

Maginatics: Beyond NAS limits with the performance and agility of the cloud

Prepared for: Maginatics
August 2014

TABLE OF CONTENTS

TABLE OF CONTENTS	1
EXECUTIVE SUMMARY	2
CHALLENGES	4
Users, devices and IT	4
Traditional storage models are broken	5
Performance is not where it is supposed to be	6
THE SOLUTION	7
Leveraging the cloud	7
Deploying the performance	8
Introducing Maginatics Cloud Storage Platform	9
BOTTOM LINE	11
JUKU	13
Why Juku	13
Author	13

EXECUTIVE SUMMARY

Both business and consumer technologies have drastically evolved, and so have the ways we access, produce and consume data. Access patterns and workloads have changed both from the end user point of view as well as at the storage infrastructure level. Traditional NAS boxes are no longer adequate to sustain performance requirements at scale of distributed environments due to a variety of bottlenecks. Some of the more visible ones include archaic chatty protocols designed to work on local area networks (LANs) and lack of smart caching capabilities at the client level. Common Internet File System (CIFS), for example, is a chatty network protocol used to access files on an enterprise network. Accessing files over a LAN with latencies of less than 10ms does not typically incur productivity issues. However, if these files are accessed over a WAN, where latencies are frequently over 20ms, productivity decreases due to longer wait times. In a CIFS communication there is a lot of traffic generated between the client and server with acknowledgements going back and forth. These are only some of the most visible constraints. As we dig deeper, it is easy to find that all the design principles of classic NAS architectures are neither effective nor agile when simultaneously coping with a multitude of devices and increasing distances between them.

In fact, the concept of “distributed enterprise” is outdated and is evolving into the “dispersed enterprise.” A “dispersed enterprise” involves a wider range of devices and a different type of end user, who wants to seamlessly access data anytime and everywhere without performance impact. Data must be always and instantly available, ready to be consumed or modified and shared with the rest of the team. Even though the number of users is not necessarily increasing much, they want access to data far more frequently and they demand a certain level of device parity in terms of features and, above all, performance. This actually means a more demanding audience asking for improved throughput and latencies. Now, for example, a single mobile device can download data at an astonishing speed. Today’s users require access to terabytes of information at the click of a button no matter where they are

located or what device they are using. Is your current infrastructure flexible and powerful enough to properly respond to many of these requests?

Maginatrics has developed a brilliant next generation cloud-based solution which is suitable for covering all these needs (and many more) for organizations of any size and for a wide range of file-based applications without having to rewrite those applications. Modern design and an object storage backend enable unmatched performance, while integration with standard infrastructure services (such as Microsoft Active Director) and complete IT control over data integrity and access control grant the end user maximum flexibility with a secure migration path from legacy NAS-based architectures.

CHALLENGES

Users, devices and IT

Each user now has access to several different devices, ranging from smartphones and tablets to notebooks and PCs. Users demand a seamless user experience, high performance and ease of use. The “user” can also be a machine such as an application server or even a VM requiring access to a highly reliable shared storage system.

All devices are increasing in size (now phones have displays of up to 5.5”), they have plenty of RAM and CPU power, and they sport the latest networking technology (LTE can download files at more than 100MB/sec). Users are able to view, create, edit and share documents while on the move and they don’t want to wait for the content they need. They are well aware that poor performance is no longer a device issue.

Users can view, create, edit and share documents while on the move and they are not willing to wait for the content they need. They know that it is no longer a device issue.

Another critical requirement for seamless workflow is device parity. The workflow must continue unabated when the user reaches his or her workplace, so traditional PCs and Macs require equally rapid access to the same data store. Only then can the user’s experience be considered seamless and productivity be maximized.

IT managers have to rapidly provision storage, performance, availability and data protection where needed, while granting access to next generation mobile services for all users.

