Evolving Backbone Networks with an MPLS Supercore

Planning and Building for Varied Applications and Other Traffic Uncertainties
Executive Summary
Traffic on backbone networks continues to grow rapidly. The bulk of this growth comes from low revenue Internet traffic. Naturally, this raises interest in lowering the costs of building and maintaining service provider backbone networks.

A traditional method of building core networks revolves around the use of connection oriented time-division multiplexing (TDM) circuits. In recent years, the advantages of IP/MPLS core networks have come to the forefront, due to IP/MPLS’s natural statistical multiplexing behavior, seamless multiservice support, and flexible and scalable characteristics. Core networks are at a crossroads, and solving the uncertainties from the rising tide of traffic is a key concern for network and content service providers alike. Video, cloud, and mobile computing are changing traffic characteristics in ways that may threaten, and eventually define, core network architectures.

Still, many existing cores remain based on traditional TDM infrastructure, and implementations of the latest optical networking standard—Optical Transport Network (OTN)—are being discussed as an alternative. To compare and contrast the costs of OTN versus MPLS, Juniper Networks and others have undertaken multiple modeling exercises. Results unequivocally show that the optimal network design depends on different specific characteristics of the traffic matrix; however, MPLS networks are more cost-effective and efficient for all types of traffic except traffic that enters the network as a circuit (e.g., inter-switch voice trunks, or certain private line business services). These results have a very interesting impact on product and network architecture, and they highlight a need (now met by Juniper Networks) for a new class of network element called a packet transport switch. Accordingly, Juniper has introduced the Juniper Networks® PTX Series, which provides higher scale, lower cost, and the flexibility of traditional core routers.

The PTX Series and the flexible core network designs that can be built with it are an integral part of the “New Network.” To demonstrate this, Juniper provides planning services to identify optimal designs and identify places in the network that are either under or overprovisioned. In this way, we’re helping many service providers improve their balance sheets as they prepare for the rising tide of packet-based services.

Introduction
A variety of factors are contributing to the ongoing tidal wave of traffic across global networks, demonstrating an exponential growth curve that signifies a real threat to service provider economics, as costs appear primed to outpace revenue within the next couple of years.

Backbone providers need to provide infrastructure to cope with this growth, and there is no guarantee that these providers will earn a profit on those investments, for much of the traffic does not intrinsically bear revenue. Thus, service providers must balance the value of the technology and the solutions they deliver against the required capital and operations expenditures. Figure 1 illustrates the potential for breakage in the cost model for service providers.
If we assume the traffic growth rate that is illustrated here (a compound annual growth rate approximating 30%) and the current rate of innovation persists, the economic model breaks. One of two things will happen—prices will rise or available bandwidth will diminish until we are in economic equilibrium. Neither of these outcomes is desirable for consumers or for business. Innovation is required to support growth and proliferation of devices while enabling greater efficiencies that drive consolidation and reduction in costs.

The drivers are visible at every turn, from the proliferation of 3G/4G smart mobile devices, to content pushed to network edges supporting “anytime/anywhere” video, to machine-to-machine applications and a social networking user base that could soon reach the billions. The networks to handle this not only need increased capacity but also far greater flexibility than the ones they are replacing, because sources and destinations vary constantly, both in terms of end users and the servers comprising cloud applications.

There is no longer a linear relationship users and where they access the network, and these untethered users (Figure 2) are already consuming massive amounts of content, thus straining the design, capacities, and caching strategies of present-day networks.

Based on industry analyst forecasts, these trends are likely to continue. Gartner\(^1\) predicts that tablet sales will grow 181% in 2011 to $54.8 million, many of which are built to take advantage of mobile 3G and 4G networks. According to International Data Corporation (IDC)\(^2\), we will reach 1 billion mobile devices in 2013, and their cloud computing survey estimates that server revenue in the private cloud category will grow from $7.3 billion in 2009 to $11.8 billion in 2014, a gain of 62 percent.

