Hyperchip Inc.

3391 Griffith Ave, St. Laurent, Quebec H4T 1W5
(514) 906-2400
(514) 906-2500

Flexible Network Processing and Embedded Routing for NGI Equipment

The keys to flexible network processing for Next Generation Internet (NGI) equipment are hardware-programmable network processors for sufficient compactness and power efficiency in full Layer-3+ network processing, advanced buffering and search techniques for carrier-grade IP QoS, huge forwarding tables, route-class statistics and large ACL spaces, and an industry standard form factor that allows incorporation in a wide variety of NGI equipment. Each of these will be covered in turn, followed by examples of how this can be used to bring routing and other deep network processing to a variety of equipment and new applications beyond traditional routers through embedded routing.

H40G AMC Network Processor Cards, at one Watt and 7.5 cubic centimeters per Gb/s, have a small enough footprint to ***embed routing*** in other equipment, adding routing to new applications including:

- GPON access equipment,
- Wireless back-haul equipment
- Reconfigurable Optical Add-Drop Multiplexers (ROADMs)
- Optical cross-connects
- Ethernet switches
- Fibrechannel, InfiniBand & SCSI Storage
- PCIe Computing

Embedding routing and other higher-layer functions into lower-layer equipment amplifies the savings from the cost-effective Network Processor chip set by also minimizing within-PoP optical links.

Even when embedded with, rather than within, other equipment, H40G routing is five times more power efficient and ten times more compact than traditional routers at a fraction of the cost. New router applications this enables include:

- MicroTCA Routers - first router in MicroTCA industry-standard form factor
- ATCA Routers - first 40 Gb/s per slot routers in an industry-standard form factor

The H40G Network Processors can, in addition to routing, use their flexibility and ample reprogrammable logic, CAM and RAM resources for other higher-layer functions such as:

- deep packet inspection,
- firewalls,
- cryptography and compression
- law enforcement applications and other in-skin monitoring.

Embedded Routing:
Embedded Routing will be commercially successful by *not* competing head-to-head with stand-alone routers such as those from Cisco and Juniper. Instead, embedded routing will succeed by adding routing functionality where routers cannot go today.

Cisco's and Juniper's big core routers are the mainframes of the routing world, and their edge and aggregation routers are the minicomputers. Microprocessors succeeded not by competing head-to-head with mainframes and minicomputers and their massive installed bases of legacy software. Instead, microprocessors succeeded by *embedding computing* in much smaller devices with initially much simpler functions.

Similarly even though the heart of the H40G is a network processor as capable as (and more flexible than) the network processors in Cisco and Juniper routers, the H40G will not compete with those routers. Instead the H40G will succeed by *embedding routing* in other devices, enabling new applications where traditional routers cannot go.

The H40G Network Processing AMC is an entire data plane for 40 Gb/s of routing, with self-contained traffic management and switch fabric, on a pocket-sized card in an industry-standard form factor. The H40G NP AMC lets routing be embedded ***into*** other equipment, which eliminates the interface cards that would have been on both ends of links to a separate router, savings additional cost, space, and power. Rather than spend hundreds of watts to send packets to a dedicated router that consumes over 1000 Watts, an H40G AMC can be added to provide routing and/or other higher-layer functionality (security, Deep Packet Inspection, etc.) in-skin. This consumes only one watt per Gb/s (25x less power), and takes less than 7.5 cubic centimeters per Gb/s (200x less space), and costs ten times less than a stand-alone router.

Even when coupled with off-the-shelf optical interfaces in its own chassis, an H40G as a router is more analogous to a dedicated industrial control computer than to a general-purpose computer with a huge set of applications. An H40G chassis that can be held with one hand is small enough (tem times smaller), and low enough in cost and power consumption (fives times lower) to be embedded ***with*** other equipment, with a dedicated (and thus simpler) set of control-plane software features, avoiding the need to have a separate much lager and more expensive general-purpose router.

Although initial H40G applications will typically need less complex control planes, the H40G is compatible with general-purpose control planes as well. And while traditional routers only run control-plane software from a single company, the H40G has an open interface for hosting control planes, making it easy for partners who have their own control planes to embed the H40G with or into their equipment. The Open control plane interface can even run control plane software from more than one company simultaneously, allowing mixing and matching best-of-breed control-plane features (for example, this has been used to add IPv6 from an off-the-shelf Nexthop control plane to IPv4 from Hyperchip's resilient and scalable IPv4 control plane).

