Part VI: Advanced Topics (Bonus Material on CD-ROM)

This part includes additional material that are related to Part IV and Part V; it consists of two sub-parts.

In the first sub-part, three chapters (Chapter 21, Chapter 22, and Chapter 23) cover functions and components of a router in further detail as a continuation of Part IV. First, different approaches to architect the switch fabric of a router are presented in Chapter 21. Second, packet queueing and scheduling approaches are discussed along with their strengths and limitations in Chapter 22. Third, traffic conditioning, an important function of a router, especially to meet service level agreements, is presented in Chapter 23.

In the second sub-part, we include two chapters (Chapter 24 and Chapter 25). Transport network routing is presented first in its general framework, followed by a formal treatment of the transport network route engineering problem over multiple time periods, in Chapter 24. The final chapter (Chapter 25) covers two different dimensions: optical network routing and multi-layer network routing. In optical network routing, we discuss both SONET and WDM in a transport network framework; more importantly, we also point out the circumstances under which a WDM on-demand network differs from a basic transport network paradigm. Furthermore, we discuss routing in multiple layers from the service network to multiple views of the transport networks; this is done by appropriately considering the unit of information on which routing decision is made and the time granularity of making such a decision. We conclude by presenting overlay network routing and its relation to multilayer routing.

25

Optical Network Routing and Multilayer Routing

Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference.

Robert Frost

Reading Guideline

The basic background on optical networking is included in the chapter in order to understand the routing problems. To understand the relevant routing problems, basic background on network flow modeling (Chapter 4) and some background on transport network routing (Chapter 24) are helpful. The material on multilayer routing requires knowledge about a variety of networking technologies covered throughout the book, and how they are related.

D. Medhi and K. Ramasamy, *Network Routing: Algorithms, Protocols, and Architectures*.

Optical network routing is an important problem domain in communication networking. Optical networking is usually used for transport services. For such services, we describe the routing problems for a representative set of scenarios for synchronous optical networks (SONET) or synchronous digital hierarchy (SDH), and wavelength division multiplexed (WDM) networks.

The second area we cover in this chapter is multilayer routing where routing coordination is introduced between two layers such as IP and WDM. With the perspective of optical networking, such a multilayer routing environment provides a perspective on how the interaction can work and future possibilities for dynamically reconfigurable networks and services.

25.1 SONET/SDH Routing

Before we discuss the routing problems, we start with a brief overview of SONET/SDH.

25.1.1 SONET/SDH Overview

A widely deployed technology for transport networks is synchronous optical network (SONET), or synchronous digital hierarchy (SDH). SONET is widely deployed in North America and SDH is deployed in the rest of the world—but both provide the same functionality. We present a brief overview of SONET/SDH technology pertinent to our discussion. The interested reader is directed to books such as [77], [509], [510], [580], [748] for additional details about SONET/SDH.

Nodes in SONET or SDH networks are equipped with devices such as terminal multiplexers (TM), digital cross-connects (DXC), and add-drop multiplexers (ADM). TMs and DXCs are used in transmission networks with mesh topology, while ADMs are typical nodes of ring networks. There are several data rates available for SONET/SDH that are given as a synchronous transfer signal (STS) for SONET and a synchronous transport module (STM) for SDH (refer to Table 25.1 for these rates). In SONET standard, optical carrier (OC) levels are also defined corresponding to electrical equivalents in STSs. To complicate this further, SONET/SDH standard allows subrates for carried demand. These subrates are referred to as virtual tributaries (VTs) in SONET and virtual containers (VCs) in SDH (see Table 25.1). Furthermore, old-style rates such as T1 and T3 can also be connected to SONET/SDH nodes through service adapters.

SONET/SDH technology can be used either as a mesh or ring. An important alternative to the mesh SONET/SDH networks discussed is the SONET/SDH ring networks where the restoration mechanisms are intrinsic to the system. This is contrary to the mesh case where restoration requires inter-DXC signaling, for example, using GMPLS. Self-healing SONET/SDH rings have been heavily deployed around the world due to its < 50 millisec restoration capability for any single-link failure. The nodes of a SONET ring network are also called ADMs and are capable of inserting or extracting any VC or VT container of the set of all containers circulating around the ring. Figure 25.1 depicts a bidirectional line-switched self-healing ring (BLSR) with four optical fibers (because of four fibers, they are also referred to as BLSR/4). Now assume that this ring is based on an OC-48 transmission system, i.e., the system that can hold 16 OC-3s. The ring is divided into two pairs of fibers, one basic pair and one protection pair. OC-3s destined for a particular node are extracted from the incoming

TABLE 25.1 Transmission rates for SONET/SDH, and subrates (VC for STM and VT for STS).

(a)

SONET Signal	SDH Signal	Bit Rate (Mbps)
STS-1 (OC-1)	–	51.84
STS-3 (OC-3)	STM-1	155.52
STS-12 (OC-12)	STM-4	622.08
STS-48 (OC-48)	STM-16	2,488.32
STS-192 (OC-192)	STM-64	9,953.28
STS-768 (OC-768)	STM-256	39,812.12

(b)

VC Type	Bit Rate (Mbps)
VC-11	1.728
VC-12	2.304
VC-3	48.960
VC-4	150.336

(c)

VT Type	Bit Rate (Mbps)
VT-1.5	1.728
VT-2	2.304
VT-3	3.456
VT-6	6.912

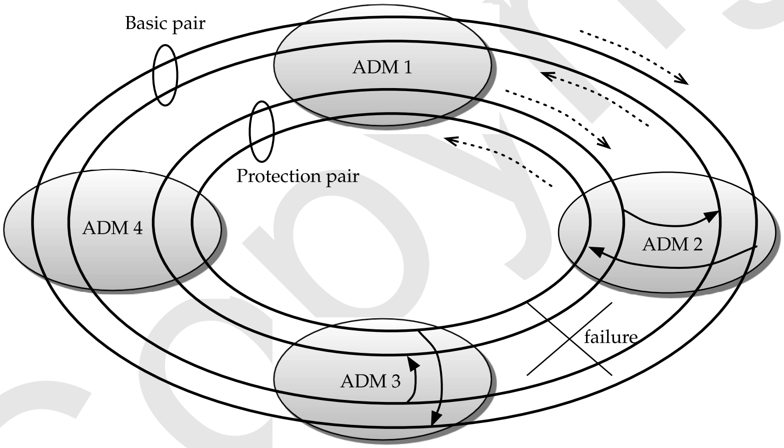


FIGURE 25.1 Bidirectional line-switched ring (BLSR) with four-nodes.

basic fiber (for example, the outermost fiber in Figure 25.1), while the originating OC-3s are inserted into the outgoing (second outer) basic fiber.

Each OC-3 can, in turn, contain 84 VT-1.5s; note that VT-1.5s are designed to map a T1. Thus, an OC-3 can effectively carry 84 T1s worth of demand. Now each OC-3 may or may not be completely filled with T1s, while to the SONET ring it sees only the OC-3s, *not* T1s that may reside. Thus, a transmission hierarchy can be built; you can start seeing its multi-layer nature due to the transmission rate hierarchy. It is, however, worth noting that data rates lower than 50 Mbps are starting to go away due to increase in demand for higher data rate optical demands.

