# Packet Aware Transport for Metro Networks

T. Afferton, R. Doverspike, C. Kalmanek, and K. K. Ramakrishnan

*Abstract*— **Today's metro networks have evolved from the need to support traditional voice and private line services. However, the tremendous growth in access to Frame Relay, ATM, IP and Ethernet services, coupled with the desire of enterprise customers to interconnect via Ethernet interfaces, suggests the need for a new approach. This paper proposes a new architecture for Packet-Aware Transport Networks (PATN) which supports both packet and traditional TDM services and which leverages an assemblage of emerging technologies to provide efficient aggregation and switching of packet traffic in metro networks. The PATN has the potential to provide significant cost savings to carriers by reducing the number of network elements, reducing transport costs through statistical multiplexing, and eliminating the need for redundant multiplexing operations.**

*Index Terms*— **Ethernet services, MPLS, packet switching, time division multiplexing.**

## I. INTRODUCTION

Today's metro networks evolved to support traditional voice and private line services and, as a result, were optimized to support Time Division Multiplexing (TDM) services. However, the growth of private line services is dominated by access (also called "backhaul") to packet switches that provide Frame Relay, ATM, IP and Ethernet services. In addition, the dominant link layer used in enterprise networks is Ethernet. Since Ethernet interfaces to network equipment have historically been significantly less expensive than TDM interfaces of similar bandwidth, enterprise customers have an incentive to deploy Ethernet interfaces to their network service provider. As a consequence, major carriers are exploring methods to provide Ethernet interfaces and services in addition to traditional private line TDM interfaces and services. However, traditional TDM transport does not allow statistical multiplexing of bursty packet data.

This paper provides a coherent architectural vision for metro access networks, which we call the Packet-Aware Transport Network (PATN), and articulates its benefits. A key requirement for the PATN is support for both packet and traditional TDM services. The PATN leverages several emerging technologies, including new protocols such as Resilient Packet Ring and IETF Pseudo-Wire Encapsulation (PWE), as well as continual improvements in silicon, to allow functions that were traditionally performed in separate network elements to be integrated into a single network element by incorporating state-of-the-art vendor capabilities. As a result, the PATN has the potential to provide significant cost savings to carriers.

The PATN is designed to support packet and TDM transport within metro areas in a scalable manner, and to interface to inter-city packet and TDM transport backbone networks. Since the PATN supports PWE for all packet access traffic, it is synergistic with the move towards a converged MPLS-based multi-service packet backbone network.

Some of the features of the PATN architecture include:

- Incorporation of Ethernet switching into metro nodes to provide a more efficient architecture for access and transport of rapidly growing Ethernet services.
- Enable statistical multiplexing of packet traffic by packetization at the entry point into the metro transport network.
- Use of Ethernet interfaces, which are less expensive than their TDM counterparts
- Introduction of a Gateway Multi-Service Switch (G-MSS) between metro and core networks, which interfaces to packet-aware rings in the metro network and converts the packet payload of TDM access circuits to Layer 2 virtual circuits. The G-MSS provides a simple and efficient hand-off to a converged MPLS-based multi-service packet backbone network.
- Reduction of the number of network elements and elimination of redundant TDM multiplexing and demultiplexing.

Manuscript received November 14, 2003..

T. Afferton, AT&T Labs, 200 Laurel Ave, Middletown, NJ, afferton@att.com
R. Doverspike, AT&T Labs Research, 200 Laurel Ave, Middletown, NJ, rdd@research.att.com
C. Kalmanek, AT&T Labs Research, 180 Park Ave, Florham Park, NJ, crk@research.att.com,
K. K. Ramakrishnan, AT&T Labs Research, 180 Park Ave, Florham Park, NJ, kkrama@research.att.com,

For Ethernet services and interfaces to substantially penetrate into the metro and intercity environment, there has to be an economic incentive for metro carriers to deploy the appropriate infrastructure given the existing service mix. Our proposed architecture accelerates this deployment because, in addition to providing an efficient transport of Ethernet services, it provides significant cost savings and efficiencies for transporting existing packet and legacy TDM services. Customers can access Ethernet 'virtual circuits' and transparent LAN services at arbitrary port speeds.

The paper is organized as follows. Section II describes the architecture of current metro transport networks and motivates the need for a new approach. Section III describes the Packet-aware Transport Network (PATN) architecture and its key principles. Section IV summarizes.

## II. CURRENT METRO TRANSPORT NETWORKS

Today's terrestrial telecommunications transport network in the United States can be loosely described as comprising three segments: access, metro, and core. The ***access*** network connects customers to transport equipment that is owned and operated by a telecommunications carrier. The ***metro*** network generally includes transport equipment and facilities in metropolitan areas, but can vary greatly in size, extent and complexity. The ***core*** network consists of long-haul switching and transport equipment providing transport services between major cities. Packet switches and routers are typically deployed in large metro areas. However, since packet switches and routers are operationally more complex than transport network equipment, packet switching equipment tends to be consolidated in a smaller number of locations than transport equipment. Packet traffic is "backhauled" through the metro transport network to a packet switch or router where it is either switched locally to another customer location within the same metro area, or switched to another core location.

Figure 1 depicts an example of today's metro network. Transmission and multiplexing equipment resides in central offices (buildings) and is transported between offices by fiber optics. Most carriers use SONET/SDH today. Furthermore, in metro networks, transport equipment is configured into SONET self-healing rings to provide reliability. Broadband Digital Cross-Connect Systems (B-DCS) are interconnected by the core optical transport network and interface between the metro and core networks. The network elements that serve as the "gateway" between metro and core networks are shown in colored background in Figure 1.

When one considers the end-to-end service provided to a customer, this picture becomes more complex, since portions of a customer's transport circuit or packet service may be provided by a combination of carriers, Internet Service Providers, etc. For simplicity, Figure 1 depicts one carrier that provides both metro and core transport services and packet services, as well as another metro transport provider in the same metro area. This simplified diagram illustrates how integration of packet-aware transport might occur within one carrier's metro and core networks and between two different carriers.

