

Update on the ASDEX Upgrade Data Acquisition and Data Management Environment

K. Behler, H. Blank, A. Buhler, R. Drube, H. Eixenberger, K. Engelhardt, A. Lohs, R. Merkel, G. Raupp, W. Treutterer, ASDEX Upgrade Team, Rechenzentrum Garching

Max-Planck-Institut für Plasmaphysik, Boltzmannstr. 2, D-85748 Garching bei München

Abstract

It has been a while since it had been reported on the status of ASDEX Upgrade Data Acquisition (DAQ) and data management environment. An update on changes, expansions, and enhancements applied in the last years will be given.

- *The acquired amount of data per shot increased from 4 GiB to 40 GiB in eight years.*
- *Network, storage, and archive challenges have been managed by stepwise improvements.*
- *New DAQ techniques have been introduced to replace outdated technologies.*
- *Real-time diagnostics speed-up data provisioning and contribute to feedback control.*

Information technology applied to ASDEX Upgrade is under permanent change. Recent and future steps are outlined.

Keywords: data acquisition, data management, real-time, diagnostics

Introduction

When ASDEX Upgrade went into operation in 1991 its data acquisition and data management was designed to deal with tenth of diagnostics up to the megabyte range [1]. Today nearly 150 diagnostics deliver up to 40 gigabytes of data for each shot. Continuous adaptation and development was necessary to adjust the data acquisition and handling facilities to the needs of a growing and demanding experiment. Legacy as well as advanced new DAQ systems are managed in compatibility with the DAQ design chosen at the very beginning.

Stored data per shot

Figure 1 shows the stored data amount per shot in semi-logarithmic scale over shot number. 30000 “shots” - as the plasma discharges are called in a historic reminiscence to the era when discharges went on with a bang - have been conducted so far. As the number of shots stayed more or less constant between 1200 and 1500 per year over the last 22 years, a linear relation between time and shot numbers can be assumed [2]. With this assumption the diagram yields a nearly exponentially data growth over time with a growth factor of 2 in 2.1 years or 10 in approximately 7 years. Since

