

RENA-CHIP Hardware Configuration Technologies

Kenji Kawai[†]

Abstract

NTT's RENA-CHIP enables network-adapter processes such as packet transferring/filtering, QoS (quality of service), and IPsec (Internet protocol security) processing to be performed in hardware. This approach enables gigabit-per-second-level processing speeds to be achieved with low power consumption.

1. Introduction

NTT's RENA-CHIP is a large-scale integration (LSI) chip that performs network-adapter processing in hardware [1]. This chip performs most of the packet processing traditionally performed by the central processing unit (CPU), making it possible for a network adapter to achieve 2-Gbit/s processing (1 Gbit/s upstream and 1 Gbit/s downstream) without its CPU being upgraded to a higher-performance one. Dedicated hardware like this can also perform tasks more efficiently than a general-purpose CPU, resulting in significant power savings. However, the application of gigabit-per-second interfaces to network adapters having low packet-processing capacity may unexpectedly result in packet loss, which may in turn degrade the quality of VoIP (voice over Internet protocol), streaming media, and other services that are sensitive to packet loss. In this regard, there are times when packets are received at short intervals even when the average transmission rate is low. In such a case, the packets must be stored in a buffer until the processing of previously received packets has been completed, and packet loss will occur if the buffer capacity is insufficient. Buffer overflows may even be caused intentionally by attacks from the outside.

2. Overview of RENA-CHIP

The chip's main specifications and functions are listed in **Tables 1** and **2**, respectively. This chip has a processing capacity of 2 Gbit/s or 3 million packets per second (3 Mp/s). This means that the functions listed in Table 2 can be carried out even under maximum load conditions for packet processing (when minimum-length packets are being received consecutively from the WAN (wide area network) and LAN (local area network) ports simultaneously). The power consumption of the chip is at most 2 W, which is low enough to eliminate the need for a fan or large heatsink. This chip can therefore contribute to the provision of compact and high-reliability network adapters.

The chip also features an NAT/NAPT (network address translation, network address port translation) table, a routing table, and a classifier table. Each of these three tables can hold up to 256 entries. Each entry consists of processing information that specifies what type of processing to perform on a packet and criteria information that specifies the conditions under which to apply it. For each table, the chip finds the entry whose conditions are satisfied by the received packet and performs the processing specified by that entry such as packet transferring, filtering, or quality-of-service (QoS) processing. The classifier table provides a fragment tracking function so that processing information can be applied to all of the packet fragments of a single packet divided up by the IP (Internet protocol) fragmentation process. Although the leading fragment of a fragmented pack-

[†] NTT Cyber Solutions Laboratories
Yokosuka-shi, 239-0847 Japan
E-mail: kawai.kenji@lab.ntt.co.jp

Table 1. Main specifications.

Process	0.13 μm 8-layer metal CMOS
Circuit scale	7.3 mm \times 7.3 mm core 96M transistors (1.7M gates + 1.4-Mbit memories) (M: million)
Interfaces	GMII x2 (WAN, LAN) 32-bit DDR-SDRAM (packet buffer) 32-bit SDR-SDRAM (packet CPU transfer) 32-bit CPU IF (control)
Clock	External input: 33 MHz (internal 66/100/133 MHz)
Voltage/power	1.5/2.5/3.3 V (core 1.5 V) about 2 W max.
Package	27 mm \times 27 mm 456-pin plastic BGA

CMOS: complementary metal oxide semiconductor

GMII: gigabit medium-independent interface

DDR-SDRAM: double data rate synchronous dynamic random access memory

SDR-SDRAM: single data rate synchronous dynamic random access memory

IF: interface

BGA: ball grid array

et must be input before the other fragments, this function can simultaneously track up to 16 packets.

In addition to the tagged VLAN (virtual local area network) and PPPoE (point-to-point protocol over Ethernet) layer-2 virtual interfaces, the chip can also handle the IPsec (Internet protocol security) layer-3 virtual interface. It incorporates logical interface tables for receiving and sending that map these virtual interfaces to up to 16 logical interfaces that conceal individual protocols. Settings in the routing and classifier tables provide these logical interfaces.

The chip supports ten QoS classes on the uplink and four on the downlink. It also enables weighted fair queuing (WFQ) to be performed in each class in addition to priority queuing (PQ) and allows an upper bandwidth limit to be specified for each class (traffic shaping). A packet of a certain class, as specified in the classifier table, is placed in that class's queue (which can store up to 1024 packets), but in times of congestion, weighted random early detection (WRED) in addition to "tail dropping" can be performed as packet-drop processing. It may also be possible to change the class of a packet slated for discard (markdown) and place it in another queue. The chip also includes a remark function that enables the Type of Service (TOS) value in the IP header to be modified according to class type, markdown, etc.

