Innovium™

WHITEPAPER

The Most Scalable, Autonomous & Modern Data Center Networks

Fueled by Customer-Focused Innovation

Overview

Part 1: Data Centers are Transforming the Digital Economy

- Massive Secular Growth of Data Centers

Part 2: Customers Highlight Major Challenges with Network Infrastructure

- Explosive Network Traffic Growth
- Autonomous Networks: Telemetry, Analytics and Automation
- Programmability
- Other Key Requirements

Part 3: Innovium: Born in the Cloud Era and Focused on Data Centers

- World-Class Data Center Network Silicon and Software Team, With a Proven Track Record
- Ground-up, Customer Focused Innovation

Part 4: TERALYNX Delivers the Most Scalable and Autonomous Data Center Networks

- Target Customers and Use Cases
- The World's Most Scalable Switch with Unmatched Performance of Up to 12.8Tbps
- Comprehensive Telemetry & Analytics Enables Best-in-class Data Center Automation
- Line-rate Standards-based Programmability
- Modern, Open and Highly Resilient Software

Part 5: Summary

Data Centers are Transforming the Digital Economy

Massive Secular Growth of Data Centers

Economics, agility, and flexibility are driving customers to move IT to the public and private cloud. Enterprises are rapidly moving to cloud data centers in an unprecedented fashion. Cisco predicts that by 2020, 92 percent of workloads will be processed by cloud data centers.[1] In addition to growth driven by enterprises, consumers world-wide are driving growth as they continue to embrace mobile applications like search, maps, social networking and photo/video sharing, which are all served from large data centers.

Looking forward, machine learning and artificial intelligence are expected to be increasingly pervasive across all applications and device types. Enormous amounts of data are collected in everything from web browser clicks and cell phone photos to IOT sensors and autonomous cars. Businesses are investing heavily in mining and running data analytics on these large datasets to gain competitive insights, make intelligent decisions, and offer value-added services. These analyses are run on massive server farms inside large data centers using modern artificial intelligence techniques, imposing significant throughput and low-latency demands on the network.

As a result of these trends, Intel predicts that 70-80 percent of all systems will be deployed in "large scale data centers" by 2025.[2] Cisco expects that 485 hyperscale data centers will represent 47 percent of all installed data center servers by 2020.[3] This secular move to the cloud can be seen from the massive year over year growth of public cloud services such as AWS and Azure, and private cloud vendors like Nutanix.

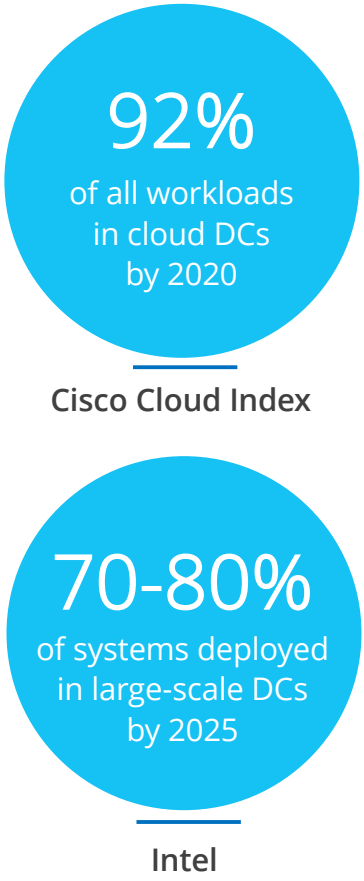

Figure 1. Secular Data Center Growth

Customers Highlight Major Challenges with Network Infrastructure

Explosive Network Traffic Growth

Within large data centers, network traffic is exploding. For example, network bandwidth demands inside Google's data centers are doubling every 12-15 months.[4] And AWS has highlighted that data center networking is a "red-alert situation" as cost escalates relative to other infrastructure.[5] Cisco predicts that traffic within large scale data centers will quintuple by 2020.[6] Industry players and analysts forecast that 100G port shipments will exceed 40G in 2017 and further 400G deployments will start in 2018 and become commonplace in data centers by the end of the decade. Accelerated adoption by hyperscalers drives much better unit economics early in the lifecycle.