At the same time the IT department needs to keep data safe and secure, lower costs in terms of both total cost of acquisition (TCA) and total cost of ownership (TCO), and prepare everything to be audited and controlled. IT managers

must rapidly provision storage, performance, availability and data protection where needed, while granting access to next generation mobile services for all users in a secure manner.

This is getting increasingly difficult given exponential data growth and unpredictable user access and workload patterns in the context of a larger and more widespread enterprise.

Traditional storage models are broken

Total Cost of Acquisition (TCA) and Total Cost of Ownership (TCO) are two important key factors when we talk about storage infrastructures, but not the only ones.

Legacy storage approaches simply cannot cope with new business and user needs. The list of potential issues is long, and getting longer day after day:

- NAS cannot be consolidated in huge repositories for technical and practical reasons. For example, CIFS/NFS are not natively efficient over the WAN. So you have to maintain different NAS appliances for remote and branch offices (ROBOs).
- Limitations of single system scalability, as well as the need for forklift upgrades, can easily introduce migration costs as well as various performance and provisioning issues, not to mention service down-time.
- Traditional storage protocols are not designed for WAN access while mobile devices are not designed to work with NFS or CIFS at all. Furthermore, end-to-end encryption is typically not a native option.
- Classic NAS file systems have limits in the volume sizes and number of files per volume they can handle. This often leads to higher complexity and management costs.
- Backup and remote replication techniques add additional costs, complexity and rigidity to the whole infrastructure.

These are but a few examples but they can illustrate what many large organizations face when managing file services today.



Performance is not where it is supposed to be

NAS systems are very data-center centric: remote users rarely if ever receive the fast access times they need to be successful. In fact, in a very dispersed environment, storage infrastructure must support high performance for both local and dispersed users.

While local performance is relatively easy to achieve (it hinges only on the IOPS/throughput the NAS box can deliver), one cannot say the same for the distributed performance, especially when one must access the storage system from several devices dispersed all around the world with various access methods/patterns.

Even though the aggregate performance capacity of your traditional storage infrastructure is huge:

- it is difficult to load balance that capacity across endpoints;
- bottlenecks exist between the end-point devices and the storage system;
- the protocols used are neither efficient nor optimized;
- caching techniques, both at ROBO sites and at the client level, are not sufficiently smart.

To avoid performance issues with traditional NAS systems, IT managers are forced to over-provision resources, but at the cost of decreased efficiency and increased expense.

In a very distributed environment, you need both local and distributed performance.

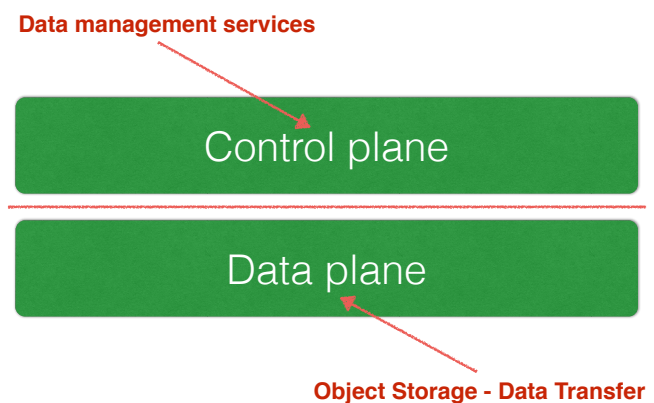
But that's not all. Even temporary or fast provisioning are not possible simply because NAS typically relies on hardware that must be acquired before it can be used. In most environments, lack of flexibility and agility are usually counted as poor indices of performance.

THE SOLUTION

There are many possible approaches to solving traditional NAS infrastructure issues, but most fail to provide a comprehensive answer for mobile devices, PCs, ROBOs and central sites. The risk is that a customer implements a new infrastructure which is not capable of managing all the data in a single namespace. As a consequence, over time, characteristics such as unlimited scalability, uncompromised performance and improved productivity will once again fall short of expectations and needs.

