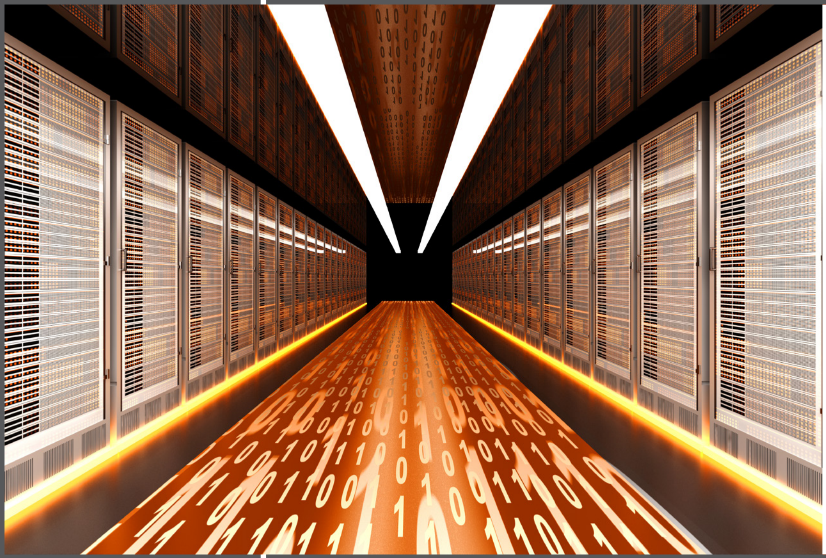
Simply a Better Network.™

Optimizing Hyperconverged Infrastructure with Plexxi Hyperconverged Network



Issue 1

Gartner.

Optimizing Hyperconverged Infrastructure with Plexxi Hyperconverged Network

Introduction

Enterprise IT is undergoing massive transformations due to the impact of on-demand rental infrastructure ("cloud") and increasingly more scalable and powerful software architectures and development/deployment tools. More than $1 trillion in IT spending will be directly or indirectly affected by the shift to cloud during the next five years, said Gartner, Inc. This will make cloud computing one of the most disruptive forces of IT spending since the early days of the digital age.

Yet some Enterprise organizations still have powerful and complex applications that cannot easily be re-built as "cloud-native" applications to fully take advantage of rental resources, or have specific custom functionality that is not available from SaaS offerings. Other Enterprises have rebuilt their applications to be cloud native, but prefer the control and security advantages they have with their own internal infrastructure, or have the skills and usage patterns to fully warrant an ownership versus rental model.

To help drive similar levels of cost, scale, automation, and utilization that cloud providers enjoy, a new breed of Enterprise IT infrastructure and software tools have emerged to help support on-premise cloud-like capabilities. Increasingly, these new solutions offer much higher levels of integration of key IT components to help these companies avoid having to do the hard and error prone job of putting all the disparate pieces together. After all, the new job of Enterprise IT professions is to help their organization deliver application experiences to their constituents, whether internal or external. This mission requires a drastic move away from the old "silo'd" approach to IT infrastructure, and a whole-hearted embrace of convergence, simplicity, and automation.

Hyperconverged Integrated Systems Reduce IT Cost and Complexity

In this vein, many IT organizations are looking to hyperconverged integrated systems[1] (HCIS) to reduce infrastructure cost and complexity so they can get to the business of delivering applications. Gartner projects that hyper-convergence revenue will grow at a compounded annual rate of more than 60% in the next few years.

In a Gartner survey, 74% of respondents indicated they are using, piloting or actively researching HCIS solutions. Next-generation HCIS solutions from vendors like Nutanix and SimpliVity pack compute and storage resources into compact x86 building blocks that are fully virtualized and uniformly administered. Hyperconverged systems reduce CAPEX by collapsing compute and storage silos and eliminating SANs. And they contain OPEX by reducing IT sprawl and lowering recurring power, cooling, rack space and administrative expenses.

By consolidating and unifying the compute and storage domains HCIS solutions can reduce TCO, accelerate time-to-value and simplify operations. But many enterprises fail to consider the networking implications of HCIS. Current data center networking constraints can hinder IT service agility, impair the performance of contemporary applications and hamper HCIS initiatives.

