

Switching Architectures for Cloud Network Designs

Networks today require predictable performance and are much more aware of application flows than traditional networks with static addressing of devices. Enterprise networks in the past were designed for specific applications while new cloud designs in the data center can address a multitude of applications. This is clearly a radical departure from today's oversubscribed networks in which delays and high transit latency are inherent.

Predictable Network Performance Based on Applications

Unlike past client-server designs based on classical web (256KB transfers), e-mail (1 MB) or file transfers (10Meg), new cloud networks in the data center are required to offer deterministic performance metrics. Modern applications can be specific and well defined, such as high frequency algorithmic trading or seismic exploration analysis that requires ultra low latency. Other examples include the movement of large volumes of storage or virtual machine images, or large-scale data analytics for web 2.0 applications. These datacenters demand non-blocking and predictable performance. A key aspect of switching architectures is uniformity of performance for application scale- out across physical and virtual machines. There must be equal amounts of non-blocking bandwidth and predictable latency across all nodes. Newer multi-core processors are also stressing network bandwidth. Therefore, consistent performance with a balance of terabit scalability, predictable low latency, non-blocking throughput, and high speed interconnects driving multiple 1/10GbE and future 40/100GbE, are all essential characteristics of cloud networking architectures.

Foundations of data center switching architectures

Two switching architectures are emerging in the cloud switching data center. Cut through switching offers ultra low latency in high performance compute cluster (HPC) applications, and store and forward switching with deep memory fabrics and Virtual Output Queuing (VOQ) mechanisms provide the necessary buffering for web based data center applications (Figure 1). The Arista 7000 Family of switches is ideally suited for low latency, two-tier, leaf and spine HPC network designs. The Arista 7048 is optimal for heavily loaded next-gen data centers using asymmetric 1 & 10GbE connections to support storage and web based applications.

