

The Case for IP Backhaul

Jeff Loughridge

Principal Consultant, Brooks Consulting LLC

January 23, 2012

In any hierarchical network, designers must specify how the access layer delivers traffic to the core. In Mobile Network Operator (MNO) networks, the transport of voice and data from the cell sites to the wireless MNOs' core networks is called backhaul. Time Division Multiplexing (TDM) backhaul has dominated backhaul deployments since the inception of wireless communication. Leasing the backhaul access of multiple T1s/E1s for every cell site becomes prohibitively expensive in terms of operating expenses, particularly for providers that do not own the last mile. Today's 3G/4G cellular technologies have spurred a major change in the backhaul network: the transition from TDM to packet backhaul.

Ethernet is the most widespread packet-based backhaul technology. While this service is a vast cost and scale improvement over TDM backhaul, carrier Ethernet is a stepping stone in the evolution of backhaul networks. Expect MNOs to move to true IP backhaul networks to meet the scalability needs of their expanding networks. In this article, we will explain mobile backhaul evolution, shortcomings in carrier Ethernet backhaul, and how evolving service requirements will motivate cell site backhaul vendors to add IP-awareness to their networks.

Legacy Backhaul

Cellular systems were initially designed to carry only voice traffic. Since transporting digitized voice was a mature and well-understood technology, there was no need to take a divergent path for the backhaul of voice traffic in early cellular systems. Using TDM had obvious advantages among those being:

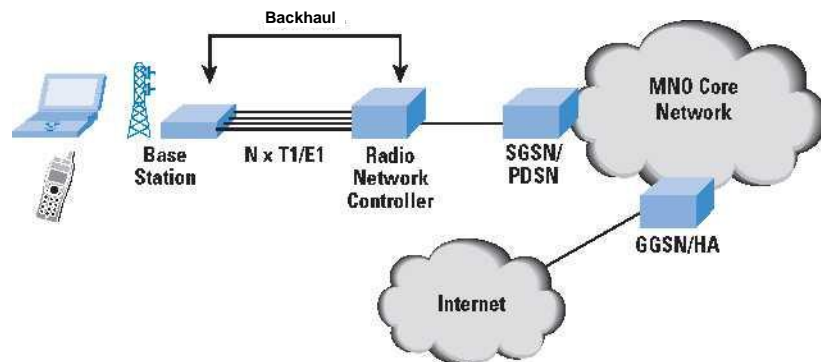
- Use of the same equipment used in wireline voice transmission
- Technical staffs' familiarity with TDM concepts and troubleshooting
- Ability to use existing Operations, Administration, Maintenance, and Provisioning (OAM&P) systems
- Ubiquity of the T1/E1 service

The initial work to offer data service on cellular systems naturally focused on adding data transmission to the existing voice infrastructure. Standards such as Global System for Mobile Communications (GSM) and Interim Standard 95 (IS-95) took similar approaches in borrowing TDM time slots for data. The data services of the 1990s were very slow, even when compared to consumer modems of the time. Standards developed in the late 1990s and deployed in the early 2000s (Enhanced Data rates for GSM Evolution [EDGE] and CDMA2000) improved data transfer speeds.

TDM was clearly entrenched as a foundational technology for data communication in cellular networks going into the early 3G technology deployments (Universal Mobile Telecommunications System [UMTS] and Evolution Data Optimized [EV-DO]).

Figure 1 depicts the backhaul portion of the MNO network and how it fits into the broader architecture.

Figure 1: The Backhaul Network in the MNO Architecture



As data traffic usage for 3G networks grew, shortcomings of TDM backhaul began to materialize. The two prominent areas were bandwidth and cost. Cell sites with TDM access are typically equipped with multiple T1/E1s. With faster radio interfaces, the backhaul became the bottleneck in the network. Some smartphones became consumers of multi-megabyte data rates. User experiences were poor on some wireless networks as a result of a dearth of bandwidth in the back-haul segment. Continuing to increase the number of TDM lines or increase their capacity was not a viable option since the growth increments were too small and the operating expenses were too high.

