Label Switch Router
Supporting ATM and SONET Interfaces
- CSR: Cell Switch Router -

**S.Matsuzawa, H.Esaki, A. Mogi, K. Nagami, T. Jinmei,
T. Kon'no, Y. Katsube
{hiroshi,shigeom,mogi,nagami,jinmei,toru,katsube}@isl.rdc.toshiba.co.jp
Tel. +81-44-5492238, Fax.+81-44-520-1806
R&D Center, TOSHIBA Corporation,
1 Komukai-Toshiba-cho, Saiwai-ku, Kawasaki, 210, Japan**

Abstract

Architecture overview of Cell Switch Router (CSR) and the CSR prototype system supporting standard ATM interface and SONET interface is described. CSR can operate both with PVC (Permanent Virtual Connection) and with SVC (Switched Virtual Connection) as the VC for cut-through packet forwarding. CSR contains cell switch fabric and IP packet switch fabric to achieve high throughput IP forwarding.

IP packets are forwarded either through a cut-through packet transmission, in which packet are forwarded without reassembling IP packet nor IP header processing, or through a conventional hop-by-hop IP packet forwarding. This paper describes and proposes the mechanism to forward the connectionless IP packet flows at the CSR using SVC over ATM.

CSR system supports three cut-through path establishment policies. Those are (i) Data-Driven, according to the actual packet arrival, (ii) Topology Driven, according to the routing entry information, and (iii) Reservation Driven according to the explicit resource reservation request.. In the CSR system, the SVC connections for cut-through packet transmission can be established on demand using ATM Forum UNI 3.0 or UNI3.1.

1. Introduction

ATM (Asynchronous Transfer Mode) is realized as one of platforms to provide high speed datalink layer service with certain QOS (Quality of Service), with a unique user network interface [I.150][AFUNI]. Sometimes, it is said that ATM network will provide a router-less large cloud datalink platform, which does not have any routers within the large cloud network. However, we will need routers, even large ATM [RFC1937] cloud datalink platform will be provided [Esaki]. Therefore, physical or logical datalink network segments (i.e., IP subnets) will be interconnected through routers (i.e., through the network layer entities), even when the ATM becomes a major datalink platform. Recently, some high performance router architectures based on label switching methodology have been proposed. Those are tag switch [RFC2105], ARIS (Aggregated Route-Based IP Switch) [ARIS], IP switch [Ipsilon] and CSR (Cell Switch Router) [CSR][RFC2098]. [RFC2098], [RFC2105] and [Ipsilon] propose a router architecture containing both conventional IP processing module and cell-switching module. As discussed briefly in this paper, all architecture models excepting CSR do not have interoperability with the standard ATM switch and can not use the standard ATM links to interconnect the routers. Since the CSR supports the standard ATM interface (i.e., ATM UNI), CSR has the interoperability with standard ATM switch and can use the standard ATM links to interconnect CSRs.

The CSR system described in this paper can support SVC (Switched Virtual Connection) and PVC(Permanent VC). The CSRs described in this paper are interconnected through the standard ATM switches using both SVCs and PVCs. As discussed in section 3.4, through the support of SVC, as well as the support of PVC, we can obtain the following advantages compared with pure PVC operation.

(1) Effective use of VPI/VCI resource
(2) Reduction of required bandwidth between the CSRs
(3) On demand exclusive bandwidth allocation for the application flow

The following is the structure of the rest of this paper. Section 2 briefly discusses the label switching router paradigm. Sections 3 proposes the IP over ATM architecture using CSR (Cell Switch Router). Section 4 gives a brief conclusion.

2. High Performance Routers using Label Switching

2.1 Concept Summary of Router using Label Switching

For label switching architecture paradigm, some layer 3 packets are forwarded either using the datalink label (e.g., VPI/VCI) or using the tag label between datalink header and packet header (e.g., tag for Ethernet frame in tag switching architecture [RFC2105]).

The label is used as the information to forward the packets, without analyzing the internetwork layer address (e.g., IP address). This means that the label represents the destination address of internetwork layer packets. In the CSR architecture, packet forwarding using the label is called as the cut-through packet forwarding [CSR]. By using the label instead of internetwork layer address for packet forwarding, the router (e.g., CSR) does not need to look up the best-match policy based routing table, whose search key is the internetwork layer address.

