White Paper

Using Flash to Optimize Cloud Software Performance, Agility, and Density

Sponsored by: Kaminario

Eric Burgener   Amy Konary
June 2016

IDC OPINION

The software market is being forever changed by the advent of cloud services and as-a-service delivery models. Application software vendors are offering cloud-based consumption models for their software to meet customer demand for IT agility, "pay as you go" pricing, and improved efficiencies in managing the overall IT infrastructure. Cloud software services also offer significant advantages for software vendors; they reduce packaging, distribution, and support costs and make it easier for vendors to move their customers to the latest releases for improved functionality and reliability.

Because of their consistently high performance and ability to support dynamic environments, all-flash arrays (AFAs) are routinely used in service provider infrastructures. However, these arrays must support more than just flash performance. They must also meet the enterprise-class availability, reliability, manageability, and scalability requirements that customers of Web-based services demand. And as most cloud-based businesses look to automated operations to help keep ongoing administrative costs low, the all-flash array systems must also offer excellent integration support for the virtual infrastructures that form the foundation of cloud-based services.

The Kaminario K2 AFA has been achieving considerable success in the cloud software market over the course of the past year. Identified by IDC as one of the top AFA solutions in the market, Kaminario K2 provides the industry's most comprehensive combination of both scale-up and scale-out architectures, supporting unprecedented configuration flexibility that provides compelling value in a dynamic and aggressive market (see *IDC MarketScape: Worldwide All-Flash Array 2015-2016 Vendor Assessment,* IDC #US40721815, December 2015).

SITUATION OVERVIEW

"Cloud software" is the service enablement of software products in all three primary software markets: applications as a service, system infrastructure software (SIS) as a service (which combines to form software as a service [SaaS]), and application development and deployment (AD&D) or platform as a service (PaaS). IDC expects that by 2019, $1 of every $4.59 spent on packaged software, and $3.44 spent on applications, will be consumed via the public cloud. The cloud delivery model accounts for almost half of projected growth in the software industry through 2019.

The market for cloud software reached $48.8 billion worldwide in IDC's most recently published forecast and is expected to grow to $112.8 billion by 2019, a five-year compound annual growth rate (CAGR) of 18.3%. Revenue associated with cloud software grew more than 10 times faster than revenue in the traditional software delivery market for 2013–2015, although that difference will drop after 2015 as growth in cloud begins to level out.

Key Tenets of SaaS Delivery Models

Cloud software services are based on a service composition and delivery model made up of a utility computing environment in which unrelated customers share a common application and infrastructure that is managed by an independent software vendor (ISV) or a third-party service provider and code or intellectual property (IP) of the service, which is typically owned by the cloud ISV. There are many emerging models for providers of this software code to leverage third-party infrastructure, business services, and other providers as hosting, selling, fulfillment, or support partners, and many new models are forming far beyond the comparatively well-understood direct versus tiered distribution models of packaged software. These new models provide customers with access to and consumption of software and application functionality built specifically for network delivery that is hosted, provisioned, and accessed by users over the Internet.

The cloud model goes well beyond prior online delivery approaches — combining efficient use of multitenant (shared) resources, radically simplified "solution" packaging, self-service provisioning, highly elastic and granular scaling, flexible pricing, and broad leverage of Internet-standard technologies — to make offerings dramatically easier and generally cheaper for customers to consume.

Creation and delivery of cloud services rest on the growing adoption of two key tenets — virtualization and multitenancy — built with technologies based on Internet standards using a business model and architecture of shared services. Cloud software services are by definition turnkey services, with application, presentation, data tiers, and all associated services in a single service that can be accessed and provisioned over the Internet.

