

POSITION PAPER

Future Nationwide Optical Network Architecture for Higher Availability and Operability Using Transport SDN Technologies

Yoshihiko UEMATSU^{†a)}, Shohei KAMAMURA[†], Hiroki DATE[†], Hiroshi YAMAMOTO[†], Aki FUKUDA[†], Rie HAYASHI[†], *Members, and* Katsutoshi KODA[†], *Nonmember*

SUMMARY An optical transport network is composed of optical transport systems deployed in thousands of office-buildings. As a common infrastructure to accommodate diversified communication services with drastic traffic growth, it is necessary not only to continuously convey the growing traffic but also to achieve high end-to-end communication quality and availability and provide flexible controllability in cooperation with service layer networks. To achieve high-speed and large-capacity transport systems cost-effectively, system configuration, applied devices, and the manufacturing process have recently begun to change, and the cause of failure or performance degradation has become more complex and diversified. The drastic traffic growth and pattern change of service networks increase the frequency and scale of transport-capacity increase and transport-network reconfiguration in cooperation with service networks. Therefore, drastic traffic growth affects both optical-transport-system configuration and its operational cycles. In this paper, we give an overview of the operational problems emerging in current nationwide optical transport networks, and based on trends analysis for system configuration and network-control schemes, we propose a vision of the future nationwide optical-transport-network architecture expressed using five target features.

key words: *optical transport network, common infrastructure, supervisory, failure recovery, deployment, network architecture*

1. Introduction

Along with the spread of high-speed fixed/mobile Internet access and high-performance terminals, people's lives and work have improved at least in terms of efficiency, and communication services have been deeply integrated into human society as lifelines. Their traffic amount will continuously increase as people's lives and work become further dependent on information communications technologies. Communication services will also be diversified, as Internet of Things (IoT) or Internet of Everything technologies enhance the efficiency of many industrial areas and a huge number of devices become attached to communication service networks.

An optical transport network is a common infrastructure to accommodate diversified communication services with drastic traffic growth [1]–[3] and is composed of optical transport systems deployed nationwide in thousands of office-buildings [4]. It is necessary not only to continuously convey increasing traffic but also to achieve high end-to-end communication quality and availability and provide flexible controllability in cooperation with service layer networks

[5]–[7]. High quality and availability is achieved by appropriate resource allocation, redundancy design, rapid failure detection, hitless protection switching, and rapid maintenance for failed function blocks. Controllability in cooperation with service layer networks enables the rapid execution of procedures such as capacity increase, network reconfiguration, and system renovation without service interruption. In this paper, we give an overview of the operational problems that emerge in current nationwide optical transport networks, and based on technology trends analysis, we propose a vision of future nationwide optical-transport-network architecture expressed using five target features.

The structure of this paper is as follows. In Sect. 2, we discuss the operational problems in the current nationwide optical network in terms of its fundamental operational cycles. In Sect. 3, we analyze related technology trends surrounding optical-transport-system configuration and network-control schemes. In Sect. 4, we discuss target features in terms of the fundamental cycles, and propose a vision and architecture of future nationwide optical transport networks in Sect. 5.

2. Operational Problems in Current Nationwide Optical Transport Network

Figure 1 shows an optical transport network supplying transmission path-sets to various service networks, and each service network operator constructs a service network using service systems connected via path-sets. Service systems include IP routers, Ethernet switches, circuit switches, and so on, and service networks include fixed/mobile/WiFi Internet access services, telephone services, leased line or VPN services for business users. Various IoT-device control service networks will also be created along with IoT technology in many industrial areas. As the number and diversity of service networks increase with drastic traffic growth, optical transport networks are required to have larger transmission capacity and provide a larger number of path-sets. To achieve high-speed and large-capacity optical transport systems cost-effectively, system configuration, applied devices and the manufacturing process have begun to change, and the cause of failure or performance degradation tends to be more complex and diversified. Drastic traffic growth and traffic-pattern change of service networks tend to increase the frequency and scale of transmission-path-capacity increase/decrease and path-sets reconfiguration in cooperation

Manuscript received December 28, 2016.

Manuscript revised June 11, 2017.

Manuscript publicized August 8, 2017.

[†]The authors are with NTT Network Service Systems Laboratories, Musashino-shi, 180-8585 Japan.

a) E-mail: uematsu.yoshihiko@lab.ntt.co.jp

DOI: 10.1587/transcom.2016EBN0013

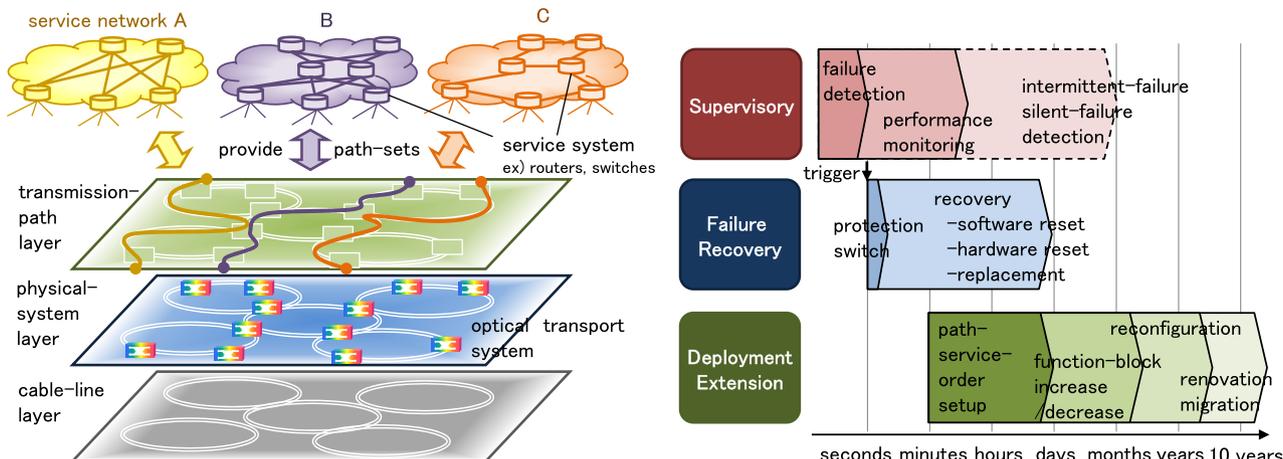


Fig. 1 Role of optical-transport-network infrastructure and fundamental operational cycle.

with service networks.

In this section, we give an overview of transport network problems that emerge in each fundamental operational cycle, i.e., supervisory, failure recovery, and deployment and extension, as shown in Fig. 1. The supervisory cycle includes failure detection, performance monitoring, and intermittent or silent failure detection. The failure-recovery cycle includes working to protection switching and recovery operation of failed functions, such as software reset, hardware reset, or replacement. The deployment-and-extension cycle is composed of path-service-order setup from each service network, function-block increase/decrease, network reconfiguration, system renovation, and migration to new technologies.

We make the following assumptions about the future performance improvement of optical transport systems. The latest optical transport systems achieve optical cross connect (OXC) capability with 100 Gbit/s optical-paths * over 80 wavelength/fiber multiplexing, and achieve Tbit/s electrical path cross connect (EXC) capacity [8]–[11]. However, the traffic volume of mobile and fixed communication in Japan is nearly 10 Tbit/s in 2016 and will continue to increase by several Tbit/s every year [12], [13]. Thus in the backbone network, the future system should handle Tbit/s-class optical paths with corresponding EXC capability, and eventually IP routers and Ethernet switches will also be required to have Tbit/s-class interface cards [14]–[17]. Meanwhile in the aggregation network, the simple and low cost system which may introduce passive sub-lambda multiplexing [18]–[20] should handle 10–100 Gbit/s paths flexibly to deliver/collect the traffic to/from thousands of office-buildings. The size and power consumption of the future systems need to be the same level as those of the current systems, because optical transport networks are fundamentally just a nation-wide simple wiring substrate that provides solid connections anywhere upon request.

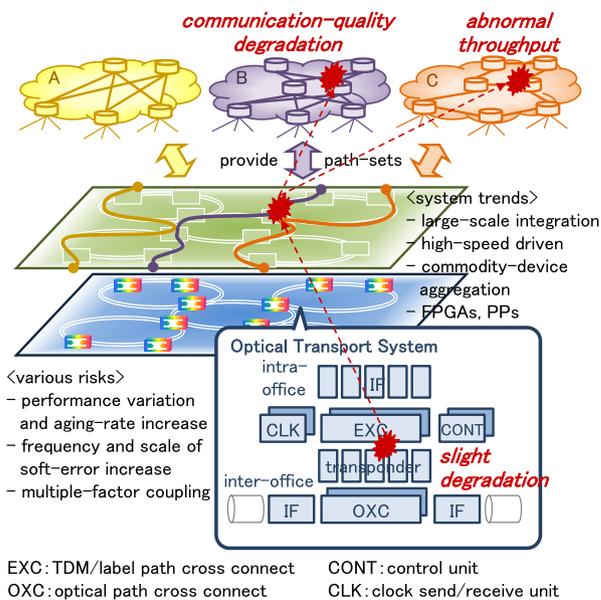


Fig. 2 Problems emerging in supervisory cycle.

