



9th TF-Storage meeting
(Thursday-Friday) 16-17 June, 2011
Athens, Greece

Table of contents

Table of contents	1
Minutes	1
1. Welcome	1
2. Approval of agenda and minutes.....	1
3. Fibre Channel over Ethernet.....	2
4. National updates.....	2
5. Storage benchmarking Work Item	4
6. Infrastructure as a Service clouds	6
7. Open discussion	8
8. Summary, next meeting, AOB, and close.....	9
List of actions	10
List of participants.....	11

Minutes

1. Welcome

The 9th TERENA Storage Task Force meeting was held on 16-17 June, 2011 hosted by GRNET in Athens, Greece. Maciej Brzezniak (PSNC), the chairman of the task force, welcomed the 27 participants attending the meeting in person and remotely via VC (list of attendees is enclosed). Brian Boyle (HEAnet), Vugar Musayev (AzScienceNet), and Jaroslav Kremenek (CESNET) apologised before. The TERENA secretary was Peter Szegedi.

2. Approval of agenda and minutes

The minutes of the last meeting (to be held in Budapest) was approved and all the actions were reported to be done. The proposed meeting agenda was also agreed with a minor change in the national update session; Peter Stefan (NIIF) presented their latest developments instead of Jaroslav Kremenek (CESNET) who was unable to attend the meeting.

The first day of the meeting was dedicated to national storage updates and specific actions of the task force work items. The second day was dedicated to NRENs' cloud related activities, especially in the area of infrastructure clouds (IaaS). The presentations are available on the TF-Storage website:

<http://www.terena.org/activities/tf-storage/ws11/agenda.html>

3. Fibre Channel over Ethernet

Andy Vallely (Cisco Systems) gave a technical overview on Fibre Channel over Ethernet (FCoE) protocol. Compared to iSCSI, Infiniband, and native Ethernet a clear technical case was made to verify the use of Cisco's FCoE solution in data centres.

FCoE is an extension of Fibre Channel onto a lossless Ethernet fabric. From the FC point of view it is FC connectivity over a new type of cable (called an Ethernet cloud) and from the Ethernet point of view it is yet another ULP (Upper Layer Protocol) to be transported. There is no native FC support in iSCSI environment because of the TCP/IP processes, Infiniband is still a knowledge gap, while everyone runs Ethernet networks. In order to enable FC over Ethernet, the Ethernet layer needs to be made lossless. The lossless Ethernet behaviour is guaranteed in FC by buffer-to-buffer (B2B) credits. FCoE remains Fibre Channel at the host and switch level. As Andy explained, FC does not drop packet. If FC sends the packets over Ethernet and Ethernet pauses FC can continue transmitting on a different buffer port. B2B credit control notifies back the source switch to slow down with transmission.

The main use case is direct attached FCoE storage; providing a unified FCoE network environment that preserves SAN architecture for high availability SAN A and SAN B. Actually, storage cloud is a series of data centres and that should be fast provisioned, that is why FC is efficient. Answering questions from the audience Andy explained that an IS-IS protocol expert may be needed to set up the whole network. It is not as easy as STP (Ethernet Spanning Tree Protocol) although performs much better. There is no typical Ethernet trace route mechanism; MAC addresses can be changed along the path. Fault management is also a bit more complex but the FCoE benefits are significant.

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110616-fcoe-andy.pdf>

Additional material:

<http://www.terena.org/activities/tf-storage/ws11/slides/Cisco-background-materials.pdf>

4. National updates

In the national update session on Day 1, GRNET, CARNet, and NIIF presented their latest developments and updates since the last meeting in Budapest.

- **Pitos v.2 storage system of GRNET**

Panos Louridas (GRNET), on behalf of the development team, presented the Pitos v.2 storage system of GRNET. Pitos storage system has 9883 users with 645,343 objects stored that exceed 3646.2 GBytes of data. The second version of Pitos is taking a more standard based approach as it uses block based storage (i.e. content-based addressing for blocks) instead of files. It has

Jettison J2EE (Java 2 Enterprise Edition) platform-independent Java-centric environment and OpenStack API (that was not available when Pitos v.1 was designed). Python programming language was chosen mainly because of better code portability to other projects (such as Okeanos, the SSaaS platform of GRNET). Deduplication, synchronisation, and partial updates are also new features that are supported by GRNET from now on. Web based access will continue to be offered in Pitos2 as well as the command line interface (CLI) that is also useful for running scripts. In the future Dropbox-like native clients will be developed. Mac OS and Android support is also in the development roadmap.

