

EGEE

TRAINING PROGRESS REPORT

WITH UPDATES AT M15 AND M24

EU DELIVERABLE: DNA3.3.1

Document identifier: **EGEE-DNA3.3.1-533762-v0-2.doc**

Date: **12/12/04**

Activity: **NA3: User Training and Induction**

Lead Partner: **UEDIN**

Document status: **FINAL**

Document link: <https://edms.cern.ch/document/533762/1>

Abstract:

The purpose of this document is to report on the progress of the training activity for the first 9 months of the project. The EGEE training is the responsibility of the 22 partners of NA3, led by UEDIN, working in conjunction with all other EGEE partners and activities. The first 9 months of this training have been very successful, exceeding planned targets by substantial margins. Overall, 1033 people have attended 39 training events that have delivered a total of 15,797 participant days. This breaks down into 11,584 participant days at induction events, 286 participant days at developer courses, 3,732 at advanced courses and 481 at workshops. We measure the views of the participants using a standard QA form, and the results show we are always in the upper half of scores and also show a positive trend of improvement.

This Report will be updated again at M15 of the EGEE project.

Copyright (c) Members of the EGEE Collaboration. 2004.

See <http://public.eu-egee.org/partners/> for details on the copyright holders.

EGEE (“Enabling Grids for E-science in Europe”) is a project funded by the European Union. For more information on the project, its partners and contributors please see <http://www.eu-egee.org>.

You are permitted to copy and distribute verbatim copies of this document containing this copyright notice, but modifying this document is not allowed. You are permitted to copy this document in whole or in part into other documents if you attach the following reference to the copied elements: "Copyright (C) 2004. Members of the EGEE Collaboration. <http://www.eu-egee.org>".

The information contained in this document represents the views of EGEE as of the date they are published. EGEE does not guarantee that any information contained herein is error-free, or up to date.

EGEE MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, BY PUBLISHING THIS DOCUMENT.

Delivery Slip

	Name	Partner/Activity	Date	Signature
From	David Fergusson	UEDIN/NA3	12/12/04	
Reviewed by	Moderator and reviewers			
Approved by	PEB			

Document Log

Issue	Date	Comment	Author
0-1	12/12/04	First draft	David Fergusson, Elena Slabospitskaya, Peter Kacsuk, Christos Aposkitis
0-2	10/1/05	Document updated to take note of reviewers comments and new information from partners.	David Fergusson, Elena Slabospitskaya, Peter, Christos Aposkitis, Steve Fisher, Martin Polak, Licia Florio, Malcolm Atkinson, Roberto Barbera, Wojtek Wislicki, Pawel Wolniewicz, Ludek Matyska, Miroslav Dobrucky, Ruediger Berlich

Document Change Record

Issue	Item	Reason for Change
-------	------	-------------------

0-2	Remove Appendices	comments by reviewers
0-2	Change “..a collection of tutorials available with GENIUS..” with “..a collection of tutorials available with GILDA..”. In fact, the t-Infrastructure is GILDA while GENIUS is only one of its tools.	comments by reviewers
0-2	Change “The version of e-Infrastructure may correspond...” with “The version of t-Infrastructure may correspond...”	comments by reviewers
0-2	Change “See: https://genius.ct.infn.it ” with “See: https://grid-demo.ct.infn.it ”	comments by reviewers
0-2	Change “...using the KNOPPIX LINUX distribution.” with “...using the KNOPPIX LINUX distribution (https://gilda.ct.infn.it/live-cd/gilda-live-ui.iso).”	comments by reviewers
0-2	Please cite the new GILDA tutorial request web form available at https://gilda.ct.infn.it/request_tutorial/ .	comments by reviewers
0-2	Change “(+Turkey within the period to end of December 2004)” with “(plus Turkey and Israel within the period to end of December 2004)”	comments by reviewers
0-2	Change “Spain, UK and Switzerland” with “Spain, UK, Switzerland and Venezuela”	comments by reviewers
0-2	Change the section number from 4.1.1 to 4.4.1	comments by reviewers
0-2	Table 6 does not contains the Quality Assessment columns while this is cited in the caption	comments by reviewers