Growth Drivers and Inhibitors

Not only does market momentum impact cloud spending through an increase in cloud maturity on the demand side, but large, incumbent ISVs are becoming more willing and more capable of selling their software as a service. In addition, companies that are not historically software providers are shifting their focuses from hardware to software-based cloud services to generate subscription revenue that compensates for shrinking margins on commoditized products. Development of new SaaS applications also falls under the category of market momentum, which will especially affect less mature secondary markets. In some of the markets, customers are willing to buy SaaS, but adoption is limited simply by a lack of available solutions. As more SaaS applications are developed and made commercially available, adoption will accelerate. However, while there is a lot of room to grow in terms of market momentum, the SaaS market is already maturing. And IDC does not expect any future years to match the growth rates the industry saw in 2014. This means that the market momentum, while an accelerator, will have a decreasing impact over time.

A key driver of SaaS applications has been the ability of business units (e.g., sales, marketing, corporate communications, and customer support) to directly source their own IT solutions. Many line-of-business (LOB) executives feel constrained by IT in terms of the ability to implement new business applications, and adopting SaaS applications can enable them to more quickly purchase solutions directly from external providers. The shift of purchasing power and control over IT sourcing to

the LOB, along with the LOB preference for public cloud services, is a major driver of demand for SaaS application. At the same time, IT buyers have appreciated the ability to adopt
cloud software strategically within their organizations, offloading certain activities and enabling the organizations to avoid capital expenditures and reduce internal staffing requirements.

IDC's research shows that, by far, the greatest inhibitors to public cloud services adoption are the perceptions of security/privacy concerns and performance/availability concerns (see Figure 1).
If vendors fail to demonstrate steady improvements in addressing these concerns — or worse, if there are a significant number of high-visibility failures — these inhibitors could seriously retard growth rates.

Service reliability is another instance where the cloud may perform better than internal IT, but the perception remains that downtime and response time can be issues when moving to the cloud. As reliability increases and this perception diminishes, this will inhibit adoption less in future years.

FIGURE 1

Top Cloud Software Inhibitors

Q   *Which of the following best describe your organization's main concerns about cloud and are important inhibitors for your organization in considering public cloud services or private cloud technologies?*
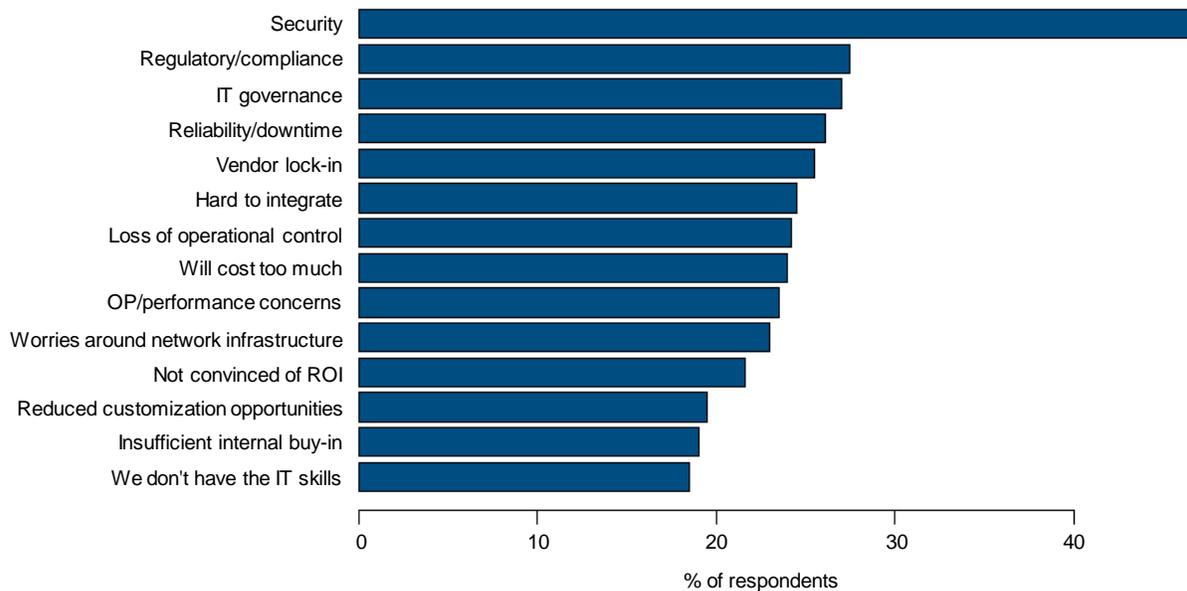
n = 6,159 worldwide respondents

Note: Data is weighted by GDP and company size.

