FDTN Architecture: Functionally Distributed Transport Networking Architecture

Kyota Hattori, Hiroki Date, Kentaro Ogawa, Takahiro Kobayashi,
Kenichi Higuchi, Hideaki Iwata and Shinichiro Chaki
NTT Network Service Systems Laboratories, NTT Corporation
9-11 Midori-Cho 3-Chome, Musashino-Shi, Tokyo, 180-8585, Japan
+81-422-59-7699
{hattori.kyota, date.hiroki, ogawa.kentaro, kobayashi.takahiro, higuchi.kenichi,
iwata.hideaki, chaki.shinichiro}@lab.ntt.co.jp

ABSTRACT
We demonstrate a prototype system of functionally distributed transport networking (FDTN) architecture that separates the control-plane processing part (control element, CE) from the forwarding-plane processing part (forwarding element, FE) of a router. In an FDTN, one path-control process in the CE consolidates and processes the path computations and path settings for two or more routers. This reduces the convergence time of the routing information. Moreover, if the CE fails due to disasters, such as earthquakes, the FDTN can reestablish operation of the routing function by switching from the failed CE to a standby one, which is geographically separated from the primary CE. We introduce the fundamentals of an FDTN and discuss our prototype system with a micro telecommunications computing architecture.

Categories and Subject Descriptors
C.2.1 [**COMPUTER-COMMUNICATION NETWORKS**]: Network Architecture and Design; C.2.3 [**COMPUTER-COMMUNICATION NETWORKS**]: Network Operations

General Terms
Management, Design, Reliability, Experimentation

Keywords
router, architecture, functionally-distributed, fail-over

1. INTRODUCTION
IP traffic volume is growing exponentially because of the rapid spread of broadband access and the continuing introduction of new applications and services. It is, therefore, important that IP networks be easily scalable without degrading in reliability. However, as Internet service providers and network carriers expand their IP networks and provide a wider variety of IP-based services, the convergence time of routing information is increasing.

The routing protocols used in IP networks are typically classified as either interior gateway protocols, which are used to exchange information inside a network (e.g., the open shortest path first (OSPF) protocol [1]), or exterior gateway protocols, which are used to exchange information between networks. OSPF is a link-state protocol that maintains the consistency of the link-state database at every node by exchanging information about the state of the network in the form of link state advertisements (LSAs), which are the basic communication means of the OSPF protocol. LSAs are periodically exchanged with neighboring nodes as well as when a network-state change is detected, and the link state database (LSDB), which is a tree-image of the network topology, is built. This exchange of LSAs causes the convergence time of the routing information to increase with the number of routers in the network, which limits the scale of the network.

The FDTN architecture was developed to overcome this problem [2]. This architecture configures a functionally distributed router (FDR), which is composed of one control-plane processing part (control element, CE) and several forwarding-plane processing parts (forwarding elements, FEs). This architecture reduces the convergence time of routing information because the number of times routing information is exchanged is reduced by the consolidation of the CE.

2. RELATED WORK
The SoftRouter architecture [3] is quite similar to our proposed router architecture. It has a control-plane processing server with high processing performance and facilitates adding functions to the transport stratum. However, the reliability of the CE has not been investigated.

With an architecture in which one CE consolidates the connected FEs, like the SoftRouter architecture, all the FEs continue to transmit data packets even if failures caused by disasters, such as earthquakes, occur in a CE because they are physically separated from it in this architecture, but they cannot receive the computed paths and path settings from the CE. That is, the FEs can no longer respond to changes in the network topology such as link-up/down of ports, links, and FEs. Therefore, it is critical that the CE in this architecture be highly reliable.

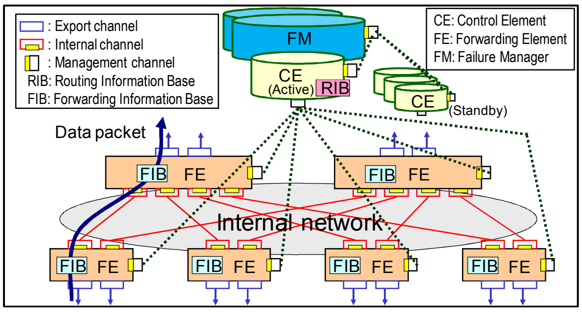

Figure 1. Configuration of FDR.