In longer term, Pitos2 will provide the underlying storage platform for other GRNET projects. For instance, the Scientific Software as a Service (SSaaS) will start next month and will be offered to the community as a software store based on Pitos2.

The storage service is designed in Pitos2 as a data base for hash tags and the blocks of data. Redundancy is not in the design yet. Pitos v.1 system uses files, so all the data has to be migrated into blocks to be stored in Pitos2. The service front-end was developed by a third-party based on GRNET's recommendations. SAN or NAS can be used as storage backend. GRNET does not charge their customers for this service; actually GRNET is funded by the government to provide this service for free. Although in the future, some kind of virtual charges have to be introduced for scheduling purposes.

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110616-pithos2-panos.pdf>

Pitos2 code is open source under BSD licence and available at:

<http://code.grnet.gr/projects/pithos>

- **sys.backup - Centralized open source backup solution for CARNet members**

Branko Radojević (CARNet) introduced an open source backup solution for CARNet community. He acknowledged that some of the basic concepts and ideas were picked up from their recent major involvement in the task force's activities.

The motivations behind the service development came from the usual practice of CARNet's member institutes in how they handle local data backup using expensive tape drives and manually exchanged tapes by the local administrators as a daily business. The solution offered by CARNet would be a centralized backup system for members (targeting system administrators and not end-users) over their existing connection to CARNet network. This solution should be easy to set up, secure, redundant (first on disk then on tape library), and cheap.

Sys.backup service is based on Bacula, an open source enterprise-ready software package. The storage and tape library hardware is a multi-tiered enterprise level backup system. The backup service is provisioned in less than 24 hours after receiving an e-mail request. User support is provided by CARNet help desk. The current limitations though are: the platform dependency

(Bacula runs only on Debian Linux), the capacity limits (100 GB per user), and the network bottleneck (running longer time on slow links). The software is open source, available on request, licence issues to be solved.

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110616-sysbackup-branko.pdf>

- **NIIF update**

Péter Stefán (NIIF) gave an update on NIIF's IaaS platform development. The initiative was first introduced at the Budapest meeting. Since then, it can be said that the service became popular among the NIIF users, so NIIF has considered upgrading it to a public infrastructure cloud service in production.

The service is based on OpenNebula core that is surrounded by NIIF developed cloud components. Its features include; self-service private virtual network management, live migration of VMs between sites (7 sites, 45 machines), integrated storage management, console access to VMs, and sophisticated access control system. Before the service can actually go to production the backend storage integration has to be finalised and the driver has to be deployed. 5 of the 7 host sites are already deployed but the proper network configuration has to be done. The operation manuals have to also be finalised.

IaaS GUI was demonstrated live. Slides are available at:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110616-niif-peter.pdf>

5. Storage benchmarking Work Item

- **Storage benchmarking by Cisco**

Andy Valley (Cisco) talked about the storage performance benchmarking issue from a vendor's perspective. Product testing and accreditation cost lots of money for vendors. Storage is often a multi-vendor environment (there is no single vendor that offers all) that increases the complexity of the performance benchmarking moreover there are no appropriate standards available. Either vendors test equipment according to their standard and NRENs are "suspicious" about it (e.g., in case of complex multi-supplier systems) or test it according to the customers' standard and cost extra money that will be built into the price, none of them is optimal.

Andy put the question "*Should the NREN community (NRENs and vendors/suppliers) work together to develop a standard for testing a complete system?*"

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110616-benchmarking-andy.pdf>

Discussion went on that practically a number of attributes can be chosen and changed during the benchmarking. Vendors need to know the exact goal of the benchmarking in order to pick the right attributes (there is no unified approach). Customers usually are able to test individual equipments during the tendering process (most likely at the vendors' lab) but complex systems cannot be checked in that way. A global "storage performance council" exists but they do not do end-to-end system testing either.

The meeting participants felt that perhaps *"The NREN community (NRENs and vendors/suppliers) should agree on best common practices and recommendations for testing storage systems."*

Storage benchmarking practices collected on the TF-Storage Wiki is the first step to achieve this goal. Andy offered to review the actual Wiki content from a vendor point of view and report back to the community about it.