ASDEX Upgrade Diagnostic Data over Shot

GiB per shot averaged over 1000 shots

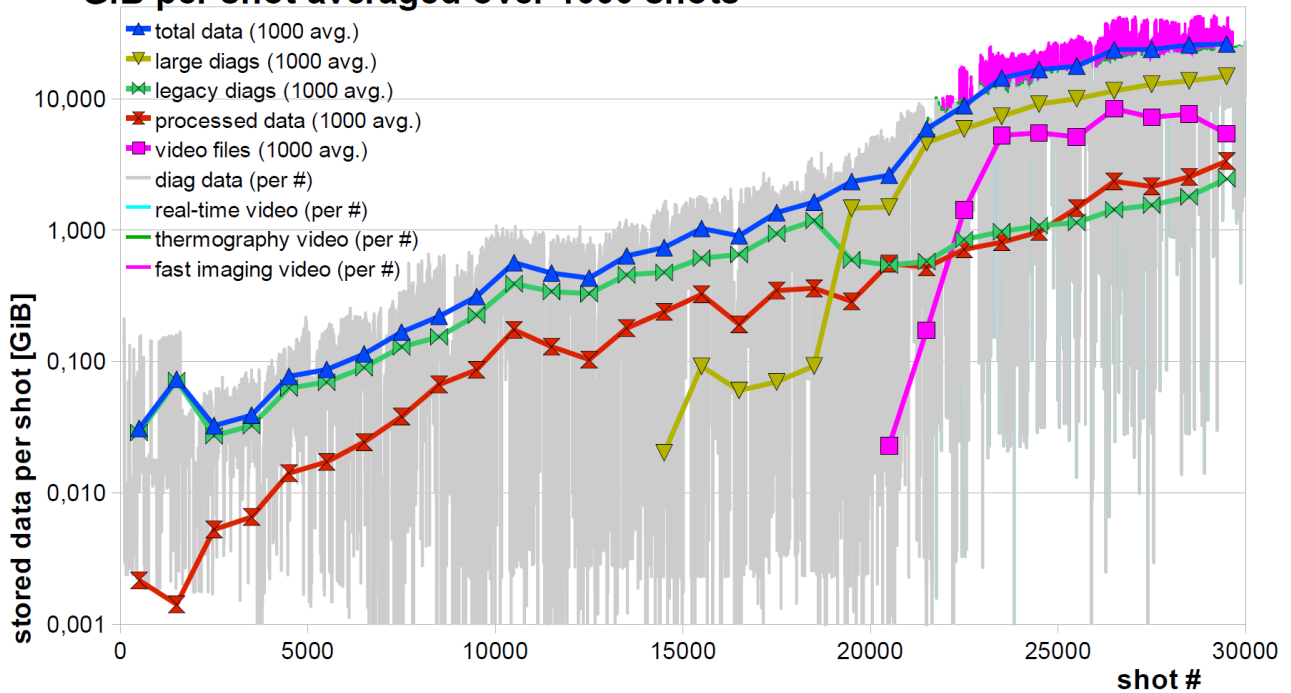


Figure 1: Development of stored data amount per shot (see detailed discussion in the text)

progress in information technology shows an exponential behavior itself, the exponential growth here represents a constant investment of funds and effort in applying information technology to extend the DAQ possibilities of the experiment. This may be new diagnostics, refurbished diagnostics, additional channels for existing diagnostics, higher sampling rates, and prolonged data acquisition times.

The noisy grey background shows the shot by shot variation of stored data amount while the imprinted curves being averaged over 1000 shots depict the long-term development of certain categories of data.

The topmost curve (triangles tip up, blue) shows the averaged total data.

The curve starting next to it (marked with horizontal double triangles, green) illustrates the legacy or small diagnostics - files below 100 MiB per shot. A couple of years ago we started to refurbish them one by one using our own Serial Input/Output (SIO) technology [3][4]. Not always, but in many cases this went along with the enhancement of analog channels demanding a higher bit resolution, higher sampling rates, and/or more measuring channels. Some of those diagnostics subsequently fell over into the class of “large diagnostics” explaining the dip in the curve of the small diagnostics near shot number 18000.

Due to the replacement of Transputer and CAMAC by HotLink and SIO DAQ configurations large diagnostics (triangles tip down, light yellow) with file sizes from 300 MiB to 2 GiB per shot are

taking over the lead. With the new technology Mirnov and Soft-X-Ray were suddenly able to collect the full discharge in full time resolution instead of only a short time interval. Further new diagnostics with SIO based acquisition contribute to the steep increase of this class. These are e. g. extended UV diodes (diagnostic names: XVR, XVS), Doppler reflectometry (PRA), electron cyclotron emission radiometer (RMC) and the corresponding real-time analysis (RRC).

Processed data files further are a considerable fraction of the data amount (vertical double triangles, red). These are files in diagnostic data format generated from measured data by analysis processes.

Recently the video diagnostics added a remarkable contribution which is mainly delivered by high resolution, high speed cameras (square marks on pink curve and pink spikes sitting on top of the background). The video streams or image collections are stored in their genuine video or file format and handling of these files differs from that of those in ASDEX Upgrade data format [4].

Management of diagnostics

ASDEX Upgrade diagnostics (DIAGs) are designed as independent systems. Today typically 100 to 150 DIAG programs, depending on the actual physics program, run distributed on a comparable number of individually configured diagnostic hosts. Being autonomous by design, nevertheless, DIAGs are manageable from a central application called diagnostic control center (DCC).

Communicating via “Diag Sync” agents, which are system tasks supervising the various diagnostic hosts and providing task administration rights to an authorized requester, the DCC collects status information and sends commands from and to the DIAGs. Thus the DCC application with its graphical user interface provides manageability to the DAQ supervising officer. The concept of the Diag Sync agent and the DCC has been described earlier [1]. It should only be mentioned here that details have been refined over the years leading to a design allowing to fully operate a large number of DIAG programs from a central place.