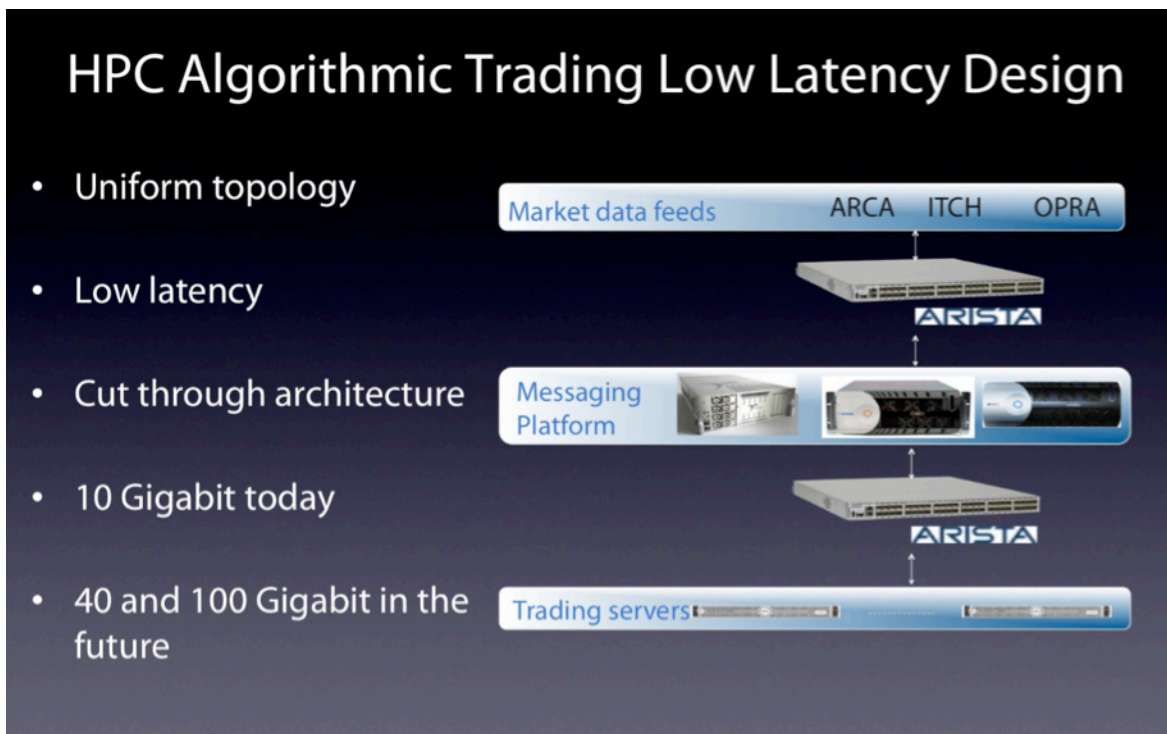


Figure 1: Left to Right: 7050SX, 7050QX, 7050TX

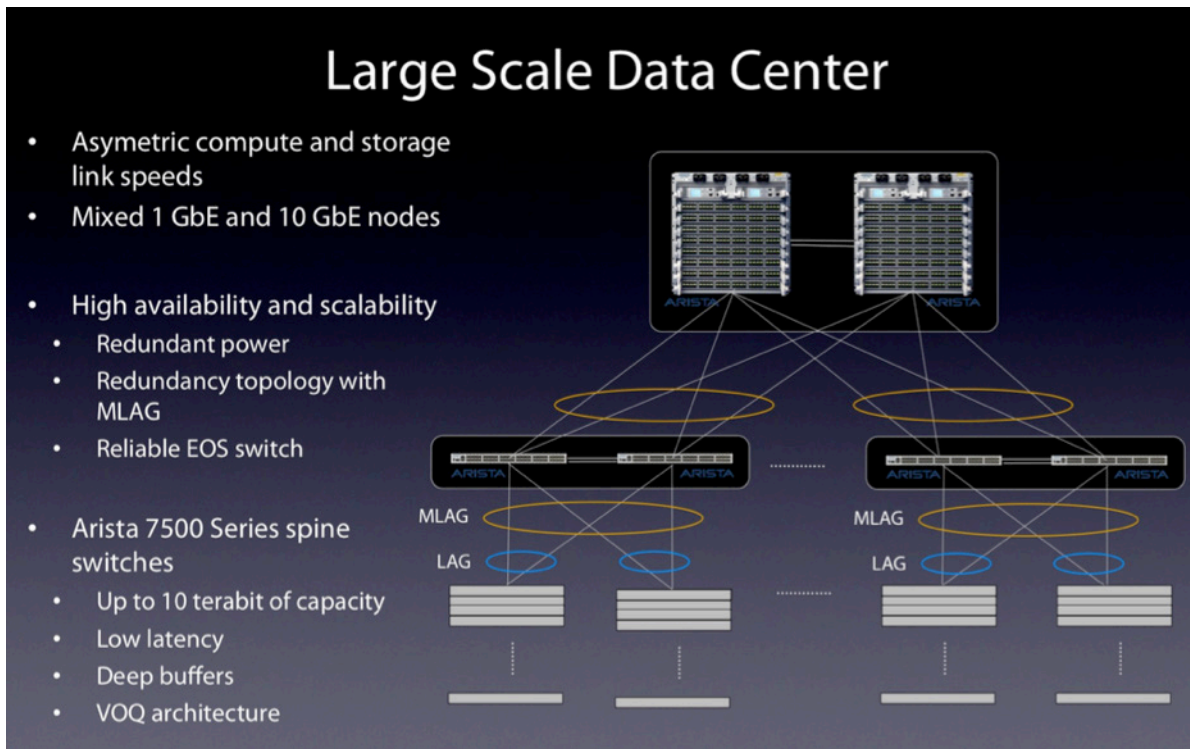


Figure 1b: Large scale, Asymmetric data center design

Low Latency, High Performance Compute (HPC) Clusters:

