EE Times Connecting the Global Electronics Community

designlines WIRELESS & NETWORKING

News & Analysis

Startup Tips 12.8-Tbit Switch

Rick Merritt

3/14/2017 08:45 AM EDT

Post a comment

Innovium leapfrogs Broadcom Tomahawk

SAN JOSE, Calif. — Startup Innovium aims to leapfrog Broadcom in Ethernet switch chips with a part that sports twice the density at similar area and power consumption. The company also closed a Series C round for $38.3 million, bringing its total funding to $90 million, and named a veteran data center executive to its board of advisors.

Innovium's TeraLynx packs a 12.8-Tbit/s switch into a single 16-nm die — twice the aggregate bandwidth of Broadcom's Tomahawk II announced in October. The chip sports 70 Mbytes of buffer memory, more than 50% greater than the buffer in Broadcom's switch.

In addition, TeraLynx supports 50-Gbit/s serdes to drive 200- to 400-Gbit Ethernet links, rates that current chips can't handle. The startup expects to cull out lower-yielding parts to create a product line that spans from 3.2 to 12.8 Tbits/s.

Innovium's co-founder and chief executive, Rajiv Khemani, declined to give details of the architecture for which the company has filed 30 patents. He did say that it uses "a clever way of packing information … some say that we must use compression, but if you do all kinds of compression, you have to do it in a way that doesn't impact latency," he said, noting that TeraLynx sports less than 350-nanosecond port-to-port latency independent of protocols used.

Besides supporting more throughput, the chips' large number of ports opens a door to topologies beyond widely used Clos networks used in today's bandwidth-starved data centers. In addition, Innovium claims that it has unique enhancements for tracking data flows and packet errors, a hot button in an era of software-defined networks.

The chip faces several emerging rivals, including Nephos, a spin-out of Mediatek that debuted last week, claiming a valuation of more than $300 million. Cavium now claims design wins for its XPliant chip with Arista and Brocade. China's Centec Networks, Marvell, Mellanox, and startup Barefoot Networks are also offering competing chips seeking a slice of the Ethernet switch chip market that Broadcom dominates.

As a result of the rising competition, switch prices are expected to fall from recent highs of $60 per 100-Gbit/s port, or almost $2,000 per chip in 2016. "We think that will fall to $36/port in 2020 … the competitive landscape is certainly heating up," said Bob Wheeler, principal network analyst for The Linley Group.

The competition comes at a time when data centers at web giants such as Amazon, Facebook, and Google are strapped for bandwidth, leading one veteran to call last month for quickly setting a standard for 800-Gbit/s Ethernet.

Next page: TeraLynx 'very close' to tapeout

CAPABILITIES	28nm SWITCHES	16nm SWITCHES	INNOVIUM (16nm)	INNOVIUM ADVANTAGES
Performance	3.2 Tbps	6.4 Tbps	12.8 Tbps (3.2, 6.4, 12.8)	Up to 6x for CLOS architectures
Power (W)	~ 0.6X	X	< X	2x better perf/watt
Buffer Size	16 – 24MB	20 – 40MB	70 MB	Superior network quality; fewer packet drops
Latency	450 – 2000+ nsec	400 – 600+ nsec	< 350 nsec	Lower latency
50G SerDes, 200/400GbE	N	N	Y	Higher bandwidth & scale, Lower cost/bit
Visibility & Analytics	Basic	Enhancements	Most Advanced	Must have for next-gen data centers
Programmability	Protocols (some)	Protocols (some)	Protocols & I/Os (line-rate)	Line rate programmability

TeraLynx aims to leapfrog Broadcom's 28- and 16-nm offerings. *Click to enlarge. Source: Innovium*

TeraLynx 'very close' to tapeout

One of the key questions for Innovium is whether it will be able to deliver TeraLynx. The chip, which is "very close" to tape out, is expected to sample in the fall and be in volume production early next year.