### Service Provider Challenges

The main challenges that service providers face are to reduce costs, and to enable new value-added services that were previously not possible. The traffic explosion in the core comes in unpredictable volumes, characteristics, and patterns. Yet the transport sectors of some of the world’s largest networks still run over circuit-based SONET/SDH equipment, originally developed in the early 1990s and only gradually being upgraded to cope with increasing voice and data demands.

---

1. www.gartner.com
2. www.idc.com
Although this approach has ensured stable service delivery for a large array of applications, and has kept the technology pains of early Internet growth away from the majority of business and residential customers, it has also resulted in high overhead on top of packet services. Thus, it represents one of the largest opportunities for cost reduction in service provider networks.

When you “peel open” the circuits in the core of service provider networks, you often will see packets from the applications described above. In addition, this packet traffic growth is rising according to most Tier 1 providers—estimates range that the increase is from less than 50/50 (packet/circuit) to 70/30 over the next couple of years. Providers are often coping with this growth today by establishing parallel networks to handle the new data, but this only adds to the operational complexity of having multiple network layers (the lower ones focused exclusively on circuits) in the first place.

The result is multiple management systems and backbone systems, as well as an underutilization of provisioned capacity. Many operators are aware of being overprovisioned in parts of their transport network, and underprovisioned in others, and they know they can’t continue to build out their core networks as if traffic were fully deterministic: “It is the nature of stochastic “bursty” traffic that peak demand will be much larger than average demand. Providing for the peak would be very expensive, and certainly against good engineering economic principles.”

The Arithmetic of Uncertainty

Holding down cost while dealing with exponential traffic growth presents conflicting business requirements. And today’s connected culture brings a new dimension to the problem in the form of multiple uncertainties on the services plane (Figure 3).

![Figure 3: The “Arithmetic of Uncertainty”](image)

In and of themselves, the propagation of new applications brings an element of uncertainty, and when combined with the consolidation of data centers, and resources migrating from location to location (based on where they are consumed), the change in traffic patterns presents a moving target for service provider executives. Moreover, the economic model is further disrupted by a kind of disintermediation, caused by the fact that content providers often build their own infrastructure, establishing direct connections between consumers and content and “cutting out” the service provider.

Existing Core Design Strategies

Due to this pervasive business tension, the design of telecommunications networks is a key task in today’s economy. Many design trade-offs are being actively considered in the aggregation and core layers of these networks. The factors identified above increase the amount of bandwidth that must be provisioned in edge networks, and network designs from Tier 1 providers are already calling for upstream links well in excess of 100 Gbps. Cost versus benefit is difficult to estimate when so much of the traffic entering the network does not generate revenue.

The mobility issue in particular affects aggregation networks because, as currently built, they are static and are based on knowing where the users are, where the content is stored, where peering points are located, etc. All of this is now fluid and dynamic.

The legacy TDM model still plays a role in the buildout, even for new services. Because SONET/SDH equipment is already in place in network aggregation layers, service providers may currently (in many cases) aggregate Ethernet edge interfaces into (for instance) OC-48 interfaces. In these instances, OTN may turn out (temporarily) to be the transport technology of choice even for packet-based services, since the overhead of a circuit-switched paradigm is already in place.

For this reason, it is estimated that circuit handled traffic may grow 3-5X, whereas packet traffic is expected to grow on the order of 10X during the next five years (2011-2016), even though the rise of traffic that begins as Ethernet will be much greater than these numbers indicate. On the other hand, network service providers that are more broadband-based than private line focused (e.g. business VPN) will see even greater relative packet growth.

And this is where many internal debates are centered. Provisioning backbone networks is a key task in the design of the communications infrastructure. Network operators have to cope with an ever-changing market, which makes it difficult to choose the most appropriate and cost-effective technology. The industry is proposing several solutions to this, with many of them focusing on OTN (sometimes referred to as a “hollow core”) versus IP/MPLS in a packet-switched paradigm. There are many factors involved in choosing between these options, and current infrastructure weighed against anticipated traffic characteristics and growth always have to be taken into consideration.

