The Italian Grid Infrastructure (IGI)
CGW08 Cracow

Mirco Mazzucato
Italian Grid Infrastructure Coordinator
INFN CNAF Director

EGEE -> EGI

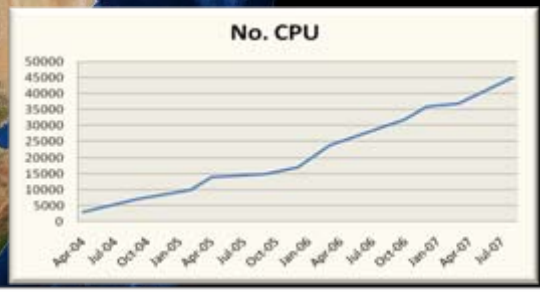
IGI is a currently a JRU of EGEE which operate the world largest production quality grid infrastructure for e-Science

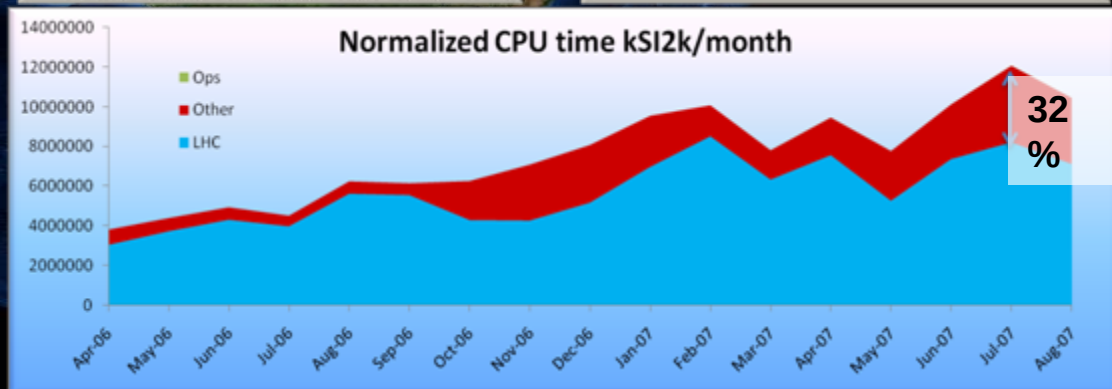
250 sites (IGI:40)
45 countries
90,000 Cores
15 PetaBytes
>5000 users
>100 VOs
>100,000 jobs/day

IGI ->
National Grid Initiative

Applications:

Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences
...





The future evolution

- **At National level IGI is making the steps to:**
- **become a legal recognized national organization (NGI) acting as a single national point-of-contact for e-Research**
- **Take over the operation of the national e-Infrastructure**
- **represent and support all national Research user communities and resource providers (application independent)**
- **In coordination at European Level with:**
- **The European Grid Initiative (EGI)**
  - EGI.Org Central Organization
  - National Grid Initiatives (NGI)
- **Prepare permanent, common production EU Grid infrastructure**
- **Coordinate the integration between National Grid Infrastructures (NGIs)**
- **Operate the production Grid infrastructure on a European level for a wide range of scientific disciplines**

IGI today

- The talian Grid Infrastructure (IGI) is
  - an EU Joint Research Unit
  - Based on a MoU signed in December 2007
  - Recognized and supported by the Italian Ministry of the University and Research
  - Recognized by the EU Commission
  - Providing a unique International interface for what concern the Italian grid infrastructure
  - Providing a common coordination of the Italian Grid infrastructure for e-Research by public Institutions
  - Open to new partners

IGI partners

- Istituto Nazionale di Fisica Nucleare (INFN),
- Ente per le Nuove tecnologie, l'Energia e l'Ambiente (ENEA),
- Consiglio Nazionale delle Ricerche (CNR)
- Istituto Nazionale di Astrofisica (INAF),
- Istituto Nazionale di Geofisica e Vulcanologia (INGV),
- Università degli Studi di Napoli Federico II,
- Università degli Studi della Calabria,
- Sincrotrone Trieste S.C.p.A. (ELETTRA),
- Consorzio COMETA,
- Consorzio COSMOLAB,
- Consorzio SPACI,,
- Consortium GARR
- L'Universita' di Perugia
- L'Universita' del Piemonte Orientale
- …..ongoing discussions with Compunting Centres

