INFN CNAF Contribution

M.Morandin,
L.dell'Agnello, D.Salomoni, C.Vistoli

EGI TF, Lyon – 19-23 September, 2011

The starting point: the present customers and setup ….

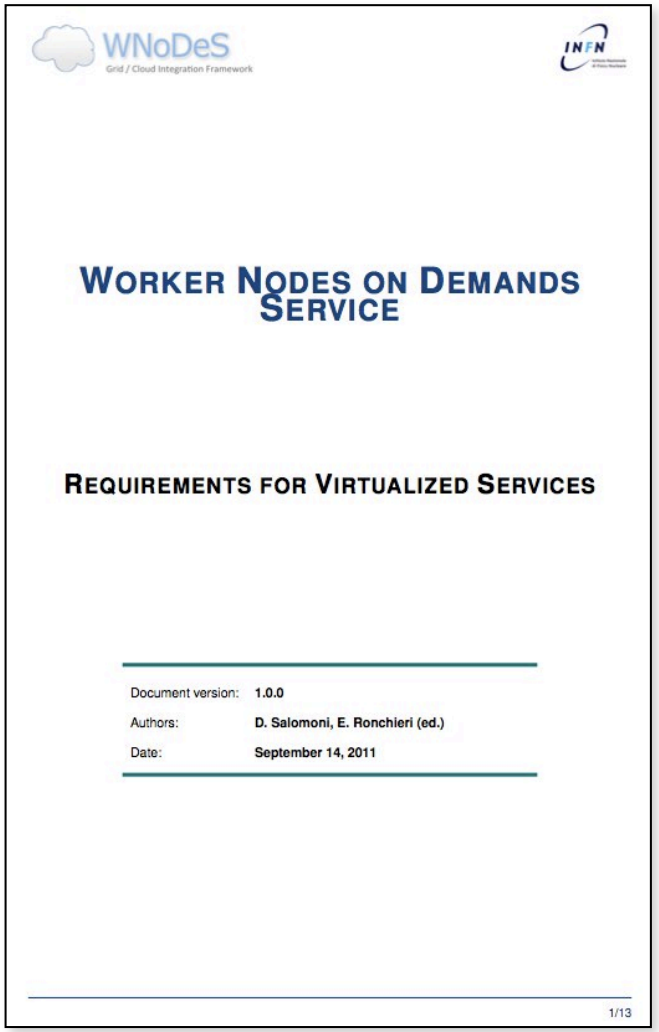
- INFN CNAF (located in Bologna, Italy):
  - Currently hosts about 8,500 computing cores (10,500 by September 2011), 9 PB of disk space, 10 PB of tape space
  - Runs 40,000 to 50,000 jobs every day, with several thousands of jobs pending at any given time
  - Supports many international experiments of astro-particle physics (e.g., AMS2, Argo, Auger, Fermi/Glast, Magic, Pamela, Virgo) with about 14% of its computing resources
  - Supports the Italian Tier-1 for CERN-based LHC experiments of high energy physics (i.e, ATLAS, CMS, LHCb, Alice) and is Tier-0/1 for several others, T2 for LHCb
  - All resources accessible through grid interfaces
    - E.g. unique farm (not partitioned), uniform storage management
  - Fabric level based on industrial standards
    - Platform LSF as batch system, IBM GPFS as shared FS integrated with TSM (as tape manager)
  - As Tier-1 we have 24x7 service (~ 20 FTEs for all the CC)

… and requirements

- A first implicit (rather obvious) requirement: unique management of all resources for all our "customers" (as it is now)
  - Less man-power required
  - Lower system complexity
  - Long term sustainability
- Common users' requirements include
  - flexible workload / job management systems
  - efficient use of storage systems
  - trivially parallel workloads possibly exploiting many-core systems
  - cloud-like instantiations
    - See WNoDeS (http://web.infn.it/wnodes) for a possible solutions to implement virtualized service
      - in production since months on a large scale (~ 2K cores)

