Terabit Node for Next-generation IP Networks

Itaru Mimura
Takeki Yazaki
Norihiko Moriwaki

OVERVIEW: IP networks have become an important communication infrastructure. Such services as virtual private networks (VPNs) for enterprises, electronic-commerce, and voice over IP are being deployed over them. Accordingly, the volume of traffic carried over IP networks is growing dramatically. To support this growth, higher speed line-interfaces and larger capacity switches will be the key components of next-generation IP network nodes. Carrier-class reliability and quality of service guarantees are strongly demanded as well for mission-critical business support. In this paper, we discuss the next-generation IP network architecture and the major requirements for the terabit IP nodes needed in backbone networks of the next generation. A distributed routing-processor architecture, scalable large crossbar switch, and high-speed line interface are the key technologies of terabit IP nodes.

INTRODUCTION

IP (Internet protocol) networks have become a global network infrastructure, and many services are deployed over such IP networks as the Internet. Accordingly, IP networks have become one of the most important telecommunication infrastructures. Since many services, for example, virtual private networks (VPNs), electronic commerce, and voice over IP, are using IP networks, the traffic loads are increasing dramatically. Traffic loads are now four to seven times those in 1999. To support this heavy usage of bandwidth, telecommunication carriers, Internet service providers (ISPs), and Internet data centers (IDCs) will soon need to implement terabit-class IP nodes in their backbone networks.

Hitachi's GR2000 gigabit router[1)] has multi-gigabit switching capacity and supports high-end edge-router functions. The GR2000 thus offers a total IP network solution, from the edge node to the gigabit core network. To meet the demand for more bandwidth, Hitachi is now developing a terabit-class IP network node.

In this paper, we introduce our terabit IP node for next-generation IP backbone networks.

REQUIREMENTS FOR NEXT-GENERATION IP NETWORKS

In this section, we first discuss the next-generation IP network architecture to clarify the key requirements for terabit-class IP nodes.

Architecture

The functions needed in next-generation networks can be logically divided into at least four categories, as shown in Fig. 1: subscriber termination, line aggregation, IP services (filtering, usage parameter control based on flow classification, security, flow-to-label mapping, etc.), and high-speed and large-volume transport.

The edge functions (subscriber termination, line aggregation, and IP services) can be implemented in

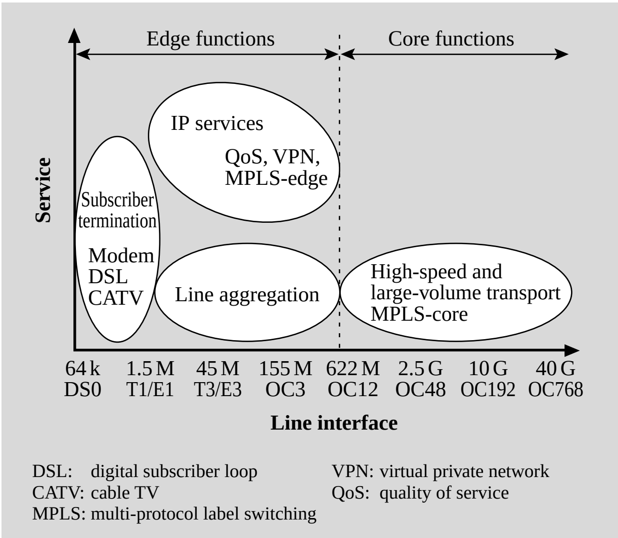

Fig. 1—Categorization of Functions Relationship in Next-generation IP Networks.

an access/edge network node. Since flow-based operation and precise quality of service (QoS) control are required, these functions should be implemented in the access/edge node, where the number of flows and transmission speed are relatively smaller and slower than those of backbone network nodes.

The high-speed large-volume transport function is the main function (the core function) of the backbone network node. The backbone node should thus provide wire-speed packet forwarding, such as OC192 (10 Gbit/s) and beyond [OC768(40 Gbit/s)]. We believe that backbone nodes should focus on high-speed packet forwarding. Therefore, only selected functions such as Diff-Serv[2] (differentiated services) core, MPLS[3] (multi-protocol label switching) core should be implemented in backbone nodes.