Most metro transport carrier's interface to customer equipment (routers, servers, voice switches, etc.) via the legacy TDM signal rates (DS0, DS1, DS3, SONET OC-3, OC-12 in North America). For the purposes of this paper, we classify the services offered by metro carriers into five broad classes. ***Private line*** is TDM transport of a signal at the above rates between two customer (or other carrier) interfaces. ***Packet network access*** is transport of packets between a customer interface and packet network service that is provided by the same or another carrier. ***Ethernet point-to-point connections*** are a virtual circuit service, usually between two Ethernet interfaces at different customer locations***.*** ***Ethernet transparent LAN service*** is a multipoint-to-multipoint service supporting virtual LAN bridging based on IEEE 802 specifications among customer sites. ***Voice access*** can be TDM DS0-based access to voice switches or Voice-over-IP (VoIP). Today's metro networks have evolved to transport private line and voice TDM services, but are limited in their ability to efficiently serve packet network access and Ethernet-based services.

To illustrate the limitations of today's network for packet services, we show two example packet network access circuits in Figure 1. The circuit in the middle of the figure illustrates a DS1 interface from a customer router connected to an access router in an ISP network. In the example shown, the metro transport and ISP services are provided by the same carrier. The access router de-encapsulates packets from the DS1 signal (the point of packet extraction) and performs Layer 3 routing and switching. The second circuit (at the bottom) illustrates an access circuit to a Frame Relay network. This circuit also uses a DS1 interface from a customer router, but the metro carrier provides an nxDS0 service, which terminates at the Frame Relay switch. The Frame Relay switch de-encapsulates the packets from the nxDS0 channel and maps them to PVCs. This access circuit is provided by a different metro carrier from than the Frame Relay service provider.

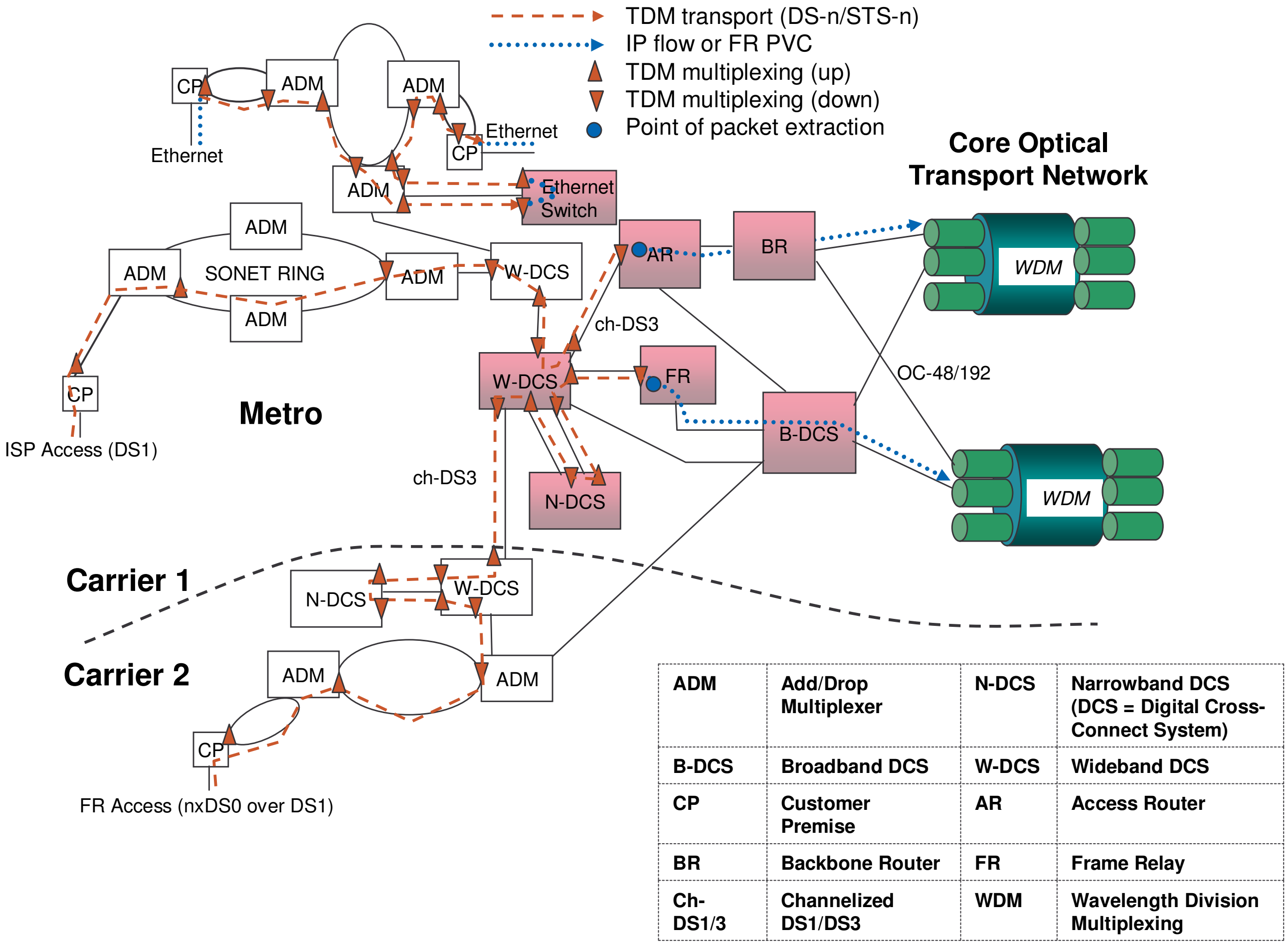
There are some obvious limitations in the present network structure:

- Demand for Ethernet interfaces is growing. To provide Ethernet service in the present network, packets are encapsulated into TDM circuits and homed to a dedicated Ethernet switch at the gateway office ("hub and spoke" architecture). However, since much of the demand for Ethernet services is intra-metro, this is inefficient. This is illustrated in the top circuit flow of

Figure 1, where a point-to-point Ethernet Virtual Circuit is provided between two customer Ethernet interfaces at different customer equipment on the same ring. The Ethernet signal is packed from each ADM into a TDM “pipe”, such as an STS-1, and routed to a separate Ethernet switch. The Ethernet switch must then de-encapsulate the packets from the two interfaces and switch the Ethernet frames between the two interfaces.

