



Broadcom Trident 3 – Programmable, Varied And Volume

Greg Ferro June 14, 2017

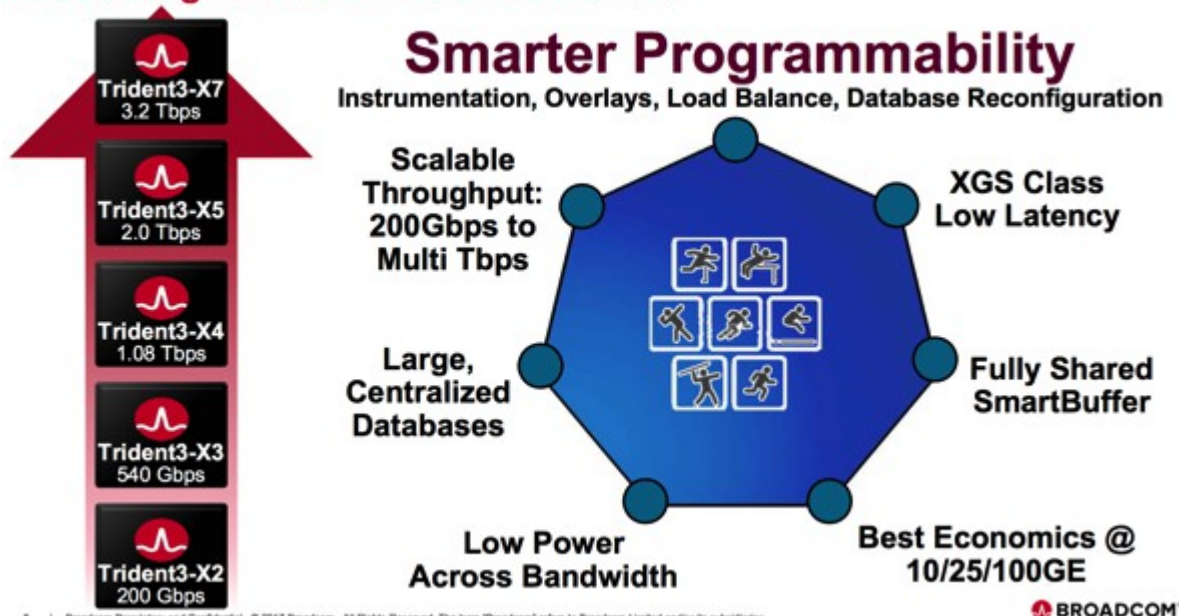
Broadcom is announcing the Trident 3 ASIC.

Takeaways:

- Trident 3 ASICs have increased performance.
- 5 versions of the Trident 3 ASIC target wireless, campus, and data center Ethernet networks.
- 5-stage programmable pipeline increases the flexibility for packet operations in the device application stack.
- Support for ISSU operations (up to software owner to implement).
- Improved access to CPU/Memory via PCIe bus upgrades (32Gbps BiDir) & other features.
- Switch hardware pricing for 32 x 100GBE to under \$3000 per unit.

Let's summarize the Trident 3 announcement from Broadcom in an image:

Introducing Trident 3: A New Generation



Broadcom says its ASIC design is a series of tradeoffs in seven dimensions and that each ASIC model represents a particular direction. Five different ASIC versions means that switch makers have a wider choice of silicon to make switches for different types of networks. This means that Broadcom is reaching into the campus and wireless switch market while advancing its position in the data centre.

Big In The Data Center

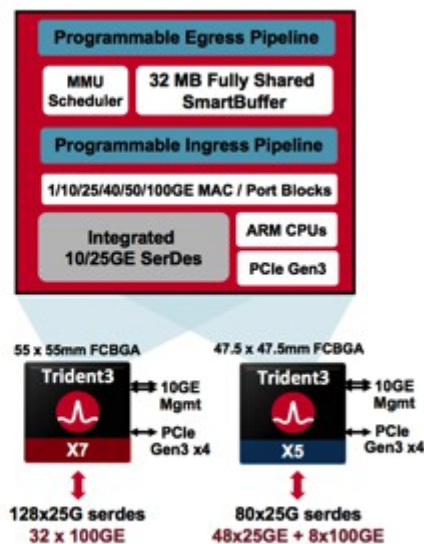
One of Broadcom's biggest customers has been the mega-scale data centres (MSDCs), which continue to increase their participation in the supply chain for their hardware.

- Trident 3 X7 ASIC at 3.2 Tbps, up to 128 port for 100G only spine networks.
- Trident 3 X5 ASIC at 2.0 Tbps, up to 128 port for 25/50G Top-of-rack deployment.
- Integrated SerDes reduces the cost of boxing the ASIC.

It seems to me that this slide highlights technologies that would be important to MSDCs. Today, the TCAM tables are often statically defined and may require reboot to modify. Some of this is due to the switch applications as coded by vendors but there are limits in the hardware. The Trident 3 removes most the hardware limitation (some issues around timing and refresh remain but should easily solved).