Modern applications, using real time market data feeds for high frequency trading in financial services, demand cut through and shared memory switching technologies to deliver ultra low latencies that are measured in microseconds and sometimes even hundreds of nanoseconds. The advantage of this 10GbE architecture is best latency with minimal buffering at a port level. This guarantees near instantaneous information traversal across the network to data feed handlers, clients and algorithmic trading applications. Cut through switching is an ideal architecture for leaf servers when traffic patterns are “well-behaved” and symmetric: such as HPC, seismic analysis and high frequency trading applications. It assumes the network is less than 50% loaded and therefore not congested, and that low latency is critical. Cut through switching can shave off several microseconds, especially with large and jumbo frame packets. This scale of low latency can save millions of dollars in a “time is money” environment. The Arista 7100 Series is ideally suited for ultra low latency; because packets are forwarded as they are being received instead of being buffered in memory. It also enables rapid multicasting while minimizing queuing and serialization delays (Figure 2).

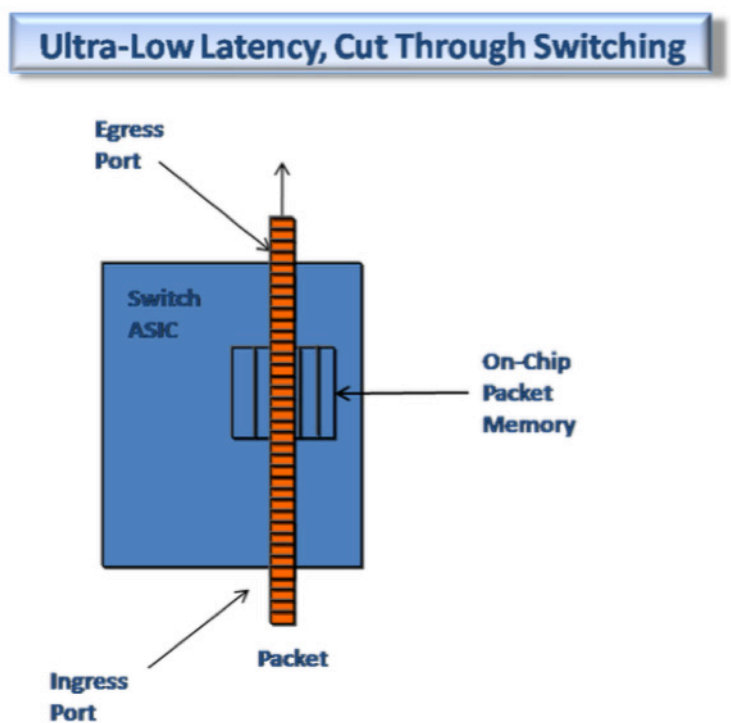


Figure 2: Low Latency Cut-through Switching for HPC Clusters

Deterministic Performance for Next Generation Datacenters:

For heavily loaded networks, such as “spine” applications, seamless “leaf” access of storage, compute and application resources, and datacenter backbones, predictable, uniform performance in a switched network scale-out is a key requirement. Applications demanding large blocks of data, such as map-reduce clusters, distributed search, and database query systems are typical examples. Performance uniformity is a mandatory requirement. Slightly higher latencies of 3-6 microseconds are acceptable since legacy switches deliver orders of magnitude poorer performance with 20-100 microsecond latencies. A switching architecture providing increased buffers, on the order of many megabytes per port, in a well-designed store-and-forward system is optimal for these applications. Modern store and forward switching architectures utilize Virtual Output Queuing (VOQ) to better coordinate “any to any” traffic. VOQ avoids switch fabric congestion and the “head of line blocking” problems that often plague legacy switches. Combining VOQ techniques with expanded buffering brings additional flexibility to application and overall network behavior. Large buffers mitigate congestion when traffic is bursty or highly loaded by devices simultaneously converging on common servers. An example of the latter occurs when an application server receives data from a striped bank of storage servers, and all the responses happen simultaneously. In this case, the switch must have adequate buffering to consistently hold the storage data without data loss. Deep buffering is also important for asymmetric transfers between 10G to 1G networks to accommodate link speed mismatch.