0-2	Why language is not considered as one of the metadata describing a presentation ?	comments by reviewers
0-2	At the end of the section a discussion about the need of a digital library for EGEE training material (come up at the last EGEE Conference) is missing	comments by reviewers
0-2	Change “..a collection of tutorials available with GENIUS..” with “..a collection of tutorials available with GILDA..”. In fact, the t-Infrastructure is GILDA while GENIUS is only one of its tools.	comments by reviewers
0-2	The last paragraph: the URL of the NA3 web site should be indicated	comments by reviewers
0-2	Figure 3 is not readable	comments by reviewers
0-2	NA3 web portal address is missing	comments by reviewers

CONTENT

1. INTRODUCTION.....	7
1.1. PURPOSE OF THE DOCUMENT	7
1.2. APPLICATION AREA	7
1.3. REFERENCES	7
1.4. DOCUMENT AMENDMENT PROCEDURE.....	7
1.5. TERMINOLOGY.....	7
2. EXECUTIVE SUMMARY.....	11
3. OVERVIEW	12
3.1. TRAINING IN EGEE	12
3.2. QUALITY ASSESSMENT	14
3.3. ACTIVE PARTNERS	15
3.4. COOPERATION IN EGEE	16
3.5. TRAINING RESOURCES	16
3.5.1. T-Infrastructure	17
3.5.2. Web portals	17
3.5.3. GENIUS/GILDA.....	17
3.5.4. NA3 web portal.....	18
3.5.5. eLearning	18
4. TYPES OF COURSE.....	19
4.1. INDUCTION	19
4.1.1. Delivery.....	19
4.1.2. Locations	19
4.1.3. Numbers	19
4.1.4. Material.....	19
4.1.5. Current and projected demand.....	20
4.2. APPLICATION DEVELOPER	20
4.2.1. Delivery.....	20
4.2.2. Locations	21
4.2.3. Numbers	21
4.2.4. Material.....	21
4.2.5. Current and projected demand.....	21
4.3. ADVANCED.....	22
4.3.1. Delivery.....	22
4.3.2. Locations	22
4.3.3. Numbers	22
4.3.4. Material.....	22
4.3.5. Current and projected demand.....	22
4.4. ADVANCED RETREATS (WORKSHOPS).....	23
4.4.1. Delivery.....	23
4.4.2. Locations	23
4.4.3. Numbers	23
4.4.4. Material.....	23
4.4.5. Current and projected demand.....	23
5. MATERIALS ARCHIVE.....	24
5.1. DESIGN	24
5.2. IMPLEMENTATION.....	27
5.3. ISSUES	28
5.3.1. Copyright.....	28

5.3.2. Maintenance 28

6. CONCLUSION..... 29

1. INTRODUCTION

1.1. PURPOSE OF THE DOCUMENT

The purpose of this document is to review the performance of the training activity NA3 over the first 9 months of the EGEE project.

1.2. APPLICATION AREA

The application area of this document is the training activity NA3.

1.3. REFERENCES

[This subsection provides a complete list of all documents referenced elsewhere in the document.]

[R1]	EGEE Technical Annex (TA) https://edms.cern.ch/document/400278
[R2]	NA3 Execution Plan https://edms.cern.ch/document/461605
[R3]	EGEE Induction & Training: Planning an Event https://edms.cern.ch/document/475027
[R4]	EGEE Induction & Training: Running an Event https://edms.cern.ch/document/475028
[R5]	NA3 Activity Quality Assurance Plan https://edms.cern.ch/document/483910

1.4. DOCUMENT AMENDMENT PROCEDURE

Requests to amend the Training Plan must be made to the main author (David Fergusson, email dfmac@nesc.ac.uk) in the first instance.

If the amendment relates to specific training or induction events, appropriate revisions will be made to the on-line tables of events in the NA3 Intranet pages (see <http://www.egee.nesc.ac.uk/schedreg/>), and the EGEE external website; these revisions will also be publicised as appropriate through the normal channels for advertising training, induction and other events. At present these channels consist of the training pages on the external EGEE site (<http://public.eu-egee.org/>), the home page in the internal EGEE site (<http://egee-intranet.web.cern.ch/egee-intranet/>), and the NA3 pages in the internal EGEE site (<http://www.egee.nesc.ac.uk/>). Since such changes *will* occur, we recommend referring to the on-line tables at <http://www.egee.nesc.ac.uk/schedreg/future.html> for the latest information. Other requested changes to the text of this document will be recorded, and will be taken into account prior to the document's submission to the EGEE internal review process or to the EU as deliverable DNA3.1.1; or else the requested change will be taken into account when preparing the next update of the document, at month M9 or M15 of the Project.