When the internetwork layer packet arrives at the node (i.e., intermediate router or host) located at the egress point of cut-through path represented by the label, the conventional packet processing (i.e., analysis of internetwork layer header) is performed.

2.1.1 Packet Forwarding Route

In label switching architecture, it is not generally assumed all nodes within the datalink have a datalink-level full-mesh connectivity. This means that nodes are interconnected through links with some topology.

Packets are forwarded along the given topology, even for cut-through packet forwarding. This is different architectural paradigm from NHRP [NHRP]. In NHRP, the short-cut virtual connection (i.e., ATM-VCC for ATM) is established irrelevant to the nodes' topology. When a short-cut VC is established to some node, packets to the node does not travel the intermediate router(s), that is(are) given by the routing protocol information. On the other hand, in label switching architecture, packets are forwarded along the exactly the same route, that is given by the routing protocol information.

The route, that packets travel, is exactly the same route indicated by routing protocol. With cut-through packet forwarding, just the conventional packet forwarding process is bypassed, while the conventional route calculation is as it is and the packet's forwarding route is also as it is.

2.1.2 Cut-through Path Establishment Policy

There are two trigger types to invoke cut-through path establishment. One is according to the appearance of packet flow, and the other is according to the appearance of routing entry in the routing table. The former mode is called as the flow driven, and the later mode is called as the topology driven.

In the flow driven mode, the cut-through path is established on demand, according to the actual appearance of packet flow at the node. In the topology driven mode, the cut-through path is established in advance, according to the topology information of routers. Therefore, in the topology driven mode, the cut-through path is established, even if there is no actual packet flow.

2.1.3 Cut-through Path Management Policy

The maintenance of cut-through path is performed based on soft-state, hard-state or the combination of both. In the soft-state management policy, the state of cut-through path is periodically refreshed. When there is no state refreshment, the cut-through path is released. In the hard-state management policy, the state of cut-through path is securely established and is not refreshed.

2.1.4 Datalink Protocol

Datalink protocol (e.g., MAC protocol) has to be modified to apply the label switching technology for all proposed architectures excepting CSR. For example, the tag switching with Ethernet link, the frame format has to be modified (i.e., the thin tag is inserted between the Ethernet header and the internetwork layer packet header).

3. Architecture of Cell Switch Router

3.1 Architecture Abstraction of Cell Switch Router (CSR)

The following is the feature of CSR, compared with the other label switch router architectures [Ipsilon] [RFC2105][ARIS].

+ Soft-state cut-through path management
+ Any connection oriented datalink, including standard ATM link, using FANP
+ Full compatible with standard ATM interface including SVC availability
+ Dual mode operation; supporting both Flow-driven and Control-driven
+ Any granularity of flow aggregation

CSR establishes the cut-through path using FANP (Flow Attribute Notification Protocol) [RFC2129]. The trigger to establish the cut-through path is invoked by the up-stream node to the up-stream node, according to the appearance of packet. The route of cut-through path is determined using the routing table obtained by routing protocol (e.g., BGP). For the scaling purpose, any granularity of flow aggregation can be defined.

Since the CSR supports the standard ATM interface (i.e., ATM Forum UNI 3.0/3.1), CSR has a full-interoperability with the standard ATM switch and can use the standard ATM links (i.e., ATM-VC) as the link to interconnect CSRs.

3.2 FANP and ATM-VCC Management Architecture

Figure 3-1 shows an example of the configuration of ATM-VCCs and IP packet flows with the cut-through packet forwarding at CSRs. All ATM-VCCs, i.e., VCCs between CSRs and between CSR and host, are PVC (Permanent Virtual Connection) or SVC (Switched Virtual Connection). During the boot procedure of each node (CSR and host), the default-VCs are established to the adjacent nodes. The default-VC transfers both control packets (e.g., the control packets of FANP described in the following subsection) and hop-by-hop based user packets. Dedicated-VC transfers only the user packets that are forwarded through the cut-through packet forwarding mode. With PVC mode, the bundle of PVCs are established between the adjacent CSR in advance (e.g., during a boot procedure of CSR), and pick up one of un-used PVC for the newly appeared packet flow, that should be forwarded with cut-through mode. On the contrary, with SVC mode, the dedicated-VC is established on demand.

