A SOLIDFIRE PAPER

A Service Provider's Perspective: Designing the Next Generation Data Center

NetApp®

Table of Contents

Why Read this Guide? 3

Chapter 1: The Service Provider Business is Changing 4

Chapter 2: Why Invest in the Next Generation Data Center 5

Chapter 3: The Five Principles of a Successful NGDC 7

Chapter 4: Scale-Out 8

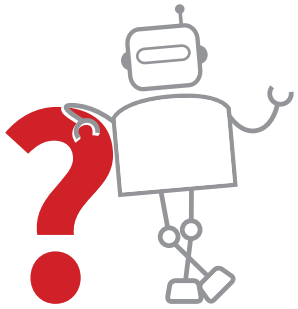
Chapter 5: Guaranteed Performance 10

Chapter 6: Automated Management 12

Chapter 7: Data Assurance 15

Chapter 8: Global Efficiencies 17

Conclusion 18



Why Read This Guide?

At the turn of the 21st century, virtualization promised a flexible, centralized data center. But the global IT community wasn't aware of the need. Early adopters and innovators realized business advantages well before virtualization gained mainstream adoption. Today virtualization is pervasive, reaching market maturity in its lifecycle. Cloud is the new disrupter, offering innovators a ride on its growth trajectory.

Consider the following studies:
451 Research states that only 32% of enterprise applications are running in multi-tenant cloud hosting environments.[1]

IDC forecasts public cloud spending will more than double to $127.5 billion by 2018 — $82.7 billion in SaaS, $24 billion for IaaS, and $20 billion in PaaS.[2]

Computerworld's 2016 Forecast Study indicates an ongoing trend of moving business-critical applications to cloud, with cloud projects trumping every other technology initiative.[3]

Cisco's global cloud index report predicts that by 2018, more than 78% of workloads will be processed by cloud data centers; 22% will be processed by traditional data centers.[4]

If you are a service provider (SP) thinking about your next generation cloud and hosting services, this paper will help you navigate a go-forward strategy. It provides next generation data center (NGDC) trends and business strategies in the context of five architectural principles. It examines key market influences fueling IT transformation in the SP market and the technological trends that support them. The five architectural principles are scale-out, guaranteed performance, automated management, data assurance, and global efficiencies. Cloud infrastructure delivery models such as IaaS, private clouds, and software-defined data centers (SDDC) are foundations for the NGDC.

In an era where SPs are expected to ensure production-grade infrastructure and support while their enterprise customers deploy a plethoric array of new applications and data, these models demonstrate how to help enterprise customers eliminate bottlenecks, increase self-service, and move their business forward. The NGDC applies a software-defined everything (SDx) discipline in a traditional hardware-centric business. This approach is helping SPs gain significant business advantages in a highly competitive market.

The five architectural principles comprise a fundamental, integrated approach affecting the software, processes, and people that support each layer of the data center stack. They represent a paradigm shift from enabling technology at the speed of business to enabling business at the speed of technology. Driven by the unchained cloud economy offering innovations in infrastructure design, service design, and fundamental cloud services, SPs are well-positioned to embrace this shift.

Cloud economics can be described as the discipline of maximizing network, compute, and storage resources in conjunction with cloud computing to satisfy customer needs. NGDC cloud economics offers a new discipline: one that uses SPs as the glue for enterprise IT cost containment, the force for enterprise IT innovation, and the architect of new revenue streams.

1 451 research. Cloud Computing - Wave 7 2014. Available at: https://451research.com/report-long?icid=3186

2 IDC. "IDC Forecasts Public IT Cloud Services Spending Will Reach $127 billion in 2018 as the Market Enters a Critical Innovation Stage." Available at http://www.businesswire.com/news/home/20141103005113/en/IDC-Forecasts-Public-Cloud-Services-Spending-Reach

3 IDG. Computerworld Forecast Study 2016 Available at: http://www.idgenterprise.com/resource/research/computerworld-forecast-study-2016/

4 Cisco. "Cisco Global Cloud Index: Forecast and Methodology, 2014–2019 White Paper." Available at http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud- index-gci/Cloud_Index_White_Paper.html

Chapter 1

The Service Provider Business is Changing

Offering a cloud is not enough

Today's service provider landscape is causing a dramatic shift in how services and solutions are being delivered to enterprise customers. As the dominance of AWS (Amazon Web Service) and Microsoft Azure continue, traditional SPs are fighting to deliver higher value narratives with a focus on managed services or vertical market solutions to differentiate their business. The decision of which data center technologies are deployed is more important than ever before.