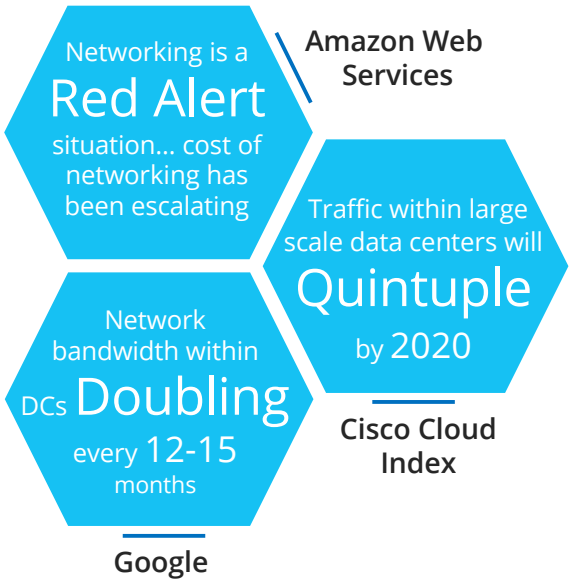

Figure 2. Huge Growth in DC Network Traffic

This remarkable traffic growth of server-server (or east-west) traffic as well as user-facing traffic is being driven by multiple trends:

East-West Traffic Drivers

User-Facing Traffic Drivers

Significant Drivers of Traffic Growth

- **Cloud Computing:** Enterprises are moving rapidly to the cloud, deploying SaaS, PaaS or IaaS models
- **Artificial Intelligence & Analytics:** AI & analytics operate on large datasets with distributed, offloaded compute requiring significantly higher network throughputs and lower latency
- **Microservice-based Applications:** New applications are adopting container-based architectures causing a dramatic increase in east-west traffic between containers

- **Distributed Storage:** Storage architectures are embracing cost effective, distributed, direct-attached models with a widespread adoption of flash and NVMe
- **Video:** Increasing growth of high-resolution HD and 4K video (e.g. YouTube, Netflix), as well as applications such as telepresence and continuous video surveillance
- **5G Networks:** Mobile operators will drive a mega-boost in traffic with new types of applications
- **Virtual & Augmented Reality:** VR & AR are expected to drive massive new high-volume platforms fueling additive kinds of high bandwidth traffic

As a result, Cisco forecasts that global data center IP traffic will grow to 15+ Zettabytes per year by 2020, representing a 27 percent CAGR from 2015 to 2020.[7]

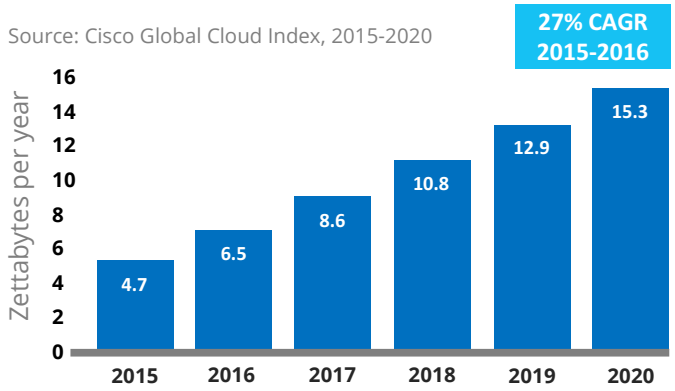

Figure 4. Global Data Center IP Traffic Growth

To address these stringent demands on the network, large data center customers need to deploy optimal network architectures in order to provide the best application performance. Larry Ellison, Oracle Founder and CTO, recently described the key network design principles for the Oracle Cloud as:

- Guaranteed bandwidth, with no over-subscription
- Very low latency from a flat, low-latency network, with a maximum of 2 hops between two compute nodes
- High bandwidth and fault-tolerant network

These design principles are not unique to Oracle. Rather, they are echoed industry-wide by other data center customers.

Autonomous Networks: Telemetry, Analytics and Automation

As the scale of data-centers increases, customers have a tremendous need for telemetry, analytics, and automation to trouble shoot problems, automate manual steps, and simplify operations. Cloud operators like Microsoft Azure, analysts like IDC, and cloud-centric network infrastructure providers like Arista Networks have all highlighted this critical need in public papers and presentations.

Figure 5. Need for Telemetry & Analytics
Source: Microsoft, IDC & Arista Networks

In addition, customers are beginning to use artificial intelligence and machine learning techniques on visibility and telemetry data from the network infrastructure to run analytics and automate the networks. The goal is to self-run, self-optimize and self-heal their networks. This is a foundational step towards autonomous data centers which will simplify network operations and lower operational costs. But this move requires network infrastructure to provide real-time, granular, accurate and comprehensive telemetry data.