- The metro network has many points where circuits are subject to TDM multiplexing (upward triangle in Figure 1) and de-multiplexing (downward triangle). Transport networks evolved to this structure because the TDM multiplexing hierarchy creates circuit bundles at different rates and it is necessary to multiplex them to aggregate circuits going to the same destination in order to achieve high link utilization. Three types of Digital Cross-connect Systems (DCSs) have been developed to handle this task, i.e., Narrowband, Wideband, and Broadband DCSs cross connect signals at the DS0, DS1 (or SONET VT-1.5 rate), and DS3 (or SONET STS-1) rate respectively. This is shown in Figure 1, in which the Frame Relay access circuit traverses an N-DCS and W-DCS to create DS1s (which are then packed into DS3s) which are bound for the Frame Relay switch and have reasonably high DS0 fill.
- Because of the TDM encapsulation, the network cannot take advantage of statistical multiplexing across multiple packet access circuits.
- The customer has only a coarse granularity of bandwidths from which to choose, based on the TDM multiplexing hierarchy. Thus, customers must purchase a TDM access circuit large enough to accommodate their peak demand, even if this means running it at low utilization. Likewise, encapsulation of Ethernet into standard SONET payloads like STS-48c, STS-12c and STS-3c can be inefficient.
- Packet switches and routers must support channelized TDM interfaces that can demultiplex down to low rates. These functions consume precious space on customer-facing interface cards, which ultimately increases cost.

The goal for the Packet-aware Transport Network is to provide a more efficient architecture for packet services, while continuing to support existing TDM interfaces and service capabilities.



| | | | |
|---|---|---|---|
| **ADM** | **Add/Drop Multiplexer** | **N-DCS** | **Narrowband DCS (DCS = Digital Cross-Connect System)** |
| **B-DCS** | **Broadband DCS** | **W-DCS** | **Wideband DCS** |
| **CP** | **Customer Premise** | **AR** | **Access Router** |
| **BR** | **Backbone Router** | **FR** | **Frame Relay** |
| **Ch-DS1/3** | **Channelized DS1/DS3** | **WDM** | **Wavelength Division Multiplexing** |

**Figure 1: Current Transport Network**

## III. PACKET-AWARE TRANSPORT NETWORK

The PATN architecture is motivated by the following network drivers:

- enable a scalable transition to emerging packet-based services and customer Ethernet interfaces,
- reduce the number of gateway network elements,
- simplify the transport architecture,
- leverage more efficient and cost-effective interfaces to packet switches, and
- continue to support TDM customer interfaces and existing TDM services.

Figure 2 shows two carrier networks. Carrier 1 has both a packet-aware-transport metro network and a core network, while Carrier 2 has a traditional TDM metro network. Carrier 1's PATN provides aggregation and transport of TDM and packet traffic from multiple central offices in the metro area into a large gateway central office containing a *Gateway Multi-Service Switch* (G-MSS). The G-MSS hands off packet traffic to a Multi-Service Edge platform (MSE) which connects to the core packet network which is assumed to be a converged, multi-service network based on MPLS [7]. For example, the MSE may support Frame Relay, ATM, IP and Ethernet services. The G-MSS hands off TDM traffic to a B-DCS which connects to the core optical transport network.

Each office in the PATN contains one or more *Packet-aware Multi-Service Switches* (P-MSS's) which interface to customers and/or other carriers. TDM and packet traffic from P-MSS to P-MSS or from P-MSS to G-MSS are carried over metro access rings that terminate on the G-MSS. The PATN supports both electrical interfaces (e.g., DS3, Fast Ethernet) and optical interfaces (e.g., OC-n, GigE). In particular, customers may interface to a Customer Premise (CP) device that connects to the P-MSS via a packet-aware interface or a dedicated TDM circuit. Access rings to customers and between P-MSSs are typically SONET rings.

Packet traffic can be carried over the metro access rings using many different approaches. For example, the P-MSS's may support the IEEE 802.17 Resilient Packet Ring (RPR) protocol [3], either on top of a dedicated OC-n physical layer or in a "cutout" consisting of a Virtual Concatenation Group (VCG) of N STS-1s as part of a larger SONET pipe [2].