Many SPs are embracing NGDC approaches; others have been slow to adopt. Below we examine some ideologies and business structures that impede — and some that feed — the changing role of SPs in the new cloud economy:

What's not working for service providers:

- Undercutting AWS and Azure pricing in an effort to build a successful business
- Building a platform without a clear understanding of current and future enterprise cloud hosting needs
- The inability to rapidly adjust services and solutions when customer needs evolve
- Overlooking opportunities in their partner ecosystem for new markets and revenue streams

What is working for service providers:

- Full application managed hosting lifecycle support. From migration, onboarding, and configuration support to maintenance, reporting, and application-level support, SPs are delivering complete lifecycle services on top of their IaaS/private/hybrid infrastructure.
- Participating in managed onboarding services to AWS and Azure for enterprise workloads or applications that are a good fit. Successful SPs are then cross-connecting/integrating these workloads with other private/hybrid hosted infrastructure platforms.
- Broad services delivery. SPs are attracting a larger subset of enterprise applications by delivering a broader set of platform services. These services are made available through one portal or a "single pane of glass" with standardization of the same managed services across the SP infrastructure.
- Expanding revenue opportunities via strategic partnerships. SPs are stitching together service solutions with systems integrators (SIs) and value added resellers (VARs) offering specialties such as SAP and Oracle. These partnerships are creating new revenue streams in nonstandard environments as well as filling voids for customers.

Regardless of size, successful SPs are recognizing similar success formulas. Those that offer next generation infrastructure services are reducing churn and attracting new markets. But even in today's cloud economy where agility, scalability, and automation rule, successful SPs follow this basic business maxim: They seek to understand their customers' needs first and then demonstrate their readiness to fulfill those needs.

Chapter 2

Why Invest in the Next Generation Data Center?

Pressure to remain relevant in an evolving market

The SP landscape today is highly competitive with an existential threat toward rapid commoditization. Public clouds like AWS and Azure have been demonstrating the possibilities for next generation cloud infrastructure deployment: elasticity and high performance at lower costs.

The challenge for traditional hosting and cloud providers is that these hyperscale SPs have become customer benchmarks for infrastructure agility and pricing. Non-hyperscale SPs are forced to defend their higher prices, slower provisioning times, and less agile infrastructure while being compared to their mega competitors. The transformational power of new IT platforms is explained in Accenture's Technology Vision 2016 report:

> "Tech companies and enterprises that are born digital, such as Amazon, Google, and Alibaba, have long understood the power of digital technologies. But look a little closer. Many of these companies' most groundbreaking innovations are not products or services; they are the platforms on which these products and services are built and the business models that these platforms enable. Such platform-based business models fundamentally change how companies can do business."[5]

Remaining relevant as a service provider in an evolving market means moving forward. Recent innovations in infrastructure design is changing how both hyperscale and non-hyperscale cloud hosting companies do business. The key is a fundamental, integrated approach across the entire data center infrastructure. Rapid elasticity, high efficiency, and cost-containment is not exclusive to AWS or Google. NGDC technologies are helping SPs of all sizes elevate their business narratives and differentiate their service offerings.

NGDC technologies are helping SPs of all sizes elevate their business narratives and differentiate their service offerings.

Automation enablement

The ability for SPs to cost-effectively deliver infrastructure that supports evolving customer needs — especially at scale — is directly correlated to automation enablement. In traditional data centers, each infrastructure component has an independent set of management software tools and inherent integration points between the hardware and associated applications. Time-consuming processes, human error, and lack of agility often impede an SP's ability to build, deploy, and support customer applications. Dynamic workloads simply cannot scale in a traditional infrastructure without reconfiguration and additional hardware investments, putting more pressure on a constrained budgets and support staff.

NGDC service offerings are abstracted from the underlying infrastructure. These services interface via APIs, translating the business requirements into resource offerings. Automation eliminates complex and rigid hard coding to siloed projects. Resources can be dialed up or down and are rapidly modified at the application level to meet changing application requirements. This control plane shift allows resources for compute, network, and storage to be deployed faster, more accurately, and with a higher level of quality control.

5 Accenture Technology Vision 2016. Available at https://www.accenture.com/us-en/insight-technology-trends-2016

Automation enables service providers to deliver more value for enterprise customers and higher profit margins for themselves using less resources. In an automated environment, staff changes will naturally evolve as the ratio of virtual servers to administrators increases. There is also a shift away from traditional UI-focused storage administrators to API-focused developers and cloud architects. In the NGDC, developers spend less time on repetitive tasks and more time doing what they were hired to do: innovate, accelerate outcomes, and provide complete solutions for their enterprise customers.

In the NGDC, developers spend less time on repetitive tasks and more time doing what they were hired to do: innovate, accelerate outcomes, and provide complete solutions for their enterprise customers.

Innovation enablement

Delivering innovation at cloud scale is a formidable achievement for many SPs, leaving some with more questions than answers. What technologies will add significant value to our current customers and support revenue growth for new ones? How can we seize opportunities that leverage the cloud economy and further differentiate our hosting business? Do we fight directly with AWS or Azure or partner and resell their services, or both?