Programmability

Data center customers are looking for ways to make the network infrastructure future proof and software-driven with programmability. For example, they expect the network infrastructure to be able to support new protocols (e.g. new tunneling types), connectivity rates (e.g. 200G & 400G) and flexibly allocate tables, without the need for ASIC spins. However these capabilities should not compromise the primary goal of the network which is to move data fast, reliably, efficiently and cost-effectively across the data-center.

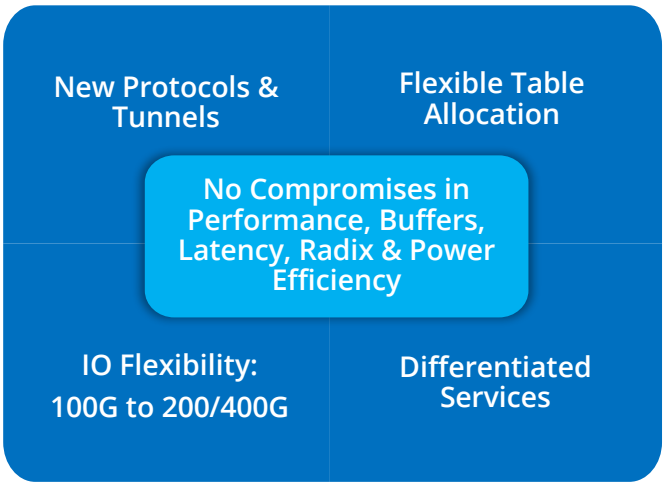

Figure 6. Programmability Requirements

Other Key Requirements

With the increasing size of data centers also comes rapid growth in the cost of powering and cooling them. These data centers are emitting CO2 at an alarming rate, which contributes to global warming. As a result, cloud data center customers are pursuing green initiatives to cut down on their respective carbon footprints. To achieve this goal, customers need ways to make their infrastructure as power efficient as possible.

Customers want the best user experience for applications running in the data centers. Hence, they want networks to have low latencies, along with high quality and resiliency. They don't want to experience any downtimes, especially as the top-of-rack switch is a single point of failure. Hence, network infrastructure needs to be highly robust and resilient, both from hardware and software perspective.

Finally, data center customers are demanding solutions to meet requirements discussed above within extremely tight constraints of Capex and Opex. As a result, solutions need to meaningfully drive down cost/bit while delivering breakthrough levels of performance and features.

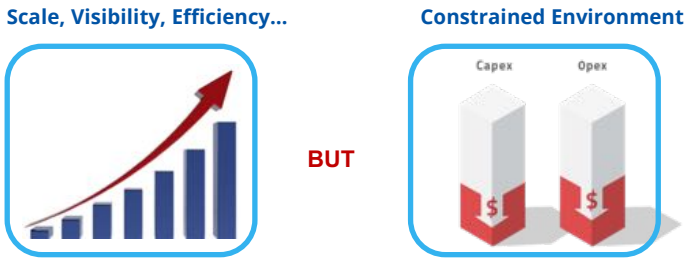

Figure 7. Operate with constrained Capex & Opex

Innovium: Born in the Cloud Era and Focused on Data Centers

World-Class Data Center Network Silicon and Software Team, With a Proven Track Record

The semiconductor industry is littered with failed organizations that started with great ideas and intentions but lacked the domain knowledge or track record to deliver timely, high quality products with essential capabilities. Further, many companies try to leverage architectures from prior projects or research unrelated to customer priorities. Large scale data centers began to emerge about 5-7 years ago, but even today a majority of networking products use architectures originally created for service provider or enterprise / campus networks. Some companies are developing products based on research they believe might be useful without any clear value proposition leading to highly sub-optimal trade-offs.

Innovium was born in the cloud era with an intense focus on the data center. It is not encumbered by legacy architectures. Innovium has also assembled a strong, world-class data center networking switch team, with many members that delivered multiple generations of network silicon products pervasively deployed in today's data centers. Innovium's team members have a track record of engineering execution with the ability to take silicon to production and ultimately deliver customer satisfaction. Data center customers recognize that this capability is an important predictor of compelling products and value along with dependable roadmaps.