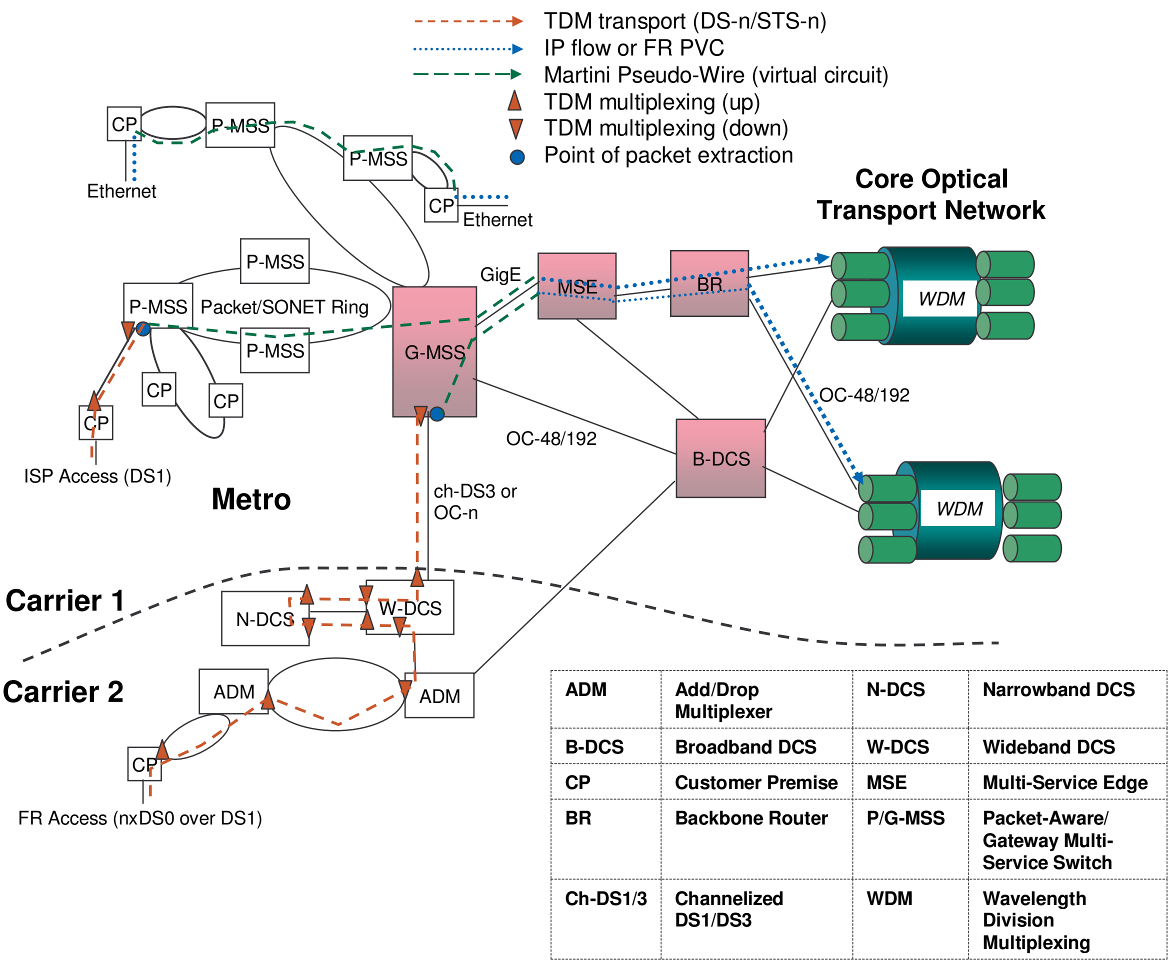
The PATN eliminates the need to multiplex and de-multiplex TDM channels multiple times and reduces the total number of network elements. P-MSS and G-MSS nodes de-multiplex TDM channels down to the lowest granularity TDM channel that is needed to extract and switch ("groom") the embedded TDM channels. Since this may require de-multiplexing a high rate TDM interface all the way down to DS0, we refer to this as ***deep channelization***. The PATN aggregates the functions of the W-DCS, N-DCS and the ring interface of the ADM in Figure 1 into the G-MSS. This aggregation of multiple cross-connect functions into one high-capacity platform has become feasible as a result of the recent improvements in silicon.

The PATN takes advantage of a converged MPLS-based multi-service backbone network by handing all of the metro/access packet traffic to a single MSE platform. Traffic from multiple services is aggregated onto a single large-capacity link between the PATN and the MSE, thus achieving high utilization on that link.

The PATN also takes advantage of the IETF Pseudo-Wire Encapsulation (PWE) standards [1,4,5] to support virtual circuits (VCs) from P-MSS to MSE and from P-MSS to P-MSS. The PATN multiplexes traffic from multiple PWE virtual circuits over the P-MSS ring to get the benefit of statistical multiplexing for packet traffic destined to the MSE. This is shown in the middle flow from the CP to the MSE in Carrier 1's metro network in Figure 2. For Ethernet services, the inefficient "hub and spoke" architecture and the dedicated Ethernet switch of today's network (Figure 1) are replaced with simpler, more efficient virtual circuits. For example, intra-ring Ethernet services can be transported over one ring as virtual circuits (top flow of Figure 2). Inter-ring Ethernet services (same metro) are switched via the G-MSS.

The proposed PATN architecture ensures that legacy interfaces are retained. As shown in Figure 2, the G-MSS (of metro carrier 1) interfaces to another carrier (metro carrier 2), who may still have the traditional TDM infrastructure. Packet extraction and deep channelization functions are performed at the G-MSS.

The key architectural principles and associated network capabilities required to realize the PATN are described in the following subsections.

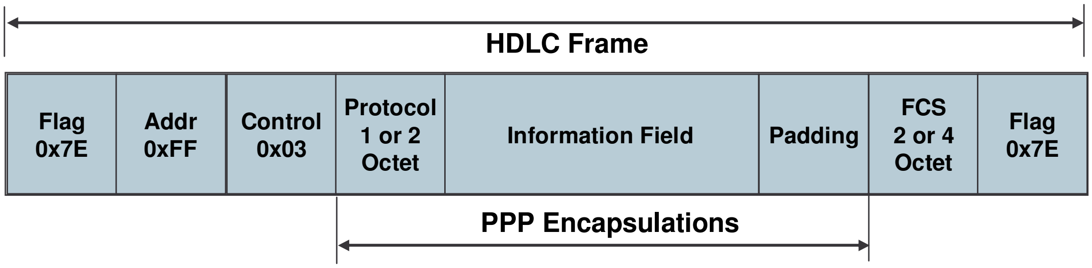


| ADM | Add/Drop Multiplexer | N-DCS | Narrowband DCS |
|---|---|---|---|
| B-DCS | Broadband DCS | W-DCS | Wideband DCS |
| CP | Customer Premise | MSE | Multi-Service Edge |
| BR | Backbone Router | P/G-MSS | Packet-Aware/ Gateway Multi-Service Switch |
| Ch-DS1/3 | Channelized DS1/DS3 | WDM | Wavelength Division Multiplexing |

