3 routers, RPR enabled switches, and ATM switches.

Multi-Protocol Label Switching (MPLS) with Policy Based Routing or Switching

Adding MPLS to a policy based routing or switching solution provides improved latency through the transport network. Label Switched Paths (LSPs) are utilized to provide hardware-based packet forwarding, which is considered key to video and voice applications. A software-based router can take up to 50 ms to process a packet, and the minimum generally accepted round trip latency in a VoIP network is approx 300 ms. This means that the VoIP traffic could traverse only three software-based routers before the quality became objectionable! Employing MPLS can reduce the device processing time to <10 ms, decreasing overall latency.

Transport Network Requirements

It is useful to define a roadmap for cost-effective migration from a simple network architecture that meets initial open access requirements to more complex architectures that meet future needs. One of the key roadmap issues is the cost-effective migration of layer 3 routing from the core of the network, the headend or regional data center, to the edge of the network, the hub. Many operators are currently evaluating network enhancements to deploy both open access and DOCSIS 1.1 and examining the tradeoff between centralized and distributed routing architectures. In order to better understand this key tradeoff, it is important to examine how layer 2 and layer 3 infrastructures support the fundamental open access connectivity requirements, and how each has a role in providing a complete solution.

Table 2 Transport requirements for Layer 2 and Layer 3

Transport Requirement	**Layer 2 Implementation**	**Layer 3 Implementation**
Multiple ISP connections	Layer 2 CMTS at hubs (requires Layer 3 router at Headend)	Layer 3 CMTS or router at each location
IP address management	Multiple IP subnets extended at Layer 2 via RPR, MPLS, VLAN	Multiple IP subnets
QoS	1) Wire model transport (Layer 1) 2) MPLS 3) Resilient Packet Ring (RPR)	1) Wire model transport (Layer 1) 2) MPLS (Layer 2) 3) Policy Based Routing
Traffic separation	1) VLAN 2) MPLS Layer 2 VPN 3) RPR	1) Multiple IP subnets 2) MPLS Layer 2 or Layer 3 VPN 3) VLAN (Layer 2) 4) RPR (Layer 2)
Latency management	Hardware based forwarding	1) MPLS

		2) Hardware based forwarding (Layer 2)
Rate Limiting and Filtering	MPLS, ATM, Resilient Packet Ring (RPR)	1) Policy Based Routing 2) Layer 2 MPLS, ATM, RPR
Interoperability	Ethernet, ATM, PPP, Frame Relay, RPR	1) Standard IP routing protocols 2) Layer 2 Ethernet, ATM, PPP, Frame Relay, RPR

Ability to provision multiple providers

The transport network should be capable of connectivity to multiple ISPs, potentially at multiple locations. This requirement is normally accomplished with a Layer 3 border router, typically located at a central location (eg. Headend.) The Layer 3 router provides inbound/outbound path awareness in addition to IP address translation and control for security. Connectivity to an ISP is established at a mutually agreed upon capacity (eg. GigE, 100bT, DS-3, etc.)

IP address management

The transport network has to be able to support multiple IP subnets which may be accomplished with distributed Layer 3 routers or by extending subnets via Layer 2 transport. A variety of technologies including RPR, MPLS, and VLANs may be deployed to enable separate broadcast domains for these IP subnets. VLANs can be implemented across Layer 2 switches, however the theoretical limit is 2048 separate VLANs. RPR enabled switches provide transparent LANs through Label Switched Paths (LSPs) as well as VLANs. MPLS can be operated over a routed network or ATM to provide the necessary traffic separation.

QoS

Packet classification enables the operator to differentiate packets by a variety of attributes and with queuing mechanisms, ensure service levels. In a DOCSIS 1.1 environment it is likely that a packet will be classified based on its service flow ID (SFID) which can be used from CMTS to cable modem to indicate the user's application. QoS can be provided by a Layer 2 transport (eg. ATM, RPR) or Layer 3 IP router.

Traffic separation

Packet security is provided at Layer 2 with MPLS VPN, or transparent LAN, or ATM virtual circuits, or at Layer 3 with IP subnet, routing protocols or tunneling.

Latency management

Hardware based forwarding limits the examination of the packet header in order to speed intermediate processing between ingress and egress thus facilitating the transport of delay sensitive traffic such as voice or video. Intelligent Layer 2 technologies can employ packet filtering and classification of services while retaining hardware based forwarding of packets. Layer 3 routing can be deployed with MPLS or RPR and can also take advantage of hardware based packet forwarding.

