Utah's UTOPIA: An Ethernet-Based MPLS/VPLS Triple Play Deployment

Ken Moerman and Jeff Fishburn, DynamicCity
Marc Lasserre and David Ginsburg, Riverstone Networks

ABSTRACT

Broadband in North America is now at critical mass, passing the inflection point from the early adopter to the early majority. In fact, Strategy Analytics [1] predicts that by 2010, over 78 million U.S. households will have some form of broadband connectivity, much of it capable of delivering the triple play of voice, video, and high-speed Internet access. Their providers will consist of local exchange carriers, multiservice operators, and municipalities that have deployed their own triple play networks. Why the municipalities, and what do they bring to the table? In many instances, smaller cities and towns have been bypassed by the first phases of the incumbent operators' triple play deployments. For reasons of economic development, quality of living, education, and the retention of skilled workforces, the government has stepped in to fill the void. This article looks at the most ambitious municipality-driven triple play deployment in North America, the UTOPIA MetroNet in Utah. The system integrator for UTOPIA is DynamicCity, which early on defined a set of design principles that would make the network successful. These included the requirement that the network be open to multiple service providers, be carrier class, be scalable and based on future-proof technology, and that the architecture be based on open standards while keeping costs to a minimum. The article first reviews some of the considerations behind the choice of the technology and deployment approach based on these guidelines, and then describes in detail the network topology and services architecture, the current status of the deployment, and future plans. Specific topics covered include the use of multiprotocol label switching/virtual private LAN service, IP multicasting, and traffic engineering.

TECHNOLOGY AND DESIGN SELECTIONS

The fundamental technology selections when planning new triple play deployments relate to the control and protocol architecture, the last mile technology, and whether the network is designed as an open or closed network. We now look at each of these in the context of the UTOPIA architecture.

A LAYERED APPROACH

One of the goals in building the UTOPIA MetroNet was to build a network that is as simple and direct as possible without compromising the scalability and robustness of that network. Another goal was to build a system that is as open and "transparent" as possible to attract many service providers. For every layer of the network's architecture you interact with on the service provider's behalf, you add a level of complexity that translates into management overhead, equipment cost, and standards issues. Keeping the network design at the lowest possible level reduces the overall complexity for service providers and increases the interoperability of the network. These goals led to a decision to focus on layer 2. Multiprotocol label switching (MPLS) [2] has most of the attributes required to construct a network at this layer, including support for tunneling, traffic engineering, quality of service (QoS), and link protection.

LAST MILE ARCHITECTURE

The next issue is the selection of the last mile architecture. The industry has four major physical medias available for access infrastructures: wireless, twisted copper pairs, coaxial cable, and fiber. Each of these physical media types was analyzed to discover which best met the municipalities needs as represented by the design principles. The requirements for an open scalable network are complementary drivers with respect to the physical medium technical decision. For the network to deliver services transparent to service providers, an infrastructure capable of delivering nearly limitless capacity is required. Using a sample footprint (a couple of kilometers across) we added up the total throughput each physical medium was delivering in existing deployments.

As depicted in Fig. 1, the dedicated nature of fiber's bandwidth combined with its ability to support 100 Mb/s to 1 Gb/s speeds allows a fiber-

0163-6804/05/$20.00 © 2005 IEEE

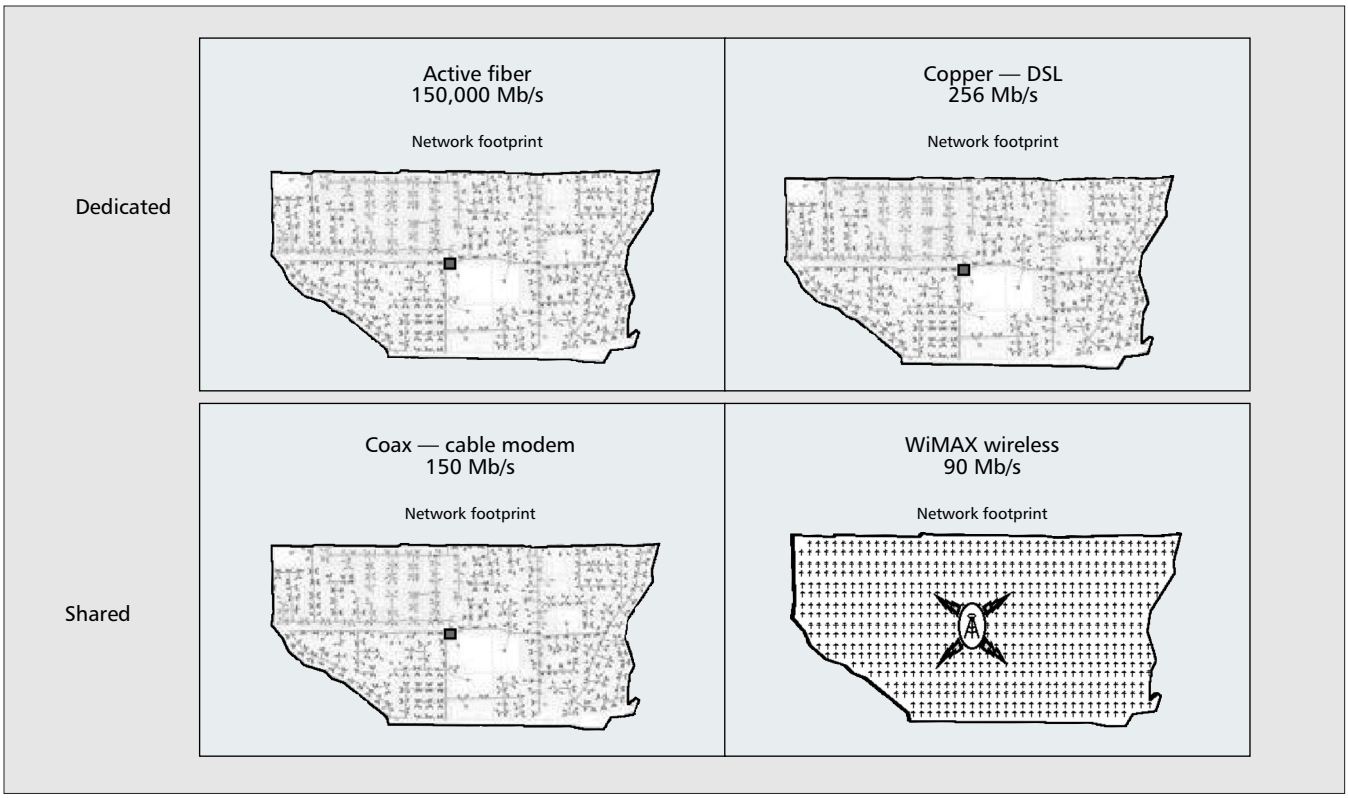

Figure 1. *Total symmetric bandwidth capacity available per footprint by physical transport medium.*

based solution to exceed alternate physical transport media by several orders of magnitude. Various fiber optic topologies, designs, and strategies exist for delivering services to subscribers.