**Figure 2: Packet-aware Transport Network**

### A. *Deep Channelization and Idle Packet Suppression*

Deep channelization allows packet traffic embedded in TDM channels to be extracted, encapsulated and switched through the PATN. For example, existing customer TDM traffic often enters an add/drop port of an ADM at the DS3 rate. The DS3 can be ***clear-channel*** or ***channelized***, the latter consisting of multiplexed DS1s (each of which may be clear channel or channelized with DS0s). When the P-MSS interfaces to a channelized DS3, it de-multiplexes the DS1 and DS0 channels to either extract packets (if provisioned as packet-network access service) or cross-connects them as TDM signals (for TDM services). For packet-network access circuits (e.g., access to the MSE), the P-MSS terminates the link layer protocol carried on the TDM channel, strips idle packets, and encapsulates each packet onto a PWE virtual circuit for switching. Figure 3 shows the typical Packet-over-SONET (POS) encapsulation used today. Packets are encapsulated in an HDLC frame which is delineated by a flag byte. Empty frames (consisting of only flag bytes) are inserted to match the bit rate of the TDM signal. The Point-to-Point Protocol (PPP) [8] is used for link management between the customer interface and edge switch. When doing Idle Packet Suppression (IPS), the P-MSS transmits the PPP, but discards the empty frames and identifies the packets of the original TDM circuit with a virtual circuit, as defined by the PWE. Thus, by combining deep channelization and idle packet suppression, the PATN architecture represents a new application of PWE to metro access networks. IPS provides significant efficiency improvement, while maintaining complete service transparency for the customer's existing packet-over-TDM access links. Additionally, IPS eliminates some issues that can arise with the utilization and integrity of the bit-stream when using POS encapsulation (more generally, Packet-over-TDM).



***Protocol Field*: one or two octet; identifies PPP encapsulated datagram, values specifed in RFC 1700 "Assigned numbers"**
***Information Field*: datagram for the protocol specified in the protocol field. Max length specified by Maximum Recieve Unit (MRU). Defaults to 1500 octets.**
***Padding*: arbitrary padding up to MRU**

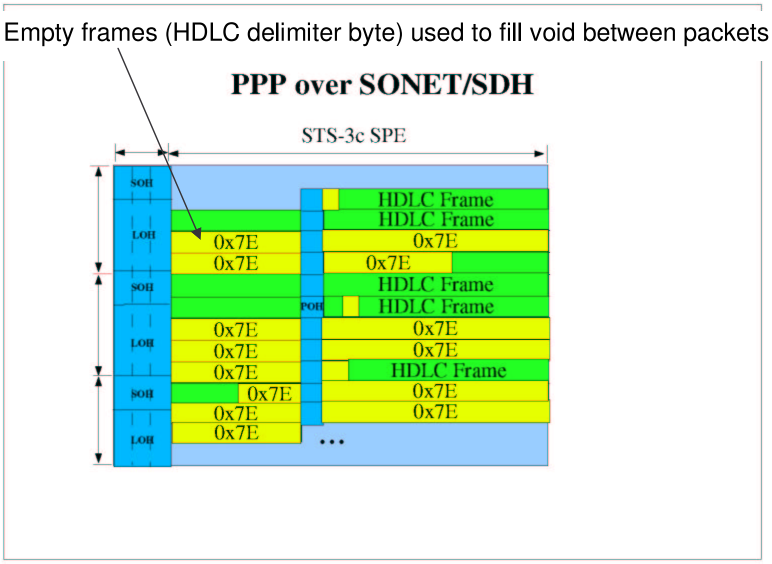


**Figure 3: Packet-over-SONET Encapsulation**

### *B. Pseudo-Wire Encapsulation and Virtual Circuit Transport*

This section describes the protocol layers in the PATN and across the interface to the MSE. Service transparency, scalability to a large number of customer ports, a rich access network topology, and reliability are important characteristics met by our proposed protocol framework.

Figure 4 shows the protocol layers, focusing on support for packet traffic. The physical layer on the metro rings uses SONET framing, although Gigabit Ethernet (GigE) or 10 GigE are also possible. SONET Virtual Concatenation (VCAT) is used to carve out a portion of the channel capacity for packet traffic, and Generic Framing Procedure (GFP) [2] provides packet framing. The IEEE 802.17 Resilient Packet Ring (RPR) provides media access control for the packet aware nodes on the metro/access rings. Layer 2 and 3 packets are encapsulated using the IETF PWE for transport and switching through nodes in the PATN up to the MSE. PWE virtual circuits are "port mapped" i.e., all traffic associated with a customer port is encapsulated on the same virtual circuit. These encapsulations introduce additional overhead. Therefore, the access rings must be engineered with enough capacity to accommodate these overheads, especially if the traffic is dominated by small packets. The MSE performs a look-up based on the VC label to perform the appropriate service function on the packet. PWE is defined for frame relay, ATM and Ethernet services; transport of IP packets within the PATN may require a new PWE code point to be defined. While the interface between the G-MSS and MSE can be OC-n (POS), the figure shows a GigE interface between the G-MSS and MSE. Link aggregation, as defined in IEEE 802.3AD provides protection capabilities equivalent to traditional SONET 1:1 protection.

Each virtual circuit has an associated Class of Service (CoS) defining the treatment appropriate for the traffic carried on the VC. Since CoS treatment requires policing and shaping in P-MSS and MSE nodes, and there may be implementation limits on the number of policers and shapers, we allow aggregation of VCs requiring the same CoS treatment onto a single MPLS label switched path (LSP) between the G-MSS and MSE. Thus, the MSE is only required to have a policer / shaper per LSP, thus reducing the number of CoS contexts it has to maintain.