Hardware Programmable Network Processors:

Data plane functions of higher-level network equipment have repeatedly evolved to increase performance and flexibility while reducing cost, space, and power consumption. CPUs were first replaced by fixed-function ASICs, but these lacked the in-field programmability needed to keep up with changing network needs. Software-programmable network processing ASICs were developed to regain flexibility, but off-the-shelf NPs have lacked the performance for extensive processing, and proprietary versions cost tens of millions of dollars per generation to keep up with Moore's law.

But software is not the only way to achieve programmability. Hyperchip pioneered a more efficient way, the hardware-programmable network processor, starting in 1999. Hardware programmable network processors capitalize on the inherent programmability of FPGAs, eliminating the need for instruction fetch and decode pipelines, register sets, programmable ALUs, branch/address management, translation buffers, contention management, out-of-order processing, etc., which give software-programmable CPUs their performance but cost die area and power. Commercial Network Processors dispense with some of these, but pay the price of increased parallel programming headaches. And all software NPs need to dynamically translate each instruction to figure out what to do, and then have a flexible processor do it, while in a hardware-programmable NP each stage is set up to simply do the proper operation.

Hardware-programmable network processors are ideally suited to deep pipelines. Each pipeline stage executes the equivalent of one instruction, so each pipeline stage is very simple, and can be as wide as that particular operation needs. Hyperchip has pioneered an FPGA-optimized pipeline architecture that simplifies deep pipelining where all stages process their individual operations in parallel without requiring parallel software.

A hardware-programmable NP running at under 200 MHz can deliver 40 Gb/s of full-router network processing from a single pipeline. In contrast Cisco's flagship 40 Gb/s SPP network processor requires *188 cores* running at 250 MHz to deliver 40 Gb/s of network processing, and burns more power just for the NP[i] than Hyperchip's entire chip-set (including TCAM, table RAMs and buffer RAMs as well as the NP) consumes.

Another way that hardware-programmable NPs beat software NPs is that FPGAs have a lithography lead of almost two generations, which is a four-year technology lead. While a decade ago FPGAs lagged ASICs, once logic passed DRAM as a driver of Moore's law, FPGAs became the ideal vehicle for pushing lithography because they are the most regular and repetitive logic chips. At 90 nanometers FPGAs passed most ASICs, and at 40 nanometers FPGA lithography passed even Intel's leading edge 45-nannometer CPUs for the first time, while most software NPs are at 90 nanometers.

Another advantage is ease of programming. Massively multi-core software NPs are not ordinary CPUs, and high-performance parallel programming is a notoriously hard problem. Software NPs compound that challenge with unique, highly-customized instruction sets that only a few people in the world can obtain maximum performance with. In contrast, an FPGA-based hardware-programmable NP is a single pipeline, so a high-tech hub has thousands of fluent FPGA programmers who can add features to it.

In addition to being more compact and power efficient any given generation, hardware-programmable NPs have a huge advantage over software NPs in the cost of progress. Software NPs bear the ever-increasing cost of keeping up with Moore's law; the average ASIC cost rose from $4 million at 180 nm to $10 million at 130 nm and $25 million at 90nm[ii], and high-performance network processors are more complex than average ASICs. In contrast, with hardware-programmable NPs the Moore's Law cost is borne by an FPGA supplier who has already reached 40-nanometer lithography and will amortize that cost over billions of dollars in FPGA sales.

Far from embedded routing competing with Cisco and Juniper, the overwhelming advantages of hardware-programmable network processor goal of hardware-programmable network processors may make them customers. Even for Cisco, spending tens and soon hundreds of millions of dollars on each chip each generation to keep up with Moore's law is onerous, and for smaller Juniper this is a massive expense, and it will become unnecessary. Similarly main-frame giant IBM has become a customer for Intel microprocessors, first in low-end systems and later even in high-end servers.

Advanced Buffering and Search Techniques:
Some network processors offer high bandwidth performing only very limited processing on packets, or by only supporting tables or buffers small enough that they fit in internal memory. But this greatly limits flexibility by restricting the operations that such network processors can perform.