Open and Closed Networks

An important consideration, particularly for municipal owned networks, is whether the infrastructure is designed as "open" or "closed." By open, we refer to a network designed as a utility, over which multiple service and content providers may operate. This model is ideal when the network is installed by the municipality, and is operated as a utility much like water or electricity. The actual service providers, whether they provide voice over IP (VoIP), business virtual private networks (VPNs), or video, have the direct relationship with the customer, and a given customer is not bound to the services offered by a single provider. While it is very traditional for the government to be responsible for the deployment of infrastructure (water, sewer, highway, airports, telephone), it is generally accepted that government should not be involved in what is considered the private content and services arena.

A closed network is designed to be deployed and operated by a single service provider. Here, the local exchange carrier (LEC) or multiservice operator (MSO) not only builds the network, but is also responsible for delivering the services. Although these networks theoretically support open access by other providers, in reality the technical, legal, and cost barriers to entry are too great to make this a viable model. One only need look at the MSOs or LECs and their percentage of their customers using Internet service providers (ISPs) controlled by the LECs and MSOs to draw this conclusion. This reality is not expected to change as part of their ongoing triple play deployments.

The UTOPIA architecture enables affordable high-speed communications by leveraging the ability of municipalities to access low-cost capital which in turn permits a sustainable return to service and content providers by vastly decreasing their capital expenditures. Residents and businesses contract with service providers for video, voice, data, or other services, while these service providers pay the city transport fees to use the advanced communications network. In doing this, the UTOPIA network balances technology, financing, and regulatory constraints with business and consumer demands.

Adopting a model developed by DynamicCity, the consulting firm for the UTOPIA project, the UTOPIA MetroNet operates under a model called the Open Service Provider Network™ (OSPN™) [3]. Under the OSPN model, a city constructs and maintains an advanced communications network as basic city infrastructure, an independent private entity designs, operates, and maintains the city infrastructure, and private industry service providers offer services over the network (Fig. 2).

Definitions and Physical Architecture

The UTOPIA network architecture is segmented into functional layers, and Ethernet routers bridge the frontiers between these layers. Figure 3 depicts these layers and routers (referred to as switches in the figures). Although the network as

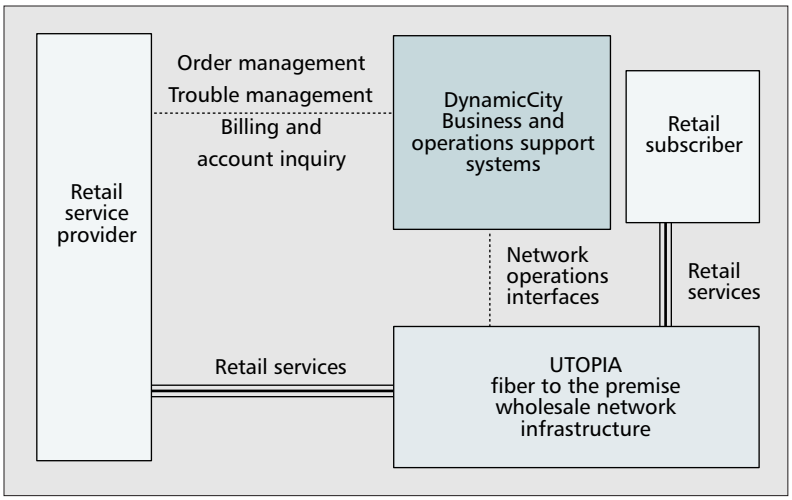

Figure 2. *Open Service Provider Network™ operations model.*

initially deployed is designed to scale to 150K households over a three hundred mile area in Utah, there is no reason why the architecture could not scale into the millions of subscribers by adding additional network elements.

The UTOPIA MetroNet has several basic divisions: the subscriber (customer premises equipment, CPE) layer, access layer, distribution layer, local core layer, provider access layer, provider layer, and regional core layer. Customers, both business and residential, connect from their CPE, an access portal (AP), which then connects via a Fast Ethernet fiber connection to an access distribution switch (ADS) located in a streetside cabinet. Each ADS is sized for on the order of 200 subscribers, and there are up to six ADSs in a cabinet. The cabinet is located up to 2 mi from the subscriber. Redundant Gigabit Ethernet fiber links extend up to five miles to a distribution core switch (DCS) serving on the order of 10,000 subscribers. This then connects via 10 Gb/s fiber to a regional core switch (RCS) pair serving up to 40,000 subscribers. DCSs are dual-homed to the RCSs, and the RCSs are meshed. The RCSs then connect to the service entry points in the network. This is the role of the core provider switch (CPS) and provider access switch (PAS), analogous to the DCS-RCS pairing on the customer side. Service provider equipment (SPE) provided by the actual service provider connects to the PAS. As additional service providers connect to the network, additional CPS-PAS pairs may be added. Later, we describe the logical topology following this physical topology.

Subscriber Layer

Drawing an analogy, just as a residential driveway is the final destination point where passengers load and unload in a road system, the subscriber layer is the part of the MetroNet where end users "load" and "unload" their particular network service. There are three types of equipment that fit in this layer: the AP, UTOPIA supplied devices (video gateways), and CPE.

Access Portal — The AP is an electronic device attached to the exterior or interior of a building much as telephone demarcation boxes are currently attached. The function of the AP is, like a phone demarc, to terminate the city's fiber and provide a connection point into which the building (business building or private residence) can connect its devices. As the demarc point, it is also the point at which the city's responsibility for many of the network services ends, and most devices connected to or located beyond the AP will be the responsibility of the end user or service provider, with the exception of the video gateway, which will be UTOPIA's responsibility.