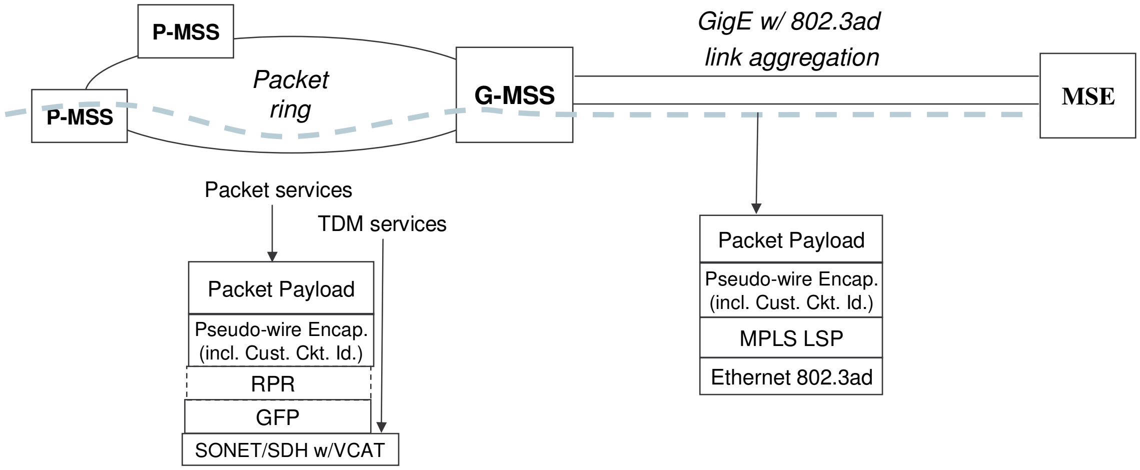


**Figure 4: Protocol Layering in the PATN**

### *C. MSE Interface and Statistical Multiplexing*

By converting TDM interfaces carrying packet traffic to their native packet form at the point of entry into the access network, carriers can take advantage of statistical multiplexing in the PATN, thus reducing the required transport capacity. This also reduces MSE interface capacity requirements. The MSE interface is simplified by eliminating hardware-based TDM deep-channelization functions. These efficiencies result in significant cost savings on the MSE customer-facing line cards, which is the dominant cost for the MSE.

To quantify the gains obtained from statistical multiplexing with the PATN in comparison to carrying packet services using TDM access, we analyzed SNMP data collected from a large number of T1 links carrying packet traffic from business customers coming into the access routers of a large ISP. The SNMP data is the aggregate number of bytes carried over the individual T1 links, measured over 5 minute intervals. The instantaneous multiplexing gain is obtained as the ratio between the sum of the peak utilizations across all the T1 links in the aggregate and the peak of the sum, as follows:

$$\text{Statistical Multiplexing Gain} = [\, \Sigma_i \max_k T_{ik} \,] \,/\, [\max_k \Sigma_i T_{ik} \,] \,,$$

where $T_{ik}$ = the five minute average of link $i$ in interval $k$. Figure 5 shows the gains measured over an entire day, across several days of a week for different levels of aggregation, ranging from 28 T1 links (that would comprise a fully populated T3) to 336 T1 links (in an OC12 link). The gain ranges from approximately 5.3 (day 4, a Sunday), to nearly 3.7 on a weekday. We also observe that the gain is obtained even with relatively small aggregates (e.g., with 28 T1s). We note that this is a theoretical lower bound on the multiplexing gains, because the numerator assumes a perfect sizing of the TDM pipe. In fact, customers can only purchase their pipes in the TDM granularities defined earlier. This makes the numerator quite a bit lower than the provisioned capacity of the TDM pipe. Thus the ***actual*** multiplexing gain tends to be much larger.

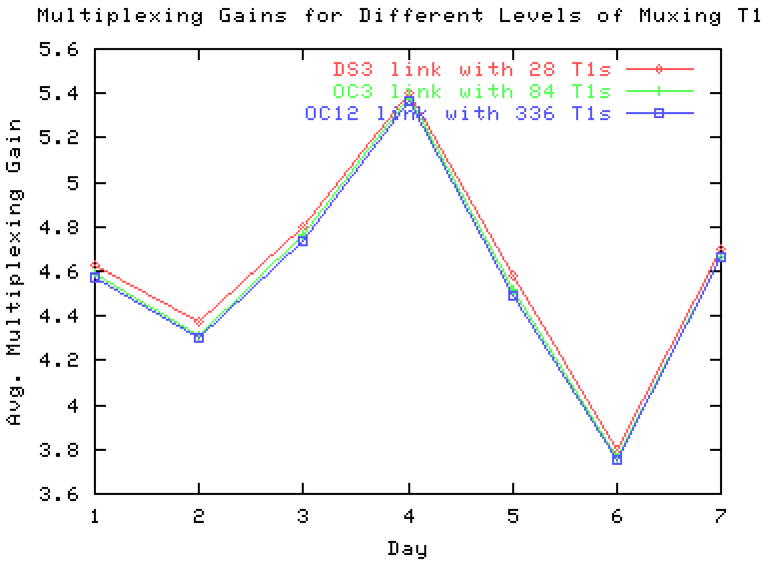


**Figure 5: Lower Bound on T1 Multiplexing Gains**

### *D. Transport of Ethernet services*

Given the growth in Ethernet services, including Ethernet virtual circuit and transparent LAN services, the PATN has been designed to support these services more efficiently than today's access methods. Ethernet virtual circuit service involves a range of service features, including port mapped service as discussed earlier, but also allowing VLAN-tagged traffic from customers to be mapped onto different virtual circuits for switching to different locations. Customers typically specify the traffic profile for a port in terms of a Committed Information Rate (CIR) and Peak Information Rate (PIR) or Excess Information Rate (EIR = PIR - CIR). Customers may also specify the CoS treatment associated with all of the traffic associated with a port. Ethernet virtual circuit services are directly supported by P-MSS nodes, using PWE encapsulation for transport of virtual circuits between P-MSS nodes and across the core network.