There is an important difference between concatenated and non-concatenated STS signals. A SONET STS-1 rate frame is organized as 9 rows and 90 columns, i.e., as 810 bytes of information. The first 3 columns are set aside for transport overhead; thus, the effective data rate (payload) is 783 bytes ($= 9 \times 87$). A non-concatenated STS-N is formed by byte-interleaving of N STS-1 signals and has N distinct 87×9 byte payloads; on the other hand, the concatenated STS-Nc, where "c" denoted concatenated, has one payload of $87 \times 9 \times N$ bytes. It may be noted that most deployment of SONET are STS-Nc based. Because of concatenation, it is also preferable to write as OC-48c, to distinguish from non-concatenated OC-48; here, we use them interchangeably.

25.1.2 Routing in a SONET Ring

We consider an OC-48 SONET ring with four nodes where we want to route OC-3 demands. We will use the topology shown in Figure 25.1 to illustrate this example; note that the protection pair is not considered in this illustration. Traffic demand volume between two SONET nodes in terms of OC-3s is given as follows:

node i \ node j	2	3	4
1	4	4	8
2	–	4	8
3	–	–	8

Note that demand is bidirectional and is shown in the upper diagonal part of the traffic matrix. The capacity of the OC-48 ring in terms of OC-3s is 16. Consider the demand between nodes 1 and 2. In this illustration, we assume that the entire demand for a pair of nodes can be routed either on a clock-wise or a counter-clockwise direction (the general case in which the demand is allowed to be split is left as Exercise 25.1). Thus, we can use a decision variable to indicate the choice of one over the other:

$$u_{12,12} + u_{12,1432} = 1,$$

where $u_{12,12}$ stands for the clockwise direction while $u_{12,1432}$ stands for the counterclockwise direction for demand pair (1:2). A similar situation exists for the other five demands:

$$\begin{aligned}
u_{13,123} &+ u_{13,143} &= 1 \\
u_{14,1234} &+ u_{14,14} &= 1 \\
u_{23,23} &+ u_{23,2143} &= 1 \\
u_{24,234} &+ u_{24,214} &= 1 \\
u_{34,34} &+ u_{14,3214} &= 1.
\end{aligned}$$

Now consider the link segment 1-2 on the ring. This will contain the following decision variables, if chosen:

$$u_{12,12},\ u_{13,123},\ u_{14,1234},\ u_{23,2143},\ u_{24,214}, \textit{and}\ u_{14,3214}.$$

Note that only one for each demand pair would be considered for each link. Now each decision variable for each pair, if chosen, will need to bear the demand for that pair. In addition, the capacity of the ring may not be exceeded. Thus, for link 1-2, we can write

$$4\,u_{12,12} + 4\,u_{13,123} + 8\,u_{14,1234} + 4\,u_{23,2143} + 8\,u_{24,214} + 8\,u_{14,3214} \le 16.$$

Similarly, we can write constraints for link segments 2-3, 3-4, and 1-4. A goal in ring network routing is that links are load balanced. To do that, a load-balancing variable is introduced that is to be minimized. Thus, we use instead

$$4\,u_{12,12} + 4\,u_{13,123} + 8\,u_{14,1234} + 4\,u_{23,2143} + 8\,u_{24,214} + 8\,u_{14,3214} \le 16\,r,$$

where *r* is the load-balancing variable to be minimized. Putting everything together, we can write the routing problem with load balancing as the goal as follows:

$$\begin{array}{ll}
\textbf{\textit{minimize}}_{\{r,u\}} & r \\
\textbf{\textit{subject to}} & \\
u_{12,12} + u_{12,1432} = 1 & \text{(pair 1:2)} \\
u_{13,123} + u_{13,143} = 1 & \text{(pair 1:3)} \\
u_{14,1234} + u_{14,14} = 1 & \text{(pair 1:4)} \\
u_{23,23} + u_{23,2143} = 1 & \text{(pair 2:3)} \\
u_{24,234} + u_{24,214} = 1 & \text{(pair 2:4)} \\
u_{34,34} + u_{14,3214} = 1 & \text{(pair 3:4)} \\
4\,u_{12,12} + 4\,u_{13,123} + 8\,u_{14,1234} + 4\,u_{23,2143} + 8\,u_{24,214} + 8\,u_{14,3214} \le 16\,r & \text{(link 1-2)} \\
4\,u_{12,1432} + 4\,u_{13,123} + 8\,u_{14,1234} + 4\,u_{23,23} + 8\,u_{24,234} + \ u_{14,3214} \le 16\,r & \text{(link 2-3)} \\
4\,u_{12,1432} + 4\,u_{13,143} + 8\,u_{14,1234} + 4\,u_{23,2143} + 8u_{24,234} + 8u_{34,34} \le 16\,r & \text{(link 3-4)} \\
4\,u_{12,1432} + 4\,u_{13,143} + 8\,u_{14,14} + 4\,u_{23,2143} + 8\,u_{24,214} + 8\,u_{14,3214} \le 16\,r & \text{(link 4-1)} \\
\text{all } u\text{s are 0 or 1} & \\
r \ge 0. &
\end{array} \tag{25.1.1}$$

On solving the above problem, for example using CPLEX, we find that $r = 1$. This means that at least one segment of the ring is completely occupied. The optimal solution is $u_{12,12} = 1, u_{13,123} = 1, u_{14,14} = 1, u_{23,23} = 1, u_{24,234} = 1, u_{34,34} = 1$. On checking, we can see that two segments, 2-3 and 3-4, are fully occupied.

A general question is what does it mean if $r > 1$ at the optimal solution? It means that there is not enough bandwidth on the ring to carry all demands. Thus, capacity expansion is necessary. The above model is useful both for routing decisions and to indicate if capacity expansion is needed.

The general model for Eq. (25.1.1) can be written in a similar way. To write the general model, consider a ring with *N* nodes. Let h_{ij} be demand between node *i* and node *j*; as before, we will consider $i < j$. For ease of notation, we will identify the binary variable *u* as clockwise or counterclockwise, i.e., no superscript shown if clockwise or marked as "counter" in the superscript for counterclockwise. We need an indicator to identify that only one of two paths for each pair is to be chosen when considering capacity constraints. Specifically, δ_{ij}^{ℓ} takes the value 1 if for pair *i*:*j*, the clockwise path uses link ℓ. Finally, we will use *c* to denote capacity. Then, the general model takes the following form:

$$
\begin{array}{lllll}
\textbf{\textit{minimize}}_{\{r,u\}} & r & & & \\
\textbf{\textit{subject to}} & & & & \\
 & u_{ij} + u_{ij}^{\text{counter}} & = 1, & i, j = 1, 2, \ldots, N, i < j, & \\
 & \displaystyle\sum_{(i,j),\ i<j} \Big(h_{ij}\delta_{ij}^{\ell} u_{ij} + h_{ij}(1-\delta_{ij}^{\ell})\, u_{ij}^{\text{counter}} \Big) & \le c\, r, & \ell = 1, 2, \ldots, L, & (25.1.2) \\
 & u_{ij}, u_{ij}^{\text{counter}} = 0 \text{ or } 1 & & & \\
 & r \text{ nonnegative.} & & &
\end{array}
$$

You can compare this one with the four-ring example above to see how a general model can be represented, and how specific problems can be represented.