1.5. TERMINOLOGY

Glossary

Ant	Ant tool from the Apache Software Foundation that automates packaging of class files.	Ant tool from the Apache Software Foundation that automates packaging of class files.
API	Application Programmer Interface	Application Programmer Interface
AUEB	AUEB University of Economics and Business, Greece	AUEB University of Economics and Business, Greece
BGConsortium	Bulgarian Grid Consortium	Bulgarian Grid Consortium
BUTE	Central Europe Federation	Central Europe Federation
CEGTO	CEGTO European Grid Training Organisation	CEGTO European Grid Training Organisation
CERN	European Organization for Nuclear Research, Geneva, Switzerland	European Organization for Nuclear Research, Geneva, Switzerland
CESNET	CESNET, partner 4: CESNET, Prague, Czech Republic;	CESNET, partner 4: CESNET, Prague, Czech Republic;
CIS	Commonwealth of Independent States, the 12 Independent States, of the former Soviet republics	Commonwealth of Independent States, the 12 Independent States, of the former Soviet republics
CYFRONET	CYFRONET Akademia Komputerowa, Krakow, Poland	CYFRONET Akademia Komputerowa, Krakow, Poland
ELUB	ELUB partner 6: Eotvos Lorand University, Budapest, Hungary	ELUB partner 6: Eotvos Lorand University, Budapest, Hungary
EMI	EGEE Member Induction	EGEE Member Induction
EPCC	Edinburgh Parallel Computing Centre	Edinburgh Parallel Computing Centre
ERA	European Research Area	European Research Area
EU	European Union	European Union
FP6	Sixth Framework Programme of the EU	Sixth Framework Programme of the EU
FZK	FZK, partner 28: Forschungszentrum Karlsruhe, Germany	FZK, partner 28: Forschungszentrum Karlsruhe, Germany
GAI	Grid-Aware Induction	Grid-Aware Induction
GENIUS	GENIUS portal made available by EGEE by training	GENIUS portal made available by EGEE by training
GILDA	GILDA testbed made available by EGEE by training	GILDA testbed made available by EGEE by training
GridPP	GridPP Collaboration of Particle Physicists from CERN, who are building a Grid for Particle Physics	GridPP Collaboration of Particle Physicists from CERN, who are building a Grid for Particle Physics
GRNET	GRNET partner 51: Greek Research and Technology Network, South-East Europe Federation	GRNET partner 51: Greek Research and Technology Network, South-East Europe Federation
GT3	Grid middleware	Grid middleware
GUP	GUP, partner 2: Institut Johannes Kepler, Central Europe Federation	GUP, partner 2: Institut Johannes Kepler, Central Europe Federation
IAG	Israel Academic Grid	Israel Academic Grid
ICI	ICI, partner 53: National Institute for Research and Development, Bucharest, Romania; South-East Europe Federation	ICI, partner 53: National Institute for Research and Development, Bucharest, Romania; South-East Europe Federation
ICM	ICM, partner 11: Warsaw University of Technology, Central Europe Federation	ICM, partner 11: Warsaw University of Technology, Central Europe Federation
IG-BIGEST	IG-BIGEST Grid for Business, Industry & Geology, E-science & ...	IG-BIGEST Grid for Business, Industry & Geology, E-science & ...
IHEP	IHEP, partner 41: Institute of High Energy Physics, Beijing, China	IHEP, partner 41: Institute of High Energy Physics, Beijing, China
INFSO-RI-508833	INFSO-RI-508833	INFSO-RI-508833
II-SAS	II-SAS, partner 13: Ustav Slovenskej Akadémie vied, Bratislava, Slovakia; Central Europe Federation	II-SAS, partner 13: Ustav Slovenskej Akadémie vied, Bratislava, Slovakia; Central Europe Federation
IMPB RAS	IMPB RAS, partner 42: Institute of Mathematics of Biological Sciences of Russian Academy of Sciences, Moscow, Russia	IMPB RAS, partner 42: Institute of Mathematics of Biological Sciences of Russian Academy of Sciences, Moscow, Russia
INFN	INFN partner 31: Istituto Nazionale di Fisica Nucleare, Italy	INFN partner 31: Istituto Nazionale di Fisica Nucleare, Italy