Source: IDC's *CloudView Survey,* January 2016

Current and Future Challenges Faced by SaaS Providers

Cloud service providers achieve economies of scale via shared, standard services built for a market, not for a specific customer. The shared service model offers customers and suppliers both enormous operating efficiencies and upgrade/enhancement velocity. As a result, this is a fundamental attribute that differentiates the cloud service model from many traditional customer-unique outsourced or hosted offerings. A key success characteristic for a cloud service provider is the ability to instrument a range of built-in configuration options that meet the most common set of customer needs for personalization and integration.

In addition, one of the most obvious user benefits of the cloud service model is that it is presented as an all-in, "turnkey" solution: The customer can access the offering without the need to own, manage, or understand any underlying resources required to support the offering. As a result, the cloud service provider bears the burden of selecting and maintaining a wide range of scalable resources to provide the best combinations of technology, service, and economics for customers.

Rapid and flexible expansion (and contraction) of service usage is one of the major benefits of cloud services for users, and cloud services' dynamic provisioning (and deprovisioning) capability — including the ability to access resources in finer-grained increments — also dramatically reduces the need for costly overprovisioning. In addition, this characteristic substantially reduces the burden on users to come up with demand plans for resources (CPU, storage, network bandwidth, and support staff), which is a major challenge for organizations and typically drives companies to greatly overprovision IT.

Customer expectations are high for cloud service providers that must be able to provide elastic scaling capabilities on a very detailed level such as the number of server instances, storage capacity, and data transfer volume. Customers want services not only scaled to need but also priced to use, whether that's in proportion to usage, the number of users, transactions, screen views, or some other consumption metric. As a convenience to some customers, providers may mask this pricing granularity with long-term, fixed price agreements, but suppliers must design their offerings so that they have the capability to do fine-grained metering and pricing for customers that need these services.

Infrastructure Requirements for Cloud Storage

Vendors delivering cloud software to their customers have a very specific set of infrastructure requirements, and this has implications for the underlying storage. Competition is fierce in this space, and cost is a key purchase criterion for end users. As a result, vendors need scalable infrastructure that can cost effectively and predictably deliver enterprise-class services. The ability to handle dense multitenancy is a must — more highly utilized infrastructure delivers better price/performance for vendors, giving them the freedom to pass those savings on to their customers. Cloud environments are characterized by unpredictable demand, and flash-based storage can deliver predictable performance across highly variable loads, giving vendors the ability to reliably execute against different service tiers. High availability is required, and vendors need the tools in the storage infrastructure — dual parity–plus RAID implementations, snapshots, replication, and hot pluggable components in arrays with no single point of failure — to transparently ride through failures while guaranteeing data integrity. And since cloud environments are based on virtual infrastructure, the storage tier must integrate well with that infrastructure, offering support for various APIs that enable array services to be utilized with automated operations.

Flash-Optimized Cloud Storage: The Kaminario K2

Kaminario is a Needham, Massachusetts-based AFA vendor that sells an enterprise-class storage platform targeted for use in medium-sized to large enterprises. Kaminario's solution, the Kaminario K2, is broadly used as a general-purpose primary storage platform and supports a comprehensive array of capabilities that enable very efficient and dense mixed workload consolidation. Primarily sold through indirect channels, the Kaminario K2 has achieved success in a variety of vertical markets, including financial services, healthcare, retail, manufacturing, transportation, and utilities. In the wake of Kaminario's 5.0 release of the platform in the summer of 2015, the company's momentum in the cloud space began to grow at a considerable pace, and this is now one of the company's strongest vertical markets. Today, Kaminario counts major SaaS businesses such as Nuance Communications, Payoneer, Emergency Communications Network, and Clarizen among its customer base.