The second limitation of TDM in 3G networks is cost. Although the cost of T1/E1s decreased considerably over the years, the costs piled up given the number of cell sites and number of T1/E1s per site. This figure became the highest contributor to the cost of the backhaul network. The MNOs that owned the last mile were at a distinct competitive advantage compared with the carriers who had to pay another party (often in a minimally competitive marketplace) for TDM access. For MNOs to continue their incredible traffic growth rates, a new access model was needed.

Carrier Ethernet Adoption

Ethernet quickly emerged as the most popular backhaul technology to replace TDM access infrastructure (other providers moved forward with microwave access with varying levels of success). The various iterations of Ethernet from 1970s to 2000s had trumped other LAN technologies in the market, and at the turn of the century gigabit Ethernet leveraged its success in the LAN to become popular in the WAN. The technology had several major advantages:

- Large drop in cost per bit: Ethernet would allow providers to drastically alter their access cost model by supplanting the aging and costly TDM infrastructure. With the price that consumers were willing to pay per month of data service staying relatively stagnant, this adjustment to the cost model was critical.
- Ethernet can be carried over more underlying technologies: Synchronous Optical Networking/Synchronous Digital Hierarchy (SONET/SDH), Generic Framing Procedure (GFP), Dense Wavelength Division Multiplexing (DWDM), and Multiprotocol Label Switching (MPLS) are a few examples. A key benefit Ethernet's ability to operate over these technologies was that many providers could consolidate their wireless access with their existing and speedier wireline access networks.
- Ethernet interfaces ubiquitous and inexpensive: Ethernet won the battle for LAN dominance. The technology was not restricted to traditional personal computers and servers—printers, phones, game consoles, Digital Video Recorders (DVRs), and home media center hubs are some examples of other equipment that often included Ethernet interfaces. This ubiquity in the business and consumer spaces results in a diverse supplier set and economies of scale for the vendors and suppliers.
- Ease of bandwidth upgrade: TDM circuits have an implementation time measured in months. This slow turn-around time for upgrades is a poor fit for an environment in which data usage is increasing at fast rates. Ethernet is much different. An increase in bandwidth to a network endpoint will not require a change in equipment unless moving between the established tiers of 10,100, 1000 Mb/s. Since the Ethernet service vendor likely uses a "policer" to keep customers within the purchased bandwidth level, a change in software configuration is usually all that is required to upgrade bandwidth. Another advantage is that bandwidth can be upgraded in granular increments. With the right back-end systems, an upgrade will take a matter of minutes. For companies looking to increase the velocity of service deployment, the ability to quickly move to high speeds is very favorable.

Established in 2001, the Metro Ethernet Forum (MEF) played a critical role in the acceptance of carrier Ethernet by wireless and wireline providers. The MEF is not a standards organization like the Internet Engineering Task Force (IETF). Instead, the MEF builds upon the work of standards bodies to establish common terminology, service requirements, and network interface requirements. The MEF created an architecture framework along with measurement and testing specifications. Although the MEF did not eliminate wireless providers' concerns about