We close this section with the comment that in a ring, when capacity is to be expanded, it is for the *entire* ring, not just a segment. Thus, there can be one OC-48 ring, two OC-48 rings, and so on, around the entire ring. How is this related to the above problem? Let us assume that we have two OC-48 rings; then c needs to reflect that in the formulation as 32 if demands were in OC-3s. If we now solve the model with the new capacity, it is possible that when we identify the optimal flows, one would have to be split to go on one ring and the rest to go on another ring. This brings up the issue of whether demand can be split. We assumed above that demand cannot be split. If we continue with this assumption, we face the situation of split demand from the solution to the above integer linear programming problem. Instead, what we can do is to solve for one ring; this will result in $r > 1$. Now, identify the link segments that overflow the capacity of the ring, and then identify the minimum amount of demand that can be taken out, but would still result in feasible flows for the rest of the demand. Now, the leftover demand can be considered, and the above model can be used assuming the capacity this time is for the second ring. The case in which rings allow demand split is left as an exercise.

25.1.3 Routing in SONET/SDH Transport Cross-Connect Networks

In Chapter 24, we presented the need for and the role of transport networks and briefly discussed the technology they use. We now discuss routing in SONET/SDH cross-connect networks for a Type B classification (see Table 17.1).

Examples of typical services that create demand for the transport provided by SONET/SDH are trunks for digital circuit-switched networks, IP network trunks, and private leased-line/virtual network services. It may be noted that while the SONET/SDH standard did not originally address interfacing with IP network routers, it has been possible to use SONET/SDH as a transport for IP network links between two routers through an interfacing mechanism called Packet over SONET/SDH (PoS).

The design questions for SONET/SDH transport networks are a bit complicated because of the actual data rates and interfaces available for a particular SONET/SDH network. An input demand (sometimes at subrate) could come into one of these interfaces depending on the type of node functionality deployed in a network.

For illustration, we consider the case in which a SONET network is used as the transport for an IP backbone network. The demand is assumed to be at OC-3 level for IP network trunks. Thus, we can count demand between an ingress cross-connect node and an egress cross-connect node in terms of OC-3 demand. The links interconnecting the transport nodes

TABLE 25.2 Notation used.

Notation	Explanation
Given:	
K	Number of demand pairs with positive demand volume
L	Number of links
M	Modular capacity of a link
h_k	Demand volume of demand index $k = 1, 2, \ldots, K$
c_ℓ	Integral capacity units of link $\ell = 1, 2, \ldots, L$
$c_{\ell n}$	Integral capacity untis of link ℓ for type n
P_k	Number of candidate paths for demand k, $k = 1, 2, \ldots, K$
$\delta_{kp\ell}$	Link-path indicator, set to 1 if path p for demand pair k uses the link ℓ; 0, otherwise
ξ_{kp}	Nonnegative unit cost of flow on path p for demand k
Variables:	
x_{kp}	Flow amount on path p for demand k

are composed of optical transmission systems OC-n, where $n = 12, 48, 192, 768$ (Table 25.1). Capacity c_ℓ of transport link ℓ is expressed in terms of multiples of OC-3s.

First, we assume that the entire network has links of only one type, say, OC-12s. We use the same notation we introduced earlier in Chapter 4. For ease of reading, notations are summarized in Table 25.2. Then, the minimum cost routing problem for the SONET cross-connect transport network can be written as follows:

$$\begin{aligned}
&\textbf{\textit{minimize}}_{\{x\}} && F = \sum_{k=1}^{K}\sum_{p=1}^{P_k} \xi_{kp} x_{kp} && \\
&\textbf{\textit{subject to}} && \sum_{p=1}^{P_k} x_{kp} = h_k, && k = 1, 2, \ldots, K, \\
& && \sum_{k=1}^{K}\sum_{p=1}^{P_k} \delta_{kp\ell} x_{kp} \le M c_\ell, && \ell = 1, 2, \ldots, L, \\
& && x_{kp} \text{ nonnegative integers,} &&
\end{aligned} \tag{25.1.3}$$

where $M = 4$ and c_ℓ means number of OC-12s on link ℓ; ξ_{kp} is the unit cost of path p for demand k, and h_k is the demand volume for demand identifier k; and P_k is the set of possible candidate paths pregenerated for consideration in the above formulation, which can be generated using a k-shortest paths algorithm. Compare this formulation (and the notation) with the general formulation, presented earlier in Eq. (4.4.7). They are the same except that the capacity constraint takes into account the modular factor for OC-12s.

How does the problem change if the links are a mix of different types, such as OC-48 and OC-192? The problem formulation changes slightly as shown below:

$$\begin{aligned}
&\textbf{\textit{minimize}}_{\{x\}} && F = \sum_{k=1}^{K}\sum_{p=1}^{P_k} \xi_{kp} x_{kp} && \\
&\textbf{\textit{subject to}} && \sum_{p=1}^{P_k} x_{kp} = h_k, && k = 1, 2, \dots, K, \\
& && \sum_{k=1}^{K}\sum_{p=1}^{P_k} \delta_{kp\ell} x_{kp} \le \sum_{n=1}^{4} M_n c_{\ell n}, && \ell = 1, 2, \dots, L, \\
& && x_{kp} \text{ nonnegative integers.} &&
\end{aligned} \qquad (25.1.4)$$

In this model, the summation on the right side of the capacity constraint captures M_1, M_2, M_3, and M_4, which refer to capacities of OC-12, OC-48, OC-192, and OC-768, respectively, counted in multiples of OC-3s; similarly, $c_{\ell 1}, c_{\ell 2}, c_{\ell 3}, c_{\ell 4}$ refer to the number of OC-12, OC-48, OC-192, and OC-768, respectively, on link ℓ.

The above two models are still somewhat simplified models. Often, demands might need to be diversified or protected from a failure. For an example of how protection can be incorporated, see Section 24.4. For a discussion on how to incorporate more complicated constraints, see [564, Chapter 4].

Finally, you may note that the transport network routing problems for both SONET ring and SONET cross-connect networks can be formulated in the MCNF framework while the objective considered can be different and, certainly, the number of path choices does differ.

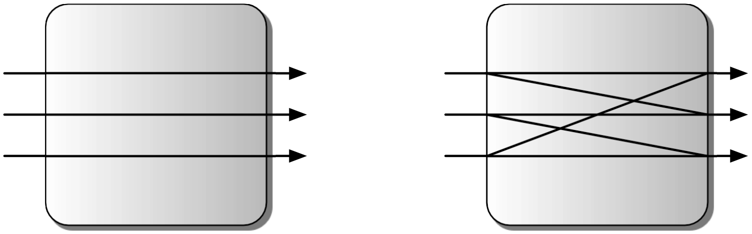
25.2 WDM Routing

We next consider routing in wavelength division multiplexed (WDM) networks. We first present an overview of WDM.

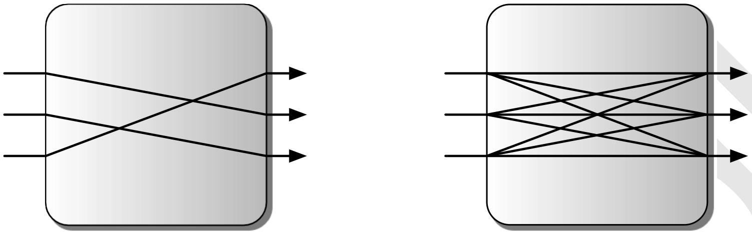
25.2.1 WDM Overview

In the past decade, WDM has received much attention [510], [580]. In WDM networks, traffic demand corresponds to *wavelengths* called *lambdas*. Capacities directly correspond to optical fibers. One wavelength is typically capable of carrying 10 Gbps, while one optical fiber can typically realize up to around 100 different wavelengths. The nodes of the WDM networks are called wavelength cross-connects (WXCs).