The nature of the various diagnostics can be very different: old or new technology, delivering small or large data volume, acquiring postmortem from transient recorder memories or real-time into main memory, delivering data with low latency or late after the shot has finished. While differing in many details all diagnostics follow the same major operation cycle and finally deliver data to a central instance the shot file monitor (SFMON). Figure 2 gives a schematic view how diagnostic data is managed via the SFMON. From top to bottom the progress of data from a diagnostic system into the archive is shown. From left to right the principal time behavior is shown. The top box describes the activity of a typical diagnostic (DIAG). From left to right the basic operation phases are depicted: **prepare**, **measure** (real-time activity or wait for external devices), **acquire** and **store**. Phases are started by time marks distributed from the Discharge Control System (DCS) over various

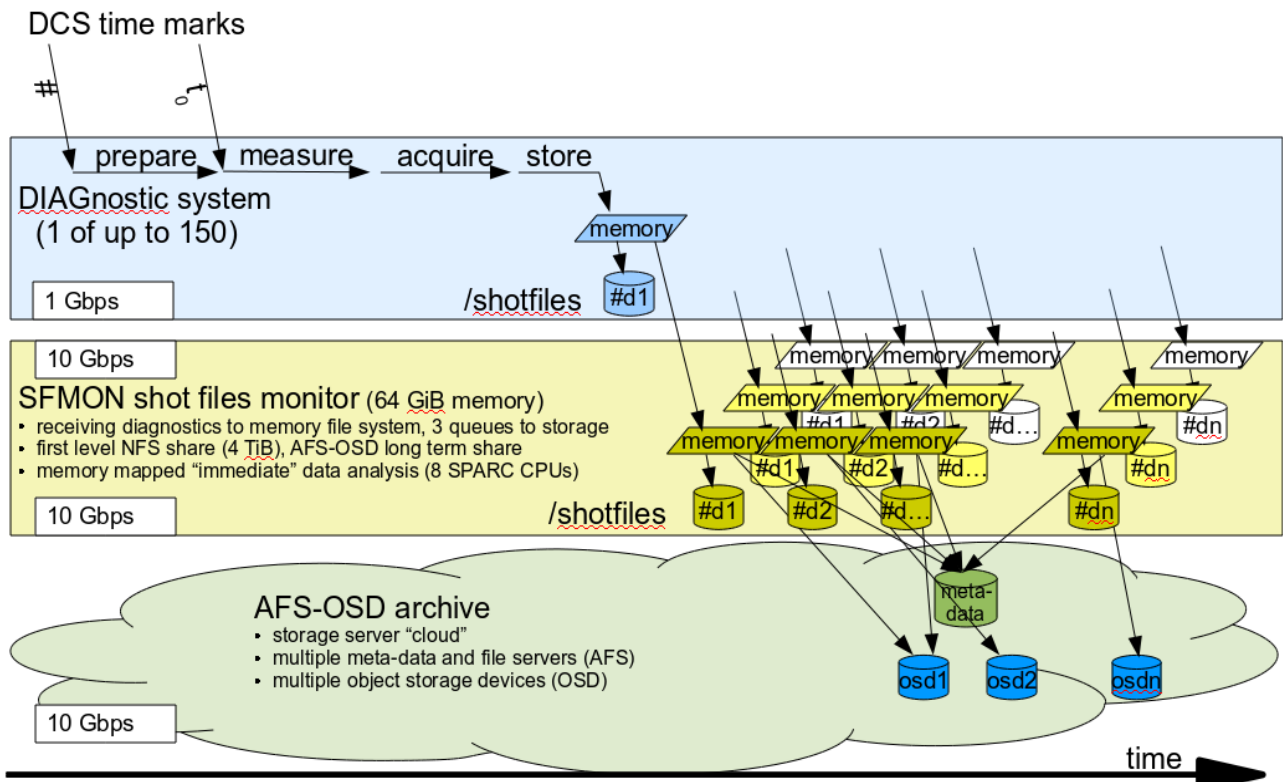


Figure 2: Schematic timing and data management of diagnostics, the role of the central shot file manager (SFMON), and final archival of shot files.

paths (central timing system, event broadcasting system, DCS progress bar, etc.) or by detecting the end of the previous phase.

Prepare is generally started when a new shot number indicating a new experiment run is distributed. **Prepare** not only checks and arms all devices, but on real-time diagnostics further configuration information with DCS may be exchanged to establish a real-time cooperation between DIAG and DCS. At the end of **prepare** the DIAG is ready to measure.

The **measure** phase is usually started with plasma ignition t_0 . (A couple of predecessor signals to t_0 exist and various methods are available to start diagnostics at the optimal point in time. These are not shown to avoid confusion. However, t_0 is the common time reference for all diagnostics.) While old type DIAGs (e.g. equipped with CAMAC devices) just wait for the hardware recorders to signal end of data acquisition, new types of diagnostics actively transport data from DAQ periphery into computer memory. From the DIAG view there is no big difference between these both types except that in the new ones during the **measure** phase real-time threads managing DMAs (direct memory accesses) are active. **Measure** normally ends with an intrinsic event as "CAMAC activity finished", "all configured samples received", or simply "configured measuring time expired" which starts the next phase.