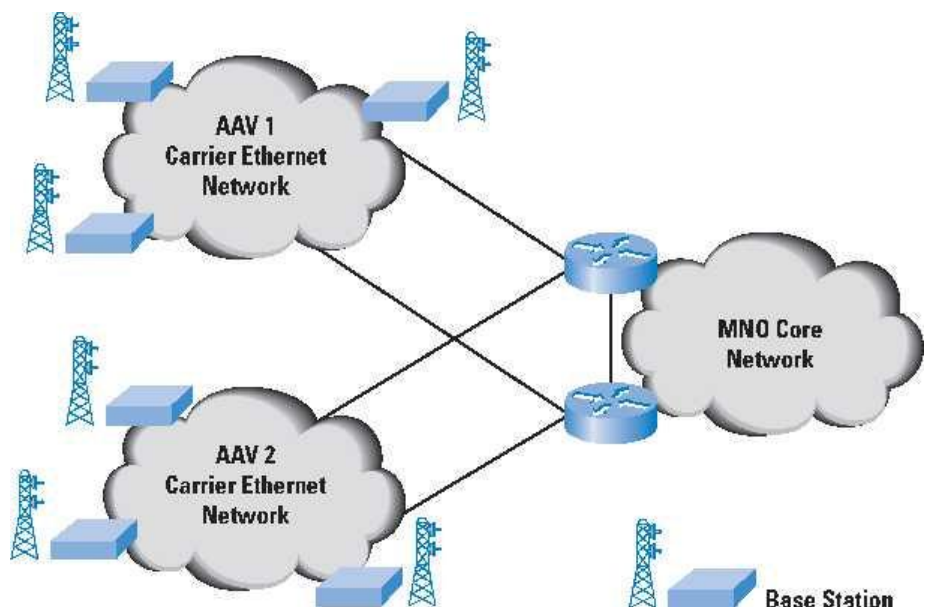
packet backhaul—particularly in the areas of jitter, delay, and packet delivery, the forum did increase the comfort level associated with metro Ethernet services. The MEF's E-LINE service definition established a connection-oriented path, a concept much more pleasing to traditional telcos than the perceived "anything goes" nature of packet switched networks. For more detail on the MEF's service definitions, see [0].

By the second half of the 2000s, many wireless providers were planning the deployment of Ethernet-based backhaul for new High Speed Packet Access (HSPA), Worldwide Interoperability for Microwave Access (WiMAX), and Long-term Evolution (LTE). In making this radical change, the providers often had to consider protecting existing revenue streams from voice and data (providers selecting to move forward with greenfield deployments were at a luxury). Pseudowire technologies enabled the carriage of TDM traffic over IP/Ethernet networks, thus preserving investment in existing infrastructure.

Rather than build carrier Ethernet infrastructure, the MNOs that were not facilities-based (or had limited last mile footprints) purchased services from other parties, known as Alternate Access Vendors (AAV) in telco parlance. In the United States, the Local Exchange Carriers (LECs) and cable companies were well positioned for this business. MNOs often used multiple AAVs in a given market to cover the cell site footprint. Getting fiber to cell sites outside of major metropolitan areas was not always possible, which led some MNOs to use hybrid backhaul solutions that included microwave and TDM inverse muxing in addition to carrier Ethernet.

Figure 2 illustrates how MNOs rely on AAVs to cover their cell site footprint in a given market.

Figure 2: Alternative Access Vendors



The adoption of carrier Ethernet services by MNOs was not without challenges. Mobility gear such as Radio Network Controllers (RNC), base stations, and Home Location Registers (HLR) historically relied on T1/E1 interfaces for connection to the network. Telecom vendors had to implement Ethernet interfaces along with IP stacks. The providers had to completely revamp provisioning, service monitoring, performance monitoring, and service assurance systems and processes. Consider the following example.

For years, operations groups at telcos counted on near-immediate notification with an alarm indication signal in the Time Division Multiple Access (TDMA) frame. TDMA frames arrive every 125 microseconds (8,000 times a second). Packet-switched networks do not share the synchronous nature of TDM and do not have OAM fields in framing bits. The operators now had to rely on nascent specifications such as Y.1731 and 802.1ag for service monitoring. Timing and synchronization—necessities in mobile networks—are gleaned from the physical layer in TDM networks. Asynchronous networks such as Ethernet/IP do not have an inherent mechanism for timing and synchronization. Keeping a single T1/E1 at the cell site is one method to ensure timing and synchronization in a carrier Ethernet scenario; however, the use of upper layer protocols is more appropriate, particularly for new builds that have no legacy TDM circuits. Synchronous Ethernet (SyncE), Precision Time Protocol (PTP, also known as IEEE 1588v2), and Network Time Protocol version 4 (NTPv4) were deployed in backhaul networks to provide timing and synchronization. Note that SyncE transports timing information over the physical layer much like the TDM timing model, while PTP and NTP use IP for transport and are not dependent on an Ethernet physical layer.