Key Features

Key features of next-generation IP networks will be resource management and failure recovery using MPLS.

(1) Resource management in IP networks

Quality of service or class of service (CoS) is a key function for mission-critical business support and service differentiation. For QoS/CoS support, network-wide resource management is necessary[4].

Hitachi's next-generation IP network architecture targets the most advanced QoS/CoS control. As shown in Fig. 2, an access node will classify flows by using the IP header information. Packets sent from user terminals will be terminated and multiplexed at the access node, then forwarded to an edge node. To provide the QoS/CoS function, a traffic engineering approach is used. For instance, bandwidth or priority resources should be allocated to each flow after it is classified by the access node. Paths between user terminals and edge nodes should be allocated to guarantee QoS/CoS.

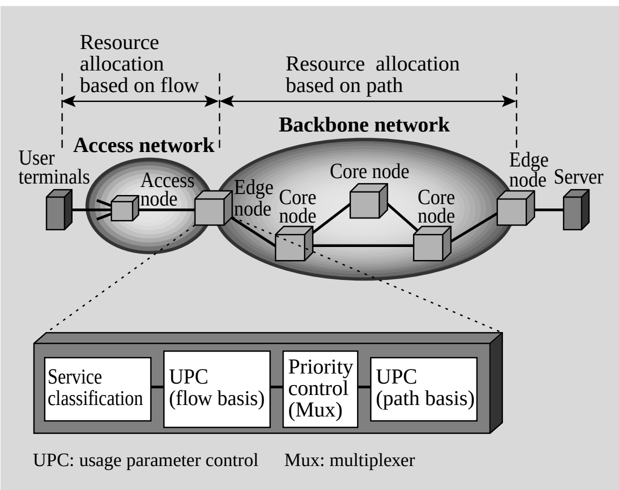

Fig. 2—Next-generation IP Network Architecture.

Between the edge nodes of backbone networks, MPLS label paths are allocated in advance, based on the services or the destination addresses. A Shim-header in a packet over SONET (POS), a virtual channel identifier (VCI) / virtual path identifier (VPI) in an ATM (asynchronous transfer mode) header, or the lambda of wavelength division multiplexing (WDM) are used as the MPLS label. At the boundary between the backbone and access networks, edge nodes classify the services or flows based on the IP header and payload. The edge nodes then select appropriate label paths and forward the packets along with the label paths. The terabit core node of the backbone network transports incoming packets based on their MPLS labels.

(2) Failure recovery with MPLS

In addition to the traffic engineering, MPLS can be used for failure recovery, so-called "MPLS fast restoration." In MPLS fast restoration, a backup label path is allocated in advance for protection. When a link or node failure is detected in transmission equipment, or when an alarm indication from the network management system is received, the fast restoration selects a backup path for continuing packet transmission. A link or node switch-over can thus be completed within several ten milliseconds. When a network-wide path substitution is required, only a couple of seconds is needed to switch over from a working path to a protection path. This MPLS label switch-over is applied to backbone networks. Since one label path for backup can be shared among several protection label paths, effective path usage is expected with MPLS fast restoration.

KEY REQUIREMENTS FOR TERABIT-CLASS IP NODES

Fig. 3 shows the relationship between the current GR2000 router and the next-generation terabit IP node in terms of functions. The GR2000 has a switching capacity of 45 Gbit/s and a line interface-speed from 64 kbit/s to 2.5 Gbit/s (OC48). It also supports fine-granularity QoS guarantees (Diff-Serv core/edge), filters, and MPLS edge/core functions. These capabilities enable the GR2000 to be applied not only to high-performance edge nodes but also to multi-gigabit-class backbone nodes. The terabit node is

