Essential Guide: Optimising hybrid IT infrastructure

In this e-guide

In this Essential Guide:

Many organisations are turning to a hybrid IT infrastructure - using a mix of public and private cloud, often alongside legacy on-premise systems. This hybrid environment presents new challenges for optimising performance, cost, application development and IT operations.

In this essential guide, we include articles looking at best practice and key technologies in server, storage, networks, DevOps and hyperconverged systems to make the most of your hybrid IT.

In this e-guide

Servers

Optimising servers and compute

What does it hybrid IT mean for datacentres, servers and on-premise compute power? We examine the issues.

New model datacentres require hybrid approach

Bernt Ostergaard, guest contributor

The datacentre vision of just a decade ago was focused on a consolidated infrastructure located in a single datacentre - with a hot backup site - consisting of a small number of highly scalable servers servicing a large application pool.

Today, IT organisations are faced with the task of optimising that legacy infrastructure and aligning it closer to fast-changing business strategies.

In this e-guide

IT must provide solid change control and all the predictability of traditional IT, while also providing cloud-like speed for the new world of mobile, big data, and cloud-native apps.

To maximise agility and minimise complexity, CIOs must bring together disparate public and private cloud environments to create a centrally managed and optimised hybrid cloud.

Resources must be allocated efficiently with simplified and integrated management. With DevOps, many new apps are being created on the go, so the corporate IT infrastructure must also reduce error-prone manual processes and automate IT operations

Success is measured by the ability to deliver the benefits of software-defined infrastructure leveraging both internal and third-party cloud services. Datacentre remodelling has both software and hardware components.

Why upgrade legacy at all?

Legacy platforms in industry verticals such as finance and insurance survive because of the risk and cost of replacing them. Replacement projects can fail, with follow-on business losses, lost confidence in the IT department and threats to people's careers. The time and cost involved in system testing and the prospect of a big user retraining programme must be factored in. Against this, a stable platform with tried-and-tested processes remains appealing.

In this e-guide

However, recent research from Temenos in the banking sector found that 14% of costs in banking are IT-related compared with a cross-industry average of 7%. This is caused by multiple factors including redundant, outdated, and/or siloed applications. Something has to give.

Some may choose to mix old and new using tools such as IBM's WebSphereMQ message broker. This can handle multiple operating systems, providing a convenient way of passing messages between applications running under incompatible operating systems. Similarly, virtualisation can disaggregate legacy systems software from the hardware on which it was originally designed to run, allowing users to replace their hardware and consolidate legacy systems on a single server.

The cost of such mix-and-match products comes with continued siloed information, more bandwidth consumption, problems with application security due to lack of software patching, and simply maintaining legacy skillsets. Not many young IT people want to get into Cobol programming or Windows NT4 these days. At some point, the cost and inconvenience of maintaining a legacy system will force even banks to remodel their datacentres.

All businesses are data-driven now. And for many organisations, customer-generated data is being created more quickly than it can be analysed. Any company thinking about internet-of-things data gathering to facilitate better decisions and business growth will certainly need to spend a lot of effort to improve real-time analysis and availability. The challenge is to clean, manage

In this e-guide

and store it, visualise it so it is easy to understand, and automate data processes so they run efficiently and without error.

Customer insights

Data visualisation tools and dashboards, such as the ones provided by Oracle BI Cloud Service, IBM Watson Analytics, SAP Lumira and Microsoft SSRS (SQL Server Reporting Services) in Azure can identify patterns and underpin reliable data-driven decisions. This way, companies can find the insights that unlock performance. The right insights can improve business performance and strengthen customer relationships.

More and more of the corporate datacentre is being virtualised across all servers to ensure seamless merging between on-site private computing and off-site public cloud. IT now faces the challenge of managing and monitoring swift infrastructure changes driven by virtualisation. Typically, this involves consolidation, automation and orchestration, workload replatforming and migration. Companies see their leading-edge peers driving increased efficiency and performance across their entire enterprise with these strategies. But how do they make it happen?

The good news is that corporate management is often willing to allocate significant resources to this transition, and a host of systems integrators are ready to provide support encompassing assessment, design, build, testing,

implementation, support and managed cloud transitions. What are the key steps in this transition, and how is it managed?

Taking the right steps

To address the key question ‘how can we use the cloud?', corporate management must formulate its cloud and data strategy (key phrases could be product and tech innovation, customer insight and market expansion), soliciting input from the IT and networking groups, as well as lines-of-business product and marketing executives to unite operational
and business insights from IT data in real time.

Directly involving lines of business is important because more and more application deployment projects are initiated directly by them and circumvent the IT department - so-called shadow IT.

The cloud and data strategy should aim to transform business decisions, enhance customer relationships and create new revenue opportunities in line with corporate governance, risk and compliance policies.

Any datacentre remodelling strategy must also ensure that the IT group has the tools to monitor, analyse and automate IT operations at a level that supports cross-application data analysis. The IT and network group audits and supports the business needs for big data and analytics infrastructure by providing the right data mining, warehousing and business intelligence capabilities.

In this e-guide

Using reference models

The ‘customer journey' is often used to describe the ongoing relationship between a business and its customers. The corporate cloud and data strategy must support this customer dialogue any time, any place, on any device and across any channel, and deliver value-added services in real time. When planning datacentre remodelling, it is a good idea to use a reference model to ensure consistency and comprehensiveness. There are many vertical industry reference models available from the leading cloud platform providers: VMware, Microsoft and OpenStack providers such as Red Hat, IBM and HP Enterprise.

But technology is only the first step. The key enabler to driving business value from the cloud is people and processes. Companies may need cloud managed services to transform their operational service. Hybrid cloud offers the best of both worlds, optimising existing datacentre space on-premise, with the ability to burst to public cloud when business scenarios demand it.

Datacentre remodelling also entails a hardware side of the equation. Processing and connectivity constraints determined by limits in power, cooling and space as well as lower capital expenditure are forcing firms to converge their infrastructure and move away from siloed servers, storage, data and processes to use available power and space more efficiently.

This converged infrastructure groups multiple IT components (servers, data storage devices, networking equipment and software) into a single, optimised

on-site computing package for IT infrastructure management, automation and orchestration. With a pool of centrally managed computers, storage and networking resources, they can be shared across a wider range of corporate applications. Policy-driven processes can be used to manage these shared resources.

Composable infrastructure

HPE coined the term composable infrastructure to denote its hyper-convergence approach to datacentre architecture (now augmented with its recent Simplivity acquisition). This offers the ability to compose and recompose fluid pools of compute, storage and fabric for any application or workload. In a similar vein, Cisco extends the functionality of its unified computing system management with a software-defined infrastructure (SDI) that treats infrastructure as code, disaggregating computer resources so they can be programmed and automatically managed more efficiently. In this way, infrastructure resources become fluid pools that can be composed dynamically.

Hyper-convergence may ultimately lead enterprises to adopt a serverless computing model. This does not mean getting rid of the hardware on which a service or application runs, but rather the capacity to provide a function as a service in a unified environment.

The first real instantiation of this serverless computing model is demonstrated by Amazon Web Services' (AWS) Lambda product. Launched last year,

Lambda lets users run code without needing to provision servers on AWS's EC2 platform. Once the code is loaded, Lambda takes care of all the rest - provisioning the resources to run the workload, monitoring and managing the dynamics of maintaining these at the right levels. Competing offerings are Microsoft Azure Functions and IBM's Bluemix OpenWhisk, plus Google's Platform Cloud Functions.

Few companies today are willing to undergo a complete datacentre remodelling. The vast majority are on a phased journey from legacy mission-critical, siloed on-site applications to a distributed and disaggregated datacentre supporting mobile users, ad hoc locations and customers 24/7. Future IT datacentre function will entail less application management and more support of lines-of-business apps development and deployment.

Bernt Ostergaard is a service director at Quocirca.

Next article

In this e-guide

Hybrid cloud strategy offers best of both worlds

Cliff Saran, managing editor

According to an IDC report commissioned by Cisco, published in September 2016, cloud adoption continues to grow, but few companies have mature cloud strategies in place.

Hybrid cloud is a significant part of the mix, with 73% of organisations surveyed saying they are pursuing a hybrid cloud strategy, subscribing to multiple external cloud providers and using a mix of cloud and dedicated (on-premise) IT resources.

The study also showed that 78% of organisations are using or planning to implement some form of cloud - a 61% increase from the previous year. Yet only 31% of IT decision makers said they had repeatable, optimised or managed strategies. Although this is an increase from 2015, there is still room for improvement.

Automated management

Automated datacentre management, along with machine learning to understand the events that lead to degradations in system performance, are set to drive datacentre strategies in the coming years.

Analyst Gartner expects that, by 2020, the operating costs of enterprise datacentres will have been reduced by 30%. As a result, the analyst firm suggests that to attract buyers, technology product management leaders at datacentre infrastructure providers should differentiate themselves by providing targeted automation and interoperability software to reduce the cost of operating their products.

According to Gartner's IT Key Metrics Data 2016 report, system costs make up between 18% and 24% (in the case of Linux and Windows servers, respectively), whereas the cost of software and administration ranges between 68% (Windows) and 72% (Linux).

Reducing operating costs applies equally to in-house datacentre operations, and the latest thinking from the large server, virtualisation and cloud providers is to offer businesses a way to bridge their own datacentres with public infrastructure as a service (IaaS) using a hybrid cloud that can move workloads automatically across public and private cloud infrastructures.

In this e-guide

Federated cloud management

An organisation is unlikely to choose just one public and one private cloud. Often, several public and private cloud services need to work together to provide the IT environment the business needs. In many instances, the IT department does not have overall responsibility for the cloud services that business users procure and deploy.

And, as is often the case in IT, getting the constituent parts to work together is a complex and costly matter.

Buying off-the-shelf products can go some way to simplifying this integration. A report from Principal Technologies commissioned by EMC said organisations could gain the savings and advantages of a cloud environment sooner. ‘By taking 1.3 years longer to build your own solution, your organisation runs the risk of not realising the benefits and savings offered by a hybrid cloud consumption model during that period. These include $1.76m savings based on our hypothetical organisation, the report says.

Dell EMC offers Virtustream and Pivotal Cloud Foundry to support both traditional and cloud-native workloads.

In October 2016, VMware said it was working with Amazon Web Services (AWS) to bridge private and public cloud infrastructures. At the time, Mark

Lohmeyer, vice-president for products at its cloud platform business unit, said in a company blog: 'VMware Cloud on AWS will be powered by VMware Cloud Foundation, a unified SDDC [software-defined datacentre] platform that integrates VMware vSphere, VMware Virtual SAN and NSX virtualisation technologies, and will provide access to the broad range of AWS services, together with the functionality, elasticity, and security customers expect from the AWS cloud.

Technology tie-up

At the VMworld Europe conference in Barcelona last October, the company showed how the technology tie-up could be used. Lohmeyer demonstrated how VMware could be used with AWS to move workloads dynamically between the private and public cloud. 'We leverage AWS Direct Connect to use vMotion to migrate private cloud VMware workloads onto AWS, he says.

Also at the event, CEO Pat Gelsinger announced VMware Cross Cloud, which he said solves the challenge of using any cloud for any application. One VMware customer facing this challenge is betting company Ladbrokes, which has built automated provisioning on its VMware system and created an abstraction layer to enable it to run on-premise or in the cloud.

Payment processing company Worldpay, another VMware customer, is looking to use software-defined IT to enable it to jump-start its strategy for software-defined automation. Gelsinger added: 'Users want freedom to choose any

cloud, and you are responsible for control. This is the world of a hybrid environment for decades to come - this is the software-defined datacentre.•

The traditional server companies appear very much on the on-premise side of hybrid clouds. Nevertheless, automation has an important role to play in making on-premise hardware as easy to manage as cloud-based infrastructure.

HPE's The Machine

Hewlett Packard Enterprise (HPE), for instance, has a long-term research project dubbed The Machine, which is a server architecture where storage and memory combine to form a so-called memory fabric. Kirk Bresniker, chief architect at HPE, says: 'We want to change the basic economics, going from relatively scarce amounts of memory.•

In the past, says Bresniker, hardware architects worked out elaborate schemes to speed up access to relatively slow storage devices using cache memory. 'The only time caches work is when we hit the cache, he says. In HPE's vision for The Machine, applications would have direct access to vast amounts of memory.

Some technology from this project is finding its way into next-generation Hewlett Packard Enterprise products. For instance, the latest HPE SAN networking components support embedded intelligence called HPE Smart SAN. This technology is embedded into HPE StoreFabric switches and directors and

enables fully automated SAN orchestration directly from within HPE 3PAR StoreServ Storage arrays.

According to HPE, Smart SAN cuts provisioning and change management time by 90% and reduces the risk of human error, assuring resilience and even enabling proactive risk mitigation when network changes are made. HPE also provides Helion, a secure hybrid cloud with OneView management, which, it claims, can help organisations speed up networking provisioning times by 10 times while reducing configuration errors and operational costs by 50%.

On Microsoft operations management suite Azure, the Microsoft Operations Manager is used for hybrid cloud management. This was recently extended with Service Map, a tool to provide real-time dependency discovery and mapping.

Nick Burling, principal programme manager on Microsoft's Enterprise Cloud management team, wrote in a blog post that the tool helps IT administrators to eliminate the guesswork of problem isolation, to identify surprise connections and broken links in the corporate IT environment, and to perform Azure migrations knowing that critical systems and endpoints will not be left behind. As part of the cloud management capabilities in Operations Management Suite, wrote Burling, Service Map helps IT administrators track dependencies in their hybrid cloud environment, making it easier to manage the complexity of multiple clouds.

In this e-guide

Businesses are demanding more from IT, and among the areas CIOs need to consider is what happens when back-office IT becomes customer-facing. This is one of the challenges that global ticket distribution company Amadeus is facing. Previously, its systems were geared towards supporting travel agents' business-to-business operations. Now that people book flights and hotels directly, the company's IT operations need to meet higher user expectations. If the service is slow or unresponsive, a consumer will find an alternative site through which to book their travel arrangements.

During a roundtable discussion at VMworld in October, Wolfgang Krips, executive vice-president for global operations at Amadeus, said: 'We are automating to take people out of the system. In the future, my people will become automation engineers.•

This is a big transformation. IT has to think about how it delivers services, and applications must adhere to newer paradigms. 'You have to find ways to operate in a hybrid world, says Krips. 'Today, if we have a problem, the supplier fixes it. In a cloud-based world, your applications have to deal with the problem.•

One of the areas Amadeus is exploring is using artificial intelligence and best practices derived from fraud detection techniques in financial services to flag up potential problem areas based on deep analysis of telemetry data from running IT systems.

In this e-guide

On-premise IT still the only way to run certain tasks

Marcus Austin, guest contributor

The cloud comes in a range of shapes and forms. There is the private cloud, which gives you flexibility, but is costly to maintain and configure and has limited scalability.

Then there is public cloud, which gives you immense amounts of flexibility and scale, but sits outside of the firewall.

Many IT directors feel it's not secure. Lastly, there's hybrid cloud, which has all the advantages of the public cloud and the private cloud and none of the disadvantages.

Unfortunately, nobody seems to have cracked the hybrid cloud.

In practice, what you get is a 'hybrid cloud that consists of new applications built in the public cloud, and old legacy apps sitting in the private cloud.

What you need to make hybrid cloud really hybrid is cloud bursting.

Seamless movement

A hybrid cloud that incorporates cloud bursting will allow you to take a workload and spin it up on the private cloud, but if the workload needs more resources, it can be seamlessly moved out to the public cloud and easily work with data sources, no matter where they are - in the cloud or on-premise.