Cloud Networking Cases:

Cloud networking designs can be constructed from two tiers of leaf and spine switches from Arista (Figure 4). Take an important and familiar social networking application such as Facebook. Sources show they have constructed a cloud network of 30000 servers with 800 servers per “memcache” cluster, generating 50 to 100 million requests while accessing 28 terabytes of memory! Instead of using traditional database retrieval schemes that would take five milliseconds access time, Facebook utilizes a memcache architecture that reduces access time to half a millisecond! Reduced retransmit delays and increased persistent connections also improve performance. In this environment, large buffers with guaranteed access should be a key consideration. With their large buffers, advanced congestion control mechanisms, and VOQ architecture, the Arista 7048 Switch and 7508 Series of switches are a natural fit in applications with high volumes of storage, search, database queries and web traffic.

Consider this proven case for low-latency, high frequency trading (HFT) applications. It uses programs that automatically execute financial trades based on real time criteria such as timing, price or order quantity. These applications are widely used by hedge funds, pension funds, mutual funds and other institutional traders. As the application runs, it reacts to any input or piece of information and processes a trade in a fraction of a microsecond: literally in the blink of an eye. Financial protocols are widely used for real-time international information exchange of related securities and market transactions. It is expected that more applications will become multi-threaded in the future, making low latency interconnect across compute cluster nodes a future requirement for cloud and switching architectures. The Arista 7000 Family of switches is a natural fit in these ultra low latency HPC designs.

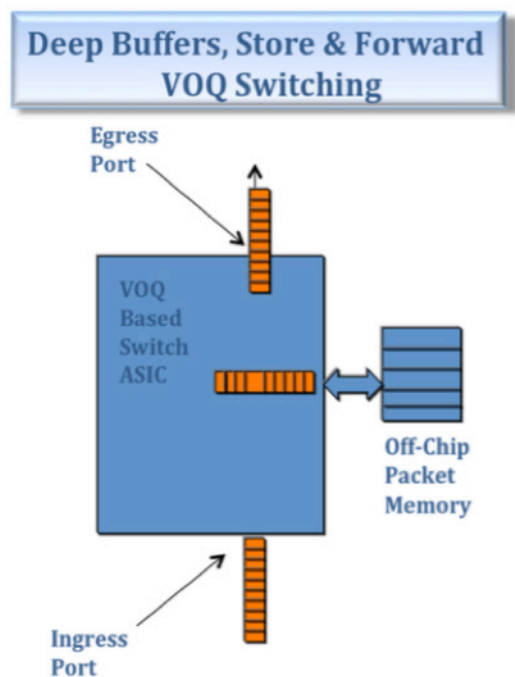


Figure 3: Large Scale Data center demand Large Buffers and VOQ/Store & Forward Switching for asymmetric traffic