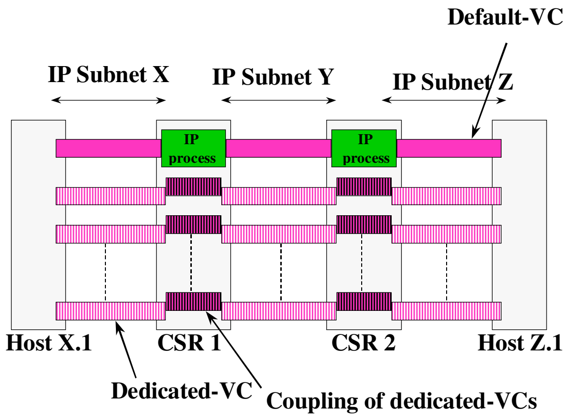

Figure 3-1. ATM-VCC Configuration Example for CSR

3.3 Cut-Through Operation Modes in CSR System

CSR supports the following three operational modes. Each mode has its benefit, therefore CSR simultaneously supports all three modes.

3.3.1 Data-Driven

The operation of data-driven is shown in figure 3-2. When the first IP packet (of the session or after the idle period) arrives at a CSR through the dedicated VC, the next hop router is decided using the IP header. CSR analyzes TCP/UDP header, as well as IP header. When the received IP packet belongs to the IP flow, that is likely to be long life session, a dedicated-VC is established for the IP packet flow through Q.9231 with SVC operation and CSR establishes the mapping table between the incoming VPI/VCI and the outgoing VPI/VCI. If the activity of the IP packet flow becomes low, the mapping information (i.e., the cached information) is pushed out.

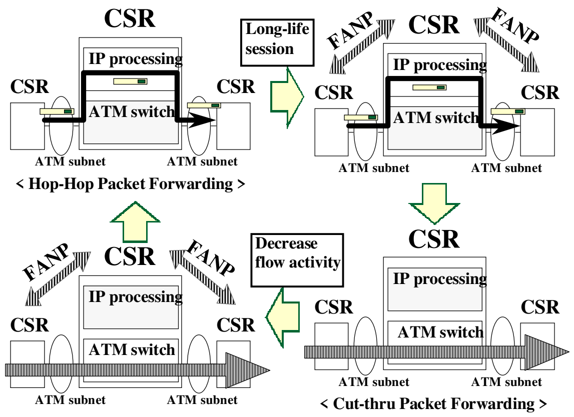

Figure 3-2. Packet Forwarding in CSR

3.3.2 Topology Driven

The cut-through paths are establish, according to the appearance of routing entry in the routing table. Each router obtains routing information, e.g., router's addresses and network addresses (i.e., network prefix plus net-mask) through the routing protocol. When the newly appeared routing entry is the one where the cut-through path should be established, the cut-through path to the router corresponding to the routing entry is established.

3.3.3 Reservation Driven

The cut-through path is established, according to the reception of control message of resource reservation protocol. In the case of RSVP, the cut-through path is established , according to the reception of RESV message [RFC2205].

3.3.4 Comparison of Operational Modes

Table 3-1 shows the comparison among three operational modes. Here, data-driven and reservation driven basically have the same feature.
In order to provide exclusive resource allocation to provide a QoS for some application flows, we should apply the flow and reservation driven. And, both for scalable and efficient use of VC resources, we should apply the topology driven. Due to applying the flow, reservation and topology driven simultaneously, the system can be scalable, while providing the exclusive resource allocation to some application flows (e.g., premier service flow or voice/video flows).

3.4 Benefit of Supporting SVC

The CSR prototype system described in this paper can support both SVC and PVC. Here, the prototype system described in [CSR] could not support SVC connection to forward the IP packets. Through the support of SVC, as well as the support of PVC, we can obtain the following three additional advantages.

(1) Effective use of VPI/VCI resource
Since the VC is established on demand, CSR does not need to have un-used VCs, that is required for PVC operation. VPI/VCI resource is obtained from the ATM network, on demand, according to the actual packet arrival for each application.