The Kaminario K2 is particularly well suited for cloud-based deployments. Built around an architecture that is a blend of both scale-up and scale-out designs, the system offers a low entry price point and cost-effective, nondisruptive scalability across both performance and capacity that allow a single system to grow to over 1 million IOPS of performance and over a petabyte of effective storage capacity. System building blocks are called K-Blocks, and each K-Block is fully redundant with dual, active/active controllers and up to 48 solid state disks (SSDs). Customers can start with a single, partially populated K-Block and easily expand flash capacity by just adding drives, or they can add additional K-Blocks to scale performance as well without interrupting application services.

Kaminario stands out for the efficiency of its design in ways that matter to vendors building cloud-based infrastructures. The scale-out nature of its architecture enables nondisruptive, balanced system expansion over a wide range, which is critical for dynamic businesses in high-growth environments. System metadata and user data are distributed across all resources in the system, a data layout that ensures extremely consistent submillisecond response even in the face of widely varying and unpredictable workloads. This type of data distribution also supports very even wear leveling across the flash storage, improving media endurance and reliability. Its use of adaptive rather than fixed block sizes minimizes the amount of work the system must perform to handle any given workload, supporting increased efficiencies of resource utilization and more predictable scalability over time. The K-RAID data protection algorithm is flash optimized to require fewer writes (thus improving media endurance) and provides better protection than dual-parity RAID (RAID 6), with only a 12.5% capacity overhead (very cost efficient compared with conventional RAID 6 implementations). And the system offers inline data reduction that lowers the effective dollar-per-gigabyte cost of flash capacity – dropping it to under $1/GB – while enabling a fully configured system to support up to 1.5PB of effective storage capacity.

AFAs deliver much better and more consistent performance than HDD-based systems, and the K2 is no different. With flash performance, time spent tuning storage drops to zero. Not only is production I/O faster, but other common operations such as storage provisioning take seconds. The K2 automatically handles many of the arcane tasks associated with storage provisioning, such as LUN creation and RAID definition, once an administrator defines how much capacity is required. System performance can be upgraded by adding more K-Blocks. Kaminario allows mixing of older- and newer-generation technologies in the same system, and customers can harness newer, faster, and more efficient controller and disk technologies without any downtime – no forklift upgrade or data migration is required. Once new resources are added (or subtracted), the system redistributes the data and metadata automatically, adjusting system performance to its new level without any operator intervention. The speed and ease of use of the K2 give time back to busy administrators, particularly in dynamic environments where virtual machines are often being spun up and down over time.

The ability to support dense multitenancy is a critical consideration in cloud software environments. Systems that can continue to meet performance, availability, and reliability requirements in growing businesses while taking up minimal real estate conserve power, cooling, and floor space. The K2 offers a number of features that support multitenancy. Data services include dual parity-plus RAID (K-RAID); read-only and read/write snapshots; AES-256 bit encryption, replication, and inline compression and deduplication; and other storage efficiency technologies such as thin provisioning, space-efficient snapshots, and delta differential-based replication. The K2's use of global adaptive block sizes lends an ability to concurrently handle disparate workloads – OLTP, OLAP, VDI, and virtual server infrastructure (VSI) – while continuing to deliver consistent performance, even at very high-throughput levels. Data services can be applied selectively, providing additional flexibility in accommodating a variety of workloads with differing requirements on the same storage infrastructure. And the K2 packs a lot of punch in a small space: 1 million IOPS and up to 1.5PB of effective storage capacity (assuming a conservative overall data reduction ratio of 4:1) in just 26U of rack space.