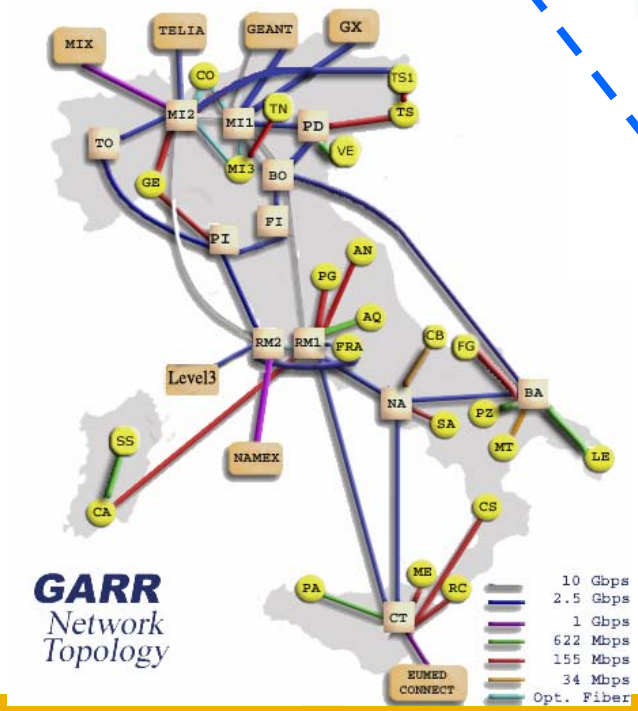

EGI
European GRID Infrastructure
IGI = A sum of grids
Italian GRID Infrastructure
Several Universities
Campus Grid
Computing Centers
INFN-GRID
ENEA-GRID
Roma
CRESCO
Progetto
SCOPE
BRINDISI
LECCE
TRISAIA
CYBERSAR
GRISU'
SPACI
COMETA
PI2S2

