12th Optoelectronics and Communications Conference/16th International Conference on Integrated Optics and Optical Fiber Communication (OECC/IOOC2007)
Technical Digest, July 2007, Pacifico Yokohama

13A1-4

Generation-Free-Platform architecture with flexible physical and logical links using optical technology

S. Yanagimachi*[a], Y. Hidaka*[b], J Suzuki*[c], J. Higuchi*[d], T. Yoshikawa*[e], A. Iwata*[f]
System Platforms Research Labs., NEC Corporation
1753 Shimonumabe Nakahara-ku, Kawasaki, 211-8666, Japan
E-mail: *[a] s-yanagimachi@ab.jp.nec.com, *[b] y-hidaka@bq.jp.nec.com, *[c]j-suzuki@ax.jp.nec.com, *[d]j-higuchi@ax.jp.nec.com, *[e] yoshikawa@cd.jp.nec.com, *[f] a-iwata@ah.jp.nec.com

1. Abstract

To achieve scalable flexible platform, we propose a stackable and modular system (GF-PF: Generation-Free-Platform) with flexible physical link using optical interconnection and logical link using our ExpEther technology.

2. Introduction

Recently, changes in business environment and process are rapidly. Especially, in the field of IT/NW, platforms needed to be replaced every two or three years to keep pace with innovations in technology and these replacements cause considerable capital investment. To solve this problem, we have already proposed a new concept, Generation Free Platform (GF-PF) [1]. The objective of our research is to pursue a scalable flexible platform that can correspond flexibly to rapid changes in business conditions. To achieve this scalable flexible platform requires modularity and reconfigurability of hardware resources and the application resources running on that hardware. In addition, system designs are required that can flexible interconnect resources and provide scalable expansion of those resources. In this paper, we describe the detailed hardware architecture of scalable flexible platform, GF-PF; especially, we focus on flexibility and reconfigurability of that platform that can be achieved by using of a stackable system and a modular system with flexible physical and logical link technologies.

3. GF-PF architecture

GF-PF is the platform which provides various services, using flexible interconnection of any resources among different shelves, and it has two main features.

First, the GF-PF logical service platform has a virtualization function that hides the complexity of hardware resources, such as physical connections and locations.

Second, the GF-PF physical service platform has following features: a stackable system that allows flexibly expanding physical resources, and a modular system that can flexibly reconfigure its hardware platform. Figure 1 shows the architectures of our system.

The followings are requirements of the interconnection for a scalable flexible platform.

- Switch capacity scalability
- Physical link scalability
- Logical link (protocol level connection over physical link) scalability and reconfigurability

4. Modular switch architecture

Several switch architectures, e.g., one-stage switching or multi-stage switching, have been widely used. Multi-stage switch architecture has been used for large-scale nodes because systems with one-stage switches lack the flexibility necessary for maximum switch port from the initial deployment. To achieve linear expansion of switch capacity, we use modularized multi-stage switch based on CLOS architecture where switches are connected with optical interconnect. The modular switch card is illustrated in Fig. 2. This switch card consists of a mother card with four switch modules on it. The switch module has an LSI chip surrounded by the four optical transceivers (PETIT) [2]. Since a switch capacity of 40-160 Gbps can be implemented in a single switch module and the switch card has four elements per switch module, the switch capacity of a single switch card is 160-640Gbps. In our design, three switch cards are used to achieve 480-1.92-Tbps system capacity.

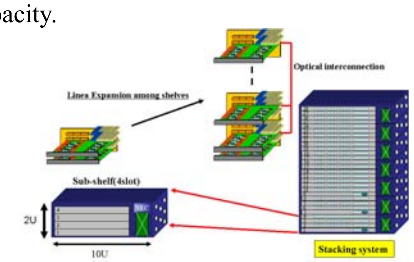

Fig 1. Stackable system and modular system with optical interconnection

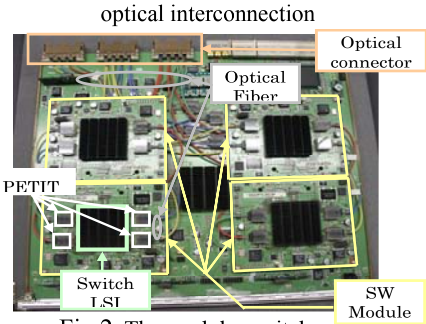

Fig 2. The modular switch card

5. Physical link technology (Optical interconnect)

Usually several interconnection among physical resources is electrical. When electrical interconnection is used, transmission speed is limited because of crosstalk between adjacent wiring, attenuation of signals, and the reflection. And also the number of backplane layers is increased and the connector area is larger in proportion to the increase in the number of slots. Furthermore, electrical wiring corresponds to a particular target transmission speed so flexible system to an increase in transmission speed is difficult. To solve these problems and to achieve scalable physical link, we use optical interconnection among physical resources. Furthermore, using optical interconnection makes backplane redesign unnecessary up to 20Gbps (10Gbps is available now, and 20Gbps is under development), and a flexible GF-PF platform can be achieved.

Figure 3 presents an eye diagram of the 10-Gbps optical output of the PETIT when it is driven by 10 Gbps. Good eye openings can be seen in the figure. Figure 4 presents the performance of optical interconnection. Transmission distance up to 50 meter can be achieved.

Meanwhile, since the number of LSI on a service card is increasing due to increasing functions, the possible area for optical components assembly becomes also smaller. We consider compact assembly of optical components to be a key technology for using optical interconnection efficiently on high-density cards.

Our optical interconnection system has following features.

