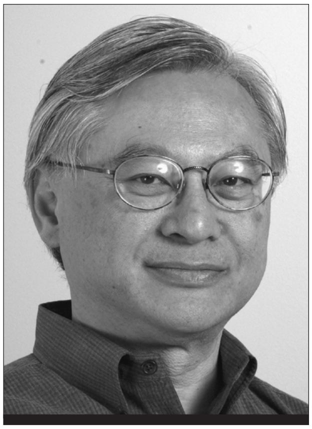
Near-Term Future of the Optical Network in Question?



Vincent Chan

In this editorial, we have often talked about and debated the research directions of the optical network of the future. We have discussed futuristic transport mechanisms such as burst and flow switching, and, in our most expansive and creative moments, even optical packet switching. However, there is a near-term architecture problem that needs our attention as well. The problem is: "As the optical fiber network moves toward the next generation (i.e., 40 and 100 Gbps per wavelength), what should be the best architecture in the core and access networks?"

It seems there is general agreement that the future growth in network capacity will mostly support data traffic. Even for real-time and streaming applications, IP and IP-like traffic (most likely originating in Ethernet format) will be the norm. Thus, at least in name, there seems to be agreement that the next-generation fiber network will be 40-100 Gbps Ethernet per wavelength in the core. However, there is ongoing debate as to what that will mean exactly. At one extreme, router companies envision simple Ethernet switches in the core operating at very coarse granularity; whereas, traditional telecom companies often take the other extreme of "carrier-class Ethernet switching in the core," which is almost an altered-ego of SONET switching with all its trappings of timing accuracies, clock distribution, synchronization, as well as a full complement of alarms for protection and restoration. Which of the two extremes is pursued will have a first-order impact on network cost, mostly due to different partitioning of the grooming functions in the network and the associated electronics complexity and manufacturing cost. While the ongoing debates will be settled soon in standards bodies through near-term investment of equipment companies, the research community can think of the 'right' architecture without near-term pressures and the politics of market positioning and standardization. Thus, I will pose a generic research problem that can shed some light on this issue.

Before examining the network architecture design, we should understand where the new traffic will be coming from. Let us assume, for the time being, that the near-term (3-7 years) increase in network capacity in CONUS and Europe will be one order of magnitude. (I am excluding in this exercise the developing countries, which have little network infrastructure in place to date.) A simple-minded estimate can come from the change from 10 to 100 Gbps per wavelength, but that is technology driven and not application driven. If we examine user traffic in developed countries, the increase in an order of magnitude in total capacity will not come from new user entries but will come from an increase in application data rates. Thus, the cost of network service per bit must also come down by almost an order of magnitude as well; otherwise, the prohibitive access cost will not support the transition. To first order, then, individual user streaming rates and transaction sizes will scale up by roughly an order of magnitude.

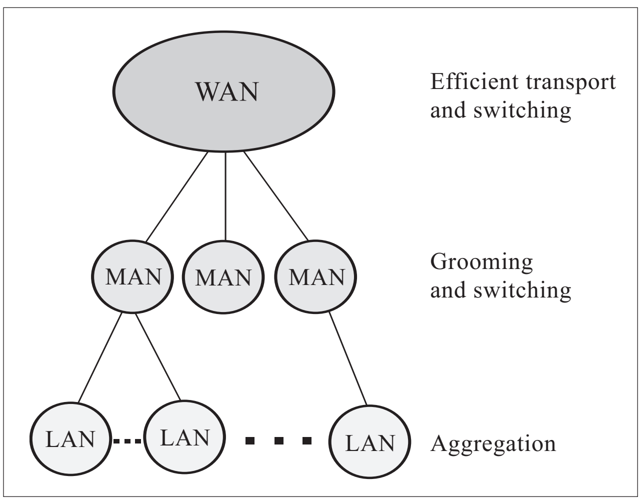


Figure 1.

The primary function of the LAN is aggregation (see Fig. 1). For the future, Ethernet is as good as any transport mechanism and it has the property of being cheap. For future high-end applications (i.e., those characterized by large transactions) there may be a feature that allows the access switch to instruct the network interface card (NIC) at the computer to hold-off on transmission until the switch is ready to accept the data, thus providing a form of admission control to prevent congestion downstream in the network. (The links in the LAN should be typically lightly loaded; and any congestion will be in the MAN or WAN.) This feature is particularly attractive — in fact, a must, unless data loss internal to the network is acceptable — if a significant fraction of the traffic does not use TCP or a TCP-like protocol in Layer 4 for congestion control.

In the MAN, grooming is done at grooming switches. This function is necessary to use expensive wavelengths in the core efficiently. SONET and Ethernet over SONET

(Continued on next page)

(EoS) are mostly used today. MPLS and GMPLS provide circuits for VPN and other circuit services. There is nothing to prevent a hybrid system with SONET, EoS, and G709. However, to reduce cost and complexity, it makes sense to do away with some of the redundant functions. Thus, Ethernet over wavelengths may be the transport mechanism of the future.

In the long-haul, there is little debate that eventually there will be rates up to 100 Gbps per wavelength. However, as mentioned earlier, there is contention as to the switching granularity in the core. The coarsest is wavelength switching only. Some traditional telco equipment manufacturers and carriers want SONET-like clock stability, fine grain cross-point switches, and the full spectrum of network management tools such as SONET/G709 alarms so that they can do fine grain switching in the core nodes as in the telco days. One manufacturer claims that these types of Layer 2 switching ports cost only 20% of router ports. The opposing view, however, is that router companies can and will achieve lower prices since they employ essentially the same technology in their routers. Router companies believe that the carrier approach drives up the cost of transport and negates the benefits of the Ethernet philosophy of simplicity, albeit with less clock stability and management, thus leaving most network management functions to the higher layers. So who is right?

The interesting, overarching research problem is "based on some parametric model of the costs of components and subsystems, what should be the right network architecture in the core network?" Some important questions, whose answers may lead to the right architecture, are:

1. What part of switching and routing should be done at Layer 2 and what part should be done at Layer 3? In addition to wavelength switching, will there be finer grain Layer 2 switching (some say 1 Gbps granularity makes sense)? Is fine grain switching necessary at all in the core? After all, transaction sizes may increase by an order of magnitude. If we need fine grain switching, perhaps there is not enough traffic to transition to 100 Gbps per wavelength.

2. What are the costs of the different architectures? We should examine all the advantages and disadvantages from the manufacturing perspective. Do SONET-like stability and pedigrees drive up electronic costs tremendously? Is SONET-like timing stability at 100Gbps difficult (yield, heat and power issues, at 65 nm or smaller scale size) and thus costly? If fine grain switching is done in Layer 3, will it be even more expensive?

3. Last but not least, we have been led to believe less layers are better (i.e., cheaper) and we thus got rid of ATM, frame relay, and now SONET. Why are we reintroducing Ethernet in the core? Why don't we just do wavelength switching only in the core and be totally blind to the format within each wavelength and let the MAN and the higher layers do the finer grain grooming? If the traffic has been well aggregated, why shouldn't the core only change circuits in a quasi-static fashion, leaving major highways (of wavelengths) connecting major MANs and only add or subtract wavelengths based on slow changes in average loads?

The full scale problem is very complex, involving all aspects of network design from the data plane to the control plane, and complexities of electronic functions and costs of chips that support these functions. However, a simple parametric study, untethered by political and standard-based constraints, may be illuminating.