The AP will have a number of connection points (Ethernet ports) and a number of built-in VoIP gateway analog voice ports. All devices within the building that need to connect to the MetroNet will aggregate and connect to the AP through one of those ports. The capacity of the Ethernet ports is 100 Mb/s, equal to the speed of a typical business LAN. All ports are configurable, and specific types of data traffic will be routed to the AP ports to deliver the services the consumer purchases.

Customer Premises Equipment — Some services to which customers may choose to subscribe will require a piece of equipment inside the customer's house. For example, subscribers wanting wireless connectivity in their home will have a wireless access point connected to the AP. In that case the wireless access point is owned by the subscriber or service provider. These types of devices are referred to as CPE. Each service provider requiring some type of CPE will be responsible for making it available to the customer, and maintaining and repairing it.

CPE and other equipment within the home owned by the subscriber or service provider may operate at open system interconnection (OSI) levels up to layer 7. But by terminating ownership of the MetroNet at the AP, VoIP gateway, and video gateway, UTOPIA is able to keep the OSPN principle satisfied, since no individual service provider owns the on ramp or gateway to the fiber network. As voice and video CPE evolves to support direct connection to Ethernet networks (like computers today), the voice and video gateway functionality will no longer need to be deployed to maintain subscribers' independence from the service provider.

Distribution and Access Layers

Access distribution switches at the edge of the access layer are designed to handle network traffic from the APs spread throughout a neighborhood or business district. The function of the distribution core switches at the edge of the distribution layer is to collect the traffic from the access layer and feed it upstream to the next layer, called the local core layer, which is designed to handle heavier traffic. The reverse is also true. Traffic from the core layers is channeled through the distribution and access layers for delivery to the AP. The design of the core layers "oversubscribes" traffic the same way residential and city streets oversubscribe the number of cars they can handle at any point in time, while at the same time supporting the requirements for guaranteed QoS traffic.

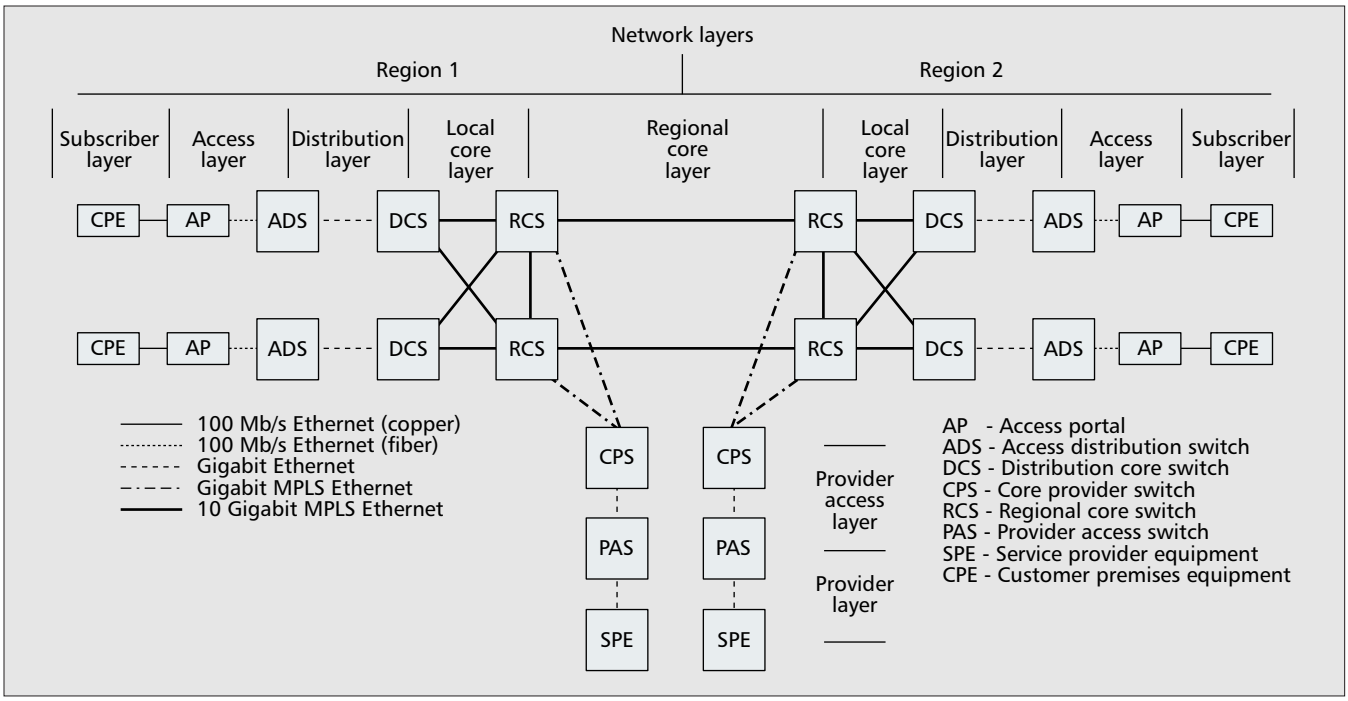

Figure 3. *UTOPIA physical architecture.*

CORE LAYERS (REGIONAL, LOCAL)

The local and regional core layers of the MetroNet take all the data aggregated from the distribution and access layers, and passes it along at the highest speeds and capacity of the network. The core layers handle all traffic from all parts of the network. As with the other layers, it is oversubscribed, but has tremendous capacity and can be upgraded by adding routers and/or ports.

Sitting in the intersection between the distribution layer and the core layers are distribution core switches. In the architecture for the UTOPIA MetroNet, the distribution core switches have fully redundant control, power, and switch fabrics. Core switches will initially support 1 Gb/s and 10 Gb/s Ethernet connections. These then connect to the regional core switches located at the edge of the regional core layer.

PROVIDER ACCESS LAYER

The function of the MetroNet is to provide a high-speed high-capacity backbone across which service providers can deliver their services. While the core layer is the freeway across which service providers will transport their services, they need a different kind of access point to the freeway than residential and business customers use: they are coming on to the freeway from another location, not from within the cities' neighborhoods. The provider access layer is like one freeway (provider access) merging into another freeway (core) for the service providers so that they can put their traffic directly onto the core.