Cloud bursting is therefore a great way for businesses to handle peaky demand patterns, such as e-commerce providers with big peaks in sales at Christmas, or news and sports websites with steady demand that spikes when something big happens, for example the World Cup or the Olympics.

Cloud bursting can also be a useful tool for businesses that need to carry out analysis on large datasets, and for traditional applications such as month-end accounting runs where the demand is predictable but requires servers and storage to sit idle most of the time.

However, not all peaky demand applications are suitable for cloud bursting. Cloud bursting works best for applications that don't depend on a complex application delivery infrastructure or integration with other applications, components and systems that are housed on-premise.

Cloud bursting barriers

There are many reasons why cloud bursting is most suited to non-complex applications and why many applications remain either public or private, but not a

mixture of both. The biggest barrier is compatibility between on-premise and cloud platforms.

Microsoft Azure, Amazon Web Services (AWS), Google and VMware are not interchangeable and probably never will be, so any application that bursts needs to have some translation element to enable the app to talk to different platforms. One of the reasons Azure cloud platform is doing so well in the market is the company's commitment to making Azure compatible with on-premise systems and the platform's ability to run Linux and Windows workloads. This makes it easier to create truly hybrid applications that can burst.

The other major issues with cloud bursting are around bandwidth and latency. Application workloads tend to be relatively small and can be moved quickly, but the associated data is not usually so small and any attempt to move it across the internet in one go will take time.

The upshot of this is that your network must be able to cope with, and prioritise, the extra traffic caused by the increased inbound and outbound data, and your cloud provider must be near enough not to cause problems with latency.

Once again, Microsoft's decision to put two new datacentres for Azure in the UK means it is becoming one of the best bets for cloud bursting. Note that any extra costs for bandwidth from your cloud provider need to be factored into the equation, otherwise the cost benefits of cloud bursting may be considerably less.

Licence considerations

Licences are also an issue. If you are relying on commercially licensed software to power your application, then to get cloud bursting working, you need to be able to extend that licence to the cloud and/or change the contract to allow short-term extensions to an existing licence to support additional use.

Also, there is a lack of management tools that can provide a single overview of the multiple environments and, more importantly, decide when it is appropriate to burst, and that can initiate the move.

Lastly, there are the inevitable security issues that are created by moving data to and from the public cloud and delivering an application on a shared environment.

Solving the cloud bursting problem can - and has - been approached in many different ways, including software-defined storage (SDS) such as Avere's FXT Edge Filers; products that solve the platform incompatibility problems, such as Cloudian's HyperStore; new cloud bursting tools, such those from US startup Velostrata; management systems, such as those from CA and RightScale that can trigger and manage cloud bursting; and systems that treat cloud bursting as a load balancing problem.

Application delivery controllers

Application delivery controllers (ADCs) are next-generation load balancers that are proving to be fundamental building blocks for advanced application and network platforms. They enable the flexible scaling of resources as demand rises and/or falls and offload work from the servers themselves. They also provide a number of other services that are essential to the effective operation of on-demand applications, including:

- Network traffic compression - to speed up transmission.
- Data caching - to make sure regularly requested data is readily available.
- Network connection multiplexing - making effective use of multiple network connections.
- Network traffic shaping - a way of reducing latency by prioritising the transmission of workload packets and ensuring quality of service (QoS).
- Application layer security - the inclusion of web application firewall (WAF) capabilities to protect on-demand applications from outside attack.
- Secure sockets layer (SSL) management - acting as the landing point for encrypted traffic and managing the decryption and rules for ongoing transmission.
- Content switching - routing requests to different web services depending on criteria such as the language settings of a web browser or the type of

device the request is coming from.

- Server health monitoring - ensuring servers are functioning as expected and serving up data and results that are fit for transmission.

ADC products have been on the market for some time and are available from a wide range of suppliers, including A10 Networks, Brocade, Citrix, F5 Networks, Kemp Technologies and Radware.

Incompatibility problem

To solve the platform incompatibility problem, US business Cloudian has created 100% S3-compatible storage that allows users to seamlessly burst data from an on-premise private cloud, using standard server hardware, into an Amazon S3 public cloud storage bucket and back using the standard Amazon S3 API on-premise, private and in the cloud.

To help maintain data and workloads across the different clouds (public, private) and on-premise, CA has produced a unified infrastructure management service, which monitors a number of on-premise services (Nutanix, OpenStack, VMware) and public clouds, including AWS. It then displays the data on a single desktop, showing cost and performance and automatically triggering cloud bursting to additional cloud-based resources based on predefined constraints.

Avere's FXT Edge Filers concentrate on storage, particularly network-attached storage (NAS). The system allows users to burst data and computation to

Amazon's EC2, while cloud compute and on-premise servers can access data in private, on-premise and public cloud with low latency and security similar to on-premise.

US startup Velostrata has a new approach to cloud bursting. It decouples the compute part from the storage, promising no changes to the apps. The compute portion of the app is streamed to the cloud and is up and running in 20 minutes, with the data migrated in the background and the data interconnect between on-premise and cloud optimised to keep latency to a minimum.

As yet, there is no single product that allows bursting of all the main computing building blocks - network, storage and compute. That requires software-defined infrastructure (SDI), where the application runs on the most appropriate systems for the current workload, the data is stored where it is needed, and the network directs traffic to the right place with the right amount of bandwidth. Unfortunately, SDI just isn't there yet, and until it is, IT directors will have to rely on cloud bursting as a stopgap.

Marcus Austin is a service director at Quocirca.

Next article

In this e-guide

CW500: The rocky road to software-defined everything

Lis Evenstad, management editor

The birth of virtualisation brought about a transformation in IT infrastructure. Suddenly, companies were given the flexibility to optimise the use of hardware within a software-defined environment.

That same flexibility is now spreading. From datacentres and storage to networks, the world of software-defined everything is here to stay.

But moving towards this new world takes time, effort and some serious convincing. At the latest CW500 club, experts gave advice and tips on how to do just that.

Rob White, executive director, global database group at Morgan Stanley, said that, done properly, software-defined everything could be a key strategic enabler, opening up new capabilities to the organisation.

‘It has the potential to let us optimise, both in terms of costs, but also agility in terms of deployment across those software-defined things we have in the infrastructure, whether it be storage, networks or datacentres, he said.

‘If you have the control, you are not waiting for the supplier to come in, you are not waiting for boots on the ground to come in and deploy what was traditionally all hardware-based. When you have a level of control where you can tune things, you can also optimise your usage and turn parts of it off in low-demand times.

Software defined networks, said White, meant you could separate the control plane from the data plan, and have a network that evolves with the organisation.

‘In a software-defined control plane, you usually get much greater flexibility to customise, he said.

The long, winding journey

At PaddyPower Betfair, they figured that out a while ago. According to principal automation engineer Steven Armstrong, the company has been on a long journey to get there, but has now more or less moved to a software-defined everything environment.

‘We have taken what we’ve done with virtualisation in the server space and extended it to switching, networking and looking to control the whole datacentre in a programmatic way, said Armstrong.

The company began its journey because of a need to increase the scale and resilience of its infrastructure.

‘Some of the real benefit we have found from it was the ability to scale the datacentre horizontally without compromising the performance of the whole system, said Armstrong. But he added that when the company started out on its journey, it was hard to find other companies that had done the same thing. ‘We couldn’t find any use cases, he said.

Quick scale-up

Anthony Guethert, senior principal at Quintiles IMS, has begun a similar journey. The organisation, which focuses on clinical research and pharmaceuticals, was created by the merger of IMS and Quintiles, which created some real challenges.

One of the company’s biggest projects at the moment is collecting data on the pharmaceuticals treating cancer across Europe. One of issues it encountered was that when it was gathering information, there were delays in the information flow. Moving to a software-defined environment improved its performance significantly.

‘In terms of performance, we were running a lot of databases, collecting data from disparate sources and securing that, said Guethert. ‘It is important for us to increase the value we are getting from suppliers providing the datacentre service, but also from the hardware we are using in the datacentres.

‘Also, by using products that virtualise and abstract things like storage to multiple parallel processing algorithms, we are able to improve performance significantly. At a minimum, we are looking at a four times improvement in processing speed, and that’s using very simple commodity hardware.

‘We have also realised that we need agility to stand up and scale out platforms in different markets. If the opportunity arises to collect data in market areas we haven’t previously done, we would like to act on that quickly and be able to capitalise on those opportunities.

Huge learning curve

But although the benefits are clear, moving to a software-defined environment brings challenges and learning curves.

‘You need to overcome the fear of the unknown, said Guethert, adding that a period of budget restraint is actually a good time to make changes.

‘Convince people by putting together two or three exceptionally good options that are financially compelling and you find that people are motivated themselves, he said. ‘When there is budget pressure, there is actually a massive opportunity.

Guethert’s view is echoed by Morgan Stanley’s Rob White, who said one of the disadvantages is that the market is still moving very slowly towards commodity. Another issue is that organsiations are often silo-based, there are different

budgets for different parts of the organisation and the culture is just not there yet, he said.

‘Our organisations were more or less created in an era when you had silos, he said. ‘Software-defined is an amalgam of those resources, so you have to plan the roll-out of them across organisational boundaries and force teams to work together in ways they’re not used to. There will be growing pains.

White added: ‘Cultural change in any organisation is the most difficult change you’ll ever do. You need to find the right levers. Most boards, I would imagine, are not going to want to do this full scale right away. They’re not just going to sign off on your plan to move to software-defined everything next week.

‘What you have to do is start small, pick your battles and prove yourself.

PaddyPower Betfair’s Armstrong said organisations should start off with an initiative ‘at the base level, using it to solve a specific problem.

‘Let’s say the network team is having the same problems every day, he said. ‘Start small and show them how a different approach could tackle that problem in a way that could really help change the way they do lots of work.

‘A lot of people have a fear of change, so if you can bring them along individually on that journey, that helps.

In this e-guide

Networking

Optimising networks

Using hybrid cloud, with its mix of on-premise and public cloud systems, puts new demands on the corporate network. We look at the key strategies for hybrid IT networking.

How to move to a network hybrid cloud architecture

Andrew Froehlich, guest contributor

Uniformity is the key to any successful hybrid cloud migration. Without consistency from a performance, policy and scalability standpoint, you're going to bump into issues at the network level when managing the various private and public cloud segments in the overall enterprise network. These are the types of issues with the potential to harm the growth and technical prowess of businesses for years to come. Fortunately, network hybrid cloud architecture problems are easy to avoid -- if you plan ahead.

Understand the network hybrid cloud architecture

When it comes to developing a strategy to migrate applications to the public portion of a hybrid cloud, a common mistake is to ignore the impact the migration will have on the performance of the overall enterprise network. To many network design engineers, the public cloud is often thought of as a separate entity from the private cloud it connects with. In truth, when the decision is made to use public and private cloud services in a hybrid architecture, the public cloud becomes nothing more than an extension of the privately managed network. It just happens that this part of the network is managed by a third-party provider. Once you have this mindset, questions surrounding the impact of the network in a hybrid setting become crystal clear.

Scalability and elasticity

Beyond Capex benefits that the public cloud can provide, companies that move to a hybrid cloud model will also want to tap into scalability and elasticity benefits. The biggest challenge with a hybrid cloud architecture from a network perspective is maintaining similar scalability and elasticity characteristics in on the private cloud side. While public clouds have plenty of resources to fall back on to provide scalability and elasticity, private data centres are far more static in nature. Fortunately, we have virtualization to help even the playing field. Virtualization of network infrastructure components -- including routers, switches and security tools -- can greatly improve the ability to spin up, down and move networking resources wherever they are required.

High availability and accessibility

Providing the appropriate level of data centre resiliency is also a part of the network hybrid cloud architecture process that needs to be considered prior to moving production apps and data into a hybrid cloud model. Connectivity between the private and public portions of the hybrid cloud should include sufficient redundancy to adhere to service levels the company demands. For some organizations, redundant internet connections used to connect to the public cloud may be sufficient. For others, it may require multiple private WAN connections from different providers to allow for full path diversity.

Geographic accessibility is also something that should be addressed at a network level. To provide a uniform experience, it's important to understand where end users will most likely be accessing your computing resources from a geographic perspective. If most end users work out of an office connected via the corporate LAN, then it makes sense to maintain the applications and services they use within the private portion of the hybrid cloud. If, however, end users will be spread around the world, a larger, public cloud is better suited for providing geographically distributed cloud entry points into the public-facing portion of your hybrid cloud.

Application data flows

With today's modern applications that use distributed architectures, having a precise understanding of data flows is crucial when determining the optimal

deployment location of various resources. For example, if you migrate the front-end application to a portion of the public cloud but maintain databases and authentication mechanisms within the private data centre, WAN traffic between the public and private clouds will increase significantly. Creating an application dependency map will help to identify distributed services each application interacts with.

Once this is complete, an application profile baseline can be used to collect and quantify throughput over a specified time frame. Armed with this information, decisions can be made as to where applications should be located. Ultimately, this will help reduce overall bandwidth requirements across expensive WAN connections.

Managing hybrid and multicloud environments

Lastly, network administrators must understand that operating an enterprise network in a hybrid mode is going to be more challenging to manage. Network hybrid cloud architecture designs require the creation and management of IP subnets, virtual LANs and network and security policies. Often, the decision is made to build and manage the network framework using manual processes. Essentially, the public cloud is built to look and function as similarly to the private cloud as possible.

However, as companies continue their expansion into the hybrid cloud, it becomes increasingly challenging to maintain uniformity between private and

In this e-guide

public cloud spaces. In situations where an organization is expecting significant growth into the public cloud, now is the time to investigate cloud automation and orchestration, multicloud management and cloud access security broker platforms. Executing a hybrid cloud management strategy prior to migration is a decision that could save you hundreds -- or even thousands -- of administrator hours each year.

In this e-guide

Adoption of cloud computing means updating networks

Amy DeCarlo, guest contributor

Predictions of rapid-fire growth in the adoption of cloud computing are coming to fruition, with an astonishing number of large- and smaller-scale projects on the docket. Enterprises are flocking to the medium -- not just as a way to cut expenses and help them scale better, but as a means to become more agile and competitive. The expectation is demand -- and cloud traffic -- will only continue to soar. Cisco, for example, said more than four-fifths of all workloads will be processed by cloud data centres by 2019, with global IP cloud data accounting for 83% of all IP traffic.

At this fast and furious pace, there is a real concern that organizations racing toward the adoption of cloud computing aren't taking a methodical approach to updating their network infrastructures. Given the tremendous growth in hybrid clouds -- where workloads are moved between and among multiple clouds, including on-premises facilities -- not having a network optimized to support cloud traffic can be an unfortunate oversight that could impede performance and introduce instability.

Adding bandwidth is only one consideration; IT also needs to consider how design elements and execution strategies may help or hinder efforts to get the maximum benefit of working with the cloud. Among those considerations:

- Does network gear optimize traffic based on the type of workload?
- Is network infrastructure robust, and does the organization have tools in place to ensure network stability and security?
- Does the network provide a consistent and high-performing experience for the end user?
- Has the organization committed to emerging technologies, such as software-defined networking, to enhance its investment in the cloud?

Better by design

One of the keys to having a network that facilitates optimal cloud performance is manageability. This can be a challenge, as organizations look to connect and migrate on-premises application workloads to resources in off-site clouds. From an infrastructure perspective, enterprises are looking to streamline their networks to get a better handle on the environment.

Employing techniques such as implementing a flatter network design -- which removes the aggregation layer and applies a one- or two-tiered architecture -- can simplify both planning and administration. Removing a switching tier allows

direct connectivity between servers and nodes, thus reducing latency and shaving operational support expenses.