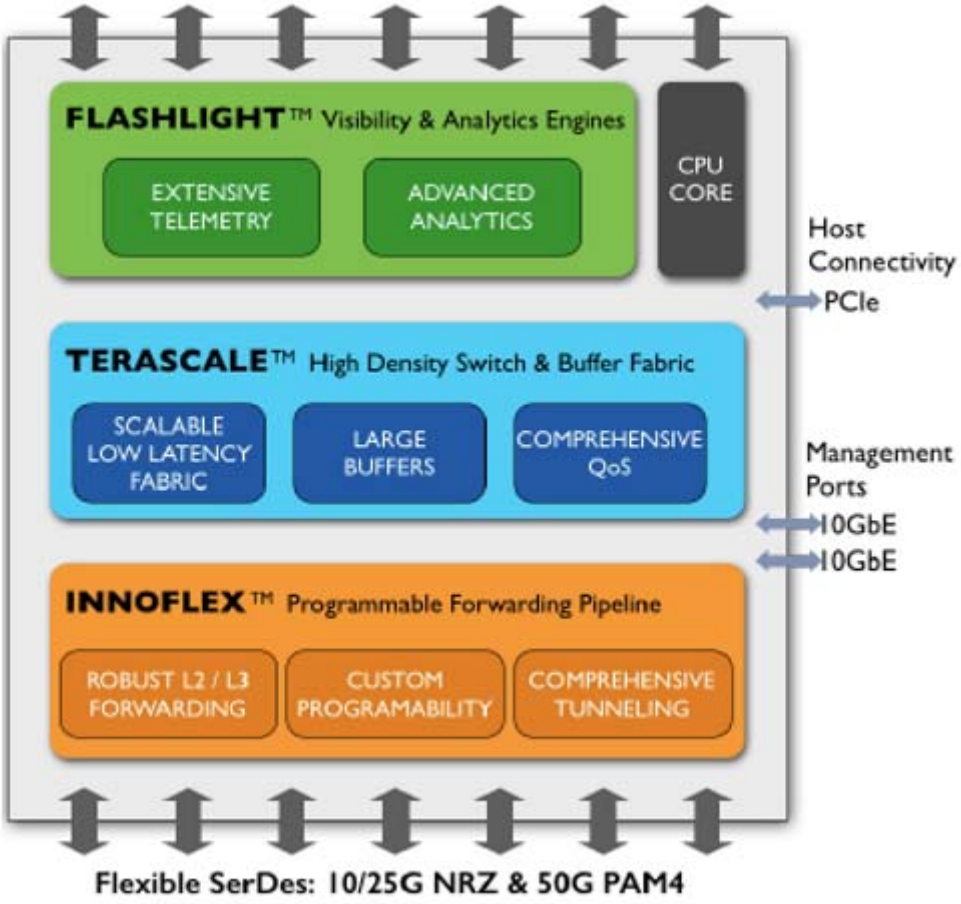
"In the last decade, several companies have defined and announced switching products, but execution missteps have proved costly in terms of market adoption," said Wheeler of Linley Group.

"We believe that we have a decent time-to-market lead and an optimized architecture — others will need to go to 7-nm technology," said Khemani.

Key to Innovium's success so far was attracting two top engineering managers who helped design Broadcom's Trident and Tomahawk switches — Puneet Agarwal and Mohammad Issa. Khemani and Agarwal were former classmates at the Indian Institutes of Technology who stayed in touch over the years.

After leaving his post as chief operations officer at Cavium in 2013, Khemani was working as an investor. He found out that both Agarwal and Issa were interested in forming a startup. Khemani "sought funding on the basis of our resumes … [and we] got funded without having a product defined," he said.

Broadcom sued the startup a few months after it was formed in early 2015. The companies settled in May 2016 for undisclosed terms, leaving open the question of whether Broadcom might sue again on patent infringement once a product becomes available.



TeraLynx claims advantages in density, programmability, and analytics. (Source: Innovium)

Growing public cloud services from the likes of Amazon, Facebook, and Rackspace, along with web giants such as Facebook are driving the need for Ethernet speed. Innovium's Khemani said that he expects large data centers to consume well over half of the world's servers in the next few years.

In a step toward winning that business, the startup named to its advisory board Yuval Bachar, a principal engineer for LinkedIn's data centers. Bachar previously led network system design at Facebook.

"Innovium's TeraLynx is a game-changer in the industry … [that can be used] in all tiers across the data center from top-of-rack to core fabric," Bachar said in a press statement, citing the chip's port and buffer density.

Other members of the startup's advisory group included Pradeep Sindhu, founder and vice chairman of Juniper Networks, and representatives of software companies Metaswitch and VMware. Software partners are key at a time when data centers are driving networking functions off of ASIC-heavy routers to x86 servers.

Initially, Innovium will use the P4, a network programming language launched by rival Barefoot Networks. Innovium uses P4 with extensions of its own for sending switch data upstream for aggregation. However, it will use JSON (JavaScript Object Notation) for programming its forwarding plane. "P4 is still immature but we can add support for it there later if needed," said Khemani.

Overall, analysts are upbeat about Innovium's chances.

"Whereas every other new entrant is merely matching Broadcom on bandwidth, Innovium is leapfrogging Broadcom with a next-generation design. If they execute to schedule, they

should have a significant lead in enabling single-lane 50 GbE as well as 200 GbE and 400 GbE," said Wheeler of Linley Group.

"The Ethernet Switch Market in the data center is undergoing rapid transition as the demand for higher speeds by cloud providers is pushing the market quickly toward 400G," said Alan Weckel, founder and technology analyst at 650 Group. "12.8-Tbits/s fabrics will start getting deployed by mid-2018 and become commonplace in the data center by the end of the decade as the cloud becomes a larger segment than the enterprise market."

— Rick Merritt, Silicon Valley Bureau Chief, EE Times

Related posts:

- Startup Taps TSMC to Attack B'com
- Cavium Attacks B'com in Switches
- Startup Aims to Simplify Networks
- Data Center Calls for 800GE Spec

More Related Links

LinkedIn Group Preps Server Specs

Google Revs Faster TPU

HPE Builds Big Data Server

Optical Industry Needs Automation

Nvidia Rolls Volta GPU For 'AI Revolution'