- Small optical transceiver
- High density optical connector
- Small fiber assembly area

We use our ultra small optical module (PETIT) in the GF-PF. This module has 4×3.125 Gbps or 4×10 Gbps TX/RX channels. The LSI size (12 x 12mm) is much smaller than those of conventional optical modules (e.g., the size of SNAP12 is 45 x 14.5mm).

In addition, we use multi-fiber optical connectors. Their dimensions are 53 x 13 mm; it accommodates 96-optical fibers, and achieves a throughput of 1Tbps.

To avoid degraded insertion loss and ensure reliability, the bending radius of the conventional optical fiber is kept within 25-30 mm. To make small fiber assembly area, in the GF-PF, we use high-Δ and low diameter MMF that can tolerate a bending radius within 5 mm. Figure.5 compares the data of insertion loss for both types of MMF. The insertion loss of high-Δ MMF is the same as for conventional MMF. Figure 6 is a photograph of a fiber assembly on a service card using a high-Δ, low diameter MMF. Due to achieving fiber assembly with small bending radius fiber, fiber assembly area can be much smaller than that of a conventional design.

6. Logical link technology (ExpEther)

To achieve flexible logical link over physical link, we already proposed ExpEther [3] as Ethernet-based virtualization technology. It groups modularized resources interconnected by an Ethernet, and transports a PCI Express (PCIe) packet between the grouped modules by encapsulating it into an Ethernet frame. ExpEther, by utilizing an Ethernet technology, can achieve scalability and easy reconfigurability of hardware resources. The configuration of ExpEther is illustrated using Figure 7. A PCIe-to-Ethernet bridge (PEB) performs the encapsulation and decapsulation of a PCIe packet. A server and an I/O assigned to its server belong to the same VLAN. Note that the assignment of an I/O to a server can be flexibly changed by setting the VLAN of the I/O to that of the server.

Figure 8 shows an ExpEther card which performs a PCIe packet encapsulation on a server side. This card accommodates x4 PCIe interface(total throughput is 10Gbps). By using our ultra small optical module (PETIT) for 10G Ethernet optical ports, we achieved compact design and high performance of the ExpEther card where the maximum throughput of the PCIe interface can be up to x16.

7. Conclusion

We propose the GF-PF architecture with a stackable system and a modular system with optical interconnection (for physical link) and our ExpEther (for logical link). In this design, the system can continue to grow and reconfigure to correspond with rapid changes in business conditions. We also describe in detail the optical interconnection used in our GF-PF. By using ultra small optical transceiver (PETIT) and high-Δ, low diameter MMFs for fiber assembly, its area can be much smaller than that of a conventional design.

Acknowledgements

This work was partly supported by the Ministry of Internal Affairs and Communications (MIC).

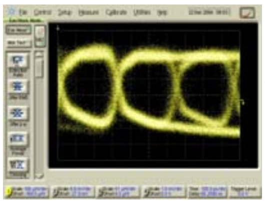

Fig 3. Optical waveforms at 10Gbps operation

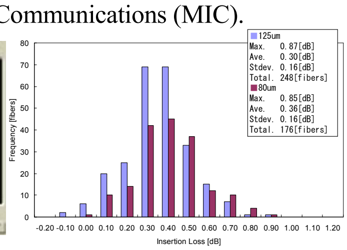

Fig 5. Insertion Loss (High -ΔMMF VS. Normal MMF)

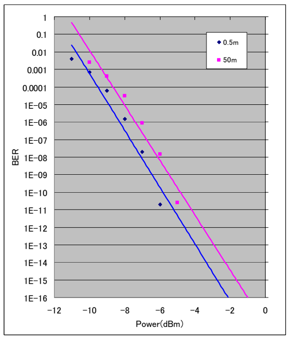

Fig 4. Performance of optical interconnection

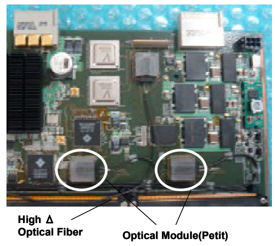

Fig 6. Fiber assembly using high-Δ and low diameter.

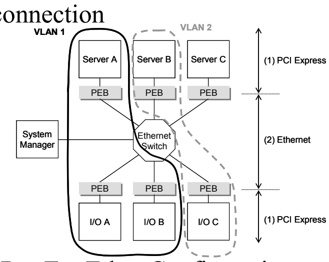

Fig 7. ExpEther Configuration

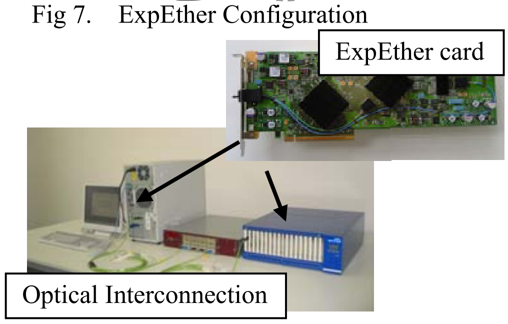

Fig 8. ExpEther card with optical interconnection

Reference

[1] N.Kami et al, "Next Generation System Architecture – Generation Free Platform", Technical Report of IEICE, NS2004-234, NS-689, pp.189-192.

[2] I.Hatakeyama, et al, "A 400 Gbps Backplane Switch with 10 Gbps/port Optical I/O Interfaces", SPIE2006

[3] J.Suzuki et al, "*ExpressEther* – Ethernet-Based Virtualization Technology for Reconfigurable Hardware Platform", IEEE Hot interconnects 14.