Rate limiting and filtering

Hardware-based routers, and MPLS/RPR-enabled CMTS and transport devices perform rate limiting at wire speed. Traditional software-based routers offer capability to rate limit and filter services to manage network congestion and ensure access fairness. However, in traditional networks, turning on these features may compromise performance by increasing latency. A software-based router can add as much as 50 ms of latency due to look up tables and filtering.

Interoperability

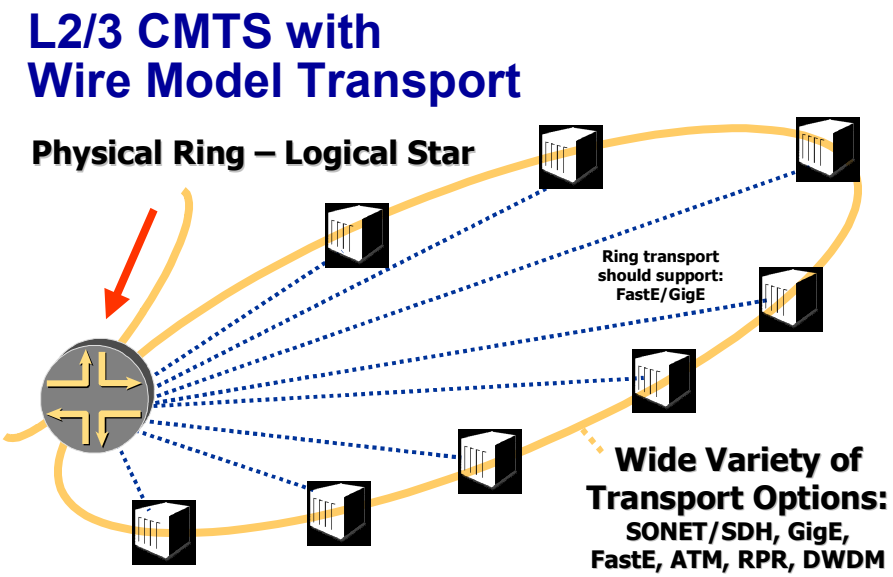
Standards based interfaces (GigE, FastE, OC-12, OC-48, etc.) are available for both Layer 2 and Layer 3 transport solutions. Establishing independent IP routing capability at the Hub requires distributed Layer 3 IP routers which may be required for some enterprise business traffic. For example, if a user requires the operator to manage their IP addresses at multiple locations off of one Hub, a Layer 3 router would prevent that user's traffic from having to traverse the network to a centralized router to perform IP address management. Layer 2 solutions are interoperable with a variety of protocols (Ethernet, RPR, ATM, Frame Relay), and can provide transparent carriage of Layer 3 services.

By building today's network with a vision of future requirements, operators can avoid potential pitfalls that can result in false starts or costly rework as the network evolves to meet future needs.

Build a Transport Network that Migrates

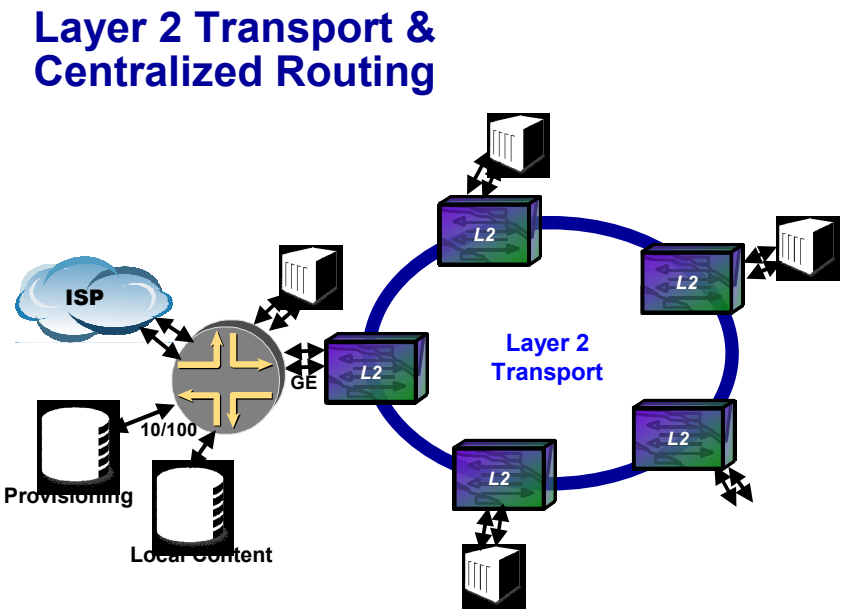
The operator can choose a starting point that meets or exceeds the initial needs of the network according to immediate business requirements, growth projections, capital constraints, and the desired risk profile for the project. In other words, how much capital is the operator willing to invest in the network based on cost of deployment vs. revenue plans and projected penetration of services. For some this may mean starting with simple physical layer connectivity between headends and hubs. Others will elect to deploy fully decentralized solutions with large hub routers supporting clusters of CMTS's at each hub. Most will choose a point in the migration between these two extremes. In fact, a single MSO might choose two or three different starting points for different systems according to expected demand for services in each system.

