Optimising networks for cloud computing and virtualisation

CW BUYER'S GUIDE CLOUD NETWORKING

Networking is not about applications or technology – it is about users. The business pressure to consolidate IT, support empowered users, and provide customised experiences to customers is stressing today's architectures and reshaping networks. The old approach of building networks around applications is too static. Application delivery networks (ADNs) are not designed to support the user-centric collaboration tools, Web 2.0 tools and cloud services that are needed to solve today's business challenges.

Contents

These articles were originally published in the Computer Weekly ezine.

Benefits of networking with a user focus

Andre Kindness looks at how a user experience network focuses on monitoring, controlling and optimising the quality of user experience

CW BUYER'S GUIDE CLOUD NETWORKING

Forrester has identified a new architecture called the user experience network (uXn) which connects users to services that are relevant to the moment, aggregated at the point of use, and originate from multiple locations.

Networking is not about applications or technology – it is about users. The business pressure to consolidate IT, support empowered users, and provide customised experiences to customers is stressing today's architectures and reshaping networks. The old approach of building networks around applications is too static. Application delivery networks (ADNs) are not designed to support the user-centric collaboration tools, Web 2.0 tools, and cloud services that are needed to solve today's business challenges.

Just as enterprise human resources are morphing from static and dedicated pools of resources to ones that come and go with the ebb and flow of the business, transports of information will be created and disassembled within nanoseconds.

IT operations teams must design their networks to accommodate this rapidly changing user environment. Today's application delivery networks must be transformed into uXns. As with the web, enterprise networks will be understood not as a single instance solving a specific issue, but as a fluid and intertwined set of functions leveraging the most optimised set of capabilities and resources for the users. Forrester defines a uXn as a network architecture that focuses on monitoring, controlling, and optimising the quality of user experience.

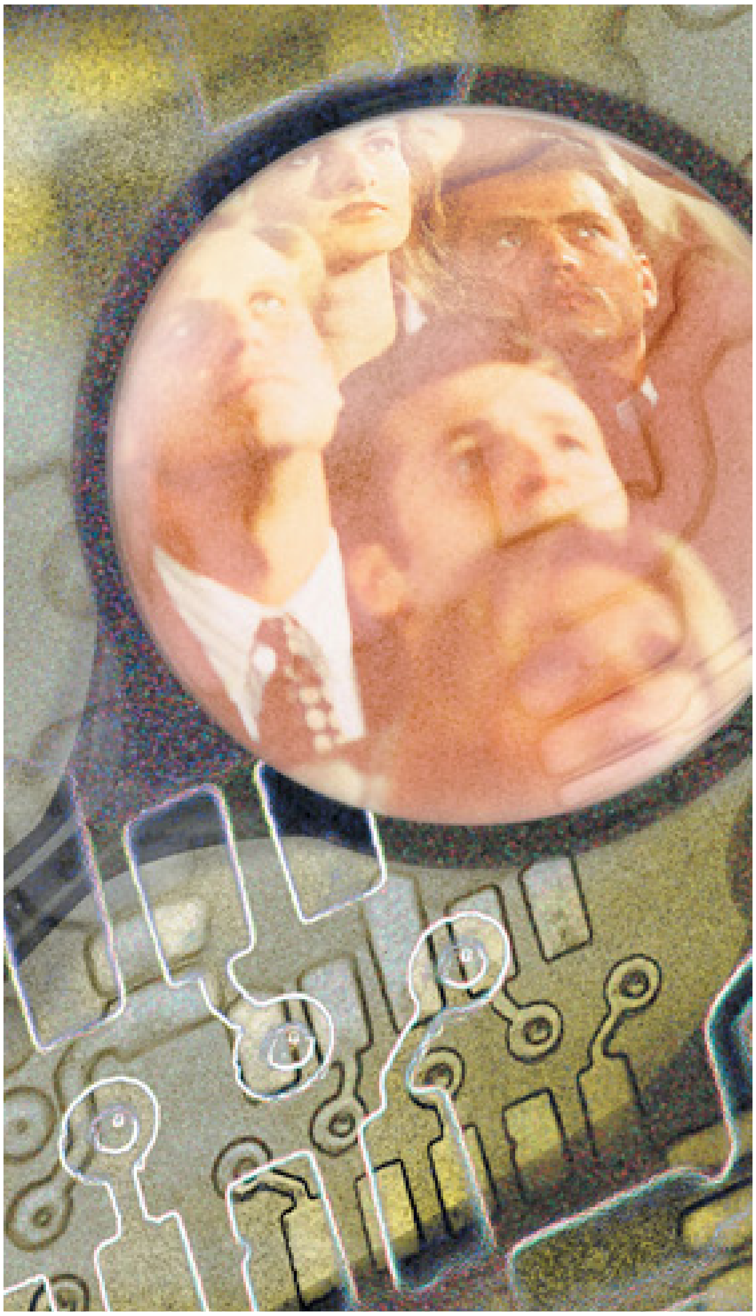
Users could be defined as internal personnel who are working remotely on the road, within the office, or at a branch office – basically, anywhere. They may be connecting and creating a virtual work or personal environment depending on their immediate needs; endless options of hardware and operating systems will be the door that opens into that world.