The K2 leverages hardware redundancy to ride through failures. Hardware components such as controllers, SSDs, power supplies, and fans are hot pluggable, allowing them to be easily replaced upon failure. Controller and disk firmware can be upgraded nondisruptively, providing more freedom to schedule maintenance operations when it's convenient for administrators. HealthShield, Kaminario's telemetrics implementation, uses predictive analytics to proactively resolve a number of potential issues before they can cause an acute problem that requires operator involvement. Kaminario tracks the uptime of its systems across its entire installed base using HealthShield, and the empirical data shows "five-nines plus" availability in production usage across Kaminario's entire installed base.

The K2 is very adept at supporting the virtual server infrastructure that serves as the foundation for most cloud-based environments. Management options include a secure (SSH-based), scriptable command-line interface, an HTTP/HTTPs GUI, a RESTful API, and SNMP, with a road map for SMI-S support in 2017. VMware integration was a focus for Kaminario early on, and the product supports key vSphere APIs, such as VAAI, VASA, VADP, and SRM as well as a vCenter management plug-in. VVOLs support is scheduled for later in 2016. Microsoft Volume Shadow Copy Service (VSS) is also supported. This extensive set of APIs allows the K2 to be easily integrated into automated workflows that increase operator productivity and the reliability of administrative operations. Host connection options for its block-based storage include 8Gb and 16Gb Fibre Channel (FC) as well as 10Gb Ethernet (GbE). Each K-Block can support up to 3.2GBps of bandwidth, or a fully configured four-node system can support 12.8GBps.

Kaminario leverages its architecture to deliver predictability to customers around the four key areas – cost efficiency, easy scalability, performance, and availability – of storage most important to cloud service providers of storage. Today, the company offers an Effective Storage Capacity Guarantee that ensures customers see at least a 3.3:1 data reduction ratio across their combined workloads. Kaminario has raised the guarantee ratio over the past year to reflect advancements in data reduction technology. Note that across Kaminario's installed base (tracked through the company's HealthShield analytics platform), most Kaminario customers enjoy an average data reduction ratio of greater than 4:1 though. Any inability to meet this guarantee is rectified by the provision of additional Kaminario hardware at no charge to the customers. The company plans to roll out similar guarantee programs aligning with scalability, performance, and availability in the future.

CONCLUSION

The migration to cloud will continue to occur. In IDC's view, hybrid cloud will ultimately become the primary IT infrastructure for most businesses, although this journey will take the better part of the next decade. On-premise infrastructure will not go away, but the agility and economics of cloud infrastructure provide the right complement to the flexibility that most businesses will require for most applications. Cloud-ready storage solutions will become increasingly popular as hybrid cloud becomes the mainstream deployment model. Application vendors are moving en masse to make their software available as cloud-based services to meet customers' evolving requirements for easier access, better flexibility, and more cost-effective pricing options with a lower touch.

Flash media will play a large part in the storage infrastructure that supports cloud-based software. As competitive pressures continue to drive prices down, systems that make flash performance available in an enterprise-class package most efficiently will have an edge. Market requirements have changed, and a much higher bar for performance, availability, reliability, manageability, and scalability must be met in cloud-based infrastructure. Storage systems targeted for use in these environments must efficiently support dense multitenancy, enable easy integration into automated operations strategies, and offer a range of storage efficiency technologies that increase storage density, reducing energy and floor space consumption and effective dollar-per-gigabyte capacity costs.

The Kaminario K2 AFA is a system with a strong legacy of successfully supporting SaaS infrastructures, offering the feature set necessary to provide cost-effective storage services to dynamic and growing cloud-based software businesses. With its reputation for high and consistent performance across a wide scalability range, the K2 is a good fit for these types of demanding environments. Kaminario K2 was one of the first AFAs that began to be broadly deployed for mixed workload consolidation, and Kaminario's experience in this respect has driven the K2's standout multitenant support capabilities. Kaminario monitors customer satisfaction closely in its installed base, and this attention to detail has paid off in allowing it to sustain high revenue growth rates in recent years: 80% of Kaminario customers have increased their spend with Kaminario by 2–2.5 times within 12 months of their original purchase. That kind of repurchase rate is enjoyed only by vendors that understand their customers' requirements and consistently meet them over time.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com