**Option 1: Same Architecture, Current Core Routers**

The first option involves building the core with bigger IP routers. This is the “safest” path to take in many respects and involves the simplest topology.

The capital costs of this approach may be higher than the circuit- or bypass-based options described below, but the operational costs are predictable and likely to be lower because it is the “present mode of operation” in so many cases. But, as one CTO remarked in 2009, “We need to fundamentally rethink how we’re carrying traffic in our networks. We have to rethink how we interoperate, how networks are constructed, how routing is done. How we move content off hours.”

The potential implications of bigger IP routers in the same architecture could include an increasingly cumbersome and expensive infrastructure, as well as mismatched functionality of IP for simply and economically moving large numbers of packets. This complexity can compromise capabilities, escalate power and space consumption, and make convergence difficult (as many network layers are needed under the IP layer).

---

*See comments by John Donovan, CTO of AT&T: [http://news.cnet.com/8301-30686_3-10382598-266.html](http://news.cnet.com/8301-30686_3-10382598-266.html).*
Option 2: OTN in Core, Routing at the Edge
Due to the caveats noted in option 1, others are proposing a focus on OTN in the core, and to reserve routing for the edge of the network.

The theory here is that you confine all the complexity of routing, queuing, and service handling to the edge of the network. This architecture is actually reminiscent of the trend in the 1990s to build out core networks with ATM and keep IP at the edge. The static provisioning of the point-to-point connections throughout the ATM core resulted in N-squared scalability problems and a lack of the necessary flexibility. Operationally, managing the two types of network equipment also proved to be burdensome. Thus, pushing this complexity to the edge also means adding costs to the edge, which offsets the savings in the core. As one global service provider noted, “Basing the network on OTN doubles the complexity of planning and modeling.”

A core architecture based on OTN means that one must provision for peak traffic, and when the core depends on OTN circuits, it is ill-suited for bursty and dynamic traffic. The complexity shifts to the edge and increases costs even more.

---

Option 3: The “Bypass” Route
A third option is to continue to maintain multiple layers and to bypass core routers whenever possible.

As with an OTN core, this two-layer approach requires provisioning for peak traffic and works only with highly static traffic. It also requires multiple centralized network management systems for the IP and the OTN/dense wavelength-division multiplexing (DWDM) layers, and these inhibit scale and add technical and operational complexity to the network.

Although optical bypass does not efficiently handle varying traffic patterns or bursts, it can (on rare occasions) be used effectively in parts of the network where it is known that transit traffic is both plentiful and identifiable. However, the risk of pursuing this too aggressively is that you have to overprovision, and you build a very static network. The content of transit traffic changes from node to node, and there is a very high risk of setting up many more circuits than are needed.

Architectural Considerations for a Better Core
Historically, the choice between a packet-based network and one based on circuit technology was heavily influenced by the system cost, but if a system is MPLS-optimized, there is no intrinsic reason why the cost should not be approximately that of circuit-switched (OEO) devices. Thus, cost optimization of modern core transport is driven by the forwarding plane.

Furthermore, one should note that the control plane of MPLS also has some built-in advantages. It can be either static or dynamic, with traffic engineering capabilities for secondary and standby tunnels. Fast reroute mechanisms can be used for detour and bypass, or can use the newer facilities of local repair.

Some of the advantages of a dynamic control plane include:

- Ability to scale to over 100,000 nodes reaching edge to edge
- Faster recovery and rerouting mechanisms leading to 50 ms restoration
- Automatic bandwidth for MPLS tunnel elasticity
- Integrated Layer 1 to Layer 3 protection for seamless recovery
- Aggregation of variable traffic patterns with statistical multiplexing
- Autodiscovery of shared risk link groups
- Optimal handling of multiple bursty, variable speed traffic flows

---


---
These MPLS characteristics are well known and documented in the literature cited here. The main issues that service providers are confronting in utilizing MPLS as a transport technology revolve around the cost differences of IP/MPLS routers as opposed to less expensive and lower functioning circuit switches. Fortunately, there are new classes of packet transport routers on the horizon that address the problems inherent to the emerging traffic matrices.