The learning and flooding aspects of all Ethernet networks present inherent scaling challenges for very large networks. Spanning tree and its derivatives are commonly used to address these issues at low and medium scale. For larger networks that provide service to multiple customers, the service must scale in terms of its ability to offer service to multiple entities and in terms of the many switches required for an expansive footprint. Many protocols have arisen to solve one or both of these challenges. Examples are Virtual Private LAN Service (VPLS), Multiprotocol Label Switching-Transport Profile (MPLS-TP), and Provider Backbone Bridging-Traffic Engineering (PBB-TE). Being relatively new technologies, these can and do present challenges for operations groups. The breakages can occur in ways that are very difficult for the Carrier Ethernet provider and wireless provider to jointly troubleshoot.

The Next Step - IP Backhaul

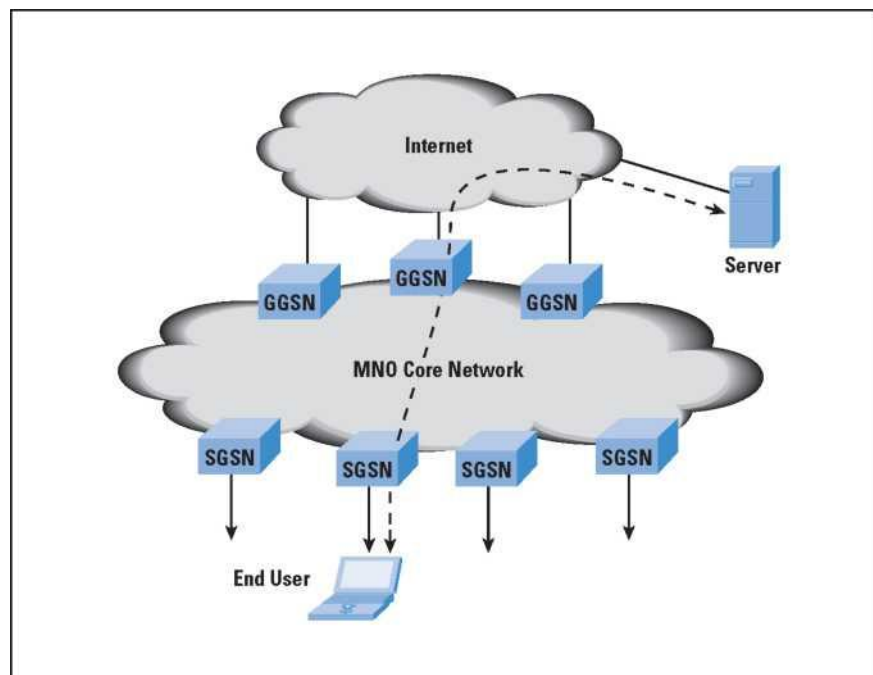
The phrase "all-IP" is frequently used to describe the most recent wireless technologies such as HSPA+, WiMAX, and LTE. This is applicable as the majority of network elements, including the handsets, are IP enabled. The existence of large-sized carrier Ethernet networks in the network architecture undermines the IP-centric argument. IP has superior scaling properties over Layer 2 networks. The footprint and number of nodes for carrier Ethernet networks continues to expand rapidly as the MNOs deploy 3G and 4G networks. The author sees evidence that

protocols used to overcome Ethernet scalability issues will become increasingly complex and push MSOs and AAVs toward Layer 3-centric backhaul networks.

Before delving into the drivers of IP backhaul, let's examine a typical data traffic flow for today's wireless networks. We'll use the 3GPP's GSM Packet Radio System (GPRS) as this is the most common in worldwide deployments. Data flows are very centralized in this architecture. Macro-level mobility is controlled by two types of GPRS Support Nodes (GSN): Gateway GPRS Support Nodes (GGSN) and Serving GPRS Support Nodes (SGSN). GGSNs are typically deployed within the mobile core network at locations with Internet access. This is often at centralized mobile switching centers. SGSNs can be deployed closer to the network edge and multiple SGSNs can be served by a single GGSN.