There are four types of wavelength conversions for a WXC (Figure 25.2): (1) no wavelength conversion, (2) fixed wavelength conversion, (3) limited wavelength conversion, and (4) full wavelength conversion. From Figure 25.2, we can see that a WXC without conversion can only serve as a pass-through device; other forms have some conversion, and, finally, some have full conversion, which is then like a crossbar switch. The reason for different types is that their costs are different. Thus, a provider might be able to afford one or the other type of device based on its traffic demand. The illustrations shown in Figure 25.2 are for 2-degree nodes, i.e., nodes that connect two locations. It is now increasingly popular to consider higher-degree nodes. For instance, a 3-degree node means that a wavelength coming from one of the three locations can be routed to either of the other two locations using a wavelength-selective cross-connect or a wavelength-selective switch.



(a) No conversion (b) Limited wavelength conversion



(c) Fixed wavelength conversion (d) Full conversion

FIGURE 25.2 Wavelength conversion: (a) no conversion, (b) limited wavelength conversion, (c) fixed wavelength conversion, and (d) full conversion.

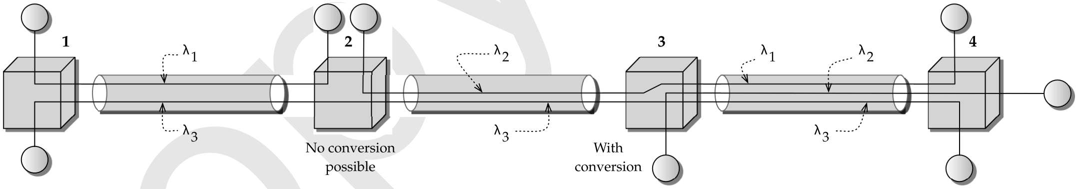


FIGURE 25.3 WDM network, with and without conversion.

What then is a route in a WDM network? It is a lightpath between two nodes that may go through multiple intermediate cross-connects. If there is no conversion, the lightpath must stay on the same wavelength; if there is conversion, some switching to another wavelength is possible. In Figure 25.3, we show a set of lightpaths through a linear WDM network where an intermediate node has conversion capability and the other does not. Because of the association with lightpath, the WDM routing problem is commonly known as the *routing and wavelength assignment* (RWA) problem.

It may be noted that there are certain practical issues to consider in a routing problem. For example, if a path is too long, it may require to have regeneration. For a detailed discussion on impairments and constraints in optical layer routing, refer to [674].

25.2.2 Routing in WDM with Full Conversion: Transport Mode

As you can probably realize, the routing problem for transport service in a WDM network is a minimum cost routing problem of integer MCNF type. Below, we present the routing problem identifying where and how this is different from the general MCNF.

In a WDM network each lightpath is identified with a demand to be routed. There may be many different distinct demands between the same two endpoints (see Figure 25.3); for each distinct demands, the path chosen need not be the same. In a full conversion WXC environment, it can take any path. If we consider all the distinct demands in the network, then each session (regardless of its endpoints) must be routed on a lightpath. Thus, for the purpose of formulation, we can list all distinct demands simply identified as $k = 1, \ldots, K$, without specifying what the endpoints are. What is the capacity of a link then? It is the number of lambdas allowed on a link. Thus, the problem can be formulated as follows:

$$\begin{array}{lll} \textbf{\textit{minimize}}_u & F = \sum_{k=1}^{K} \sum_{p=1}^{P_k} \xi_{kp} u_{kp} & \\ \textbf{\textit{subject to}} & \sum_{p=1}^{P_k} u_{kp} = 1, & k = 1, 2, \ldots, K, \\ & \sum_{k=1}^{K} \sum_{p=1}^{P_k} \delta_{kp\ell} u_{kp} \le c_\ell, & \ell = 1, 2, \ldots, L, \\ & u_{kp} = 0 \text{ or } 1, & \end{array} \tag{25.2.1}$$

where u_{kp} is the path decision variable for the specific distinct demands to be routed if path p is selected and c_ℓ is the capacity of a link in terms of number of wavelengths allowed. The rest of the notations are the same as summarized in Table 25.2. As mentioned earlier, the candidate paths to be considered need to take into account impairments and other constraints [674].

We need to make an important comment about K. Note that K is the total number of sessions to be routed, regardless of its endpoints. Consider a network with N nodes; then there are $N(N-1)/2$ demand pairs. Assume on average that there are J number of distinct demands for each pair. Then, $K = J \times N(N-1)/2$ is the total number of sessions. Thus, K can be a large number for a network with a large number of nodes. Note that Eq. (25.2.1) is an integer linear programming problem. Thus, it can be time consuming to solve for large K. This is when you want to determine how often such a routing configuration should be done for transport networking and whether the computation can be done off-line. If the answer is yes to both these questions, then a canned integer linear programming solver may suffice.

25.2.3 No Conversion Case

The no conversion case is somewhat more complicated to model. Note that a ligthpath must stay on the same wavelength for the entire path. We present here a formulation discussed in [675]. In addition to the path selection variable u_{kp} for each session k, we want to assign this session to only one wavelength i; we thus need another variable w_{ki} to relate this requirement. Furthermore, for each link, it must be the same wavelength for a particular session; this means the product $w_{ki}u_{ikp}$ should not be more than one when considered for each link ℓ and each wavelength i. Formally, we can formulate the problem as follows:

$$
\begin{aligned}
&\textbf{\textit{minimize}}_{u,w} && F=\sum_{k=1}^{K}\sum_{p=1}^{P_k}\xi_{kp}u_{kp} && \\
&\textbf{\textit{subject to}} && \sum_{p=1}^{P_k}u_{kp}=1, && k=1,2,\ldots,K, \\
& && \sum_{i=1}^{I}w_{ki}=1, && k=1,2,\ldots,K, \\
& && \sum_{k=1}^{K}\sum_{p=1}^{P_k}\delta_{kp\ell}w_{ki}u_{kp}\le 1, && \ell=1,2,\ldots,L, \quad i=1,2,\ldots,I, \\
& && u_{kp}=0 \text{ or } 1 && \\
& && w_{ki}=0 \text{ or } 1. &&
\end{aligned}
\tag{25.2.2}
$$

The difficulty with the above problem is that it is a *nonlinear* integer programming problem due to the product term; these types of problems are the hardest to solve in general. Certainly, heuristic approaches can be developed. Another possibility is to linearize the above problem by defining a third variable to replace the product term. See [675] for further details.

25.2.4 Protection Routing

A WDM transport network can be set up with protection routing. With GMPLS signaling, FAST-REROUTE can be used for fast restoration to a backup path in case there is a link failure. Thus, any demand between two nodes would need to have a primary path and a backup path. Second, if there are different demands for customers requiring either full or partial protection, these would need to be accommodated by the transport provider as well. For this purpose, the transport network routing design problem presented earlier in Section 24.4 is applicable; thus, we refer you to this section for how the routing problem can be formulated. Note that if all demands are to be protected, instead of some being partially protected, the same model can be used. In this case, the value for protection-level parameter, α_k^s, is needed to be set to 1, and again the model presented in Section 24.4 is applicable.