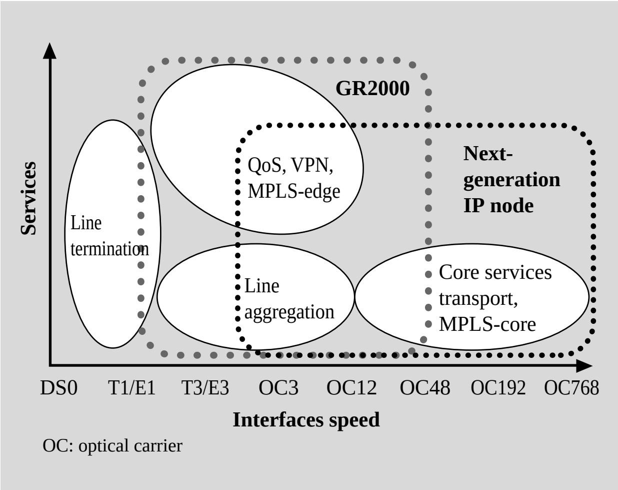

Fig. 3—Function Allocation to Terabit Node.

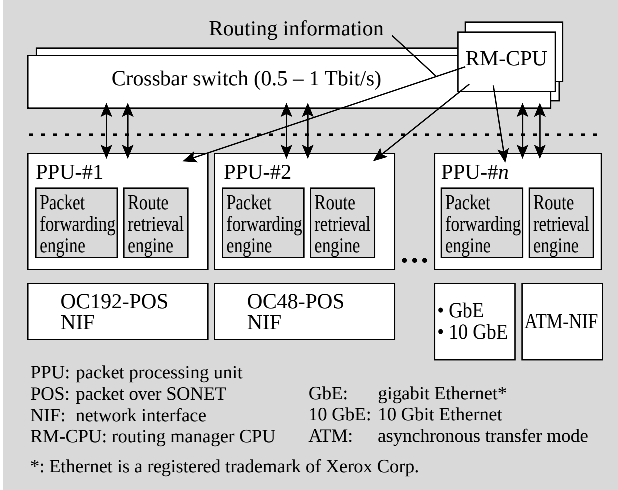

Fig. 4—System Block Diagram of Terabit Node.

targeted for application to the higher speed core portion of multi-terabit backbone networks.

Many high-speed lines [OC48 (2.5 Gbit/s) - OC192 (10 Gbit/s)] are aggregated in the core node. Therefore, a very large switching capacity and a high port density are key features to be supported. In addition, several functions such as VPN, MPLS core, Diff-Serv core, and filter functions are needed to support edge-to-edge QoS guarantee and rapid failure recovery.

Other basic requirements for the next-generation IP node are as follows.

(1) Scalable capacity up to terabit

To support multiple OC192 lines, the node should have terabit-class switch capacity, and incremental capacity enhancement (scalability).

(2) High-speed network interface up to OC768

Highly aggregated traffic would be effectively transported with high-speed interfaces such as OC192 and OC768. Higher speed line interfaces would improve the usage efficiency of installed optical fibers.

(3) High port density

The higher the port density, the smaller the footprint and the fewer the number of interface slots needed in the chassis. A smaller footprint saves central office building space, a strong need in metropolitan areas.

(4) Carrier-class reliability and dependability

A redundant architecture is needed for carrier-class reliability. The crossbar switch, power supply, and cooling components should be fully redundant. An automatic protection system (APS) is needed to protect against line failures.

NEXT-GENERATION IP NODE ARCHITECTURE

As shown in Fig. 4, Hitachi's next-generation IP node consists of a large-capacity scalable crossbar switch (CSW), packet-processing units (PPUs), network interfaces (NIFs), and a routing manager CPU (RM-CPU). This is the same architecture as in the GR2000 router. Since the packet forwarding processes are distributed at each forwarding engine of the PPUs, this structure provides greater packet-forwarding performance with relatively slower hardware (i.e., LSIs).

We are developing several enabling technologies for terabit nodes. One in particular is a crossbar switch architecture for terabit-class switching. The crossbar switch merges several variable-length IP packets that have the same destination output port into one large fixed-size packet. This merger reduces packet-header overhead and switching frequency because the common packet header represents the destination of all the merged packets. The architecture thus generates a longer arbitration* interval, and a higher performance arbitration algorithm can be applied during this longer interval. This architecture will enable a 0.5 - 1 Tbit/s switch capacity in a single chassis.