During **acquire** the captured samples from the various external and internal memories are collected

and demultiplexed into signals which then are aggregated together with header information into a diagnostic shot file structure in memory. The subsequent **store** phase finally saves the shot file structure from memory first to local disk and afterward via socket connection to SFMON. While a single DIAG arrived at this point reverts to the beginning, SFMON manages centrally the further provisioning of the whole number of diagnostic files for "immediate" accessibility and long term storage. (see next section)

Beside the described creation of shot files from diagnostics, there exists a real-time framework for low latency access, processing and sharing of short time intervals of diagnostic data in real-time [5] [6]. While this is beyond the topic of this article it may be allowed to just mention its main goals. The simplest could be an internal feedback algorithm to the local DAQ process e.g. to adapt the input gains of ADC channels. More sophisticated scenarios contribute evaluated diagnostic results in real-time to the DCS. Even the involvement of multi-node physics analysis or simulation applications for higher physics control purposes is possible. This concept is used e.g. for the NTM stabilization project [7].

Role of the central shot file monitor

The last transaction in DIAG cycle is forwarding data to the central shot file monitor (SFMON) which is meant to serve as a central collecting, first level sharing, and immediate synoptic analysis platform. All diagnostics forward their data to this one central computer. Due to its 8 processor fast SPARC architecture, its 64 GiB memory, and its fast 10 Gbps interfaces SFMON is able to handle multiple incoming connections efficiently. Figure 2 gives an overview of the functionality while Figure 3 illustrates the timing of SFMON managing in this case 110 diagnostics in a typical experiment shot cycle.

As indicated in Figure 2 SFMON receives in parallel and in arbitrary order transfer requests from all diagnostics currently participating in experiment operation. The incoming diagnostic files at first go into memory (in fact a memory based tmpfs file system) to make the transfer as fast as possible. After the transfer each file is immediately stored into a 4 TiB RAID partition named /shotfiles and in parallel files are put on three queues to be send into the AFS-OSD archive "cloud".

For data stored in /shot files a three fold of services is performed by SFMON: 1st - "Immediate" shared memory mapped data access for analysis processes running locally on SFMON. 2nd - NFS sharing of /shot files providing an early access path to recent shot files on the local network. With its 4 TiB capacity SFMON serves a short term storage (one week or two) for ASDEX Upgrade data. 3rd - Scheduled and serialized archival of shot files into the AFS cloud for wider sharing and long-term access.