It is worth noting that in addition to GMPLS, there are hardware-based and control-plane mechanisms are also available. For instance, automatic protection switching is available for protection. For additional discussions, see [77]. Also, diversity can be used as an alternative to backup paths, which serves as a mechanism to provide some level of connectivity if one of the paths fails where each path is limited in what it can carry due to diversity requirements. This is another type of constraints that can be incorporated in a modeling framework.

25.2.5 On-Demand, Instantaneous WDM services

In recent years, there have been efforts to provide on-demand, instantaneous WDM services. This means that the customer request arrival is similar to a voice call arrival, and a request blocking cannot be ruled out. Then, in the WDM network, the routing problem will be on demand, unlike in transport mode discussed earlier. Since the request requires a dedicated wavelength, the on-demand problem is essentially similar to the dynamic routing circuit-switched routing problem. One major difference is the conversion capability of nodes; if nodes have full conversion capability, then this is similar to dynamic call routing, certainly allowing for multilink paths, which we discussed and analyzed earlier in Chapter 10 and Chapter 11, as well as QoS routing presented in Chapter 17. Thus, issues such as trunk reservation are important to consider in routing decision to minimize request blocking. When the nodes do not have full conversion, the general issue is similar—the main difference is some paths are not allowable due to this restriction. In any case, we refer you to these chapters for understanding routing and control implications, which would be similar in an on-demand, instantaneous WDM routing network.

25.3 Multilayer Networking

25.3.1 Overview

Within the context of the transport network, we can see that a transport network provider has its own domain to meet the demand requirement through transport node equipment and transport network links. It is important to point out that three different ISPs could conceivably use a single transport network provider as shown in Figure 25.4, or an ISP network may be carried by multiple transport network providers as shown in Figure 25.5. Furthermore, it is possible that a transport network provider would carry customer requirements for Internet, telephone network, or private-line customers' networks (as shown in Figure 25.6). Regardless, note that routing within its own network remains the responsibility of each provider, be it an ISP, a telephone service provider, a virtual private network provider, or a transport network provider.

It is becoming apparent that the overall conglomerate of these various networks gives rise to a *multilayer network* environment where each layer has its own definition of traffic, link capacity, and node gears (i.e., functionalities provided by the equipment in a node).

To put it simply, the architecture of communication networks can be complicated; this is due to not only the large number of nodes that can form a particular network, but also the

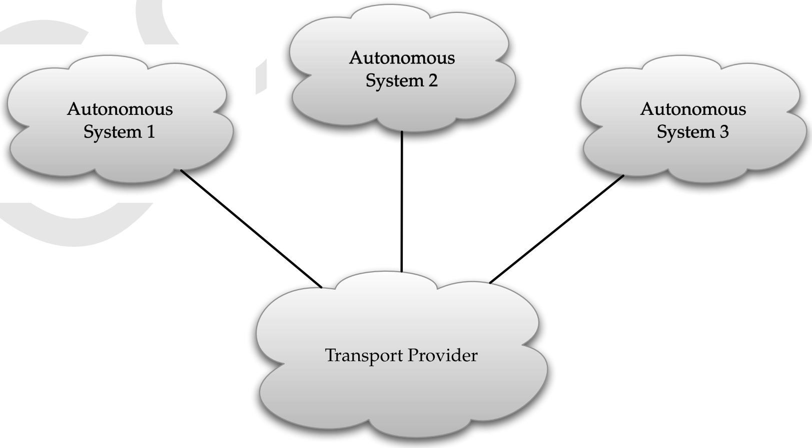


FIGURE 25.4 Three different administrative domains using the same transport provider.

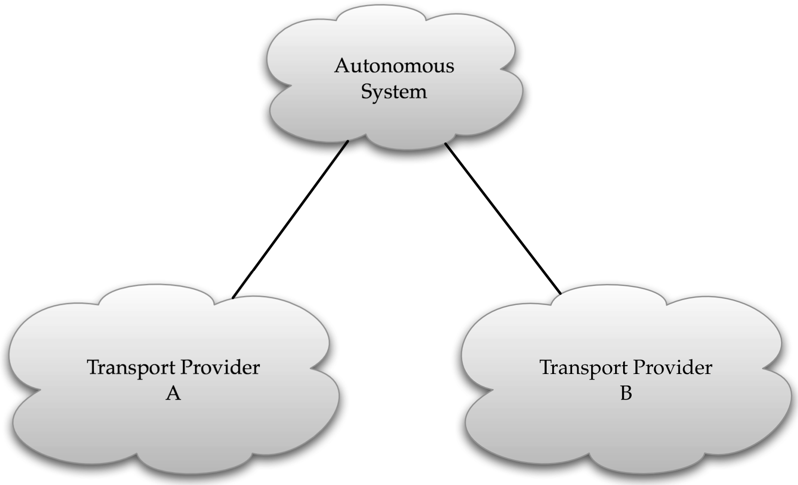


FIGURE 25.5 An administrative domain using multiple transport providers.

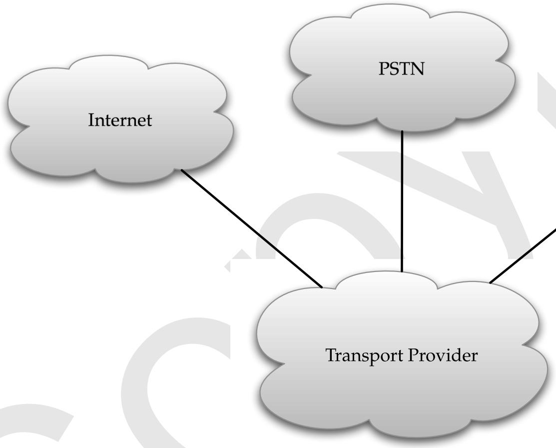
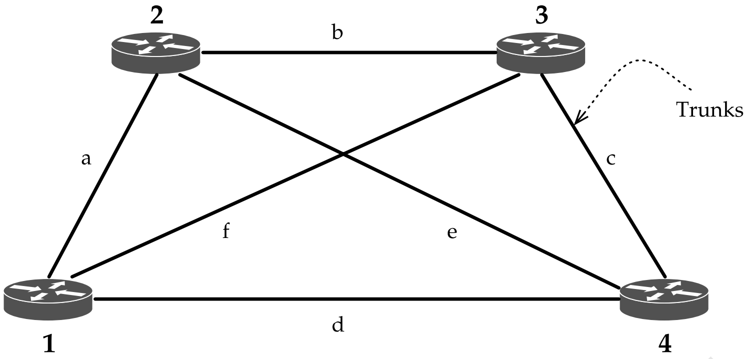


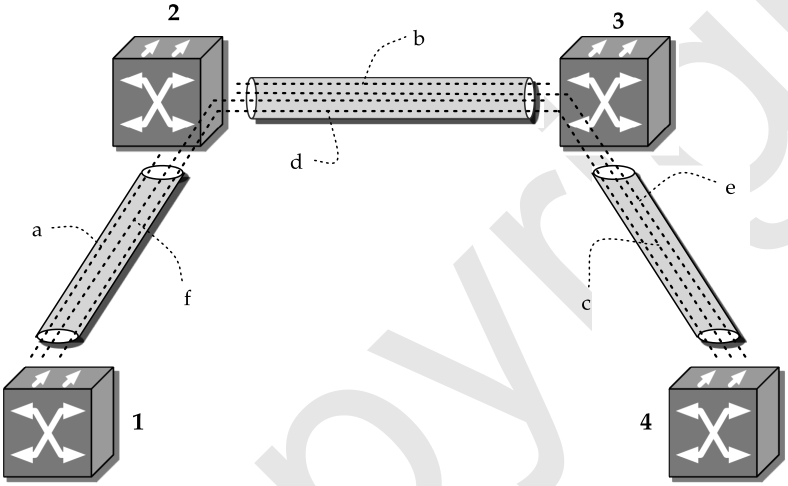
FIGURE 25.6 Multiple service networks over one transport provider.

traffic network such as the Internet and PSTN, and the transport network such as SONET or WDM for carrying these traffic networks. In essence, a network (or layer) rides on another network, i.e., a traffic network needs a transport network to connect the links needed for the traffic network; then, within the transport network, multilayers are possible due to different data rates. From a service point of view, a user of a traffic network does not "see" the dependency on the transport network.