There is a lot of uncertainty and unpredictability in the emerging next generation of networks. And service providers, on both the “service” (typically IP) and “transport” (typically optical switching) side of the organization, want the uncertainty removed.

In particular, a traditional paradigm wherein bandwidth is “thrown at the problem” will no longer work. The economic model for the next several years, as demonstrated in the Introduction above, will break under this mode of operation. The allocation of bandwidth needs to be flexible in such a way that it is not dependent on peak traffic, because traffic flow sizes vary widely based on many factors such as routes becoming more dynamic, and will be even more unpredictable going forward. The more elasticity to handle peak flows, the better.

Network capacity planning amidst traffic uncertainty forces a trade-off:

- Circuit switching at a lower initial cost (at the risk of having too much or too little capacity)
- Packet routing that manages statistical variation well

As customers look at the total cost of ownership (TCO) of MPLS, OTN, and hybrid networks, the benefits of statistical multiplexing where it is indicated—for most of the traffic coming onto the network—become apparent. Juniper’s modeling shows savings ranging from 40–65% depending on traffic characteristics such as flow size, peak-to-average ratios, and simultaneous traffic peak occurrences. These factors impact the decision on whether to keep traffic in the MPLS domain or move it to the OTN or optical domain. Belotti et al. highlighted the importance of statistical multiplexing on multilayer planning, and the importance of considering the different bandwidth provisioning rules applied to circuit and packet networks.

From a technological point of view, the core network infrastructure must be flexible enough to accommodate changes in traffic characteristics. This will be a critical capability as traffic patterns in the core of the network change due to a number of factors—new applications, data center consolidation, cloud computing, and of course the rise of mobile and video content.

Detailed modeling demonstrates that the core of the network needs to be packet-based to be most efficient. Aggregation and multiplexing need to happen at the packet-level. Integration of the packet layer and optical transport layer will improve the economics of the core network, both CapEx and OpEx.

Although intrinsically more expensive (at the network level) and based on an older paradigm that can no longer dominate viable core networks, OTN can be deployed in a supportive role, i.e., in the transport layer for services that originate as circuits (Figure 7).
The key to successful implementation of this model will be software. Seamless integration of the IP/MPLS and optical control planes will allow for coordinated provisioning, management, and restoration of these multiservice networks.

The goal of a unified packet transport network is to consolidate service delivery infrastructure. Today’s networks need to carry both legacy (circuit-based) services as well as rapidly growing packet-based services. In many cases, for both historical and operational reasons, both traffic types run on totally separate networks.

OTN transport (long-haul transmission, switching, and multiplexing) can help to consolidate these networks by bringing the following functionality:

- The ability to multiplex L1, L2, and L3 networks over 100 Gbps wavelengths
- Improving wavelength reach and Operation, Administration, and Management (OAM)
- Migrating SONET/SDH links onto 100 Gbps wavelengths
- Private line services
- MPLS, on the other hand, is critical in areas that OTN cannot support well, specifically:
  - Handling dynamic bursty variable-speed traffic and dynamically aggregating traffic. The market forces outlined above create thousands of variable-speed traffic flows that can only be economically processed at the packet level.
  - Providing protection. This has both an operational and a financial impact and can be optimized in either the OTN layer or IP layer. The exact mix of services and the layer at which they are provided has great significance into the ultimate choice.

Thus, an emphasis on MPLS forwarding yields the most flexible and cost-effective network.

The product and network architecture depicted in Figure 7 allows a promising approach for multilayer network planning, a service provided by Juniper Networks. The goal is to avoid having large IP flows traverse multiple nodes in the IP domain whenever the associated cost is not justified. In essence, it promotes selective offloading of big IP flows over the optical transport network, reducing the number of core routers traversed.