Ground-Up, Customer-Focused Innovation

Innovium delivers the industry's first network switch architected from the ground-up to optimally address the needs of large scale data centers. Alternative products in the market have been designed for multiple disparate segments resulting in poor trade-offs for any single segment. From day one, the Innovium team focused its attention on data center customer pain points and trend lines and oriented its innovation around addressing those problems. Others have taken a more academic approach where innovations were born in the lab and are not optimally designed to address the real-world problems of large-scale data centers.

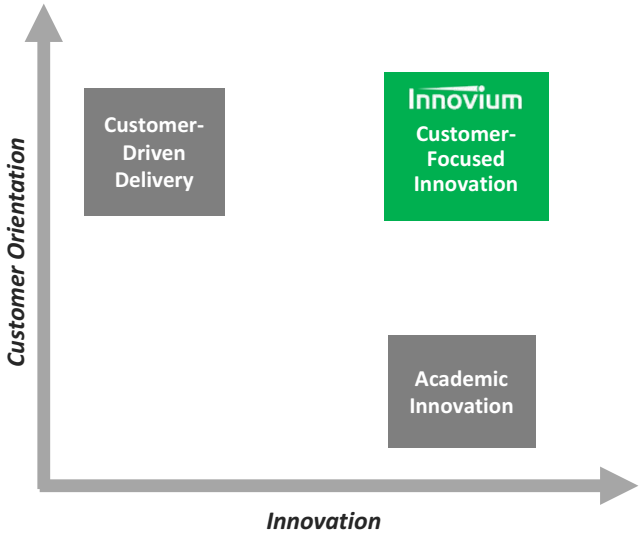

Figure 8. Innovium delivers customer-focused innovation

TERALYNX™ Delivers the Most Scalable & Autonomous Data Center Networks

Innovium's TERALYNX, fueled by customer-focused innovation, delivers the world's fastest and most scalable Ethernet switch silicon family with industry-leading analytics, programmability and power efficiency. The TERALYNX 7 family offers up to 12.8 Tbps single-chip performance; 2x performance per watt and 3x buffer size versus alternatives; breakthrough telemetry and analytics; and support for new 200G and 400G Ethernet standards.

The modern and scalable TERALYNX switch silicon architecture offers highly innovative data center focused capabilities in a modular and area-optimized fashion to scale from 2Tbps to 51.2Tbps and beyond to address the bandwidth needs over the next decade. There are three key areas of innovation:

- **INNOFLEX™:** Line-rate, programmable forwarding with best-in-class, predictable latency
- **TERASCALE™:** High performance, high capacity switch and buffer fabric
- **FLASHLIGHT™:** Comprehensive visibility and analytics engine to enable autonomous data centers

In addition, TERALYNX features industry-leading, proven SerDes that supports 10G and 25G NRZ as well as 50G PAM4 I/O standards. This provide customers a variety of connectivity choices ranging from the widely deployed 10/25/40/50/100G Ethernet to the upcoming 200/400G Ethernet.

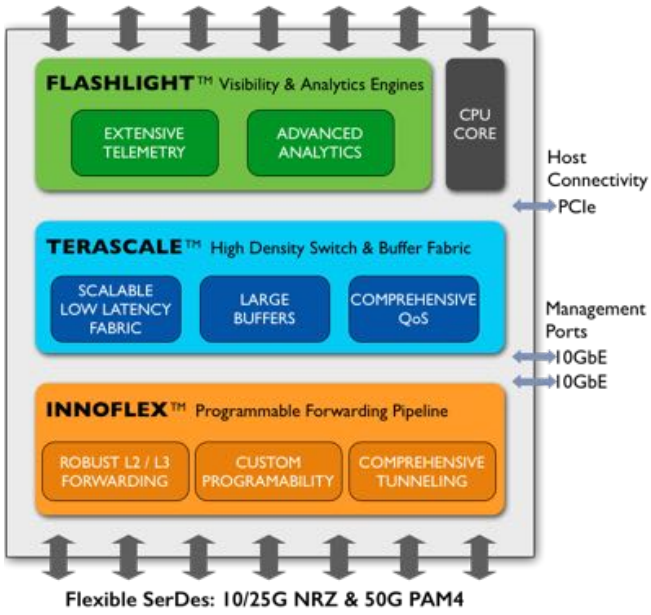

Figure 9. TERALYNX Switch Architecture

Target Customers and Use Cases

Innovium developed the TERALYNX Ethernet switch silicon family to support high performance, autonomous and programmable data center networking infrastructure for public and private clouds, HPC and service provider data centers.