Transparent LAN service is also readily supported by the PATN architecture, by integrating transparent bridging functionality into P-MSS and G-MSS nodes. IEEE 802.17 defines *flooding* procedures which support broadcast on the ring, while avoiding packet duplication even in the presence of ring topology changes. Customer VLAN tag information is carried transparently through the provider-based bridged LAN. The PATN encapsulates the customer traffic using a provider-based VLAN tag which is provisioned at the P-MSS and G-MSS nodes. Thus, traffic from all ports associated with a specific customer are bridged over the provider-based bridged LAN within the metro area. Transparent LANs may be extended across the core network utilize the emerging Virtual Private LAN Service standards [9].

### *E. Resilient Packet Ring (RPR)*

The Resilient Packet Ring (RPR) protocol being developed in the IEEE 802.17 group [3] is designed to carry both packet traffic and constant bit rate traffic on metro access rings. RPR is a Medium Access Control (MAC) protocol supporting dual counter-rotating rings that can potentially replace traditional SONET rings. RPR has been designed in particular for supporting metro Ethernet transport. Nodes on an RPR ring transport frames from a source to a destination node by encapsulating the higher layer frame with an RPR header. RPR supports spatial re-use, which increases overall network capacity by enabling multiple sources to send traffic to destinations simultaneously as long as their traffic does not use the same span (a link between neighboring nodes).

To support a range of performance requirements, RPR defines three QoS classes: Class A, B and C, with strict priority between them. Class A supports traffic requiring bandwidth and jitter guarantees. Class B supports traffic requiring bandwidth guarantees, specified as a CIR and EIR. Class C supports best effort traffic. These classes of service support services similar to those supported by the IETF Diffserv classes (e.g., EF, AF and BE classes) and for support of service classes defined in IEEE 802.1p. As with other MAC protocols, the RPR MAC is constructed to achieve high ring utilization while also ensuring fair access to the channel for contending sources. Class A traffic and the CIR portion of Class B traffic are subject to admission control. Class C traffic and the EIR portion of Class B traffic are considered to be *fairness eligible* (FE). RPR defines a fairness algorithm that allocates the available bandwidth among stations sending fairness eligible traffic under congestion. The RPR standard process in IEEE 802.17 is still work-in-progress, with the first release targeted to be completed by the end of 2003.

### F. *Transport of TDM services*

The PATN architecture allows a range of implementation alternatives for P-MSS and G-MSS vendors and multiple options for transporting TDM traffic across metro access rings. Since P-MSS nodes implement deep channelization, the de-multiplexing and switching of TDM channels is distributed throughout the PATN. One option for the transport of TDM traffic uses circuit emulation to carry TDM traffic through a packet-only access network. In this case, (following one direction of transmission) the P-MSS extracts embedded TDM channels, encapsulates the TDM payload in a format supporting circuit emulation [6], and switches encapsulated packets through the PATN with the appropriate quality of service. The G-MSS converts the packets back to TDM and multiplexes the TDM channels onto a channelized TDM interface to the B-DCS (Figure 2). For this implementation, RPR can be used for packet transport: it is highly reliable, capable of transporting circuit-emulated traffic to carrier-grade specifications, and a standard, which allows inter-working of different vendors with common performance specifications.

A second option subdivides the capacity on the metro access rings using a "cutout" for the packet traffic. Here, Virtual Concatenation [2] is used to allow packet traffic to be carried in a subset of the STS-1s, with TDM traffic carried in the remaining STS-1s. Additional implementation options relate to the switch fabric in the P-MSS and G-MSS nodes. One example is where the P-MSS and G-MSS contain both packet and TDM fabrics, with the TDM fabric supporting TDM grooming down to at least VT1.5 or DS0 granularity.

## IV. BENEFITS AND REALIZATION OF PACKET-AWARE TRANSPORT

This paper presents a new architecture for access networks, introducing simple yet powerful packet switching capabilities into metro transport networks. In particular, the PATN architecture creates an efficient and scalable transport infrastructure for emerging wide-area Ethernet point-to-point virtual circuits and transparent LAN services. The low cost and ready availability of Ethernet interfaces suggests that Ethernet services may represent a new growth opportunity for metro transport carriers. In addition, the PATN provides significant cost savings for packet services with traditional TDM interfaces by eliminating the need for channelized MSE interfaces and by supporting statistical multiplexing of packet traffic. The PATN does not migrate all packet switching functions to the metro transport network, but rather extends the traditional definition of transport services to include support for packet aggregation and transport accommodating different traffic classes. This approach is synergistic with the migration of many packet service providers towards a converged MPLS packet network.

Furthermore, the PATN retains support for legacy TDM interfaces and features, but simplifies the implementation of the metro transport network by reducing the number of network elements and eliminating the number of multiplexing and de-multiplexing operations that need to be performed. This simplification is possible because advances in silicon technology now allow functions such as deep channelization to be performed on a line card of a metro network element, such as the G-MSS, rather than requiring separate cross-connect platforms for each rate.

Our investigation of the state-of-the-art in multi-service switching suggests that the functionality and scale of network elements needed for the PATN is now technically feasible, with multiple implementation alternatives possible. To better validate our vision, we have assembled end-to-end network testbeds of leading-edge metro technologies that represent the components of the PATN architecture. The key principles of the PATN, such as virtual circuit creation, idle packet suppression, deep channelization, circuit-emulation, Ethernet transport, QoS and fairness have been evaluated for feasibility. The key to implementation is a network transition strategy that evolves to the PATN, yet leverages the existing network in a cost-efficient way. This is for future work.

## ACKNOWLEDGMENTS

We gratefully acknowledge the contributions of several people to the PATN architecture: Krishna Pattabhiraman, Luis Aguirre-Torres, Shane Eleniak, John Wei, Jorge Pastor, and Dongmei Wang.