Illustrated in Drawing 1 is a wire model implementation over a ring-based architecture that will work with either Layer 2 or Layer 3 CMTS. As a baseline architecture, the MSO can deploy a centralized border router to manage the IP addressing and connections to multiple ISPs, and implement low cost extensions using DWDM or SONET that connect Hubs with Layer 2 CMTS. The typical interface to a CMTS is 100bT or GigE, although there are Packet Over SONET (POS) interfaces as well. If sophisticated services such as Layer 3 VPN, deep packet filtering, or MSO-managed, routing protocol interoperability is required, a Layer 3 CMTS, or an edge router connecting Layer 2 CMTS, may be deployed strategically at individual hubs. As the network scales to accommodate increased penetration or additional hubs, the transport network may be migrated to a statistically multiplexed backbone.



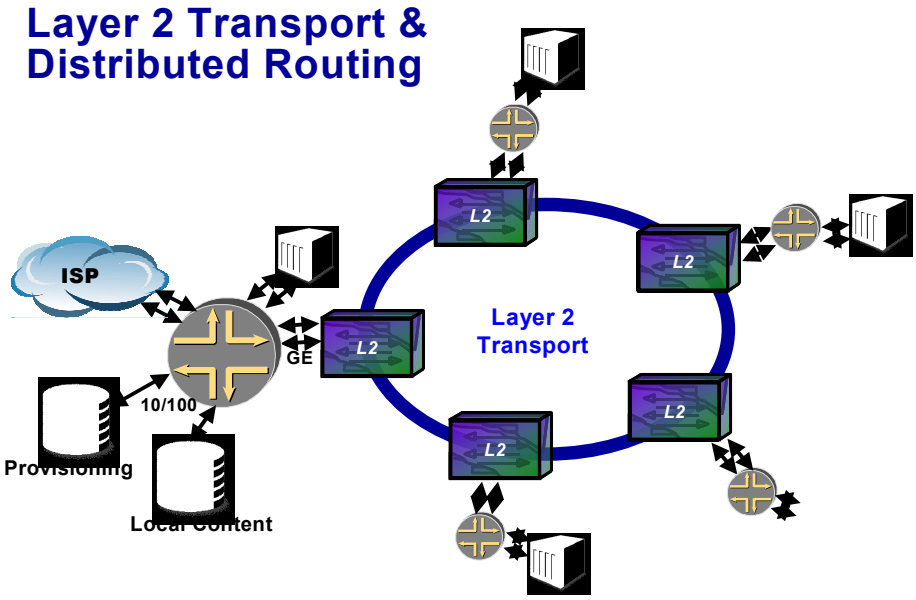
Drawing 1 Layer 2 or Layer 3 CMTS with Wire Model

Drawing 2 shows a statistically multiplexed Layer 2 transport with a centralized border router. This architecture scales better than a wire model, and offers over-provisioning in the backbone to take advantage of typical cable modem data applications that are bursty and not delay sensitive.



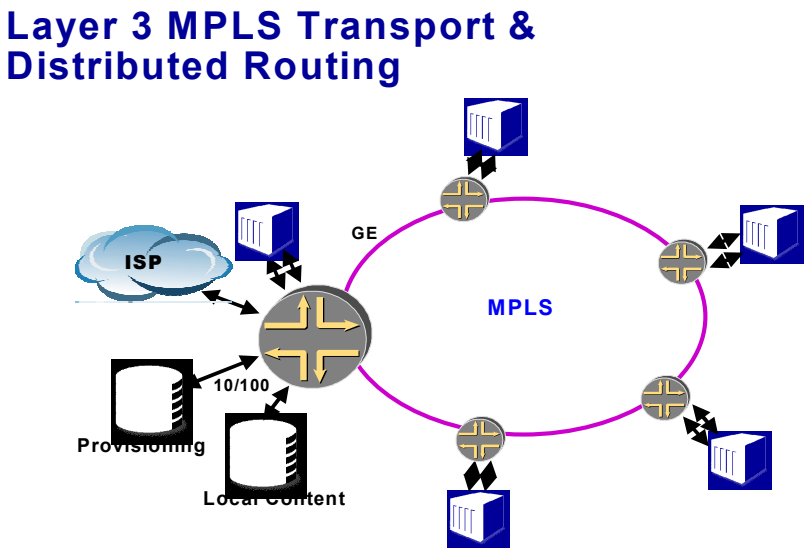
Drawing 2 Centralized Routing with Layer 2 Transport