The Italian e-Infrastructure

+Several Universities
+ Compute Centers

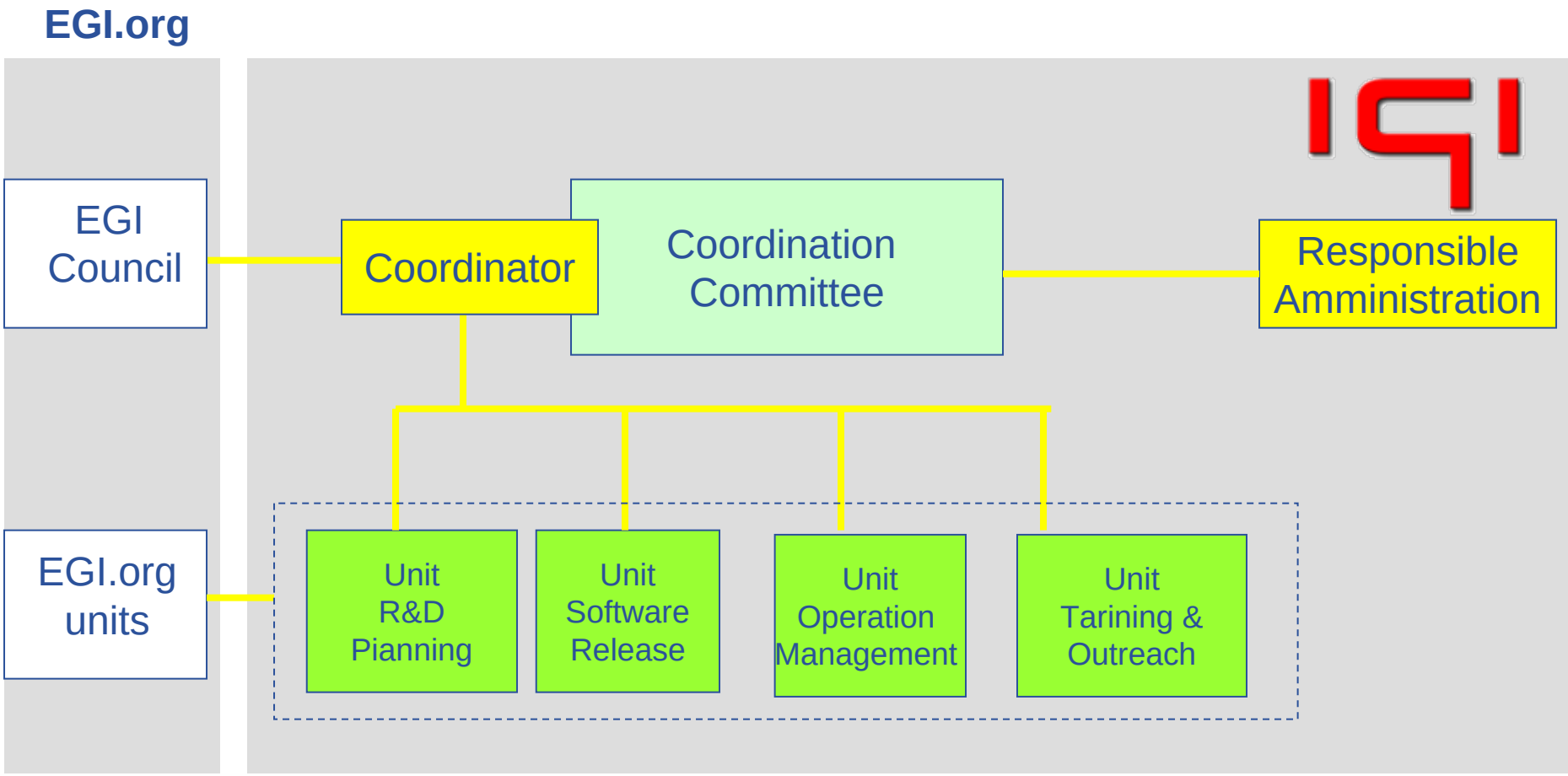
INFN production GRID

GARR Network



The IT Grid

IGI: Initial Managerial structure



Strategy

- Keep initially the structure as simple as possble to make it sustainable
- Allow representation of all active Institutions both at technical and managerial level
- Leverage from what already exists (from EGEE) and adapt to a broader scope
- 4 Units with technical responsibilities:
- 1 Coordination Committee
  - Coordinator
  - Responsible of the administration

IGI Units: taks and services (1/2)

- **Unit Operation Management (Infrastructure and Grid Services)**
  - Management of Grid central services and *catch-all* for production infrastructure and test
  - Planning of infrastructure and interoperability
  - Help desk for users site manager
  - Monitoring, accounting
  - Management of other tools and services supporting opertions
  - Grid infrastructure management
  - Grid security (security and incident response coordination in the region)
  - Contribution to EGI operations, technical collaboration with EGI.org (Unit Operations)
- **Unit software Release (Release mw and certification)**
  - Adapt UMD (Unified Middleware Distribution) to national needs , add required missing components
  - Integration and certification of missing components componenti aggiuntive
  - Technical Collaboration with EGI.org (Unit Developments)

IGI Units (2/2)

- **Unit Planning (Planning, R&D):**
  - *Scope:* verify user requirements and plan grid infrastructure enhacement
  - *Partecipants*: 1 rappresentative per IGI Institution
  - Promote partecipations to R&D national and international Calls
  - Plan common R&D activitites (mw and management tools)
- **Unit Training and Application support (Applications and Training)**
  - Support to porting of applications
  - Collect new requirements and user feedback
  - Training for users and site managers
  - Dissemination and contacts with industrial partners
  - → All IGI Units collaborate and coordinate their activities with EGI.org to get common things done with a common unique effort

IGI tasks and issues

Ongoing merge from INFN + Other Grids -> IGI :

- **Provide common services for e-Research**
- **Common middleware releases**
- **General Guidelines**
- **Common management of the Italian Grid infrastructure**
- **Initial set of supported services:**
  - National CA
  - Provision and support of the Grid portal
  - Management of Grid Services (also for applications)
  - Monitoring tool
  - Accounting tool
  - Management and control infrastructure
  - Users support

The middleware issue in Europe

- **EGEE Grid: focus on production quality**
  - gLite based, strong hierarchical model. Include Globus/ Condor
  - Guarantee stable and high quality services
- **NorduGrid**
  - ARC based, many commonalities with EGEE, more flexible model
- **DEISA as the supercomputing Grid**
  - UNICORE based, less overlap with ARC/gLite
- **Smaller Grid islands**
  - Mostly Globus and Condor based
  - Default choice for general grid R&D publications and prototypes
  - Some serving specific application communities
  - Includes also activities of some European countries
- **Commercial products**
  - Usually single components, basic functions, limited use in academic world
  - Cloud: Amazon S2 and S3 services still not used by Science