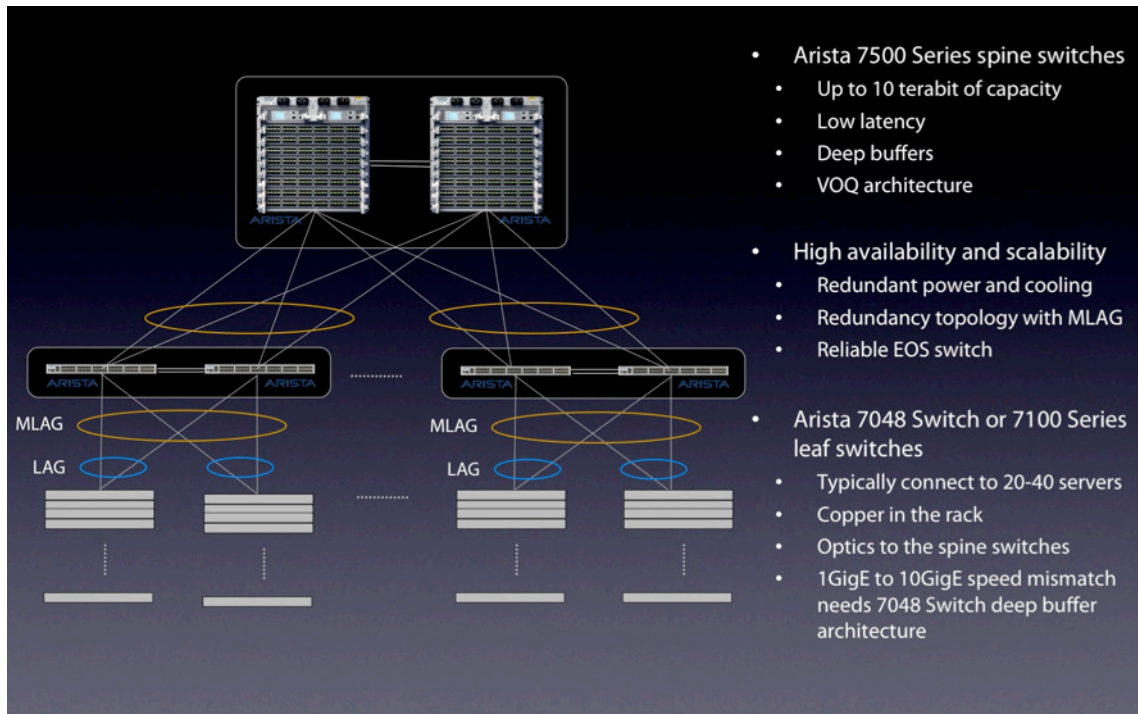


Figure 4: Cloud Networking Designs based on Arista 7000 Family

Summary

A growing number of “killer” cloud applications can take advantage of Arista’s new switching architectures for the datacenter. These include:

- High frequency financial trading applications;
- High performance computing (HPC)
- Clustered compute applications
- Video on Demand
- Network storage access
- Web analytics, Map-Reduce, Database, or Search Queries
- Virtualization

Networks designed in the late 90s primarily addressed static applications and email. Today’s new applications and traffic patterns are dynamic and demand new switching approaches for real-time application access. The future of cloud networking optimizes the following variables: guaranteed performance, low latency and any-to-any communication patterns. Arista’s switching architectures and expanded 7000 Family are designed to deliver the optimized cloud networking solution.

Santa Clara—Corporate Headquarters

5453 Great America Parkway,
Santa Clara, CA 95054

Phone: +1-408-547-5500

Fax: +1-408-538-8920

Email: info@arista.com

Ireland—International Headquarters
3130 Atlantic Avenue
Westpark Business Campus
Shannon, Co. Clare
Ireland

Vancouver—R&D Office
9200 Glenlyon Pkwy, Unit 300
Burnaby, British Columbia
Canada V5J 5J8

San Francisco—R&D and Sales Office
1390 Market Street, Suite 800
San Francisco, CA 94102

India—R&D Office
Global Tech Park, Tower A & B, 11th Floor
Marathahalli Outer Ring Road
Devarabeesanahalli Village, Varthur Hobli
Bangalore, India 560103

Singapore—APAC Administrative Office
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

Nashua—R&D Office
10 Tara Boulevard
Nashua, NH 03062



Copyright © 2016 Arista Networks, Inc. All rights reserved. CloudVision, and EOS are registered trademarks and Arista Networks is a trademark of Arista Networks, Inc. All other company names are trademarks of their respective holders. Information in this document is subject to change without notice. Certain features may not yet be available. Arista Networks, Inc. assumes no responsibility for any errors that may appear in this document. 09/15