It is up to the infrastructure to understand the users and vehicles they are using to create the new world. There are customers, suppliers and partners outside your organisation who will demand their own instantaneous virtual world.

To do this they will need a user experience network that has granular visibility to customise services for each user, optimises the transport mechanisms by combining and accelerating the required service and can control the flow based on policies of the business. All three of these capabilities are tied together using a common policy framework designed to set user SLAs.

Building the end-to-end uXn

IT operations must take a page from the virtualisation playbook. Much »



Three companies leading the uXn charge: Citrix, F5 and Riverbed

Companies are creating the necessary hybrid appliances and virtual software that will create streams of services that are optimized and controlled by management software with visibility into the user, location, device, and required services. Three vendors are leading the charge in delivery and end-to-end uXn: Citrix, F5, and Riverbed. What sets these suppliers apart? Each approaches the architecture from different strengths. Specifically, we see that:

Citrix
Citrix's solutions are intimately tied to the user experience. Citrix's pedigree of delivering a virtual world to end users gives it a leading edge in understanding what users expect, and it is able to translate that into a strong uXn architecture. Its ability to deliver a complete portfolio today is still a work in process, and its experience on the virtual side of Citrix might be the step function in being able to deliver a tightly integrated set of optimisation technologies following a user.

F5
F5 provides customers a variety of optimisation, monitoring, and control tools. Its ADC is still the most widely deployed and takes the lead in having a rich tool set that can offer control over the user experience. F5 also has the richest policy interface with iRules and a rich developer network on DevCentral. I&O managers can tap into a large knowledge base to help them connect F5's solutions with the users and create unique uXn services.

Riverbed
Riverbed rounds out WAN optimisation with a solid cloud gateway. Riverbed Whitewater enhances enterprises' storage options by allowing its users to store data in the cloud securely. Riverbed's acquisition of Mazu offers customers an end-to-end uXn that is starting to focus more on the users than applications. The firm also provides an aggressive set of virtual form factors and even a virtualisation platform — Riverbed Services Platform — from which user-centric services can be launched.

The uXn delivery network must be optimised for the user and not the specific application

«like what virtualisation did in the datacentre by freeing applications tethered to specific servers, today's network must deliver a set of services untethered from the application infrastructure.

This delivery network must be optimised for the user and not the specific application. Internal personnel, customers, suppliers and partners will connect to business and leverage multiple parts of the infrastructure to get the service they need.

There may be unique considerations and ignition of technology, but fundamentally, there is a large, overlapping set of capabilities among WAN optimisation controllers, application delivery controllers and cloud gateways – such as caching, compression, quality of service and other services – to ensure optimum experience.

Thus, building a uXn requires that you look at all three of these in tandem. Traditionally, companies have deployed these as separate point solutions. Moving forward, a uXn requires universal policies around delivering an end-to-end user experience across these three delivery technologies.

How do you integrate and rationalise this overlap? Focus on virtual appliances, not hardware appliances. WAN optimisation controllers, application delivery controllers, and cloud storage are evolving from point solutions in hardware form into a hybrid appliance that runs software in a virtual machine dispersed over the infrastructure.

This way you can specialise the uXn virtual machine (VM) on the unique aspect of environment in that area but leverage common functionality from another piece. IT operations can then use policy to focus on building end-to-end uXn services for a user rather than on turning on features in a hardware appliance located in a specific datacentre or branch office. ■

This is an extract from Forrester's *Focus Your Network Strategy On User Experience, Not Application Delivery* (February 2011) by Andre Kindness, senior analyst at Forrester Research. He contributes to the blog for IT infrastructure and operations professionals: http://blogs.forrester.com/andre_kindness

Making Ethernet part of the fabric

Cliff Saran looks at the practicalities of optimising an Ethernet network for virtualisation and cloud computing

CW BUYER'S GUIDE
CLOUD NETWORKING

The latest thinking on datacentre design recommends that businesses deploy virtualisation, where virtual machines can be started (or spawned) and stopped dynamically. But this can have a huge impact on the network.