## REFERENCES

[1] Stewart Bryant and Prayson Pate, "PWE3 Architecture", October 2003, work in progress.< http://www.ietf.org/internet-drafts/draft-ietf-pwe3-arch-06.txt>

[2] D. Cavendish, K. Murakami, S. Yun, O. Matsuda and M. Nishihara, "New Transport Services for Next-Generation SONET/SDH Systems," IEEE Communications Magazine, May 2002.

[3] IEEE. IEEE Standard 802.17: Resilient Packet Ring. http:/ieee802.org/17 (standard specification in progress).

[4] Luca Martini, Matthew Bocci, Jeremy Brayley, Ghassem Koleyni, and Eric C. Rosen," Encapsulation Methods for Transport of ATM Cells/Frame Over IP and MPLS Networks", October 2003, work in progress. http://www.ietf.org/internet-drafts/draft-ietf-pwe3-atm-encap-03.txt

[5] Luca Martini, Nasser El-Aawar and Eric C. Rosen, "Encapsulation Methods for Transport of Ethernet Frames Over IP/MPLS Networks", October 2003, work in progress. < http://www.ietf.org/internet-drafts/draft-ietf-pwe3-ethernet-encap-04.txt>

[6] M. Riegel, et al., "Requirements for Edge-to-Edge Emulation of TDM Circuits over Packet Switching Networks", June 2003, work in progress, < http://www.ietf.org/internet-drafts/draft-ietf-pwe3-tdm-requirements-01.txt>

[7] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture", Request for Comments 3031, January 2001, <ftp://ftp.rfc-editor.org/in-notes/rfc3031.txt>

[8] W. Simpson, Editor, “The Point-to-Point Protocol (PPP)”, Request for Comments 1661, July 1994, < ftp://ftp.rfc-editor.org/in-notes/std/std51.txt>

[9] M. Lasserre, V. Kompella, et al., “Virtual Private LAN Services over MPLS,” November 2003, work in progress, <http://www.ietf.org/internet-drafts/draft-ietf-ppvpn-vpls-ldp-01.txt>

**Robert D. Doverspike** (IEEE SM 1997) received his undergraduate degree from the University of Colorado and Masters and Ph.D. degrees in Mathematics from Rensselaer Polytechnic University. He started with Bell Labs in 1979 and, upon divestiture of the Bell System, went to Bellcore (now Telcordia). In 1997 he went to AT&T Labs (Research) where he currently manages the Transport Network Evolution Research Department. He has extensive experience with optimization of metro networks and, in particular, optimization of multi-layered transmission and switching networks. His current work includes advanced transport and IP network architectures, network restoration methods for optical cross-connects (for which he has numerous patents), and methods for IP over optical-layer integration. He has published broadly journals in telecommunications, optical networks, mathematical programming, IEEE Communications Society, operations research, applied probability, and network management. He is member of the Mathematical Programming Society, Optical Society of America (OSA), INFORMS, IEEE, and serves as associate editor of Heuristics Journal and Operations Research Journal. He has held key leadership positions of multiple international telecommunications societies and conferences, most prominently the INFORMS Technical Section on Telecommunications.

**Charles R. Kalmanek** (IEEE M 1997) received his undergraduate degree from Cornell University, and M.S. degrees in Electrical Engineering and Computer Science from Columbia University and New York University respectively. He joined AT&T Bell Labs in 1980 and has managed advanced development and research groups since 1989. He currently manages the Networking Research Division in AT&T Labs Research. He has extensive experience in network architecture, protocols and distributed systems, and has worked on ATM switch and host interface design, congestion control, routing, Voice over IP and multimedia streaming, access network architecture, wireless networks, and management of large-scale IP networks. His current work includes advanced transport and IP network architectures, IP network measurement and analysis, and network-based applications such as content distribution networks and storage networking. He is co-chair of the IEEE Internet Technical Committee.

**K. K. Ramakrishnan** (IEEE M 1983) is a Technology Leader at AT&T Labs.- Research in Florham Park, New Jersey. He joined AT&T Bell Labs in 1994 and has been with AT&T Labs. Research since its inception in 1996. Between 2000 and 2002, he was at TeraOptic Networks, Inc., as Founder and Vice President. Prior to 1994, he was a Technical Director and Consulting Engineer in Networking at Digital Equipment Corporation. At AT&T Labs. Research, he is involved in several technical and strategic activities in networking, including Quality of Service, Congestion Control, Signaling, Virtual Private Networks, IP Telephony and Metro Area Networks. He has published nearly 100 papers and has 58 patents issued. His contributions on congestion control, channel access protocols (for Ethernet and Packet Cable), network interfaces, operating system support for network I/O, signaling and IP Telephony have been adopted and implemented in the industry. K.K. has an M.S. degree from the Indian Institute of Science (1978), and an M.S. (1981) and Ph.D. (1983) in Computer Science from the University of Maryland, College Park. K.K. has been on the editorial board of the IEEE/ACM Transactions on Networking and IEEE Network Magazine and a member of the National Research Council Panel on Information Technology for NIST. He has participated in numerous standards bodies working on communication networks.

**Thomas S. Afferton** (IEEE SM 2001) received his undergraduate degree from the University of Virginia and his M.S. degree in Electrical Engineering from Stanford University. He has been with AT&T since 1991 and has held a variety of positions in AT&T Labs and AT&T Network Services. His responsibilities have spanned network architecture, optical technology planning, program management, certification testing, operations planning and network reliability management. He is currently a Division Manager in the Global Network Planning and Development organization in AT&T Labs. He has been a frequent speaker at conferences on carrier applications for optical network technologies. He is a member of the IEEE Communications Society and represents this society on the Steering Board of the Optical Fiber Communications (OFC) Conference. He was a co-Technical Program Chair of OFC 2003 and will be a co-General Chair of OFC 2005. He is also a member of the board of the Optical Internetworking Forum.