**The New Supercore with the PTX Series**

In response to the market needs described above, Juniper has developed the PTX Series (Figure 8) with a new cost-optimized platform design and a unique combination of hardware and software. The PTX Series provides a solution inspired by the challenges faced by service providers today. With PTX Series, the services dictate network connectivity and bandwidth, not the other way around.

![Figure 8: PTX Series Packet Transport Router](image-url)
The PTX Series (which includes the 8-slot PTX5000) is based on Juniper’s patented Junos Express chipset, which is the foundation of the packet transport platform and which provides:

- 8 Tbps in the PTX5000
- 480 Gbps/slot in 10, 40, and 100 GbE interfaces
- 720 Mpps/slot (5 times the packet processing of the nearest competitor)
- Up to 65% cost reduction over OTN solutions
- State-of-the-art silicon fab: 40 nm
- Full function MPLS traffic engineering
- Built in “no packet drop” assurance

The PTX Series provides a unique combination of hardware (fastest switching silicon, MPLS forwarding, long-haul optics) and software (unified switching and transport control and management planes, operational software tools) to enable service providers to manage their core networks more efficiently.

Built from the ground up for speed, scale, and optimal cost, the PTX Series flexibly adapts to rapidly changing traffic patterns caused by video, mobility, and cloud-based services, and is the industry’s first Converged Supercore packet router, supporting single chassis systems with 8 Tbps of capacity. Finally, the modular power system of the PTX Series allows industry-leading power efficiency of 1 W/Gbps Energy Consumption Rating (ECR).

The PTX Series (which fits easily into existing IP core networks) allows a service provider to build a core network that is flexible enough to accommodate dynamically changing traffic patterns caused by factors described above—new applications, data center consolidation, mobility of content users, and increased video consumption.

Only a packet-switched core, which the PTX Series provides in the most scalable and cost-effective manner, is flexible enough to tackle this challenge. The interfaces on PTX Series devices support OTN framing for management and greater signal distance. The transport level resiliency of the PTX enables migration of the Circuits to Packets with the same Service Level Agreements.

The seamless integration of the IP/MPLS and optical control and management planes allows coordinated design, modeling, planning, simulation, provisioning, management, and restoration of multiservice networks built using the PTX Series.
Introducing Junos Express

As stated above, the PTX Series is built with the new Juniper Networks Junos® Express chipset. This is the fastest chipset that Juniper has ever introduced in its long history of industry-leading silicon.

Figure 10: Junos Express

Junos Express is purpose-built for speed, scale, low cost, and power efficiency. Built with a 40 nm process geometry (first release 480 Gbps/slot), platforms (such as PTX5000) built with Junos Express can comfortably scale to petabit levels.

The following table highlights the scale, differentiation, efficiency, and reliability of the Junos Express chipset.

Table 1: Characteristics of the Junos Express Chipset

| Scale | • Extra high density to satisfy today’s bandwidth requirements  
|       | • Extreme scale to meet tomorrow’s bandwidth growth demand |
| Differentiation | • Rich built-in traffic engineering for differentiated handling of different traffic types  
|       | • Full delay bandwidth buffer support for network congestion management |
| Efficiency | • Advanced architectural and integration techniques that optimize system components  
|       | • Focus on core switching to enable low power operations |
| Reliability | • In-system “end to end” cyclic redundancy check (CRC) that protects customer data at all times  
|       | • Embedded error detection capability and redundancy for fast recovery |

Junos Express is four times as fast as the packet forwarding chipset of competing platforms, and uses one-third the power, while providing ten times the ultimate system scale.
Towards a Converged Supercore

Figure 11 illustrates the PTX Series in a Converged Supercore.