Re-architecting Data Center Networks for HCIS Implementations

Hyperconverged systems consolidate diverse applications, workloads and traffic types onto common infrastructure, introducing network engineering challenges for IT planners. First, while new virtual machines (VMs) and storage services can be easily provisioned from the HCIS

[1] Often referred to as hyperconverged infrastructure solutions

management system, the provisioning of the network services (e.g. the Virtual LAN or other port configurations on the network switches) still need to be manually coordinated with a separate networking team. Also, conventional data center networks are not always designed to handle IP-based storage traffic on the same common network, and certainly the constraints around a distribute file or block-based storage system such as those employed in HCIS's need to be engineered by the network team to ensure the proper SLAs and quality-of-service attributes are designed. In many traditional data centers, storage traffic is kept on a separate network (such as a Fibre Channel SAN, or even a separate IP network for IP-based storage traffic). In HCIS deployments, the inter-node storage traffic is co-mingled with the compute traffic and unless the network team is made aware, will be treated the same which could lead to storage performance or cluster coherence issues. And because all workloads are treated equally, a data-intensive or bursty application can monopolize network capacity, impairing the performance of other applications (the so-called noisy neighbor problem). As Gartner points out "Mixing user access, node-to-node application traffic, VM mobility, storage access and back-end storage traffic on a single network can lead to unpredictable performance, availability and data integrity."[2]

When implementing HCIS solutions, IT planners must re-architect data center networks to ensure adequate performance and service quality for all applications. Gartner's "Leverage Networking to Ensure Your Hyperconverged Integrated Systems Can Support Demanding Workloads" report examines key HCIS networking challenges and provides recommendations for ensuring adequate performance and reliability for diverse workloads.

The research firm identifies four distinct classes of enterprise workloads based on their resource needs:

- CPU-dependent workload
- Memory-dependent workload
- I/O-dependent workload
- Storage-dependent workload

Gartner concludes "the first two, typically, do not require high-performing networking services; however, the latter two do, especially for highly variable or rapidly growing workloads."

An easy way to accomplish this without a full re-architecture of the network is to deploy HCIS with an integrated hyperconverged network (HCN) such as Plexxi HCN.

Source: Plexxi

[2] Gartner Report G00312901 "Leverage Networking to Ensure Your Hyperconverged Integrated Systems Can Support Demanding Workloads," August 19, 2016.

Research from Gartner:

Leverage Networking to Ensure Your Hyperconverged Integrated Systems Can Support Demanding Workloads

As adoption of hyperconverged integrated systems for mainstream workloads grows, network architects must ensure networking is more than a set of fat, dumb pipes. Demanding workloads such as enterprise-class databases and event/stream processing will drive HCIS networking into uncharted territory.

Key Challenges

- Most HCIS deployments today have largely ignored the network. Adoption of HCIS has built a new silo, resulting in new coordination challenges for the network team.
- Mixing user access, internode application traffic, internode storage traffic, VM mobility and storage access traffic on a single network switch without bandwidth control can lead to unpredictable performance and system stability, and compromised data integrity.
- Treating the network as fat, dumb pipes makes it harder to troubleshoot application performance problems from an end-to-end perspective.
- Lack of integration into the network topology and network management of data center leaf/spine networks increase cost and complexity, while reducing performance and availability.

Recommendations

Network architects should:

- Make network design an integrated part of your HCIS cluster design and implementation to ensure performance and reliability requirements are met.
- Reduce the number of required switch ports and cables by consolidating traffic on higher-speed (10/25 Gbps) ports.
- Require that your HCIS vendor's system management integrates with your switch's network operating system or network fabric controller to provide an application/workload-specific view of system performance.
- Connect your HCIS nodes directly to your data center leaf switches.
- Map network design to the unique internal characteristics of your chosen HCIS and supported applications.

Strategic Planning Assumptions

By 2018, more than 10% of hyperconverged integrated system (HCIS) deployments will suffer from avoidable network-induced performance problems, up from less than 1% today.

Through 2018, 25% of HCIS deployments will be plugged into the wrong tier of the data center network.

By 2018, 60% of providers will offer integration of networking services, together with compute and storage services, inside their HCIS products, up from less than 10% today.

Introduction

This research targets network architects (and their colleague HCIS platform architects) who are responsible for designing and deploying HCIS that will support very demanding workloads.

HCISs tightly couple compute, network and storage hardware that dispense with the need for a regular storage area network. Storage management functions, plus optional capabilities such as backup, recovery, replication, deduplication and compression, are delivered via the management software layer and/or hardware, together with compute provisioning. Examples include Nutanix, Scale Computing, Cisco HyperFlex and SimpliVity.

