SOLUTION BRIEF

Affinities in Action

Plexxi and CALIENT

At Plexxi, we believe that the network exists primarily to allow conversations to happen between resources in and around the network (from applications to physical and virtual infrastructure). The fundamental issue with today's networks is that they are disconnected from the workloads they serve. The result is an overly complex, static, and brittle amalgamation of protocols and technologies cobbled together to treat a rash of symptoms.

Plexxi makes datacenter infrastructure that aims to tighten the connection between the network and the workloads for which it is designed using a concept we call *affinity networking*. Plexxi's controller-based solution with optically-connected switches is purpose-built to optimize workloads across infrastructure relationships, or *affinities*. Through connectors on our controller, Plexxi can exchange information with other elements of the infrastructure and then optimize for the conversations across the network.

Plexxi offers unparalleled traffic engineering capabilities in its optical backplane. The Plexxi-CALIENT integration allows even more flexible topologies and additional low-latency optical flyway capacity – all driven through Plexxi's Affinity engine. Together, Plexxi and CALIENT provide an excellent example of how real-time flow information can be used to drive network optimization in hybrid packet-optical datacenter architectures.

WHAT DOES CALIENT DO?

CALIENT makes high-density, low-latency photonic switches leveraging 3D Optical MEMS technology. Their S320 Photonic Switch delivers a sweet-spot of high reliability, small form factor, low power consumption, and ease of use that brings the benefits of true photonic switching to the datacenter.

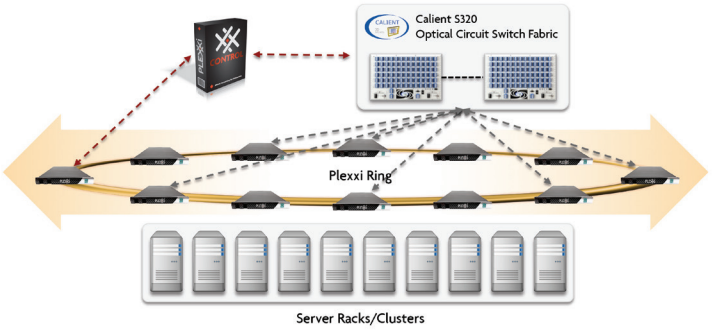
In modern datacenters, the rise of big data, video, and cloud applications is driving massive growth for datacenter network traffic. Operating at significantly low oversubscription rates, traditional networks can quickly become congested, leading to degraded performance for all applications in the datacenter. Countering this growth by designing the datacenter for worst case requires a capital outlay that is prohibitively expensive.

CALIENT's photonic switches make possible a hybrid packet-optical architecture that can scale to support datacenter growth. This architecture helps traditional datacenter networks scale by introducing additional optical capacity, and it allows the network to gracefully reconfigure based on scheduled or real-time events.

LAYING OUT THE HYBRID ARCHITECTURE

At the center of the hybrid packet-optical datacenter architecture remains the packet network. The packet network provides any-to-any connectivity between racks or clusters in the datacenter.

In parallel to the packet network, the hybrid datacenter features an Optical Switch Circuit (OSC) network consisting of a fabric of optical switch elements. Each of the top-of-rack switches in the packet network are connected to the optical circuit fabric.



The key to the hybrid packet-optical architecture is that the packet and optical elements must co-exist and be coordinated by higher-level management layers. In SDN terms, this coordination happens through a contoller. In the Plexxi-CALIENT integration, Plexxi Control serves this role.

PERFORMING THE HANDOFF

In the hybrid architecture, sections of the Plexxi switch ring are connected to the CALIENT optical fabric via 10GbE or 40GbE ports. Plexxi switches are connected to each other via an optical interconnect. When there is a large flow, to offload the traffic from the Plexxi network, the Plexxi switches pass the traffic from ingress switch to the S320 optical fabric, and then the fabric forwards traffic to the egress switch, which sends it to its destination.

The objective in offloading the large flow is twofold: protect the rest of the network from congestion, and guarantee a high-bandwidth, low-latency path for the flow. To do this, two things must happen:

- Once the ingress and egress Plexxi switches are known, a connection between them must be established on the CALIENT S320.
- The offloaded flow must be specified in the Plexxi network so that it can be optimized.

For the two systems to work together, there must be a means of exchanging information. The exchange of information between a Plexxi network and any infrastructure or application occurs through Plexxi *connectors.*

A connector is a lightweight data collection service that can reside either natively inside Plexxi's controller, or anywhere in or around the network that is reachable by the controller. Plexxi Control ships with several built-in connectors, but can also support third-party or user-defined additions. The primary distinction is one of packaging and distribution.

In the Plexxi-CALIENT solution, the connector is an off-box connector that can be deployed on any device that has a Python interpreter and has IP access to the CALIENT S320. The connector must be executed for information to be exchanged between the CALIENT fabric and the Plexxi controller. This can be done manually (as with an initial deployment of either a Plexxi switch or a CALIENT S320), or as part of some other workflow.