Many businesses also opt to segment their networks along business lines. By taking this approach, each business line gets the computing and application resources it needs.

Service management and security

IT execs must also have the right service management tools, so they can track usage and administer resources across their environments. Providers such as ServiceNow can meter usage across multiple locations and the cloud. These tools employ automation to support dynamic resource allocation as capacity needs fluctuate. Service management options also ensure the departments using the network are charged accurately and efficiently.

Security is critical -- not just to ensure cloud resources are protected, but that they are also available. The virtual and highly distributed nature of cloud computing obscures visibility and makes identifying and isolating threats challenging.

A new crop of cloud security tools has emerged that address not just the visibility challenge, but also the need for faster detection and mitigation in what can be very complex, sprawling environments. Software like the Incident Response Platform from Resilient Systems -- now owned by IBM -- works with

firewalls, as well as other detection and prevention appliances and software, to detect potential issues and initiate a response.

Security products are one part of the equation. Policies and practices are just as important. Organizations must ensure they have crafted effective policies, and then communicate those policies to staff and external partners who may have access to network resources.

All of the security elements, along with service management, are critical to maintaining a stable and highly available network infrastructure that provides the conduit for the adoption of cloud computing. It is crucial that enterprises make the investment upfront in design to support a network that is optimized to carry cloud traffic. As too many organizations are learning, simply adding a bigger switch isn't enough. Instead, IT has to look at network design holistically and consider future innovations before it invests.

Next article

Enterprise SDN adoption rapidly approaching tipping point, claims report

Alex Scroxton, networking editor

The vast majority of IT decision makers will have begun to eliminate their legacy networking infrastructure in favour of adopting a software-defined networking (SDN) model by 2021, suggesting that the technology is fast approaching a tipping point, according to a study conducted for software-defined wide area network (SD-WAN) specialist Riverbed.

Its *Future of networking* study canvassed 1,000 IT leaders in nine countries, and found virtually unanimous agreement that legacy networks were starting to cause enterprises significant pain because they were holding back wider digitisation programmes and cloud adoption.

In total, 98% of those quizzed agreed that next-generation networks were critical to keep up with business, 97% agreed that legacy networks were struggling to keep pace with changing demands, and 91% said their cloud strategies could only reach their potential with a next-generation networking underlay.

Just under 60% said they were having to firefight cloud issues resulting from legacy network failures at least once a month.

As a result of this, 93% said their organisations planned to trade up to SD-WAN in the next four years to help manage cloud and hybrid networks, something just 4% were taking advantage of right now.

To some extent, this data reflects the findings of the most recent Computer Weekly/TechTarget IT Priorities report, which indicated that the idea SDN was struggling to gain traction outside of major telecoms operators and enterprises with widely distributed sites or campus networks - something that now looks set to change.

‘This survey and the resulting data reveal the tremendous pressure that IT decision makers are under to execute their cloud strategies, achieve digital transformation and keep pace with the speed of innovation that is the norm in today’s hypercompetitive markets, said Jerry Kennelly, co-founder and CEO at Riverbed.

‘It was almost unanimous that to have a successful cloud strategy, organisations must adopt next-gen software-defined networking immediately to support it. Riverbed is dedicated to ensuring this transformation is successful and that organisations are able to achieve the operational agility needed to thrive through the use of SD-WAN for cloud networking.

Riverbed’s report found that the gap between the understanding of the need to refresh legacy networks and the actual deployment of SDN played out consistently across multiple verticals, from financial services and retail to

energy, manufacturing and utilities, and even communications and IT services businesses.

The survey also showed that IT decision makers would be willing to sacrifice a lot if it meant they no longer needed to spend so much time tinkering with dysfunctional network infrastructure; 54% said they'd suffer a longer commute, 53% would take a shorter lunch break, and 42% would give up coffee. The full survey results can be downloaded here.

Next article

What is intent-based networking, and what can it do for your business?

Rene Millman, guest contributor

Cisco made a big push into the world of intent-based networking this summer, and some are touting it as the next big thing in enterprise networks. In its announcement in June 2017, the firm said this new type of network would 'recognise intent, mitigate threats, and learn over time.

But how does it do this and, more importantly, if an organisation needed it, how would it implement it?

According to Gartner, intent-based networking looks to be the 'next big thing in networking.

In a blog post, Gartner analyst Andrew Lerner says it is not a product or a market, but instead 'a piece of networking software that helps to plan, design and implement or operate networks that can improve network availability and agility.

Amol Phadke, who is global network strategy and consulting practice lead within the Communications, Media & Technology (CMT) industries for Accenture, says that intent-based networking is essentially about the network deterministically

following the intent of the operator/operations through the use of automation and analytics.

"Intent-based networking has been around for a while but it is growing in importance given the real-time nature of services and devices that sit on top of the pervasive network, he says.

According to Kireeti Kompella, chief technology officer of Juniper Networks' development and innovation team, intent is characterised by the high-level, declarative nature of its statement, and the absence of implementation detail that leads to complexity and inflexibility.

He says that at a basic level, you say 'what you want rather than 'how to do it. The implementation details are filled in as the intent is instantiated in the network.

'This is important as it allows networks to be more agile - the implementation can be changed on the fly, in response to changing network conditions, he says. 'Intent-based specification also means that much of the mundane operational tasks can be eliminated, allowing IT staff to work at a higher level, meaning lower costs, fewer errors, greater security, reliability and optimality.

Ronan Kelly, chief technology officer for EMEA and APAC at Adtran, says intent-based networking is 'not a new concept.

‘In fact, it has been in existence and operation in its early forms in operator networks for many years. One of the firms that Adtran acquired almost a decade ago, Luminous Networks, was one of the pioneers in this space, where in a similar architecture to today’s emerging software defined networks (SDN), they had separated control and forwarding plane functions, with the control plane residing on commodity compute resources in the cloud, he says.

Minimal human interaction required

Kelly adds that in this environment, the human interaction required to build services and circuits across networks with hundreds of network elements, was minimal.

‘An intent was entered into the system to build a service from point A to point B, with a particular set of characteristics, and the cloud based computer resources would calculate the available paths that could deliver on the requirements specified, with ultimately, a primary and fall-back path being identified, he says.

Once the path was chosen, the cloud-based control took responsibility for programming each of the forwarding plane network elements to deliver the service and reprogram them in the event of a network path failure. He says this early implementation demonstrated the potential of an intent based approach.

‘The level of detailed interaction with the network elements was greatly reduced if not eliminated, as was the scope for human error.

Problem solving

The networking industry is full of seemingly good ideas that eventually go nowhere. One has to ask whether intent-based networking is going to help enterprises.

Phadke says that in the enterprise, most companies are looking to simplify their network by utilising SDN-based infrastructures. As their requirements change, they are looking to leverage intent-based networking to manage the infrastructure.

The use cases are diverse - proactive/predictive network-assurance, elastic bandwidth deployment, real-time network topology change, dynamic service orchestration and modifications, he says. ‘The technology will either help with bottom line benefits, top-line growth, enhance the customer experience, or any mix of the three.

Phadke sees intent-based networking being used by service providers and, over time, in enterprises as a key tool in the simplification of the network, operational systems and operational processes.

The potential benefit for enterprises will be the ability for them to command network resources from their carrier, as and when they are needed, right down to application flow levels of granularity. ‘This approach will ensure they are paying for what they need, but not restricted when their needs change, he says.

Going mainstream

Phadke believes the technology will see significant acceleration over the next two-to-three years, ‘although it’s hard to predict because there are a number of factors as play such as technology maturity, operator pace of transformation, operational readiness and availability of investment.

Kompella sees the industry ’inching toward autonomous networks as we infuse automation frameworks with machine learning and big data analytics. He says the path to these networks relies on telemetry, automation, machine learning, and programming with declarative intent.

‘Take automation for example. Today, we automate topology discovery, path computation, and path installation. We have bandwidth reservation that is responsive to traffic changes, but we need smarter auto-bandwidth, for example, is that traffic spike due to downloading the latest Beyoncé song or a DDoS attack? says Kompella.

He adds that when networks are self-driving, there will be automatic service placement and service motion; specific upgrades based on configured services; and inductive network response based on machine learning.

But he warned that while significant barriers exist to developing fully autonomous networks, they should motivate the development of intent-based

networks. ‘It took 10 years for self-driving cars to advance from vision to prototype. I don’t expect an autonomous, self-driving network to take this long.

Kelly says that some of the early implementations of intent-based networking in the datacentre industry are already in place and operational today. A lot of the early initiatives have been spearheaded by open source projects headed up by the likes of Google and Amazon.

‘The impacts of this are far reaching where datacentre switching, optical transport, and fixed and wireless access will all be impacted, says Kelly.

SDN’s role in intent-based networking

Software-defined networking is critical to intent-based networks, says Phadke. It’s the underlying network that needs to be software driven for any of the higher layer intent-based algorithms to be successful.

“As network actions become more algorithmic and deterministic, machine learning plays a key role both in optimising costs and enhancing customer experience. Similarly, advanced automation is important to fully leverage the ‘softwarisation’ of network as well as the operational processes.”•

Kelly adds that machine learning coupled with Open SDN control and standardised data models amplify the potential for intent-based networks where insights about both normal and abnormal traffic patterns can be compiled. This permits the optimal operation of the network using the available assets. For

example, every network goes through a 24-hour cycle of peak and off-peak traffic.

‘When using technologies like NGPON2, the potential exists during off peak time to migrate all optical networking terminal (ONT) endpoints onto a single wavelength, permitting the shutdown of large portions of the access network equipment when it is not needed, he says. This reduces power drawn, heat generation, and accordingly necessary air-cooling requirements.

Kelly says that with AI implementation, greater intelligence can be applied to such scenarios.

‘It will take into account the day of the week, whether it is a holiday period or a major sporting event is on. It considers whether the current weather’s likely to increase data usage, or if there has been a major news event that will drive usage, he says. ‘All of this permits real time optimisation across entire networks that wouldn’t be possible without AI [artificial intelligence] and Open SDN control.

New infrastructure planning

Network infrastructure planning for the 2020 Tokyo Olympics started in 2015, but Kompella says that through an autonomous, self-driving network, planning for events such as the Super Bowl will take only a few days.

‘As IT infrastructure is able to order and deliver itself, then self-organises and optimises on site, he says. ‘Maintenance will be proactive rather than reactive (as it usually is today), meaning that components and systems will be fixed before an outage happens instead of the failure occurring and causing traffic disruption.

In the next 12 to 18 months, the intent capabilities embedded into the leading Open SDN controllers will continue to evolve, permitting ever more flexible use cases.

‘The most exiting piece is the potential innovation that emerges from outside our industry, thanks to the Open API approach and the use of Open SDN controllers with standardised data models, he says.

An improvement in network availability and agility could mean intent-based networking being an essential part of an organisation’s digital transformation plans.

Next article

In this e-guide

Storage

Optimising storage

Organisations need to manage data across in-house and cloud-based systems in a hybrid IT environment. Find out what to do and the technologies you need.

Hybrid cloud storage: What it is and how to deploy it

Bryan Betts, guest contributor

Despite all the hype around cloud storage, it still faces hurdles when it comes to commercial uptake. Those hurdles include concerns over access speed and latency, as well fears around security, compliance and data portability.

So, storage and cloud suppliers have come up with a potential solution. With hybrid cloud storage you do not have to use cloud-hosted storage for all - or indeed any - of your data. Instead, data can reside on-site, in a private cloud or

in a public cloud - as appropriate for your performance needs, the economics involved, regulatory compliance and, of course, your risk assessment.

The most common enterprise use for cloud storage today is for off-site backup and archiving, as a relatively inexpensive way to help protect against technology and site failures. These applications are also less sensitive to latency and bandwidth limitations, especially if backup can be done from a snapshot.

However, while this can mean using a public cloud alongside a private cloud, it is about as hybrid as using disk for your primary data and tape for the backup. That is, the two are not integrated, but perform separate roles in the overall IT infrastructure.

Hybrid cloud storage more accurately means using on-premise storage and storage in the public cloud to create a greater overall value - as a kind of mash-up. You could have some data on one and some on the other, depending on its risk classification or its latency and bandwidth needs. Alternatively, you could federate a private storage cloud with a public cloud, using public cloud storage for archive, backup, disaster recovery, workflow sharing and distribution.

This hybrid approach can allow an organisation to take advantage of the scalability and cost-effectiveness of cloud storage without exposing mission-critical data.
The challenge is to integrate and govern such a system, preferably without altering the existing on-premise infrastructure or the applications. That is

especially true when you consider services must be provisioned from different sources, yet must act and interact as a single system. This, in turn, means you need common data and software management tools.

Different suppliers try to solve this in different ways, including accessing everything via an Internet Small Computer System Interface (iSCSI), integrating primary storage with the cloud or via a cloud gateway of some sort, for example.

One of the most popular routes is a hybrid cloud storage appliance, which has intelligence, software and local storage built into it. Application servers communicate with the appliance and never directly to the cloud. By caching data locally the appliance provides more bandwidth than the wide-area network, reduces bandwidth and storage costs, and minimises the effects of link latency. The appliance can also deduplicate and encrypt data before staging it in the cloud.

Another route - or another element of the hybrid jigsaw - is to hybridise the application. Most mission-critical applications are vertical in nature, with the task moving through a stack of functionality and usually ending up in a database. While this database might be too sensitive (and large) to host in the cloud, other elements - most notably the web-based graphical user interface (GUI) - may be ideal candidates for cloud hosting.

For example, most modern applications are designed with a web front-end process that uses a browser or a series of RESTful application programming

interfaces (APIs) to present information to users and obtain updates. This model makes it easier to accommodate different mobile devices or changes to the language, and it could also be cloud-hosted.

> A vertical view of your systems opens new ways to take advantage of the cloud, and can also improve the cost and performance of those mission-critical applications

If the application does not have a web front end, you can follow the information flow through the stack and find the specific software component where formatting meets information processing. This is the logical GUI/application service boundary and it is where you could use that component's interfaces and APIs to connect a web front end.

Of course, this all assumes you can host the data yourself and provide adequately fast access to it, most probably via a query-server model. This allows cloud services to send database requests to a server and have it return only the specific data needed. That reduces traffic, delay and cost. You may also need to add load balancing at the service boundaries, deploying extra copies of each service as needed.

For transaction processing systems, however, speed and latency issues mean that running on remote storage is rarely an option.

One route is to do what NetApp has done with its NetApp Private Storage, which is to partner with colocation (colo) provider Equinix to access the high-speed, low-latency local connections that the major cloud providers make available to nearby datacentres. This enables an organisation to host its filers in a colo and connect them directly into the likes of Microsoft Azure or Amazon Web Services (AWS). Your applications can then be cloud-hosted while your data remains on private storage. Alternatively you can replicate your on-premise filers to cloud-connected filers for disaster recovery, so if the primary datacentre is lost the applications can be spun up in the cloud instead.

Again, that relies on segmenting the application into layers and figuring out which layers - and which data - can safely be cloud hosted and which you want to retain control over.

Many planners consider hybrid cloud for mission-critical applications purely for the failover or cloud-bursting opportunities it brings, but a vertical view of your systems opens new ways to take advantage of the cloud, and can also improve the cost and performance of those mission-critical applications.

Hybrid cloud storage suppliers

Several of the major storage suppliers and cloud providers have specific products targeted at building and operating hybrid clouds.