2.1 Supervisory Cycle

Figure 2 gives an overview of the problems with the supervisory cycle. To achieve Tbit/s-class transponders, interface cards and corresponding EXCs, the clock rate and parallel-processing degree of hardware should be improved using a finer LSI manufacturing process. Because it is difficult to achieve a clock rate and parallel-processing degree inversely proportional to the process, the driving margin of each device has decreased steadily, and performance variation and aging rate of each device have relatively increased [21], [22]. As many commodity CPUs, chip-sets, or optics used in LANs or data centers are mounted to various function blocks to reduce system costs, the performance, availability, and life-cycle of the total system become dependent on these commodity devices.

To reduce system development cost and risk, pro-

programmable engines such as field-programmable gate arrays (FPGAs) or packet processors (PPs) are introduced into the system. For example, FPGAs are introduced into transponders or EXCs [23]–[26]. PPs are used mainly in EXCs because standardized optical transport network (OTN) multiplexing is consistent with packet multiplexing, and destination search and scheduling based on packet labels in the multi-protocol label switching-transport profile (MPLS-TP) are consistent with packet forwarding [27]–[30]. FPGAs store reconfigurable wiring information and logic circuit information in large-scale and high-speed static random access memories (SRAMs), and PPs introduce SRAMs and content-addressable memories (CAMs) into packet-forwarding tables or schedulers, which require high-speed processing [31]. Generally SRAMs and CAMs are weakly resistant to soft-errors, and if uncorrectable soft-errors occur in SRAMs and CAMs for search or reference, such as forwarding tables or wiring information, system malfunction occurs and continues until the erroneous information is overwritten [32], [33]. According to ITU-T Recommendation K.124 [34], the error rate in the latest FPGAs including SRAMs is larger than 10,000 Failure In Time without mitigation measures. If a function block consists of 2 FPGA chips, its mean time between failure (MTBF) would be 50,000 hours. Although internal error correction functions overwrite some erroneous information, soft-errors will certainly increase the frequency of serious failures.

Therefore system failure or performance degradation is probably increasing gradually along with the increase in performance variation and aging rate of each device, frequency and scale of soft-errors, and multiple-factor coupling. Slight degradation may largely impact on quality of services, and silent or intermittent failure may occur more frequently. Therefore if software-defined networking (SDN) is introduced to optical transport networks in the near future, it needs to reveal performance degradations due to diversified and complex factors and signs of large-scale failure earlier.

2.2 Failure-Recovery Cycle

Communication services that service networks provide are generally required to achieve end-to-end availability of 99.999% (unavailable time of 5 minutes/year) [35]–[37]. This value includes the availability of access networks in which it is difficult to have redundancy, and advanced service control functions processed by applications. So service networks composed of service systems and optical transport systems are supposed to achieve 99.9999% availability (0.5 minutes/year unavailable time) for each end-to-end connection. To achieve this, it is important to first properly make each function redundant and execute protection switching in case of failure to maintain all end-to-end communications. It is also important to repair failed functions and recover redundancy to prevent multiple simultaneous failures and disruption in the communication of each user.

Redundant connections are composed of some

protection-switching segments as shown in Fig. 3(a) to maintain the connectivity even in simultaneous failures on the multiple routes. End-to-end unavailability for a connection composed of k protection switching segments, r redundant routes, and $l_{i,j}$ functions in segment i and route j is expressed as;

$$P = 1 - \prod_{i=1}^k \left(1 - \prod_{j=1}^r (1 - (1 - \alpha)^{l_{i,j}}) \right). \quad (1)$$

$$\alpha = MTTR / (MTBF + MTTR)$$

Individual failures are assumed to occur randomly [38], [39], and due to the failure increasing trends in Sect. 2.1, MTBF is assumed to be 100,000 hours for all the functions, which is the conservative value in previous work [40]. MTTR is the mean time to repair.

We will make some assumptions about the number of cascaded functions for end-to-end connections. As for the backbone network, if optical transport systems are deployed in all the nodes of JPN48 [41] with optical repeaters placed at 80 km intervals in each link, connections between Tokyo and the other cities pass through 15 systems and 15 repeaters at most, except for submarine sections that require special repeaters. Considering that a connection passes through multiple non-redundant functions in a system as shown in Fig. 3(a), a connection across a backbone network is assumed to pass through 45 functions. As for the aggregation network that handles intra-prefecture connections, the necessity to deploy repeaters is low. If 80 buildings are deployed in a prefecture and connected to be a multi-ring structure as shown in Fig. 5, connections between the center and the edge buildings pass through 9 systems at most, which corresponds to 18 functions. The following discussion assumes the number of cascaded functions to be 80 for simplicity. As for $l_{i,j}$, Although P can be reduced as $l_{i,j}$ decreases, allowable $l_{i,j}$ values depend on the physical network topology and the cost to deploy protection switching functions. Thus, $l_{i,j}$ is treated as valuables.

Figure 3 shows an unavailable-time-calculation example, which is P multiplied by $60 \times 24 \times 365$ to show the average minutes per year. (b) shows the results where all $l_{i,j}$ values equal 10 or 20, and (c) shows the results where the $l_{i,1}$ and $l_{i,2}$ values are different, in the case of $r=2$ and $k=8$. The unavailable time tends to increase depending on the extension of failure-recovery time and decrease as r increases. To achieve 99.9999% availability (0.5 minutes/year unavailable time) for end-to-end connections, the failure-recovery time should be less than several hours and the $l_{i,j}$ values should be less than 20 if r equals 2.

To complete the repair operation of failed functions distributed in thousands of buildings in several hours, sufficient maintenance organization should be maintained under severe conditions. Therefore, for minimizing total life-cycle costs, transport SDN needs to optimize the balance of the number of functions in protection switching segments, end-to-end redundancy, and failure-recovery time considering MTBF of each function, and control the related systems properly. If

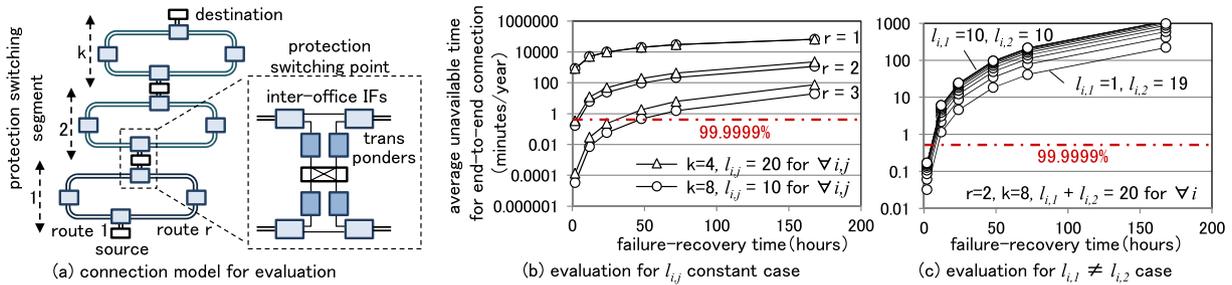


Fig. 3 Problems emerging in failure-recovery cycle.

	procedure	example	conditions
path-service order	-capacity increase /decrease -creation /deletion	POI router, core router, edge router, optical transport network	full-time redundancy
reconfiguration	-system install /remove -topology change -POI change	change in network topology	millisecond-order switching no route frapping
renovation migration	-software update -hardware renewal	EOL, new system	roll-back in case of error

Fig. 4 Fundamental procedures in deployment-and-extension cycle.

the MTBF is shorter than 100,000 hours due to the factors discussed in Sect. 2.1, the availability should be maintained by redundancy gain or severe recovery time. If IoT-device control services require further availability, both redundancy gain and severe recovery time are required.

2.3 Deployment-and-Extension Cycle

Each service-network operator should determine the configuration of his/her own service networks based on the target number of users, target end-to-end communication quality, target network availability, and cost/performance balance of the service systems and provided path-sets. From the viewpoint of long-term and stable service delivery and total cost minimization, service-network operators should deploy and extend their own service networks at appropriate times, considering changes in external conditions, in cooperation with optical transport networks. Figure 4 summarizes the operation procedures in the deployment-and-extension cycle. If the service network requires an increase/decrease in traffic volume between service systems, the optical transport network increases/decreases or creates/deletes the corresponding paths (path-service order). If the service network requires a new system install, topology change, and point-of-interface (POI) change, the optical transport network executes large-scale reconfiguration or reallocation of path-sets (reconfiguration). Moreover, the service and optical trans-

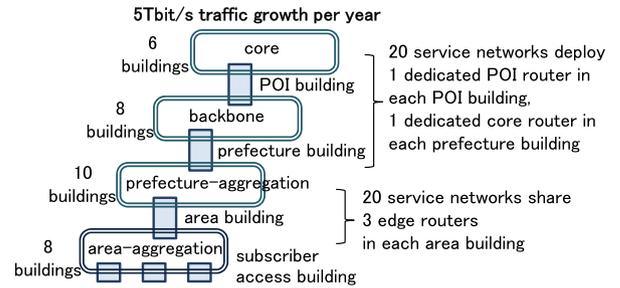


Fig. 5 Nationwide optical transport network model for deployment-and-extension frequency estimation.

port systems have their own lifetimes and should be renewed or migrated to new technologies at their end-of-sale (EOS) or end-of-life (EOL) times (renovation/migration). The renovation/migration procedure also includes system-software updates.