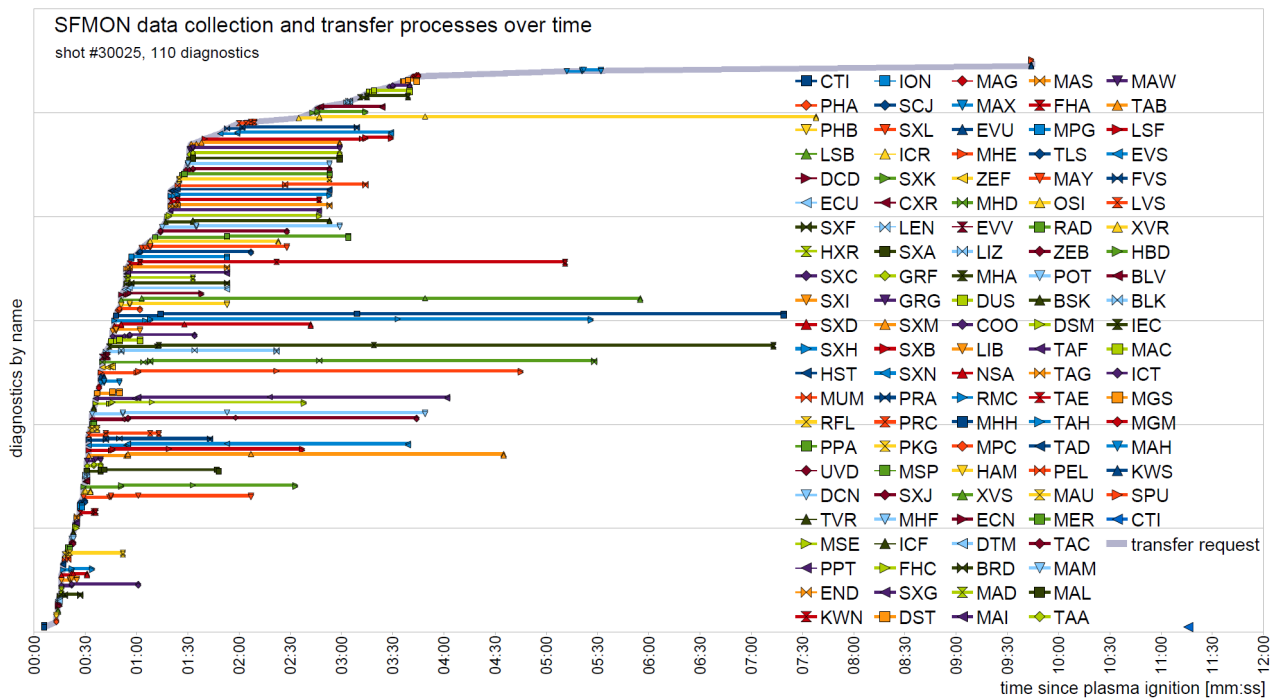


Figure 3: Timing log of diagnostics as they arrive at and are stored and forwarded by the ASDEX Upgrade shot file monitor (SFMON). (The full spreadsheet for this diagram is available as downloadable addendum to this paper. This as well documents the size of the particular diagnostic files.)

Figure 3 illustrates the arrival and forwarding of diagnostics at and by the SFMON. It shows a recent shot (30025) featuring 110 diagnostics. The horizontal axis represents wall clock time relative to experiment pulse (shot) start in 30 second intervals. The vertical axis is the sequence number as the diagnostics request transfer to SFMON. No particular sequence is predetermined. Large and small diagnostics come in and are received as they become ready on a first come first serve principle. The enveloping curve of this diagram, shown as a thick transparent line, illustrates the intrinsic data acquisition latency of the various diagnostics. Diagnostics are listed by name in the incoming order (y-axis from bottom to top, diagram legend from upper left to lower right). The meaning of the abbreviated diagnostic names is of no particular importance for this paper. Horizontally a line is drawn for each diagnostic beginning with the transfer-request-time and ending with the archival-to-AFS-finished-time. A little step is inserted on these horizontal lines at the time the transfer from diagnostic host to SFMON memory is finished and the store operation to DISK and AFS is started. The time between the step and the next symbol on the bar or the end of the bar is the additional global NFS respectively AFS latency of the diagnostic. The initial steep slope of the over all transfer curve indicates many diagnostics arriving in a short time. So it is of particular importance to run SFMON on a powerful platform gracefully responding to high peak loads. In fact most data is available within 2 minutes after the shot. This as well is a consequence of tuning the individual diagnostic systems to deliver data as fast and early as possible.

(A few diagnostics having special tasks which take their time come late by design. This causes the long tail in Figure 3, but does not harm the general operation.)

Storage challenges

While AFS still has its advantages for managing manifold data archives[8], some aspects of development seemed to be behind the speed of today's network and storage technologies development. To enhance the ASDEX Upgrade archival procedure, a modified AFS called AFS-OSD (OSD standing for Object Storage Device) is used which is intended to extend the classical OpenAFS [9]. AFS-OSD is featuring a parallel storage concept and was developed at the Computer Center Garching based on openAFS to make storage of big data files (e.g. from Supercomputers or large experiments) in AFS more efficient. While small objects still are stored in their respective AFS volumes, objects above a certain size are stored separately in object stores. The file distribution to the object stores is managed by configurable rules considering e.g. the filling degree or group assignment of object stores. The advantage is that an AFS server can be extended flexibly by an arbitrary number of object stores. These being separate file servers offering a significantly extended number of data paths for parallel requests or even a striped writing algorithm when configured. The as well modified AFS-OSD client provides direct transfer of bulk data to and from the object storage. The OSD storage philosophy also foresees rules for background data duplication and check-sum management for data integrity and verification. By duplication of data to a hierarchical storage system long-term tape storage can be attached transparently.

Currently it is not clear, if the AFS performance development over the next years will meet the prospected needs of the experiment. The still exponential growth of ASDEX Upgrade's storage requirements may come to a point where a storage paradigm shift may become unavoidable. Looking at the price per disk capacity development over time [10] and big-data concepts recently featured on the business market may suggest to change over to a completely commercial storage and database solution.