If Layer 3 services (Deep packet filtering, Layer 3 VPN, routing protocol interoperability) are a requirement at a hub, the architecture depicted in Drawing 3 could be implemented. Still taking advantage of an intelligent Layer 2 transport with statistical multiplexing, routers or Layer 3 CMTS are deployed at hubs.



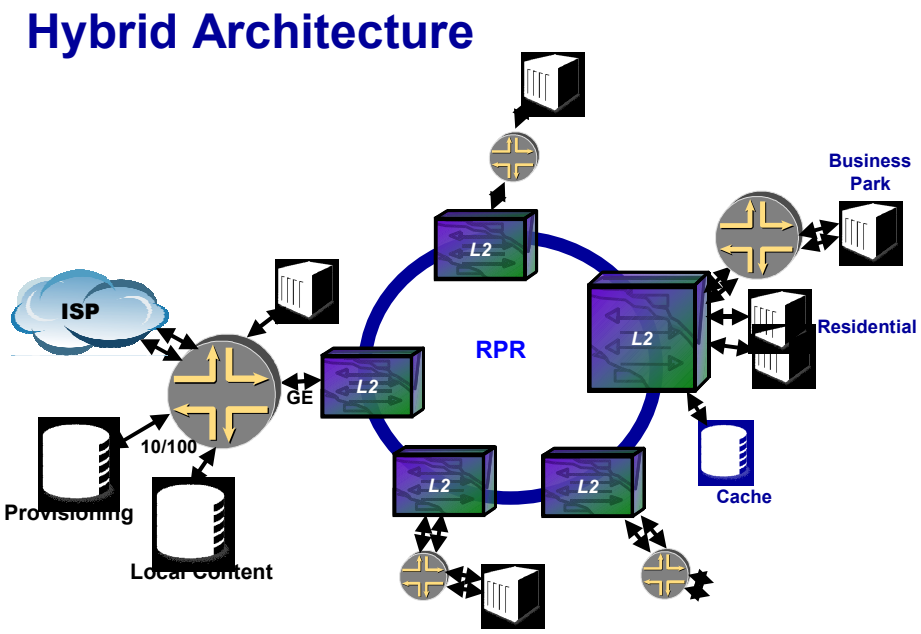
Drawing 3 Distributed Routing with Layer 2 Transport

Running MPLS on a Layer 3 transport network improves the speed of packet processing by providing hardware-based packet decisions, and ensures low latency for delay sensitive services such as video or voice. Drawing 4 shows an intelligent, statistically multiplexed Layer 3 transport with MPLS with Layer 3 CMTS deployed at the Hubs. The MSO has to balance the higher cost of this architecture with revenue generating Layer 3 services.



Drawing 4 Distributed Routing with MPLS Layer 3 Transport

It is possible to implement a hybrid architecture where Layer 3 CMTS or edge routers are deployed strategically throughout the network where required to create Layer 3 service offerings. An intelligent Layer 2 transport running RPR can forward packets to a border router or edge router based on Service Flow ID (SFID) using VLAN tagging.



Drawing 5 Hybrid Architecture with RPR Transport

Conclusion

By focusing on the fundamental problem of open access connectivity, carefully evaluating alternative architectures and technologies in the context of the central problem, beginning with the end in mind, and investing in accordance with projected growth while minimizing risk, MSOs can confidently plan for the cost-effective introduction of open access into their networks while allowing for scalable migration to meet future needs.

Glossary of Terms

Border Router

This term is generally used to describe the device that defines a boundary between autonomous systems. As opposed to a “network router” operating in the interior of an autonomous network, the border router operates at the edge, and makes the final determination of the external destination of a packet that has traversed the internal network. The boundary between networks is generally determined by ownership of the network devices.