A typical configuration of a network adapter equipped with the RENA-CHIP is shown in **Fig. 1**. Here, most packets being transferred between a WAN and LAN have their processing completed within the RENA-CHIP, but those that require out-of-the-ordi-

Table 2. Main functions.

Layer 2	IEEE 802.1Q VLAN, PPPoE Up to 16 connections may be specified.
Transfer (routing table)	IPv4/IPv6 or multicasting may be specified. Up to 256 entries in combination with ARP table
Filtering/classifying (classifier table)	Conditions: input/output logical interfaces, TOS, IP address, L4 protocol, port number, TCP flag, ICMP type/code, etc. Processes: forward/discard, specify queue, remark, markdown, specify VLAN/WRED priority, etc. Up to 256 entries; automatic fragment tracking
NAT/NAPT	NAT: max. 8 entries; NAPT: max. 256 entries
QoS	Number of queues: 10 (WAN), 4 (LAN), 3 (CPU) Number of schedulers: 3 (WAN), 2 (LAN), 1 (CPU) Scheduler: 4 inputs; PQ/WFQ mode selectable Shaping is possible at module-output time; connection may be changed.
IPsec	Encryption: NULL/AES-CBC (key: 128/192/256 bits) Authentication: HMAC-SHA-1/HMAC-MD-5 ESP tunnel mode (IPv6 encapsulation); max. 16 SAs

TCP: transmission control protocol

ICMP: Internet control message protocol

ESP: encapsulating security payload

SA: security association

nary processing can be processed by the CPU. Some examples of the latter are address resolution protocol (ARP) packets, PPPoE/IPsec control packets, and packets for dynamically updating tables.

3. Configuration of RENA-CHIP

A block diagram of the chip is shown in **Fig. 2**. A packet received at the WAN port is input to a MAC (media access control) block where it is subjected to MAC termination processing including a cyclic redundancy check (CRC) to check for Ethernet-frame errors. The packet is then input to the IPsec block to determine whether it is a target of IPsec processing. If so, it is subjected to authentication and decoding and then input to the memory control block after its IPv6 encapsulated security protocol (ESP) capsule has been removed. A packet received at a LAN port is also input to the memory control block via a MAC block. On the other hand, a packet coming from the CPU is input to the memory control block from the CPU interface (IF).

The memory control block stores packet data in a DDR-SDRAM (double data rate synchronous dynamic random access memory) and inputs the data to the parser block. The DDR-SDRAM memory area is partitioned into 2-kilobyte sections, each of which

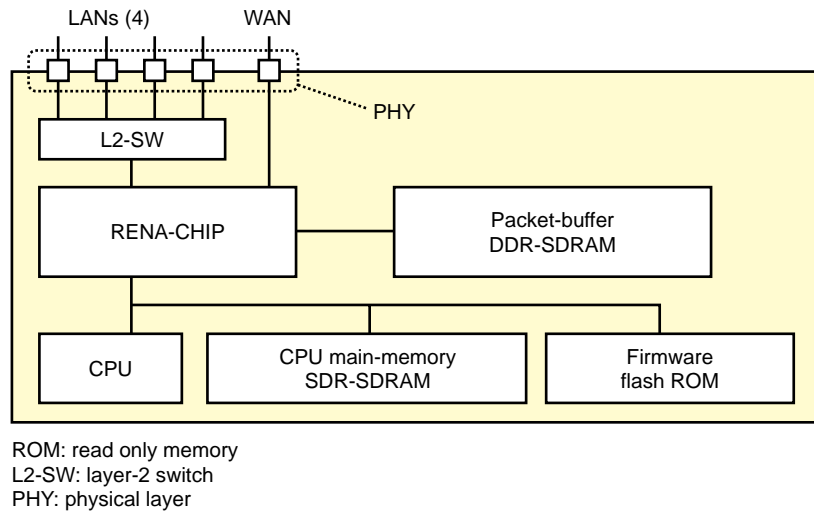


Fig. 1. Typical configuration of a network adapter equipped with the RENA-CHIP.