IHEP	NA3, partner 41: Institute of High Energy Physics, Protvino Moscow Region, Russia; Russia Federation
II-SAS	NA3, partner 13: Ustav Informatiky, Slovenska Akademia vied, Bratislava, Slovakia; Central Europe Federation
IMPB RAS	NA3, partner 42: Institute of Mathematical Problems of Biology of Russian Academy of Sciences, Puschino, Moscow Region, Russia; Russia Federation
INFN	NA3, partner 31: Istituto Nazionale di Fisica Nucleare, Italy; Italy Federation
IITEP	NA3, partner 43: Institute of Theoretical and Experimental Physics, Moscow, Russia; Russia Federation
ITI-CERTH	Information and Telematics Institute of the Center for Research and Technology Hellas, Greece
IUCC	Israel's NREN
JINR	NA3, partner 44: Joint Institute for Nuclear Research, Dubna, Russia; Russia Federation
JRA1	EGEE Joint Research Activity 1: Middleware Development Activity
JRA2	EGEE Joint Research Activity 2: Quality Assurance Activity
KU-NATFAK	NA3, partner 34: Faculty of Science, University of Copenhagen, Denmark; Northern Europe Federation
MTA SZTAKI	NA3, partner 8: Magyar Tudomanyos Akademia Szamiastecnikai es Automatizalasi Kutato Intezet, Budapest, Hungary; Central Europe Federation
NA2	EGEE Networking Activity 2: Dissemination and Outreach Activity
NA3	EGEE Networking Activity 3: User Training and Induction Activity
NA4	EGEE Networking Activity 4: Application Identification and Support Activity
NEG	Northern European Grid
NeSC	UK National e-Science Centre, Edinburgh, Scotland
NGUI	Non-Grid User Induction
NICE	CERN computing service
NOA	National Observatory of Athens, Greece
NREN	National Research and Education Network
NTUA	National Technical University of Athens, Greece
PEB	EGEE Project Executive Board
PNPI	NA3, partner 46: Petersburg Nuclear Physics Institute of Russian Academy of Sciences, Gatchina, Leningrad District, Russia; Russia Federation
RDIG	Russian Data Intensive Grid – Russia federation in EGEE
RoGrid	Romanian Grid consortium
RRC KI	NA3, partner 47: Russian Research Centre "Kurchatov Institute", Moscow, Russia; Russia Federation
SA1	EGEE Service Activity 1: European Grid Support, Operation and Management
SEE	South-East Europe: an EGEE federation including Bulgaria, Cyprus, Greece, Israel and Romania

SOAP	Simple Object Access Protocol
TA	EGEE Technical Annex
TAU	NA3, partner 52: Tel Aviv University, Tel Aviv, Israel; South-East Europe Federation
TCR	NA3 Training Component Repository
UEDIN	NA3, partner 17: University of Edinburgh, Scotland; UK & Ireland Federation; lead partner for NA3
UNINNSBRUCK	NA3, partner 3: Institut für Informatik der Universitaet Innsbruck, Austria; Central Europe Federation
UoC	University of Crete at Heraklion, Greece
UoI	University of Ioannina, Greece
UoPatras	University of Patras, Greece
UoPiraeus	University of Piraeus, Greece
WSDL	Web Services Description Language
XML	Extensible Markup Language, the foundation on which Web Services are built

2. EXECUTIVE SUMMARY

The EGEE training is the responsibility of the 22 partners of NA3, led by University of Edinburgh, working in conjunction with all other EGEE partners and activities. The first 9 months of this training have been very successful, exceeding planned targets by substantial margins. Overall, 1033 people have attended 39 training events that have delivered a total of 15,797 participant days. This breaks down into 11,584 participant days at induction events, 286 participant days at developer courses, 3,732 at advanced courses and 481 at workshops. We measure the views of the participants using a standard QA form, and the results show we are always in the upper half of scores and also show a positive trend of improvement. Full details are presented below.

In every case except developer courses, this is well in excess of the planned targets for the first 9 months. This has been achieved because of the enthusiasm and hard work of members of EGEE drawn from many activities across all regions. The developer courses have taken longer to prepare, have required higher trainer to participant ratios and have more diversity than originally anticipated. The increasing functionality and resources of the operational EGEE systems and the growing diversity of user communities from more and more disciplines guarantee that demand for developer courses will significantly increase – it already exceeds supply. Therefore, in the next year increased effort and attention will be devoted to developer training. Our analysis of demand shows that all other aspects of training will continue to be in great demand – see the Revised Training Plan DNA3.1.2.