At the intersection of the core layers and the provider access layer is the core provider switch. CPSs are deployed in pairs to provide service providers the option of fully redundant connectivity into the UTOPIA infrastructure. As large numbers of service providers sign up to deploy services across the MetroNet, additional CPS and PAS switches can be added. Before service providers access the provider access layer freeway, they have all their equipment lined up and ready to go using their own on ramps to the freeway. Quite literally, the service provider brings in SPE that will facilitate the distribution of services across the MetroNet. This equipment accesses the provider layer by hooking up to the provider access switch. The SPE may include large storage vaults of video programming, large interactive gaming servers, routers, and the like.

LOGICAL ARCHITECTURE

Logical data, voice, and video flows map across the physical topology described above. This relies on a combination of MPLS-based virtual private LAN services [4] and VLAN tagged Ethernet packets at the access and distribution layers to create logical Ethernet MPLS or VLAN-based forwarding topologies. Native untagged Ethernet packets are presently used at the subscriber demarcation to an AP element. A given VLAN is used to create individual service topologies that emulate a single redundant geographically disperse Ethernet switch for a particular service. The network will initially be configured to support three logical topologies (multipoint-to-multipoint, point-to-point, and point-to-multipoint) within broadcast domains or service VLANs to encompass all broadcast and unicast taffic.

The network has a full mesh of two physically diverse MPLS traffic engineered "tunnel" label switched paths (LSPs) between all PE elements (DCS and CPS), while the data network uses virtual circuit (VC) LSP load sharing across both tunnel LSPs between all PE elements. This eliminates the need for fast reroute

Business intranet and extranet Transparent LAN Services will also use Q-in-Q stacking to map service VLANs into existing VPLS instances. This technique permits users of the service such as municipalities and corporations to maintain their existing VLAN ids across the network.

or backup LSPs for data services. The MPLS tunnels are established via Resource Reservation Protocol (RSVP), while the VC tunnels use LDP signaling.

Topology Types

Multipoint-to-multipoint (mpt2mpt) service topology VC LSPs are configured to support seven QoS levels through MPLS EXP bit mapping. All mpt2mpt service VLANs will be assigned one of the seven QoS levels and configured through the entire MPLS cloud to the ADS element link aggregation group (LAG) based on the IEEE 802.3ad standard, and there is a one-to-one relationship between VPLS instances and QoS. This architecture limits VPLS domains to seven (one for each QoS level), reducing complexity and LDP signaling. When new services are requested, a service VLAN is created and simply added to the appropriate VPLS instance. This is accomplished by mapping multiple VLANs into the same "vc-identifier." When subscribers request a service from a service provider, the service VLAN only needs to be provisioned on the ADS subscriber facing port and AP.

Point-to-point (pt2pt) service topologies use the same seven VPLS instances as mpt2mpt topologies, also by mapping VLANs. However, pt2pt services will use Q-in-Q (based on IEEE 802.1ad) stacking on the ADS ingress port connected to the AP. The outer VLAN tag/id of a packet will be added to the appropriate port/VLAN forward equivalence class (FEC) for a specific VPLS instance, defining a service in the same way as non-Q-in-Q services. This again greatly reduces the number of VC LSPs configured through the cloud. The other benefit is that a single one of the 4096 VPLS core service VLANs is capable of supporting 4096 pt2pt circuits defined via the inner VLAN tag. Committed information rate (CIR) pt2pt VLANs using the same inner tags will be supported through a VPLS instance configured with QoS seven. Service providers may use these CIR pt2pt service VLANs to support backhaul of circuit emulation devices (T1, DS3, etc.). Business intranet and extranet transparent LAN services (TLS) will also use Q-in-Q stacking to map service VLANs into existing VPLS instances. This technique permits users of the service such as municipalities and corporations to maintain their existing VLAN IDs across the network.

Point-to-multipoint (pt2mpt2) service topologies are also possible, but these are custom configured as requested. Given these different MPLS topologies, we now look at actual delivery of voice, data, and video across the network.

Voice and Data

Revisiting the physical architecture, Fig. 4 depicts unicast data or voice traffic originating at a subscriber and traversing the network to the service provider. The next sections describe the role of the various network elements in this flow.

Access Distribution Switch — ADS elements are each configured based on specific service VLANs provided through APs to subscribers. Service VLANs are based on unique VLAN IDs, which are configured for specific topologies and QoS. As noted above, the topologies supported are pt2pt, pt2mpt, and mpt2mpt. Topologies on the ADS are constructed by configuring combinations of layer 2 bridging filters and VLAN IDs. The ADS supports four QoS levels: three (high, medium, and low) used for retail services/VLANs and one (control) used for management. Mapping into the MPLS EXP bits is as follows: Control = 7 = Reserved; High = 6 = Circuit Emulation; High = 5 = Voice; High = 4 = Video; Medium = 3 and 2 = Data; Low = 1 and 0 = Data.

All ADS configuration instances have specific ingress layer 2 filters to protect the wholesale transport network from an attack on all service VLANs, and limit ICMP traffic (pings) as well as broadcast traffic. Retail service providers may also request supported layer 3 access control list (ACL) filters to mitigate specific retail denial-of-service attacks such as L2 and L3 misses. Certain TCP/IP and/or UDP port numbers and IP addresses may also be filtered if requested by retail service providers. ACLs to block MS file and print sharing on the ingress to configuration may also be requested from service providers. ADS configurations also include global configuration filters that include blocking ingress (subscriber sourced) multicast traffic on all service VLANs, and also perform ingress port/VLAN to medium access control (MAC) address locking for learned MACs on the AP management VLAN and video gateway management VLAN.

Service providers may also request other supported specific retail configuration enhancements on service VLANs like Dynamic Host Configuration Protocol (DHCP) Option 82, which allows them the ability to trace or designate dynamically allocated IP addresses and correlate them to specific physical geographic subscriber addresses. All service VLANs also collect billing information (ingress and egress bytes) for each service VLAN, which is passed to the network operation support system/business support system (OSS/BSS) applications.

Physically, ADS elements include two physical Gigabit Ethernet connections with diverse paths to separate blades on a single redundant DCS element. These physical ports to DCS elements are configured as a LAG. LAGs aggregate multiple physical ports into a single logical port. LAGs also have all service VLANs configured on them, and include layer filters configured to protect the MetroNet.