The EGI - IGI choice

- **Ian-Carl definition in the Anatomy of the Grid:**
- ***Grid : set of services which enable dynamic ICT resource sharing (CPU, Storage, Data, Archives, Instruments…) in multi-administrative domains for the benefit of multi-institutional virtual organizations***
- **Did we practically achieve this after 8 years?**
- **Not yet!!!!**
- **->A set of different implementations not inter-operating each other**
  - Globus, Condor, gLite, Unicore, Arc, Naregi….
- **A shared large scale production infrastructure like IGI or at larger EGI could be theoretically achieved**
  - Trying to interconnect “free” individual Grid islands – however no proofs that scalability and quality problem can be solved – can not remove barriers (lack of widely accepted standards) for building easily global collaboration
  - Accepting a unified “standard” compliant middleware certified solution – no barriers, quality and scalability at the price of less flexibility, some danger of middleware o(g)ssification
- ->**Problems pushed at the borders but EGEE is there to show this strategy can be successful**
- **IGI has certainly valued the high quality and no barrier the best choice in line with EGEE**
- **Put effort on making the grid access simpler**
- **IGI is ready to support an EU effort to improve standards and interoperability**
- **Will support a Unified Middleware Distribution**

Guidelines for IGI: Operations Model

- Aims
  - Autonomy → using a common infrastructure for local and international work for greater efficiency
  - Sustainability: Pass from project activities to useful services paid by Reasearch Institutions and Resource Provders
  - Subsidiarity: do things at as local a level as possible (EGI.org pulls things together)
  - Increased reliability through pushing responsibility down to sites
  - Preserving current scalability in presence of more middleware stacks to be supported, more non-EGEE Grids integrated, ...

EGI Guidelines for NGIs

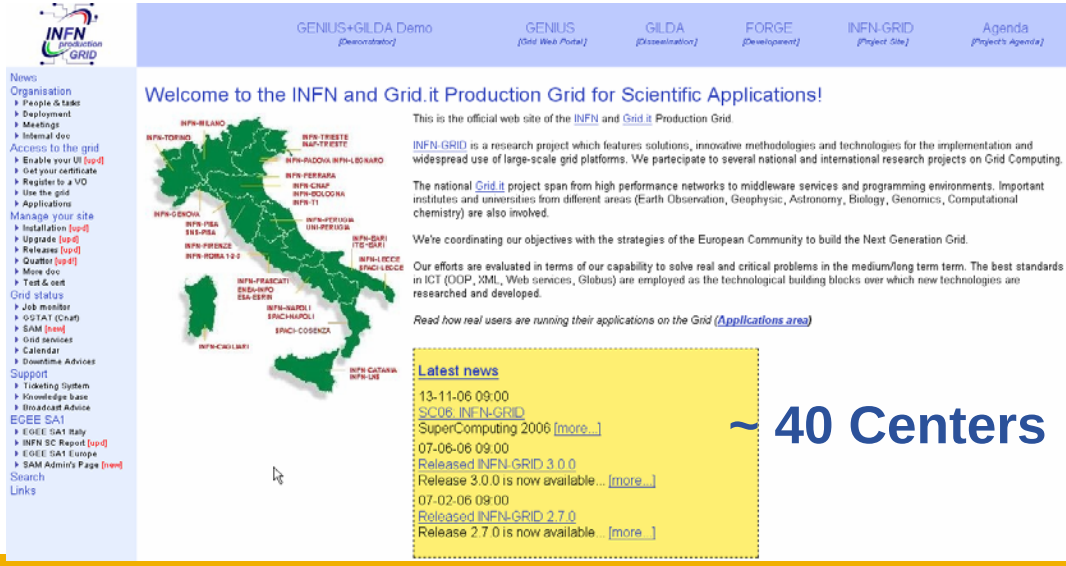
- **NGIs are expected to:**
  - operate in autonomy secure Grid infrastructures in the countries
  - coordinate Grid operations in the countries
  - collaborate to the definition of common operational procedures, policies, standards/specifications
  - adhere to standards/specifications to ensure interoperability
  - support users and operational problems
- **Very much in line with IGI strategy**

Established GRID Management

Overall Grid management is essential to guarantee quality. Well consolidated practice in Italy from INFN