We will now illustrate a simple network example to illustrate the distinction between different layers in a network topological architecture and highlight the relationship. Consider a four-node network environment for an IP network within an administrative domain. For this network, we have four routers that are connected as shown in Figure 25.7 (top); links



(a) Logical view of IP network links



(b) Transport view through WDM nodes

FIGURE 25.7 Trunking view (IP or PSTN) and transport network view.

(trunks) have the capacity to carry the traffic, possibly with mixed capacity, T1, T3, or OC-3. Note that links are logical in the traffic network (IP network in this case).

We now need the aid of a transport network to route these logical links and the associated capacity (see Figure 25.7, bottom). For example, the link capacity unit for the logical link, f, between nodes 1 and 3, in an IP network is connected using the transport network route 1-2-3; similarly, the demand unit for logical link 1-4, between nodes 1 and 4 in the traffic network, is connected via the transport route 1-2-3-4.

Based on mapping between just two layers in the network hierarchy, an important picture emerges. For example, in the IP network, we see three node-diverse and link-wise logically diverse routes between nodes 1 and 4; they are 1-4, 1-2-4, and 1-3-4. By *diverse* we mean there is no common link (in the logical view) from one route to another. In reality, the actual topology view can be different at a different layer. This is shown at the bottom of Figure 25.7 where we see that the logical links are actually all routed on the same transport network path,

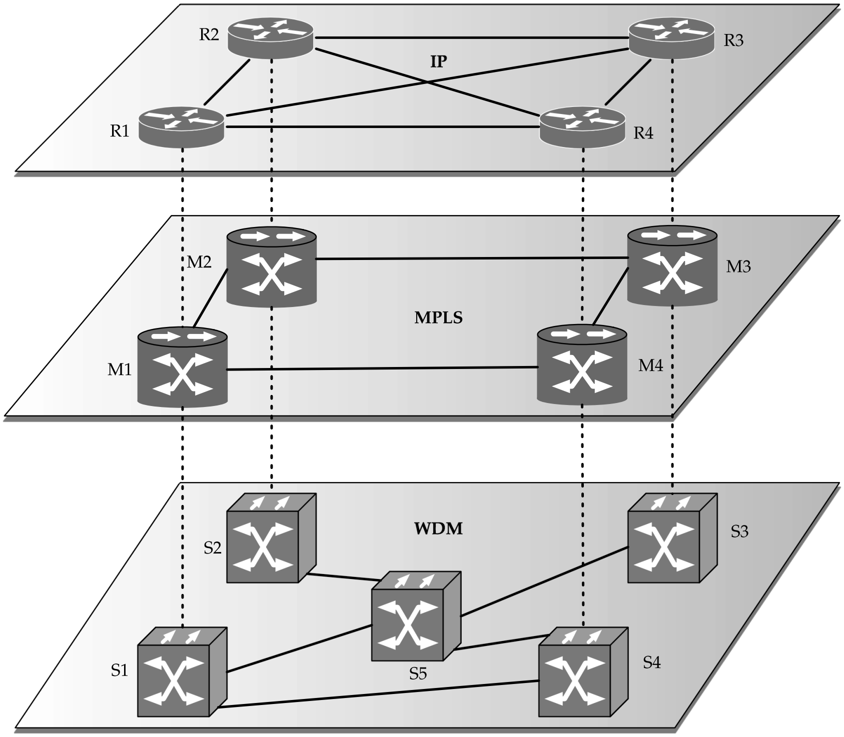


FIGURE 25.8 IP over MPLS over WDM: a three-layer architectural view.

i.e., there is no diversity. Thus, a network may look logically diverse in one layer but may not be diverse in another layer; this also has implications in protection and restoration design (*network robustness*) due to the interrelation between layers. Thus, multilayer network design is an important problem to consider. For instance, it needs to address which layer would be responsible for restoration. There are speed issues, which can affect any coordinated effort. For example, if upper layer takes time to converge, and the lower can do it in less than a sec, then the upper layer may not perceive it. Thus, we can see that coordination between layers is an important issue to understand to avoid undesirable behavior when both layers try to solve the restoration problem at the same time; for additional details, see [564].

As pointed out earlier, there are different traffic networks possible, e.g., Internet, PSTN. Also, service networks such as VPNs can also be considered along with the traffic networks over transport networks. However, there can be multiple transport functionalities, one stacked over another. For example, an MPLS network can be a transport network for IP; in turn, the MPLS network can use a WDM network for transport. These may be stacked in a physical network architecture. Thus, from a network architectural view, a simple picture to consider is an IP or telephone network at the top layer; this uses a first-layer transport network such as MPLS, which, in turn, uses an optical network; in our illustration, we show IP over MPLS over WDM (Figure 25.8).

25.3.2 IP Over SONET: Combined Two-Layer Routing Design

We have discussed so far why the communication network infrastructure is inherently multilayered and how different layers of network resources are related, either in a traffic-over-transport or in a transport-over-transport manner. In this section, we will discuss a two-layer routing design problem for a network consisting of the traffic (IP) and the transport (SONET) layer. As you will see, the routing and capacity design gets intertwined in a multilayer framework.

Recall that in Chapter 7 we discussed IP traffic engineering; in doing so, we have shown how IP traffic flows depend on the link weight (metric) system with protocols such as OSPF or IS-IS that use the shortest paths for routing data packets. In Section 25.1.3, we considered another technology, SONET/SDH, for the transport network with DXC capabilities. Consider now an IP network and suppose that the IP links connecting IP routers need to be physically realized as transmission paths in a SONET network using DXCs. Thus, we have the IP-over-SONET network with a two-layer resource hierarchy, using PoS technology. A pictorial view of this hierarchy is shown in Figure 25.9. Then, the two-layer routing design question we want to address is as follows: given an IP intradomain network and the fact that the IP links are realized as transmission paths over a capacitated SONET network, how do we determine the capacity required for the IP links and the routing of these links in the SONET network in an integrated manner to meet a traffic engineering goal?

