

Jean-François Labourdette &
Zhenheng Zhang
Editors

The Interconnect Penalty of Small Size Switches

Jean-François Labourdette
Tellium

Carriers are often faced with the need to scale the switching capacity of a central office site by installing multiple switches at that site. We address in this column the question of whether to interconnect multiple switches within the same site, and we assess the penalty incurred due to ports reserved for interconnection, the so-called “interconnect penalty.”

We consider core optical networks consisting of central office sites with optical switches interconnected by point-to-point WDM fiber links in a mesh configuration. Each WDM fiber link carries multiple wavelength channels. Multiple fiber links from adjacent sites may be incident on a given site. Optical switches enable re-configurable optical networking by rapidly provisioning end-to-end circuits called lightpaths between the client edge equipment. Figure 1 illustrates such a core optical network.

Such core networks are sparsely connected with a small average number of fiber links incident at each site (say three). Because of the sparse connectivity, a typical lightpath travels several hops, and as a result, a *large portion* of the traffic at a central office site is pass-through traffic, and a *small portion* of the traffic is originating traffic. A typical node could have 70% of pass-through traffic and 30% terminating traffic.

Carriers architect their network and dimension the optical switches and WDM fiber links based on a de-

mand forecast over several time periods. Initially all sites may contain a single optical switch equipped according to the traffic forecast for the initial period(s). As traffic at the site (both originating and pass-through) grows, the site switching capacity needs to be scaled accordingly. Typically, it is done in the following fashion. First, the optical switch is equipped in-service up to a preset threshold (say 70%) of switch capacity. Then, additional switches are installed at the site and interconnected to yield a larger switching complex.

In one possible configuration, channels from each WDM fiber link could be terminated on a different but *single* switch. Traffic passing through the central office would necessarily consume interconnection capacity to be routed from its incoming WDM link onto its outgoing one. Since pass-through traffic is a significant proportion of the total traffic, such a configuration would be very inefficient in the way switch ports are used. Therefore, a better configuration would have WDM channels from each fiber link terminating on all (or a majority of) switches as illustrated in Figure 2, showing a site with three interconnected switches. Channels from each WDM system are terminated on all switches.

How to terminate WDM channels in an office and to interconnect the switches together is driven by traffic forecast and network planning. Actual traffic demand is different from expected demand due to uncertainties in traffic forecast. For example, consider an unforecast orig-

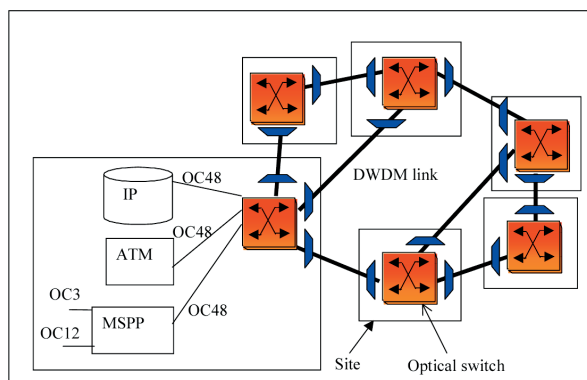


Figure 1: Core optical network.

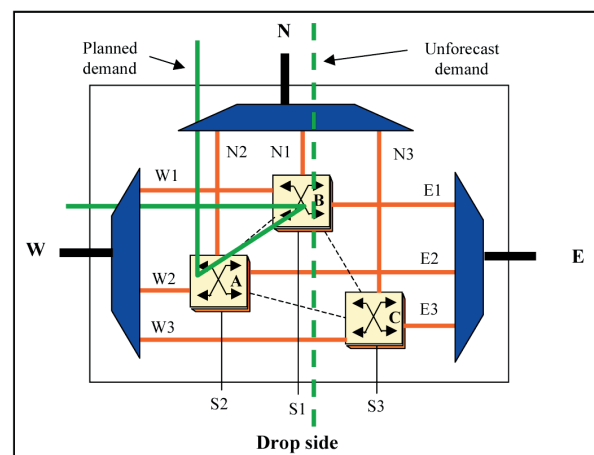


Figure 2: Model of a site with multiple switches; WDM links are spread among switches.

inating demand from S1 routed through switch B to N1, exhausting capacity on N1. Consider now that, because of similar occurrences at other sites, W2 and W3 capacity is also exhausted. With interconnect capacity between switches A and B, future planned traffic demand from **W** to **N** can still be routed from W1 to N2. The example illustrates that unless interconnections are provided between switches A and B, future planned traffic demand from **W** to **N** would block even though network capacity is available on W1 and N2.

To quantify this phenomenon, we define Q to be the fraction of the total traffic that has ingress and egress ports (including drop side ports) on the same switch, and therefore does not require any interconnection hops at a site. This is also a measure of forecast accuracy. When $Q = 1$, no interconnection is theoretically required. The assumption that all traffic will have ingress and egress on the same switch at a site, however, is unrealistic, as discussed above. Actual traffic will deviate from planned traffic patterns, requiring the use of interconnection channels between the switches at a site. As a result, the fraction of total traffic, Q , which has ingress and egress ports on the same switch, is less than 1 in practice, creating the need for interconnections.

Figure 3 illustrates the number of interconnection ports required (as a fraction of the total ports at the site) as a function of Q , the traffic prediction accuracy for different values of the number N of switches per site. Here we assume that traffic/connections are never blocked.

In the extreme case when $Q = 0$, i.e., for completely arbitrary traffic patterns (no forecast), and no blocking, the above result indicates that approximately $2/3$ of the ports at a switch would have to be dedicated for intercon-

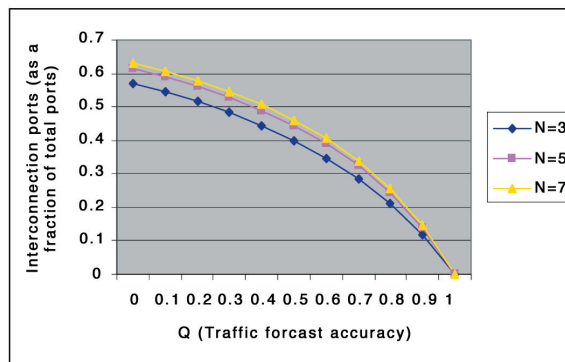


Figure 3: Interconnection ports (as a fraction of the total ports) as a function of Q .

nections with other switches. The number of useful ports would then be at most (about) a third of the total ports at the site. When $Q = 0.7$, i.e., 70% of traffic is accurately forecast and does not need interconnection hops, about 30% of the ports still need to be interconnection ports.

These findings have implications on both desirable switch size and network architecture. This will be the subject of a future column.

Acknowledgment

I would like to acknowledge Ramu Ramamurthy for carrying out the analysis of this problem.

References

R. Ramamurthy, J-F. Labourdette, S. Chaudhuri, "Scaling Switching Capacity with Multiple Switches," to be submitted for publication.