EMC

EMC has a number of offerings related to hybrid cloud storage, such as CloudArray. Derived from its TwinStrata acquisition, this works as a cloud storage gateway running both on-premise and in the cloud to provide capacity expansion, data protection and so on, with 256-bit Advanced Encryption Standard encryption. It presents cloud storage as an iSCSI or networked-attached storage device, and can replicate to the cloud and archive cold data there while maintaining on-site access to hot data.

Hybrid storage is also an element of EMC's recently announced Enterprise Hybrid Cloud, which can provide hybrid access between its on-premise storage and clouds belonging to VMware's vCloud Air, Microsoft Azure and AWS. Also, EMC has acquired Maginatics, the developer of a global namespace that can overlay multiple public and private clouds for unified data management.

Although not normally thought of as a storage provider, Microsoft's acquisition of StorSimple gave it a local storage appliance that also works as a cloud storage gateway. Now branded as Azure StorSimple, the device provides local storage for primary data, while moving infrequently-accessed data and snapshots to Azure cloud storage. A cloud-based snapshot can be mounted as if it were a local file system and accessed remotely.

Hitachi

Hitachi's object storage software, the Hitachi Content Platform (HCP), allows enterprises to build multi-tenanted private clouds hosting up to 80PB and to

automatically tier data to public clouds. Supported targets include Microsoft Azure, Amazon Simple Storage Service (S3), Google Cloud, Hitachi Cloud Services, and any other S3-enabled store. HCPs can be globally distributed and synchronised for better performance and availability. The company also offers HCP-based file sharing and data ingestion products.

IBM

IBM has the goal of making public and private clouds seamless, for example via its Elastic Storage on Cloud (Esoc) service which offers hybrid options. Hosted on IBM's SoftLayer bare-metal cloud and designed to scale beyond 1PB, Esoc - which also supports OpenStack Swift - works as a control plane able to automate snapshots, backups and movement of older data off to cheaper storage. It forms part of the SoftLayer-hosted IBM Platform Computing Cloud Service. IBM also has some hybrid cloud storage capabilities elsewhere in its range, notably in its StorWize and XIV scale-out NAS families.

NetApp

NetApp promotes its hybrid concept NPS, which allows customer-owned filers to be hosted in colo facilities that have direct low-latency connections to the nearby datacentres of major cloud providers. It also now has a version of its OnTap storage management software that works in the cloud - on an AWS virtual machine, for example - and interoperates with OnTap on-premise to provide dynamic data portability.

Dell

Other companies, such as Dell, focus on working with the likes of VMware and OpenStack, and on providing the underlying cloud hardware and software, whether for mid-range private clouds or for enterprise and public clouds.

Red Hat

There are also startups and software developers that tackle hybrid cloud storage. For example, Red Hat says its software-based Red Hat Storage Server can bring together private cloud storage and the Amazon public cloud, unifying data access and creating a hybrid storage cloud. The company also owns Inktank, developer of Ceph Enterprise, an enhanced version of the open-source Ceph massively-scalable storage system.

Avere

Avere offers edge filer technology, either as hardware or a virtual server, which makes cloud resources addressable as NAS. The local filer minimises latency to the cloud, and can include Flash, NVRAM and DRAM to further accelerate performance. Existing NAS can also be integrated with cloud into a seamless single storage resource.

Ctera

On the hardware side, Ctera offers a cloud storage gateway that provides local NAS, plus backup and replication to the cloud and a virtual cloud drive. Ctera Portal adds the ability to manage, synchronise and backup local and public cloud storage, allowing the use of cloud and on-premise storage depending on requirements. Desktop and mobile apps also allow local folders and files to be shared and synchronised to the cloud.

Nasuni

Nasuni provides local filers that act as cache and gateway for a unified hybrid cloud storage service. Files are moved to the cloud for long-term storage, as are regular snapshots. Both file (NAS) and block (SAN) access are supported, along with web-based access and mobile sync.

Panzura

Taking a different tack, Panzura's Quicksilver cloud storage controllers use a global file system to cover remote cloud storage and local disk or Flash. Quicksilver devices can be federated with others at different locations, while the controllers can work with a wide range of public cloud storage platforms.

Amazon

Amazon's AWS Storage Gateway can provide hybrid functionality, either caching hot data locally or storing specific primary volumes locally for low-latency access, with asynchronous snapshots backed up to the cloud.

Alternatively, the gateway can be configured as a virtual tape library, storing backups in the cloud.

HP

Several other companies also use public cloud storage as a backup or replication tier, including HP Autonomy, which uses the public cloud to back up a private cloud. HP's wider private and hybrid cloud strategy has been firmly based on OpenStack, but in September 2014 HP acquired Eucalyptus, an open-source tool for building AWS-compatible private clouds that can seamlessly burst to Amazon.

Next article

Software-defined storage: The pros and cons, and what is available

Chris Evans, guest contributor

Software-defined storage, in which storage array features are delivered by software products, is gaining popularity.

It promises cost savings, because software-defined storage is run on commodity server hardware and can use spinning disk and flash drives to provide high-performance, fully featured storage for organisations ranging from small companies to enterprises.

But is software-defined storage suitable for all? We weigh up the pros and cons, but first let's look at its key characteristics.

Typically, one or more of the following features are part of a software-defined storage deployment:

- **Commodity hardware** - The use of non-proprietary components that allow systems to be built by the user. Software-defined storage should be able to consume standard hard drives, SSDs and work within a typical server chassis.

- **Hardware abstraction** - Separation of the logical aspects of data storage from the physical components, such as HDD/SSD performance and RAID. Software-defined storage should use more general terms that define latency, IOPS and throughput, independent of the hardware used.
- **Automation** - The ability to drive the configuration (both for provisioning and policy) at an API or CLI level. The key benefit here is the use of abstracted policies that deliver customer business-focused requirements.

The rise of software-defined storage has been made possible by two major factors - the standardisation and commoditisation of hardware components. Standardisation has seen the industry settle on x86 as the platform of choice for applications and storage. Almost all suppliers have migrated their hardware platforms to exploit the x86 architecture and its associated ecosystem, such as PCIe and NVMe.

Meanwhile, the commoditisation of components means spinning disk and flash drives (as well as other hardware components) are reliable, predictable and easily available to the user to build storage platforms.

Build or buy?

With the rise of software-defined storage, is there any advantage to buying storage from an array supplier? Can't the user simply build their own storage hardware more cheaply?

The cost argument is definitely worth considering, but let's look first at the technical pros and cons.

From a hardware perspective, the components used to build all but the highest-performing systems are readily available. Array suppliers put a premium on the price of their hardware, despite having the buying power to get lower wholesale prices than their customers.

However, the storage suppier may add hidden value that is not always obvious. For example, the components chosen will have gone through significant testing to identify edge cases and scenarios that stress component hardware. Suppliers work closely with suppliers and can influence firmware upgrades that optimise disks, SSDs and adaptors for their storage platform.

Suppliers also take feedback from the field that collects data on thousands of hardware deployments. This ensures issues are addressed in future code releases in a virtuous feedback loop. This same process does not exist for software-defined storage suppliers, who may get critical feedback from customers only when products fail to work or lose data.

Regression testing issues

But the ability to use any hardware for software-defined storage can actually be problematic. Although hardware has standardised, it is possible to build from a huge range of configurations based on multiple server supplier products, with

various generations' and manufacturers' components, each of which could be running one of many versions of code. This could bring significant regression testing issues.

Finally, we should consider the problems involved in maintaining software-defined storage.

With software-defined storage, the user becomes responsible for sourcing hardware components, testing new configurations and firmware, and for liaising with the software supplier for patches, updates and fixes. Much of this work, including actual upgrades, would normally be done by the storage array supplier.

This brings us back to the cost discussion.

We can see that hardware suppliers do add value and can justify the higher cost of their products.

That said, smaller customers may feel it more cost-effective to acquire hardware themselves and simply buy software to run on top. Large customers may feel the economies of scale are such that they can afford to be both builder and consumer.

One thing is sure - users that directly control the hardware cannot be held to excessive maintenance charges in the three to four years after initial purchase.

Multiple tracks

As the software-defined storage market evolves, two contrasting approaches are being taken by storage suppliers. Specialist software-defined storage suppliers have moved to offer standardised hardware platforms for their products, whereas array suppliers have started to produce software versions of theirs.

Software-defined storage products that are available with standardised hardware configurations include Maxta with its MaxDeploy configurations, Atlantis Computing with HyperScale, and Dell-EMC with its ScaleIO Ready Nodes.

Some customers simply do not want to design their own hardware, so by providing a software and hardware solution with less mark-up, these suppliers have found a middle ground and used the opportunity to pivot more towards hyper-convergence.

Meanwhile, Dell-EMC and HPE provide software versions of their hardware offerings, including HPE StoreVirtual (Equallogic), HPE StoreOnce, Dell-EMC Unity and Data Domain Virtual Edition.

These are fully supported platforms that have minimal capacity and additional paid-for licences.

NetApp used to offer a software-only implementation of Data ONTAP, but that appears to have been discontinued. The company does offer a software-only version of the SolidFire operating system that can be deployed on specific hardware configurations.
Another area of software-defined storage adoption is object storage. Object stores are well suited to being deployed on commodity storage, where throughput, rather than latency, is an important metric.

Almost all object storage suppliers - including Scality, Cloudian, Caringo, Cleversafe/IBM, OpenIO and NooBaa - can be deployed as software, either onto bare metal or as a virtual machine.

In the cloud, we see offerings from existing storage suppliers, both array makers and software-defined.

NetApp offers Data ONTAP as ONTAP Cloud for AWS, SoftNAS has CloudNAS and Zadara Storage has VPSA, a hardware-based SDS offering. Meanwhile, Cloudian HyperStore is available as an AWS AMI (Amazon Machine Image) and Panzura offers its Global Cloud Storage System.

Container-based storage

The world of containers is seeing an increase in storage offerings, such as Hedvig's Universal Data Plane, which can be used for container-based storage.

Portworx and StorageOS both offer storage systems for containers that are also built on containers. This is a novel approach considering the fact that storage was typically the persistent layer for transient containers sitting above it.

Finally, we should not forget a range of other commercial software-defined storage systems from DataCore (SANsymphony), Starwind (Windows-based), Datera (distributed storage) and Storpool (distributed storage).

There are also open source platforms from Ceph (scale-out storage), CoreOS (Torus) and Gluster. Both Ceph and Gluster are supported commercially by Red Hat.

Looking forward

This is not intended to be a comprehensive round-up of all the products on the market, but it is clear there are many options available for users and IT departments. Lines of deployment are being blurred between buying hardware and software combined, or buying software and using commodity hardware.

Possibly the biggest benefit of moving to a software-defined model is the future transition to hybrid and multi-cloud operation.

Many storage systems can already be deployed in public cloud environments and provide equivalent functionality to that already available on-premises. This means IT departments can begin their transition to hybrid cloud and manage the big issue of data mobility.

Data can be migrated between on-site and cloud-based platforms using whatever replication techniques the physical/virtual appliance offers.

This means we are likely to see hardware-only storage being reserved for high-end performance requirements or niche applications such as the mainframe.

Software-defined storage deployments will continue to increase as suppliers improve commodity hardware support. Ultimately, this can only benefit the customer, with an embarrassment of choice in available systems.

The most difficult transition will be for the traditional hardware suppliers because they have to adapt to software-based licensing models and a different way of selling.

Next article

How to tune storage performance for different workloads

Chris Evans, guest contributor

For many years, tuning storage performance (particularly shared storage arrays) has been seen as a dark art that requires a degree of skill and experience to achieve.

But as the market evolves with new technology, it might be assumed that less effort is needed to match the requirements of specific application workloads.

However, storage still needs to be ‘tuned.

In this article, we discuss some of the issues involved and what can be done to optimise storage hardware to the demands of modern applications.

External storage has typically been used as a permanent store for application data, with the downside that the speed of persistent media is vastly slower than main memory.

Hard drives have latency (response) times measured in milliseconds and are good for sequential rather than random input/output (I/O) performance.

NAND flash offers a performance boost with good random I/O handling, albeit at the expense of the lifetime of the media. Flash storage also has an issue with the need to perform background tasks like garbage collection that can add temporary spikes to response times.

DRAM provides high performance and, as we will discuss, can be used to improve performance.
Finally, we should remember that for shared arrays, servers and storage are connected via a network, and with scale-out node-based solutions (as in hyper-converged infrastructure), there is a network connecting the nodes that is essential to maintain data integrity.

With all that said, here are the areas where we can configure tuning options.

- **Data layout** - If you distribute data across physical media I/O performance can be improved. Individual disk and flash drives have limited I/O capability, so ‘striping across multiple devices can spread I/O across many concurrent read/write streams. With RAID data protection, stripe width can’t be extended indefinitely as there is a trade-off in resiliency of the RAID group and the rebuild time with extended RAID group sizes. RAID-6 schemes extend protection at the expense of extra space and parity calculation overhead. An alternative is to use erasure coding, but this is suited more to object-type data.

- **Caching** - Caching data on flash or in DRAM allows I/O latency to be improved by serving read requests from the cache in a shared array or in the application host. Write I/O can also be accelerated but needs to be

protected against hardware failure by replication and/or writing to a persistent cache device. Modern caching solutions like Nimble's Adaptive Flash or HPE 3PAR's Adaptive Flash Cache look to optimise the use of more expensive resources while improving performance.

- **Network tuning** - In shared storage environments, Fibre Channel and Ethernet networks can be tuned to improve performance. For Fibre Channel, this can mean looking at settings like buffer credits and for Ethernet looking at packet size. Obviously, having non-blocking switches ensures that point-to-point throughput is guaranteed for each port on the switch. Network design also has an impact. Historically, Fibre Channel networks were designed using a range of topologies, based on saving cost. Today, Fibre Channel and Ethernet are approaching individual port speeds that are hard to saturate (32Gbps for FC, 40Gbps for Ethernet), so port sharing isn't a big issue. However, if ports can be dedicated as much as possible then this helps eliminate bottlenecks.

- **Tiering** - The use of tiering typically offers cost savings but is a performance option too. Data can be placed on the most appropriate tier of storage based on I/O performance needs and cost effectiveness of media. Tiering algorithms have developed rapidly over the years, moving from LUN to block-based tiering. Getting a tiering algorithm and data layout right can improve performance without resorting to purchasing additional hardware.

To get the best out of tuning, the starting point is to know the I/O profile of the application.

This can be quite variable, but we can break it down into a number of categories.

Structured data - This is typically represented by SQL (Oracle, SQL Server) and NoSQL (MongoDB, CouchDB) databases and has a mixed I/O profile. Data stored typically has an I/O profile that demands random I/O (excluding full table scans) whereas data writes are logged as small append-style writes. Traditional wisdom has been to place data on RAID-5 storage and logs on RAID-10, but this was really only relevant where there was little I/O caching. With modern storage arrays, most database loads (except the most intensive) are easily managed without manual placement of data. For more intensive workloads, placing logs on high performance storage is a good strategy.

Virtual servers - Server virtualisation introduces the 'I/O blender effect that randomises even sequential I/O workloads. This is because data is distributed across a LUN or volume from many virtual machines, each of which acts independently of the other, generating a random workload profile. Improving performance for virtual servers means deploying faster media (to reduce individual I/O latency) or introducing caching. Both vSphere (VMware) and Hyper-V (Microsoft) allow caching to be implemented for individual virtual machines. There are also third-party caching solutions that integrate into the hypervisor to improve I/O performance. For HCI, VMware's Virtual SAN offers an all-flash option that uses a mix of high performance and high capacity flash to optimise I/O workloads.