Some examples of “new” requirements

- Using custom execution environments at both the Grid and local level
- Self-allocation of custom resources for the instantiation of virtual pools
  - Keeping in mind scalability and efficiency of resource usage
- Integrated, flexible authentication and authorization
- Long-term data access
- Efficient coexistance of real and virtual services
  - Exploiting many-core systems, GPGPUs, high-performance workloads with I/O-bound processes
- Provide additional, ad-hoc services


WNoDeS
Grid / Cloud Integration Framework

WORKER NODES ON DEMANDS SERVICE

REQUIREMENTS FOR VIRTUALIZED SERVICES

Document version: 1.0.0
Authors: D. Salomoni, E. Ronchieri (ed.)
Date: September 14, 2011

1/13

The present situation with Grid and the LHC approach

- Our main user community is WLCG
- The success in exploiting the Grid by the LHC experiments depends on many developments that they have accomplished in a rather independent way; two aspects of such process have turned out to be rather unpleasant for the resource providers, i.e. :
  - to a large extent, the solutions that have been designed and developed are experiment specific (or specific for WLCG experiments!)
  - the evolution of the experiment usage models has been pointing towards a more and more direct management of the computing resources (possibly allocated rather statically by the computing centers)
- This approach comes with drawbacks:
  - security issues, resource efficiency exploitation issues, maintainability issues, support at CC issues
- For sure this cannot be a sustainable model if we want to increase the number and the variety of communities using the Grid
  - Obviously build on experience and feedback from WLCG to enhance grid (cloud?)

Integration of new and old requirements and communities

- Implies significant investments in terms of research and development
  - Research: follow and act on market trends, invest in collaborations with industry and with other academic centers
  - Development: increase efforts on strategic sectors that are not currently adequately covered by available products
    - E.g., StoRM, WNoDeS, multi-CPU/GPU support
- Must be financially viable
  - Minimize duplication of efforts, changes to key workflows, manpower needs
  - Reduce involvement in activities less relevant for the center
  - Exploit resource usage and increase flexibility
  - Attract new customers (certainly academic, but also possibly from the public sector) to tap new sources of funding

Middleware

- Expand beyond the concept of “Grid middleware”
  - A layer mediating flexible access to computing, storage, network resources through Grid, Cloud interfaces (or else)
    - This is not the middleware we have today.
    - **Judicious** use of virtualization may help on this.
- The one thing we **do not** want:
  - To support multiple interfaces, middlewares, stacks in order to satisfy the needs of “generic” communities having requirements largely different from ours.
    - The financial sustainability considerations shown above apply: we will eventually have to choose what to support and for whom, if convergence is not realized.
    - The risk is of course to partition potentially large amounts of distributed resources into community-specific pockets.

Some considerations and conclusions (1)