HCIS has treated networking as a fast, dumb cluster interconnect, while networks have treated HCIS clusters like big monolithic servers. This has been acceptable since HCISs were introduced as application-specific environments, hosting less demanding workloads such as virtual desktop environments, application development and remote office/branch office infrastructure. However, that is changing.

As HCIS are used to support more demanding applications from a performance (compute, storage and input/output [I/O]) and availability perspective, and as those clusters grow beyond a handful of nodes, networking must be an integral part of the cluster design.

Buyers are expanding their use cases to cover more demanding enterprise workloads. We characterize enterprise workloads into four types based on their resource needs:

- CPU-dependent workload
- Memory-dependent workload
- I/O-dependent workload
- Storage-dependent workload

The first two, typically, do not require high-performing networking services; however, the latter two do, especially, for highly variable or rapidly growing workloads. For example, network function virtualization applications or a typical MapReduce job can be considered I/O-dependent workloads, and database workloads and stream/event processing can be considered storage and I/O-dependent workloads. Networking performance in these categories will directly impact the application performance; thus, a compelling integration with networking services (such as reducing the I/O path) inside HCIS is important.

Scaling HCIS clusters with I/O-intensive workloads using an end-to-end network design results in a more robust, cost-effective and reliable HCIS deployment.

Analysis

Reduce the Number of Required Switch Ports and Cables by Consolidating Traffic on Higher-Speed (10/25 Gbps) Ports

HCIS clusters require large numbers of network ports (see Table 1). Many HCIS vendors use small form factor servers with four nodes in each 2 rack units (U). By default, each node uses at least two 1 Gbps ports, which is eight ports/2U. With 16 four-node clusters per rack, 128 ports per rack are required. If the systems separate networking for storage and app traffic, the number jumps to least 256 ports per rack. Out-of-band management raises the number to 320 ports per rack. Networking costs and cabling complexity quickly become challenges for customers. In many cases, systems are deployed with at least two 10 Gbps ports, in addition to the one port for management. For demanding workloads, additional ports can be required.

While these prices may not seem excessive considering the price of the fully configured HCIS, many clients comment to Gartner that they hadn't anticipated the level of investment required to support networking.

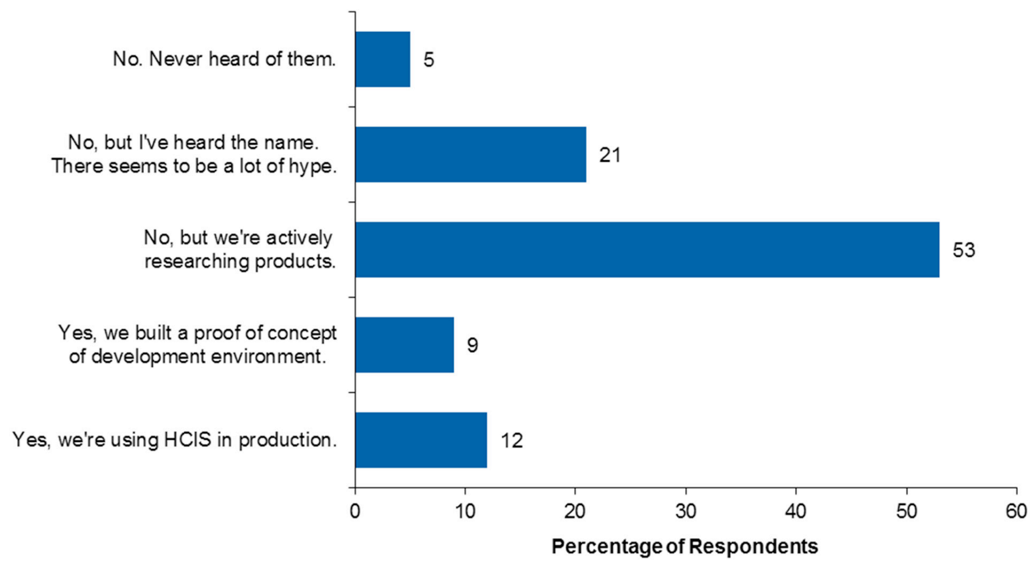
Total port count can be reduced by using higher-speed network interface cards (NICs; 10 Gbps or 25 Gbps), with virtual interface technologies (for example, Cisco UCS Virtual Interface Card 1240 and QLogic FastLinQ 45000) to support multiple virtual network adapters/NIC with quality of service (QoS)/flow. When combined with the appropriate bandwidth/QoS configuration on switches, it is possible to reduce the number of cables per node.