Sampling Now: Trident3 X7 and X5 Devices

- **3.2 and 2.0 Tbps Ethernet Switching**
 - Up to 128 front panel ports
 - Ideal for 25/100GE Data Center Top-of-Rack and Fixed/Chassis Spine
- **Optimized 25 / 50 / 100 GE Configurations**
 - Configurable 10/25G LR serdes
 - Integrated CL74, CL91 and CL108 FEC on all ports
- **Programmable Architecture for Software-Defined DC**
 - ✓ Future-proof support for extensible overlay encap/decap
 - ✓ Flexible Database reconfiguration with large table scales
 - ✓ Dynamic Load Balancing on LAG/ECMP/HG
 - ✓ Enhanced control-plane (PCIe Gen3)
 - ✓ Enhanced MPLS segment routing
 - ✓ Enhanced BroadView™ v2 Instrumentation
- **Ease of System Design**
 - Low Power 16nm design
 - 55x55mm and 47.5x47.5mm FCBGA Packages, 1mm pitch



The load balancing of elephant flows in larger networks is a serious concern. Enterprise networks increasing their deployments of software-defined storage platforms will be pleased to see improvement in the load balancing across links as the width of ECMP spines increase over time. Segment routing will be popular with the MetroEthernet and SP crowd who continue to hug their MPLS control planes ever tighter and remain hopeful that SR will stave off the wave of SDN that will soon replace it.

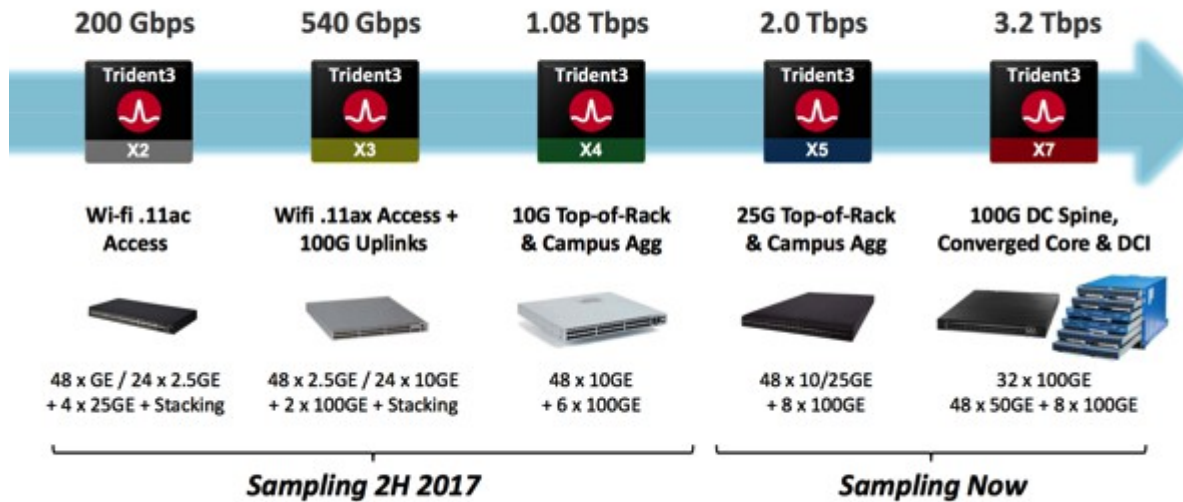
Finally the instrumentation is less obvious for users, but for network apps it's key to debugging & troubleshooting the operational systems. If vendors implement Broadcom's instrumentation in their code we could see an improvement in reliability and stability.

Whitebox In The Campus

Broadcom is producing a wider range of ASICs than previous generations. The first generation Trident was intended for the Top-of-Rack data center switch market where prices are highest and to test the market. The mega-scale cloud companies were quick to adopt the Trident 1&2 chipsets for their data center needs.

Trident 2 was suited to the Data Centre more generally including ECMP spines. The Trident 3 expands data centre scale & performance but also starts the obvious move into the Campus & Wireless networks. The Trident 3 X2 & X3 are suited to lower density/speed of 1/2.5/10G and stacking in the Wireless and Campus Edge.

Trident 3 2017 Series: Five Optimized Solutions from 200G to 3.2T Unified Programmable Architecture from Campus Access to DC Core



14 | Broadcom Proprietary and Confidential. © 2017 Broadcom. All Rights Reserved. The term "Broadcom" refers to Broadcom Limited and/or its subsidiaries.



For the Data Center, the X4 is suited for 10G TOR, X5 for 25/50G TOR and top-of-the-range X7 for the 100G ECMP spine.