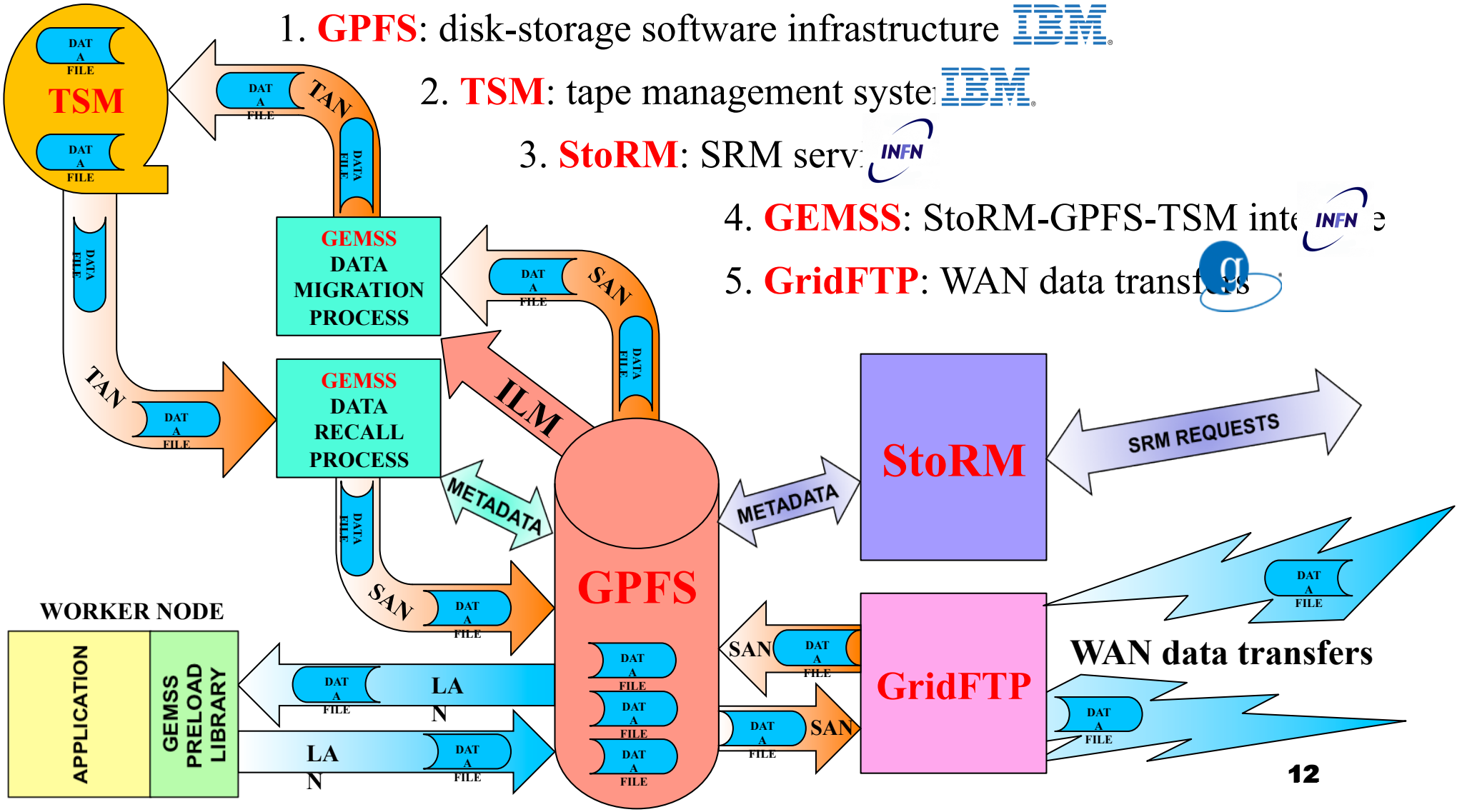
- An evolutionary path toward new services is probably not well defined at the European level
  - On the one hand, there are "traditional" services like Grid computing
    - Not much innovation here, anyway – several of our customers seem to agree that they have to keep on complementing Grid middleware with specific services
  - On the other hand, there seems to be focus on "definining Cloud profiles"
    - Conceptually interesting, but done many times in the past
    - No clear, sustainable path toward scalable, multi-center implementations and service offerings
  - Some areas where capabilities are lacking:
    - Workload management systems; authentication systems; flexible information systems (take for example MPI, GPGPU, VM availability and instantiation – for Grids and Clouds); inter-cloud connectivity.
- We need the Grid, the Cloud, the local and site-specific services (now)
  - But again, we can't afford to build "test infrastructures", or to offer separate services, or to reduce service levels for existing key customers in the hope to attract others.