In his blog "Hey Enterprise IT Vendors! Your Customers Need Help!," Oracle's Chuck Hollis writes: "The very nature and role of the corporate IT function is being questioned, evaluated, and debated in a way that we haven't seen in a very long time."[6] IT has become a core competency in business. Thanks to the cloud, every company — no matter the industry — is a de facto technology company, and enterprise IT needs more help than ever.

Enterprises are looking for service partners to fill the gaps between what they need and what they can cost-effectively deliver internally. Customer-centric service providers adopting NGDC innovations are filling those gaps to become valued, trusted extensions of the enterprise.

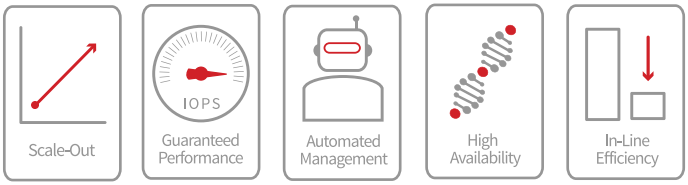
Reduced risk and complexity

Enabling innovation without expanding risk is paramount. Each time a service provider can't react rapidly to address customer needs, operational and financial risks are heightened. Legacy infrastructure is inherently complex, involving the challenge of managing multiple, disparate platforms and relying on manual operations. Traditional platforms and systems hold SPs back from delivering the solutions enterprise customers expect from their IT partners.

The NGDC offers a software-defined converged approach to infrastructure management, a highly organized design methodology, and iterative, agile processes measuring all aspects of the ITIL lifecycle. This transformational approach not only reduces business risks and operating costs, it fuels innovation and competitive advantages.

The very nature and role of the corporate IT function is being questioned, evaluated, and debated in a way that we haven't seen in a very long time.

6 Chuck's Blog. "Hey Enterprise IT Vendors! Your Customers Need Help!" Available at http://chucksblog.typepad.com/chucks_blog/2013/10/hey-enterprise-it-vendors-your- customers-need-help.html



Chapter 3

The Five Principles of a Successful NGDC

NGDC design involves a unified approach, applying five architectural principles across each layer of the data center framework. This approach affects the software, processes, and people supporting the entire delivery stack. The five principles are:

Scale-Out: In the NGDC, resource pooling provides nondisruptive horizontal expansion across the data center layers. Projecting business needs two years, three years, or five years out is a non-issue because your infrastructure can scale incrementally as customer requirements change. Linear, predictable growth of capacity and performance with guaranteed quality of service (QoS) is an essential operative in the NGDC, which makes a scale-out design critical to success.

Guaranteed Performance: Raw performance is only half the solution in the NGDC. QoS controls must be incorporated across the entire infrastructure or else any guarantee is only as good as the weakest link. Legacy infrastructures were not designed to balance increased network capacity simultaneously with expected levels of service in large-scale cloud deployment. In the NGDC, resources like CPU, memory, bandwidth, storage capacity, and storage performance are dynamically managed to deliver the application experience required and expected.

Automated Management: The core business goal of service providers is to enable innovation and enterprise growth. Employing software automation and API control to take charge of decision points changes the paradigm from the physical limitations of hardware to the unbounded design capabilities of software. NGDC automation maximizes business results with policy-driven provisioning and resource allocation and eliminates endpoint-centric administration. Automating tasks and orchestrating workflows are fundamental for SPs if service delivery needs are expected to be met at scale.

Data Assurance: NGDC data assurance includes seamless infrastructure resiliency without application configuration. Engineering for data assurance can be likened to designing buildings for earthquakes; there is no building that is guaranteed to withstand all earthquake activity, but earthquake-resistant structural design exponentially lessens the probability of disaster. NGDC architects plan for failure while mitigating its likelihood of occurrence with a self-healing, fault-tolerant architecture.

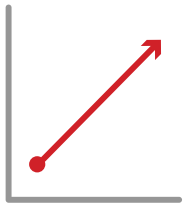
Global Efficiencies: The traditional data center model was built on the premise that more capabilities would require more resources — whether they be physical, financial, or human resources. In the NGDC, service providers are not burdened with excess IT resources. Enabling global efficiencies begins with improved utilization of server platforms, networks, and storage protocols as well as improved utilization of vendors and staff that support each layer.

THE EVOLUTION OF SERVICE PROVIDER INFRASTRUCTURE

Legacy data center	Next generation data center
Single tenant	Multi-tenant
Isolated workloads	Mixed workloads
Dedicated infrastructure	Shared infrastructure
Scale-up	Scale-out
Preprovisioned capacity	Capacity on-demand
Hardware-defined	Software-defined
Project-based	Self-service
Manual administration	Automation

Chapter 4

Scale-Out



The foundational tenet of software control in the NGDC is represented in a scale-out design. Scale-out is positioned to become a predominant architecture in cloud infrastructure deployment because of its ability to offer seamless, transparent resource expansion without the cost and complexity of traditional infrastructure migrations.

Scale-up versus scale-out

Scale-up architectures allow expansion of a limited set of variables because the capability of each component limits their collective performance.