Let us estimate the frequency of the procedures assuming a nationwide optical transport network when the total traffic volume increases 5 Tbit/s per year [12], [13]. The network consists of 4 hierarchical sub-networks as shown in Fig. 5: an area-aggregation network including 8 subscriber-access buildings, a prefecture-aggregation network including 10 area buildings (each of which represents each area-aggregation network), a backbone network including 8 prefecture buildings, and a core network including 6 POI buildings. The number of buildings is $8 * 10 * 8 * 6 = 3840$. For simplicity, 20 service networks deploy one dedicated POI router in each POI building, deploy one dedicated core router in each prefecture building, and share 3 edge routers deployed in each area building for cost reduction. So the number of routers is

$6 * 20 + 48 * 20 + 480 * 3 = 2520$ and the number of optical transport systems is 3840. We assume traffic matrix to be only between 6 POI buildings and 3840 subscriber-access buildings equally, the path topology of the core network to be mesh, and those for the other sub-networks to be star.

(path-service order) According to the orders from service networks, the optical transport network increases the paths in 10 or 100 Gbit/s units and allocates them to service networks. Table 1 shows the amount of capacity increase of a source-destination pair, assumed path-increase unit, and the number of path-increase events for each sub-network in the case of all the paths duplexed. For example, the path-

Table 1 10/100G path increase frequency for sub-networks

sub-network	no. of sub-networks	path topology	no. of s-d pairs/sub-network	capacity increase/year for s-d pair	path increase unit	increase event /year
core	1	mesh	15	139G	100G	42
backbone	6	star	7	104G	100G	88
prefecture	48	star	9	10.4G	10G	900
area	480	star	7	1.3G	10G	875

increase events for backbone network are calculated as

$$6 * 7 * 104 \text{ Gbit/s} / 100 \text{ Gbit/s} * 2 = 88 \text{ events/year.}$$

Accordingly, the 10/100 Gbit/s-based path-increase frequency is 1905 events/year for the nation-wide network, and the frequency increases drastically if we consider the larger volume traffic growth, or the division and re-allocation processes of 10/100 Gbit/s paths to individual service networks.

(reconfiguration) Because the traffic volume of each service network increases 250 Gbit/s per year, additional routers should be installed, cut-through paths between routers established, or new POIs deployed in accordance with the changes in traffic matrix or service conditions. If a service network operator executes the events three times a year, the reconfiguration frequency for the nation-wide network is 60 events/year. This estimation assumes the yearly or quarterly network planning. If in-operation planning [42]–[44], which reconfigures each service network dynamically and in real time, is required for further cost reduction, the frequency will increase drastically.

(renovation) The operating software of routers tends to be released yearly to supply new functions and fix the bugs of its previous version [45], [46]. Meanwhile, the hardware of routers released in the early 2000s has passed its EOS/EOL time [47], [48]. If the software up-date interval of routers is 5 years and the EOS/EOL interval of routers, as well as optical transport systems, is 15 years, the renovation frequency for the nation-wide network is $(3840 * 1 + 2520 * 3) / 15 = 760$ events/year.

If the introduction of commodity devices discussed in Sect. 2.1 shortens the EOS/EOL cycle of the systems, or if some service networks require service-specific edge routers, the frequency will increase drastically.

Service interruption or quality degradation is not acceptable in the above procedures because they are circumstances of service-network operators and optical-transport-network operators. Therefore, they should be in compliance under severe execution conditions; maintaining redundancy during the entire process, millisecond-level switching time, IP/Ethernet route flap suppression in each service network, and roll-back in case of error. Enormous verification cost, deployment cost and time are required to achieve nation-wide deployment under such severe execution conditions.

The external conditions in each service network have recently been changing more rapidly. It is likely that various service networks will need to execute the procedures under their own conditions and times. If each procedure execution is delayed due to the enormous cost and time,

the overall network configuration or resource allocation is likely to deviate from the optimized solution. A new optical transport network control architecture in cooperation with service networks using transport SDN is required to execute the procedures rapidly without service interruption and in accordance with the various external conditions.

3. Related Systems and Control-Technology Trends

To develop larger-capacity optical transport systems, key devices should be manufactured using finer LSI processes. The performance variation and aging rate of each device will further increase due to various causes [49]–[52], leading to an increase in system failures or degradations without more severe device screening. Along with the larger scale of systems, more commodity devices will be introduced into the system, such as commodity CPUs and chipsets mounted on every function block and all blocks connected via Ethernet LANs in the system. Along with a finer LSI-manufacturing process, enormous cost will be incurred to prepare an application-specific-integration-circuit (ASIC) development environment, and it will be easier to develop high-speed key devices in an FPGA- or PP-based environment. The frequency and scale of soft-errors will further increase with the finer manufacturing process and the number of FPGA- or PP-based key devices will increase [53]. Therefore the failure and performance-degradation problems discussed in Sect. 2.1 become more severe. Although a device-level fail-safe is of course necessary, system- or network-level supervision and fail-safe will also be important.

To enable rapid redundancy recovery (Sect. 2.2), network reconfiguration and system renovation (Sect. 2.3), flexible cross-layer networking between service networks and an optical transport network is necessary. In generalized multi-protocol label switching (GMPLS), cross-layer control technologies for optical paths and TDM/label paths have been studied and standardized [54], [55]. Flexible path allocation design and control has also been studied based on the recent discussion of SDN for optical transport networks [56]–[58]. These studies can be classified as connectivity virtualization or path virtualization. Regarding function virtualization, network functions enabled by software on common computer hardware is being studied based on the discussion of network function virtualization (NFV) [59]. These virtualized functions can be easily migrated to other hardware, assisted by layer 2 or 3 connectivity change. Although discussions on NFV now focus on higher-layer functions, discussion on the possibility of common hardware for functions under layer 3 have begun, such as large-scale IP routers composed of commodity switches and controllers [60] or open hardware architecture for optical transport systems [61]. Furthermore, advanced research and development have begun, such as resource portability and migration for IP routers, or N+1 protection switching among transponders in optical transport systems [62]–[65]. To solve the issues introduced in Sects. 2.2–2.3, it is necessary to extend and integrate the

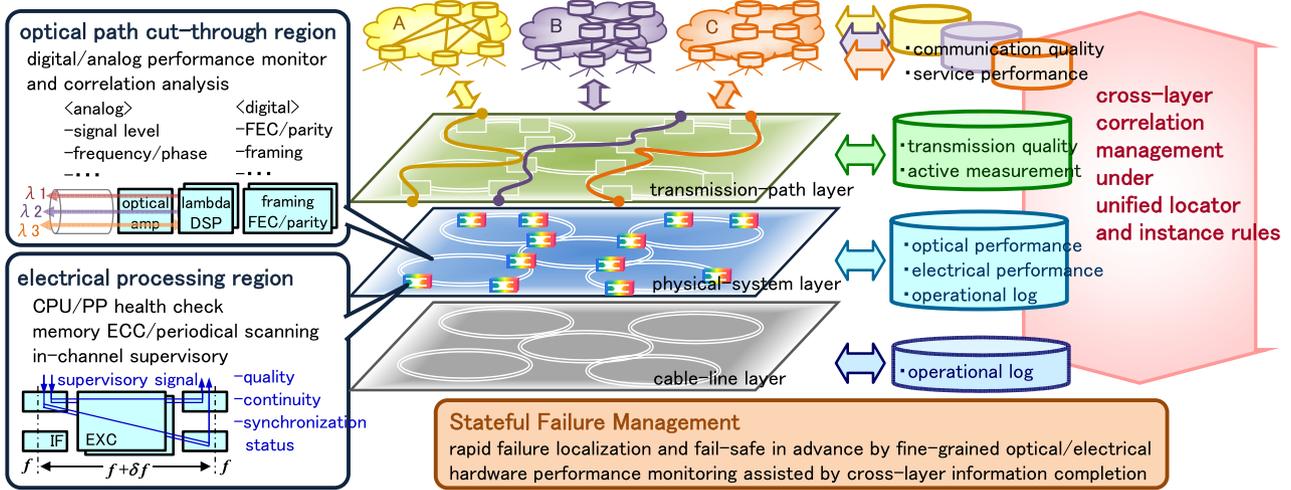


Fig. 6 Target features for supervisory cycle.

above-mentioned technologies and properly use them according to the situation of service and optical transport networks.

4. Target Features

In this section, we describe the target features of an optical transport network and systems based on current operational problems and technology trends. The fundamental direction is that the software-based cross-layer supervisory and control mechanisms clarify each hardware status and operate it in accordance with the overall situation to achieve high end-to-end quality and availability, and high interoperability with service networks. Hereafter an optical transport network includes not only an inter-office transmission-path network, but also an intra-office wiring network that serves connections between service systems or those between service systems and optical transport systems.

4.1 Supervisory Cycle

To reveal various performance degradations due to diversified and complex factors and detect signs of large-scale failure earlier, end-to-end supervision should be achieved mainly by fine-grained physical system performance monitoring assisted by cross-layer information completion. The implementation of this feature is shown in Fig. 6.

