

Caringo Swarm 7: beyond the limits of traditional storage. A new private cloud foundation for storage needs at scale

Prepared for: Caringo
May 2014

TABLE OF CONTENTS

TABLE OF CONTENTS	1
EXECUTIVE SUMMARY	2
UNDERSTANDING THE CHALLENGES	3
Design limitations of legacy storage	3
Next generation architectures	3
STORAGE FOR THE CLOUD ERA	4
Why deploy Object Storage	4
The power of gateways	4
HOW CARINGO SWARM 7 HAS IMPLEMENTED IT	6
A description of Caringo Swarm 7	6
The power of “Scalers”	7
BOTTOM LINE	8
JUKU	9
Why Juku	9
Author	9

EXECUTIVE SUMMARY

Traditional storage infrastructures are having a hard time keeping pace with data growth and next generation workloads. And the outlook is even worse!

Everything is digital now, your next phone will take bigger pictures than current one, users create richer content for their documents; your company stores everything and wants to keep every single bit of it! At the same time data accessibility is another issue: anytime, anywhere and from any device is not simply possible, or too costly, if you don't have the right kind of infrastructure.

Facing next generation business needs with legacy infrastructures doesn't make sense: Web applications, Big Data, Cloud, long term archiving and so on, are all high demanding applications which require highly automated, scalable, available, reliable, durable and open infrastructures. Every single service disruption, even when planned, has a tremendous impact on business: five nines, or more, of availability is what users usually expect from IT now.

New high challenging issues will become more and more common in the near future: more and more data awaits around the corner (think about machine-to-machine communication, internet of things, sensor recordings) and we will have to transform stored data into a valuable source of information (e.g. Dark data, Big data).

Legacy storage is simply not designed to cope with all of that. On the contrary, next generation object storage solutions, like Caringo Swarm 7, are the right choice to overcome the limits of traditional infrastructures and to implement a new platform capable of serving different storage needs at any scale.

Caringo Swarm 7, with all its characteristics, a series of integrated gateways (CloudScaler, BlockScaler and FileScaler) and a strong ecosystem of partners, is a compelling solution on which to build a next generation unified storage platform that can be considered a key element of a mature cloud strategy.

UNDERSTANDING THE CHALLENGES

Design limitations of legacy storage

Legacy storage was not designed for today's needs. Its basic concepts were devised in the last century, when nobody was aware about what was going to happen in terms of the quantity of stored data, diversification of access, workloads, and so on. Unstructured data growth is dramatic and its forecasts increment year after year.

Poor scalability, scarce manageability, lack of automation and RAID are showing all their limits. Especially now, that we talk in terms of Petabytes, all these constraints drive up costs



while we should be shrinking them to maintain infrastructure sustainability at a reasonable level.

Adding to the list of issues, traditional data protection mechanisms like backups and data replication become harder to manage, sometimes impossible, and their effective value is questionable.

Next generation architectures

The storage industry is trying to move on and modern infrastructures are adopting new architectures:

- Scale-up is slowly passing away in favor of scale-out,
- Software is winning over hardware,
- New data protection schemes are displacing traditional RAID,
- Automation mechanisms are aimed at placing data where they are effectively needed.

But it's not always enough. Business is asking for “always-on” infrastructures, next generation services, fast provisioning, agility and flexibility. This is why cloud computing is becoming more and more relevant for enterprises.

STORAGE FOR THE CLOUD ERA

Designing a private cloud, without thinking about a modern storage as well, risks producing a poorly balanced architecture and, consequently, diminishing the advantages of the cloud computing approach.

Why deploy Object Storage

Organizations interested in implementing a successful cloud computing infrastructure should think about Object Storage as a fundamental piece of their strategy. In fact object storage, when properly implemented, is an excellent storage system with state of the art capabilities:

- Automation and management: policy based storage;
- Availability and reliability: distributed storage system ;
- Scalability: scale-out architecture;
- Next generation data protection: erasure coding and/or multiple copies of objects;
- Efficiency: power management and balanced data placement;
- Meta data: standard and custom tags to describe objects;
- Security: encryption, hash fingerprints, WORM capabilities and so on;
- Always-on: no forklift upgrades, no migrations, no downtime.

Object storage characteristics combine all together to build a strong enabling horizontal platform.

The power of gateways

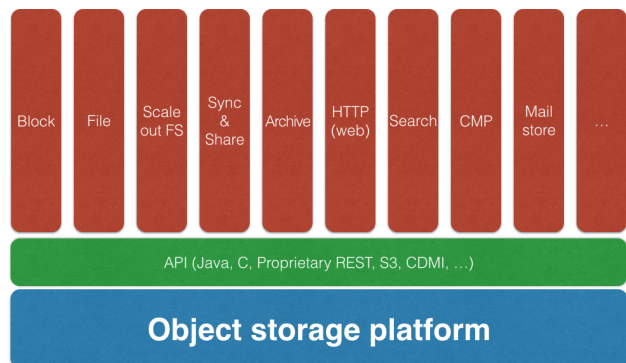
Integrating the Object Storage system with gateways and applications allow the delivery of next generation cloud storage services like:

- Files (with no more backups, ROBO, worldwide consolidation);
- Sync&share for mobile devices and PCs;
- Block devices (e.g. iSCSI volumes);

- Active archiving;
- App and Web storage;
- Dark Data and Big Data Analytics (metadata indexing and Hadoop);

ROI of these kind of services is extremely positive while the overall TCO of Object Storage is very low when compared to legacy solutions, especially with the growth of the number of stored objects and space.

