Affinity Networking for Data Centers and Clouds
Solution Overview

**DATA CENTERS AND CLOUDS NEED A NEW NETWORKING MODEL**

Rapid advances in information technology are fundamentally transforming the way enterprises deliver applications and services, and profoundly impacting the way data center and cloud networks are designed and administered.

Server virtualization, consolidation and cloud computing initiatives, clustered computing, and distributed application architectures are reshaping data center traffic flows, escalating bandwidth and performance demands, and introducing network engineering and management challenges for IT organizations.

Legacy data center networks originally designed to enable conventional client-server applications are too complicated, costly and rigid for today's highly-virtualized, on-demand computing environments. To achieve any-to-any server connectivity they employ complex networking protocols and convoluted hierarchies and rely on over-subscribed every-to-every network designs. This puts untenable constraints on workload placement and provides no flexibility for changing workload needs. Simply put, the network has become a barrier to business innovation.

The data center has evolved significantly from the computer rooms of the early client-server days in the 1990s. The monolithic application silos of the past are being supplanted by virtualized, highly distributed application environments. Today's data centers require a fundamentally new networking model driven by the needs of data center application resources and their interconnection requirements. To solve the problem Plexxi has created a new networking model based on the following core premises:

Workloads are served by compute and storage resources that have specifiable affinities to each other and to other workload resources. Understanding the needs of those resources, i.e. these “affinities” is a prerequisite to building an interconnect solution that can serve their needs.

- These affinity needs are knowable as “first facts.” The underlying network interconnect substrate can use this knowledge to compute network affinities directly and make optimal use of network, compute and storage resources.
- Workload needs are dynamic in nature. Their underlying resources cannot be assumed to be stationary within the data center. The network must adapt to shifting workloads and fluctuating traffic flows.
- Data centers will continue to scale. The number of computational entities (i.e. processors/VMs) and the speeds by which those entities are connected will continue to increase. The related network substrate must be architected for multi-generational, cost-effective scalability.

In Plexxi's new networking model, application requirements direct network capacity in real-time, providing direct pathways for machine-to-machine communications, directed by changing workload demands.

---

**What You’re Paying For**

Legacy networks use meshed structures that aim to connect all servers to all servers all the time, creating complex, costly and rigid networks for today’s on-demand computing environments, when in fact applications want specific network connectivity where and when they need it.

**What You Actually Need**

In Plexxi's new networking model, application requirements direct network capacity in real-time, providing direct pathways for machine-to-machine communications, directed by changing workload demands.
INTRODUCING PLEXXI – NETWORKS FOR DATA CENTERS AND CLOUDS

Plexxi has created a scale-out networking product family to build, orchestrate and manage network capacity and functionality directly from data center resident workloads. These products employ a top-down approach that starts with application configuration and ends in the connectivity needed to achieve those needs, rather than the other way around. Whether you are supporting a single application for a single user base or a massive internet-scale data center or public cloud, a Plexxi network will:

- Provide direct server-to-server (east-to-west) network capacity.
- Allow that capacity to be controlled by intelligent software that understands application interconnection needs.
- Make the entire network manageable from the application perspective, either manually via centralized software, or programmatically directly from application orchestration tools.

To achieve these goals, the Plexxi solution exploits a unique and revolutionary concept: Affinity Networking™ with Virtual MultiCore™ technology.

AFFINITY NETWORKING

The notion of affinities is not new. For years, data center architects and application architects have understood that in order to maintain peak application performance, the optimal interconnection of interdependent application resources (such as servers, VMs, databases, network services, storage, etc.) is paramount. Many architects implement affinity concepts manually by co-locating instances of a workload on the same physical switch or adjacent switches. However, as workloads become larger, require more diverse types of resources, and become more portable via virtualization solutions, the task of discerning and manually implementing affinities becomes difficult to nearly impossible.

The Plexxi solution streamlines discovery of workload resources – gathering critical resource information from existing repositories – and organizing those resources into relevant “Affinity Groups.” The process of grouping can be achieved administratively by using a sophisticated GUI or a scriptable shell, or programmatically.

Once workload resources have been grouped, specific policies can be expressed that represent the needs of those workloads for precise interconnection mapping within the data center.

VIRTUAL MULTICORE FABRIC

In order to implement the desired workload needs expressed by the Affinity Group policies, we also need a network that is designed in a fundamentally different manner than existing hierarchical networks. Today’s networks are analogous to single-core processors in that they functionally behave as a large monolithic block in which individual application needs cannot be expressed. Single-core processors used to be the norm in computing, and CPU vendors raced to build the highest clock rates so that application performance could be improved. However as clock rates increased, the complexity of the CPUs became unmanageable, and as application architectures changed from large monoliths to distributed composites, it was no longer useful to employ a single large processing core as a means of scaling performance.