(2) Reduction of required bandwidth between the CSRs
With only PVC operation, we can not dynamically modify the bandwidth between the adjacent CSRs and for each PVC. Therefore, with PVC operation, we have to prepare the bandwidth between CSRs, in advance. How much bandwidth should be prepared is subject to network engineering, however, it would be likely the bandwidth required for the busiest period in the operation. Therefore, the required bandwidth between the adjacent CSR would be larger than with SVC operation.

(3) On demand exclusive bandwidth allocation for the application flow
When the application requires the exclusive bandwidth (e.g., CBR pipe) allocation, only with PVC operation would have some difficulty.

Table 3-1. Comparison of Operational Modes

	Flow & Reservation Driven	Topology Driven
Advantage	1. Exclusive QOS allocation 2. Cut-thru across routing domain 3. On-demand cut-thru path Establishment	1. Less VC resource 2. Advanced cut-thru path establishment
Disadvantage	a. Large VC resource b. Cut-thru establishment latency c. Some hop-by-hop forwarding	a. Sharing resource b. IP processing at routing domain boundary c. Full-mesh VC d. Simultaneous cut-thru path establishment

This should be harder, when the required bandwidth by the application is not the same for all applications. Also, since we can establish SVC according to the actual appearance of application, the blocking possibility to provide the exclusive bandwidth allocation for the application could be smaller than that with only PVC operation.

4. Conclusion

Overview of CSR architecture supporting SVC, that is the only label switching architecture interoperable with the standard ATM switch and with the standard ATM links, is discussed in this paper. Default VC and dedicated-VC can be provided both by PVC and by SVC, in the system proposed in this paper. With using SVC, dedicated-VC among the CSRs can be established through the standard ATM signaling (i.e., Q.2931), on demand.

CSR system supports three cut-through operation modes, i.e., flow-driven, reservation-driven and topology-driven. Due to the support of three cut –through operation modes, the CSR system can be scalable, while providing the exclusive resource allocation for some packet flows.

References

[AFUNI] ATM Forum : "ATM User Network Interface Specification version 3.1", September 1994

[ARIS] R.Woundy, A.Viswanathan, N.Feldman, R.Boivie : "ARIS: Aggregete Route-Based IP Switching", IETF Internet-Draft, draft-woundy-aris-ipswitching-00.txt, November, 1996.

[CSR] S,Masuzawa, K.Nagami, A.Mogi, T.Jinmei, Y.Katsube, H.Esaki: "Architecture of Cell Switch Router and Prototype System Implementation", IEICE Transactions on Communications, Vol.E80-E, No.8, August, 1997.

[Esaki] H.Esaki, K.Nagami, M.Ohta : "High Speed Datagram Delivery over Internet using ATM Technology", IEICE Transactions on Comm., Vol.E78-B No.8, August 1995

[Ipsilon] P.Newman, T.Lyan, G.Minshall: "Flow labelled: Connectionless ATM Under IP", Engineer Conference, Networld+Interop'96 Las Vegas, April, 1996.

[I.150] ITU-T Recommendation I.150: "B-ISDN Asynchronous Transfer Mode", 1995

[NHRP] J.V.Luciani, D.Katz, D.Piscitello and B.Cole : "NBMA Next Hop Resolution Protocol (NHRP)", IETF Internet-Draft, draft-ietf-rolc-nhrp-10.txt, October 1996

[RFC1937] Y.Rekhter and D.Kandlur: "Local/Remote Forwarding Decision in Switched Datalink Subnetworks", IETF RFC1937, May, 1996.

[RFC2098] Y.Katusbe, K.Nagami, H.Esaki : "Toshiba's Router Architecture Extensions for ATM : Overview", IETF RFC2098, Feb., 1997.

[RFC2105] Y.Rekhter,B.Davie,D.Katz,E.Rosen,G.Swallow : "Cisco Systems's Tag Switching Architecture Overview", IETF RFC2105, Feb., 1997.

[RFC2129] K.Nagami, Y.katsube, et.al : "Toshiba's Flow Attribute Notification Protocol (FANP) Specificatiion", IETF RFC2129, April, 1997.

[RFC2205] L.Zhang, R.Braden, S.Berson, S.Herzog, S.Jamin :"Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", IETF RFC2205, October, 1997.