While 25 Gbps switches and NICs that are priced at 10 Gbps levels are available, switch vendors often lead with 10 Gbps switches (see Note 1). HCIS vendor support is highly varied, with some such as VMware offering support today, and many others delaying until as late as 2H17. Because of low volumes and cabling complexity, 40 Gbps will remain an expensive option with diminishing support. When deploying new systems, consolidate functions on 10 Gbps links only when doing so won't compromise performance or system availability and integrity.

Select nonblocking switches that offer 10/25 Gbps to minimize congestion-based performance problems and to provide investment protection.

HCIS can be implemented on blade systems with integrated switches to minimize cabling, although I/O slots will be consumed by switch modules, and onboard storage will be limited. Leveraging low cost 25 Gbps networking is valuable for blade systems as well.

FIGURE 1 HCIS Adoption Trends



Survey question: Have you ever deployed hyperconverged integrated systems?
n = 58
Source: Gartner (August 2016)

Table 1. Typical HCIS Network Requirements for 10 and 1 Gbps Switch Ports and the Associated Costs at $400/10 Gbps NIC, and Switch Port and Cable

Four-Node Clusters per Rack	Two 10 Gbps Ports/Nodes	Add Two 10 Gbps Ports/ Nodes for Storage Net	Add One 1 Gbps Port for OOB Management ($100)
4	32 ($12,800)	64 ($25,600)	80 ($26,000)
8	64 ($25,600)	128 ($51,200)	160 ($52,000)
16	128 ($51,200)	256 ($102,400)	320 ($104,000)

NIC = network interface card; OOB = out of band
Source: Gartner (August 2016)

Require That Your HCIS Vendor's System Management Integrates With Your Switch's Network Operating System or Network Fabric Controller to Provide an Application/Workload-Specific View of System Performance

Treating the network as fat, dumb pipes makes it harder to troubleshoot end-to-end application performance problems. Top-of-rack switches can provide very detailed information about network traffic, and emerging merchant and proprietary silicon enables you to gather data on, and act upon, virtually any byte pattern in an Ethernet packet. This will enable you to gather critical statistics, often on a per-VM/traffic-type basis that you can use to optimize and troubleshoot your clusters. This can reduce total cost of ownership because these systems require less time to manage/maintain, and it allows them to be options for workloads that require more stringent SLAs.

Optimizations can include bandwidth and buffer allocation, as well as priority and latency. HCIS vendors are beginning to link their provisioning and management with switch vendors' software. For example, SimpliVity has integrated with Cisco UCS Manager. This allows mapping network performance to traffic type, enabling easier management and troubleshooting. Nutanix has linked its management system with networking from Plexxi, which allows Plexxi's low-latency optical network to dynamically adjust network topology to maximize performance.

Connect Your HCIS Nodes Directly to Your Data Center Leaf Switches

This reduces initial system costs and operational complexity. Many HCIS providers assume that some portion of the ports on their switches will be dedicated to end-user access traffic, and these ports will be connected into the data center fabric (see Figure 2).

At a minimum, connect user-facing ports directly to leaf switches (see Figure 3). This approach decreases latency (one switch hop versus three) and complexity (no configuration parameters to match), while increasing reliability (removes two switches in line and reduces the number of cables).

Using leaf ports for intracluster traffic enables the network team to troubleshoot problems all the way to the HCIS node. In the event of a switch failure, a spare switch from a common pool can be substituted.

Map Network Design to the Unique Internal Characteristics of Your Chosen HCIS and Supported Applications

Applications (VMs and containers) and HCIS storage produce many different traffic patterns and loads that are vendor-specific. This requires an end-to-end network design to support storage (replication and sharding, for example) and VM networking requirements. Mixing user access, node-to-node application traffic, VM mobility, storage access and back-end storage traffic on a single network can lead to unpredictable system performance, availability and data integrity (see Figure 4).

Workload-specific requirements such as switch latency, buffer depth, queueing/scheduling, bit serialization delay and OS/hypervisor/TCP stack latency must be considered when selecting switching hardware and networking software for HCIS. Switch latency and buffer capacity, allocation and management can have significant effect on application performance of some workloads, particularly those that are I/O-bound or lack locality of data reference. Data striping/sharding algorithms for storage functions can be adversely affected by buffer depth and congestion. Failure to match the network topology and technology can result in significant and availability issues.