It can be used to develop a range of fixed and modular (chassis) switches using single or multiple chips for Top-of-Rack (TOR), Leaf, and Spine switch system designs. Some desirable system design examples include:

Form Factor	Example Options	Key Benefits
1U: 32 x 400G QSFP-DD / OSFP	32 x 400G 64 x 200G 128 x 100G	▪ Single chip 12.8T with best performance per watt ▪ First 400G solution with 50G PAM4 SerDes I/O ▪ 128 x 100G available with breakout
2U: 64 x 100/200G QSFP	64 x 200G 32 x 400G 128 x 100G 64 x 100G	▪ Single chip 12.8T with best performance per watt ▪ Also delivers 128 x 100G in 2RU form factor ▪ Future proof for 200G with 50G PAM4 SerDes I/O
1U: 32 x 100/200G QSFP	32 x 200G 32 x 100G 48 x 25/50G + 8 x 100G	▪ Single chip 6.4T with best pps and perf/W metrics ▪ Also delivers 64 x 100G in 1RU ▪ Future proof for 200G with 50G PAM4 SerDes I/O
4U: 128 x 100G QSFP	128 x 100G	▪ Single chip 128 ports of 100G ▪ Up to 3x lower power, cost and latency

Figure 10. Single Chip Fixed Switch Design Examples

In addition, customers are beginning to use artificial intelligence and machine learning techniques on visibility and telemetry data from the network infrastructure to run analytics and automate the networks. The goal is to self-run, self-optimize and self-heal their networks. This is a foundational step towards autonomous data centers which will simplify network operations and lower operational costs. But this move requires network infrastructure to provide real-time, granular, accurate and comprehensive telemetry data.

The World's Most Scalable Switch with Unmatched Performance of Up to 12.8Tbps

The TERASCALE switch and buffer fabric in Innovium TERALYNX delivers the industry's highest performance, radix and buffers along with best-in-class low latency. Key capabilities include:

- 12.8Tbps, 9.6Tbps, 6.4Tbps and 3.2Tbps single-chip performance options, which deliver up to a 6x advantage over alternatives using widely adopted CLOS network architecture
- Single flow performance of up to 400Gbps at 64B packet size – a 4x advantage versus alternatives
- 70MB of on-chip buffer using a breakthrough InnoShare™ architecture for superior network quality, fewer packet drops, and substantially lower latency compared to alternate buffering options
- Up to 128 ports of 100GbE, 64 ports of 200GbE, or 32 ports of 400GbE which enables flatter networks for reduced CAPEX and fewer hops
- Support for multicast, cut-through, and store-and-forward with best-in-class cut-through latency of less than 350ns

Compared to alternatives currently deployed in the market (3.2 Tbps switches) and even other announced 6.4 Tbps products, the TERALYNX 12.8 Tbps switch offers a 6x - 12x advantage as illustrated in the diagram below. It would take twelve 3.2 Tbps switches or six 6.4 Tbps switches to build an equivalent CLOS network that matches a single TERALYNX 12.8 Tbps switch. This translates to huge cost savings for customers.

There are several additional key advantages:

- **Savings from optics:** Optics cost is a huge component of network infrastructure costs, that could run up to 3-5x of switch costs. Cost per bit for 200/400G optical transceiver solutions is expected to be significantly more compelling than 100G optical solutions. Hence, data center savings from optics are expected to be huge.

- **Saving from other system components:** Not only is there savings from switch silicon, there are significant savings in CPU, PCB, memory, power supply, fans etc.
- **Savings from power & cooling:** With reduction in number of switches and associated system components, there are significant savings from in on-going power and cooling costs.
- **Software & Management Simplification:** Reducing the number of networking elements reduces route convergence times significantly. Also, reduction in number of networking elements reduces the complexity of managing switches.

The breakthrough performance and scalability of TERALYNX, with best-in-class latency enables customers to design the most optimal network architectures that are not possible with other solutions today. These architectures provide guaranteed bandwidth, no over-subscription, and a flat, low-latency network with a minimum number of hops between two compute nodes. TERALYNX advantages give customers the freedom to place and dynamically shift workloads anywhere in the data center without network constraints.