Data and applications will usually proliferate and, ultimately, maximum capacity will be reached in a scale-up architecture. As applications are added, the infrastructure resources are spread out and performance degradation creeps in. With a scale-up architecture, the hardware you operate determines your limits, and faster hardware likely requires significant migration efforts.

By contrast, a scale-out architecture is a distributed architecture. It is not limited to the resource capacity of a single machine. It delivers linear expansion of all variables, allowing you to scale up or down by clustering resources. Scale-out offers a wide range of resource expansion options, allowing the service provider to leverage investments consistently across the data center infrastructure and over the longer-term.

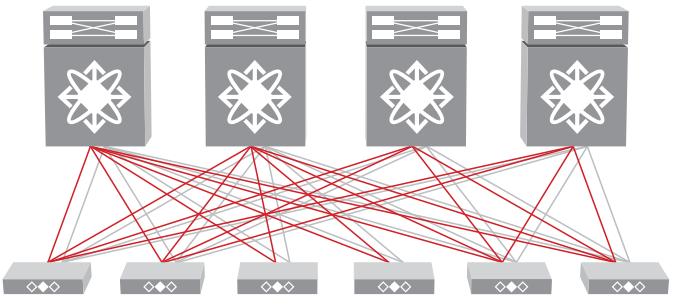
The key to a successful scale-out architecture is making sure the performance from additional nodes can be intelligently provisioned by QoS capabilities. If workloads cannot be protected from each other in terms of performance, service providers will fail to attract new enterprise customers. Predictable utilization is critical in solving the longer-term SP challenges related to consolidation and infrastructure optimization.

Trends enabling scale-out

Resource clustering

The traditional, siloed data center relies on dedicated, physical servers with stranded resources. Those servers, along with their storage footprint, dictate the maximum data that can be managed or stored. Virtualization was built on the concept of sharing compute, memory and IO resources between multiple workloads. Technologies such as KVM, ESXi, and HyperV take this further, enabling resource clustering.

The strategy of resource clustering is key for parallel processing applications, load balancing, and fault tolerance (redundancy) in the cloud. Pooling resources helps service providers enable quick, non-disruptive modifications with singular orchestration software. Adding another server to the pool effectively expands CPU and memory, and resources can be scaled in or out to meet changing business demands.



Leaf-spine

Before the adoption of Ethernet fabric, network topologies were built in layers that resembled a hierarchical structure and were typically segmented into pods that inhibited the location of VMs and other devices. Originally designed for use in enterprise networks, the traditional model consisted of core routers, aggregation switches, and access layer switches that traditionally required the use of Spanning Tree Protocols (STP) to facilitate scaling. This topology added significant complexity to both network design and troubleshooting.

Leaf-spine design has only two layers: leaf and spine. Access switches in the leaf layer connect to devices such as VMs, firewalls, edge routers, and load balancers. The backbone of the network is the spine layer with switches that perform routing. Each leaf switch is interconnected with every spine switch via dynamic Layer 3 routing and the use of Equal Cost Multi-Pathing. Adjustments and determination of the best path are based on responses to network change.

As leaf-spine architectures scale to thousands of nodes, latency and bottlenecks between access-layer switches — common in traditional architecture — are minimized because these switches are no more than a single hop away. Network vendors like Arista are enabling speed, density, predictability, and scalability with this NGDC topology.

Solid-state arrays

Even as compute and network administrators began leveraging the efficiencies of cloud, storage technology remained mostly static. In legacy storage, service providers had little choice but to use separate storage networks and distinct storage pools to deliver the appearance of traditional “tiers” of performance to its customers. Storage administrators managing traditional scale-up architectures were required to add a fixed number of drive shelves to add capacity. These disk arrays required upfront capital outlays for all resources anticipated for future expansion.

Scale-out solid-state arrays are expanding the once siloed, static storage ecosystem. In the NGDC, pools of storage (GB, IOPS) are shared resources. The ability to cluster these resources provides incremental, as-needed, on-demand scaling. Clustering allows service providers to more effectively manage cash flow and align purchases more closely with storage consumption.

NGDC scale-out storage provides the flexibility to independently, nondisruptively scale both capacity and performance in a predictable linear pattern over time. This means SPs can strategically scale in or out and distribute data and traffic over any number of nodes while increasing the scope of data services.

Chapter 5

Guaranteed Performance

In the NGDC, raw performance is only half of the solution; just delivering a simple storage service is not good enough. QoS resource controls must be utilized across the entire infrastructure or else any guarantee is only as good as the weakest link. CPU, memory, bandwidth, storage capacity, and storage performance is dynamically managed in the NGDC to deliver the application experience expected and required.

Trends enabling guaranteed performance

Monitoring tools

Information is power. Automated performance monitoring provides the critical intelligence needed to manage usage and availability of all customer deployed resources. Service providers are using software to monitor, report, and analyze live and historical data for both internal and customer use. These tools are enabling service-level improvements, reducing security risks, and cutting operational costs.