For physical system monitoring of optical path cut-through regions, OTN termination points store standardized digital performance attributes [66], such as frame-alignment status, forward error correction (FEC) statistics, and bit interleaved parity (BIP) statistics. Intermediate components including optical amplifiers and wavelength switch store analog information fluctuations such as input/output level, phase, frequency of optical signals, temperature, and power consumption. When detecting a slight degradation at OTN termination points, the physical system collects the intermediate analog information and analyzes the correlation be-

tween the stored digital and analog information for rapid failure localization [67], [68]. For electrical processing region such as electrical cross connect or control unit, it is necessary to check the health of CPUs and PPs and periodically scan high-speed SRAMs and CAMs that are weakly resistant to soft-errors and have poor error correction functions. Furthermore, the in-channel supervisory scheme is quite efficient for failure localization. In cross-connect functions for example, the hardware-driving clock rate is set slightly higher than that required to forward entire client signals, and in-channel supervisory signals are inserted periodically to continuously monitor port-to-port communication quality, or synchronization status between working/protection entities [69]. If a failure or degradation is localized, the system or maintenance staff should rapidly scrub the failed information and reboot or replace the degraded function block in advance.

Moreover, the transmission-path layer or service network may first detect degradations in physical systems, because they include degradations below the alarm thresholds, uncorrectable soft-errors, and so on. Therefore, the cross-layer supervisory mechanism manages the performance attributes of individual layers under a unified locator and instance rules, including service performance and communication quality in service networks, transmission-path quality and active-measurement in the transmission-path layer, operational logs in the cable-line layer, and physical-layer status discussed above. If this mechanism detects degradation at one layer, it immediately collects performance information from all the function blocks along the suspected route and analyzes their correlation for rapid localization. Multi-lateral and long-term correlation analysis is possible because the rich CPUs and memories mounted on each function block can store a large amount of information. This leads to understanding end-to-end degradation statuses and failure signs without leakage. Rapid failure localization and fail-safe in advance by fine-grained optical/electrical hardware performance monitoring assisted by cross-layer information

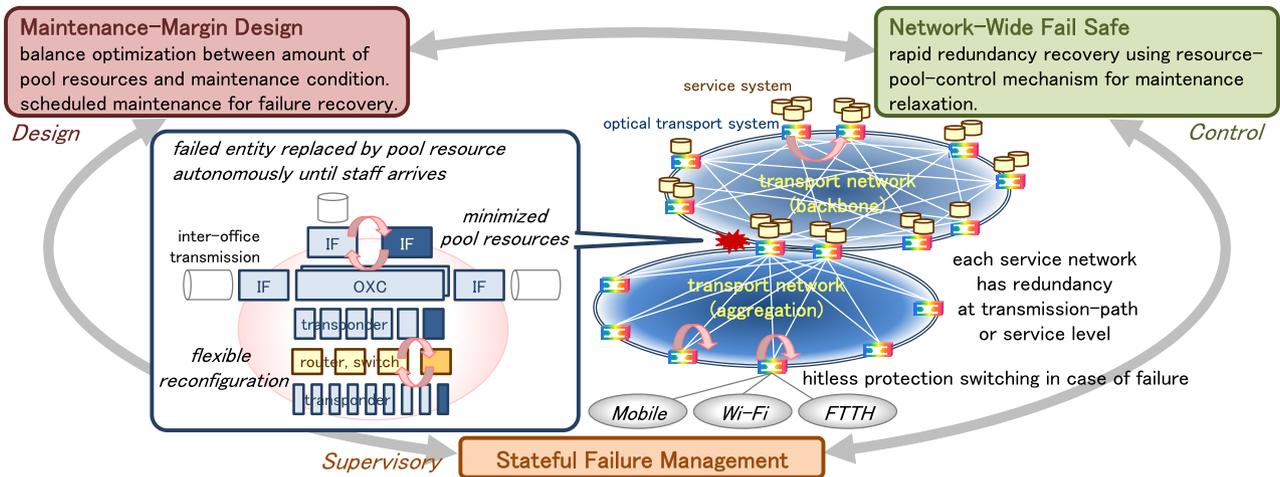


Fig. 7 Target features for failure-recovery cycle.

completion is the concept called “**Stateful Failure Management**”.

4.2 Failure-Recovery Cycle

To reduce maintenance organization and relax maintenance conditions while maintaining high end-to-end availability, rapid redundancy recovery or maintaining full-time redundancy should be achieved using a transport-resource-pool control mechanism, and the balance between redundancy and maintenance conditions should be optimized continuously. The implementation of this feature is shown in Fig. 7.

To minimize the impact of each physical failure on service performance or quality, each service network should be redundant on the transmission-path level or service-connection level. In case of failure, all communications should be maintained by hitless protection switching of transmission paths or address-based rerouting in the IP/Ethernet layer. Furthermore, it is effective to thinly and widely deploy pool resources to the overall network, recover redundancy by replacing failed function and pool resource autonomously, and avoid long-term network running without redundancy, as in the left balloon in Fig. 7. This is the concept called “**Network-Wide Fail Safe**”. Each office building has at least inter-office transmission interfaces corresponding to the number of inter-office fiber routes, transponders corresponding to the number of wavelength paths terminated at the building, and packet switches and routers corresponding to the number of services. Therefore, the amount of pool resources should be reduced by deploying minimized pool resources against the same type of function group. The replacement control occurs only within each building, so it is unnecessary to assign additional-wavelength resources in the inter-office transmission line. As for rapid redundancy recovery, many transmission-path-based restoration schemes have been proposed for improving dual failure restorability [70]–[72]. Our proposal is a function-based restoration, which achieves flexible and robust redundancy recovery regardless of failure locations. Both schemes will be used in

combination depending on the end-to-end network configuration.

Although failed function repair or redundancy recovery is currently the task of maintenance staff, we argue that pool resources and their control mechanisms should temporarily act on behalf of the manual repair or recovery process. From the viewpoint of rapid redundancy recovery or control on behalf of the manual process, immediacy or hitless is unnecessary for replacement operation. Steady replacement control, even if it takes a few minutes, drastically eliminates the risk of multiple simultaneous failures and disruption in user communication. This leads to expanding the degree of freedom in the maintenance time and locations of maintenance staff. Therefore, in the resource-pool-control mechanism discussed above, it is important to establish a resource-pool design scheme, which optimizes the balance between pool-resource amount and maintenance conditions in view of the end-to-end network structure or MTBF of each function, and finally aims at a full-periodic maintenance scheme including the failure-recovery process. This is the concept called “**Maintenance-Margin Design**”.

The control and design schemes for optical transport networks discussed above can be commonly used in backbone and aggregation areas. However, in aggregation networks distributed widely and close to users, it is first necessary to minimize failure and recovery frequency by simplifying network systems and thoroughly reducing the number of active devices [18]–[20]. By deploying pool resources just for the remaining failure-frequent functions, full-time redundancy will be achieved, independent of fiber-route topology or failure locations.

We previously proposed a fundamental resource-pool control scheme and balance design scheme between pool-resource amount and maintenance conditions [73]. We showed that thinly and widely deployed pool resources can relax maintenance conditions and drastically reduce costs associated with maintenance. Figure 8 shows a system composed of the number of working entities in function k , nk , and the number of pool entities of function k deployed as pool

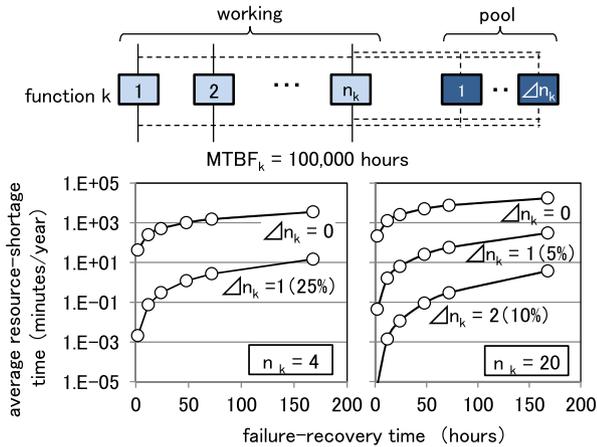


Fig. 8 Evaluation example of balance between pool-resource amount and maintenance conditions.

resources, Δnk . It also shows the dependence of resource-shortage time of function k on Δnk and maintenance condition (i.e. failure-recovery time). **Appendix A** shows a detailed calculation scheme. Each % value after Δnk indicates the total percentage of resource increase by adding pool resources. The resource-shortage time when $\Delta nk = 0$ and failure recovery time = 2 hours is larger than that when $\Delta nk = 1$ and failure-recovery time = 72 hours, regardless of nk . Therefore, a small amount of pool resources can drastically relax maintenance conditions.

Figure 9 shows a maintenance scheme and estimated number of maintenance staff members required for failure-recovery times of 2 and 72 hours. Maintenance staff is normally engaged in another job and address a failure with the highest priority when it occurs. Assuming that on-site failure-handling time is generally 1.5 hours, when the recovery time is 2 hours (a), dedicated staff is required for each building and the staff should immediately address failures if they occur. When the recovery time is 72 hours (b), it is possible for staff to avoid mid-night and weekend work and to address failures of multiple distributed buildings. If 10 hours are set aside for travel time, the remaining 60 hours can be used as a margin for actual execution-time planning. Figure 9 also shows the average time for which unhandled failure occurs due to there being no staff; (a) and (b). **Appendix B** shows a detailed calculation scheme. It is assumed that the number of working/pool functions in a building = 100/10 and the number of buildings that staff can reach within the traveling time is 30. The staff-unavailable time for (a) with 60 staff members assigned is equivalent to the time for (b) with 14 staff members assigned. Furthermore for (a), it is necessary to keep staff so as to act quickly even at mid-night or weekends, so the staff members needed per week increases to $60 * 7 = 420$ people*day. For (b), there is no need to keep staff at mid-night or weekends, so the number of staff members needed per week is $14 * 1/2 * 5 = 35$ people*day, which is 1/10 or less than that for (a).