The GGSN is the mobility anchor, much like the home agent in wireless networks that use Mobile IP. The SGSN is akin to the foreign agent in Mobile IP. GPRS network tunnel traffic between SGSN and GGSN using an IP-in-IP tunneling protocol called Generic Tunneling Protocol (GTP). Although GTP has several purposes in the GPRS core network, our focus will be on its tunneling of packets between SGSN and GGSN (called the Gn interface). The movement of the subscriber to a region served by another SGSN will trigger a macro-mobility event. A new GTP tunnel is formed using the original GGSN for session continuity [2]. Since all traffic from the Mobile Subscriber (MS) must traverse the GGSN as the mobility anchor, the traffic flow from the MS follows a very predictable path to a centralized location. Note that there is not a 1:1 relationship between SGSNs and GGSNs. As mentioned earlier, typical deployment of GGSNs is very centralized. Figure 3 depicts the flow.

Figure 3: Data flow in a GPRS Network



Although technologies like LTE are touted as flat IP networks, this only holds true from a Radio Access Network (RAN) perspective. What if a subscriber wants to communicate with another subscriber in the same building or local machine-to-machine traffic is highly sensitive to latency? The packets will be sent to the mobility anchor, perhaps hundreds of kilometers away. Routing decisions can be made in the RAN and core network; however, the decision is restricted since traffic must traverse the predefined tunnel endpoints.

Wireless networks will gradually decentralize and distribute mobility management. In 3G networks, some providers have been extending the core network closer to the subscriber as mobile gateways (GSNs and their equivalents in non-3GPP networks) become more cost-competitive. By deploying mobile gateways at what were previously aggregation Points of Presence (POPs) and buying Internet connectivity at these locations, Internet-bound traffic exits the network quickly, consuming fewer resources for the provider. Other signs of this shift are evident in LTE and WiMAX. LTE's S1-flex interface allows the RAN to be connected to multiple core networks. The WiMAX reference model separates the Network Access Provider (NAP) and Network Service Provider (NSP). The NAP, which provides radio access functionality, can connect to multiple NSPs for Internet connectivity.

To fully realize the benefits of an IP-centric backhaul, steps must be taken to go beyond simply distributing mobility management. New solutions are needed to eliminate mobility anchoring via tunneling. Vendors, providers, and universities have already started to examine how to dispose of tunneling in the mobile environment [2].

The IP-centric backhaul network has many advantages over the carrier Ethernet networks that enable many of today's packet backhaul networks. Various advantages benefit the wireless providers, the IP backhaul provider, or both. These advantages are most prevalent when the MSOs have a highly distributed mobility management architecture.

- **Backhaul Offload:** Today's mobile elements at the cell tower have no ability to influence routing decisions; there is only one path to the core network. Adding egress points to the cell site or backhaul network reduces the distance and amount of traffic that must be backhauled. To accomplish the addition of egress points in a carrier Ethernet network, connection-oriented mechanisms such as Ethernet Virtual Circuits would require that the MSO and AAV modify multiple network elements' configurations. Offloading traffic with an IP network is substantially more simple and scalable. Offloading packets from the backhaul will represent a huge savings in access costs. The base station could be capable of hot potato routing traffic directly to an ISP

instead of backhauling commodity Internet traffic to the MSO, where the costs of equipment, power, and software licenses quickly accumulate.