Some considerations and conclusions (2)

- We are willing to contribute with our own requirements
  - Hopefully considered in plans for future European service offerings
- We will strive for reinforcing direct links between big European academic resource centers
- We can share our experience in our efforts to build integrated Grid and Cloud services within the Italian Grid Initiative (IGI)
- IGI, the Italian Grid Infrastructure (NGI) has now been activated:
  - IGI is now hosted within INFN, but it's managed by a Executive Board formed by representatives of Italian Research Institutions and Universities
- IGI is expected to participate in national and international R&D and infrastructural projects that could make it possible to secure the resources needed for setting up advanced services for a wide spectrum of users
  - A first example is the project for building the new distributed computing centers for the SuperB experiment
    - 4 sites (Naples, Bari, Catania and Cosenza) in total larger than the Tier1
    - Project pending approval (end of 2011)

Building blocks of GEMSS

- **Disk-centric system** with five fundamental components
  1. **GPFS**: disk-storage software infrastructure
  2. **TSM**: tape management syste
  3. **StoRM**: SRM serv
  4. **GEMSS**: StoRM-GPFS-TSM inte
  5. **GridFTP**: WAN data transf



- Currently the Italian Grid is supporting mainly the fundamental physics community; however there are other VOs that are quite actively exploiting the Grid (Chemists, Biologists, ...)
- however, even within the INFN community, research groups have often found difficult to use the Gird and the acceptance of the Grid paradigm has proceeded rather slowly, despite the non negligible amount of available resources;
- part of the problem has been represented by well known aspects of the Grid that have been felt as obstacles sometime difficult to be overcome (X509 auth., VO bureaucracy, hw and sw constraint, etc.) but I believe that more generally, and especially for small and medium size groups that cannot count on their own dedicated IT manpower, the complexity of the distributed environment and the lack of common tools to manage it (e.g.: global bookkeeping tools, job monitoring and management toos, etc.) have made it difficult to enlarge the ensemble of those who exploit efficiently the Grid

The Italian Grid Infrastructure (II)

- The European projects EMI and EGI Inspire will be soon migrated to IGI
- Work is well under way to prepare the Statute of the IGI Consortium that will enable it to become an independent entity
- IGI is not aimed only at supporting the usage of computing resources for data intensive distributed scientific computing, but has been proposed as a general infrastructure to be exploited also by sectors of the Italian Public Administration like:
  - School, University and education
  - Public Health sector
  - Justice administration
  - Local communities

The Italian Grid Infrastructure (III)

- the IGI staff is projected to grow up to ~75 people with a significant fraction of the manpower allocated to support new communities and to promote the developments needed to satisfy their requirements
- IGI is also expected to participate in national and international R&D and infrastructural projects that could make it possible to secure the resources needed for setting up advanced services for a wide spectrum of users
  - A first example is the project for building the new distributed computing centers for the SuperB experiment
    - 4 sites (Naples, Bari, Catania and Cosenza) >> Tier1
    - Project pending approval (end of 2011)

PON

- recently the Ministry of University and Research has approved the National Research Plan which includes the SuperB project, i.e., the construction of a high-luminosity electron-positron collider to be used initially for basic HEP research and subsequently as a synchrotron light source;
- computing resources for the experimental activities are expected to be financed using funds earmarked for the development of infrastructures in the southern part of Italy.

PON (II)

- a project for building three large computing centers in Napoli, Bari and Catania plus a smaller one in Cosenza has been submitted recently and there are good chances that the proposal will receive green light before the end of the year;
- the project, with a budget of more than 40 MEuro, will address not only the build-up of the computing centers, but also the development of the core software and middleware that will be required to give the exploit the computing resources

What we can offer today

- Contributions to the community in terms of our own requirements
  - Hopefully considered in plans for future European service offerings
- Visibility in our efforts to build integrated Grid and Cloud services within the Italian Grid Initiative (IGI)
- Access to our development platforms and expertise for what regards, for example, large storage systems (StoRM, GPFS, Lustre), virtualization and integration with Grid and Cloud computing (WNoDeS)
- We will strive for reinforcing direct links between big European academic resource centers

- **For more information**: XXX