Distribution Core Switch (DCS) — DCS elements are also configured with LAGs as described above to aggregate ADS elements. These LAGs map service VLANs into the MPLS/VPLS cloud using service-VLAN-specific port/VLAN MPLS FECs. LAGs also perform layer 2 bridging for mpt2mpt topologies across DCS Ethernet ports toward other attached ADS elements when configured with the same port VLAN FEC. DCS elements will allow the ability to restrict local bridging for pt2mpt topology service VLANs between multiple ADS elements attached to the same port VLAN FEC.

Port VLAN FECs will enforce Ethernet QoS

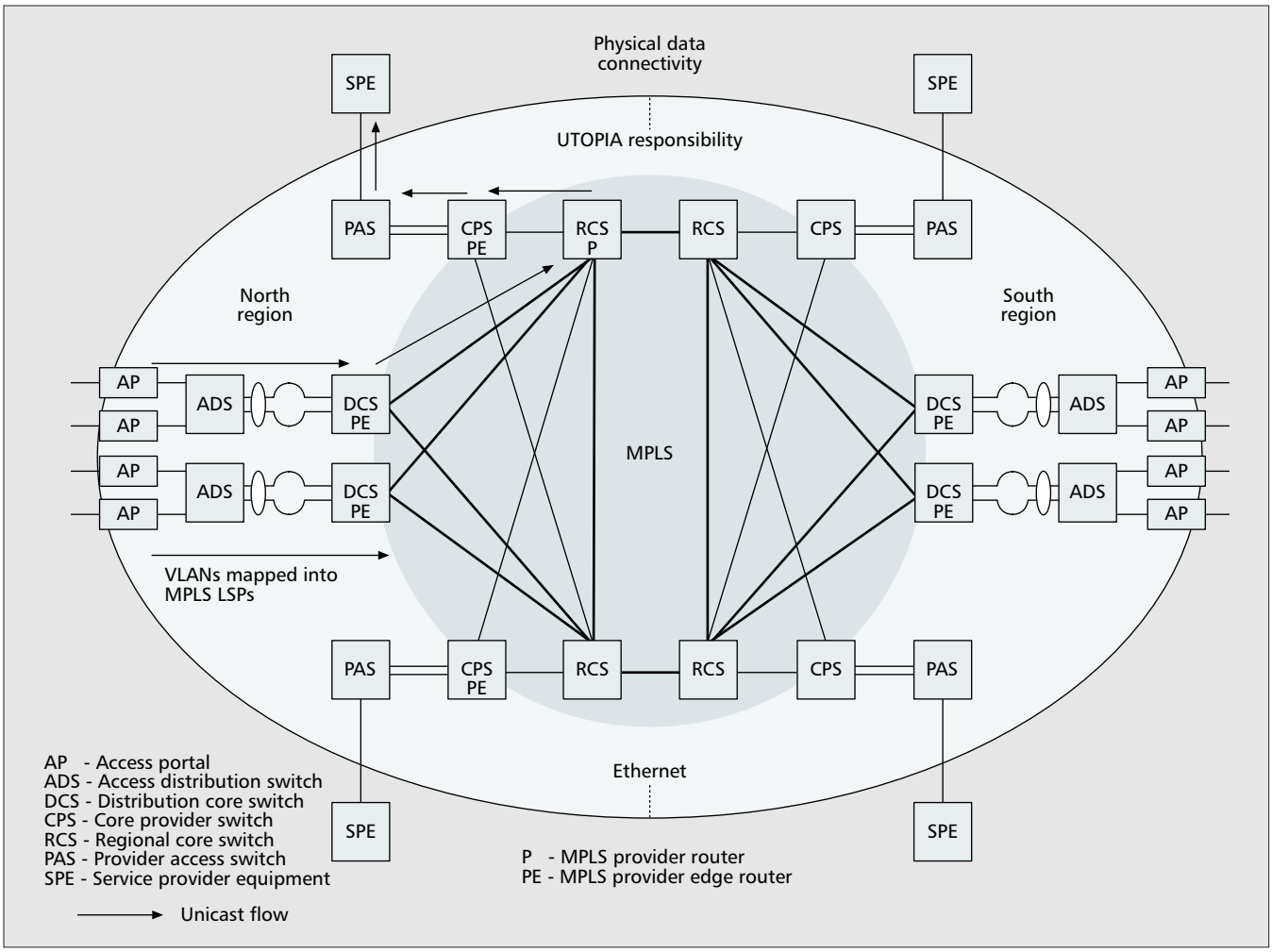

Figure 4. *Physical data and voice topology.*

based on service VLAN IDs and map into the corresponding MPLS QoS based on EXP bit correlation.

Core Provider Switch — CPS elements serve a similar function as DCS ones, but are service provider facing instead of subscriber facing and have more ports configured in their LAG connecting to the PAS.

Provider Access Switch — PAS elements serve a similar function as the ADS elements, but are service provider facing instead of subscriber facing, and has more ports configured in its LAG connecting to the CPS.

Multicast Video

The multicast physical topology (Fig. 5) differs from the unicast topology in that, at present, a multicast overlay based on Protocol Independent Multicasting (PIM) [5] routers is formed. These PIM routers connect to any service provider content (SPC) devices as well as to the RCSs. With future planned multicast enhancements, described later, these PIM routers could be eliminated.

Two redundant PIM-SM (PIM1A and PIM1B) routers are each configured to aggregate all video subnets from the head-end into a single subnet that is attached to the multicast video VLAN (McastVLAN). One router is configured as the primary multicast rendezvous point (RP) and the other as the backup (Fig. 6). The backup RP in normal operating mode prunes all multicast groups. The primary RP floods all multicast groups across both LAGs, and if the RCS element being flooded with video traffic from one of the PIM-SM routers fails, the redundant PIM-SM router begins flooding.