Such a two-layer integrated design is often possible only for network providers who own both the IP network (upper layer) and the SONET network (lower layer). Therefore, we assume that this is the case and that the capacity in the SONET network is given (and hence limited). Now, for the IP network, we need to determine the IP link capacity given that (packet) flow allocation is driven by the shortest-path routing. Suppose that we are given the

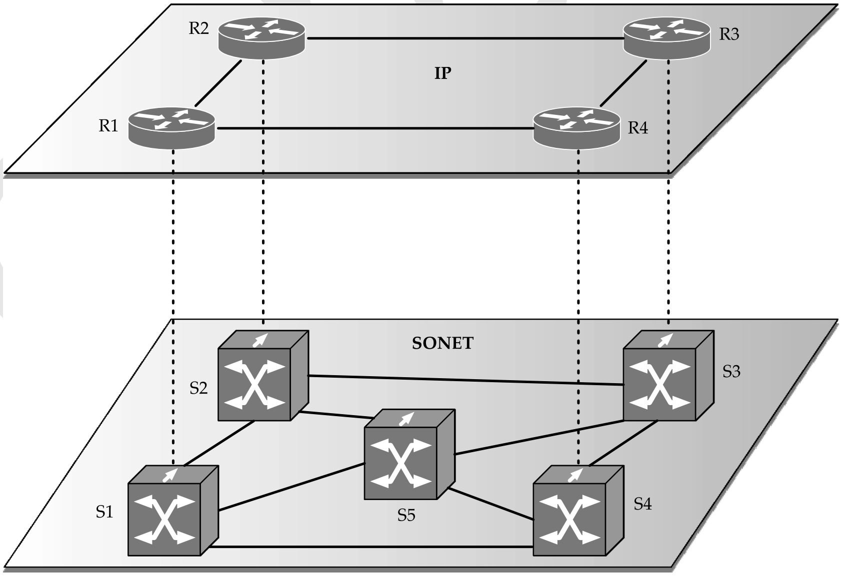


FIGURE 25.9 IP over SONET: two-layer architecture.

demand volume for the IP network (in Mbps) between different routers. We will be introducing two terms: demand volume unit (DVU) and link capacity unit (LCU). Suppose also that we use OC-3 interface cards to connect the routers; this means that IP links are modular with a speed equal to 155.52 Mbps, and the LCU of IP links is then 155.52 Mbps. If one DVU in the IP layer is equal to 1 Mbps, then the IP link module value is given as $M = 155.52$ Mbps. Now, the capacity of the IP links becomes demand volumes for the SONET layer, implying that one DVU in the lower layer is equal to one OC-3. This demand is then routed over the lower layer network using high-speed SONET transmission links such as OC-48 (or OC-192); this in turn implies that one LCU of the lower layer links is equal to $N = 16M$ because one OC-48 ($= 2{,}488.32$ Mbps) system can carry 16 OC-3 modules. Finally, observe that the capacity of an IP link is routed (realized) on a path traversing a series of intermediate DXC nodes between the end DXCs connected to the end IP routers of the considered IP link.

To summarize, the DVU for IP demands is equal to 1 Mbps, and the LCU for IP links is equal to $M = 155.52$ Mbps. The LCU from the IP network becomes the DVU for the SONET network in the two-layer architecture, i.e., DVUs for the SONET network can be thought of as OC-3s. We assume that the link capacity in the SONET network is given in multiples of OC-3s, namely, in OC-48s. Then the LCU for the SONET network links is equal to OC-48 with modularity $N = 2{,}488.32$ Mbps.

Formally, we denote the IP network traffic demand volume as h_k for demand k, $k = 1, 2, \ldots, K$. The flow on an allowable path, p, for demand k in the IP layer that is induced by the link weight (metric) system, $\boldsymbol{w} = (w_1, w_2, \ldots, w_L)$, is given by $x_{kp}(\boldsymbol{w})$, as we discussed in Chapter 7 for IP traffic engineering modeling. Here, we are interested in the IP routing *and* capacity design, subject to capacity limitations in the SONET transport layer. We use $\delta_{kp\ell} = 1$ to indicate path p for demand k if the IP network uses link ℓ ($\delta_{kp\ell} = 0$, otherwise). Then if we write the modular capacity (to be determined) on IP layer link ℓ as y_ℓ (expressed in modules M), we can see that this new demand volume, y_ℓ, induced in the upper layer would need to be routed on the SONET network. In the SONET, we will use the variable $z_{\ell q}$ to route demand volume, y_ℓ, for upper layer link ℓ on a candidate path $q = 1, 2, \ldots, Q_\ell$ in the SONET network. It is important to make a distinction between routing in the two considered layers. Routing in the IP layer is at the packet level and generates the aggregated packet flows, while routing in the SONET network is at the SONET frame level and is set up on a permanent or semi-permanent basis by setting up connection paths of OC-48 modules switched in the DXCs along the path. Note that analogous to $\delta_{kp\ell}$, we need to use another indicator to map the SONET links onto the SONET paths realizing the IP links. The candidate paths in the SONET layer for IP link ℓ would be denoted by index q, here $q = 1, 2, \ldots, Q_\ell$. Then, $\gamma_{g\ell q}$ takes a value of 1 if path q on the transport layer for demand ℓ uses link g, and 0 otherwise. Finally, we denote the capacity of link g in the SONET network by c_g expressed in OC-48 modules denoted by N.

Assume that the routing cost in the IP network is ξ_{kp} on path p for demand k; similarly, in the SONET network, we incur a cost of $\zeta_{\ell q}$ to carry demand y_ℓ on path q for demand ℓ. Then, the traffic engineering design problem can be written as follows:

$$
\begin{array}{lll}
\textbf{\textit{minimize}}_{\boldsymbol{w},\boldsymbol{y},\boldsymbol{z}} & \displaystyle\sum_{k=1}^{K}\sum_{p=1}^{P_k}\xi_{kp}x_{kp}(\boldsymbol{w}) + \sum_{\ell=1}^{L}\sum_{q=1}^{Q_\ell}\zeta_{\ell q}z_{\ell q} & \\
\textbf{\textit{subject to}} & \displaystyle\sum_{p=1}^{P_k}x_{kp}(\boldsymbol{w}) = h_k, & k = 1,2,\dots,D, \\
& \displaystyle\sum_{k=1}^{D}\sum_{p=1}^{P_k}\delta_{kp\ell}x_{kp}(\boldsymbol{w}) \le \rho M y_\ell, & \ell = 1,2,\dots,L, \\
& \displaystyle\sum_{q=1}^{Q_\ell}z_{\ell q} = y_\ell, & \ell = 1,2,\dots,L, \\
& \displaystyle\sum_{\ell=1}^{L}M\sum_{q=1}^{Q_\ell}\gamma_{g\ell q}z_{\ell q} \le Nc_g, & g = 1,2,\dots,G, \\
& w_\ell \text{ nonnegative integer} & \\
& y_\ell, z_{\ell q} \text{ nonnegative integer.} &
\end{array}
\qquad (25.3.1)
$$

Note that other factors in the objective function can be incorporated as well (refer to Chapter 7). In the above, we can see that capacity, y_ℓ, of IP layer link ℓ becomes the demand volume for the lower layer and needs to be routed on the paths in the SONET network. Note that there is a coefficient, ρ $(0 < \rho < 1)$, called the *link utilization coefficient*, used in the upper layer link capacity constraints that can be used for limiting IP link congestion. There are two cost components. The first is the routing cost in the IP layer, and the second cost component is the routing cost in the SONET layer. The second component can be used to model various situations. For instance, if we assume $\zeta_{\ell q} \equiv 1$, then we are in fact maximizing the spare capacity on the SONET links. Another example is when $\zeta_{\ell q} = \zeta_\ell, q = 1,2,\dots,Q_\ell$; then we can interpret ζ_ℓ as the cost rate (e.g., monthly or yearly cost) of one LCU of the IP link ℓ to be paid by the IP provider to the SONET network provider for carrying the IP link capacities. Additional discussion on multilayer design can be found in [564].