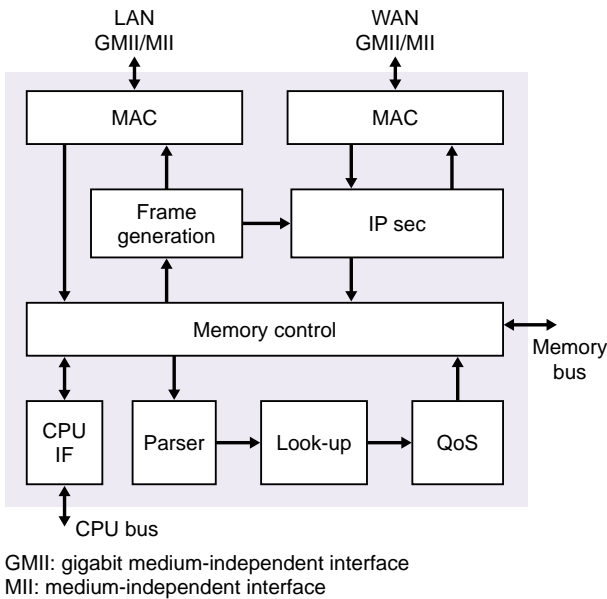


Fig. 2. RENA-CHIP block diagram.

is used to store a single packet. This scheme simplifies memory management.

Next, the parser block parses the input packet data beginning at the head of that data. Here, by following the chain of various headers from layer 2 to layer 4, the parser extracts information (MAC address, IP address, port number, etc.) from each of those headers and inputs that information to the look-up block. The flow of packet data stops here at the parser block while the flow of information extracted from the

packet headers continues on.

The look-up block searches tables such as the routing and classifier tables for entries that match the input packet. Specifically, it reads out the entries registered in a table one by one, and for each entry, it compares the information extracted from the packet headers with the criteria for that entry. If a match is found, the look-up block obtains the processing information contained in that entry. It then combines that information with other information such as packet storage location and packet length and inputs the result as packet job information to the QoS block. However, for a packet input from the CPU, the processing information prepared by the CPU is used here instead of that obtained from the parser block and look-up block. The look-up block is a key component in achieving wire-rate processing, which is a performance-related feature of the RENA-CHIP. The internal configuration of this block is described in Sec. 3.1.

The QoS block now selects a queue based on input job information and inputs that job information into the selected queue. At this time, job information about a packet slated for discard by the look-up block is discarded, tail-drop/WRED or markdown processing is performed, and job information for multicast packets is duplicated. The QoS block also informs the memory control block of the storage location of a packet whose job information has been discarded so that packet data can also be discarded. Each queue is actually a first-in first-out (FIFO) buffer for storing job information. A scheduler is also provided for each

of the WAN, LAN, and CPU outputs. This scheduler selects job information waiting to be output at the queue, establishes timing, and outputs the job information to the memory control block. The QoS block is a key component in achieving a level of flexibility comparable to software, which is a feature of the RENA-CHIP related to functionality. The internal configuration of this block is described in Sec 3.2.

Next, the memory control block reads out from DDR-SDRAM the packet data corresponding to the job information input from the QoS block and inputs this packet data together with that job information to the frame generation block except for a packet to be transferred to the CPU via the CPU interface block. The frame generation block performs packet-data processing (mainly header rewriting) based on the job information and then inputs the packet and job information to the IPsec block for a packet received from a WAN port, to the LAN-side MAC block for a packet received from a LAN port.

The IPsec block now determines, based on job information, whether the packet is a target of IPsec processing. If so, it attaches authentication data, performs encryption processing, and encapsulates the packet by IPv6 ESP. If not, the IPsec block passes it on unchanged to the MAC block, which performs MAC termination processing such as attaching CRC data. Once this has been completed, the packet is finally transmitted as an Ethernet frame.

3.1 Look-up block

A block diagram of the look-up block is shown in **Fig. 3**. A new configuration was chosen to enable look-up to be accomplished for all packets even when minimum-length packets are being received consecutively from both WAN and LAN ports simultaneously.

The look-up process begins with the receive logical-IF look-up circuit (20 entries), continues with the NAT/NAPT table (NAT: 8 entries, NAPT: 256 entries), routing table (256 entries), and classifier table (256 entries), and ends with the fragment tracking/look-up circuit (16 entries). Here, each of the 256-entry NAPT/routing/classifier tables has six comparators, each of which can compare two entries per cycle of a 66-MHz clock. Simultaneous look-up for up to six packets can be performed for each of these three tables. This configuration achieves the performance required for wire-rate look-up (3 Mp/s). In addition, the parallel arrangement of comparators without a parallel arrangement of tables reduces the amount of memory within the RENA-CHIP to two-thirds that required when a parallel arrangement is used for both comparators and tables.