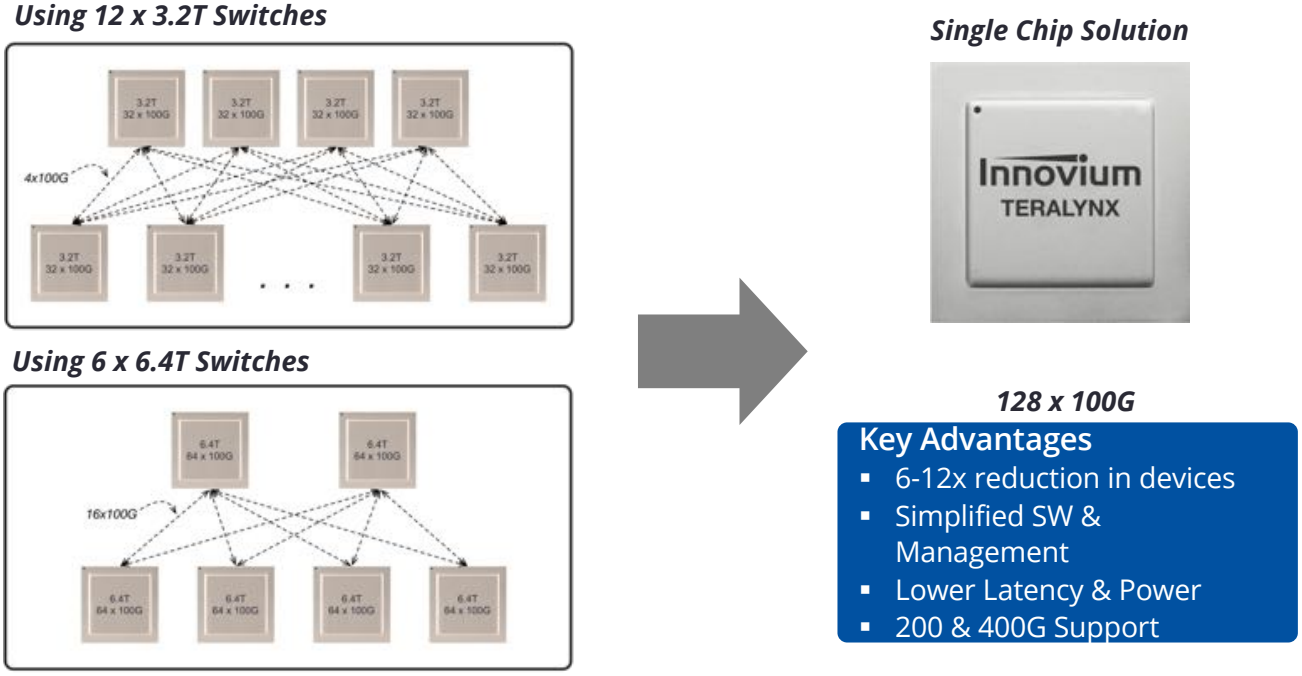

Figure 11. Key Advantages of TERALYNX 12.8Tbps Performance

Comprehensive Telemetry & Analytics Enables Superior Data Center Automation

The Innovium FLASHLIGHT™ visibility and analytics engines offers innovative and actionable telemetry and analytics to simplify network operations and facilitate autonomous data center networks. They offer the following key capabilities:

- Host CPU-based automation using high bandwidth DMA with minimal CPU overhead with 20+ Gbps Gen3 PCIe host and 100+Gbps Ethernet connectivity
- Extensive visibility and telemetry capabilities like sflow, FlexMirroring along with accurate and highly customizable 64-bit extra-wide counters
- Runtime modifiable watermarks, thresholds, and configurations to enable actions in a predictive manner
- Real-time automation enabled by an on-chip GHz ARM CPU
- Network-wide automation with in-band telemetry using P4-INT with critical extensions that can aggregate data across multiple switches and react to macro network needs

Unique, advanced analytics that are designed from the ground-up understanding customer pain-points enable optimal resource monitoring, utilization, and congestion control which allows predictive capabilities and network automation.

Artificial intelligence and machine learning techniques are being aggressively adopted by customers. These techniques require the kind of granular, real-time data that TERALYNX's extensive visibility and telemetry capabilities provide. FLASHLIGHT's capabilities allow customers to self-run, self-optimize, and self-heal their networks and build truly autonomous data centers.