Virtual desktops - The challenge of delivering I/O performance for virtual desktops brings up the same issues of randomness as virtual servers, with a

couple of differences. First, most desktops are made from a single image so there is a large amount of duplicate data when starting many desktops. Second, virtual desktops are booted frequently, potentially daily, and so there are some intensive read (startup) and write (shutdown) periods. Virtual desktop performance can be vastly improved by caching and deduplicating the desktop image in a shared array or using third-party software. Non-persistent desktops can even be cached in DRAM. This solution works out much cheaper than buying an expensive all-flash system.

Web servers - Web and other read-intensive applications (like content management systems) will benefit from the use of additional read cache. This can be implemented in shared arrays, or as dedicated cache in the hypervisor. There's an obvious trade-off here in avoiding backend I/O altogether with efficient caching in the web server itself, but these cache systems still have a limit and so need to fall back on external I/O at some stage.

Email servers - Looking back 10 or 15 years, the demand for I/O per mailbox on a platform like Microsoft Exchange would have been quite high. With successive product releases, the I/O demands per user have dropped by factor of 15x to 20x. Exchange 2016 requires around 5% of the IOPS per user of Exchange 2003. As a result, Exchange can be deployed successfully on JBOD systems rather than a SAN. Having said that, email platforms like Exchange will benefit from increased use of cache and distributed data layout like wide striping.

Analytics - Many analytics tools read and re-read the same data as they build up a profile of the data. These tools are sensitive to latency and need to be able to execute queries in parallel, hence the design of Hadoop across many physical storage nodes with multiple disk spindles, for example. So, improving performance for analytics workloads is about reducing I/O read latency. This can mean using flash media, adding more cache to external storage or hosts running the analytics software. Data placement isn't that useful as it's hard to predict exactly what data will be used when running analytics software. So, the focus for analytics is getting the right balance of storage and cache and designing for the ability to increase cache as desired.

With all of the above scenarios, having access to detailed metrics that show performance and resource usage is critical. With the right data, the impact of any change can be assessed and measured against the cost in additional resources.

Next article

In this e-guide

Private cloud: Orchestration, storage and containers

Chris Evans, guest contributor

Cloud, to most people, means public cloud infrastructure, such as that offered by Amazon Web Services (AWS) and Microsoft Azure.

In contrast, private cloud tends to refer to on-premise infrastructure that functions in the same way as public cloud, by offering access to resources on a service-based delivery model.

So, what is involved in building a private cloud and how does it impact on storage in the datacentre?

Before we dive into how a private cloud can be implemented, we should take a moment to review the features of cloud in general.

Cloud, whether private or public, is expected to meet the following criteria:

- **Elasticity** - The ability to scale up or down on demand from the user. From the perspective of the user, cloud resources are typically seen as ‘unlimited, although in private clouds where budgeting occurs, the idea of limitless is likely to be constrained.

- **Service-based** - Resources are delivered as services, with service definitions and abstraction from the hardware used to deliver them. Requests are made for virtual instances - rather than servers - or storage capacity based on performance and latency needs.
- **Multi-tenancy** - Infrastructure needs to support multiple users, without those users being aware of or impacted by others on the same platform.
- **On-demand access** - The ability to request and have resources configured automatically, typically through a portal or application programming interface (API).
- **Billing & Reporting** - Features that enable the reporting of usage statistics, billing done at a high level of granularity, typically per hour or day (note that public cloud providers charge per hour, which may not be practical or appropriate for private infrastructure).

In a private cloud, some features seen in the public arena are not necessarily required.
Some organisations may choose not to implement billing (or even chargeback), for example, as the financial mechanism for directly charging business units for their usage may not exist.

The idea of 'unlimited is typically not a private cloud requirement, as IT budgets are generally restricted, whether that refers to the spending of the business or the IT department.

Private cloud orchestration

A key part of private cloud is the orchestration framework that provisions resources based on customer requests. These include open-source platforms such as OpenStack and CloudStack, proprietary tools such as VMware's vRealize Suite Cloud Management Platform and the upcoming Azure Stack, Microsoft's on-premise automation suite implementation of public Azure.

There are also platforms, such as Mesosphere, Kubernetes and Docker Swarm, that provide open source solutions that go part way to implementing private cloud.

Finally, we should mention that suppliers, such as Zerostack and Platform 9, that are looking to optimise some of the tools we've already mentioned.

Storage for private cloud

When we look at the way storage is consumed by these platforms, we see a wide variety of implementations.

Possibly the most mature and easily understood storage options are for VMware.

VMware vSphere, which is a core component of the vRealize Suite, supports a range of existing block (Fibre Channel, iSCSI) and file (NFS) based systems.

Supplier support is widespread, and once storage is configured into the infrastructure, it can be provisioned out automatically as part of virtual machine creation.

OpenStack has a range of storage support options, described as projects, that cover block storage (Cinder), file (Manila) and object (Swift). Each project defines a set of APIs that suppliers can code to, automating the provisioning and mapping of their storage platforms to instances (virtual machines) by the OpenStack platform.
The same level of application programming interface support is also available within Kubernetes, enabling the provisioning of persistent storage to container instances.

This can include traditional protocols (such as iSCSI and NFS), but also scale-out storage such as GlusterFS and Ceph. There's more of an assumption here that users need to manage their storage manually (perhaps simply as local JBODs), but storage hardware suppliers are starting to add support for Kubernetes volumes. One example is NetApp with their Trident project to automate storage provisioning.

Docker deployments tend to use storage provisioned to the host running containers, which could be a local file system or block devices mounted to the server. Flocker is an open source tool that enables the automation of provisioning block devices to containers through a set of APIs for which suppliers can provide support. Through their EMCCode efforts, Dell EMC

supports connectivity for many of their existing storage platforms to a Docker ecosystem. Much of this work is best efforts and only community supported, so perhaps not suited to critical production environments.

Adding storage to private cloud

We can see from the way storage is consumed by private clouds that there are certain differences that must be considered. The most obvious is that of automation. Storage is expected to plug into orchestration frameworks in a way that makes it easy to auto-provision logical unit numbers (LUNs) and file shares through the orchestration tools.

This automation represents two issues.

First, there is security. Can storage be presented from a pool of resources, or does the orchestration platform get access to the whole storage array or environment?

Second, what controls are in place to limit the consumption of storage resources? We can envisage a scenario where the on-demand capability of private cloud could exhaust the storage capacity available.

This issue takes us on to the issue of maintenance. How easily can a storage solution be expanded or replaced?

Typically, storage migrations are implemented as big projects that consume significant resources. This doesn't arise as an issue in the public cloud environment, where storage issues are obfuscated from the customer.

How storage migrations be managed in private cloud?

As instances are created or destroyed, new applications can be provisioned from new storage capacity. That leaves the migration of existing instances to take care of. These may have to be migrated offline or through some sort of snapshot process. Either way, the introduction of a new storage platform needs to be implemented with a minimal impact to service.

It's worth noting that most of the above issues haven't been fully resolved. As a result, an entire storage platform may be presented to a cloud solution, for example, rather than shared with other users because the required security controls don't exist. This may be practical in large environments, but could cause issues for smaller IT organisations.

Choosing the right supplier

Picking a supplier or storage product requires thought about the orchestration platform and supportability from the storage supplier.

Older (and more likely to be classed as legacy) platforms will be harder to integrate, as they usually have no native automation capabilities. These tend to have been added on later. NetApp SolidFire is an example of a storage platform

that provides a native API, with supported integration into all the common provisioning platforms.

An alternative is to go for a scale-out storage solution, either an open source platform or one from an SDS provider. Ceph is one solution directly integrated into OpenStack that can provide object, block and file support.

There are also scale-out solutions like StorPool (block storage), StorageOS (specifically for containers), ScaleIO from EMC, StoreVirtual from HPE and of course native Virtual SAN from VMware.

The advantage of using these platforms is that storage can be deployed as part of the server, rather than having dedicated storage software. Integration of solutions like Virtual SAN, for example, are directly supported within vSphere, enabling policy-based provisioning of storage for virtual instances.

The options for building out private clouds and persistent storage are wide ranging and potentially confusing at first glance. But remember the main tenets of public cloud (elasticity, on-demand, multi-tenancy and service-based) for a good basis to help choose the most appropriate solution for your private cloud infrastructure.

Next article

In this e-guide

DevOps

Optimising with DevOps

DevOps involves combining IT operations and software development to improve productivity and deliver more flexible IT – as such, it's an important tool for optimising hybrid IT architecture.

How going cloud native can increase app agility

Clive Longbottom, guest contributor

Customer relationship management (CRM) and enterprise resource planning (ERP) are enterprise applications that were created to do a specific set of tasks that an organisation needed.

Trouble is, these nominally single-purpose applications have grown until they bear no relationship to what they once were, and the way in which they attempt

to create an ‘own and control environment no longer fits with businesses that are trying to compete in increasingly dynamic and complex markets.

Now, the world has started to look at dynamic composite applications. Rather than try to chase a single application that does everything, the search is now on for a way to create functional stubs that can be brought together in real time to support a business process at a specific time.

Intelligent FAQ system

Let’s take an example: a simple means of a prospect becoming a customer, ordering an item and paying for it. Even now, a lot of the technology that underpins this could well be being served via third-party services.

Questions that the prospect has in mind may be responded to by Transversal’s intelligent FAQ (frequently asked questions) system; the payment is unlikely to be gathered and dealt with by the application itself, but via a third-party payment clearing system, such as WorldPay or PayPal.

It is unlikely that the next multibillion-dollar company is going to be built off the back of a major new application. The clever developers are following the money - and are creating functional services that can be hosted in the cloud and used by a large number of customers paying small amounts per usage, or a constant monthly subscription. The model is no longer low-volume, high-margin, but high-volume, low-margin.

Even within an organisation, this is still the right approach: create a function that can be called by different surrounding systems as required, that can be plugged in and out as something better comes along; that can be easily optimised, updated and upgraded as required without the need for dependent, hard-coded upstream and downstream systems having to be changed.

Multi-faceted problem

It is strange that the main area for such innovation has been through tablet and smartphone devices. The easy access to small functional apps at low cost has whetted users' appetite for the same approach within the workplace. But most of these small apps are monofunctional: it is still very difficult to pull them together to create a solution to a multi-faceted problem.

Home automation via the internet of things (IoT) is beginning to drive a need here: there is a much greater use of capabilities such as IFTTT (if this, then that) programming and the use of home hubs that enable multiple different devices to work together behind a single front end, such as Amazon Alexa or Google Assistant. There are a number of techniques that developers can use to become cloud-native. We are now in the time of the application programming interface (API) economy. Any service must have an open API, so that others have a common and standardised means of accessing the capabilities of the service being provided.

A prime example of this for web-based services is to adopt a ReSTful approach. ReST is the Representational State Transfer capability - a set of common calls and responses that can operate in a stateless manner. ReST is optimised for performance while providing reliability and future-proofing through being an extendable interface.

Easily updatable

Any function or service must also be written to be easily updatable: approaching the writing of functions in a manner of cascade or waterfall projects will not be sufficient to support a business' needs. Instead, it is far better to take a well-ordered and controlled DevOps approach, using open-source tools such as Jenkins, Puppet and Chef, possibly combined with application release automation systems such as CA Automic.

Aim for the three 'CDs of continuous development, delivery and deployment to be able to get new functionality out in response to requests from the business and users within days or weeks, rather than months or years.

Ensure that the function is abstracted from underpinning resources. It is a waste of time writing a new function that has deep dependencies on, for example, the amount of CPU it requires, or a certain type of storage subsystem. Cloud allows for all of this to be abstracted, and the overall scalability of a function should then no longer be restricted by the code - it should only be constrained by the

amount of resources that the administrator wants it to have at any moment within the cloud.

Ease of provisioning

The use of code being compiled and then directly provisioned to a platform is still strongly with us, but the future is pointing towards some form of containerisation, such as Docker, rkt or LXD. By placing code within a container, it makes it easier for the function to be delivered to a range of different platforms, as there is again less dependency between the software and the underlying physical platform than has previously been the case. Systems such as Kubernetes or JuJu can help in ensuring that containers are suitably managed.

Fully define each service with metadata and context. It is likely that the future will see the payload of such containers reduce to being single-function services (microservices). Ensure that any such microservice has enough metadata provided with it for others to make sense of what it does.

Development portals will need to see the microservice and to be able to offer it for use to any individual developer, so there needs to be enough information for this to happen automatically. However, this must not be the only focus for the provision of metadata - the future of loose-coupled, high-performance, real-time composite applications is predicated on metadata and context that can be used

to bring together microservices on the fly in a fully automated and orchestrated manner.

For a cloud-native composite application to be able to be built dynamically in this way, the orchestration systems responsible for pulling everything together will have a need to be able to understand what your function does and how to use it. Advanced systems, such as EnterpriseWeb, can figure this out and create the metadata that enables it to define and create the required integrations and orchestrations for deep and end-to-end automations. Other systems may need a standardised set of data available that provides them with the information they need to be able to apply the integrations.

Don’t reinvent the wheel

There are already lots of publicly available services covering things like calendaring: use these rather than writing one that fails as soon as leap year comes along. Look to both commercially supported services available as public cloud offerings, as well as discrete code chunks from supported sites or community groups. Such sites will generally provide feedback from others saying how good the coding is and whether it does the required job or not. Use of pre-prepared code in this way can drastically cut down on coding time as well as error identification and management down the line.

If you are operating within an existing business with its own datacentre (or using a colocation facility), don’t forget to look at what you already have available to

you. Although the creation of cloud-native applications from the ground up requires a different approach in coding and the composition of the resulting overall environment, many services may already be there, hidden with existing enterprise apps.

The secret is to be able to abstract specific functions from within the overall monolith of the main application, making them available as callable functions that can be used by other systems. Here again, companies like EnterpriseWeb and CA Automic can provide the capabilities to identify where islands of data and functionality can be pulled into a far more open and responsive overall cloud-based platform.

Overall, the move to writing cloud-native apps requires a fairly fundamental change in how an organisation's IT department views its development, test and operations systems. The key is to think small, using containerised microservices as a starting point, mixed with intelligent orchestration and automation systems to provide a portfolio of services and applications that deliver far more dynamic support to the business' needs.

Clive Longbttom is founder of Quocirca.

Next article

In this e-guide

How composable applications can improve software development

Adrian Bridgwater, guest contributor

If every protagonist in the service-based, cloud-centric, software-integrated connected world is to be believed, the age of monolithic software is dead.

They now talk about the merits of ‘composable’ applications.

Composable applications is the idea that the functional blocks of an application can be decoupled from the complete applications. These individual component parts can then be more finely tuned to create a new application that is ideologically, if not also functionally, greater than the sum of its parts.

To make applications composable, it is necessary to define new parameters and perimeters for each individual constituent part. Essentially this is the world of packaged software.

Unlike commercial off the shelf (or COTS) products, that comprise of packaged software which offer a set of built-in features, composable applications are not predefined. Instead, packaged software represents granular and modular component blocks that can be brought together in the always-on world of web

and cloud to form what effectively represents an entirely new technology proposition in every use case.

Packaged software is the process of intelligently and logically componentising and containerising related and co-related software functions, tools or even complete platform elements into defined packages.

These packages can be consumed, used, implemented and deployed by developers in entirely new compute environments and in totally different ways to anything the package's original developers may have considered.

So why are packages so critical to successful software application development today and how does composable modularity in software enable developers to be more Agile?

Singularity matters

In a post on the Computer Weekly Developer Network blog, Daniel Jones, CTO of EngineerBetter, notes: 'Software that is well-packaged requires a lower cognitive load of the engineers using it, allowing them to focus on solving business problems.