ACTION 1 on Andy (Cisco) to review the actual storage performance benchmarking Wiki content from a vendor point of view and report back to the community about it.

TF-Storage benchmarking Wiki:

<https://confluence.terena.org/display/Storage/Measuring+storage+performance>

- **Related talk - Persistent IO challenges and approaches by FORTH**

Angelos Bilas (FORTH) gave a talk about persistent IO challenges and approaches from a scientific point of view. Modern application stacks (such as STREAM and CumuloNimbo) were presented that are quite complex and often IO is the significant bottleneck. To support these application stacks, scalable I/O stack (virtualized) over direct and networked storage devices must be provided. Performance and scaling analysis have to carefully be done. Today, persistence is "heavy" due to device/controller technology, not designed with multi-cores in mind and inefficient when scaling across nodes. What we should do is to get closer to the CPU, better support persistence for metadata, and understand overheads and scaling characteristics on modern systems.

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110617-forth-angelos.pdf>

Angelos and his team at FORTH have knowledge and expertise on these. Angelos offered to review the actual storage performance benchmarking Wiki content from a scientific point of view and report back to the community about it.

ACTION 2 on Angelos (FORTH) and his team to review the actual storage performance benchmarking Wiki content from a scientific point of view and report back to the community about it.

6. Infrastructure as a Service clouds

- **Ganeti cluster manager by Google**

Guido Trotter (Google) gave an overview on Ganeti, the open source cluster virtualisation manager of Google. GRNET IaaS solution is based on Ganeti core. The basic design principles were: to make the virtualization entry level as low as possible, to be scalable to enterprise ecosystems, and to be a good open source citizen. Ganeti is built on Linux and its standard networking utilities (iproute2, bridge-utils), KVM/Xen (support in progress for lxc), DRBD, or LVM, or files (SAN support is coming), Python (plus a few modules), socat, ssh, and Haskell.

Ganeti is currently used by Google Corporate Computing Infrastructure, Oregon State University Open Source Lab, GRNET, and many others. One Ganeti cluster can manage about 200 nodes. The IO performance depends on the hypervisor and the hardware. How Google monitors its hardware is not public. Before the hardware dies the data is migrated somewhere else, that is what the cluster resource manager does.

Ganeti code is available at:

<http://code.google.com/p/ganeti>

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110617-ganeti-guido.pdf>

- **VISION Cloud project by NTUA**

Dimosthenis Kyriazis (NTUA) talked about VISION Cloud, a 3-year project for content centric virtualised storage. The project started in October 2010. Its goal is to architect and implement an infrastructure for reliable and effective delivery of data-intensive storage services, facilitating the convergence of ICT, media, and telecommunications.

Today's data driven journalism is an example for data analysis and its importance. The project is not only about storage but data analysis too. The first architectural document will soon be published by the project. Multi-tenant cloud-native business application is a potential use case. The business model behind is a telco based typical cloud business model (nothing storage specific in there) contributed by Telefonica, the Spanish operator. If a data centre wants to be part of VISION Cloud that would be possible but there is no federation in place yet. A common access layer will be provided to access different data centres.

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110617-visioncloud-dimosthenis.pdf>

- **GRNET IaaS platform**

Vangelis Koukis (GRNET) introduced Okeanos platform that delivers IaaS to GRNET customers targeting IT departments as well as university students and researchers in academia. The long-term strategic concept is that all university IT will be outsourced to the cloud provided by GRNET and directly funded by the government.

The main design goal was to create a production quality Infrastructure as a Service. The software architecture is based on Synnefo, a custom-made cloud management software, and Google Ganeti backend. The virtual machines are powered by KVM, managed by Ganeti cluster manager, and accessible via web or programmable interface (OpenStack v1.1). Okeanos is integrated with GRNET AAI federation and Pitos v.2 storage service.

At the end of the presentation the Okeanos user interface and the basic service functionalities were demonstrated live.

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110617-okeanos-vangelis.pdf>

- **SURFnet IaaS pilots**

Rogier Spoor (SURFnet) presented some considerations and opportunities for creating a community cloud in the Netherlands.