Create templates for critical workload types and deploy accordingly. When designing clusters, consider end-user access, intra-application module access, storage read/write striping, storage replication/rebuild traffic and VM/container mobility.

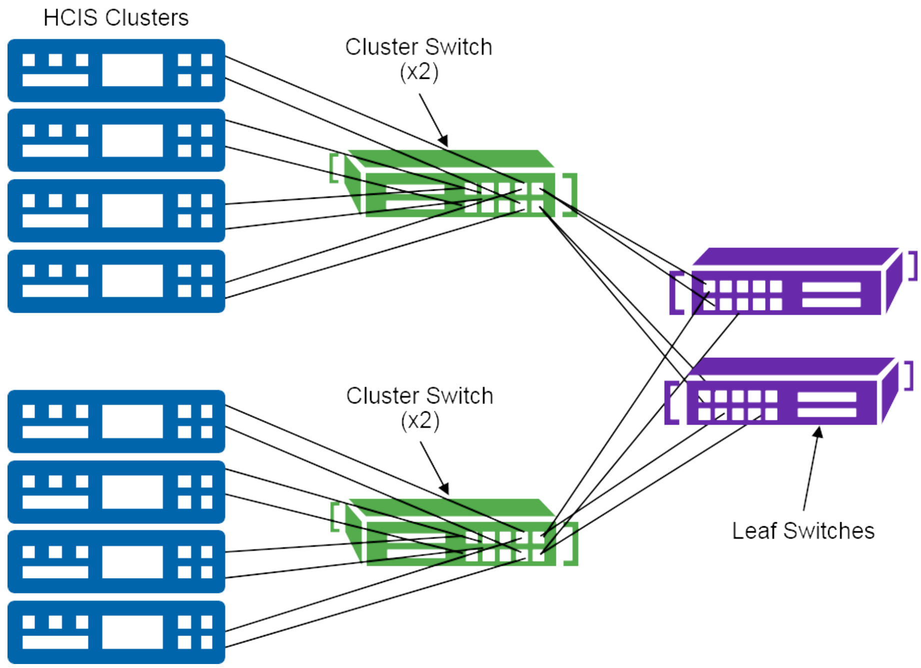
- On 10 Gbps Ethernet, isolate internode traffic by type on a separate virtual LANs (VLANs), or groups of ports, as necessary to ensure predictable behavior. If available, use 25 Gbps ports to reduce total port counts, and logically segment traffic. Tag traffic based upon bandwidth and latency requirements, and allocate bandwidth based upon maintaining overall system performance in the presence of failure.

Select switches that allow you to change VLAN/QoS policies dynamically as conditions change, and then develop QoS policies that will minimize performance problems during system failure. Integrate these capabilities into your management system so configurations can be dynamically adjusted to meet SLAs. Work with your hypervisor vendors to leverage their best practices for network IO control functions.

- For latency-sensitive internode traffic, move to 25/50 Gbps Ethernet and low-latency switches whenever possible, even when bandwidth is not an issue. This is useful because transmission rate (Gbps) directly affects bit serialization delay, which is why many lightly loaded Fibre Channel networks operate at 16 Gbps or 32 Gbps. Consider copper interfaces such as SFP+ often deliver five-times latency improvements over 10 Gigabit Base-T.

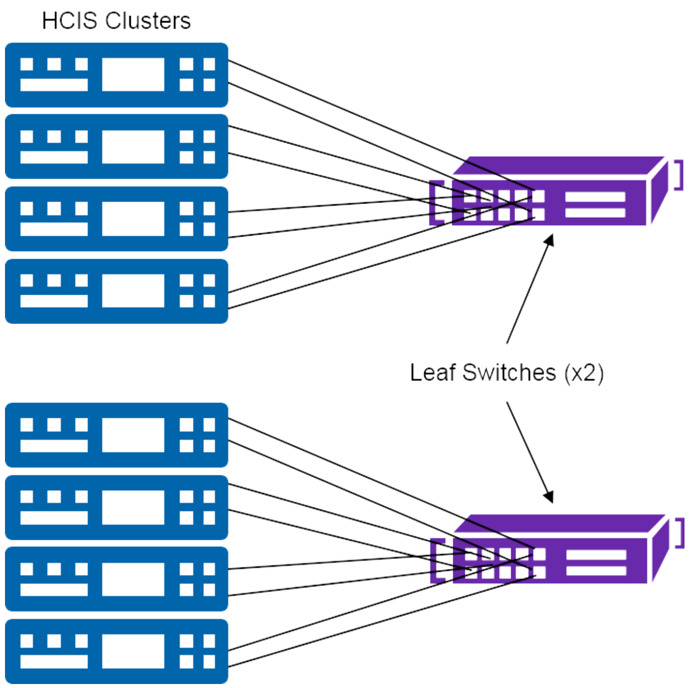
For extremely demanding workloads, consider alternative networks such as InfiniBand or switched optical networking from vendors such as Plexxi.