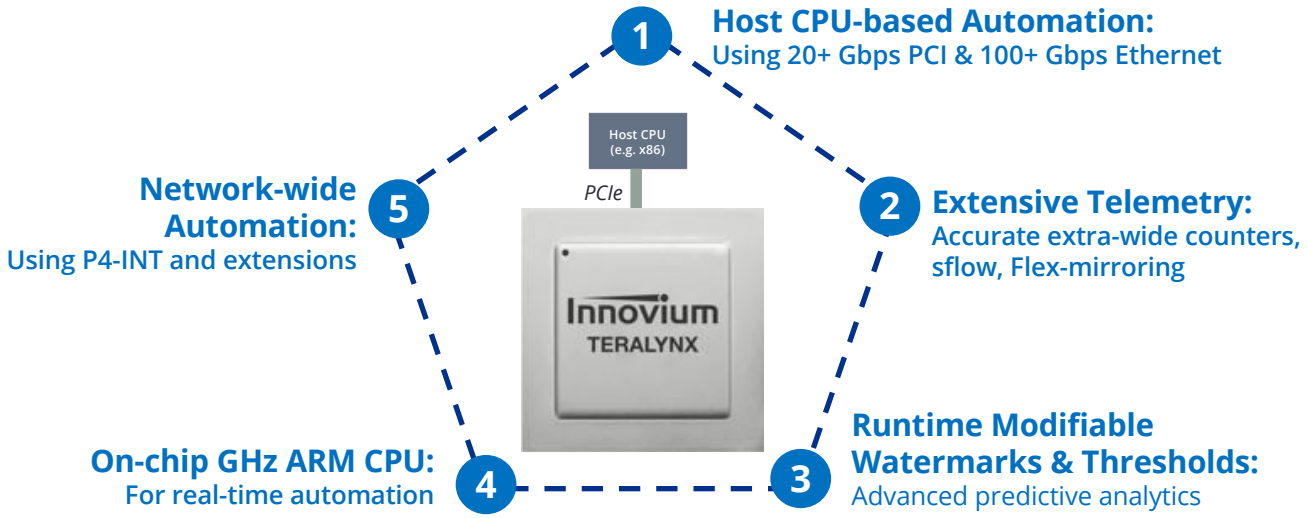

Figure 12. Comprehensive Telemetry & Analytics for Autonomous Networks

Line-rate, Standards-based Programmability

The INNOFLEX™ programmable forwarding pipeline provides line-rate, standards-based, and feature-rich flexibility. It offers the following key capabilities:

- Programming comprehensive layer 2/3 forwarding and flexible tunneling including MPLS, VXLAN, Geneve etc. from software running on the host CPU
- Large table resources with flexible allocation across L2, IPv4 and IPv6 to meet customer specific deployment scenarios
- On-chip GHz ARM CPU for real-time programming
- Rich, standards-based line-rate programmability to add new/custom protocols and features
- Programmable I/O connectivity for 10/25/50G SerDes per lane to support 10/25/50/100/200/400G per port

TERALYNX enables customers to deploy switches easily with today's 10G or 25G NRZ infrastructure for 10/25/50/100G Ethernet connectivity with up to 128 ports of 100GbE Ethernet using a single chip. This leverages the widely available infrastructure of direct attach copper (DAC) and optical cables and SFP/QSFP-based optical modules. Additionally, 50G PAM4 I/O is being urgently demanded by major data center operators to scale performance and meaningfully reduce cost per bit. TERALYNX is the industry's first switch to support PAM4 I/O and enable a seamless transition without requiring any forklift upgrades.

TERALYNX enables customers to program new protocols (e.g. new tunneling types), flexibly allocate tables, and support newer and higher connectivity rates (e.g. 200G & 400G). Unlike with other solutions, customers do not need to sacrifice any of the essential capabilities like performance, buffers, latency, radix and power efficiency.