25.4 Overlay Networks and Overlay Routing

In recent years, overlay networks and overlay routing have received considerable attention. From our discussion so far on multi-layer routing, you can see that the notion of overlay has been around for quite some time. For instance, consider the telephone network over the transport network, or Internet over the transport network; we can say that any such “service" network is also an overlay network over the telecommunication transport network. Understanding the interaction of such overlay networks over the telecommunications transport network has been studied for quite some time. One of the key issues to understand is how a failure in the underlying transport network, for example, due to a fiber cut, can impact rerouting in the service network [174], [241], [262], [468], [473], [474], [475], [498], [723], [761]. Any such routing decision also needs to consider shared risk link groups, both in terms of reaction after a failure and also to do preplanning during route provisioning through diversity or capacity expansion. For instance, consider Figure 25.8 in which MPL links M1-M2 and M1-M3 would likely to be routed on WDM routes S1-S5-S2 and S1-S5-S3, respectively; here, link S1-S5 falls into the shared risk link group category since the failure of this link will affect multiple MPLS network links; in fact, it would isolate MPLS routers M2 and M3, and thereby would isolate corresponding IP routers. Thus, to protect against such situations, the WDM network should provide diversity by adding, say link S3-S4 (not shown in figure).

The overlay concept is, however, not limited to just two layers. Consider the three-layer network architecture such as IP over MPLS over WDM. In this case, the MPLS network is an overlay over the WDM network while it is, in turn, an underlay to the IP network; in other words, the IP network is an overly over the MPLS network. It is important to recognize that each such network can employ routing within its own context; typically, however, the time granularity of routing decision in each such network could be on different time scales. Regardless, when a failure occurs, each such network might decide to react based on its own knowledge, which could lead to instability in the overall infrastructure; this point was highlighted in Section 19.3. As of now, there is very little protocol-level coordination between networks in different layers to deploy an orchestrated recovery for overall benefit.

More recent usage of overlay networking is in the context of a virtualized network on top of the Internet. In this case, nodes can be set up that act as overlay network routing nodes, where a logical path is set up between any two such nodes over the Internet, for example, using a TCP session. To convey this picture, consider Figure 25.9, but this time imagine the lower layer network to be IP (instead of WDM), and the upper network to be an overlay network (instead of IP). That is, the nodes on the upper plane will be routing nodes for the overlay network. For example, logical virtual link R2-R3 could take the path, R2-S2-S5-S3-R3, in one instance, or the path, R2-S2-S3-R3, in another instance due to change in routing in the underlay IP network. Thus, from the perspective of the overlay network, an estimate on logical link bandwidth would need be assessed frequently, so that the information is as accurate as possible in the absence of specifics about the underlying topology; this would then be useful for the benefit of services that use the overlay network [767]. Similarly, the delay estimate might be necessary to know for some applications that use the overlay network. To even out unusual fluctuations, it might be useful to smooth the available bandwidth or the delay estimate using the exponential weighted moving average method (see Appendix B.6). Such smoothed estimates can be periodically communicated between overlay network nodes using a customized link state protocol so that all nodes have a reasonably accurate view. In turn, based on the information obtained by overlay network nodes, a routing decision for services that use the overlay network would need to be considered. This would depend on the scope of the service, though. If, for example, a service requires bandwidth guarantee, then a QoS routing based approach can be employed (refer to Chapter 17), which may involve alternate routing through overlay network nodes; in this case, a performance measure such as the bandwidth denial ratio would be important to consider. If, however, services that use such an overlay network requires only a soft guarantee, then performance measures other than bandwidth denial ratio, such as throughput, would be necessary to consider [767]. In addition, understanding the interaction between overlay and underlay in terms of routing and the impact on performance is an important problem to consider [414], [626].

25.5 Summary

In this chapter, we covered two topical areas: optical networking and multilayer networking. For optical networking, there are two main classes of problems: SONET/SDH routing and

WDM routing. We discussed how these are transport network routing problems. We also pointed out that on-demand WDM routing is closer to a dynamic call routing problem.

We then discussed multilayer networking, presenting the overall architectural view in order to see how routing fits in. It may be noted that routing and capacity design are intertwined in a multilayer setting. That is, an upper layer's capacity becomes demand volume for a lower layer. Thus, if the capacity assignment can be dynamically configurable, it has many implications for network and system stability.

It may be noted that multilayer routing requires common addressing schemer for nodes, or else a mechanism so that information can be exchanged from one layer to another layer. Furthermore, a coordinated network management system is required to exchange such information [472].

Further Lookup

Historically, the first important instance of multilayer networking goes back to the development of the circuit-switched voice network as the traffic network, and the transmission system (for circuit routing of the link capacity, i.e., trunk groups, for circuit-switched voice) with rates such as T1 and T3 as the transport network, thus forming a traffic transport layering architecture. That is, in summary, this combination of circuit-switched voice traffic networks over transport networks is the first example of multilayered networks.

While this relationship has been known and has been in use for several decades [582], [583], [584], [596], [742], integrated network modeling and design considering both of these networks together was not considered initially. In earnest, it can be said that the need was not as great when the transmission system was made of co-axial cables, which is inherently physically diverse. The need became much more pronounced when the transmission network started to move from the PDH systems based on co-axial cables to fiber-based SDH/SONET systems in the late 1980s. The immediate effect was that the transmission network became sparse, with links composed of fibers of enormous capacity, capable of carrying many trunk groups between distant switching nodes. The downside of this was that a single fiber cut could affect multiple trunk groups in the circuit-switched voice networks. With the advent of IP networks, the same issues have come up over the past decade. Thus, this area has seen tremendous interest, starting in the early 1990s. Thus, for the area of multilayer routing and design, we refer you a sampling of collections: [3], [31], [32], [174], [184], [185], [187], [241], [254], [267], [268], [420], [383], [465], [467], [468], [472],[473], [475], [511], [512], [633], [723], [761].

Optical networking, particularly routing, has been an active area of research in the past decade. Accordingly, the literature is vast. There are excellent books on optical networking such as [509], [580]. A framework for IP-over-optical networks is described in RFC 3717 [573]. For discussion related to PPP-over-SONET, see RFC 2615 [440] and RFC 2823 [110]. For a historical view of IP over optical architecture at a tier-1 provider, see [426].

Several heuristic algorithms have been developed to solve the routing and wavelength assignment problem [53], [137], [446], [540], [511]. For a recent survey of various solutions of RWA problem, see [136].

Another stream of problems in optical networks is IP logical topology design and routing at IP layer in an IP-over-WDM networking paradigm; for example, see [54], [55], [196], [511],

[512], [579]. Another important factor in logical topology design is time-varying traffic, as a topology designed for a traffic demand at a certain time might not respond well for traffic matric at another time. For detailed discussion of logical topology reconfiguration, see [3], [54], [246], [576] .

Exercises

25.1. Solve the SONET ring routing problem discussed in Section 25.1.2 in which demand is allowed to be split, but still must be integer valued.

25.2. Explain the relation between routing and capacity in a multilayer setting though a small network topology example.

25.3. Consider the following demand matrix on a four-node ring (Figure 25.1).

node i \ node j	2	3	4
1	12	16	8
2	–	4	2
3	–	–	8

Determine the optimal ring routing if the goal is to balance the ring load.

25.4. Consider Figure 25.8. Determine minimum link connectivity required in the WDM network for protection again any WDM link failure.

25.5. Convert nonlinear Model (25.2.2) to an equivalent model where the constraints are linear.