FIGURE 2 Cross-Cluster Traffic Suffers Increased Latency Due to Lack of Network Integration



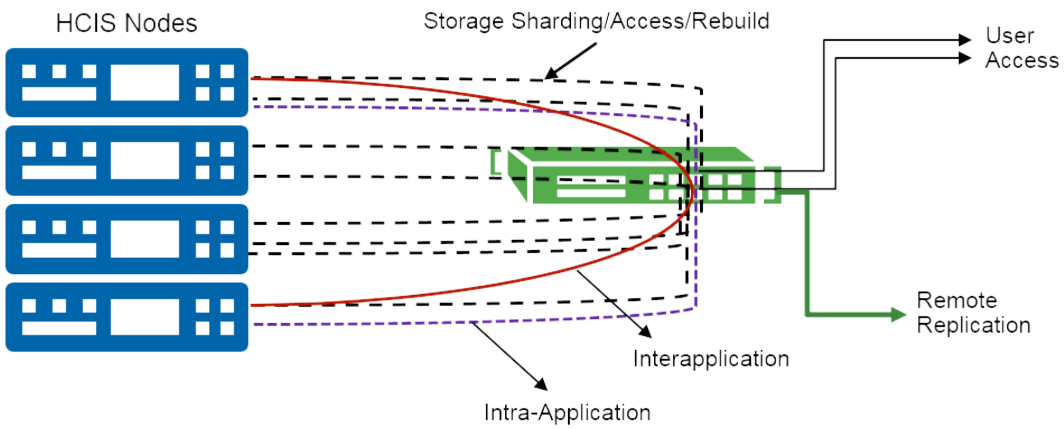
Source: Gartner (August 2016)

FIGURE 3 HCIS Clusters Integrated With the Data Center Network Integration



Source: Gartner (August 2016)

FIGURE 4 Network Traffic Patterns



Source: Gartner (August 2016)

- For latency sensitive I/O-intensive workloads, consider optimizing the entire communications stack. All-flash clusters can experience significantly different performance between data on local storage versus data that is on a different node. In this case, switch latency may play a role, but latency within the OS/hypervisor/container/TCP stack is often the culprit, and optimizing across the entire stack may be the only recourse. Some HCIS vendors are optimizing interworkload communications by creating a kernel-based fast path in Linux/KVM. Protocols such as remote direct memory access (RDMA) over converged Ethernet and RDMA/InfiniBand may also be warranted.

Note
Sample vendors that offer 25 Gbps at 10 Gbps prices: Cisco Nexus 9300 series, Dell Z9100 series.

Source: Gartner Research Note G00312901, Joe Skorupa, Andrew Lerner, Evan Zeng, 19 August 2016

Optimizing Hyperconverged Infrastructure with Plexxi Networking

IT organizations often introduce HCIS solutions to support a specific project such as a virtual desktop infrastructure initiative. Over time, most attempt to move additional applications onto the hyperconverged infrastructure to gain additional cost and simplicity savings. But networking constraints can impair application performance and service quality.

Gartner provides several recommendations for effectively designing and scaling HCIS implementations including:

- Make network design an integrated part of your HCIS cluster design and implementation to ensure performance and reliability requirements are met.
- Reduce the number of required switch ports and cables by consolidating traffic on higher-speed (10/25 Gbps) ports.
- Require that your HCIS vendor's system management integrates with your switch's network operating system or network fabric controller to provide an application/workload-specific view of system performance.
- Connect your HCIS nodes directly to your data center leaf switches.
- Map network design to the unique internal characteristics of your chosen HCIS and supported applications.