In contrast, Hyperchip has taken the opposite approach, developing advanced buffer access algorithms that support carrier-class QoS with 100 millisecond DRAM buffers, and without power-hungry SRAMs for buffer management. A novel HCAM lookup supports huge forwarding tables without cascaded TCAMs, with up to 3 million routes in the standard chip-set and 6-million-route support available. Large external statistics tables are also supported, and these can be used for additional table lookups when such statistics are not needed. And over 90% of the TCAM is free for ACLs, allowing tens of thousands of ACLs in a small TCAM and over 100,000 ACLs with a large TCAM.

SONET, Ethernet, IPv4, IPv6 and MPLS LER/LSR are all supported, and all processing is full line rate even for the shorted IPv4 packets of SONET. ACLs can contain over 30 Layer-2 to Layer-4 fields, and are 160 bits for IPv4 and 320 bits for IPv6. H40G network processors even have wire-speed private filter lookups to protect control planes from DoS and other attacks.

Compact and Convenient Industry-Standard Form Factor:
The highly optimized chip set and the low power consumption allow packing the H40G Network Processor into a very small space. The H40G is an ingress+egress implementation packed onto an industry-standard AMC card form factor. This can support 40 Gb/s even for SONET traffic, and can support Ethernet at up to 60 Gb/s without over-subscription.

While part of the ATCA family of standards, the AMC form factor is also used in some proprietary chassis. The ATCA-based H40G can thus be used in a huge variety of chassis, from 1U microTCA chassis to 4U and 6U microTCA chassis, and to 2-slot, 5-slot, 14-slot and 16-slot ATCA chassis, and also in proprietary chassis of almost any size.

The H40G is also available as an H40G MicroTCA hub, combining chassis management, switch fabric, traffic management and network processing for a whole MicroTCA chassis in a single MicroTCA hub slot.

The H40G AMC also fits in an AMC slot on an ATCA line card or fabric card. This allows one H40G (or two for redundancy) to provide switch fabric, traffic management, and network processing for an entire chassis full of moderate-bandwidth line cards, or allows an ATCA fabric card to add up to 160 Gb/s of network processing for a chassis.

With off-the-shelf AMC-based interfaces available from Ethernet to GPONs to storage to CPUs, a complete H40G-based MicroTCA or ATCA product can be rapidly assembled while achieving industry-leading density, capabilities and power efficiency. The ATCA / MicroTCA / AMC ecosystem is growing rapidly, and Hyperchip is happy to contribute to that growth with the H40G.

The industry-standard form factor and the highly-optimized chip set significantly reduce the cost of the H40G. A full 40 Gb/s router, for example will cost less than half as much as a traditional proprietary-chassis 40 Gb/s router, as well taking 10 times less space and consuming five times less power.

The ability to embed an H40G AMC into other equipment can offer even larger saving, as it eliminates the optical interfaces that would run between separate equipment in addition to the savings already discussed. An embedded H40G takes 200 times less space, 25 times less power, and is a fraction of the cost of a traditional separate router. This enables adding embedded routing to other equipment to create new applications.

Examples (these will be fleshed out after being agreed on):

1) H40G Green NGI router. Basically the highlights of the data sheet and the competitive comparables, compress down to a page.

2) H40G GPON. 50 Gb/s of GPON interfaces, or 60 Gb/s with an in-skin or external CPU. Quick spec-sheet type of highlights including routing, local content insertion, etc.

3) H40G Wireless Backhaul with routing and uplink. Quick spec-sheet type of highlights.

4) H40G ATCA Ethernet Switch (1600 Gb/s) with 160 Gb/s (4* H40G AMC) of routing.

5) Optical cross-connect with 160 Gb/s (4 * H40G AMC) of routing plus deep-packet inspection added, in a proprietary (Tellabs-clone) chassis.

Future directions:

Once volume is established on the H40G (chipset, AMC and hub), this product family will be extended in two directions. The Hyperchip NP pipelines will be split into back into separate ingress and egress NPs, allowing 100 Gb/s support in the latest FPGAs (on two AMCs). Hyperchip's scalable traffic managers will also be implemented in the AMC form factor, allowing scaling of ACTA systems at 40 Gb/s and 100 Gb/s per slot, all the way to multi-chassis systems with a non-blocking fabric.

Example:

6) Scaling and ATCA chassis at 100 Gb/s. (The PowerPoint slide with a few words).

---

[i] On the Use of General-Purpose Multi-Core Processors in Networking Devices
Patrick Crowley and Jon Turner, Washington University Department of Computer Science & Engineering
[ii] IBS Handel-Jones Report 2004