Service providers are using tools such as Nagios, Zenoss, Nimsoft, and ScienceLogic to perform monitoring and performance tracking for infrastructure analytics and visualization. Dashboards map key performance indicators (KPIs), highlight infrastructure status, and provide drill-down data for problem evaluation and resolution. Monitoring tools provide the quantitative data supporting the ability to deliver hard service level agreements (SLAs) to enterprise customers.

Network QoS

In the NGDC, enterprise architects are managing performance by segmenting network traffic into prioritized queues while simultaneously controlling how much of the pipe is used for storage, applications, replication, and DR activities. Network QoS is designed to guarantee service delivery for critical application traffic without breaking other flows and to enable granular control of bandwidth utilization.

Applying QoS at the network layer is a key enabler to facilitating convergence and end-to-end SLAs. For service providers, network QoS helps deliver the application performance experience of a dedicated cloud platform in a multi-tenant infrastructure. Intel's Marketing Programs Manager Brian Yoshinaka explains the impact of network infrastructure choices:

"Today, as IT departments prepare to deploy internal cloud environments, it's significant to evaluate how network infrastructure choices will impact the cloud's ability to meet its service level agreements (SLAs). Terms commonly used to describe cloud-computing capabilities, such as agility, flexibility, and scalability, should absolutely apply to the underlying network as well." [7]

7 Data Center Knowledge. "Tips for simplifying your cloud network." Available at Available at http://www.datacenterknowledge.com/archives/2011/12/14/tips-for-simplifying-your-cloud-network/

Storage QoS

In storage, the term QoS has lacked a consistent definition among industry vendors, creating understandable skepticism among cloud providers. Legacy storage models operate various virtual workloads on discrete pools of storage. The model was created to protect the performance and availability of a single application across a single storage array when each application has a diverse I/O pattern that must be supported.

While these systems could provide capacity on-demand, they struggled with allocating performance resources efficiently because they were not built to support the individual capacity and performance requirements of collective workloads. Service providers then purchased more storage than they needed, driving costs up and efficiencies down.

ESG Lab Analyst Aviv Kaufmann explains the legacy platform challenge:

"Traditional storage infrastructures have evolved to better meet the demands of enterprise workloads by leveraging new technologies as they become available. But 'evolving' and 'incorporating' indicate a compromise; this is not the same as being purpose-built for the task. Traditional storage architectures can be configured to meet the SLAs of most of today's consolidated workloads, but to do so often requires over-provisioning and overspending for an infrastructure that is inflexible, complex to manage, and slow to respond to the needs of the business."[8]

In the NGDC, storage QoS means that performance controls are enforced at a granular level by ensuring a guaranteed amount of storage resources to each application. Every application provisioned is assigned a guaranteed amount of IOPS and those allocated IOPS are respected consistently, regardless of any other application activity, capacity level, or I/O pattern.

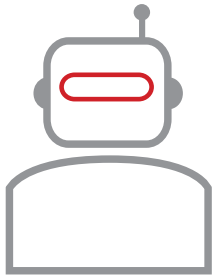
The NGDC uses a scale out, shared-nothing storage design that can consolidate workloads without increasing application risk profiles. It offers consistent, predictable performance based on the requirements of and value to the business while isolating and protecting the workload from other workloads or hardware/software faults.

The precision balancing act between performance and capacity in NGDC storage design enables SPs to productize and deliver a whole suite of new services on the same consolidated platforms. This consolidation enables application density, reduced operations costs, and streamlined management. Storage QoS also enables SPs to offer an SLA supporting guaranteed storage performance based on latency or IOPS.

What virtualization enabled in compute resources is now available in NGDC storage architecture: pooling and isolation of separate resources, reserving resources for critical workloads, and moving resources dynamically from workload to workload.

The precision balancing act between performance and capacity enables guaranteed performance and true QoS while reducing operational expenses.

8 ESG. "Quantifying the Economic Value of a SolidFire Deployment." Available at http://info.solidfire.com/rs/solidfire/images/ESG-Lab-WP-Quantifying-SolidFire-Value-Feb.pdf



Chapter 6

Automated Management

The extent to which service providers can offer rapid resources on-demand and as a service is directly correlated to how deeply they automate their infrastructure. Employing automation to take charge of decision points changes the paradigm from the physical limitations of hardware to the unbounded design capabilities of software. Automation across the stack is vital in the NGDC where speed and innovation rule. Fast, accurate provisioning means higher profit margins and increased customer satisfaction.

At Cisco Live in June 2015, John Chambers gave his final keynote as CEO of the networking market leader. In doing so, he chose to deliver a very direct warning to the audience of more than 25,000 attendees, some of whom may be slow to react to the disruptive digital renaissance. Chambers forecasted that 40% of businesses in the room would not exist in a "meaningful way" in the next 10 years.

