Pushing packet performance

Leading edge chipset powers 100Gbit/s IP router platform.
By **Roy Rubenstein**.

While the 100Gbit Ethernet interface will become an IEEE standard this summer, chipmakers have anticipated this development for some time.

Alcatel-Lucent's in house asic team started design of the FP2 chipset in 2005, with the goal of supporting packet processing at 100Gbit/s line rates. The chipset is now an integral part of its 7750 IP router and 7450 Ethernet switch platforms. And the firm is already showing a 100Gbit/s line card using the FP2 to customers before its official launch.

The FP2 has been designed with additional processing resources to accommodate the expected development of networking protocols during the chipset's design and deployment.

Alcatel-Lucent's 7750 service router platforms carry IP traffic in the core and edge of the network. According to Ken Kutzler, vp engineering with Alcatel-Lucent's IP division, while a core router handles more traffic, the packet processing tasks performed by an edge router are more taxing.

A core router transports packets at high speed, with networks engineered to minimise the buffering of traffic in the core as much as possible. In contrast, an edge router performs more packet 'touches', such as packet queuing or processing access control lists, that determine which users or processes can access what items. Kutzler said this ensures the end user receives the level of service signed up for with the operator.

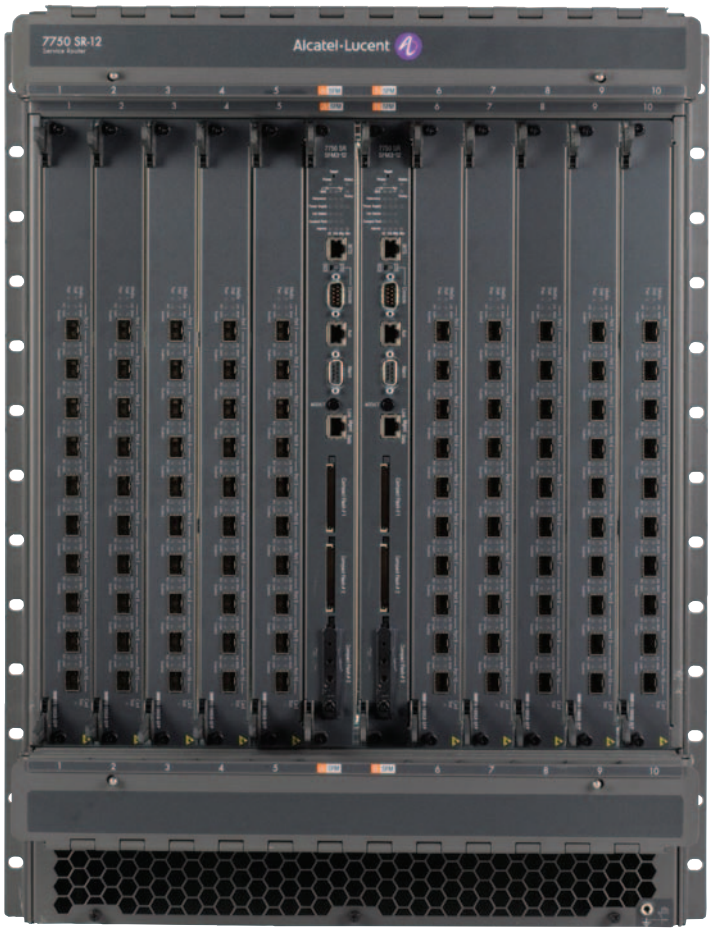
Houman Modarres, director product marketing with Alcatel-Lucent's IP division, uses a postal analogy to contrast the two tasks: sorting the different packets and their priority in terms of transportation is handled at the network edge (the sorting office), while the various delivery options (postal transport) are performed by the core routing.

Edge routers must now perform several applications, each with its own characteristics and hence packet flow requirements. For IPTV services delivered to homes, subscriber management – involving such tasks as user authentication, authorisation, accounting and security issues – is one such task.

For enterprises, supporting IP virtual private networking is common. Operators provide a variety of business connectivity services, ranging from simple point to point virtual private networks (VPNs) to more sophisticated IP VPN connections that require a greater level of packet processing.

Mobile broadband data, another growing source of traffic with its own requirements, now spans from second generation GSM to the emerging Long Term Evolution (LTE) standard. "The FP2 is designed to scale without compromising these various services," said Modarres.

The 7750 SR-12 IP router has a switching capacity of up to 1Tbit/s

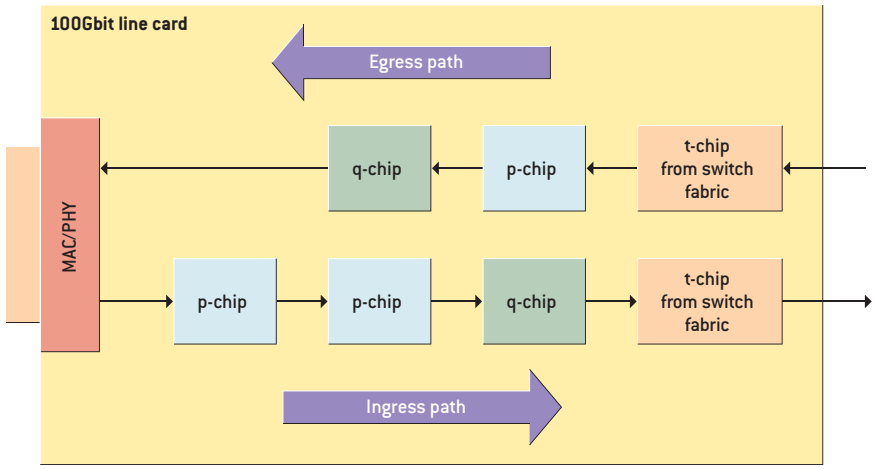


The FP2 chipset comprises three 90nm process ics: a 100Gbit/s network processor dubbed the p-chip; a traffic manager, called q-chip; and an interface to the router's switch fabric, called t-chip.