Rogier started with a poster proposal on how to position a community cloud in the overall cloud arena. Meeting attendees were asked to send comments, suggestions about the content of the poster to Rogier via e-mail. The final poster will be distributed to the TF-Storage e-mail list once it's ready.

Community cloud is a shared infrastructure for specific community. Its functional features are mostly in line with public cloud and there is an added value of the community dimension (such as control over data location, single legal jurisdiction, flexibility, standardisation to prevent lock in, network integration). In a brokered model third party resources can be shared via a broker. SURFnet prefers the brokering model.

Currently an in-house pilot is running based on DRBD, iSCSI, VMware and using CloudStack GPL version. Live migration is possible between two sites in the Netherlands. The Web GUI was demonstrated live. The next step would be to pilot a community cloud setup involving two Dutch universities. It was also found out that GreenCloud a private company in Iceland, is using the same architecture. SURFnet is planning to cooperate with them and migrate live VM instances to Iceland. This is part of the green storage/computing efforts of SURFnet.

Slides:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110617-iaas-rogier.pdf>

7. Open discussion

The open discussion about the future challenges of TF-Storage (concerning IaaS and clouds) was chaired by Maciej Brzezniak (PSNC). Maciej repeated the high-level statement that was agreed before: "Only a large-scale, pan-European cloud initiative can provide the economic benefits and technical advantages demanded by the global research and education user community."

Furthermore, Maciej asked the meeting attendees:

- Is a European-level initiative for establishing cloud services necessary in your opinion?
- If yes, what are the chances to either build or buy-in such a common cloud infrastructure?
- And last but not least, what can we (TF-Storage) do at the moment?

Maciej emphasised that the NRENs' involvement in any pan-European cloud initiative is essential because they have the network (operational experience as well as cost advantage), they can ensure sustainability in long term, and they hold the trust of their customers supported by (inter)federations.

What TF-Storage can definitely do is to continue working on the technical level towards storage, IaaS, and clouds. Peter (TERENA) mentioned that the official mandate of TF-Storage will expire in the end of February 2012. The group should start thinking about any potential continuation of the task force activities and practically about the re-chartering of the task force Terms of Reference. The actual ToR is available here:

<http://www.terena.org/activities/tf-storage/tf-storage-tor2-v2.pdf>

Maciej highlighted some potential areas in order to foster better collaboration in the future:

- Explore cloud technology on paper and in practice.
- Pilot the technologies within TF-Storage.
- Develop new platforms/services, if appropriate.
- Forum to discuss opportunities and possibly apply for EU funded R&D projects in the area of clouds to demonstrate the ability of the NREN community.
- Work together on high level architectural design issues such as outsourcing of services, security and privacy issues, maybe standardisation and best practices.

The discussion led to two practical next steps and some actions:

1. *Collect our knowledge on cloud (IaaS, storage) solutions of NRENs:*
 - a. agree on the set of the attributes of the national deployments such as hardware/software platform, total capacity, replication, encryption, etc.
 - b. create a comparison table on the TF-storage Wiki.

Andy (Cisco) mentioned that Cisco has something about architectural principles and attributes. That can be used as an initial template.

ACTION 3 on Andy (Cisco) to look after the cloud/storage architectural principles and attributes at Cisco and inform the community if there is something available.

ACTION 4 on Peter (TERENA) and Maciej (PSNC) to create the skeleton of the technology/platform comparison table on the Wiki that can then be filled out by the interested TF-Storage participants.

2. *Do some piloting/test deployment of cloud storage middleware/solutions jointly:*
 - a. test out each other's solutions,
 - b. discuss implementation options and agree on applications interesting for piloting.

Participant found that it is too early to cooperate on the infrastructure level yet. Common practices, cloud API implementations, profiles may be shared by now. Rogier (SURFnet) asked for a guest account to the GRNET cloud to test it out. NIIF will also be asked to provide the same, if possible. Faidon (GREENT) replied that GRNET is in principle open to any joint pilots.

ACTION 5 on Faidon (GRNET) to create guest account to their IaaS demo environment to the interested participants.

Maciej's slides are available at:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110617-discussion-maciej.pdf>

8. Summary, next meeting, AOB, and close

At the end of day 1, Peter Szegedi (TERENA) gave an overview on the actual statuses of the task force work items and warned participants that the end of the task force mandate is approaching fast and we need to make a clear case for continuation.