For extremely demanding workloads, Gartner recommends enterprises consider alternative networks such as InfiniBand or switched optical networking. Ideal for data-intensive, delay-sensitive workloads, Plexxi networks provide high capacity, low-latency direct machine-to-machine connectivity within and between data centers.

Enterprises can deploy Plexxi solutions for I/O-dependent and storage-dependent HCIS workloads and then gradually migrate other workloads onto the new networking infrastructure for ultimate economics. Plexxi solutions enable an agile, multipurpose network fabric that efficiently supports diverse applications and workloads. Plexxi's programmable and application-defined fabric dynamically changes the network topology, per workload, in real time to ensure data, storage and application SLAs are simultaneously met on one infrastructure. The fully converged Plexxi network fabric extends across the entire enterprise—within and between data centers—ensuring high service quality, reliability and performance for conventional applications as well as next-generation workloads.

Some HCIS vendors recommend customers implement separate networks to isolate distinct applications and workload traffic. Plexxi's software-defined control enables customers to isolate multiple, disparate workloads across a single Plexxi Fabric, eliminating the complexity, and capital and operational expense of deploying distinct, segregated networks.

Plexxi's data center networking product portfolio includes:

- **Plexxi Switch** – a family of compact, high performance scale-out switches. A unique product design combines Ethernet switching with programmable optical multiplexing to create a highly resilient, agile and scalable data center network fabric. Deployed as a single tier, Plexxi Switches eliminate cost and complexity, while enabling high bandwidth, low-latency machine-to-machine connectivity within and between data centers.
- **Plexxi Control** – a software-based distributed control plane that reconfigures the network fabric in real-time to satisfy changing workload demands. Plexxi Control models workloads and automatically reprograms the network topology to optimize application performance and service quality.
- **Plexxi Connect** – an open framework for tightly integrating Plexxi Control software with external service orchestration solutions, DevOps tools, and policy stores. Plexxi

Connect eliminates manually intensive, error-prone configuration and scripting processes, delivering a comprehensive development environment for fully automating the network infrastructure and treating it is a code. A standards-based publish-subscribe message bus makes it easy for external applications and systems to exchange policy, configuration and state data with Plexxi Control. Orchestration systems can provision virtualized compute, storage and networking resources in a holistic manner to accelerate service velocity and enable self-serve IT.

Plexxi HCIS Integration Packs Streamline Automation

Plexxi offers Connect integration packs for leading HCIS platforms including Nutanix (both VMware and Nutanix Acropolis) and SimpliVity. These pre-built software connectors automatically discover HCIS platforms and dynamically re-provision the Plexxi network fabric in response to HCIS compute and storage events, ensuring high performance and reliability for diverse applications and workloads, while optimizing network resources.

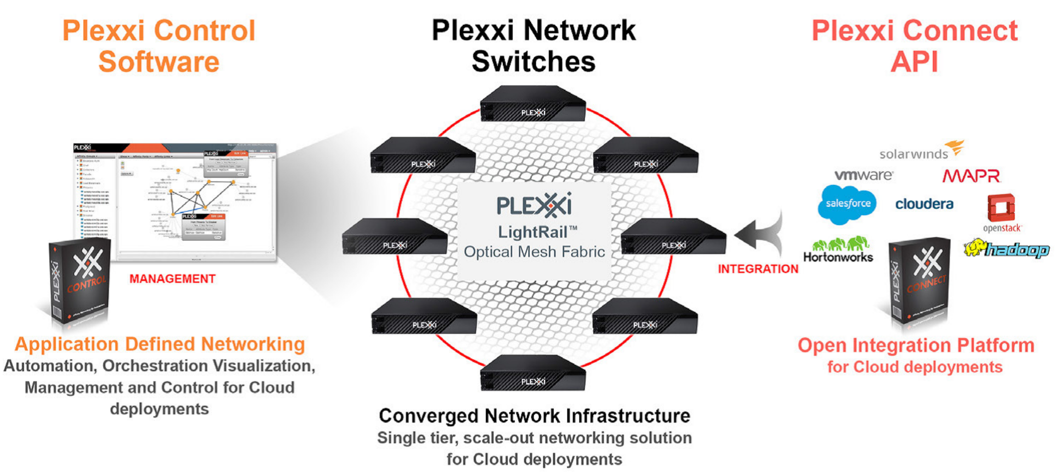
Plexxi's tight HCIS integration enables:

- **Event-driven, open integration** – When new nodes are introduced into the Nutanix cluster, Plexxi Connect automatically gathers information about the physical node, hypervisor and Controller VM.