In contrast, the chips in ZTE's latest 100Gbit/s network processor and traffic manager are implemented using a 65nm process. The ZXRIC SF600 switching fabric chip features 600Gbit crossbar switching capacity and a three stage CLOS chassis technology. The ZXRIC PFE packet forwarding processing chipset supports 100Gbit wire speed forwarding, while the ZXRIC TME multipolicy traffic management chip supports a five level hierarchical Quality of Service mechanism

The Chinese vendor announced its ZXR10 T8000 core router – said to support 1million users concurrently – last September and also uses the

Fig 1: Block diagram of Alcatel-Lucent's 100Gbit line card



chips in its metro and edge network platforms.

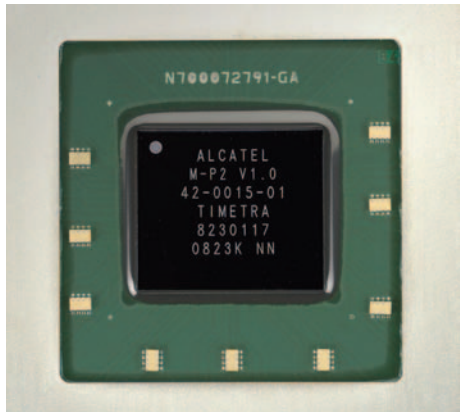
The p-chip's role is to inspect packets and perform the look ups that determine where the packets should be forwarded. This network processor comprises 112 cores arranged in 16 rows of 7 columns – with the cores clocked at 840MHz. This contrasts to the 0.15μm process first generation 10Gbit/s FP1, whose 30 cores are clocked at 190MHz.

Each processor is programmable and has its own microcode. Each row is programmed with a given task before being fed data packets, while packets from the same flow are sent in order. Thus, one row could be implementing multiprotocol label switching (MPLS) protocol tasks used to direct tagged packets between nodes, while another row could be performing IP routing.

The p-chip has spare capacity to accommodate new protocols. "The platform has enough headroom to allow new factors to be added," said Kutzler, "and entire rows are unused to allow for such factors."

The 100Gbit/s card is fed using either ten 10Gbit/s optical transceivers or a 100Gbit Ethernet CFP transceiver form factor. On the ingress path (see figure 1), two p-chips are used in series, along with the q-chip traffic manager. Though a single p-chip can process 100Gbit/s line rates, a second p-chip is included to provide additional processing for the broad range of current and future edge router processing tasks.

The p-chip determines a packet's class and the quality of service it requires and tells the q-chip traffic manager in which queue the packet is to be placed. The q-chip handles the packet flows and makes decisions as to how packets should be dealt with, especially when congestion occurs. The q-chip features four in house designed risc processors to accommodate new requirements. "It's for flexibility," says Kutzler, "You can never fully anticipate the traffic engineering and traffic management enhancements that will be required over time and, to future proof our platforms over time, we always plan our designs conservatively."



Part of the FP2 chipset, the p-chip is a 100Gbit/s network processor fabricated on a 90nm process. Other members of the chipset are the q-chip traffic manager and the t-chip interface device.

The 100Gbit/s line card uses 1.5Gbyte of RLDRAM for buffering. This allows some 200ms of traffic to be held within the router, equivalent to 16m 64byte packets. According to Kutzler, the q-chip buffers packets efficiently, based on their size. "We don't reserve big chunks [of memory] for every packet, whether it is 64 or 512byte," he said.

The interface between the two chips is based on a proprietary serdes. Kutzler said: "We use many parallel channels; nothing [available] could handle the flows we wanted – the backpressure and the signalling." Backpressure is a mechanism used to temporarily throttle traffic to minimise – and, ideally, avoid – the need to drop packets. This is because packet bursts can, over short durations, exceed the allotted capacity for a link and is a natural element of packet networks, regardless of link speeds.

One of the main concerns expressed regarding IP routers is their huge power consumption, which continues to grow. Alcatel-Lucent's 100Gbit/s line card is rated at 4W/Gbit, whereas its predecessor 50Gbit/s card consumes less than 6W/Gbit.

With regard the FP2, circuitry such as the processors are not clocked if they are not being used to reduce power consumption. Having an integrated 100Gbit/s design also saves power, says Modarres. For example, when using four 10Gbit/s network processors for a 40Gbit/s stream, each requires its own memory for a copy of the route look up tables. "By having one chip, a single memory can be used, integration cuts inefficiency drastically, minimising the need for segmentation and duplication," he says.

Given the time it takes to develop the packet processing chipset, does that mean that Alcatel-Lucent is well advanced in its next design?

Modarres answers by points out that the 10Gbit/s FP1 was available commercially when the first 2.5Git/s merchant network processors arrived in 2003, and that the FP2 sampled in 2008, a year after the first 10Gbit/s network processor/traffic manager ics.

He wouldn't say more but, speculating, this suggests an FP3 chipset – rated anywhere between 400 and 1000Gbit/s – will sample in 2012 and will follow integrated 100Gbit/s network processor traffic managers that are becoming available this year.