Peter also mentioned the GN3-ASPIRE study that is going to look at clouds

<http://www.terena.org/activities/aspire/>

as well as the new nren-cloud mailing list <nren-cloud@terena.org> that was created by TERENA upon the request of the TNC2011 "Provisioning infrastructure services on-demand: beyond Grids and Clouds" BoF participants

<https://tnc2011.terena.org/core/event/6>

Peter's slides are available at:

<http://www.terena.org/activities/tf-storage/ws11/slides/20110616-future-peter.pdf>

After the TF-Storage meeting happened

TERENA Technical Committee (TTC) had a brief discussion about NRENs' general strategy

towards clouds. TTC came up with three recommendations (relevant to TF-Storage community and the IaaS/cloud discussion):

- TTC recommended organising a face-to-face meeting (preferably in September 2011) among all the TERENA technical and non-technical task force chairs and secretaries about clouds (among many other topics). Maciej (PSNC), as the chair of TF-Storage, has been invited to this meeting to represent TF-Storage.
- TTC recommended organising a potential Cloud Workshop (preferably in November 2011) to the larger TERENA community.
- TTC recommended the NRENS to seriously consider the FP7-ICT-2011-8 EC project call about clouds (Date of publication: 26 July 2011, Deadline of submission: 17 January 2012)
<http://secaia.com/content/ict8-fp7-ict-2011-8-17january-2012>

Peter Stefan (NIIF) signalled that NIIF/HUNGARNET is seriously interested in and willing to submit a joint project proposal to this call. If any other TF-Storage participants are interested in to join, please, contact NIIF directly.

About the next TF-Storage meeting

TF-Storage participant found that the October-November period is already crowded with other storage and cloud related events (and with a potential TERENA cloud workshop) so the earliest feasible time for the next coming face-to-face TF-Storage meeting would be in January-February 2012. (Virtual meetings may happen in the meantime!)

We have two potential hosts for the next meeting; ARNES in Ljubljana and CARNet in Dubrovnik. Peter Szegedi (TERENA) will approach them in this order as soon as the meeting dates are agreed.

The meeting was closed on time. All the contributions were appreciated. Big thank to GRNET for the excellent meeting facilities was given.

List of actions

- ACTION 1 on Andy (Cisco) to review the actual storage performance benchmarking Wiki content from a vendor point of view and report back to the community about it.
- ACTION 2 on Angelos (FORTH) and his team to review the actual storage performance benchmarking Wiki content from a scientific point of view and report back to the community about it.
- ACTION 3 on Andy (Cisco) to look after the cloud/storage architectural principles and attributes at Cisco and inform the community if there is something available.

- ACTION 4 on Peter (TERENA) and Maciej (PSNC) to create the skeleton of the technology/platform comparison table on the Wiki that can then be filled out by the interested TF-Storage participants.
- ACTION 5 on Faidon (GRNET) to create guest account to their IaaS demo environment to the interested participants.

List of participants

Christos Argyropoulos	NTUA
Andrej Bagon	Arnes - remotely
Angelos Bilas	FORTH-ICS
Maciej Brzezniak	PSNC
Ivan Cabric	CARNet - remotely
Sasa Cavara	CARNet
Panagiotis Christias	NTUA
Athanasios Douitsis	National Technical University of Athens NOC
Nikola Garafolic	University Computing Centre (SRCE)
Dimitris Kalogeras	NTUA
Thorsten Kleindienst	SWITCH
Alexandros Kosiaris	GRNET
Vangelis Koukis	GRNET
Dimosthenis Kyriazis	National Technical University of Athens
Martin Kämpf	SWITCH
Faidon Liambotis	Greek Research and Technology Network
Rossend Llurba	NCF
Panos Louridas	GRNET
Apollon Oikonomopoulos	GRNET
Branko Radojevic	CARNet
Rogier Spoor	SURFnet
Peter Stefan	NIIFI
Peter Szegedi	TERENA
Guido Trotter	Google
Andy Vallely	Cisco Systems
Mario Vandaele	Belnet
Dimitris Zacharopoulos	AUTH-NOC

Apologized:

Brian Boyle	HEAnet
Vugar Musayev	AzScienceNet
Jaroslav Kremenek	CESNET