In a network it is often a good idea to know where server and storage are located. This way, the network path can be optimised to minimise the number of jumps (ie the number of network switches and routers that IP traffic must pass through) between the various server and storage components. A more direct path between server and storage components in the datacentre leads to better performance.

But to get the most from server and storage virtualisation, the physical location of a virtual machine should not be tied down, as this would affect the flexibility of the virtual infrastructure. As a result, a traditional Ethernet network cannot easily be optimised for virtualisation and cloud computing.

Fabric for VMs

A poll conducted at the Gartner Data Center Conference held in December 2010 found that 83% of respondents were using mobility to reassign new locations or shift workloads, or were using policy-based software rules for optimisation.

Gartner believes an approach called "computing fabric" will be required to support the dynamic allocation of virtual machines (VMs), where the network, server and storage act as a single unit connected using a switch. "Cloud computing and virtualisation make networking difficult. Modern blade platforms, such as HP Virtual Connect and Cisco UCS [unified computing system], are integrated with switches [to simplify networking]," said Andy Butler, a distinguished analyst at Gartner. "Fabric computing relies on the network switch being integrated with the server, the network and the storage."

According to Brocade, Ethernet networks are not designed for cloud computing. While network managers have previously been able to optimise networks by managing performance at the network's core, Marcus Jewell, regional sales director at Brocade, says virtualisation means network traffic becomes unpredictable.

Duncan Hughes, systems engineer at Brocade, said: "For the past 20 years we have been using a three-layer hierarchical layer comprising the access layer, aggregation layer and the core layer. Routing would only be performed at the core layer, so if you needed to communicate with a server on a different part of the network [subnet] you would need to travel the network up through the three layers [three network boxes] and back down again." This is inefficient and does not copy well when used in a virtualised environment.

Hughes says that if this approach to networking is used in a virtualised environment, network traffic between virtual machines will bounce up and down the network.

Brocade and other network equipment makers are now selling the idea of a network fabric, which overcomes the network problems caused by virtualisation. The storage networking company has developed what it calls a flat-layer, two-network, self-healing fabric, which it says overcomes the limitation of using a three-layer network topology. ■

Case study: De Persgroep

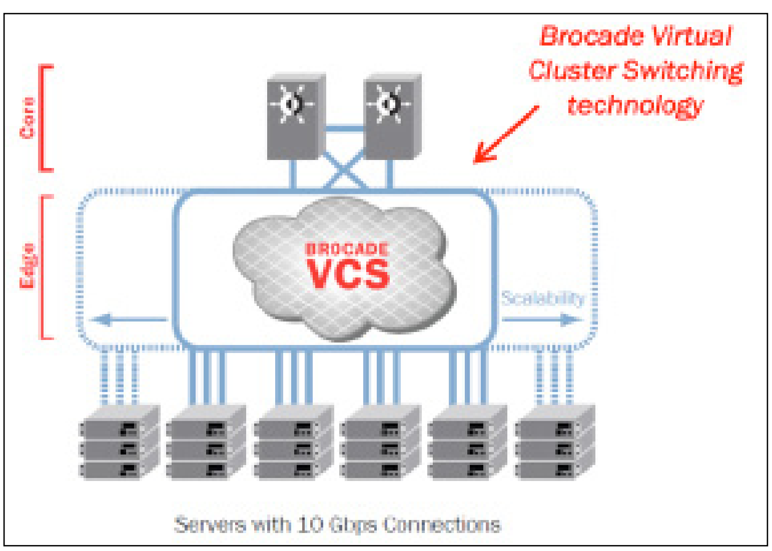
The Brocade network infrastructure is helping Belgian media company De Persgroep manage its datacentre networking. The company, which has grown rapidly through acquisition, has increased the number of servers in its datacentre from 300 in 2007, and supports around 1,000 virtual servers today. Along with office automation, the main applications are editorial and advertising systems, plus its websites, some of which support two million users at peak times, with bandwidth of 2Gbps.

De Persgroep's previous network was experiencing capacity issues due to the growth in virtual servers, and limitations on the number of physical network ports was affecting network performance. Throughput bottlenecks meant that back-up schedules would fail, and overall switch performance and stability was no longer acceptable. Following a competitive tender, the company selected the Brocade VDX 6720 as the foundation to build an Ethernet fabric for its datacentre.