In the face of his prediction, Chambers emphasized the role IT must play in successfully navigating this transition, stating: "IT no longer enables your business strategy; it IS your business strategy." Bluntly connecting the dots between IT strategy and business outcomes, Chambers suggested that "if you can't deliver on fast IT, your company will fail."[9]

Trends enabling automated management

The software-defined data center

Across new cloud consumption models, the software-defined data center (SDDC) can discover, deploy, manage, consume, release, and monitor infrastructure with fine-tuned control. Agile business systems need this technology to eliminate silos. With this approach, infrastructure operations can shift from static to dynamic and fully configurable. SDDC technologies accelerate the adoption of NGDC cloud deployment models by allowing a convergence of workload-centric architectures and services into a single operating domain.

In the NGDC, the control plane and data plane are loosely coupled or abstracted, and the control plane — which is making decisions about where traffic is sent — is implemented in software. SPs can control, deploy, and reconfigure the infrastructure via software using open API control points. This can also be presented to enterprise customers that have the need and skills for programmatic control. When all infrastructure resources are abstracted into a higher-level control plane, they can be dynamically provisioned and managed. Resources are then granularly defined to match application and/or service needs.

Think of the control plane like a restaurant menu of options. You may order a complete meal or à la carte menu items. You choose what you want from the menu of available dishes (control plane) and tell your waiter. The data plane in this analogy is the kitchen, which cooks up your order as requested. You may not see how the chef is filling the order, but you appreciate a fresh and hot meal when it arrives at your table.

Taking this analogy a step further, consider new applications in restaurant automation that are improving service delivery and increasing revenue. At Chili's Grill & Bar, patrons can choose to self-order and self-pay via tabletop tablets. Chili's cites a jump in sales as customers no longer have to flag down wait staff to refill a drink or to order dessert. In addition to increasing revenues, the company sees this as an opportunity to create stronger connections and interactivity between the restaurant and its guests.[10]

Application programming interface (API)

Today's API ecosystem is diverse and expansive. Multiple tools are leveraged to achieve diverse business goals. Offering command-line interfaces (CLIs) can be insufficient; vendors are illustrating the value and benefit of code as an instrument for day-to-day operations. Scripting common and/or disparate infrastructure tasks with APIs dramatically reduces errors and increases productivity. It streamlines service delivery by offering efficient sharing of data and processes.

9 SolidFire blog. "Cisco Live! Takeaway: Fast IT = Faster Innovation = Faster Revenue." Available at http://www.solidfire.com/blog/cisco-live-takeaway-fast-it-faster-innovation-faster-revenue/

10 Computerworld. "Automation arrives at restaurants (but don't blame rising minimum wages)." Available at http://www.computerworld.com/article/2837810/automation-arrives-at- restaurants-but-dont-blame-rising-minimum-wages.html

In the NGDC, speed and agility are fueled by repeatable methods that validate and verify tasks. A consistent process provides an auditable record of activity and a gateway to fully automated management.

Employing software automation to take charge of decision points changes the paradigm from the physical limitations of hardware to the unbounded design capabilities of software.

Configuration management tools

In the last several years, the automation ecosystem has burgeoned with explosive use of configuration management (CM) tools to automate DevOps processes across the infrastructure. These tools help orchestrate elastic workloads at scale. Configuring and maintaining from a few virtual servers to thousands is easier because the infrastructure can be described and managed via code. A flexible, policy-driven approach to operations enables enterprise cloud agility. In the NGDC, deployment systems manage the full spectrum of service provider solutions via policies that govern deployment configurations and behavior.

Puppet was conceived as a cross-platform, open-source CM solution. In 2011, with investments from partners such as VMware, Google, and Cisco, Puppet Labs introduced its first solution for commercial use. Its declarative configuration language helps define and manage each step in the delivery process and infrastructure lifecycle, including provisioning, configuration, OS management, orchestration, and reporting and quality validation.

Chef offers a similar CM tool that is also based on Ruby. With Chef, a workstation is deployed to control the master server. Its configuration uses Git, an open-source distributed version control system. Chef Director of Patterns and Practices Colin Campbell explains the product strategy: "Treating infrastructure as code produces a workflow that effectively can use the same testing process used for other types of software."[11]

Software-defined networking

In conventional networks, operators configure functionality at a low level, device by device. With software-defined networking (SDN), administrators write high-level control programs and stipulate the behavior of an entire network because the control logic is separated from underlying physical routers and switches that forward traffic. Complex tasks for integrating network functions such as resource control, prioritization, and security are specified into a single SDN control framework. Paradoxically, SDN offers the ability to create configurations that have more sophisticated policies but are much easier to manage, maintain, and secure.