Conclusions

The data acquisition and management requirements of ASDEX Upgrade have been growing exponentially and are expected to do so in future. Various improvements and developments answer the resulting demands without breaking the original diagnostic system design philosophy. The DIAG concept was extended to support real-time requirements, central management of diagnostics and a reasonable speed-up in provisioning data after a shot. Data storage and long-term data archival was adapted and will still be subject to further considerations.

References

- [1] K. Behler, H. Blank, A. Buhler, R. Drube, H. Friedrich, K. Förster, K. Hallatschek, P. Heimann, F. Hertweck, J. Maier, R. Merkel, M. -G. Pacco-Düchs, G. Raupp, H. Reuter, U. Schneider-Maxon, R. Tisma, M. Zilker, Review of the ASDEX Upgrade data acquisition environment - present operation and future requirements, *Fusion Engineering and Design* 43 (1999) 247-258, [http://dx.doi.org/10.1016/S0920-3796\(98\)00392-5](http://dx.doi.org/10.1016/S0920-3796(98)00392-5)
- [2] Albrecht Herrmann, Otto Gruber, ASDEX Upgrade - Introduction and Overview, *Fusion Science and Technology* 44 (2003) 569-577, <http://epubs.ans.org/?a=399>
- [3] K. Behler, H. Blank, H. Eixenberger, A. Lohs, K. Lüddecke, R. Merkel, G. Raupp, G. Schramm, W. Treutterer, M. Zilker, ASDEX Upgrade Team, Real-Time Diagnostics at ASDEX Upgrade - Architecture and Operation, *Fusion Engineering and Design* 83 (2008) 304-311, <http://dx.doi.org/10.1016/j.fusengdes.2008.01.015>
- [4] K. Behler, H. Blank, A. Buhler, R. Cole, R. Drube, K. Engelhardt, H. Eixenberger, N.K. Hicks, A. Lohs, K. Lüddecke, A. Mlynek, U. Mszanowski, R. Merkel, G. Neu, G. Raupp, M. Reich, W. Suttrop, W. Treutterer, M. Zilker, ASDEX Upgrade Team, Real-Time diagnostics at ASDEX Upgrade - Architecture and operation, *Fusion Engineering and Design* 85 (2010) 313-320, <http://dx.doi.org/10.1016/j.fusengdes.2010.03.001>
- [5] Wolfgang Treutterer, Gregor Neu, Gerhard Raupp, Thomas Zehetbauer, Dieter Zasche, Klaus Lüddecke, Richard Cole, ASDEX Upgrade Team, Real-time signal communication between diagnostic and control in ASDEX Upgrade, *Fusion Engineering and Design* 85 (2010) 466-469, <http://dx.doi.org/10.1016/j.fusengdes.2010.04.031>
- [6] Wolfgang Treutterer, Gregor Neu, Gerhard Raupp, Dieter Zasche, Thomas Zehetbauer, Richard Cole, Klaus Lüddecke, Richard Cole, ASDEX Upgrade Team, Management of complex data flows in the ASDEX Upgrade plasma control system, *Fusion Engineering and Design* 87 (2012) 2039-2044, <http://dx.doi.org/10.1016/j.fusengdes.2012.01.038>
- [7] M. Reich , A. Bock, K. Behler, L. Giannone, M. Maraschek, E. Poli, J. Stober, W. Treutterer, ASDEX Upgrade Team, Progress on ECCD-based NTM rt-control at ASDEX Upgrade, 38th EPS Conference on Plasma Physics 2011 35G (2012) 1704-1707, <http://ocs.ciemat.es/EPS2011PAP/pdf/P5.102.pdf>
- [8] A. Maslenikov, Storage Technology Overview, <http://www-fusion-magnetique.cea.fr/tmiaea2009/website/data/bydate.html?b=3#3> (visited 2014-04-09)
- [9] H. Reuter, Integration of AFS/OSD into OpenAFS, <http://conferences.inf.ed.ac.uk/eakc2012/slides/IntegrationAFS-OSD.pdf> (visited 2014-04-09)
- [10] M. Komorowski, a history of storage cost, <http://www.mkomo.com/cost-per-gigabyte> (visited 2014-04-09)

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.fusengdes.2014.04.030>.