Deep Packet Filtering

Deep packet filtering is a method of classifying and policing packets by identifying their source and purpose in the network. The word “deep” implies that the device performing the filtering is looking beyond the source and destination address available in an IP or Ethernet packet, to the Session Layer (5), Presentation Layer (6), or Application Layer (7) from the Open Systems Interconnection (OSI) reference model.

Enterprise Network

An organization's internal communications network. Usually carries multiple types of traffic (data, voice, and video) between multiple sites within the enterprise, and integrates all the computing systems within an organization, including DOS, Windows®, Mac®, UNIX® workstations.

Ethernet

The most popular of several LAN types, often used by desktop computers and servers to access networks. In particular, Ethernet dates from the early 1980s, when a consortium of DEC, Intel, and Xerox published the Ethernet "DIX" protocol definition.
The real advantage of Ethernet LAN technology is that its computer interface is a very inexpensive commodity item.
Originally, Ethernet was implemented as a coaxial cable copper pair as a "shared bus," in which all packets were sent onto the bus one at a time. As is necessary, each packet was available to all computer interfaces on the bus; the particular computer is addressed by a 48-bit address administered into every interface and in the Ethernet packet header. The address format is termed a Media Access Control address, or MAC address. Note that the shared bus acts like a switch, but not in the modern sense.
In order to "share" the bus an access scheme, called the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol, was devised. This defined a procedure allowing a sender to access the bus in competition with others also seeking access. This required both a minimum (64-octet) and maximum (1520-octet) Ethernet packet size, and bus limited to 500 meters in length.
Subsequently the IEEE took over the standardization of LANs, and DIX Ethernet became modified and termed "802.3." In reality, all Ethernet is still compliant with the DIX.

Frame Relay

A simple connection-oriented Layer 2 protocol specified by the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) for the transfer of information between two compatible endpoints. The Frame Relay protocol defines a simplified protocol with frame delimiters (flags), virtual connection identifiers (called DLCIs), congestion indication and discard eligibility bits and error detection

capability. The variable-length frame can typically be up to 4,096 bytes long. Frame Relay does not include control procedures such as retransmission or flow control and is optimized for low error-rate networks. Frame Relay is specified in ITU-T Recommendation Q.922.

Label Switched Path (LSP)

A LSP is the path established by the higher layer protocol used to deterministically forward traffic across an MPLS network. By definition, a LSP is unidirectional and return traffic must have a separate LSP defined, which may mean unequal delay, jitter, and other QoS attributes.

Layer 2 (Data Link Layer)

The part of the OSI Reference Model that is responsible for transparent transport of Layer 3 information between adjacent nodes in a network across an individual physical link. Defines formats for data transmission (e.g., frames or cells.)
In connection-oriented implementations, includes procedures for establishing and maintaining the Layer 2 connection between adjacent nodes as well as the format and procedures for data transport including error detection. May also include facilities for flow control and retransmission of corrupted data. An example would be High-level Data Link Control (HDLC.) In connectionless implementations, includes a frame format and an addressing scheme. Examples include the Media Access Control (MAC) protocols of Ethernet, Token Ring, and FDDI.
Frame Relay and ATM are also considered to be variations of Data Link Layer protocols which support multiplexing.

Layer 3 (Network Layer)

The part of the OSI Reference Model dealing with network addressing, routing and switching of data. Includes acknowledgments that an entire message is correctly received. May include breaking a Layer 4 message into packets of suitable transmission size. IP and X.25 are examples.

Multi-Protocol Label Switching (MPLS)

IETF specification for transporting Layer 3 data units over a network of Layer 3 routers and/or Layer 2 (label) switches, which shortcuts the full Layer 3 routing process in each router. Therefore, when MPLS protocols and procedures are added to a network of routers and switches, it is expected to improve performance relative to traditional Layer 3 routing by reducing processing time.
The major feature of MPLS is its ability to identify flows of traffic to a common destination, attach an extra protocol header (label) to packets of those flows, then switch those packets by consulting the label, not the entire Layer 3 destination address.

Over-provisioning or Over-subscription

The ability to assign more aggregate traffic to a network than can be carried on it. Packet switched networks can statistically multiplex many applications and/or users onto a single line, or trunk between packet switches. By over-subscribing the network, the operator can maximize the efficiency of the network facilities. In an over-subscribed network it is statistically unlikely that all of the access interfaces will demand their maximum bandwidth all at the same time, thus allowing a single shared network to efficiently supply bandwidth to multiple services.