Jones believes if a software-powered service only has a single thing to do then the developers who need to use that service in their own code only need to understand a few things about it.

Decoupling application code

Cloud Foundry's buildpacks. are a good example of where there is a clear separation of responsibilities: the developer provides the app code and the platform provides the operating system, base filesystem and language runtime.

This separation allows the application code to have a 'change cadence' independent of the underlying software layers, meaning kernel patches can be applied below deployed applications, without interference.

However, technologies such as Docker have blended these responsibilities. This means developers now need to care about what operating system and native libraries are available to their applications. After years of striving for more abstraction, the industry appears to shifting towards greater coupling.

Source: Daniel Jones, CTO, EngineerBetter

'Single-purpose tools are more efficient, Jones explains. As an analogy he says:' Although we can make dual-necked musical instruments that combine a guitar and a bass, you can't play both necks at the same time, but we do still have to carry the weight of both even if you only need to use one.

Simplicity is key. Software that is capable of performing multiple functions means that people that use it need to remember more about its capabilities and its state. He argues that this is the reason why interfaces with multiple modes are dangerous in critical applications.

‘Humans prefer things that always work singularly, says Jones. ‘How many times have you failed to log into a system because you accidentally had CAPS lock on? He says that software with a single responsibility can be reused more effectively and by defining well-thought-out boundaries with clear interfaces it can be changed and iterated upon with minimal impact on other elements of the system.

Package abstraction

In the past, every language and platform had its own packaging system. Ruby used gems, .NET had Nuget, Java’s was Maven. With different packaging systems it becomes impossible for developers to truly share all of their code, according to Burke Holland, director of developer relations at Progress Software.

Writing on CWDN, Burke acknowledges the benefits of JavaScript to aid code portability and the ability to share with other developers.

Holland recommends developers take advantage of the rise of server-side applications and combine them with front-end JavaScript software, giving developers the scope to package packages as finely as they require. According to Holland, NPM has done a remarkable job of package abstraction, making applications more modular than was ever possible previously.
‘Such ease in package management doesn’t come without trade-offs though. When developers use NPM packages they then have a dependency on them,

meaning that if the creator decides to remove their NPM package, this could cause problems, Holland warns.

Avoiding dependency hell

Dependencies, the link between different software packages, is among the challenges facing compossible applications.

While the theory of packaging software looks good on paper, according to Arianna Aondio field engineer at content delivery specialist Varnish Software it can lead to so-called 'dependency hell', where internal software elements depend upon each other.

Aondio has worked in second-line technical support, where the software being supported needs to run on more than one operating system and there is the added complexity of having to support several versions of the product.

She explained the problem in a CWDN guest blog post: 'When a customer calls with a problem, I often need to install the version he or she is working on. Invariably, the customer's installation has multiple dependencies. This means I need to download the same libraries to replicate the problem... and so on and on.

In software, components evolve all the time, along with dependencies. Cutting through these dependencies is unnecessarily time consuming. Containers like Docker are the best remedy according to Aondio says, for managing the

dependencies efficiently. ‘Once you’ve invested the time defining the provision file for each of your containers you can then install your favourite software, avoiding the 'hell', she explains.

‘Each container runs a different piece of software/application and it's up to you to decide if you want them to interact or not, she adds.

Microservices as a standard

A microservices architecture seems to be the approach of choice for many companies.

Modernising IT applications using an elastic microservices architecture, which offers advantages like better performance, scalability and fault isolation along with minimising application downtime.

Chetan Manjarekar, senior VP head of Syntel’s Digital One practice says the key characteristics that make microservices an attractive proposition for developing new software, is the complexity associated with the old ways of doing things.

Writing on the CWDN blog Manjarekar notes: ‘In recent years, there has been a profound shift from legacy programs and applications to flexible, cloud-native applications that can deliver the speed, scale, and security required for the digital world. Traditional monolithic application designs tend to be extremely

complex, with enormous code bases built to encapsulate all functionality into a single block.

Takeaway package

There is, unfortunately, something of a dependency and connectivity tradeoff when it comes to software packages and packaging. When developers break their software apart into smaller component parts, these components still need to be connected together in order to perform something useful.

Given the views of the experts on CWDN, it seems a considered and methodical approach with a full appreciate for the total architectural well-being of the entire software stack appears to be the most prudent way forward when considering composable applications.

Next article

In this e-guide

Container technologies for cloud-native apps

Charlie Dai and Dave Bartoletti, guest contributors

Container technologies allow enterprises to develop highly differentiated applications and services more quickly, with better quality and geographic reach, to create compelling customer experiences.

They have quickly become an important element of digital business transformation because they promise faster software delivery, tremendous scale, higher resilience, greater flexibility, and broader implementation options. Everything about enterprise app infrastructures, development styles and architectures is changing, and containers play a key role in each area.

The basis of infrastructure platforms, both on-premise and in the cloud, is moving from virtual machines (VMs) to infrastructure and container services as developers want to consume smaller fit-for-purpose app components via application programming interfaces (APIs). Development styles are speeding toward true continuous integration and continuous delivery - DevOps - and app architectures are moving from multi-tiered monoliths to more agile collections of microservices.

In this e-guide

Holistic view

Architecting a container-based platform for enterprise applications requires a holistic view of various capabilities and how they fit together. Regardless of whether enterprise app professionals use the container-native or integrated-platform pattern, they must still choose components and assemble them seamlessly. The reference architecture of container-based application platforms can be defined as consisting of eight layers, each with specific componentised technical capabilities.

Many open source options for the container-native pattern normally focus on one capability component (or a few) for assembly, while options for integrated patterns usually embed the container engine, container orchestration, external integration and operations management layers together.

The first layer is the container engine, which provides the foundational execution environment. This environment should support predefined configuration file formats of mainstream container images, such as those from Docker, App Container and the Open Container Initiative. The environment should not only support container runtimes, such as runC, Docker containerd, rkt, and cri-o, but also runtimes of VMs, such as runV, Clear Containers and LXD.17. It should then be able to execute a set of standard build-ship-run operations for container images and instances. These can range from building an image from configuration files to tagging an image in the repository.

Key capabilities

The second layer, container orchestration, enables key capabilities for enterprise adoption. Task scheduling with high-availability (HA) support components such as Mesos and Marathon ensures efficient, robust cluster resource use for running applications. App configuration management components such as Docker Compose and the Topology and Orchestration Specification for Cloud Applications (TOSCA) simplify service composition and app configuration. Service discovery and configuration management components such as etcd can register and discover services across container instances to allow storage and lookup for service configurations.

Container cluster management components like Docker Swarm and Fleet handle cluster membership, detect and recover failed container instances, and propagate custom events. Container networking components such as Docker Networking, Weave Net and Canal provide abstract, unified networking features. Container storage management components such as Fuxi manage data volumes for containers. Only a few, such as Kubernetes, cover all capabilities in this layer.

Diversified use scenarios

Next is the external integration layer, which allows extensive support for diversified use scenarios. Big data integration components help mainstream big data frameworks, such as Apache Hadoop YARN or Apache Spark, leverage

the power of containers to maximise performance in the cloud while using their own schedulers. Message queue integration components deploy third-party message queue and streaming platforms, such as Apache Kafka, into the container environment. And infrastructure-as-a-service (IaaS) integration components help IaaS platforms such as OpenStack embrace container orchestration.

The fourth layer is for operations management, which streamlines operations or maintenance processes. Container monitoring and data collection components such as Logstash, OpenTracing and Fluentd monitor containers, trace API invocation in distributed environments, and collect data in a unified manner. Metrics storage, analysis and alerting components such as Prometheus and Elasticsearch analyse aggregated data using persisted data model metrics and send alert notifications. Management portal components such as Kibana and Grafana provide dimensional dashboards of analytics results for operational management.

Next up is the container infrastructure layer, which allows adaptability of operating environments. Container operating system (OS) components such as Red Hat Enterprise Linux Atomic and VMware Photon OS provide minimalist OS features that support containerised application execution. And generic bare-metal hypervisors such as ESXi, KVM and Hyper-V, which are generic components for virtualisation and cloud deployment, allow virtualisation on physical servers for hardware resource pooling.

Unified control and value co-creation

The container image management layer ensures unified control and value co-creation. Private image repository components provide centralised and scalable building, discovery, distribution and change management of container images; firms can deploy components such as Docker Registry and CoreOS Quay Enterprise on-premise or use cloud-based services such as Docker Cloud, Google Container Registry and Amazon EC2 Container Registry. Public image marketplace components, such as Docker Store and CoreOS Quay, provide a trusted platform with reputation indicators for sharing container images.

The penultimate layer concerns container security, which safeguards end-to-end security. This layer provides protection for the complete technology stack, from container infrastructure to container operations management and image management, and for the entire container lifecycle from image creation to runtime execution, including image signing and verification, role-based access control, security policy management, and enterprise LDAP integration. For example, Docker Content Trust enforces client-side signing and verification of image tags and CoreOS Clair provides static analysis of vulnerabilities in container runtimes.

Application lifecycle acceleration

The final layer is DevOps automation, which allows application lifecycle acceleration. This layer not only helps enterprises simplify the deployment of

components in other layers, but also integrates with existing cloud environments and DevOps tools to ensure seamless collaboration. For example, Docker Machine can help install Docker Engine on virtual hosts atop major cloud service providers, enabling management capabilities.

Enterprise application professionals must assess the capability coverage for each layer against business requirements. They must take the lead in mapping business needs to technology requirements and evaluate the ability of each component to address these requirements. For example, if your large-scale web applications translate to proven clustering management capabilities over 5,000 nodes, Mesos might be a better choice than Kubernetes. However, the Kubernetes project has more than 950 contributors on GitHub , offering a more active ecosystem.

This is an extract from the Forrester report Vendor landscape: container solutions for cloud-native applications by Charlie Dai and Dave Bartoletti.

Next article

Insurance giant Allianz opens up about how DevOps success is fuelling its move to cloud

Caroline Donnelly, datacentre editor

Allianz has revealed how a trip to Silicon Valley inspired its senior management team to embrace DevOps, and paved the way for the German financial services giant to ramp up its use of cloud.

Speaking at the DevOps Enterprise Summit in London, Andrea Hirzle-Yager, head of the IT department at Allianz Deutschland AG, said - on their return in early 2016 - the management team set a challenge for each of the organisation's business units to create an app within a year.

'They came back with lots of great ideas about what we had to do inside just a few months to continue to focus more on our customers, she said.

'It's not like we hadn't focused on our customers before, but we realised we needed to get closer to the customer. We understand their needs and we can quickly deliver products to them anytime or anywhere.

The challenge was initially greeted with a degree of scepticism - particularly from individuals working inside the organisation's IT security and compliance teams - but Hirzle-Yager said she was determined to support what the company's senior leaders were hoping to achieve.

'I raised my hand. I'll take this project on and we will deliver those apps by the end of the year to make this possible, she said.

One of the reasons why she was so keen to get involved was to protect her operations team from finding themselves in a situation where their views and needs were not factored into a wider discussion about how the project should proceed.

I've been too many times on the other side of the wall where somebody went to do something [a code deploy], and just before it was nearly done, it was thrown over the wall to the operations people to sort out. I didn't want to do that again, she said.
'I also want my team to be there at the very beginning, having discussions as problems come up and decide how we're going to address them.

A pivotal moment

One of the firms the Allianz team visited during its Silicon Valley jaunt was digital transformation company Pivotal, which coaches enterprises how to incorporate agile software development processes into their business.

This led to Allianz seconding teams of software engineers to Pivotal's offices in London to help them get to grips with concepts such as paired programming and working as part of a cross- functional team.

To ensure the working practices the Allianz team in London were being schooled continued to be followed once they returned, the company set about creating some agile training centres of its own in Munich and Stuttgart in June 2016, with the first teams moving in one month later.

'We needed a space they could come back to where they wouldn't be distracted by all of the everyday phone calls or meetings they were exposed to, she said.

Achievement unlocked

The first crop of apps produced as a result of these efforts were ready for testing by August 2016, before making it on to the app store several months ahead of schedule.

'That means my team, together with the help of numerous external and internal experts, had created this platform or at least the test environment for a platform at light speed, said Hirzle-Yager. 'I have to tell you, this was something that had never been done before. It would usually take us several months to get a hardware order - let alone create something that could be used.

The first application to make it onto the App Store arrived in September 2016, and two further additions joined them in November and December of that year, too.

They included one that lets users keep tabs on their contracts and dealings with the firm from their smartphone, while another lets customers upload pictures from a road traffic accident, for example, and receive a near-instant response on how much it would cost to repair any damage caused.

‘That crazy goal, set at the beginning of the year, was made possible by working together, she said.

The move to the cloud

Having surpassed the app development goal set by the senior management team, Hirzle-Yager has now been asked to draw on her experience of overseeing the DevOps push at Allianz to help the company start moving to the cloud.

‘For an insurance company like Allianz, going into the cloud, that’s another crazy project, right, but that’s what we’re doing, she said.

‘We have looked at all the lessons we’ve learned, created a cross-functional team to create, setup and define this new environment to support more app development, for some of our legacy environments to move into and my team has been asked to lead the initiative.

Top-down management support

The top-down push from management to embrace DevOps allowed the organisation to overcome many of the hurdles enterprises commonly run into when embarking on a digital transformation of this scale, she said.

For instance, with senior management support already assured, this saved its organisation from having to repeatedly make the business case for DevOps, instead of getting on with the job at hand.

‘You need your senior leadership and board to understand what you’re trying to do as you embark on this journey or as you continue to report on this journey, because DevOps is not easy. It’s well worth it and it absolutely should be done, but it’s not an easy thing and takes a lot of patience, she said.

‘So, if you have to explain your business case to your senior executives all the way, you’re not going to have enough time to focus on all the other important things. You need to get that out of the way early and get your senior leadership aligned.

Running into difficulties

That’s not to say the order from on high to adopt DevOps proceeded without a hitch, as the company encountered some difficulties in the early days as its developers and operations staff adjusted to their new working relationship.

‘One team would ask the other team to do something, but nobody really knew what exactly needed to be done to move forward. So I called the team together and made them all literally sit at the same table together, said Hirzle-Yager.

The team she pulled together for this discussion included the individuals responsible for security, infrastructure, operations and application development.

‘It was amazing to see, that once they were all sitting there together and actually talking about what needed to be done, so many of the issues disappeared, she said. ‘Some of it was just a misunderstanding of what the person actually needed.

The security and compliance teams had misgivings about the project, which were also worked through in this forum and provided them with an opportunity to raise their ‘red flags - and ensure the developers and operations teams addressed these issues during their portion of the work.

Proving the value of DevOps.

In organisations where DevOps has emerged as more of a grassroots initiative, driven by technology teams inside individual business units, for example, Hirzle-Yager said the pressure is on them to prove the value of DevOps to win the backing of senior management.

It is also important to bear in mind that DevOps success also requires organisations to get buy-in from other parts of the business too, including - crucially - the human resources department.

‘DevOps is not just an IT thing, she said. ‘In many organisations, it starts in IT, but you need to get your HR people on board. I don’t know how many of you have the challenge of trying to find the right talent with the mindset to help you along with this journey, and if you don’t have your entire organisation aligned, it makes it that much more challenge.

Next article

In this e-guide

Hyperconverged Infrastructure

Optimising with hyperconverged systems

Hyperconverged systems combine compute, network and storage capabilities into a highly optimised unit, making it an essential technology in a hybrid cloud infrastructure.