T-Infrastructure emulates the production e-Infrastructure but has additional properties that meet particular training and student needs – such as pre-production features, short response times, light weight authorisation and convenient portals. We have recognised the vital importance of providing sufficient good quality t-Infrastructure. Due particularly to the outstanding commitment of Roberto Barbera and the INFN resources dedicated to the initial work, EGEE can boast that it has pioneered exceptionally good t-Infrastructure for introductory, developer and advanced courses requiring grids. This is a valuable collaboration between NA3 and NA4.

An unexpected requirement has been the large demand for training courses in installing and running LCG2 services. As more and more sites join in the provision of grid mediated resources, there are more staff who have to set up and operate EGEE infrastructure. Many of the operational difficulties are due to configuration and operational errors at new sites. Consequently there is not only high demand for systems administration training, it is crucial for EGEE that this training be given. Effective collaboration between SA1 and NA3 has delivered early versions of these courses. They require exclusive use of well provisioned t-Infrastructure as well as experts with experience of systems administration. This will require more investment as e-Infrastructure is rolled out to EGEE sites and the trained staff produced will be a vital underpinning of Europe's eventual e-Infrastructure.

To improve training efficiency and to support self-paced learning we are establishing a repository of training material at the NA3 web site and a collection of tutorials available with GILDA, the preferred t-Infrastructure for introductory courses. In conjunction with SA1, JRA1, NA2 and NA4 we have begun the development of integrated, user-oriented information about EGEE and its use.

3. OVERVIEW

3.1. TRAINING IN EGEE

Training is a central part of the EGEE project. While the technical thrust of the project is to put in place the e-Infrastructure required to support a production level grid throughout Europe, the success of the project will be measured by the take-up and usage of this grid within the broadest scientific community. In order to achieve this goal not only must the grid service exist but potential users must be made aware of it and be given the tools they require to use it as effectively as possible. This is the function of the EGEE training activity.

The first task of the training activity is to introduce users (and also those joining the EGEE organisation) to the technical and organisational basis of EGEE. To this end we must create, deliver and publish materials which adequately give an abstract view of EGEE but with enough technical detail to allow users to progress to more focussed view of what EGEE can deliver to them.

Following such a general introduction to the project, more specialist courses are required to maintain the momentum in developing productive use of the EGEE infrastructure and middleware. Beyond the introduction to the project training needs become more domain specific.