Leveraging the cloud

Cloud, in the form of an Object Storage platform, is the first step towards achieving separation between data management and services offered to end users. As discussed below, splitting the control plane from the data plane is also a huge benefit when the organization is moving towards a software-defined strategy.



Whether the object store is provided by a public service (e.g., Amazon Web Services S3) or a private, on-premises implementation, object storage delivers characteristics that enable the creation of a single distributed global data repository with a virtually unlimited scalability. Simultaneously, object storage does not require traditional backup and data replication mechanisms for DR because these are all embedded in the basic object storage infrastructure, allowing great savings, better resiliency and superior availability without sacrificing data durability.

Deploying the performance

Modern object storage implementations, with their scale-out distributed architectures, have already demonstrated unprecedented scalability, agility and the power to serve files to millions of

Object storage has to be considered as only the backplane. In fact, you need a transparent homogenous access layer to take full advantage of it

concurrent users at the same time. Object storage can also deliver exceptional performance. But object storage must be considered as only the backplane, so an equally scalable front end is also needed to realize the performance potential of object storage. Specifically, you need a transparent homogenous access layer to take full advantage of the power of object storage and deliver consistently high performance to all kinds of devices in all scenarios.

The cloud access layer should implement a series of characteristics that guarantee consistently high performance:

- Smart caching mechanisms at multiple levels: both at the endpoint (the mobile device, laptop or server) and at the local and Wide Area Network layers (at branch offices or a service provider, for example). Smart caching is essential to minimizing data transfers and speeding up file retrievals.
- Data reduction techniques (e.g., deduplication and compression) to optimize communication, save bandwidth and minimize the amount of unique data the must be stored in the backend object store.
- Scale-out design of all central services. Authentication, metadata and data management must not be allowed to become a bottleneck.
- Native clients for all supported operating systems to enable the benefits of the system (failsafe security, performance acceleration, etc.) to extend “from cloud to endpoint” and to maximize advantages from optimized code in terms of performance and specific features.

Introducing Maginatics Cloud Storage Platform

The Maginatics Cloud Storage Platform (MCSP), provides the abstraction layer and a set of services that enable the creation of a next generation software-defined cloud-based storage infrastructure for modern dispersed environments. MCSP is designed from the ground up to serve mobile devices as well as PC, Macs, servers and virtual machines, with maximum performance regardless of the device or location.

The foundational component of MCSP is the Maginatics Virtual Filer (MVF): a set of virtual machines that can be deployed on premises or in the cloud depending on the customer's preference. The MVF delivers:

- data and metadata management,
- cryptographic key management,
- access control (by integrating with Microsoft Active Directory, for example),
- data integrity,
- cache coherency,
- deduplication services,
- management console.

Thanks to its innovative design, MCSP can start off very small but can quickly scale up to 14.4PB of storage while serving hundreds of thousands of users in a single domain.

The MVF is part of MCSP's control plane and in fact does not store data - only metadata and related configuration information. Data is directly accessed from the object store by each endpoint device over its own, separate "data plane" which is completely split from the control plane for both performance and scalability reasons. Thanks to this innovative design, MCSP users can start off small but quickly scale up to 14.4PB of storage while serving multiple endpoints and users in a single namespace.

As noted earlier, the other key component of MCSP aside from the MVF is a set of native endpoint agents (MagFS Agents) that can be deployed on mobile devices, PCs, Macs, servers or VMs running any supported operating system, which includes Windows, MacOS

X, Linux, iOS and Android. There is also a Web client for devices running legacy or less popular OSs. These agents play an active role in the process by accomplishing much of the “heavy lifting” in the MCSP system, including WAN optimization (multi-tier, adaptive local caching, compression, pre-fetching and a host of other performance acceleration techniques), encryption/decryption, file segmentation and re-assembly and myriad other optimization tasks —all aimed at maximizing performance, ensuring failsafe data security and improving user experience.