It will also allow the node to have a fully redundant configuration. SONET-based APS will provide the carrier-grade reliability as well.

Major specifications of the terabit node are shown in Table 1.

*: Computational process to determine the combination of input and output ports.

TABLE 1. Specifications of Terabit Node

Item	Specification
Architecture	Distributed routing processor + distributed switch with large fixed-length cell
Node capacity	0.5 – 1 Tbit/s
Packet forwarding	Wire-speed
Number of slots	16 /chassis
Line cards	OC192-POS, OC48-POS, 10 Gbit-Ethernet, Gigabit-Ethernet
Functions	IPv4/IPv6 routing, MPLS, QoS (Diff-Serv), filtering, multicasting
Routing protocols	OSPF-v2, BGP-4, IS-IS, RIP-v2, IGMP, DVMRP, PIM DM/SM

CONCLUSIONS

We have described the next-generation IP network architecture and Hitachi's terabit node. Hitachi is now developing a terabit-class IP network node. This node is designed to be the core of carrier-class backbone networks. Its major features are
(1) terabit-class scalable capacity (0.5 - 1 Tbit/s),
(2) high-speed WAN interface, such as OC192 (10 Gbit/s) and OC768 (40 Gbit/s) support, and
(3) carrier-grade reliability.

In addition, the terabit IP node will have the same functionality and control interfaces as the GR2000 router. This means that policy servers and network management systems deployed for the GR2000 will run without modification on the terabit-class node. This will provide a smooth migration scenario from multi-gigabit to terabit-class IP networks. With the GR2000 and terabit-class nodes, Hitachi provides QoS-guaranteed, multi-service (VPN, fast restoration, etc.) IP networks.

To achieve more efficient next-generation IP networks, Hitachi will develop optical and IP integrated nodes. The integration of an optical-cross-connect, optical add-drop multiplexer, and IP node will be the keys to achieving an advanced optical-IP network solution. Hitachi supports a total IP network solution with the GR2000 and next-generation terabit-class IP node, enabling the provision of next-generation IP network services.

REFERENCES

(1) N. Ikeda et al., "Gigabit Routers for Advanced IP Networks," *Hitachi Review* **49**, 159-162 (Dec. 2000) (this issue).

(2) T. Aimoto et al., "Overview of Diff-Serv Technology: Its Mechanism and Implementation," *IEICE Transaction on Information and Systems (D)* **E83-D**, No. 5, 957-964 (May 2000).

(3) Multi-protocol Label Switching Architecture, draft-ietf-mpls-arch-07.txt, Internet Engineering Task Force (http://www.ietf.org/internet-drafts/draft-ietf-mpls-arch-07.txt).

(4) N. Endo et al., "Carrier Network Infrastructure for Integrated Optical and IP Network," *Hitachi Review* **49**, 151-154 (Dec. 2000) (this issue).

ABOUT THE AUTHORS

Itaru Mimura

Joined Hitachi, Ltd. in 1984, and now works at the Research & Development Group. He is currently engaged in development of digital video transmission systems, MPEG-over-IP systems, and IP routers. Mr. Mimura is a member of the IEICE, and ITE, and can be reached by e-mail at mimura@crl.hitachi.co.jp.

Takeki Yazaki

Joined Hitachi, Ltd. in 1995, and now works at the Research & Development Group. He is currently engaged in development of ATM switches and IP routers. Mr. Yazaki can be reached by e-mail at yazaki@crl.hitachi.co.jp.

Norihiko Moriwaki

Joined Hitachi, Ltd. in 1995, and now works at the Research & Development Group. He is currently engaged mainly in the development of ATM switches and IP routers. Mr. Moriwaki is a member of the ICIEC, and IEEE, and can be reached by e-mail at moriwaki@crl.hitachi.co.jp.