Compared to legacy unified storage systems, which are usually proposed with only two types of protocols (Block and File), the object storage back-end allows to provide many more options ranging from classic protocols to a next generation class of services.




HOW CARINGO SWARM 7 HAS IMPLEMENTED IT

Caringo, a leading provider of object storage solutions, has just announced Swarm 7, the latest version of its flagship product. It has the characteristics I have already introduced in the previous pages and, thanks to its modern architecture, it can be adopted by organizations of any size to build private cloud storage infrastructures.



A description of Caringo Swarm 7

Caringo Swarm 7 is a distributed scale-out object storage system, which supports APIs, block and file access. It's a software solution that can run on any linux compatible x86 hardware (or VMs), and each node can be different in size or performance. New nodes can be seamlessly added (or upgraded) in an existing cluster to provide more space and performance when needed. All operations are performed

 Caringo Swarm™ without service interruption and the system automatically rebalances itself in relationship to the available resources.

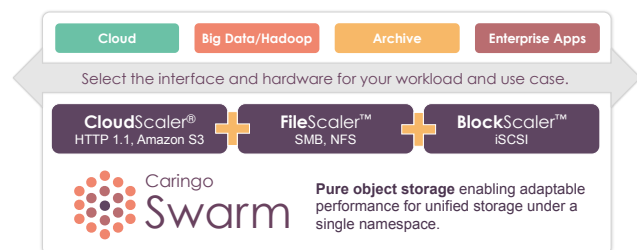
Swarm 7, doesn't have a metadata database: objects (data+metadata+policies) are stored together directly on raw disks without needing a File System. A copy of metadata index is maintained in RAM for faster access while the absence of an internal File System speeds retrievals and maximizes space efficiency.

These features are at the base of the huge scalability and advanced capabilities provided by the product. In fact, continuous data integrity checks, adaptive power consumption, aggressive recovery policies and node level caching are all present and automated. Data protection mechanisms are flexible and granular, they can be chosen based of the value of the stored data, the SLA, or the kind of needed local/geo distribution.

Caringo Swarm 7, thanks to NoSQL indexing of metadata and the interface for Hadoop, can also be an interesting companion solution for Big Data and Dark Data Analytics. Support for S3 API widens the, already large, spectrum of compatible third party applications and solutions.

The power of “Scalers”

Caringo, alongside, Swarm 7 is introducing a full set of gateways (CloudScaler, FileScaler, BlockScaler) capable of providing traditional access interfaces (like iSCSI, SMB and NFS) as well as S3 API compatibility.



These interfaces vastly extend Swarm 7 capabilities and enable every type of organization to make a seamless transition to this kind of next generation unified storage platform. For example, FileScaler can be presented as a traditional NAS to the clients but with all the power of the objects and hence the cloud.

Legacy and modern applications (web, mobile apps, Big Data analytics, etc.) can now be supported with a single secure and scalable storage platform which brings the concept of unified storage to a new level.

BOTTOM LINE

Juku has been talking about the need of next generation object-based unified storage platforms since 2010. It's happening right now: a solution like Caringo Swarm 7, with its enabling integrated interfaces (CloudScaler, BlockScaler and FileScaler), is a powerful building block to implement such a platform and it can be considered a key element on any mature private cloud strategy.

Object storage can be a valid alternative to legacy storage systems when you want to build innovative cloud services for your organization. Its advantages, especially at scale, and the overall much lower TCO, when compared to traditional storage systems, makes it a good choice to consolidate many different kinds of data in a single distributed and secure repository. Furthermore, its characteristics of high availability and unprecedented resiliency brake traditional data protection models and open new interesting scenarios for securely storing any form of unstructured data without needing costly and inefficient backup and DR infrastructures.

Caringo, with Swarm 7, consolidates its technology and market leadership with an innovative product capable of scaling from one single node to multi Petabyte archives. Its set of features enable organizations to build a cost effective software-defined platform aimed at building innovative storage services like, for example, integrated NAS and Sync&Share services. The vast number of partners and ready out-of-the-box solutions make this ecosystem one of the most complete of the industry landscape.

Consolidating data on a platform like Swarm 7 also opens up new opportunities like Big Data and Dark Data Analytics. Even though, your enterprise might not be ready to take advantage of stored data and transform them in information, storing massive amounts of data for future usage is cheaper than for most on-line or near-line systems and easier to be retrieved when needed!

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 Caringo. This paper was commissioned by Caringo. No employees at the firm hold any equity positions with Caringo. Should you have any questions, please contact Juku consulting srl (info@juku.it - <http://jukuconsulting.com>).