All multicast groups are statically joined through the PIM-SM routers to this VLAN that supports all set-top boxes (STBs) on the network. Each PIM-SM router has two LAGs configured for the McastVLAN, mapped into a VPLS instance using port/VLAN FECs. One LAG is used to connect the PIM-SM routers together at layer 2 for the McastVLAN subnet. This passes through the RCS elements without VLAN tags. It is not replicated into the MPLS cloud, and is simply passed through the RCS elements to connect the video VLAN between the PIM-SM routers so they can negotiate pruning. The other LAG uses VLAN tagged frames from the PIM-SM routers to the RCS elements that are the PEs for the McastVLAN. This LAG replicates packets based on PIM pruning.

The multicast groups (200+) are transmitted across the MPLS cloud to each DCS, and the

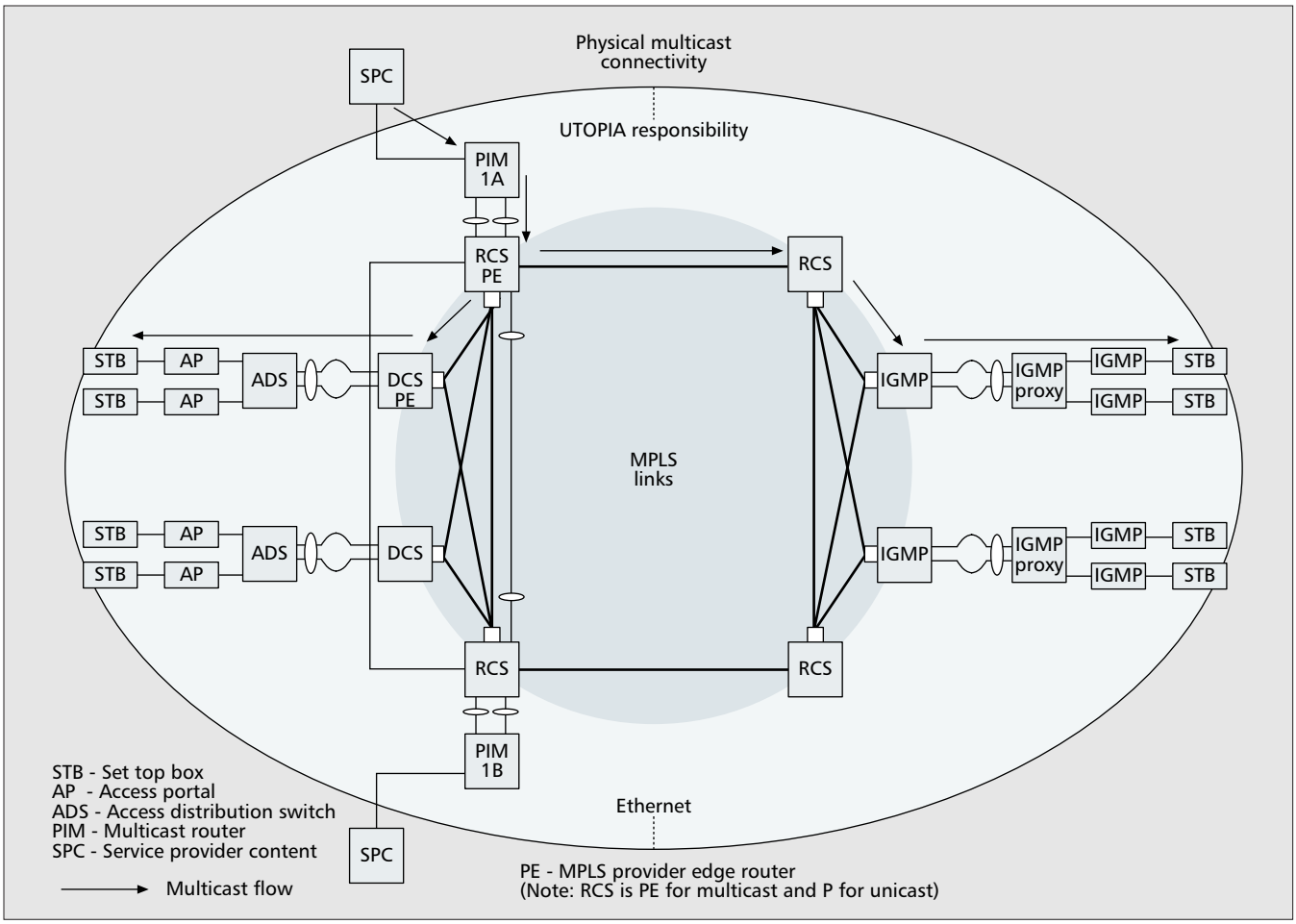

Figure 5. *Physical multicast topology.*

DCS elements are PEs for the McastVLAN/VPLS instance as well. The DCS elements perform Internet Group Management Protocol (IGMP) snooping on the Ethernet LAGs connected to ADS elements for the multicast video VLAN, while the ADS elements perform layer 2 proxy IGMP snooping [6] on the McastVLAN and the AP performs IGMP snooping for IGMP joins and leaves. The PIM-SM routers filter all IGMP traffic to reduce CPU overhead as all multicast groups are statically joined, which allows them to scale and support all IGMP hosts (set-top boxes/video gateways) network-wide. The 10 Gigabit Ethernet physical path between RCS elements is used as an MPLS hot standby or backup path protecting against an RCS to DCS primary path failure. This path is needed in the event of a primary path failure as OSPF and PIM would not be aware and would not reroute

QoS and Traffic Engineering

In order to provide transparency to higher-layer services, the network does not reorder or prioritize packets across switch elements within a service VLAN. All packets entering and exiting a service VLAN are therefore treated in a first-in first-out (FIFO) manner. Service providers may perform their own prioritization schemes before packets enter the UTOPIA MetroNet. However, the network will reorder or prioritize Ethernet packets based on VLAN IDs on Ethernet ports, and will also reorder or prioritize MPLS packets based on MPLS experimental (exp) bits through the MPLS cloud.

Link load balancing is supported across all layers with the exception of the access layer, and the distribution layer load balances Ethernet traffic flows across multiple physical (single logical) links (LAGs). MPLS links implement MPLS administration group traffic engineering and load balance VC LSPs across two tunnel LSPs between MPLS label edge routers (LERs)

Logical Redundancy Requirements

The UTOPIA MetroNet supports physical and logical link redundancy across all layers with the exception of the access layer. The distribution layer has a minimum of two links between ADS and DCS elements configured as a single logical link (LAG). If one of the physical links becomes inoperable, all traffic fails over to the operating link without any layer 3 interactions. The PAS to CPS link has links that will also be configured into a single LAG. MPLS links use traffic-engineered backup LSPs to reroute around any MPLS link failure.