Similarly today’s networks are built as hierarchies (access, aggregation, core; or leaf and spine) where the “network processing” is ultimately funneled up to a large core layer. As link speeds move from 1 GbE to 10 GbE and beyond, these core layers become more difficult to build and manage to the point where it is difficult to envision an economical and practical core layer capable of supporting even today’s 10 GbE servers, let alone emerging 40 GbE capable devices. Beyond the issue of scalability, these single core networks require a “most common denominator” approach for applications. Generally the network is designed around the needs of the most demanding application or applications, and the result is large swaths of network waste and un-needed complexity.

The Plexxi solution overcomes the limitations of conventional hierarchical network designs by leveraging a unique multi-core fabric concept. Similar to a multi-core processor, a multi-core fabric provides each application its own network core dimensioned to its particular needs – specifically encoded directly in that network. Instead of a single network core, the Plexxi Virtual MultiCore solution implements as many cores as needed by the different application groupings – each with its own size and attributes.

Combined with the knowledge of affinity groupings and expressed policy of those groupings, the Virtual MultiCore fabric represents a fundamentally different approach to networking, one that starts with the application needs and ends in the wires needed to achieve those needs.

THE PLEXXI SOLUTION

Plexxi offers the first and only networking solution specifically designed to deliver Affinity Networking and Virtual MultiCore fabric technology. The Plexxi solution includes Plexxi Control network orchestration and control software and the Plexxi Switch hardware platform.
PLEXXI CONTROL

Plexxi Control is state-of-the-art, intelligent global network orchestration and control software that provides centralized discovery, configuration and control functions based on application affinity needs. It performs three primary functions: workload modeling, algorithmic network fitting and global network control.

WORKLOAD MODELING – Control first establishes a view of the server environment. This could include physical hosts, storage and even virtual machines. It then automatically builds a model of the current underlying physical network interconnecting the server environment together with intermediate switches, routers, and/or network service devices such as firewalls, and load balancers.

Once the data center resource information has been collected and mapped to the physical network, they are grouped into “Affinity Groups” of related components, interconnected by “Affinity Links.” These Affinity Groups represent the collection of data center resources that form a coordinated workload. Affinity Links represent the most critical interactions between affinity groups (applications) such as from a web front-end to a database.

ALGORITHMIC NETWORK FITTING – Control algorithmically determines the best individual topology for each Affinity Group and applies corresponding policies to each Affinity Link. These policies express the actual interconnect requirements of each individual Affinity Group as well as the overall network topology that will best satisfy the combined needs of all the Affinity Groups.

Affinity policies can be created across a number of categories that define how the network should be optimized to accommodate the workload needs. The following optimizations are initially supported, and many more will be included with future versions of Plexxi Control:

- **Bandwidth Optimization** – optimizes the amount of network bandwidth that an Affinity Group has access to relative to other network traffic that shares the same paths
- **Hop Count Optimization** – optimizes the number of switch hops between communicating hosts (by reducing to a minimum), which is generally considered a proxy for reduced latency
- **Isolation Optimization** – optimizes the use of isolated paths that an Affinity Group traffic traverses

Hosts or other resources that are not part of an Affinity Group will be assigned load-balanced paths across the network. Control will even detect specific host-to-host communications that could benefit from a more optimized network path, but Affinity Groups with explicit policies will always have priority access to any network resource required to meet their policy.

GLOBAL NETWORK CONTROL – With its unique global view of the entire network and its operational state, Plexxi Control serves as a central network command and control center. Once Control has established the Affinity policies it then orchestrates the desired new network state within the Plexxi Switch network through a federated control process. Control communicates core topology information to the switches. The switches then implement the configuration for the physical layer (L1) as well as the link and network layers (L2/L3) across the network. This has several benefits:

- The network can scale seamlessly
- Switches maintain connectivity and the network self-heals regardless of continued communication with Control

Plexxi Control can algorithmically compute the network topology in real-time and will systematically re-fit or adjust it as Affinity Group needs change. Because Control maintains a network-wide view, it can provide simplified debugging and problem resolution capabilities.

PLEXXI SWITCH

The Plexxi Switch 1 is an innovative switching system designed specifically to enable data center networks that implement directly the needs of discovered Affinity Groups. The switch's advanced architecture integrates Ethernet switching with a unique optical interconnect system. The switch delivers high-density 10/40 GbE server access interfaces while using Plexxi's LightRail™ optical interconnect technology to provide direct rack-to-rack connectivity and capacity dynamically as directed by the changing Affinity Group needs.