How hyper-convergence is changing the shape of IT

Cliff Saran, managing editor

According to 451 Research's Voice of the enterprise survey of IT buyers from September 2016, hyper-converged infrastructure is currently used by 40% of organisations, and the analyst company expects this to rise substantially over the next two years.

In the survey of 750 IT professionals, a quarter indicated that they had hyper-converged infrastructure (HCI) either in a pilot phase, or in plans for future adoption.

The analyst firm believes hyper-converged infrastructure is evolving from supportive, edge infrastructure into a primary component of today's IT organisations. In the survey, almost three-quarters of IT professionals said their organisations were currently using hyper-converged systems in their core or central datacentres. This signals a transition in IT buyers now thinking about server purchases, says 451 Research.

Innovative technologies

Christian Perry, research manager at 451 Research and lead author of the Voice of the enterprise: Servers and converged infrastructure - organizational dynamics report, says: 'Loyalties to traditional, standalone servers are diminishing in today's IT ecosystems as managers adopt innovative technologies that eliminate multiple pain points. Innovation inherent in converged systems and in hyper-converged infrastructure, in particular, is driving process efficiencies and agility that are increasingly tangible.

For instance, lubricant company WD-40 moved from its incumbent IT hardware provider to Nutanix to provide a more flexible IT infrastructure. Explaining the switch, Mark Breed, IT services manager at WD-40, says: 'We started out by looking at adding to our traditional three-tier infrastructure. However, the solution proposed by the existing supplier would have exceeded all the spare capacity in our racks, requiring the addition of new servers, extra storage trays and SAN [storage area network] switching hardware. In the worst-case scenario, we would have needed to install more racks to cope with what was required.•

Some hyper-converged systems are based on highly integrated appliances, while others are software-based, making them hardware agnostic. Each approach has its benefits and drawbacks. In a recent Computer Weekly article, Clive Longbottom, founder of analyst Quocirca, noted: 'Being able to apply HCI software against existing hardware provides obvious cost savings in not having to purchase new hardware, but the lack of highly engineered interconnects and internal data buses counts against gaining the highest possible performance from the resulting platform.•

The Forrester Wave: Hyper-converged infrastructure (HCI) Q3 2016 market assessment report defines hyper-converged infrastructure as an approach to technology infrastructure that packages server, storage and network functions into a modular unit and adds a software layer to discover, pool and reconfigure

assets across multiple units quickly and easily without the need for deep technology skills.

The main benefit is the ability to scale infrastructure by paying for the additional hyper-converged appliances that combine storage and compute as the needs of the business grows.

Forrester's Richard Fichera explains the core principles of hyper-converged systems in the report: 'At the core of the hyper-converged value proposition is the simplicity of scaling the cluster, which, in part, derives from the co-provisioning of compute and storage resources in pre-integrated modules. This means technology buyers must rely on the know-how of the hyper-converged suppliers that the ratio of compute and storage they sell is correct, he says.

What this means in practice is that suppliers need to make well-informed assumptions with their initial offerings. To provide greater flexibility, many offer modules with differing levels of compute and storage, says Fichera. Some suppliers also offer appliances that can access and federate external storage resources.

When assessing products, one of the key criteria is the ability to manage software-defined infrastructure. Forrester says the core value of a hyper-converged system is the simplicity it brings to the management of virtual machine (VM) storage. 'Hyper-converged systems must also provide a software abstraction of the federated storage of the physical nodes and present this

storage to the hypervisor, the analyst company noted in its 2016 market evaluation report.

The third criterion identified in Forrester's market assessment is automatic discovery and configuration. 'The hyper-converged system must be capable of discovering and adding a new resource to the environment with minimal operator interaction, says Fichera.

While this is hard to quantify in terms of administrator time, as a guide, Forrester suggests it should be possible to add a new module with an additional 20-100TB of storage to a hyper-converged system in no more than 15 minutes of admin time. Also, the reconfiguration of the hyper-converged system should happen automatically, without the need for any administrative intervention on the existing nodes on the system.

Third-party management tools

Another criterion to consider is the ability to use third-party management tools. â€• Management must be largely within the purview of the hypervisor console, enhanced further with APIs [application programming interfaces] for integrating with other automation tools such as cloud management suites, the Forrester report notes.

Fichera adds: ‘Other than initial setup, avoid manually managing hyper-converged systems from the command line, as this is costly and prone to errors. Local system management must be minimal, with extensive automation of management tasks.

Forrester places Nutanix as the biggest provider of hyper-converged infrastructure, offering what the analyst believes is among the largest managed clusters, with more than 100 nodes in its largest customer clusters. Nutanix and SimpliVity sit as the market leaders in converged systems, while the analyst rates established players such as Dell EMC, HPE and Cisco as ‘strong performers.

The way ahead for HPE

Earlier this year, HPE acquired SimpliVity for $650m to strengthen its stack in the hyper-converged market. In effect, HPE bought SimpliVity’s OmniStack product, which Forrester rates as one of the market leaders. Before the acquisition, HPE’s rival offering, combining its ProLiant DL380 server, StoreVirtual IP SAN technology and OnView systems management, was considered less mature.

In a blog post about the acquisition, Charles King, principal analyst at Pund-It, said: ‘Adding the company to its quiver should be a plus for HPE, and it also complements its other recent hybrid cloud acquisition, CloudCruiser.

‘But the way ahead for HPE may not be as simple as it assumes. Like Nutanix’s solutions and those of other software-defined, hyper-converged players, OmniStack can run on virtually any x86-based server and SimpliVity was a strategic partner with several Intel-based hardware vendors, including Lenovo and Dell EMC. HPE is obviously planning for its own Proliant systems to become the platform of choice for OmniStack, but it would be surprising if SimpliVity’s former partners don’t try to secure those customers for themselves.

Dell is gaining hyper-converged capabilities through its EMC acquisition. Forrester notes that EMC’s VMware-based VxRail appliances are likely to make up most of its volume sales through direct and channel partners. The product set comprises x86 servers, VMware VSAN software and value-added EMC installation and admin software.

Meanwhile, Forrester says that while Cisco has had a strong relationship with SimpliVity, it has also entered the HCI market with its own product, HyperFlex systems, which aims to offer tight integration with its own UCS servers.

HCI has huge potential to replace small to mid-size, general-purpose disk arrays in highly virtualised, general-purpose environments. However, many mainstream enterprise IT directors do not yet trust HCI systems to deliver multi-petabyte capacity at scale for Tier 1 mixed workloads that require low latency, according to analyst Gartner.

Its 2016 report on the state of play in the hyper-converged market, Beware the 'myth-conceptions' surrounding hyper-converged integrated systems (HCIS), discusses common misconceptions about these systems.

'One of the pointed comments arising from our live chats was that, when all capacity utilisation and costs are considered, modern hybrid or solid-state-array-based deployments are likely to be more economical in the long haul than software-defined storage, says Gartner distinguished analyst George Weiss in the report.

Weiss says such IT directors tend to look for systems that can offer backup and recovery, wide geographies, large user groups, partitioned and secure access, and require supplier expertise and in-house expertise. 'HCIS might not be a great fit for all storage workloads, he says. 'The inability to predict a storage-to-compute ratio for mixed use cases in enterprises might result in inefficient utilisation of either compute or storage components in HCIS.

What is interesting from the analysts' assessments of the market is that the major server and storage providers are starting to take HCI more seriously. HPE has made a big bet with SimpliVity, while Dell is starting to piece together something of an HCI roadmap based around VxRail and VSAN.

Among the areas CIOs and IT directors need to look at is how admin skills work out when hyper-converged systems are used. 451 Research's enterprise survey of IT buyers found that converged infrastructure is a catalyst of this ongoing IT

team transformation. More than one-third (35.5%) of enterprises say they have added more VM specialists as a result of adopting standard converged systems. This is more than double the number of organisations actively adding specialists in hardware-specific areas such as servers, storage and networking, according to 451 Research.

The analyst firm's Perry says: 'Businesses expect the same flexibility from their internal IT that a public cloud service can provide. Converged and hyper-converged infrastructure is transforming the technology that underpins business and the teams that manage it. This means the make-up of IT infrastructure teams will change, says Perry, with the generalist-driven infrastructure administrator emerging as the key cog in business operations.

Next article

In this e-guide

Datacentre decisions: Converged vs hyper-converged infrastructure

Chris Evans, guest contributor

In recent years, we've seen deployment models move on significantly from the siloed IT operations of the early 2000s, when separate server, storage and networking teams were the norm.

The deployment paradigm has evolved to converged infrastructure and hyper-converged infrastructure (HCI), both of which bring their own benefits and disadvantages.

How did we get here? Why choose one over another? And who are the market leaders?

Before the advent of widespread server virtualisation, IT infrastructure was managed by separate teams. At the time, this made complete sense, with IT organisations looking to deploy best-of-breed components for each technology.

Server virtualisation radically changed that landscape, as infrastructure started to become more integrated. This change allowed the migration towards the converged and hyper-converged systems we see today.

In this e-guide

Converged infrastructure

Converged infrastructure is effectively a packaging process that delivers the offerings of a supplier (and partners) as a stack of hardware, pre-tested and validated by the supplier.

Converged offerings consist of server, storage and networking hardware, usually delivered as a single rack and sold as a single product.

From the customer's perspective, there are multiple benefits. On the technical side, the supplier takes on the job of checking components all work together and will validate updates before releasing them to the customer. Converged systems usually include management software that delivers orchestration for hardware components and virtual machines (VMs).

From a financial perspective, the customer gets a single system, fully supported by the supplier as the single point of contact, making it easier to cost, amortise and deliver systems for specific business needs, such as email services and virtual desktops.

Having a single support contract simplifies the operation of the hardware, removing the 'blame game that can be seen with large and complex infrastructure deployments and providing just one throat to choke.

We have seen the development of two styles of converged infrastructure: the product stack, where the supplier provides and ships the hardware as a single product; and reference architectures.

The reference architecture model simply describes which supplier platforms have been tested and certified to work together, allowing the customer more purchasing choice in specific products and specifications but, at the same time, providing a validated system with supplier support.

Hyper-converged infrastructure

Hyper-converged systems take the packaging process a step further by integrating the features of storage and server virtualisation together in a single hardware offering.

Networking is still delivered externally with dedicated switches, but much of the traffic between internal VMs flows over software-based networking in the hypervisor.

The form factor in hyper-converged offerings is the server or node that consists of processor, memory and storage.

Storage services are delivered out of the hypervisor that runs on the node, or in a virtual machine on the node. Resilience is achieved by replication of data across multiple nodes and so hyper-converged systems often have a minimum node-count in their initial deployment.

Hyper-converged offerings are aimed at reducing the administrative or operational burden compared with traditional or converged systems.

Storage is significantly simplified and, in most cases, integrated into the node, and hence is hidden from the administrator.

The deployment method for hyper-converged systems is much easier than traditional systems, with the addition of nodes being a case of ‘plug in and go, with little additional configuration work.

From a cost perspective, scalability means systems can be scaled one node at a time, which is very attractive for small to medium-sized enterprises (SMEs).

Converged versus hyper-converged - when and why?

Both converged options look to simplify operations, so when would one system be more appropriate over another?

From a scalability perspective, converged systems are more suited to larger applications such as Oracle, Microsoft Exchange or SAP installations.

Dedicated storage provides for higher availability. To minimise the impact of device failure on application performance, converged systems can use fully featured storage with array functionality such as predictive failure, where a disk device is replaced in a controlled fashion.

By contrast, hyper-converged systems usually rely on node-based recovery that can generate significant network traffic from storage rebuilds.

Looking at scalability from a different angle, converged systems tend to be designed to scale to a specific size with expansion being non-trivial.

In contrast, hyper-converged offerings are designed to grow with the addition of compute and storage resources in a cluster configuration. This can make hyper-converged systems more appropriate for smaller-scale deployments with smaller virtual machine requirements that expand over time.

Making the choice between converged and hyper-converged is a case of looking at the requirements of the applications being deployed. For smaller organisations, where dedicated storage and networking skills don't exist, hyper-convergence is a great choice. But, where more control over the configuration is required - such as to dedicate or partition resources - then a converged system is a better option.

Market roundup - converged infrastructure

There is a large range of converged systems in the market, covering products and reference architectures.

Hitachi Data Systems offers converged systems that support three main hypervisor systems - VMware vSphere, Microsoft Hyper-V and Red Hat Linux.

The systems offer three configurations (UCP 2000, UCP 4000e and UCP 4000) that scale from 24 to 128 servers, with a maximum of 1,536GB of memory per server (2,048GB for UCP2000). Network connectivity includes 10Gbps Ethernet and 2Gbps to 16Gbps Fibre Channel.

Storage for the systems is based on Hitachi VSP G-series. Hardware resources are managed through two software systems: UCP Director and UCP Advisor. These integrate into hypervisor management tools such as VMware vCenter to enable hardware resource provisioning.

HPE has a large range of converged systems designed to fit a variety of workloads. This includes Converged System 700, designed for mixed workloads and which uses HPE BL460c Gen9 blade servers and 3PAR StoreServ 7000 series storage.

Meanwhile, HPE Helion CloudSystem offerings are designed for cloud-based systems such as HPE's Helion software (OpenStack and Stackato). HPE also now offers composable infrastructure through Synergy, a system that enables customers to combine hardware resources (storage, network compute) to meet specific application requirements.

Since its merger, Dell-EMC has an expanded portfolio of products from the two companies. Converged systems include Vblock systems, based on the original EMC VCE partnership with Cisco and so include Cisco UCS servers and networking.

VxBlock systems use similar hardware but have started to introduce VMware NSX for networking. For storage, VxBlock/Vblock 350 systems use Dell EMC Unity arrays, 540 systems use XtremIO, and 740 systems use VMAX, all either as all-flash or hybrid (except XtremIO). Systems scale from two to 512 servers with up to 15PB of effective storage capacity.

NetApp has focused on reference architectures marketed under the FlexPod brand. Architecture offerings cover a wide range of systems, including Datacentre (for cloud-based workloads), Express (entry level), Select (Hadoop), Security, Enterprise Apps, VDI, Database and Cloud.

IBM has developed VersaStack, a set of reference designs, developed in conjunction with Cisco Systems. The systems use Cisco UCS servers, Cisco networking (Nexus and MDS) with IBM FlashSystem storage, IBM StorWize or IBM's SAN Volume Controller (SVC). Each of the systems uses Cisco's UCS Director software for hardware management and VMware vCenter with ESXi for virtualisation.

Oracle offers a range of systems that make use of the company's software products, including ZFS-based storage (including all-flash), Oracle VM Server, Oracle Linux and Oracle Solaris.

The Oracle Private Cloud Appliance provides up to 30 compute nodes per rack, with two 22-core Intel Xeon processors, 256GB DRAM (per node) and

Infiniband in-rack connectivity. Shared storage is delivered by an Oracle ZS3-ES appliance with Infiniband connectivity and hybrid flash and HDD capacity.

Market roundup - hyper-converged infrastructures

Hyper-converged systems are available from startups and the incumbent hardware suppliers.

Nutanix pioneered the hyper-converged market and now offers a wide range of hardware systems across four main product sets (NX-1000, NX-3000, NX-6000 and NX-8000).

NX-1000 series models start with four nodes per appliance and single Intel Xeon E5-2609 processors (eight cores), up to 256GB DRAM and three 3.84TB solid-state drives (SSDs).

At the high end, NX-8000 nodes offer Intel Xeon E5-2699v4 processors (44 cores), 1.5TB of DRAM and 24 SSDs (up to 1.92TB). Nutanix offerings can run VMware vSphere or Acropolis, Nutanix's own hypervisor system.