Point-to-Point Protocol (PPP)

Point-to-Point Protocol is a connection-oriented, non-switched Data Link Layer (Layer 2) protocol defined by the Internet Engineering Task Force (IETF), considered part of the IP protocol suite. PPP offers many important services concerning the management of a point-to-point serial line physical channel, including QoS metrics, authentication, the assignment of IP addresses, etc. Most common Data Link Layer protocols, like Frame Relay, ATM, MPLS, and all LAN technology (e.g., Ethernet) are switched, thereby are unable to provide these services, resulting in the "tunneling" of PPP through these as Point-to-Point Protocol over Ethernet (PPPoE), Point-to-Point Protocol over ATM (PPPoA), etc.

Policy based routing

Allows operators to define their own policies as to how a packet gets routed in their network. Policy routing provides a mechanism that selectively cause packets to take different paths, and marks packets so that certain kinds of traffic receive differentiated, preferential service when used in combination with queuing techniques. These queuing techniques provide an extremely powerful, simple, and flexible tool to network managers who implement routing policies in their networks.

Rate limiting at wire speed

Provides the ability to enable rate limiting on a packet flow without impacting latency. The term wire speed characterizes the addition of a function or process to a network device while maintaining traffic flow as if it were just traveling down a wire.

Resilient Packet Ring (RPR)

Resilient Packet Ring is an emerging network architecture and technology, currently being standardized in the IEEE 802.17 Working Group, which is designed to provide the best features of SONET, ATM, and Ethernet in order to meet the needs of service providers as they migrate their networks to accommodate the high growth of packet-based services, while continuing to support circuit-based and legacy services. RPR offers <50 ms resiliency, packet classification and QoS, traffic engineering, and simplified service provisioning.

Standard IP routing protocols

RIP – Routing Information Protocol is a protocol for exchanging network reachability (hop count) and routing information between routers in a router-based network.

OSPF - Open Shortest Path First is an alternative to RIP as an Interior Gateway Protocol (IGP). It is a link-state protocol, as opposed to RIP , which is distance-vector protocol. In a link-state protocol each router actively tests the status of its link to each of its neighbors, sends this information to its other neighbors and so on. Each router takes this link-state information and builds a complete routing table. This method is much faster then the distance-vector protocols, especially in case of changes in the links in the network.

IS-IS - Intermediate System-to-Intermediate System is an OSI link-state hierarchical routing protocol that floods the network with link-state information to build a complete, consistent picture of network topology. To simplify router design and operation, IS-IS distinguishes between Level 1 and Level 2 ISs. Level 1 ISs communicate with other Level 1 ISs in the same area. Level 2 ISs route between Level 1 areas and form an intradomain routing backbone.

BGP - The primary function of a BGP speaking system is to exchange network reachability information with other BGP systems. This network reachability information includes information on the full path of Autonomous Systems (ASs) that traffic must transit to reach these networks. This information is sufficient to construct a graph of AS connectivity from which routing loops may be pruned and some policy decisions at the AS level may be enforced.

VLAN

A Virtual Local Area Network is a group of PCs, servers and other network resources that behave as if they were connected to a single, network segment, even though they may not be. When a packet enters its local switch, the determination of its VLAN membership can be port-based, MAC-based or protocol-based. When the packet travels to other switches, the determination of VLAN membership for that packet can be either implicit (using the MAC address) or explicit (using a tag that was added by the first switch). Port-based and protocol-based VLANs use explicit tagging as their preferred indication method.

Wire model transport

Refers to a data network that is designed to provide dedicated connections from point-to-point in a star or mesh configuration. The contention for bandwidth in a wire model is only among services or users on the same connection.

Bibliography

1) “Delivering Internet Connections Over Cable, Breaking the Access Barrier”
Authors: Mark E. Laubach, David J. Farber, Stephen D. Dukes
Published: Wiley Computer Publishing, 2001

2) “MPLS Based Layer 2 Virtual Private Networks”
Authors: Kireeti Kompella, Matt Kolon, Pierre Bichon, Annette Kay Donnell
Published: Juniper Networks, 2001

3) “RPR and MPLS, in Heterogeneous Network Architectures”
Author: Lorenzo Bombelli
Published: Scientific-Atlanta, Inc., 2002

4) “Understanding Latency in IP Telephony”
Author: Alan Percy
Published: TelephonyWorld.com

5) Telecommunications Research Associates LexiCat®
Author: TRA