Within the “inner edge” or within the “outer core” (where the Junos Trio chipset is used), additional flexibility in the form of IP routing may be needed, whereas within the supercore one can simply take advantage of statistical multiplexing and dynamic, elastic, label-switched path (LSP) creation and management. In Figure 11, either Juniper Networks T Series Core Routers or MX Series 3D Universal Edge Routers connect to the PTX Series in the supercore.

At the edge of the network, operators can focus on subscriber management, services, and bandwidth management, whereas at the core of the network the focus can be on higher scale, speed, and lower cost. The Junos Trio (for edge) and Junos Express (for core) chipsets have been designed with these main capabilities in mind.

Evolutionary Steps

Ultimately, further simplification takes place with the convergence of the control and management planes. The ideal supercore shown in Figure 12 is driven by the same multilayer control plane and managed by the same management system. It uses both the optical transport capabilities and huge capacity compared to traditional core routers. The following figure illustrates the architecture of a supercore solution when used in conjunction with an optical switch under the control of a converged network management system (NMS).
This scenario illustrates a single multilayer control plane and NMS communicating with both the PTX Series and the optical switch. Service providers have the advantage of both an IP/MPLS control plane (where the network elements “know” the locations of their neighbors), and an optical control plane in which the topology does not dynamically change.

Figure 13 illustrates a pathway for the industry to get there. Today’s telecommunication transport networks (shown on the left) consist of a stack of subnetworks with different technologies, so-called layers. Every link in an upper layer is realized by one or more paths in the next lower layer. For example, an IP link between two Internet routers may be realized by one or more light paths in an underlying optical fiber network.

The PTX Series can simplify this, as it provides the functionality of all of these layers in a single platform, supported by a single Juniper Networks Junos® operating system.

In the existing IP core infrastructure shown on the left, note that separate management systems work at each of three layers—one for IP, one TDM, and one for wavelengths. The introduction of PTX Series devices, when equipped with colored optics and integrated with a reconfigurable optical add-drop multiplexer (ROADM) at the optical layer, eliminates a layer of management.

The PTX Series architecture also supports TDM services, and the realization of OTN capabilities will mean that single control and management planes handle the entire transport system. Multilayer provisioning, planning, and restoration can be supported across all layers.

Juniper will guide this offering with a professional services offering for network migration and optimization (see Appendix A).
Conclusion
The affordable way to support all electronic communications on a converged IP platform is to build smart networks that use capacity efficiently, and to avoid overbuilding whenever it is not necessary.

In order to remove the uncertainty and unpredictability from the planning process, network service providers are making an effort to better understand their customers, applications, and traffic. They can then start operating their core networks more efficiently by no longer “throwing bandwidth at the problem” (in the form of “nailed-up” circuits) and by growing their networks in such a way that they can phase their investments.

With this knowledge and strategy in hand, they can allow the services they are providing to guide the architecture and evolution of the transport layer. This will be a welcome new capability in the hands of service provider executives, who of necessity have had to take a more bottom-up approach in the past, an approach that they can no longer afford in today’s richly connected, content-hungry global culture.

Appendix A: Planning Services for Core Network Design
Juniper Networks can remove the uncertainty and unpredictability from the network planning and provisioning process. Our planning services help you understand your customers, applications, and traffic in ways that were not possible before, so that you can start operating your core network more efficiently.

The facets of this offering are:

- Existing architecture evaluation
- Network traffic analysis
- Network design and modeling based on existing and future traffic requirements
- Cost comparison analysis for various network architecture options

Our network modeling tools and services will provide build-out costs under different scenarios of load, failures, and uncertainty, and allow you to fully explore different technology options before testing and deploying them. Juniper provides options to relieve worst-case bottlenecks with optimal levels of provisioning and the avoidance of provisioning for different levels of growth in traffic.

For more information, please contact your Juniper Networks account representative.

About Juniper Networks
Juniper Networks is in the business of network innovation. From devices to data centers, from consumers to cloud providers, Juniper Networks delivers the software, silicon and systems that transform the experience and economics of networking. The company serves customers and partners worldwide. Additional information can be found at www.juniper.net.