Although the hardware cost of functions will be reduced using commodity devices, FPGAs and PPs, the fail-

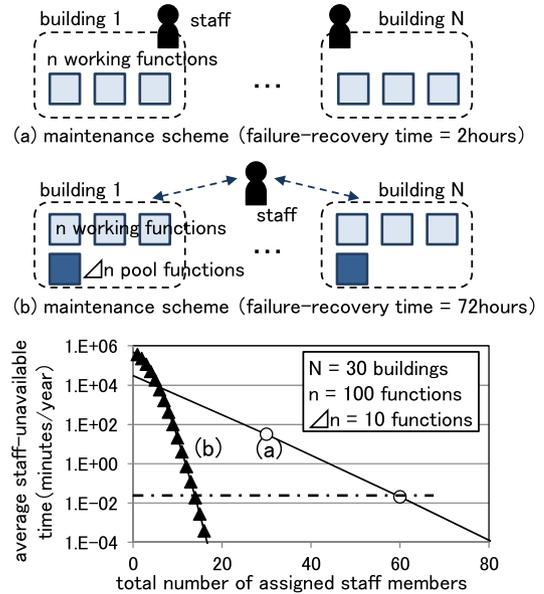


Fig. 9 Evaluation example of number of necessary staff members dependent on maintenance conditions.

ure frequency tends to increase continuously according to Sect. 2.1. On the other hand, thousands of buildings include buildings where staffs reside and buildings difficult to dispatch staffs. To reduce total cost, the availability assignment for the overall network areas and the balance between pool-resource amount and maintenance conditions for each area must be sustainably optimized. For example, it is quite reasonable to deploy minimized pool-resources and recover failures under severe maintenance conditions for the buildings in backbone areas where staff members reside, and to deploy abundant pool-resources and relax maintenance conditions for the buildings in aggregation areas. Furthermore, this resource-pool control scheme has a potential to relax the constraints on the number of functions in protection switching segments discussed in Sect. 2.2, i.e., the constraints on the deployment position of protection switching functions or service systems. This leads to expanding the degree of freedom in the configuration of each service network.

4.3 Deployment and Extension Cycle

To carry out reconfiguration, renovation and migration at appropriate times without risks, automated, flow-through, and service-uninterrupted execution procedures should be established upon the optimized transport-network-control architecture in cooperation with the service networks.

The previous studies about packet-optical integration mainly discuss flow-through multi-layer operation procedures while maintaining the network and system architecture of each layer [74]–[76]. Thus, we first consider the optical-transport and service-network architecture from the viewpoint of easy reconfiguration, renovation, migration, or switching to the pool resources discussed in Sect. 4.2 for IP service networks. Current IP routers mount two types

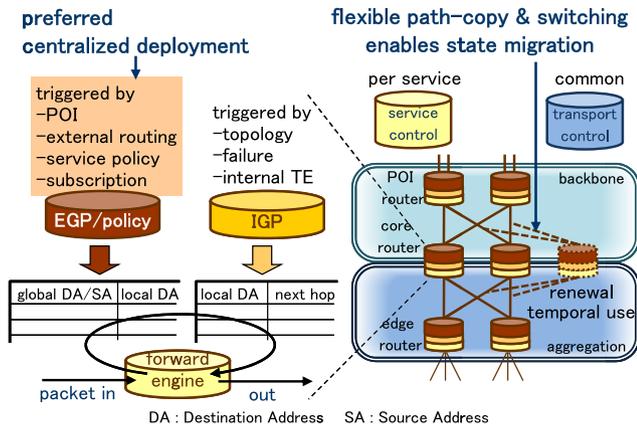


Fig. 10 Service and optical-transport-network-control architecture for easy reconfiguration, renovation, and switching to pool resources.

of autonomous routing functions as shown in Fig. 10: interior gateway protocols (IGPs) that handle only local addresses and exterior gateway protocols (EGPs) that handle global addresses or addresses assigned to end-users. An IGP conducts control triggered by internal topology change or transmission-path failure. An EGP conducts control triggered by the change in POI, received external routing information, service policy, and service-subscription status. For simplicity, we included a policy routing function to EGPs, so the latter function is indicated as EGP/policy in Fig. 10. If protection switching in an IP-based service network is executed by IP rerouting, it is quite natural that an IGP be mounted on each router and high-speed switching be carried out among related routers. For EGP/policy, there are few triggers that require urgency or immediacy, so the degree of freedom for deployment is relatively large. For network-wide consistency of the EGP/policy routing table throughout reconfiguration, it is effective for EGP/policy to be centrally deployed as route reflectors or SDN controllers, calculate master-routing table, and directly set the table to all the routers [77]. Since the IGP mounted on each router can reroute traffic flexibly without service interruption by the appropriate sub-area and cost settings, it does not constrain reconfiguration operation.

Furthermore, if routers hold a simple IGP such as open shortest path first (OSPF), the IGP status of working router can easily duplicated on the secondary router, by copying the signals including OSPF messages from adjacent routers to the working router and inserting them into the secondary router while blocking the signals from the secondary router to the adjacent routers [78]. Therefore, if copying or protection-switching points can be flexibly deployed across an optical transport network, the IGP status of working router can be easily duplicated on routers for renovation or temporal use, and service traffic can be migrated to them without service interruption, IP topology change, and additional interfaces on adjacent routers.

Therefore, the best interworking architecture between optical transport and service networks for easy reconfigura-

tion, renovation, or switching to pool resources is summarized as follows and illustrated on the right-side of Fig. 11;

- EGP/policy function is centrally deployed for network-wide table consistency
- Service systems are simple forwarding engines with simple IGP-based rerouting functions
- Optical transport network is equipped with flexible path-copy and switching functions to assist easy state migration between service systems.

Flow-through and hitless network-control procedures should be executed on this architecture. This is the concept called “**Zero-Touch Configuration**”.

For long term and stable service delivery and total cost minimization, each service-network configuration and overall resource assignment of optical transport networks are optimized sustainably. Based on quality conditions, availability conditions and predicted environmental change of each service and prediction, a true equipment-cost-minimum solution for each service network and overall resource assignment is comprehensively delivered with its alternatives. Execution cost, risk, and time are also calculated to migrate current network configuration/allocation to the candidates, and the next configuration/allocation is determined under the balance of equipment cost of candidates, and execution cost, risk, and term for the migration to candidates. This is the concept called “**Environment-Adaptive Network Design**”.

5. Future Optical-Transport-Network Architecture

Based on the discussions in Sect. 4, we summarize the architectural points of future optical-transport-network infrastructure, as shown in Fig. 12.

First, an optical transport system executes fine-grained physical-layer performance monitoring for rapid failure localization. It analyzes the digital and analog performance correlation for the optical-path cut-through region and executes engine-device health-check, memory scanning, and in-channel supervision for electrical processing region.

Second, an optical transport network including an intra-office wiring network has the capability for flexible path-copy and switching for hitless and flow-through control procedures in cooperation with service networks that include service-network extension/reconfiguration, renovation of both optical transport and service systems, and temporal pool-resource utilization.

Finally, the optical-transport-network control function has the capability of cross-layer supervision and control in cooperation with service networks. It manages and configures overall connections including intra-office connections and is responsible for executing the cooperative control procedures. It also executes failure localization by cross-layer failure and performance correlation analysis and may hold a design optimizing engine for each service-network configuration and resource assignment of the optical transport network.

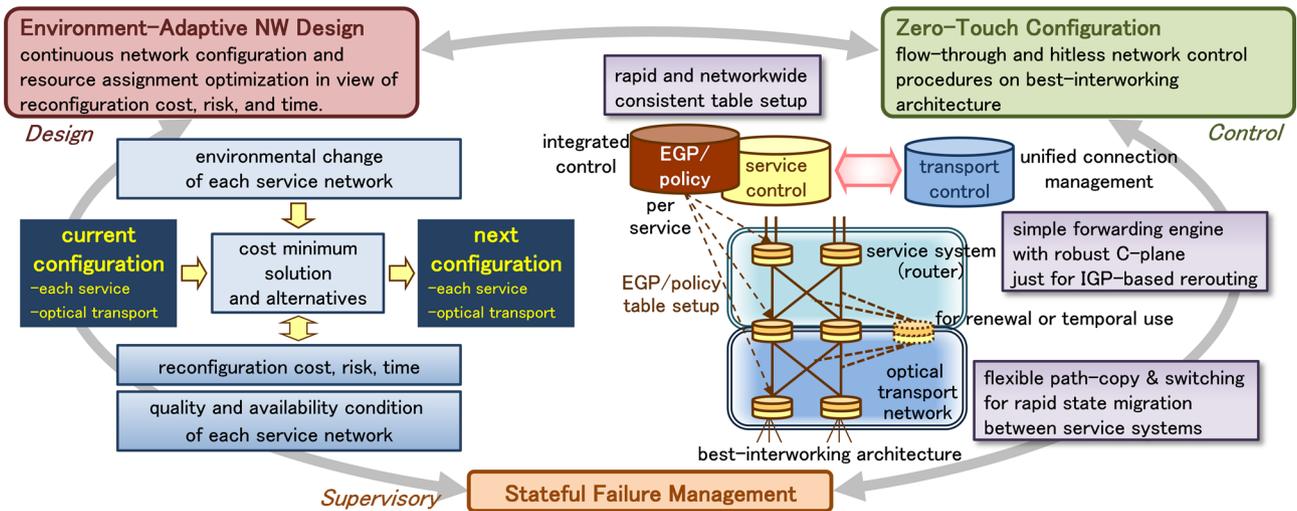


Fig. 11 Target features for deployment-and-extension cycle.