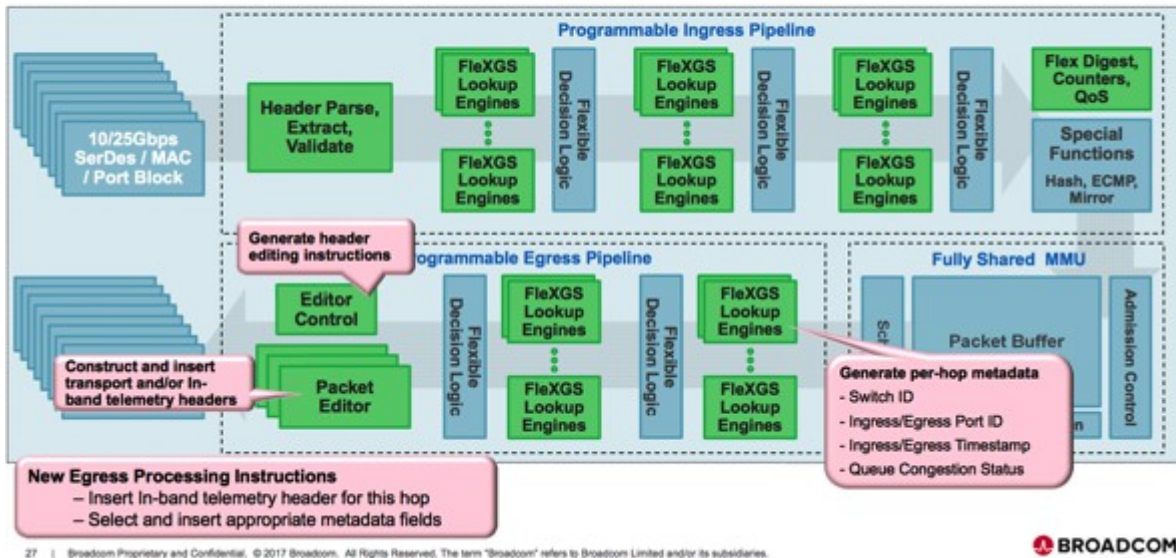
Programmable Pipelines

I have talked extensively about programmable pipelines for three or four years now since Cavium released its Xpliant switch silicon. Broadcom Trident 3 implement a five-stage pipeline, three before the switch fabric and two after.

For most enterprises this means support for new tagging and encapsulation protocols like Geneve and NVGRE through a software change instead of hardware. At the more remote extremes of possibility, vendors could interoperate with Cisco-driven formats like SGT or Network Service Headers to get customer attention.

It's also possible to create new tagging schemes to provide deeper insight into packet flows on the network. All of the silicon vendors are promoting the same story around telemetry through tagging and the ability of the switch hardware to timestamp.

Mapping In-Band Telemetry to Trident3 Programmable Pipe



Die Size


The fab process for the Trident 3 is 16nm which is near to the most modern die production. Smaller die sizes lower power consumption and improve performance. At the same time, increases in the number of transistors and clock speeds will consume extra power and needs extra cooling. In the briefing, MSDC customers have been highlighting power consumption as a key concern and this is feeding into the Broadcom design selection.

Pricing

You don't often see pricing details at product launch but this slide suggests that Broadcom knows exactly how much a 32 x 100GE switch should cost customers. **It's less than USD\$3000** .

Of course, you will still have to buy an operating system, applications and QSFP28 modules to complete the product, but the price point means that a data center class switch now costs roughly the same as an typical (non-vendor) x86 server. That is quite a transition in the price of network hardware.

Trident3: Unparalleled Switch Economics at 25/50/100GE



**VX1 32x100GE
Trident3 Based Switch**

**IX7 32x100GE
Trident3 Based Switch**

Volume List Price*

< US\$3000

* Bare Metal Configuration w/ ONIE, not including optics, for production volumes of 1Ku/year

Programmable Trident 3 ODM Systems Available Q3 2017 at Market-Leading Prices

36 | Broadcom Proprietary and Confidential. © 2017 Broadcom. All Rights Reserved. The term "Broadcom" refers to Broadcom Limited and/or its subsidiaries.

BROADCOM

Why This Matters

In the last year we have seen a rash of companies offering new silicon. Barefoot Networks and Innovium are the new entrants, with Cavium, Marvell, Centec Systems all producing silicon targeting different areas of the Ethernet switching market.

The Broadcom Trident 3 expansion into wireless and campus switches is a sound market expansion. In particular, Cisco's UADP silicon and the best-selling C3850 switches will have renewed competition as smaller vendors can focus on their applications using third party switching ASICs like the Trident 3.

The programmable pipeline is somewhat late to market. The Barefoot Networks Tofino chip has created a substantial amount of interest around its P4 programming language and the ease of use. Broadcom points to the wide usage of own StrataXGS SDK and backward compatibility will reduce the time to market for network software makers while also offering a range of new programmable features for packet handling.

Notes

Images are reproduced with permission of Broadcom.