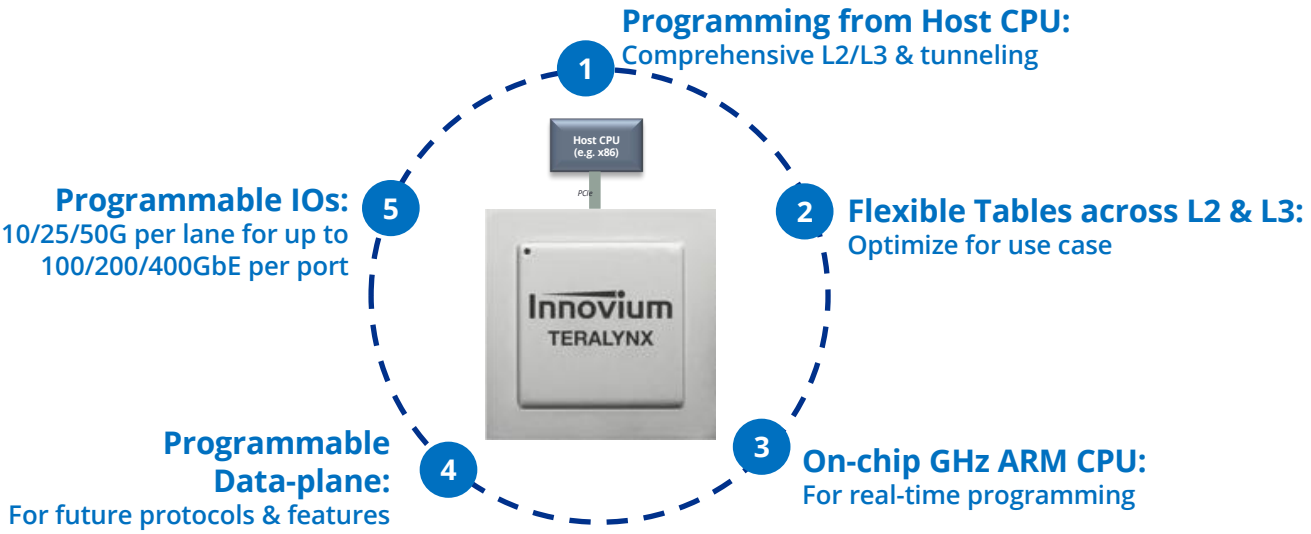

Figure 13. Compelling Programmability Options

Modern, Open and Highly Resilient Software

Innovium offers a comprehensive software solution which includes the TERALYNX Software Development Kit (SDK), example application code, standard APIs, open-source network operating system (NOS) software, software programmability and services for seamless and accelerated deployment. The SDK, developed with a clean-sheet modern design, delivers rich functionality along with high performance and high availability capabilities critical for data centers. It supports open APIs including OCP SAI (Switch Abstraction Interface) for easy integration with NOS software, as well as innovative APIs to unlock TERALYNX's FLASHLIGHT analytics and INNOFLEX programmability capabilities.

Summary

TERALYNX is the world's highest performance and most scalable Ethernet switch silicon family that leapfrogs the industry with an innovative data center optimized architecture. It is the industry's first switch to shatter the 10 terabits per second (Tbps) performance barrier in a single-chip while delivering breakthrough telemetry and automation, line-rate, standards-based programmability, and best-in-class low-latency. TERALYNX, developed by a highly experienced team that has developed multiple generations of successful switch products in deployment today, provides a 6x - 12x advantage over alternatives in the market today. For more information, please visit: http://www.innovium.com

References

1. Cisco Global Cloud Index: Forecast and Methodology, 2015–2020 [White paper]. Retrieved February 23, 2017, from Cisco Systems, Inc.: http://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf
2. Sverdlik, Y. (2016, April 22). Intel: World Will Switch to Scale Data Centers by 2025. *Data Center Knowledge.* Retrieved from: http://www.datacenterknowledge.com/archives/2016/04/22/intel-world-will-switch-to-scale-data-centers-by-2025/
3. Cisco Global Cloud Index
4. Vahdat, A. (2016, May 16). Interop 2016 Keynote, Google [Video file]. Video posted to: https://www.youtube.com/watch?v=vMgZ_BdipYw&t=691s and
   Vahdat, A. et al (2015). Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network (Presented at ACM SIGCOMM 2015). Retrieved February 26, 2017, from: http://conferences.sigcomm.org/sigcomm/2015/pdf/papers/p183.pdf
5. Hamilton, J. (2014, November 17). AWS re:Invent 2014: AWS Innovation at Scale (SPOT301) [Video file]. Video posted to: https://www.youtube.com/watch?v=JIQETrFC_SQ
6. Cisco Global Cloud Index
7. Cisco Global Cloud Index

Innovium, Inc.
6001 America Center Drive
San Jose, CA 95002
info@innovium.com

Innovium™