- **Per-workload optimization** – Plexxi Connect harvests data about individual workloads from HCIS platforms allowing Plexxi Control to determine workload performance, latency and capacity requirements. Plexxi Control adjusts the Plexxi fabric in real-time (scaling up or down) to guarantee per workload, application-level SLAs.

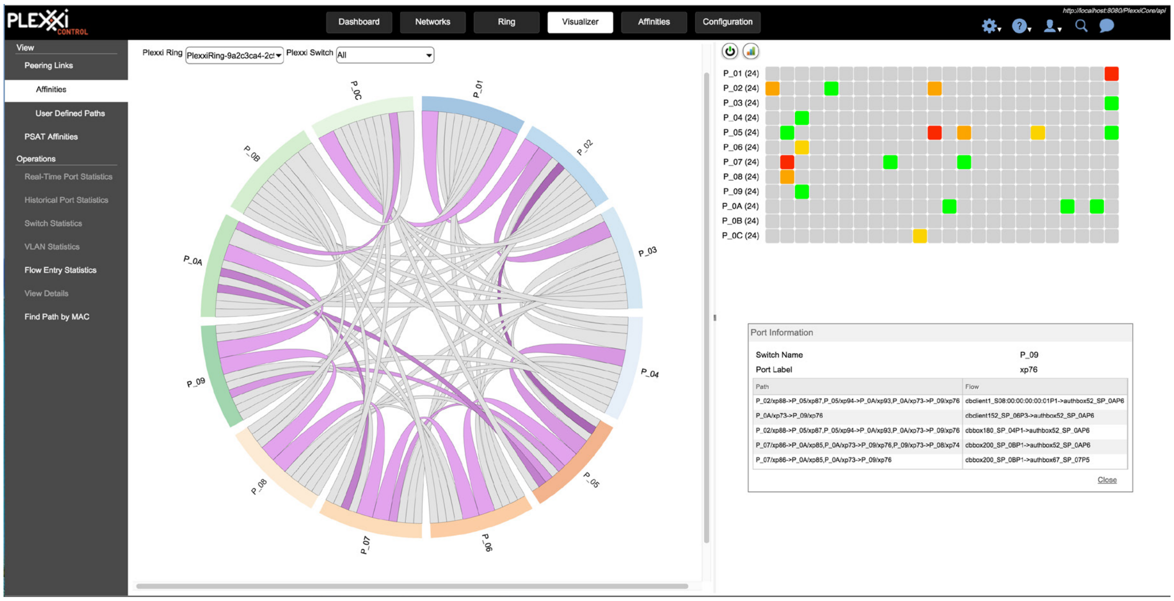
- **Per-workload segmentation** – Plexxi Control leverages unique paths across the Plexxi Switch network fabric to balance loads and make optimal use of network capacity. Unlike traditional leaf/spine architectures, which overprovision and strand capacity, Plexxi can simultaneously utilize all paths and capacity to optimize performance while ensuring service quality.

FIGURE 1 Plexxi's HCN Solution



Source: Plexxi

FIGURE 2 Plexxi Control Workload Visualization Tool



Source: Plexxi

- **Per-workload visualization** – Plexxi Control includes a visualization tool that lets IT administrators see where and how traffic is being distributed across the network fabric, on a per-workload basis. Real-time utilization statistics and graphical heat maps satisfy Gartner's recommendation to "Require that your HCIS vendor's system management integrates with your switch's network operating system or network fabric controller to provide an application/workload-specific view of system performance."

Conclusion – Achieving the Full Benefits of Hyperconvergence

HCIS solutions provide significant CAPEX and OPEX benefits and improve IT service agility by consolidating compute and storage resources, eliminating SANs and unifying system administration. But hyperconverged solutions consolidate a variety of application traffic onto the data center network, which can create performance bottlenecks and impair HCIS initiatives.

Plexxi delivers a software-defined, adaptable network fabric that efficiently supports diverse applications and workloads. By dynamically engineering the data center network around workload requirements, Plexxi networks optimize application performance while making optimal use of network resources. Ideal for today's on-demand applications and elastic services, Plexxi solutions help IT organizations fully exploit all the benefits of hyperconvergence—lower TCO, linear scalability, faster time-to-value—without sacrificing reliability or service quality.

For more information visit www.plexxi.com.