3.2 QoS block

A block diagram of the QoS block is shown in **Fig. 4**. Since the processing of out-of-the-ordinary packets by hardware will naturally increase the circuit

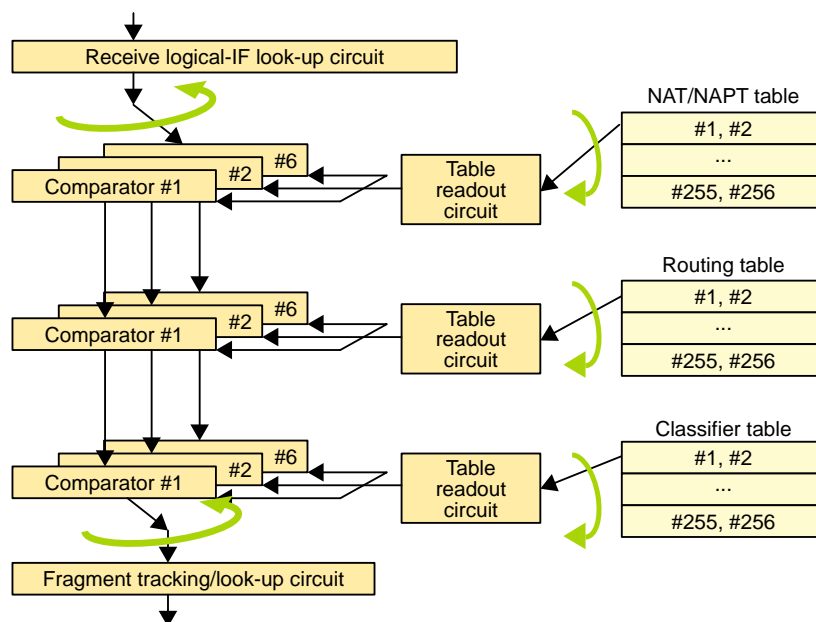


Fig. 3. Configuration of look-up block.

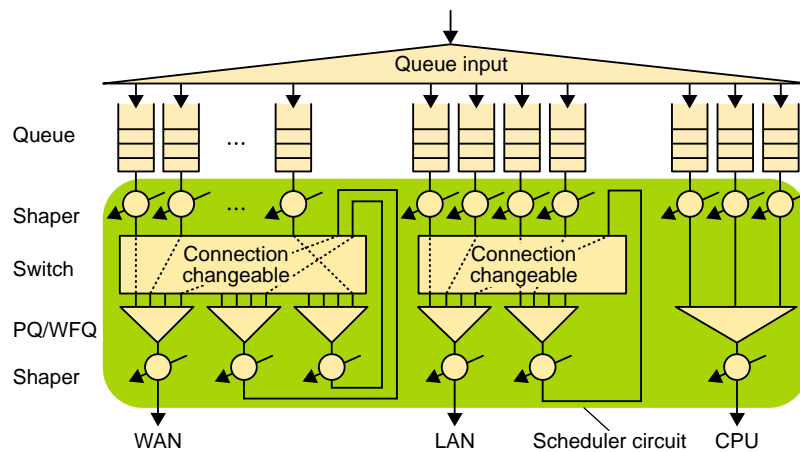


Fig. 4. Configuration of QoS block.

scale without an appreciable rise in performance, the RENA-CHIP transfers such packets to the CPU for processing by software. The volume of such packets is normally low enough that they can be handled by the CPU, but it could exceed CPU capacity, for example, as a result of an attack from the outside. To ensure that important packet processing will function as it should under any conditions, the RENA-CHIP therefore incorporates a dedicated QoS circuit for CPU transfer. In addition to classes based on classifier settings, classes that correspond to the factors leading to an abnormal packet and its transfer to the CPU may also be established.

It is desirable that the configuration of the QoS block be flexible in accordance with the characteristics of services used by users. To this end, a new type of scheduler circuit has been developed so that QoS with the same flexibility as that achieved by software can be achieved by hardware. This scheduler circuit has a switch circuit that can change the combination of multiple 4-input schedulers. Each of these 4-input schedulers features PQ/WFQ mode selection and traffic shaping at the time of output. This flexible configuration enables various types of operations such as scheduling that combines PQ and WFQ and the setting of an upper limit on the total bandwidth of multiple classes. Achieving all of the above by hardware also makes for high-accuracy operations regardless of the processing volume.

4. Future developments

Topics for future study include expanding the RENA-CHIP for business systems by increasing the number of table entries and for home appliances by enhancing security functions. Although the use of hardware for packet processing has great advantages, the subsequent addition of new functions and modification of existing functions can create difficulties. This requires further investigation.

Reference

- [1] K. Koike, "Technical Trends of Network Processors and the RENA-CHIP," NTT Technical Review, Vol. 4, No. 9, pp. 12-16, 2006 (this issue).



Kenji Kawai

Senior Research Engineer, First Promotion Project, NTT Cyber Solutions Laboratories.

He received the B.E. and M.E. degrees in electronic communication engineering from Waseda University, Tokyo, in 1989 and 1991, respectively. He joined NTT LSI Laboratories, Kanagawa, in 1991. Since then, he has been engaged in research on the design of high-speed LSIs.