Announced at VMworld in 2013, VMware's NSX was an early entrant in SDN implementation. Born out of VMware's acquisition of Nicira in 2012, NSX is a network virtualization platform that reproduces the network model according to policies set in software, enabling any size network topology to be rapidly created and provisioned. Built around a controller cluster that manages the distribution of logical network functions (switches, routers, firewalls, VPN, security) into hypervisors, NSX allows virtual networking through any platform leveraging NSX APIs. This approach helps service providers use the physical network as a pool of capacity for on-demand consumption.

11 PC World. "Chef cooks up infrastructure testing tools." Available at Available at http://www.pcworld.com/article/2455060/chef-cooks-ups-infrastructure-testing-tools.html

VMware's vice president of product marketing for networking and security, Chris King, says businesses are demanding a change:

"The transformation of the network is inevitable. The only real question is: Why would you entertain an antiquated architecture when so much of your business's success relies on your ability to deliver more speed and efficiency? Why should agility require a forklift? Virtualizing your network is the next step forward on the path to the software-defined data center." [12]

Another example of this trend is Cisco's ACI framework. ACI is an integrated network architecture for dynamic workloads that is moving the needle in network automation. The ACI automated application-centric policy model promises to meet IT market needs with embedded security, centralized management, compliance, and the ability to quickly scale.

ACI is reported to be the industry's first solution to provide a dynamic, application-aware network policy that can reduce application deployment time from months to minutes.[13] Cisco reported in June 2015 that ACI user and leading security-solutions firm Symantec projected a five-year cumulative benefit of 441% ROI and reported 87% faster application development lifecycle as a result of deploying ACI.[14]

Storage APIs and interfaces

In storage, automated management is likely to be a combination of interfaces to manage the system directly (CLIs and UIs), integrations to orchestrate higher-level tools (SDKs, plug-ins, drivers), and storage APIs to automate workflows (adding, deleting, or reporting on usage). The current paradigm treats storage as an undifferentiated resource in the infrastructure automation puzzle. Applications and workloads in a multi-tenant platform need more than just capacity-provisioned storage; they need performance profiles driven by business requirements. The options are limited with legacy storage architecture.

Storage automation in the NGDC changes the capacity-provisioned paradigm to an automated, programmable, repeatable, policy-driven environment that drives SP innovation. It eliminates device-and endpoint-centric administration and fully enables multi-tenant data management and protection. Web-scale principles of automation, QoS, and API-based access help orchestrate all workloads and their integration with NGDC delivery solutions. These attributes enable service providers to deliver higher value propositions and level the playing field.

Storage automation in the NGDC changes the capacity-provisioned paradigm to an automated, programmable, repeatable, policy-driven environment that drives SP innovation.

12 Network World. "Virtualize It! Build a data center that is defined by software not by hardware." Available at: http://www.networkworld.com/article/2359521/tech-debates/what-s-thebest-approach-to-building-next-generation-data-center-networks.html

13 Network World. "Cisco's ACI goes beyond SDN with policy-based automation." Available at http://www.networkworld.com/article/2359521/tech-debates/what-s-the-best-approach-tobuilding-next-generation-data-center-networks.html

14 Cisco. "Cisco Extends SDN Leadership; Delivers Customer Choice for Data Center Programmability and Automation; Enhnaces Software and Expands Nexus Portfolio." Available at http://newsroom.cisco.com/press-release-content?articleId=1648780

Chapter 7

Data Assurance

In the legacy data center, infrastructure was customized with specialty or dedicated hardware to ensure a failure would not negatively impact specific customer applications. The applications were saved, and the price was paid in over-engineering and over-provisioning.

Engineering NGDC data assurance can be likened to designing buildings for earthquakes; there is no building that is guaranteed to withstand all earthquake activity, but earthquake-resistant structural design exponentially lessens the probability of disaster. In the NGDC, a resilient and secure infrastructure is the expectation, and failure is managed because the self-healing architecture is designed to tolerate it. From component failure to failures in managing customer dynamics, NGDC architects plan for failure while mitigating its likelihood of occurrence.

Trends enabling data assurance

Dynamic CPU and VM allocation

VMware's DRS (Distributed Resource Scheduler) is a utility that dynamically allocates and balances computing capacity and VM placement with pooled resources from multiple ESX server hosts. Using automatically programmed, predefined priorities, hypervisors can redistribute VMs from areas of contention to other areas without workload disruption. SPs are using DRS to simplify provisioning of applications and get higher utilization by optimizing resource allocation.

The business benefits of utilities such as DRS include simplifying day-to-day IT operations, as staff are less affected by localized events in their environment. When loads on individual VMs change, automatic resource relocation and optimization reduces the need for administrators to respond. Time is dedicated to higher-level tasks, and risk of data loss or service disruption is avoided.

In the NGDC, a resilient and secure infrastructure is the expectation, and failure is managed because the self-healing architecture is designed to tolerate it.