Performed by the Italian Regional Operation Center (ROC). The main activities are:

- **Production of Italian MW release and test**
- **Deployment of the release to the sites, support to local administrators and sites certification**
- **Periodical check of the resources and services status**
- **Support at the Italian level**
- **Support at the European level**
- **Introduction of new Italian sites in the grid**
- **Introduction of new regional VOs in the grid**
- **Support to experimental services**



~ 40 Centers

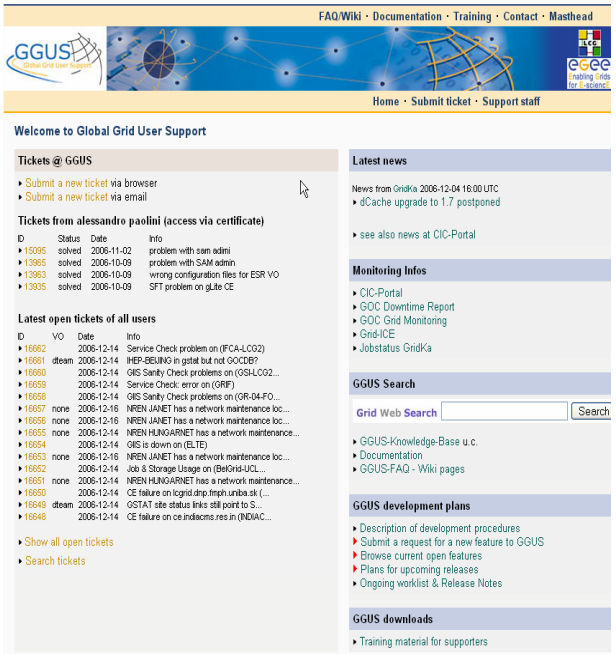
Experimental Services

- **Required to allow tests of new components at a sufficient scale before final releases**
- **Jointly made by developers, operations and users**
  - Normally run tests based on real user applications and real user feedback
  - Fast cycle: feedback, correction patch, deployment, new feedback