The MCSP architecture avoids potential bottlenecks, distributes computing and caching resources across the entire network and quickly scales when and where required. All these benefits will be realized by the end user because of the superior overall usability and performance when compared to traditional NAS-based infrastructures.

BOTTOM LINE

Today, when it comes to scalability and performance, especially if you think about unstructured data, adequate solutions (let alone optimal ones) are very hard to find. IT departments struggle to match evolving user needs with enterprise necessities and budget constraints. In a few words, legacy approaches are no longer sufficient to respond to problems faced in this era of exponential data growth and dispersed, instant access needs.

We, at Juku, strongly believe in object storage and modern approaches like the one proposed by Maginatics. You too should ask yourself if it's time to look at something different for your storage infrastructure:

- How much data are you managing today? Is growth accelerating?
- Do you have performance or latency issues? Especially on remote endpoints?
- Is the concept of the “dispersed enterprise” beginning to characterize your organization?
- Do you need to retain unstructured data for longer and longer periods of times?
- How much do you spend to back up your files? Are backup windows long enough? How much time do you spend retrieving files or data volumes?
- Do you replicate data offsite? How many copies do you maintain? How much does it cost to maintain your DR strategy?
- How many data centers and NAS boxes do you currently operate? How would you characterize their overall efficiency in the context of your organization, its workflows and workforce?
- Can your end users access data from anywhere, at anytime and from any device? Does your infrastructure ensure that such access is both failsafe-secure and highly efficient?
- Do you have storage provisioning issues? How long does it take you to fulfill your users' requests for one more terabyte of storage capacity?
- is your storage infrastructure truly reliable?

While this list of questions is long, it represents but the tip of the iceberg. The cost of legacy NAS infrastructure is too high, and the performance too low, for this post-Local Area Network era.

JUKU

Why Juku

Jukus are Japanese specialized cram schools and our philosophy is the same. Not to replace the traditional information channels, but to help those who make decisions for their IT environments, to inform and discuss the technological side that we know better: IT infrastructure virtualization, cloud computing and storage.

Unlike the past, today those who live in IT should look around themselves: things are changing rapidly and there is the need to stay informed, learn quickly and to support important decisions, but how? Through our support, our ideas, the result of our daily interaction that we have globally on the web and social networking with vendors, analysts, bloggers, journalists and consultants. But our work doesn't stop there, the comparison and the search is global, but the sharing and application of our ideas must be local and that is where our daily experience, with companies rooted in local areas, becomes essential to provide a sincere and helpful vision. That's why we have chosen: "think global, act local" as a payoff for Juku.

Author



Enrico Signoretti, consultant, trusted advisor and passionate blogger (not necessarily in that order). Having immersed into IT environments for over 20 years, his career began with Assembler in the second half of the 80's before moving on to UNIX platforms (but always with the Mac at heart) until now when he joined the "Cloudland". During these years his job has changed from deep technical roles to management and customer relationship management. In 2012 he founded Juku consulting SRL, a new consultancy and advisory firm highly focused on supporting end users, vendors and third parties in the development of their IT infrastructure strategies. He is constantly keeping an eye on how market evolves and continuously looking for new ideas and innovative solutions. You can find Enrico's social profiles here: <http://about.me/esignoretti>

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources Juku Consulting srl (Juku) considers to be reliable but is not warranted by Juku. This publication may contain opinions of Juku, which are subject to change from time to time. This publication is covered by [Creative Commons License \(CC BY 4.0\)](#): Licensees may cite, copy, distribute, display and perform the work and make derivative works based on this paper only if Enrico Signoretti and Juku consulting are credited. The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors. Juku consulting srl has a consulting relationship with Maginatics. This paper was commissioned by Maginatics. No employees at the firm hold any equity positions with Maginatics. Should you have any questions, please contact Juku consulting srl (info@juku.it - <http://jukuconsulting.com>).