Each switching element has high availability capabilities. Besides standard hardware redundancy such as dual power supplies, switch fabrics, and control modules, additional protection capabilities are available. Upon failure or crash of the main control module, the backup module,

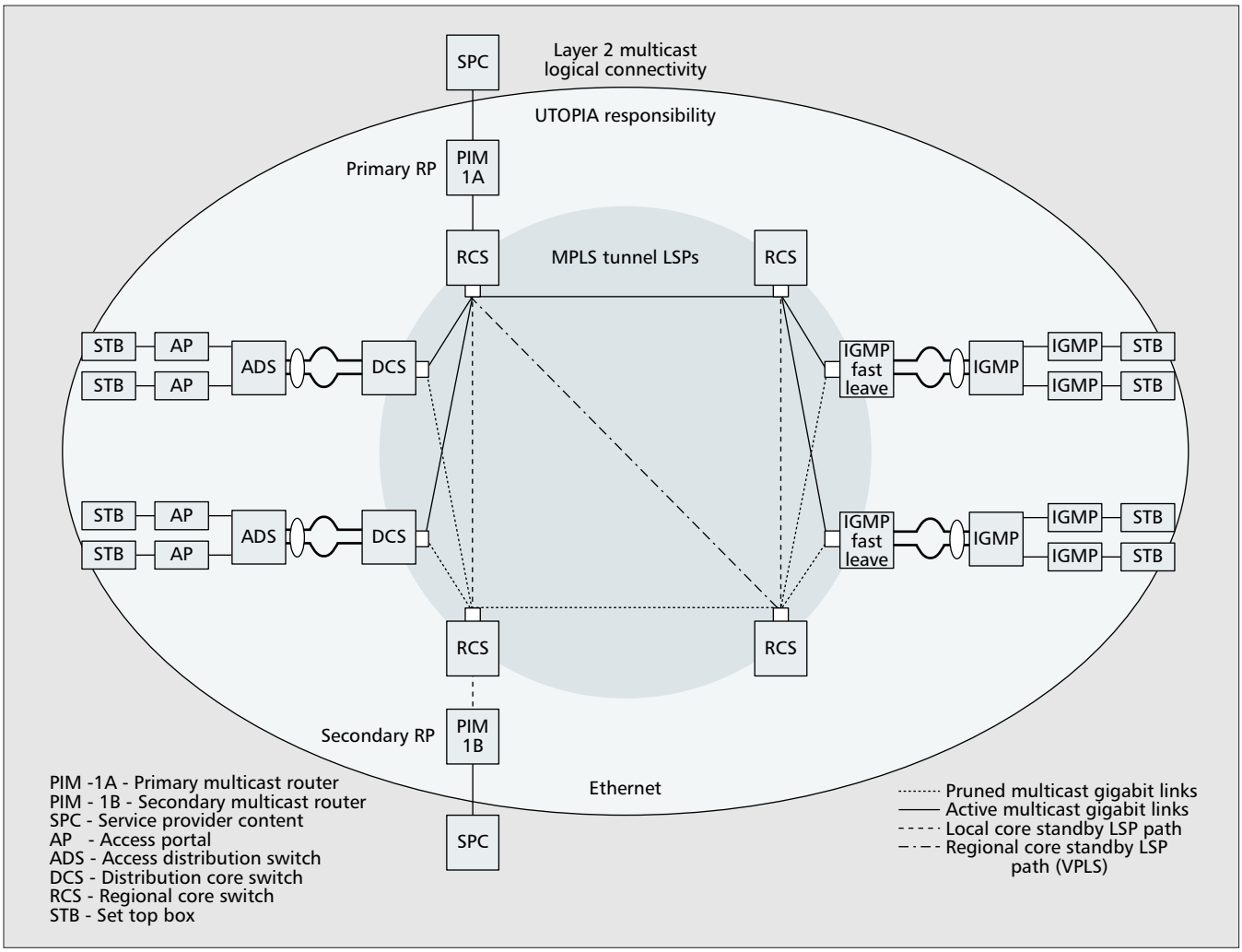

Figure 6. *Logical multicast topology.*

which keeps a synchronized state of the main processor, takes over without affecting the data traffic flows handled in hardware. This includes restart capabilities for most routing and MPLS protocols. Separation of control and forwarding functions is maintained. Control processors are dedicated to control and management functions, while application-specific integrated circuits (ASICs) and microprocessors handle the data forwarding functions. This separation enables data to be forwarded even in the case of control software failures. Graceful restart capabilities are available for both routing protocols such as OSPF or BGP but also MPLS signaling protocols like LDP and RSVP.

Future Optimizations

In the future, it is planned to move from a flat VPLS topology, where all the core switches such as RCS, DCS, and CPS nodes are fully meshed today, to a hierarchical VPLS (HVPLS) model. As our network evolves and passes more residential and business sites, this will help reduce the size of the VPLS PE mesh required with a smaller number of LDP adjacencies but also fewer LSPs to be maintained in the core of the network. Similarly, PAS, ADS, and SPE elements can act as VPLS multitenant units (MTUs) and be dual-homed into their respective core switches acting as VPLS hub PEs. HVPLS will also help with broadcast and multicast replication by spreading the replication effort across distribution switches and core switches (RCSs and DCSs).

With VPLS, replication happens at the ingress of the network (i.e., from the RCS to the DCS and then to the ADS in the current topology). With a large of amount of broadcast and multicast traffic, core bandwidth is not used in the most efficient way and might induce some undesired latency and jitter. IGMP snooping [7] helps ensure that traffic is only sent toward recipients who request specific multicast streams. Further enhancements to optimize bandwidth usage with the use of point-to-multipoint LSPs [8] will ensure that a single copy of a multicast/broadcast frame will be sent across the MPLS core. They can be used to carry multicast/broadcast traffic across various VPLS instances but can also be defined on a per-VPLS instance [9].