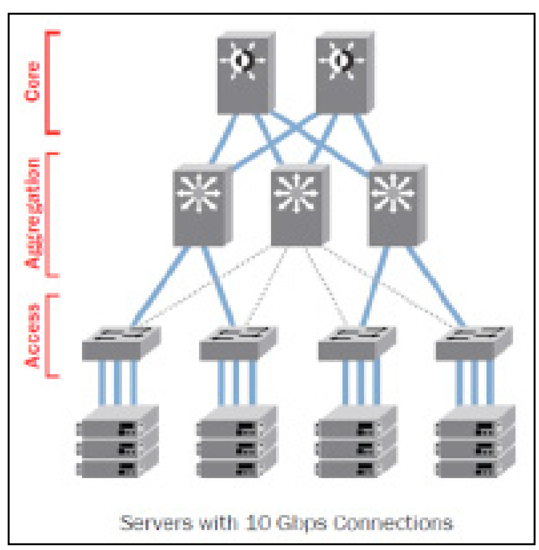
A total of 72 Brocade VDX 6720 switches have been deployed across the fabric-based architecture, creating a single logical chassis with a single distributed control plane across multiple racks of servers. This design has provided compelling reductions in capital and operating costs, while simplifying virtual machine (VM) migration. The deployment is already delivering the desired performance and resilience for De Persgroep.

The Ethernet fabric uses dual 24-port network switches fitted to the top of its server racks to enable each rack to operate its own dual redundant network. Servers in the racks are equipped with two network cards each to support the dual redundancy. The racks are connected to aggregator switches which bring the networks from the servers together and connect them to another datacentre, which is located a few hundred metres away. This provides a so-called active-active system, where both datacentres are operational.

"On our old system you could lose several seconds if a switch failed," said Wim Vanhoof, ICT infrastructure manager at De Persgroep. Such a delay would be enough to cause an application to crash. "On the new fabric, which uses a virtual switch configuration, fail-over is instant," he added.



Ethernet fabric architecture



Classic hierarchical ethernet architecture

NETWORK INFRASTRUCTURE

Datacentre operators are long used to the trade between resilience and performance

Resilience in the datacentre

Resilient switching in converged infrastructure must be able to take some beating, as **Steve Broadhead** demonstrates

CW BUYER'S GUIDE CLOUD NETWORKING

In the world of datacentres and large-scale enterprise networks there has always been some form of perceived trade-off between performance and resilience. Building in resilience is absolutely essential of course, but it has historically affected service and application ability when brought to play. And – in spite of the best planned and designed networks, quality of components and management – problems do arise.

Add in the virtual world to the physical one we've come to know and trust and the stakes are raised again. The result is that suppliers have been forced to redesign their systems to support the virtual environment, maintaining that level of resilience – or improving it – while also improving round-the-clock access to those services, applications and the data that lies beneath.

Resilient virtual switching

One such example is HP's Converged Infrastructure solution – incorporating servers, storage, networking and management. The datacentre is growing ever more critical to the enterprise, whether physical or virtual, in-house or outsourced.

From a supplier's point of view, this means creating a complete system – a converged infrastructure – based on marrying truly compatible components with the best performance/feature set and with as little compromise as possible.

At the heart of HP's Converged Infrastructure (CI) system is the key to the resilience contained within – what HP calls the Intelligent Resilient Framework – that creates a resilient, fully-redundant virtual switching fabric.

Intelligent Resilient Framework (IRF) is designed to combine the benefits of box-type devices (simple, standalone switches, for example) and chassis-based distributed devices, such as a blade switch. The argument is that box-type devices are cost-effective, but can be less reliable and less scalable, and are therefore unsuitable for critical business environments. In contrast, chassis-based devices tick all these boxes but are more expensive and considered to be more complex to deploy and manage. With IRF, then, HP is looking to merge the benefits of both approaches into one. IRF allows you to build an IRF domain, seen as one, big, logical device.

By interconnecting multiple devices through ports (regular or via dedicated stacking) it is possible to manage all the devices in the IRF domain by managing one single IP address (attached to the logical device), which provides the lower cost of a box-type device and the scalability and reliability of a chassis-type distributed device.

In a converged infrastructure environment, an IRF-based network extends the control plane across multiple active switches, enabling interconnected switches to be managed as a single common fabric with one IP address. The claim is that it increases network resilience, performance and availability, while simultaneously reducing operational complexity.