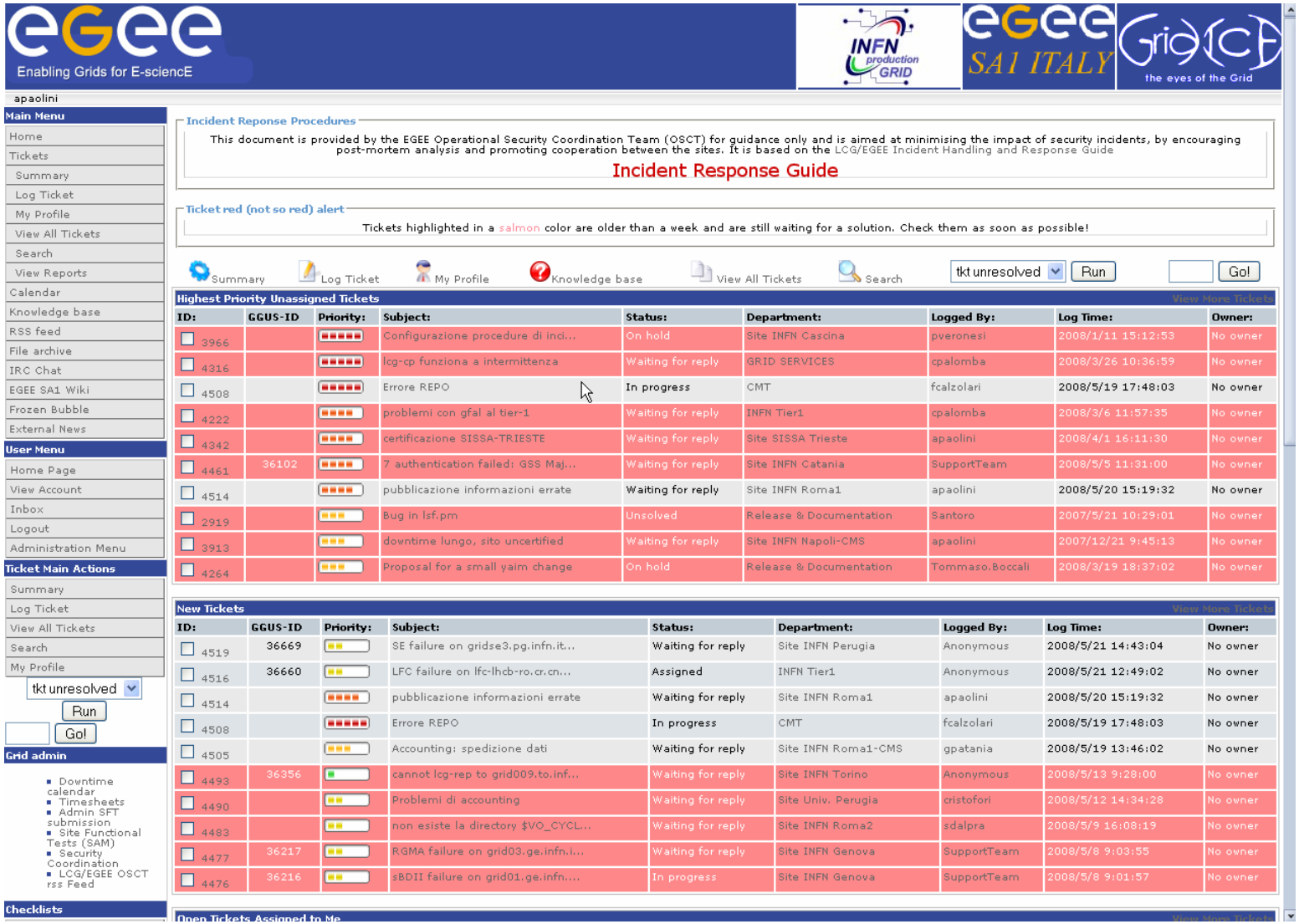
Users and sites support

- The Italian Regional Operation Centre ticketing system based on XOOPS/xHelp
- Integrated with EGEE GGUS (Global Grid UserSupport) ticketing system
- Web services are used to:
- Transfer tickets from the global to national system
- Transfer tickets from the national to the global system

- The user support group handles the tickets according to addressing
  - At national level when appropriate
  - Either send them to GGUS



The support system



Access to the GRID

- **Via User Interface (UI). Could be:**
  - A dedicated server, installed in a similar way as the others grid elements
  - UI Plug-and-Play (UI PnP), a software one can install on any pc without root privilegies
- **A web portal: https://genius.ct.infn.it/**

The General Purpose Services

- Resource Brokers
- Logging & Bookeeping
- Top Level BDII
- voms servers + replicas
- gridice server
- LCG File Catalog server
- Server MyProxy
- FTS server
- SRM Data storage service

Supported VOs

VO Name ▼	Q#	MinSlotFree	MaxSlotFree	RunJob	WaitJob	Storage Available	Storage Total	Storage Load
alice	24	625	625	203	253	154.1 TB	266.5 TB	42%
argo	20	365	365	-	-	147.2 TB	253.1 TB	42%
astro	6	78	78	-	-	-	-	-
atlas	34	881	881	10	167	3.2 PB	3.8 PB	17%
auger	7	157	157	-	-	357.1 GB	357.1 GB	0%
babar	23	396	396	173	0	59.9 TB	121.1 TB	50%
bio	18	203	203	0	82	59.5 TB	118.9 TB	50%
biomed	20	222	222	360	1778	61.8 TB	122.7 TB	50%
cdf	19	630	630	303	0	60 TB	120.6 TB	50%
cms	34	700	720	413	1187	188.7 TB	346.5 TB	45%
compassit	9	138	138	-	-	30.6 TB	74.1 TB	59%
compchem	26	289	289	2	89	59.4 TB	118.7 TB	50%
cyclops	12	165	165	-	-	38.4 TB	80.9 TB	52%
diligent	0	-	-	-	-	-	-	-
dteam	45	1150	1150	13	80	250.2 TB	407 TB	39%
edteam	4	45	45	-	-	21.2 TB	37.6 TB	44%
eela	1	45	45	-	-	21.2 TB	37.6 TB	44%
egee	2	78	78	-	-	-	-	-
egrid	21	273	273	-	-	59.8 TB	119.2 TB	50%
embrace	1	2	2	-	-	5.3 GB	58.7 GB	91%
enea	23	271					118.7 TB	50%
enmr	8	93					-	-
esr	15	217					118.4 TB	50%
euchina	10	194					39.4 TB	42%
euindia	14	348					50.3 TB	39%
eumed	5	155					39.4 TB	42%
fusion	7	78					357.1 GB	0%
geant4	11	150					37.6 TB	44%
gear	10	142					37.6 TB	44%
gilda	0	-					-	-
gridit	27	477					119.2 TB	50%
ildg	0	-					8.8 GB	83%
inaf	17	201					118.8 TB	50%
infngrid	43	1150					127.6 TB	50%
ingv	14	193					118.3 TB	50%
lhcb	30	686	686	228	82	247.9 TB	404.3 TB	39%
libi	13	189	189	-	-	51.6 TB	109.8 TB	52%
lights	14	171	171	-	-	-	-	-
magic	15	235	235	-	-	59 TB	118.4 TB	50%
na48	6	178	178	-	-	399.3 GB	2.2 TB	82%
omiieurope	1	44	44	-	-	0	0	0%
ops	41	1150	1150	76	230	250.2 TB	407 TB	39%
pamela	18	341	341	126	11	53.6 TB	112.7 TB	52%