3. OVERVIEW OF FDTN

This section explains the basic framework of the FDTN architecture. Figure 1 shows the configuration of an FDR in which the CE is physically separated from the FEs. In this model, one or more FEs connect to one CE. The CE controls the connected FEs through control sessions, which are established through management channels. The router functions are distributed among the CE and FEs. The CE performs various functions: (a) collect and manage the topology information for the network between the FEs and construct the LSDB, (b) exchange the LSDB with other CEs that control other FE groups, (c) compute the routing information base (RIB), (d) generate the forwarding information base (FIB) for each connected FE from the RIB, and (e) send each FIB to the corresponding FE through a control session. The FEs also perform various functions: (a) collect topology information about the network between adjacent FEs, (b) send the topology information to the CE, and (c) forward the data packets according to the FIB. A failure manager (FM) is connected to the active and standby CEs, which are geographically separated. Each active CE sends a heartbeat signal to the FM. If the FM no longer receives the heartbeat signal from an active CE, it determines that the CE has failed, checks the CE database in the FM, and activates the standby CE. When the failed CE recovers, it connects to the FM as a new standby CE. This architecture can configure N+m redundant configurations in which multiple active CEs (N) have multiple independent standby CEs (m) as recorded by the CE database in the FM. Since a single FM controls all CEs, the FM is itself a single point of failure. This means that full duplication of the FM is required for the system to be robust. The FM has an active and a standby body and always runs status and data mirroring between them. Also, the links between the FM and CEs are duplicated.

4. DEMONSTRATION

Our prototype FDR system is composed of micro telecommunications computing architecture (μTCA). The μTCA is a smaller chassis embodiment that uses advanced telecom computing architecture [4] standard daughter cards, called Advanced Mezzanine Cards (AMCs). The FDR has a CE, an FM, and several FEs mounted on each AMC. The physical connections used as control planes between the AMCs use μTCA backplane interfaces. The CE and FEs are configured on the basis of the ForCES framework [5]. This demonstration is composed of two scenarios for checking the effectiveness and feasibility of the FDTN architecture. The first scenario is used to confirm the basic behavior of the CE and FEs. In the initial setting, there is no FIB entry at the FEs. Then, the operator activates the CE. The CE collects the topology information of the data-plane network for use in communicating with the FEs, computes the FIBs for each FE on the basis of the topology information, and sends them to the FEs. As a result, each FE has a complete routing table. Data packets are then forwarded at each FE. The second scenario is used to confirm that CE switching following a CE failure does not negatively affect the data-plane network. In the initial setting, routing and forwarding in the system function normally. The active CE (CE-1) sends a heartbeat signal to the FM and all the FEs. It also uploads CE-1's control data, including a list of the connected FEs to the CE database on the FM. Data is uploaded every time there is an update to the data in order to keep the CE database up-to-date. The FM monitors the state of CE-1 on the basis of the heartbeat signal and stores the CE-1 data in the CE database. It also stores a table showing the correspondences between CE-1 and the standby CE (CE-2). This table is sent to the FEs via CE-1. The FEs monitor the state of CE-1 on the basis of the heartbeat signal. If there is a loss of the heartbeat signal from CE-1, the FM and FEs connected to CE-1 detect a CE failure. The FM activates CE-2 and sets the data for CE-1 stored in the CE database to CE-2. All the FEs connected to CE-1 begin trying to switch their connections to CE-2. As it may take some time before CE-2 completes preparations for reconnection, the FEs periodically retry connecting to it. Once all the FEs have switched over to CE-2, CE-2 collects the topology information for the network between the reconnected FEs and re-computes the LSDB, RIB, and FIB for each connected FE. This enables the system to respond to changing situations in the network topology during the recovery following CE-1 failure.

5. CONCLUSION

Our demonstration aims to verify the concept of the FDTN architecture, which separates the control-plane processing part of a router from the forwarding-plane processing part. We will show the basic behavior of the FDR. Its redundant configuration, in which the active CE is geographically separated from the standby CE, will be demonstrated using our prototype system with μTCA.

6. ACKNOWLEDGMENTS

A part of this research project is supported by the National Institute of Information and Communications Technology (NICT) of Japan.

7. REFERENCES

[1] J. Moy, RFC 2178 - OSPF Version 2, 1997.

[2] K. Ogawa, M. Aoki, and S. Chaki, “Verification of path computational performance in functionally distributed transport networking on next-generation network,” IEICE APSITT 2008, pp. A-1-2, Apr. 2008.

[3] T.V. Lakshman, T. Nandagopal, R. Ramjee, K. Sabnani, and T. Woo, ”The SoftRouter architecture”, Proceedings of ACM HotNets-III Workshop, San Diego, CA, Nov. 2004.

[4] ATCA (Advanced Telecom Computing Architecture) http://www.picmg.org

[5] Yang, L., Dantu, R., Anderson, T., and Gopal, “Forwarding and Control Element Separation (ForCES) Framework”, RFC3746, Internet Engineering Task Force, Apr. 2004.