When executed, the Plexxi-CALIENT connector determines how the Plexxi network is connected to the CALIENT optical fabric. The connector then sends a request (via SSH) to the CALIENT device to create the fabric connection between the Plexxi switches. Finally, the connector creates an affinity in Plexxi control between the source and destination hosts.

AFFINITIZING THE NETWORK

While it is powerful to understand relationships in and around the network, networks ultimately rely on layer 1, 2, and 3 information to function. Plexxi's networking solution must translate the affinity construct into something that resembles how networks run.

When individual hosts talk across a Plexxi network, the Plexxi switch through which they connect learns the MAC address and the ingress port. The switch passes this information to the Plexxi controller, giving it a complete port-MAC mapping for all attached hosts. If a host moves (as with vMotion), the Plexxi controller can track the new location and adjust the network accordingly.

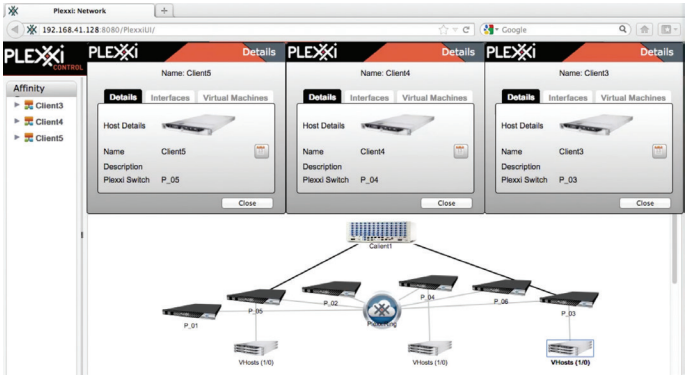
For the Plexxi-CALIENT integration, when the connector creates affinities, those relationships exist between the end hosts for a particular flow. Because the affinities are tied to the hosts and not the switches, it means that any changes due to host migration are transparent to the network.

PUTTING AFFINITIES TO WORK

Simply knowing that there is a relationship between two hosts on the network is interesting, but it does not in itself make the network function any better as a resource. The affinity must be put to work.

Typically, the objective in offloading a large flow is to guarantee performance for that flow while reducing the impact to other traffic on the network. For instance, imagine a big data transaction (as with MapReduce). The goal in using an optical bypass is to provide a high-bandwidth, low-latency path between hosts.

In this scenario, the Plexxi controller uses the affinity between participating hosts to isolate the traffic between the two. Isolation effectively creates a dedicated link, through the CALIENT optical fabric, between hosts. The dedicated path ensures that the flow has both guaranteed bandwidth and the lowest possible latency through the network. Isolating the flow has the added benefit of removing the impact of that traffic from the rest of the network, relieving all other network traffic as well.

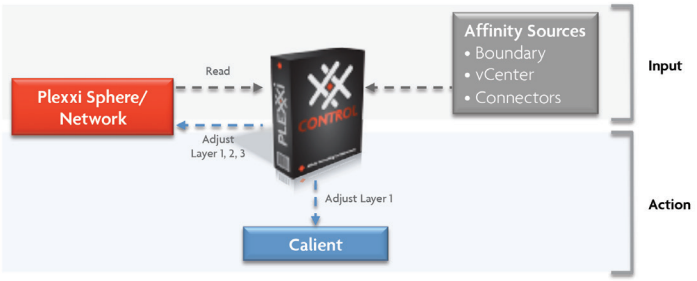


The above screenshot shows an affinity between hosts connected to switches 3 and 5. When traffic between them enters the Plexxi network, the affinity isolates that traffic on the links to the CALIENT S320.

Because the relationship is between the hosts, if the host migrates and now enters the Plexxi network through switch 4, the affinity automatically updates and preserves the optical bypass without any additional cabling or provisioning.

The Plexxi-CALIENT integration can be further enhanced to solicit real-time performance information from an Application Performance Monitoring (APM) tool like Boundary. In addition, scheduled event information about a storage or virtual resouce migration can be used to reconfigure the network to fit the resource requirements. With this type of solution, a large flow can be detected based on 3rd-party information, and that can then initiate the affinities required to optimize the network.

Ultimately, the Plexxi-CALIENT integration is about enabling collaboration across different parts of the infrastructure to satisfy a common goal: optimal application experience. By describing and then optimizing for important relationships, Plexxi and CALIENT have unlocked greater value in the hybrid packet-optical datacenter.



Plexxi, Inc.

222 Third Street, Suite 1100 ◆ Cambridge, MA 02142 ◆ +1.888.630.PLEX (7539) ◆ info@plexxi.com ◆ www.plexxi.com

March 2013