Square Pegs and Round Holes (Network and Applications)

Lori MacVittie, 2013-21-10

#DataCenterWorld #SDN #NFV One size does not fit all.

Shannon Poulin, VP Data Center and Connected Systems Group and General Manager, Data Center Marketing Group for Intel gave the keynote address at Data Center World Fall on transforming the data center for a services-oriented world. Now, that was interesting enough in itself and of course it touched on SDN and cloud. But what really grabbed my attention was a focus on processor design and how decomposition of applications and the network into services and functions is changing the way processors and board-level components are designed.

You see, it turns out that one size does not fit all, and the varying resource and processing models of different types of "things" has an impact on how you put a machine together.

I mean, it's all well and good to say commoditized x86 is the future, and white box machines are going to be the basis for our virtualized data center, but the reality is that that is not a good design idea. Why? Because applications are not switches - and vice versa.
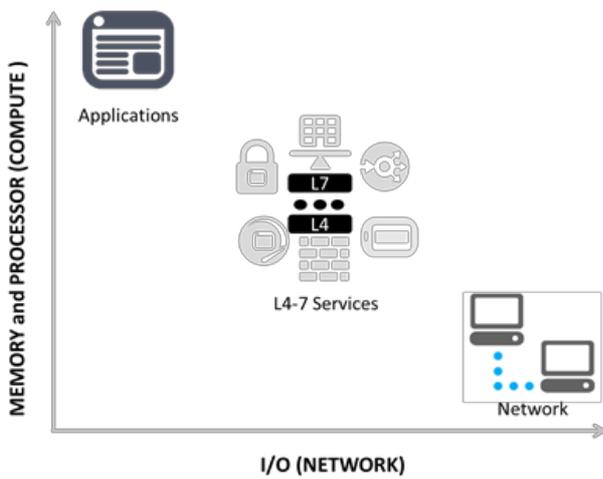
I/O versus COMPUTE

Switches are, by their nature, highly dependent on I/O. They need a lot of it. Like Gbps of it. Because what they do is push a lot of data across the network. Applications, on the other hand, need lots and lots of memory and processing power. Because what they do is mostly processing lots of user requests, each of which eats up memory. What Poulin discussed in his keynote was this diversity was not going unnoticed, and that Intel was working on processor and board design that specifically addressed the unique needs of networks and applications.

What's important for folks to recognize - and take into consideration - in the meantime is that one size does not fit all and that the pipe dream of a commoditized x86-based "fabric" of resources isn't necessarily going to work. It doesn't work as a "resource fabric" for both network and applications because the compute and network needs are vastly different for network functions and applications.

Which means no matter what you do, you can't have a homogenized resource fabric in the data center from which to provision willy nilly for both network and applications. You need specific sets of resources designated for high I/O functions and others for high processing and memory usage.

And, just to throw a wrench into the works, you've also got to consider that many "network" services aren't as "networky" as they are "application". They're in the middle, layer 4-7 services like application acceleration, load balancing, firewalling and application security. These are high I/O, yes, but they're also compute intensive, performing a variety of processing on data traversing the network.

This was somewhat glossed over in Poulin's keynote. The focus on network versus compute is easier, after all, because there's a clear delineation between the two. Layer 4-7 services, though key to modern data centers, are more difficult to bucketize in terms of compute and I/O required.

Depending on what the application service (l4-7) is focused on - an application delivery firewall needs a lot of I/O to defend against network and application DDoS while a web application firewall needs lots of processing power to scan and evaluate data for threats - each of them may have different needs in terms of compute and network resources as well.

What I heard from Poulin is that Intel recognizes this, and is focusing resources on developing board level components like processors that are specifically designed to address the unique processing and network needs of both application services and network functions.

What I didn't hear was how such processors and components would address the unique needs of all the services and functions that fall in the middle, for which "general purpose" is not a good fit, but neither is a network-heavy or compute-heavy system. Indeed, the changing landscape in application architecture - the decomposition into services and an API layer over data approach - is changing applications, too. An API focused on data access is much more network (I/O) heavy than is a traditional web application.

I'm all for specialization as a means to overcome limitations inherent in general purpose compute when tasked with specialized functions, but let's not overlook that one size does not fit all. We're going to need (for the foreseeable future, anyway) pools (and/or fabrics) made of resources appropriate to the workloads for which they will be primarily responsible.

F5 Networks, Inc. | 401 Elliot Avenue West, Seattle, WA 98119 | 888-882-4447 | f5.com

F5 Networks, Inc.	F5 Networks	F5 Networks Ltd.	F5 Networks
Corporate Headquarters	Asia-Pacific	Europe/Middle-East/Africa	Japan K.K.
info@f5.com	apacinfo@f5.com	emeainfo@f5.com	f5j-info@f5.com