Another key element of the system is HP's Virtual Connect Flex-10, comprising two components: 10Gbps Flex-10 server NICs and the HP VC Flex-10 10Gbps Ethernet module. Each Flex-10 10Gb server NIC contains four individual FlexNICs, so a dual-channel module provides eight LAN connections with the bandwidth for each FlexNIC is user defined from 100Mbps to 10Gbps in 100Mbps increments. From a practicality » «perspective, VC Flex-10 reduces cables and simplifies NIC creation, allocation and management.

HP's IRF put to the test

A series of tests based around the resilience of HP's IRF were created, inducing a series of different failures to see how the solution coped with the problems and what this meant in latency/lost packet issues. We also looked at the day-to-day management of the solution, including what happens when planned maintenance is required, in this case carrying out routine firmware upgrades involving switches reboots.

Our CI for the test was built around HPs A5820 Ethernet switches supporting IRF, then – at the back-end – a combination of the aforementioned Flex-10 technology and standard HP A6120 blade switches and C3000/C7000 server enclosures.

First IRF test

The first test involved seeing what happened when we simulated a failed link between the A5820 switch and a VC Flex-10. In this test both switches are simultaneously active, thanks to the LACP Bridge Aggregation mechanism – a key benefit of IRF being its ability to maintain an active-active state. So, in the event of a link failure, the second link of the LACP Bridge Aggregation and the second switch supports the traffic while the broken connection is repaired.

Looking at the illustration as a guide, note that we experienced 3ms failover time on this connection. However, between servers 9 and 10 we experienced no dropped packets whatsoever. Between server 11 and ESX4 we communicated with only a 1.3ms failover time. As we brought the link back up we experienced just a minor failover time, again across all server to server links, just 1ms in total.

Reverting back to the original situation and testing all connections while the second module was shut down and restarted, we recorded an aggregate failover time across all links of just 1.2ms.

Second IRF test:

For the second test we checked what happened when we simulated an additional bridge aggregation failure – potentially a traffic killer. Testing with a 64-byte ping while this was happening, we recorded just 4ms failover time and, while in recovery mode, a further 36ms between server 9 and server 10 and 23.6ms between server 11 and server 9 – easily our most significant latencies recorded yet, but still both well below our target level of 50ms.

Third IRF test

For test three we are having a really bad day at the office, simulating an additional failed link between the second A5820 switch and the second VC Flex-10, meaning we now have to repair a situation with three concurrent broken links.

Does anyone remember the song "three wheels on my wagon"? In addition to our failures induced in tests 1 and 2, adding this third failure saw us record an additional 4ms of failover time and minimal recovery time latencies.

Already down to the bare bones of communication, we then simulated a classic scenario where one of our redundant switch pair fails or has to be rebooted (maybe for unplanned maintenance, for example).

We saw total failover time of sub 6ms, 4ms on shutdown and <2ms on the reboot. As if it wasn't bad enough already, in this scenario we additionally simulate the second A5820 switch losing all its' IRF links to the first A5820 and thereby losing all connectivity, to prove that there are multiple tiers of redundancy in the solution. For our test case we cut off the IRF, with a 64-byte ping running, with a default configuration of Unit 1 as Master and Unit 2 as Slave.

All three IRF links were cut, meaning that now both units were in master status. At this point in the test we measured a failover time of just 4ms. We now merged the two units, with unit 2 rebooting and the configuration being pushed to slave mode, so no conflicts and both units back up. During this phase we recorded just 1.4ms failover time.

We then tested how the IRF stack can accommodate other virtualisation technology such as VMotion – a VMware technology that enables virtual machines to be migrated live from server to server – through the IRF links directly, to gauge the effect on performance.

The aggregated bandwidth of the IRF links (here 30Gbps) provides best-in-class network performance and low latency for any VMotion events in the datacentre that we've seen. The traffic peaks we can see in the top right and left of the highest graphs above show the VMotion traffic. This test suggests the CI solution is optimised for virtual environments.

Summary of IRF test results

Overall we found the claimed resilience of HP's IRF technology to be justified. In every case we found the system recovered successfully from induced failures, most of which were very severe. We recorded latency/lost packets at every stage of the system recovery and found extremely low failover times – generally in the low milliseconds for complete recoveries, allowing for system elements to shutdown, reboot etc. To put this into perspective, it is not very long since failover times – bringing a redundant device up after a failure of the primary device – in this type of situation were measured in seconds and where sub 30-seconds recovery time was seen as class-leading.

Our firmware upgrade test also ran successfully and recorded very low overall switchover time, just 9ms for the complete upgrade of two switches (master/slave). This augurs well for day-to-day management of what is, in theory a complex CI solution, making it a relatively straightforward administrative task. ■

Steve Broadhead is founder and director of Broadband-Testing