LAGs offer a very efficient way to protect against link failure by dynamically mapping traffic onto the active physical links. However, LAGs are point-to-point and hence need to be terminated by a single element. These LAGs are

The UTOPIA network demonstrates that Ethernet routing in conjunction with MPLS and VPLS is capable of supporting triple-play services. It provides the carrier-grade scalability, reliability, application support, standardization, and future-proofing required for mass-market deployment.

currently used between access and core switches. As HVPLS is implemented, access nodes, such as the PAS and ADS, will be able to be dual homed into separate core switches with individual LAGs.

The ability to seamlessly integrate layer 2 and 3 capabilities into the different elements provides a significant advantage. This typically involves a deep packet inspection into the IP header while making a forwarding decision at layer 2 and snooping control protocols. For instance, instead of relying on 802.1P markings, which might not always be available or trusted and are currently ignored, switching nodes can examine the type of service (ToS)/DiffServ code point (DSCP) and incoming port/VLAN combination to determine the best path toward a destination and mark the MPLS EXP bits accordingly. ACLs can also be configured based on various IP header fields in order to prevent denial-of-service attacks. Finally, snooping IGMP or PIM, acting as an IGMP proxy snooping DHCP, will allow traffic to only be sent to the proper place, avoiding unnecessary control traffic sent through the network.

Conclusions

The UTOPIA network demonstrates that Ethernet routing in conjunction with MPLS and VPLS is capable of supporting triple play services. It provides the carrier-grade scalability, reliability, application support, standardization, and future-proofing required for mass market deployment. Recent developments supporting IP multicasting, resiliency, and manageability help to make this triple play deployment a reality. For the UTOPIA network, it provides a basis on which a municipality acting as a competitive carrier may cost-effectively deploy an open services infrastructure for multiple application providers with a goal of developing local economies, increasing quality of life, and retaining skilled workers.

References

[1] "Broadband Internet Services in North America: Market Outlook & Analysis," *Strategy Analytics*, 2005, http://www.strategyanalytics.net/default.aspx?mod=ReportAbstractViewer&a0=2423

[2] E. Rosen *et al.*, "Multiprotocol Label Switching Architecture," RFC 3031.

[3] "Open Service Provider Network," *DynamicCity*, http://www.dynamiccity.com/city/ospn.html

[4] M. Lasserre, V. Kompella, "Virtual Private LAN Services over MPLS," draft-ietf-l2vpn-vpls-ldp-06.txt

[5] B. Fenner *et al.*, "Protocol Independent Multicast - Sparse Mode (PIM-SM)," draft-ietf-pim-sm-v2-new-11.txt

[6] B. Cain *et al.*, "Internet Group Management Protocol, Version 3."

[7] R. Aggarwal *et al.*, "Multicast in BGP/MPLS VPNs and VPLS," draft-raggarwa-l3vpn-mvpn-vpls-mcast-01.txt

[8] R. Aggarwal *et al.*, "Extensions to RSVP-TE for Point to Multipoint TE LSPs," draft-ietf-mpls-rsvp-te-p2mp-01.txt

[9] R. Aggarwal *et al.*, "Multicast in BGP/MPLS VPNs and VPLS," draft-raggarwa-l3vpn-mvpn-vpls-mcast-01.txt

Additional Reading

[1] "Encapsulation Methods for Transport of Ethernet Frames Over IP/MPLS Networks," draft-ietf-pwe3-ethernet-encap-10.txt.

[2] "Transport of Layer 2 Frames over MPLS," draft-ietf-pwe3-control-protocol-17.txt.

Biographies

KEN MOERMAN, chief system engineer, DynamicCity, has over 25 years experience in networking and telecommunications, engineering network architectures from bisyc to MPLS while utilizing dedicated 1200 baud analog to OC192 circuits. He has recently designed metropolitan area networks using DWDM enabling fiber-to-the-home (FTTH) networks for various public utility departments and cities across the United States and Mexico. Sample metropolitan network projects ranged from 2000 to over 150,000 nodes using 100 Mb/s, 1 Gb/s, and 10 Gb/s native Ethernet and MPLS supporting voice, video, and data. Network design customers include private and government entities, including Aerie Networks, AT&T Wireless, Touch America, the U.S. National Security Agency, and Interfibra. Previous to Dynamic City he worked 10 years at Nortel and 3Com in various engineering roles.

JEFF FISHBURN, chief technology officer, DynamicCity, has dedicated the past 15 years to the telecommunications industry, beginning with product development at AT&T Bell Laboratories, business management for Lucent Technologies, and most recently including FTTH network designs for WINfirst and UTOPIA. He has been responsible for topology and network architecture designs covering voice/video/data access architectures over DOCSIS as well as ATM/Ethernet fiber-based infrastructures, dedicated transport over SONET/CWDM/DWDM for local/metropolitan/long-haul deployments, and packet-based data networking access and metropolitan networks over ATM and IP fiber infrastructures. He was previously chief technology officer for WINfirst, an FTTH overbuilder that provided a complete voice, video, and data service package in Sacramento. Prior to and during deployment, he managed the integration of the various voice, video, and data technologies into a cohesive network which currently provides 100 Mb/s data services to individual residential subscribers. He is currently employed by DynamicCity, which brings a team that supports the financing, design, buildout, and operations of OSPNs for municipalities interested in advanced fiber networks.

MARC LASSERRE, Chief Scientist, Riverstone Networks, is responsible for standards definition and network architecture. Over the years, Mr. Lasserre has held technology leadership positions at several high-tech companies in Silicon Valley, where he has been active in the design of Ethernet switches, ATM switches, and IP routers. He has worked on system design, protocol design, and implementation for such technologies as IP stacks, IP over ATM, LANE, Frame Relay, PPP, and MPLS-based Ethernet.

DAVID GINSBURG, VP of marketing and product management, Riverstone Networks, brings more than 18 years of marketing and technical networking experience to Riverstone. Prior to his current position, he was vice president of marketing at Allegro Networks. Previously, he was vice president of product marketing at Nortel Networks, prior to which he was a founding member of the Shasta Networks marketing group until its acquisition by Nortel. He has also held a variety of management and engineering positions with Cisco Systems, Alcatel, and the U.S. Army. He has also authored several internetworking books including *Implementing IP Services and the Network Edge*, *Implementing IP Services*, and *Implementing ADSL and ATM: Solutions for Enterprise Internetworking*. He is also a former chair of the Broadband Content Delivery Forum (BCDF) and holds an electrical engineering degree from Rensselaer Polytechnic Institute.