

Cisco Cloud Scale ASIC switches – deliver a 2 year advantage over Merchant silicon switches



Dave Dhillon - July 5, 2016 - 3 Comments

Ever drive through a tract housing development where two to three models are repeated over and over again. All the houses have the same look, same layout, same square footage, etc. Over the last few years, it seems like a similar trend is happening in data center switching. Since most data center switches are based on the same Merchant silicon, there is very limited differentiation or innovation, much like the cookie-cutter housing tracts.

Sure, there's a difference in the operating system from vendor to vendor—but for the most part, switches based on Merchant silicon are limited by what the ASIC is capable of. They offer the same functionality, number of ports, speeds, power profile, and capabilities. It's like going to buy a car and only having a choice of colors. This approach is leading to stagnation and slow switching innovation.



With all the rapid shifts going on in the data center (cloud, higher density, faster speeds, containers, increased complexity, and more), Cisco recognized that current Merchant ASICs weren't going to deliver the capabilities needed for the next generation of data centers.

To break away from the cookie-cutter, me-too approach that barely addresses today's data center needs, Cisco developed the Cisco Cloud Scale ASIC to power our next-generation switches for the next-generation data center. Currently, the Cloud Scale ASICs are deployed across Cisco Nexus 9200 switches and the Nexus 9300-EX versions.

So, what's different and innovative about Cisco's Cloud Scale ASICs that puts it 2 years ahead of Merchant silicon ASICs? Let's count the ways!

It starts off with the latest in manufacturing technology. Merchant silicon ASICs are manufactured using 28nm technology. Cisco Cloud Scale ASICs use the latest 16nm technology. So, what? 16nm fabrication allows us to put more transistors in the same die size as Merchant silicon, resulting in cost optimization, higher density, higher bandwidth, larger buffers, larger routing tables, and more room for new hardware-based features.

Lower price point: With the 16nm fabrication, we're able to pack a lot more capacity and features into the same die size as Merchant ASICs and offer multi-speed 10/25G ports at the cost of 10G, and 40/100G ports at the price of 40G ports. More features, more flexibility, higher speeds, at the same price!

100% visibility: With Cloud Scale ASICs you can capture telemetry from every packet and every flow at line rate (yes, even 100G ports) with ZERO impact to the CPU (try that with NetFlow). Using the [Cisco Tetration Analytics™](#) platform, the extensive telemetry provides pervasive visibility into your applications and infrastructure and delivers operational capabilities that just can't be delivered with Merchant ASICs.



Greater bandwidth capacity: Cisco's Cloud Scale switches offer more bandwidth per rack unit and are more cost-effective than that provided by Merchant silicon-based switches. The 16nm fabrication technology has enabled Cisco to build a single switch-on-a-chip (SoC) ASIC that can support 3.6 terabits per second (Tbps) of line-rate routing capacity. It has also enabled Cisco to build the first native 48 x 10/25Gbps and 6 x 40/100Gbps top-of-rack (ToR) switch in the industry with full Virtual Extensible LAN (VXLAN) support.

Smart buffering: Cloud Scale ASICs deliver larger internal buffers (40 MB versus 16 MB) plus several enhanced queuing and traffic management features not found in Merchant silicon switches. There's a myth that the larger the buffer the better performance you get. That myth

has been debunked in a [Miercom test](#). Large buffers place all flows in a common buffer on a first-come, first-served basis, and then moved out in the same order. It does not distinguish between flow sizes; both large flows and small flows are treated the same. Therefore, the majority of the buffer gets consumed by large flows, starving the small flows. No matter how large the buffer is, the end result of deep buffers is added latency for both large and small flows, benefiting neither.

In contrast, the Cisco Cloud Scale smart buffer approach implements two key innovations to manage buffers and queue scheduling intelligently, by identifying and treating large and small flows differently.

Dynamic Packet Prioritization (DPP): DPP prioritizes small flows over large flows, helping to ensure that small flows are not affected by larger flows due to excessive queuing.

Approximate Fair Drop (AFD): AFD introduces flow-size awareness and fairness by an early-drop congestion-avoidance mechanism.

Both DPP and AFD ensure small flows are detected and prioritized and not dropped to avoid timeouts, while large flows are given early congestion notification through TCP to prevent overuse of buffer space. As a result, smart buffers allow large and small flows to share the switch buffers in a much more fair and efficient manner. This provides buffer space for small flows to burst, and large flows to fully utilize the link capacity with much lower latency times than a simple large buffer approach implemented in most Merchant silicon switches. For a more technical discussion on deep buffers versus smart buffers, check out [Tom Edsall's video](#).

Greater scalability: Beside higher bandwidth and larger buffers, Cisco Cloud Scale ASICs offer increased route and end-host scale. The Cisco Nexus 9000 switches (varies based on modular or fixed) can support up to 512,000 MAC address entries and up to 896,000 longest prefix match (LPM) entries: two to three times the number supported by popular Ethernet Merchant silicon-based switches. Cisco Nexus 9300 Cloud Scale switches are designed to support up to 750,000 IPv6 routes, compared with the 84,000 routes available with Merchant silicon.

So, no matter how you slice and dice the comparison between Cisco Cloud Scale ASICs and Merchant Silicon, you're getting a two year head-start with a lot more value and investment protection with Cisco's Cloud Scale based switches.