49 VOs supported:
4 LHC (ALICE, ATLAS, CMS, LHCB)
3 test (DTEAM, OPS, INFNGRID)
20 Regional
1 catch all VO: GRIDIT
21 Other VOs

24

Full Accounting using DGAS

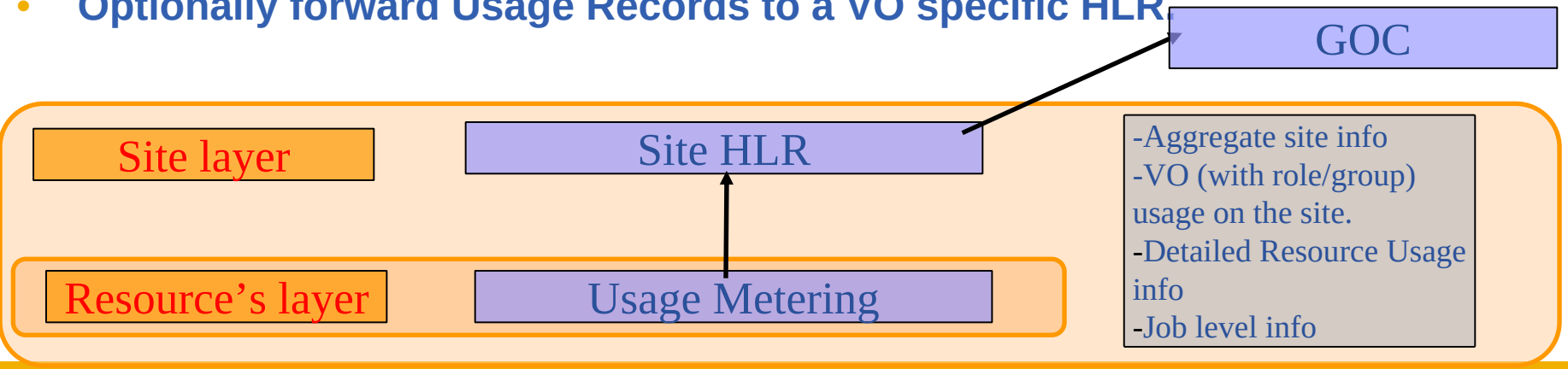
DGAS (Distributed Grid Accounting System) is fully deployed in INFNGrid (13 site (Home Location Registers) HLRs + 1 HLR of 2nd level (summaries).

The site HLR is a service designed to manage a set of 'accounts' for the Computing Elements of a given computing site.

For each job executed on a Computing Element (or a on local queue), the Usage Record for that job is stored on the database of the site HLR.

Each site HLR can:

- **Receive Usage Records from the registered Computing Elements.**
- **Answer to site manager queries such as:**
  - Datailed job list queries (with many search keys: per user, VO, FQAN ,CEId…)
  - Aggregate usage reports, such as per hour, day, month…, with flexible search criteria.
- **Optionally forward Usage Records to EGEE database.**
- **Optionally forward Usage Records to a VO specific HLR.**



Conclusions

- In Italy different grids, single centers and Research Institutions are joining in the National Grid Infrastructure (NGI) called IGI
- Following a similar process at EU level which will bring to the European Grid Infrastructure i.e.
- a central EU organization: EGI.org
- NGIs in each country
- IGI is well advanced in offering all services requested to an NGI
- Will look progressively for more sustainability introducing service charges for more consolidated services