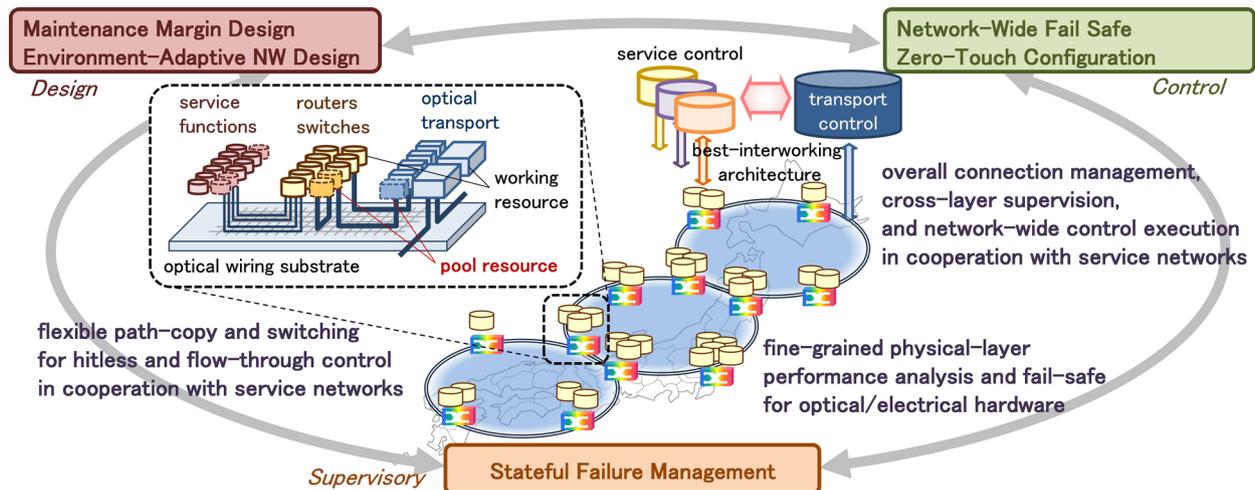


Fig. 12 Architectural points for future optical transport network infrastructure.

6. Conclusion

We gave an overview of the operational problems that emerge in current nationwide optical transport networks. Based on technology trends analysis for optical-transport-system configuration and network-control schemes, we proposed a vision of future nationwide optical transport network architecture expressed using five target features. The proposed architecture will mainly be realized by further extension and integration of software-based cross-layer network supervisory and control technologies, as well as connectivity and function virtualization technologies discussed in SDN or NFV. We expect this paper to contribute to discussions on future nationwide optical-transport-network architectures, related system architectures, and their controllability in cooperation with service networks.

References

- [1] H.V. Helvoort, *The ComSoc Guide to Next Generation Optical Transport: SDH/SONET/OTN*, John Wiley & Sons, ISBN978-0-470-22610-0, 2009.
- [2] T.J. Xia, S. Gringeri, and M. Tomizawa, "High-capacity optical transport networks," *IEEE Commun. Mag.*, vol.50, no.11, pp.170–178, Nov. 2012.
- [3] A. Gumaste and N. Krishnaswamy "Proliferation of the optical transport network: A use case based study," *IEEE Commun. Mag.*, vol.48, no.9, pp.54–61, Sept. 2010.
- [4] L. Pererson, A.A. Shabibi, T. Anshuts, S. Baker, D. Das, J. Hart, G. Palukar, and W. Snow, "Central office re-architected as a data center," *IEEE Commun. Mag.*, vol.54, no.10, pp.96–101, Oct. 2016.
- [5] R. Sabella, F. Testa, P. Iovanna, and G. Bottari, "Flexible packet-optical integration in the cloud age: Challenges and opportunities for network delayering," *IEEE Commun. Mag.*, vol.52, no.1, pp.35–43, Jan. 2014.
- [6] O. Gerstel, C. Filsfils, T. Telkamp, M. Gunkel, M. Horneffer, V. Lopez, and A. Mayoral, "Multi-layer capacity planning for IP-optical networks," *IEEE Commun. Mag.*, vol.52, no.1, pp.44–51, Jan. 2014.

- [7] T. Bauschert, C. Büsing, F.D'Andreagiovanni, A.M.C.A. Koster, M. Kutschka, and U. Steglich, "Network planning under demand uncertainty with robust optimization," *IEEE Commun. Mag.*, vol.52, no.2, pp.178–185, Feb. 2014.
- [8] T. Kawasaki, M. Inami, D. Shimazaki, and K. Koda, "Development of 100 Gbit/s high speed packet optical transport system," *IEICE Trans. Commun. (Japanese edition)*, vol.J99-B, no.4, pp.275–282, April 2016.
- [9] Fujitsu, "Revolutionary IFINITY Platform: Key Enabler of Programmable, Disaggregated Networks," <http://www.fujitsu.com/us/Images/IFINITY-platform-brief.pdf>
- [10] Ciena, "Product Brochure 6500 Packet-Optical Platform Transforming Networks into Intelligent Programmable Platforms" http://media.ciena.com/documents/6500_Packet_Optical_Platform_PB.pdf
- [11] Ericsson, "Ericsson SPO1400 Family–Packet Optical Transport Platforms (ETSI)," http://tele-a.ru/wp-content/uploads/2015/09/spo1400_datasheet.pdf
- [12] http://www.soumu.go.jp/main_content/000462459.pdf
- [13] <http://www.soumu.go.jp/johotsusintokei/field/data/gt010602.pdf>
- [14] Y. Miyamoto and M. Tomizawa, "High-capacity scalable optical communication for future optical transport network," 2014 IEEE ISSCC Digest of Technical Papers, pp.118–119, 2014.
- [15] A. D'Errico and G. Contestabile, "Next generation terabit transponder," 2016 OFC Conference and Exhibition, W4B.4, pp.1–3, 2016.
- [16] G. Raybon, "High symbol rate transmission systems for data rates from 400 Gb/s to 1 Tb/s," 2015 OFC Conference and Exhibition, M3G.1, pp.1–25, 2015.
- [17] P.J. Winzer, "Beyond 100G Ethernet," *IEEE Commun. Mag.*, vol.48, no.7, pp.26–30, July 2010.
- [18] D. Chiaroni, G. Buforn, C. Simonneau, S. Etienne, and J.-C. Antona, "Optical packet add/drop systems," 2010 OFC Conference, OThN3, March 2010.
- [19] N. Deng, S. Cao, T. Ma, X. Shi, X. Luo, S. Shen, and Q. Xiong, "A novel optical burst ring network with optical-layer aggregation and flexible bandwidth provisioning," 2011 OFC Conference and Exposition, OThR5, March, 2011.
- [20] M. Nakagawa, K. Masumoto, K. Hattori, T. Matsuda, M. Katayama, and K. Koda, "Flexible and cost-effective optical metro network with photonic-sub-lambda aggregation capability," *OECC/PS 2016 ThA2-2*, 2016.
- [21] Renesas, "Semiconductor Reliability Handbook," <https://www.renesas.com/en-in/media/support/products-common/reliability/r51zz0001ej0100.pdf>
- [22] K. Kobayashi, "[Invited Talk] How to mitigate reliability-related issues on nono-scaled LSIs," *IEICE Technical Report*, VLD2012-5, 2012.
- [23] XILINX, "Xilinx Rolls World's First Heterogeneous 3D FPGA," *Xcell journal*, issue 80 cover story, 2012, <https://www.xilinx.com/publications/archives/xcell/Xcell80.pdf>
- [24] ALTERA, "100G OTN Transponder IP solution," https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/po/ss-100g-transponder.pdf
- [25] ALTERA, "OTN Family | ODU Multiplexers for ODU Cross-Connect | TPOC244/TPOC314," https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/po/ss-tpoc244_314.pdf
- [26] XILINX, "Xilinx All Programmable Optical Transport Network (OTN) IP Subsystem Honored by 2016Lightwave Innovation Reviews," <http://press.xilinx.com/2016-03-16-Xilinx-All-Programmable-Optical-Transport-Network-OTN-IP-Subsystems-Honored-by-2016-Lightwave-Innovation-Reviews>
- [27] ITU-T Recommendation G.709/Y.1331, "Interfaces for the optical transport network," 2012
- [28] OIF, "OTN over Packet Fabric Protocol (OFP) Implementation Agreement," Nov. 2011 <http://www.oiforum.com/public/documents/oif-ofp-01.0.pdf>
- [29] ITU-T Recommendation G.8112/Y.1371, "Interfaces for the MPLS transport profile layer network," 2015.
- [30] BROADCOM, "BCM88650 Series World Most Dense 100GbE Switching Solution" <http://ja.broadcom.com/collateral/pb/88650-PB200-non-nda.pdf>
- [31] H.J. Chao and B. Liu, *High Performance Switches and Routers*, Wiley-Interscience, John Wiley & Sons, ISBN978-0-470-05367-6.
- [32] N. Kanekawa, E.H. Ibe, T. Suga, and Y. Uematsu, *Dependability of Electronic Systems: Mitigation of Hardware Failures, Soft Errors, and Electro-magnetic Disturbances*, Springer Verlag, ISBN 978-1-4419-6714-5, 2010.
- [33] M. Nicolaidis, *Soft Errors in Modern Electronic Systems*, Springer Science & Business media, ISBN 978-1-4419-6992-7, 2011.
- [34] ITU-T Recommendation K.124, "Overview of particle radiation effects on telecommunication systems," 2016.
- [35] D.R. Kurn, "Sources of failure in the public switched telephone network," *IEEE Computer*, vol.30, no.4, pp.31–36, 1997.
- [36] KDDI, "KDDI Powered Ethernet," <http://www.kddi.com/english/business/cloud-network-voice/intranet/powered-ethernet/>
- [37] 5GPP "5G Vision–The 5G Infrastructure Public Private Partnership: the next generation of communications networks and services," <https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf>
- [38] M. Tornatore, G. Maier, and A. Pattavina, "Availability design of optical transport networks," *IEEE J. Sel. Areas Commun.*, vol.23, no.8, pp.1520–1532, 2005.
- [39] S. Verbrugge, D. Colle, P. Demeester, R. Huelsermann, and M. Jaeger, "General availability model for multilayer transport networks," *IEEE DRCN 2005*, pp.85–92, 2005.
- [40] S. Verbrugge, D. Colle, M. Pickavet, P. Demeester, S. Pasqualini, A. Iselt, A. Kirstädter, R. Hülselmann, F.-J. Westphal, and M. Jäger, "Methodology and input availability parameters for calculating OpEx and CapEx costs for realistic network scenarios," *J. Opt. Netw.*, vol.5, no.6, pp.509–520, 2006.
- [41] <http://www.ieice.org/pn/jpn/jpnm.html>
- [42] L. Velasco, A. Castro, D. King, O. Gerstel, R. Casellas, and V. Lopez, "In-operation network planning," *IEEE Commun. Mag.*, vol.52, no.1, pp.52–60, Jan. 2014.
- [43] L. Velasco, F. Morales, L. Gifre, A. Castro, O.G. de Dios, and M. Ruiz, "On-demand incremental capacity planning in optical transport networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol.8, no.1, pp.11–22, 2016.
- [44] D. Mcdysan, "Software defined networking opportunities for transport," *IEEE Commun. Mag.*, vol.51, no.3, pp.28–31, March 2013.
- [45] <http://www.juniper.net/support/eol/junos.html>
- [46] http://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/ios-software-releases-12-2-mainline/prod_bulletin0900aecd801eda8a.html
- [47] <https://www.juniper.net/support/eol/>
- [48] <http://www.cisco.com/c/en/us/products/eos-eol-listing.html>
- [49] M. Yabuuchi, R. Kishida, and K. Kobayashi, "Correlations between BTI-induced degradations and process variations on ASICs and FPGAs," *IEICE Trans. Fundamentals*, vol.E97-A, no.12, pp.2367–2372, Dec. 2014.
- [50] N. Suzumura, M. Ogasawara, K. Makabe, T. Kamoshima, T. Ouchi, S. Yamamoto, T. Furusawa, and E. Murakami, "Comprehensive lifetime prediction for intrinsic and extrinsic TDDDB failures in Cu/Low-k interconnects," *IEEE IITC 2011*, DOI:10.1109/IITC.2011.594292, 2011.
- [51] M. Umeki, E. Nakazawa, R. Kitao, S. Hiroshima, T. Kurokawa, N. Furutake, H. Yamamoto, N. Inoue, Y. Tsuchiya, and Y. Hayashi, "Effects of metal-cap coverage on electro-migration (EM) tolerance for scaled down Cu interconnects," *IEEE IITC 2011*, DOI:10.1109/IITC.2011.5940294, 2011.
- [52] Y. Mitani, S. Fukatsu, D. Hagishima, and K. Matsuzawa, "Lifetime prediction of channel hot carrier degradation in pMOSFETs separating NBTI component," *IEEE ICICDT 2011*, DOI: 10.1109/ICI-