- **Multicast:** The reliance on tunneling as described earlier in this piece severely restricts the usefulness of multicast in current wire-less networks. Distributing the mobility elements controlling the tunneling closer to the subscriber will mitigate these effects as would the elimination of mobility anchoring via tunneling techniques. The implementation of a true flat IP network would extend multicast capability into the RAN and position both MNOs and IP backhaul providers to realize the efficiency gains of multicast.
- **Localized Content and Peering:** With localized egress points, local content could be reached directly rather than traversing the core network. This would position wireless providers to peer with other providers at the local or regional level, a benefit that would be substantial for wireless providers operating in countries with non-meshy Internet infrastructure and expensive wide-area communications lines. In addition, caches could be implemented much closer to the subscriber to improve the user experience for video and other content types.
- **Machine-to-Machine (M2M) and Peer-to-Peer (PtP):** When the communication is device to device in close geographic proximity, the traversal of the core network only adds latency, complexity, and cost. A distributed mobility management architecture and IP backhaul network engender an optimized path for M2M and PtP. The mobility anchor point could be placed at the cell tower or local aggregation point, providing a much improved communication path for subscribers and machines connected to the wireless network.
- **Uptime and Reliability:** Wireless providers have experienced challenges with carrier Ethernet service. Some of these problems can be chalked up to the relative newness of using carrier Ethernet for cell site backhaul. One has to wonder though, what experience exists in the industry for maintaining giant Layer 2 networks? The number of mobile devices will expand exponentially, triggering the deployment of thousands of new cell sites, microcells, and picocells. The author is less than confident that any underlying technology that enables carrier Ethernet will scale to the necessary degree while maintaining the uptime and reliability that users expect from their data service.

For large IP networks, the industry has over fifteen years' experience in designing, engineering, and operating IP networking carrying traffic at staggering capacities. The staff expertise, software maturity, and systems support exists today to maintain sizable IP networks. There are established best practices for Tier 1 ISPs that help ensure long uptime, speedy convergence upon failure, and sound network design.

Delivering an IP Backhaul Service

IP backhaul offerings could be delivered in a variety of ways. The simplest design for IP backhaul providers would be a shared IP transport network that commingles traffic between customers. The wireless providers could then use protocols such as Layer 2 Tunneling Protocol version 3 (L2TPv3) to build an MPLS/VPN-like overlay to provide logical separation and address overlap prevention. The preferred approach for MNOs would likely be a Layer 3 VPN service from the AAV, thereby offloading much of the routing complexity from the MNO.

An IP backhaul service must be capable of routing IPv6 packets, as the useful lifetime of an IPv4-only service is limited. MNOs cannot obtain new IPv4 addresses to number the base stations, and using RFC 1918 space is not a scalable approach. Using IPv6-only to address mobility equipment at cell sites (and equivalent radio interfaces) is the preferred method for overcoming the scarcity of IPv4 addresses.

The shift from carrier Ethernet to IP backhaul should not be a monumental one for many carrier Ethernet providers. The heavy lifting of installing fiber and deploying a packet switched infrastructure has already been accomplished. In addition, carriers that implement carrier Ethernet with protocols like VPLS already have an infrastructure that is ready for IP. The most challenging aspect of the transition will be the work needed to prepare OAM&P systems for an IP service. Of course, this may vary based on carrier Ethernet implementation and systems.

Conclusion

Carrier Ethernet service for cell site backhaul is a vast scale and cost improvement over TDM backhaul and has been extremely successful. OSI Layer 3 IP networks have superior scaling properties that will replace Layer 2 backhaul networks of today. Advances in wireless networking systems, the proliferation of new devices, and the development of new mobility services will be best served with a truly IP-centric backhaul network.

References

[0] Santitoro, Ralph, "Metro Ethernet Services—A Technical Overview," 2003, <http://metroethernetforum.org/metro-ethernet-services.pdf>

[1] M. Grayson, K. Shatzkamer, and S. Wainner, IP Design for Mobile Networks, Cisco Press, 2009.

[2] Distributed Mobility Management in Future Wireless Networks (DiMoWiNe), <http://conference.researchbib.com/print.php?category=event&id=10232&uid=6>

JEFF LOUGHRIDGE is the principal consultant and owner of Brooks Consulting LLC, a firm that specializes in Tier 1 ISP best practices and the design, engineering, and operations of large-scale wireline and wireless IP/MPLS networks. Prior to founding Brooks Consulting, Jeff spent over ten years supporting Sprint's global IP network in both technical and managerial capacities. He earned a bachelor's degree in computer science from Duke University and an MBA from the University of Phoenix—Northern Virginia campus.

E-mail: jeffl@brooksconsulting-llc.com

Phone: 703-229-0098