Hitachi Data Systems offers four hyper-converged node configurations, based on hybrid or all-flash storage. V240 and V240F are 2U four-node configurations with a range of Intel Xeon configurations (maximum E5-2680 v4 with 14 cores) and up to 512GB of DRAM. Storage capacities vary up to 6TB raw (hybrid) or 19TB raw (all-flash).

V210 systems scale up to dual 22-core E5-2699 Xeons, 1.5TB DRAM and 38TB (flash) or 60TB (hybrid) per node. Both configurations scale from two to 64 nodes and support 10Gbps Ethernet networking.

HPE offers hyper-converged systems based on VMware vSphere, using HPE StoreVirtual VSA to deliver distributed storage. Each HPE Hyper Converged 380 node supports two Intel Xeon E5 processors (six to 18 cores) with up to 1.5TB of DRAM. A single node can support up to three storage blocks for a maximum of 25.2TB per node (mix of HDD, SDD or hybrid) and up 16 nodes in a cluster.

HPE recently acquired Simplivity, a hyper-converged startup. Currently, Simplivity is advertising systems based on white-box servers or hardware from Cisco, Dell and Lenovo. White-box OmniCube nodes provide a range of hybrid or all-flash offerings that scale up to dual Intel Xeon E5-2600v4 processors with 1,443GB of DRAM and 40TB of storage capacity per node.

Dell EMC hyper-converged systems include VxRAIL, and it has five model series: G - general, E - entry, V - VDI, S - storage dense, and P - performance intensive.

These are delivered as either 1U or 2U nodes with dual Xeon E5-2600 processors and from 64GB to 1.5TB of DRAM. All models - except S-Series - are available as all-flash or hybrid configurations. VxRAIL systems use VMware Virtual SAN for distributed storage.

NetApp recently entered the market with NetApp HCI, a hyper-converged offering that uses SolidFire technology to deliver clusters with dedicated storage and compute nodes. Three configurations are available (small, medium, large) with up to 16 and 36 cores and 256GB to 768GB of storage available per compute node and up to six drives per storage node (480GB to 1,920GB SSDs).

Cisco now offers hyper-converged systems using Springpath technology in its HX series of appliances. Cisco Hyperflex offers four node sizes - HX220c M4 is a 1U node available in hybrid or all-flash (23.2GB all-flash or 480GB SSD and 7.2TB HDD).

HX240c is a 2U node also available with hybrid or all-flash (38.4GB all-flash or 1.6TB SSD and 27.6TB). Essentially, the model ranges provide for higher storage capacities and all-flash models.

Scale Computing focuses on delivering hyper-converged systems based on the KVM hypervisor and a proprietary distributed storage platform.

The HC3 range has several model groups: HC1000 for entry level; HC2000 for mid-range; and HC4000 for high-end requirements.

Each model type varies in storage, processor and network performance. The entry-level HC1100 has a single Intel Xeon E5-2603v4 CPU, 64GB of RAM and

four 1TB SAS HDDs. This scales to the HC4150 with two Intel Xeon E5-2640v3, 384GB of DRAM and two 400GB SSDs.

A number of suppliers also offer their products as software-only, or bundled with a hardware appliance. These include Pivot3, Atlantis Computing and Maxta.

Next article

In this e-guide

Tips for buying hyper-converged systems

Clive Longbottom, guest contributor

Hyper-converged infrastructure (HCI) is a concept that is already somewhat tarnished. Many storage suppliers are touting systems as being hyper-converged, when they should just be described as converged.

Worse still, many suppliers in the storage and server business are using hyper-converged to describe approaches to both storage and combined server, storage and network platforms.

This means trying to make sense of what is being referred to can be problematic. For simplicity, hyper-converged systems can be described by two phrases.

First, a converged system brings together all the hardware and software required for a single task - in the case of storage suppliers, this is the storage and management of data. As such, a converged storage appliance may still have a server chip in it, memory, network ports and storage systems, but you would not be able (nor would you want) to install and run, say, Microsoft Exchange on it.

Second, a hyper-converged system brings together all the hardware and software required for multiple tasks - a shared environment that runs workloads that still use data storage, but are not just focused on the storage environment. As such, alongside all the software that allows the platform to run, users can install and run applications and services that carry out a business purpose, such as enterprise resource planning (ERP), customer relationship management (CRM), big data analytics, and so on.

To muddy the waters a little more, many hyper-converged systems, while fully capable of running multiple workloads, are targeted at specific workloads - particularly those such as virtual desktop infrastructure (VDI) or big data analysis.

Hardware meets software

It's a confusing world we live in - but let's try to clear the waters.

A hyper-converged platform must start with the hardware. Servers, storage and networking need to be brought together in a single environment. Arguably, the first hyper-converged platform was the mainframe, but the first real offering in the Intel space was the Unified Computing System (UCS) platform from Cisco, launched in 2009.

This was based on earlier approaches of creating engineered systems using 'bricks, where hardware components could be built up within a specialised

chassis to create a computing system. However, most brick-based systems were still based on the use of storage area networks (SANs), so missing out on the advantages of having storage components close to the server.

By providing a chassis approach to configure various hardware components, several enhancements to the engineering could be applied. For example, whereas a standard scale-out approach uses Ethernet to create an IT platform built from different server and storage components, hyper-converged systems can use proprietary connections within the overall system.

With little or no need for the internal systems to be highly standardised, hyper-converged systems can be highly tuned to maximise performance. Only when connectivity is required between the hyper-converged system and the rest of the world do standards become necessary.

As Cisco's UCS started to show promise, other providers brought offerings to market, such as VMware Vblocks, Dell FX2, HP Hyper-converged and IBM PureSystems. The majority of first-generation HCI systems used proprietary firmware and software to support standardised operating systems running on them.

Adding value

A new group of software-focused suppliers came to market to provide hyper-converged operating systems that added more value through advanced

functions, such as virtual machine (VM) and/or container management, alongside data enhancements such as global data deduplication and compression.

Some of these suppliers provided both software that could be installed on other suppliers' hardware and a complete hardware-plus-software system of their own; others just the software.

The main players in this area have been Nutanix and SimpliVity. Since being acquired by HPE, SimpliVity is no longer seen as being the agnostic player it once was. Nutanix is making the most of this by providing a supported version of its software on HPE hardware - much to HPE's annoyance. Nutanix has partnerships with Dell EMC and Lenovo.

Another supplier is Pivot3, which has developed its own software stacks - vSTAC and Acuity - that enable advanced policy-based data management to optimise mixed workloads across HCI platforms. Alongside providing fully configured hardware/software offerings of its own, Pivot3 claims that its software allows greater flexibility in choice of third-party hardware allowed for its system.

This is both a benefit and a drawback: being able to apply HCI software against existing hardware provides obvious cost savings in not having to purchase new hardware, but the lack of highly engineered interconnects and internal data

buses counts against gaining the highest possible performance from the resulting platform.

Further changes in approach and various acquisitions later, VMware has taken a more software-based approach with its vSAN ReadyNode, Dell EMC has moved to bring together Dell and EMC's approaches via its VxRail Appliances, HPE has acquired SimpliVity to bolster its offerings, and IBM has pretty much left the market, moving instead to its cloud-first SoftLayer model. Dell was also a Nutanix partner, and this has continued through into the Dell EMC environment via Dell EMC's XC platform.

What to look for

The first check a buyer must carry out when considering HCI is to ensure that the platform is as future-proof as possible. Check that the system is guaranteed to be supported for the viable life of the platform. Ensure the supplier is still going to make building blocks in the form of incremental server, storage and network capabilities available. Ensure the supplier has a good story on supporting emerging technology - for example, will the system be able to support server-side NVMe PCI or DIMM storage? The last thing you need is a system that costs a lot to acquire and then cannot provide ongoing flexibility in the future.

Next, check on the software capabilities. Even though HCI is based on an engineered system bringing different resources together, it is unlikely to be the

only system your company will use. Management, orchestration, monitoring and other capabilities in the software must be able to traverse beyond the box itself and integrate seamlessly with other systems - whether they be physical, virtual, private or public cloud-based. Accept that the innards of the HCI system itself may be proprietary to a degree, but ensure that the system supports de facto standards where the rest of the world is concerned.

Ask for references where the system has already been used for workloads similar to those your organisation will run on it. Ask if the mix of workloads is viable, or if it will result in sub-optimal performance. As with pretty much anything, trying to design something to be all things to all people will result in it failing somewhere in its main goals. For example, a system that has been pre-tuned to run VDI loads is unlikely to be good at big data analysis. One that is targeted at supporting high-throughput transactions may not be so good for large file management, for example, as used in design work.

Choose carefully

The HCI market is still maturing, so carrying out due diligence on a specific supplier’s performance to date may not throw up anything useful. As already discussed, the main hardware providers are already on their second or third attempt at getting it right; some of the software providers have disappeared or been acquired. The best thing to do is to ask around - find those with a similar workload or workload mix to your organisation and ask whether they have had any success with an HCI platform.

Overall, HCI does have a role to play. It provides a simpler means of acquiring a total platform that is faster to provision and easier to manage. HCI can perform better than platforms built from disparate discrete components. A well-engineered system provides flexibility for adding incremental resources as required without downtime.

However, a poorly chosen HCI system could end up being a white elephant - one that cannot be upgraded or that does not interoperate well with the rest of the IT platform.

Choose carefully - or choose based on tactical needs, understanding that the platform you choose may not have a significant useful life as part of your organisation’s overall IT platform.

Next article

Hyper-convergence: the good, the bad and the ugly

Rene Millman, guest contributor

As organisations look beyond cloud for the next technology trend to boost their business efficiency and agility, an investment in hyper-converged infrastructure (HCI) is now on the radar for many.

According to market watcher IDC, sales of HCI grew by 104.3% year on year in the third quarter of 2016, generating $570.5m worth of sales, amounting to 22% of the total converged systems market value.

Meanwhile, figures from rival analyst house Gartner suggest the market for HCI will be worth $5bn by 2019, with the technology gaining traction in enterprises as more businesses investigate its potential benefits.

In a recent Computer Weekly article, Simon Robinson from analyst 451 Research noted that the presence of incumbent IT suppliers in the HCI market meant the technology was becoming more mainstream (see box, p20).

In this e-guide

Step up from convergence

But before any organisation embarks on a move to adopt HCI, it needs to get a firm grasp on what the term really means. It is generally accepted that HCI is a step up from converged infrastructure, which, in basic terms, is where compute, storage, network and virtualisation components are brought together and packaged as a pre-tested and integrated offering.

HCI builds on this concept by including components with little distinction from each other and which are further combined with software-defined functions that run on these appliances. When capacity dries up, you simply add more systems.

When upgrading or changing IT systems, disruption must be kept to a minimum, because every second of downtime represents lost money.

Albie Attias, managing director of IT supplier King of Servers, says hyper-convergence counters this by offering an out-of-the-box infrastructure that is readily scalable. ‘It’s quick, easy deployment is very attractive to enterprises as the drain on IT resource to install and configure the hardware is minimal, he says.

Tackle rising costs

Hyper-convergence can also tackle rising IT costs, says Matt Foley, director of Hewlett Packard Enterprise’s European solutions and technology group,

because entry-level commitments are very low, with users able to start scaling at just two nodes.

‘Due to their all-in-one configuration, hyper-converged systems allow you to build, scale and protect IT infrastructure more affordably and effectively than any other option available, he says. ‘And software-defined intelligence reduces operational management, providing automated provisioning of compute and storage capacity for dynamic workloads.

Neil Thurston, chief technologist of hybrid IT at service provider Logicalis UK, cites simplified maintenance as another advantage of HCI. ‘The stack self-builds itself initially, and when new units are added the stack self-patches the hardware and software layers of the platform and self-upgrades the software, he says.

The need for data protection is also catered for in HCI, says Thurston. ‘Data is typically spread across the stack using erasure coding, which is far more efficient on disk than traditional storage array methods of replication and keeping multiple full copies of the same data.

Another benefit of HCI is that the hardware is often simpler to manage, says John Abbott, founder and research vice-president at analyst 451 Research. ‘The idea is to eliminate the need for specialist expertise in storage, servers and network management, he says.

‘To avoid the introduction of yet another management console, plug-ins and APIs [application programming interfaces] will increasingly be developed for integration with existing toolsets. As support for multiple hypervisors becomes more common, more consistent virtual machine-level management tools will emerge that work in the same way across multiple hypervisors, be it ESXi, Hyper-V or KVM.

‘Infrastructure is made up of standard x86 server technology without the need for external storage. Everything is accessed in a standard way through the virtual machine hypervisor, which helps to reduce complexity.

Security concern

But although adopting HCI has many benefits, enterprises must also be mindful of the drawbacks, one of which is security, says Liviu Arsene, senior e-threat analyst at Bitdefender.

Such systems will face attacks against the control plane, data plane and management infrastructure, and mitigating these threats will not only require understanding the environment in which they operate, but also an appreciation of how this may affect the security tools’ functionality, he says. ‘A security architecture that is able to support automated infrastructures while being able to spin up and spin down, maintaining performance, needs to be an integral part of the migration toward hyper-convergence and software-defined datacentre environments.

These drawbacks can be overcome by using security tools that are compatible with any type of hypervisor technology, as this guarantees they can be deployed and scaled regardless of how many nodes are added to the hyper-converged infrastructure.

Changing landscape

Gradually, hyper-converged infrastructure software platforms are beginning to support a broader range of hardware from multiple sources and hypervisors, says 451 Research's Abbott. 'The initially limited recommended workloads are also broadening to include analytics and mission-critical enterprise applications, he says. 'Test/dev, DR [disaster recovery], Tier 3 and Tier 4 workloads are common, but adoption is also growing in more business-critical applications, such as Oracle, SAP and Exchange.

There is also an emergence of more specialised forms of hyper-converged infrastructure, optimised for specific rather than general-purpose use cases - Hadoop-based analytics, for example.

In the future, we should also expect hyper-converged platforms to support microservices and cloud-native services and also to support platform as a service (PaaS), allowing organisations to create and run applications without being troubled with the core infrastructure and platform services or location.

In this e-guide

Hyper-convergence goes mainstream

Hyper-convergence is now becoming a core weapon in the armoury of suppliers looking to demonstrate to customers that they can meaningfully change the fundamental economics of running core and edge IT infrastructure.

This is not only a lynchpin of any IT transformation strategy, but absolutely essential for any supplier that wants to demonstrate ongoing relevance in the era of public cloud.

In January 2017, Hewlett Packard Enterprise announced it would pay $650m cash for seven-year-old startup Simplivity.

VMware recently said it has exceeded 7,000 customers for VSAN, the hyper-converged infrastructure software that is optimised for VMware environments.

Meanwhile, the combined Dell-EMC has highlighted hyper-converged infrastructure as a key growth market.

Other major players are still hoping to make a run at hyper-converged infrastructure, including Cisco with HyperFlex, and there are still startups with momentum, including Pivot3 and Scale Computing.

In this e-guide

One crucial test of whether a market is becoming mainstream is the extent to which it is embraced by large incumbent suppliers. As far as hyper-converged infrastructure is concerned, we are now approaching that point.

Source: Simon Robinson, 451 Research.

In this e-guide

Getting more CW+ exclusive content

As a CW+ member, you have access to TechTarget's entire portfolio of 140+ websites. CW+ access directs you to previously unavailable "platinum members-only resources" that are guaranteed to save you the time and effort of having to track such premium content down on your own, ultimately helping you to solve your toughest IT challenges more effectively—and faster—than ever before.

Take full advantage of your membership by visiting www.computerweekly.com/eproducts

Images; stock.adobe.com