- CDT.2012.6232842, 2011..
- [53] E. Ibe, H. Taniguchi, Y. Yahagi, K. Shimbo, and T. Toba, "Impact of scaling effects on neutron-induced soft error in SRAMs from a 250 nm to a 22 nm design rule," *IEEE Trans. Electron Device*, vol.57, no.7, pp.1527–1538, 2010.
- [54] E. Oki, K. Shiimoto, D. Shimazaki, N. Yamanaka, W. Imajuku, and Y. Takigawa, "Dynamic multilayer routing schemes in GMPLS-based IP+optical networks," *IEEE Commun. Mag.*, vol.43, no.1, pp.108–114, Jan. 2005.
- [55] B. Puype, J.-P. Vasseur, A. Groebbens, S.D. Maesschalk, D. Colle, I. Lievens, M. Pickavet, and P. Demeester, "Benefits of GMPLS for multilayer recovery," *IEEE Commun. Mag.*, vol.43, no.7, pp.51–59, July 2005.
- [56] G. Parulkar, T. Tofigh, and M.D. Leenheer, "SDN control of Packet-over-optical Networks," *OFC 2015, WIG.4*, 2015.
- [57] Y. Ito, "A new paradigm in optical communications and networks," *IEEE Commun. Mag.*, vol.51, no.3, pp.24–26, March 2013.
- [58] S. Gao, X. Cao, T. Sato, T. Miyazawa, S. Yoshida, N. Yoshikane, T. Tsuritani, H. Harai, S. Okamoto, and N. Yamanaka, "Demonstration of SDN/OpenFlow-based path control for large-scale multi-domain/multi-technology optical transport networks," *IEICE Trans. Commun.*, vol.E99-B, no.7, pp.1492–1500, July 2016.
- [59] ETSI, "Network Functions Virtualization," <http://www.etsi.org/technologies-clusters/technologies/nfv>
- [60] <http://onosproject.org/>
- [61] <http://www.openroadm.org/>
- [62] Y. Wang, E. Keller, B. Biskeborn, J.v.d. Merve, and J. Rexford, "Virtual routers on the move: Live router migration as a network-management primitive," *Proc. ACM SIGCOMM*, pp.231–242, 2008.
- [63] X. Chen and C. Phillips, "Virtual router migration and infrastructure sleeping for energy management of IP over WDM networks," *Proc. TEMU 2012*, pp.31–36, 2012.
- [64] X. Liu, P. Juluni, and D. Medhi, "An experimental study on dynamic network reconfiguration in a virtualized network environment using autonomic management," *Proc. IFIP/IEEE Integrated Network Management (IM 2013)*, pp.616–622, 2013.
- [65] T.J. Xia, H. Fevrier, T. Wang, and T. Morioka, "Introduction of spectrally and spatially flexible optical networks," *IEEE Commun. Mag.*, vol.53, no.2, pp.24–33, 2015.
- [66] ITU-T Recommendation G.874, "Management Aspects of Optical Transport Network Elements," 2013.
- [67] R. Rejab, M.S. Leeson, and R.J. Green, "Fault and attack management in all-optical networks," *IEEE Commun. Mag.*, vol.44, no.11, pp.79–86, Nov. 2006.
- [68] W. Yuttachai, P. Saengudomlert, and W. Kumwilaisak, "Detection and localization of link quality degradation in transparent WDM networks," *IEICE Trans. Commun.*, vol.E96-B, no.6, pp.1412–1424, June 2013.
- [69] Y. Uematsu, "Performance Evaluation of Cell-Stream Synchronization Mechanism for Cell-based Transport Systems," *IEICE Trans. Commun. (Japanese edition)*, vol.J83-B, no.12, pp.1661–1669, Dec. 2000.
- [70] Z. Sui and V.Y. Liu, "Dual failure resiliency on single failure protected packet optical integrated networks," 2015 11th International Conference on the Design of Reliable Communication Networks (DRCN), pp.119–122, 2015.
- [71] K.N. Georgakilas, K. Katrinis, A. Tzanakaki, and O.B. Madsen, "Impact of dual-link failures on impairment-aware routed networks," 2010 12th International Conference on Transparent Optical Networks, Tu.B3.1, pp.1–4, 2010.
- [72] D. Harle and S. Albarak, "Differentiated survivability in a distributed GMPLS-based IP-over-optical network," 2008 International Conference on Optical Network Design and Modeling, pp.1–5, DOI: 10.1109/ONDM.2008.4578400, 2008.
- [73] Y. Uematsu, M. Nakagawa, H. Yamamoto, S. Kamamura, K. Genda, and M. Katayama, "Shared-resource-pool design scheme for failure-resilient optical transport network," *IEICE Trans. Commun.* (Japanese edition), vol.J99-B, no.4, pp.345–355, April 2016.
- [74] S. Gringeri, N. Bitar, and T.J. Xia, "Extending software defined network principles to include optical transport," *IEEE Commun. Mag.*, vol.51, no.3, pp.32–40, March 2013.
- [75] A. Gumaste and S. Akhtar, "Evolution of packet-optical integration in backbone and metropolitan high-speed networks: A standards perspective," *IEEE Commun. Mag.*, vol.51, no.11, pp.105–111, Nov. 2013.
- [76] C. Janz, L. Ong, K. Sethuraman, and V. Shukla, "Emerging transport SDN architecture and use cases," *IEEE Commun. Mag.*, vol.54, no.10, pp.116–121, Oct. 2016.
- [77] T. Takeda and Y. Uematsu, "Network design and control for router offloading networks," *iPOP2012 Proceedings*, 1-1, 2012.
- [78] S. Kamamura, H. Mori, D. Shimazaki, K. Genda, and Y. Uematsu, "OSPF and BGP state migration for resource portable IP router," *IEEE GLOBECOM 2015 Proceedings, CQRM3 (Traffic Control and Network Virtualization)*, 2015.