Plexxi Switches can be physically connected in a ring topology to create more scalable and more resilient network topologies than traditional tree hierarchies can achieve. Using the LightRail optical interconnections, this ring can carry multiple wavelengths of light (lambdas) so that each switch can connect to any other switch in any number of combinations.

The Plexxi Switch 1 is the first ever Ethernet switching fabric to use optical wavelength division multiplexing (WDM) to create highly fluid and dynamic topologies that can conform to the needs of the application layer.
Switch 1 presents a standard L2/L3 Ethernet switch interface (the electrical domain) to directly-attached servers and storage using 1/10 GbE (32) or 40 GbE (2) for immediate integration into a data center environment. However, instead of utilizing a 3-tier (or collapsed 2-tier) hierarchy using aggregation/core or “spine” layers, Switch 1 uses the two Plexxi LightRail interfaces (the optical domain) to provide direct east-to-west connectivity between server racks. Each LightRail interface supports 24 10Gbps optical wavelengths allowing for the creation of extremely high capacity and scalable meshed networks that can be controlled and optimized based on the workload needs.

In addition, each wavelength can be re-used many times in a given ring configuration in different spans to achieve complex network topologies for each application grouping. Wavelengths also can be bonded together to create very high capacity paths between specific racks. All of these combinations are controllable using the advanced Control software and algorithms.

**PLEXXI NETWORK APPLICATIONS**

- **Cloud/Datacenter Networks** – avoid costly aggregation or core layer expansions when migrating server population to 10 GbE/40 GbE; leverage Plexxi affinity fitting to ensure application workload requirements and policies are met by the network.
- **Multi-Tenant Datacenters** (Public/Private Cloud, Hosting, XaaS) – facilitate workload mobility; direct provisioning of “east-west” bandwidth; Layer 1 tenant isolation; direct implementation and verification of tenant-specific policies; scale-out economics; and physical infrastructure fit- and fill-rate optimization.
- **Big Data Compute and Storage** (Financial Services [HFT, Modeling], HPC [Oil/Gas, National Labs], Unstructured Data [Consumer Internet]) – enable low-latency compute-compute and compute-storage connectivity; on-demand bandwidth provisioning for dataset replication and backup; automated rack-awareness for scale-out MPP database and Hadoop installations; and low-cost, low-complexity network infrastructure.

**WAVELENGTH SPANS**

Traffic assignment and reach of optical channels is flexible and mapping is driven by application affinity.

Channels are efficiently used and re-used through the network.
THE PLEXXI RING

Plexxi switches are physically connected as rings to create networks. Rather than being cabled with “uplinks” to an aggregation layer, each switch is cabled to the others to form a closed ring. The goal of the ring configurations is to create high-capacity, low hop-count (electrical switch hops) network topologies that can be dynamically reconfigured to support changing workload requirements. Rings can be constructed as small as four switches to create very small high capacity server clusters, up to over 250 switches to create very large distributed networks. Rings can also be interconnected to achieve even higher levels of scalability. In addition rings can be connected to existing conventional hierarchical networks to protect and extend existing investments.

The cabling between switches is an inexpensive single-mode fiber. Each cable can carry up to nearly 24 Tbps of data traffic (using 10 Gbps wavelengths) and much more as higher transmission rate optics are used.

CONCLUSION

Fundamental changes in information technology are transforming the way enterprises deliver business applications, impacting the way data center networks are designed and operated. Legacy hierarchical data center networks built to enable traditional client-server applications are far too complex, costly and inflexible for today’s highly-distributed computing environments. Simply put, the network has become a barrier to business innovation.

Plexxi offers the first and only application-driven networking solution, built from the ground up for today’s highly-virtualized, on-demand data centers. Only Plexxi offers a unique Virtual MultiCore fabric that enables dynamically-orchestrated, direct server-to-server network capacity based on workload affinities. Using Plexxi’s intelligent control software and state-of-the art Ethernet switching with optical multiplexing technology, enterprises and cloud providers can build flat, low-latency, high-performance networks that readily accommodate the delay-sensitive, bandwidth-intensive east-west traffic flows that accompany today’s virtualization and cloud computing initiatives.

To learn how Plexxi can help you build a more scalable, agile and cost-effective data center network please visit us on the web at www.plexxi.com or contact a Plexxi sales representative at 1-888-630-7539.