Availability zones

The concept of availability zones for data assurance in the cloud was popularized by AWS. Availability zones protect your customer's applications from the high cost of failure. In open-source platforms such as OpenStack, availability zones are used to arrange compute hosts into logical groups, providing a form of physical redundancy and isolation from other availability zones. Customers can provision resources by specifying which availability zone they want their instance to be built in, allowing application resources to be spread across disparate machines for high availability (HA) in the event of a failure.

Multi-pathing
Similar to a utility power grid that provides multiple paths to a single location, transparent multi-pathing allows data to pass via alternate routes. This removes the likelihood — and risk — of a single point of network failure, whether it's component or environmental. In addition to enabling redundancy, NGDC technologies like multi-path routing are improving network performance because more load-balanced traffic can be distributed through more routes. Network architects are using multi-pathing to gain higher throughput, reliability, and performance in the network.

Dynamic firewalling
Another networking trend in NGDC deployment is dynamic firewalling to segment-specific instances or groups of instances. Security administrators programmatically apply firewall policy to a customer workload or group of similar workload profiles (for example, simple web server) instead of manually touching each device in the chain, which can be a significant task in large infrastructure deployments. Automating with dynamic firewalling enables data assurance by reducing the likelihood of human error and security breaches.

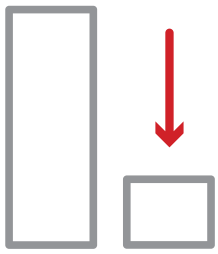
In the NGDC, service providers are not administering their customers' storage, they are automating it.

Replication-based HA in storage
Storage innovations have evolved in the last several years to ensure high availability, security, and data protection in the cloud. Storage vendors are replacing RAID-based data protection with replication-based high availability. A 30-year-old technology, RAID has not kept pace with the scale of cloud. Recovering a large disk (even SSD) from a failure with a traditional RAID-based storage design can take hours or days, risking greater faults and customer data loss.

Legacy storage companies built their management systems in the construct of using storage administrators, and, as a result, created complex, feature-rich administration tools. Data assurance at service provider scale is not realized through armies of storage administrators but through effective configuration management tools built for the cloud. In the NGDC, service providers are not administering their customers' storage, they are automating it.

Replication-based HA is a post-RAID data protection scheme based on a distributed-replication algorithm. This NGDC architecture protects against both drive and node failure. It allows the storage system to absorb concurrent failures across all levels of the storage solution.

In a failure event, each drive in the system redistributes a small percentage of its data in parallel to the free space on all other remaining drives, and no operator intervention is required. The self-healing properties of an automated, replication-based architecture enables fast rebuilds with minimal performance impact because parity reads are not required. With this architecture, faults are isolated, and QoS settings remain enforced.



Chapter 8

Global Efficiencies

Service providers are consolidating and converging the people, processes, and software that fueled their legacy hosting infrastructures. Convergence enables SPs to take charge of perceived customer activity, including seasonal and typical usage, while controlling oversubscribed resources. Like airline carriers that practice overbooking, SPs can programmatically govern thresholds to meet expectations, maximize utilization, and control costs. Fewer systems and fewer interfaces between components are pillars of NGDC global efficiencies.

Trends enabling global efficiencies

Converged networking

The cost/benefit of converged networking is gaining more attention from SP infrastructure architects seeking optimization of services at massive scale. Rapid increases in Ethernet speeds coupled with network-based QoS technologies are facilitating the convergence of previously separated application and storage traffic. Further efficiencies are gained by eliminating expensive Fibre Channel host bus adapters (HBAs) and per-port costs. In the NGDC, a single-fabric technology reduces infrastructure and power requirements, optimizes performance and utilization, and simplifies operations.

Native in-line storage efficiencies

A converged approach is also applied in NGDC storage architecture. Consolidation of storage into a simplified, scalable platform is improving utilization rates and substantially lowering capex.

The impact of all-flash storage on data center efficiencies cannot be overstated. All-flash architecture reduces storage equipment costs, carbon footprint, and power consumption. With NGDC data reduction techniques, devices can store data using fewer bits, requiring less space. Data reduction, in-line deduplication, compression, and data thin provisioning enable environmental and economic benefits at scale without performance tradeoffs.

Fewer systems and fewer interfaces between components are pillars of NGDC global efficiencies.

Conclusion

It is an understatement to describe today's cloud landscape as competitive. Service providers must continually improve TCO, shorten ROI, meet SLAs, and deliver non-generic solutions. At the same time, they're focused on creating a customer environment that ensures production-grade support for a plethoric flow of new data and applications. Hardware-defined legacy infrastructure impedes SPs from achieving these goals — and understandably so. The traditional data center was not built for the cloud.

NGDC cloud economics is changing the paradigm. Successful service providers are not chasing the curve; they are defining it. Service delivery models such as IaaS, private cloud, and SDDC will continue to disrupt the traditional infrastructure approach to hosting and cloud services. By applying a software-defined everything (SDx) discipline in a once hardware-centric business, service providers are creating opportunities to maximize resources and satisfy 21st century enterprise needs.

The traditional data center was not built for the cloud.