Appendix A: Calculation Scheme for Average Resource-Shortage Time

Let us consider a system in which working and pool entities of function block k as shown in Figure 8. The resource-shortage probability P_k can be described as follows based on Eq. (3) of [43].

$$P_k = 1 - \sum_{i=0}^{\Delta n_k} n_k + \Delta n_k C_i \cdot (\alpha_k)^i \cdot (1 - \alpha_k)^{n_k + \Delta n_k - i} \quad (\text{A} \cdot 1)$$

n_k is the number of working entities of function k in a building and Δn_k is the number of pool entities of function k in a building.

α_k is $MTTRk/(MTBFk+MTTRk)$, where $MTBFk$ is the mean time between failures of function k , and $MTTRk$ is the failure-recovery time corresponding to the maintenance condition. The average resource-shortage time in Fig. 8 is P_k multiplied by $60 * 24 * 365$ to show the average minutes per year, assuming $MTBFk = 100,000$ hours, and $n_k = 4$ and 20. The P_k tends to increase as $MTTRk$ increases and tends to decrease as Δn_k increases. When $n_k = 20$, P_k under condition $\Delta n_k = 1$ and $MTTRk = 72$ hours is 56.6 minutes/year and is lower than 210 minutes/hour of P_k value under condition $\Delta n_k = 0$ and $MTTRk = 2$ hours. Assuming that keeping the resource-shortage probability of each function block at the same level enables the end-to-end availability at the same level, a small amount of pool resources can drastically relax maintenance conditions (i.e. $MTTRk$).

Appendix B: Calculation Scheme for Average Staff-Unavailable Time

Figure A-1 shows the maintenance procedure based on the time axes shown in Figs. 9(a) and (b), where

n is the number of overall working entities in a building,
 Δn is the number of overall pool resource entities in a building, and

N is the number of buildings that staff can reach within

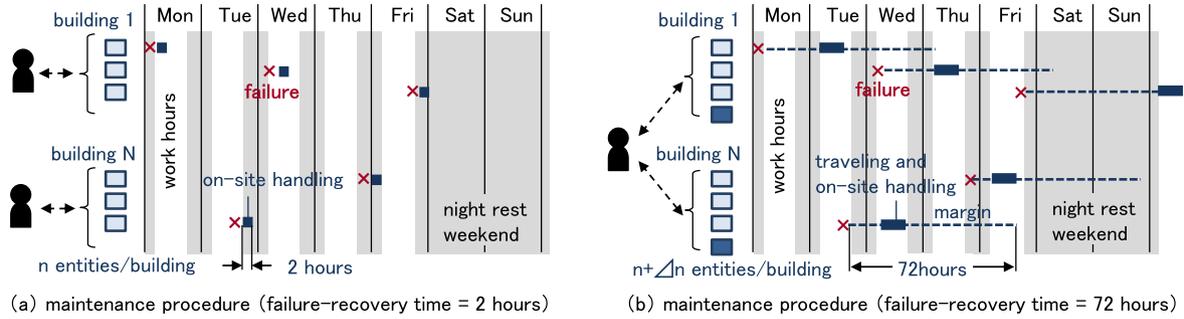


Fig. A-1 Maintenance conditions (failure-recovery time) and maintenance procedure.

the assigned traveling time.

Assuming that the on-site failure-handling time is generally 1.5 hours, dedicated staff is required for each building for a recovery time of 2 hours (a). For a recovery time of 72 hours (b), it is sufficient to ensure traveling and on-site handling times somewhere within 72 hours after failure occurrence. Therefore, the sum of traveling time, on-site handling time and margin for actual execution time planning or adjustment is 72 hours.

Let us estimate the probability that unhandled failure occurs due to staff shortage for cases (a) and (b). For case (a), assuming a system in which S staff members are assigned to a building and execute repairs for n entities failing randomly within the determined recovery time, the system should be precisely analyzed by $M/D/S$ (∞) queuing model. In this study the staff-unavailable probability was evaluated using the following simple scheme, assuming that a sufficient number of staff members are assigned against the average number of failures occurring concurrently.

The probability generation function in the number of failures occurring concurrently in a building can be written as follows.

$$X_a(z, n, \alpha_a) = (\alpha_a \cdot z + 1 - \alpha_a)^n = \sum_{i=0}^{\infty} a_i \cdot z^i \quad (\text{A} \cdot 2)$$

$$\alpha_a = \text{MTTR}(2H) / (\text{MTTR}(2H) + \text{MTBF})$$

a_i is the probability of i entities failing concurrently. Assuming S_a staff members are assigned for each building, the staff-unavailable probability in at least one building among N buildings (P_a) is written as follows.

$$P_a = 1 - \left(\sum_{i=0}^{S_a} a_i \right)^N \quad (\text{A} \cdot 3)$$

For case (b), the probability-generation function in the number of failures occurring concurrently in N buildings is written as follows.

$$X_b(z, n, \Delta n, N, \alpha_b) = (\alpha_b \cdot z + 1 - \alpha_b)^{(n+\Delta n) \cdot N} = \sum_{i=0}^{\infty} b_i \cdot z^i$$

$$\alpha_b = \text{MTTR}(72H) / (\text{MTTR}(72H) + \text{MTBF}) \quad (\text{A} \cdot 4)$$

Assuming S_b staff members assigned to N buildings, the

staff-unavailable probability (P_b) is expressed as follows.

$$P_b = \sum_{i=S_b+1}^{\infty} b_i \quad (\text{A} \cdot 5)$$

The average staff-unavailable time in Fig. 9 is P_a and P_b multiplied by $60 \cdot 24 \cdot 365$ to show the average minutes per year. It is assumed that one staff member assigned to a failure cannot be assigned to another failure until the determined failure recovery time (72 hours), which is the worst-case estimation.



Yoshihiko Uematsu received his B.S. and M.S. degrees in physics and Ph.D. degree in information science and technology from the University of Tokyo in 1988, 1990, and 2006. He joined NTT Laboratories in 1990 and has carried out research and development of multi-layer transport network architecture, end-to-end network control scheme, and network system design that includes high-speed SDH/ATM transport systems, transport quality control systems, media gateways for real-time communications, and IP routers for multi-service integrated networks. His current research interests include transport networking flexibility enhancement and resource virtualization. He is a member of IEICE and IEEE. He received the Best Paper Award from the International Conference on Communications from IEEE in 2014.



Shohei Kamamura received B.E., M.E. and Ph.D. degrees from Waseda University, Tokyo in 2004, 2006 and 2013, respectively. He joined the Nippon Telegraph and Telephone Corporation (NTT) in 2006. Since then, his research interests have included reliable IP topology and routing design. He received the Young Engineer Award from the IEICE in 2010 and the Best Paper Award from International Conference on Communications from the IEEE in 2014. He is a member of IEICE and IEEE.



Hiroki Date received the B.E. and M.E. degrees from Tokyo Institute of Technology in 2002 and 2008, respectively. In 2008, he joined Nippon Telegraph and Telephone (NTT) Corporation, Tokyo, Japan. Currently, he is with NTT Network Service Systems Laboratories, where he is engaged in research and development on transport network systems and centralized control of IP and optical transport networks. He is a member of the IEICE.



Hiroshi Yamamoto received the B.S., M.S. and Ph.D. degrees from Osaka University, Japan, in 2008, 2010 and 2013, respectively. In April 2013, he joined Nippon Telegraph and Telephone Corporation (NTT) as a researcher. His research interest includes management in backbone transport networks. He is a member of IEICE and IEEE.



Aki Fukuda received B.E., M.E. degrees from Akita University in 2007, 2009, respectively. She joined the Nippon Telegraph and Telephone Corporation (NTT) in 2009 and has carried out research of network virtualization and bandwidth control on IP network. Her current research interest includes reliable IP topology and routing design.



Rie Hayashi received the B.S. and M.S. degrees in Electronic Engineering from the University of Tokyo, Tokyo, Japan, in 2001 and 2003, respectively. In 2003, she joined NTT (Nippon Telegraph and Telephone Corporation) Network Service Systems Laboratories, where she has been engaging in research of IP-optical traffic engineering techniques. She is now researching traffic engineering techniques for next-generation backbone network. She received the Young Engineer Award in 2009, ICM English

Session Encouragement Award in 2011, and Communications Society Best Paper Award in 2012 from the IEICE. She is a member of IEICE.



Katsutoshi Koda received the B.S. and M.S. in mechanical sciences from Tokyo University of Science in 1987 and 1989, respectively. Since joining Nippon Telegraph and Telephone Corporation (NTT) in 1989, he had been engaged in the development of transport systems at the Network System Development Center and Network Service Systems Laboratories. He has developed transport network systems (such as MPLS-TP and Sync-E systems). Currently, he is engaged in research and development of future optical

networking and transport network management technology.