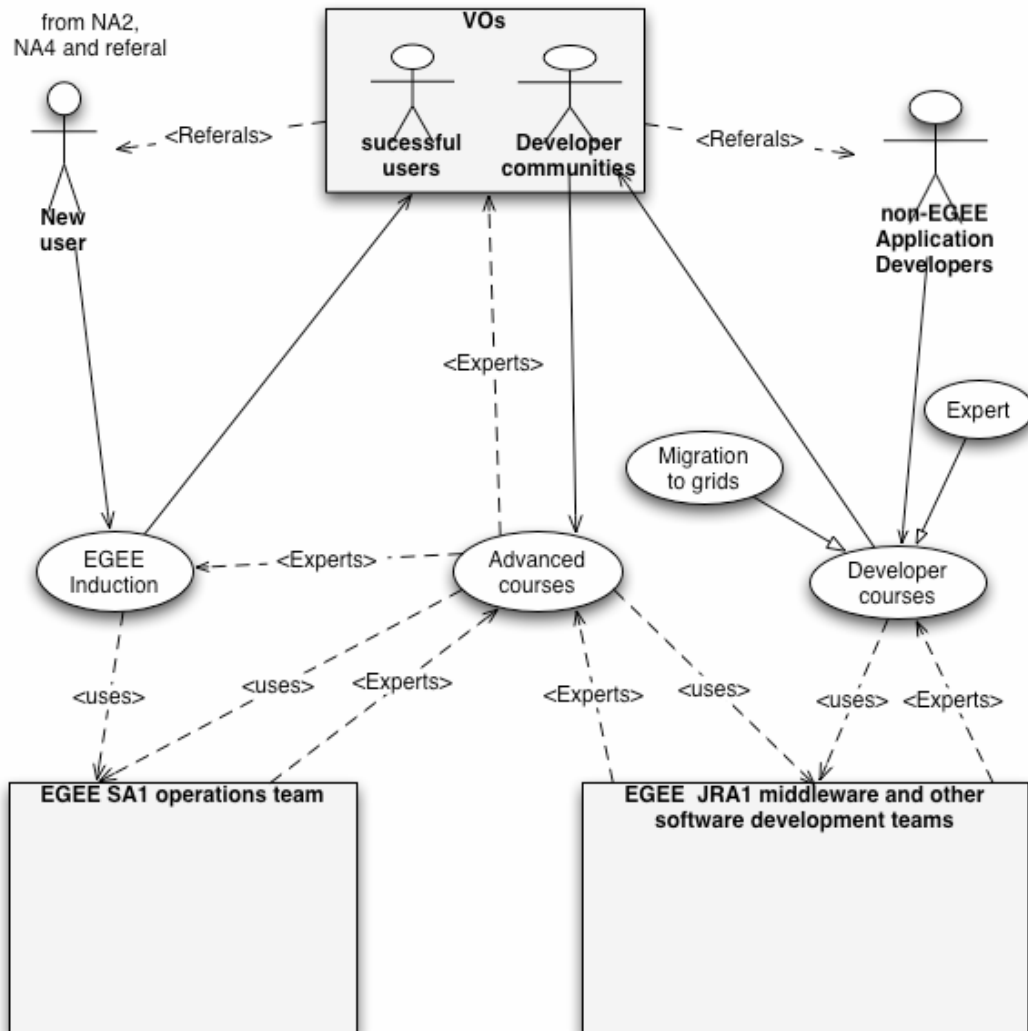


Figure 1: UML Use Case diagram of the relationships between course types, EGEE users and EGEE activities.

NB. As is customary with UML diagrams not all components and relations are explicitly shown, only those which are informative in context. For instance the Training (NA3) activity is omitted as it would be connected to all of the entities and therefore obscure the other relationships.

Once EGEE members are familiar with the services offered by the project and are productively using the applications available in their domain, then they may require more advanced courses which may focus on topics in more depth than the more necessarily utilitarian training offered in the courses discussed earlier.

A further function of training within EGEE is to foster communication within the project and the sharing of experience. To this end the training activity should be fostering the use of cross-activity workshops. This should also provide a mechanism for bringing advanced user communities into contact with the middleware development activities and grid operational activities.

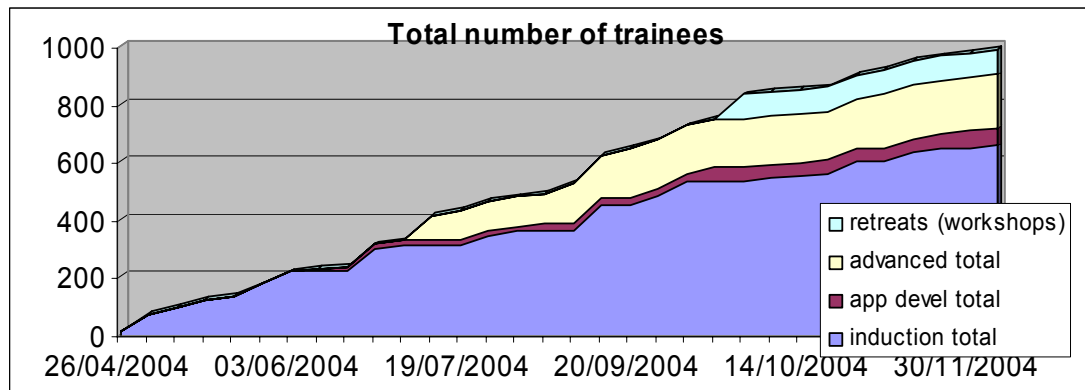


Figure 2: Graphical representation of the training events by category over time

Figure 2 shows that the initial focus in training was on induction and, while this effort is continuing, more advanced course have been added over time.

3.2. QUALITY ASSESSMENT

For each training event feedback is collected from participants. Attendees are asked to score how well the event met its goals, the individual presentations, any practical work, the facilities used and to give an overall score for the event. The scores are collected using a scale of 0 – 6 where 0 is poor and 6 is excellent.

We have always achieved better than the mid point on this scale and generally significantly above. The trend shows steady improvement.

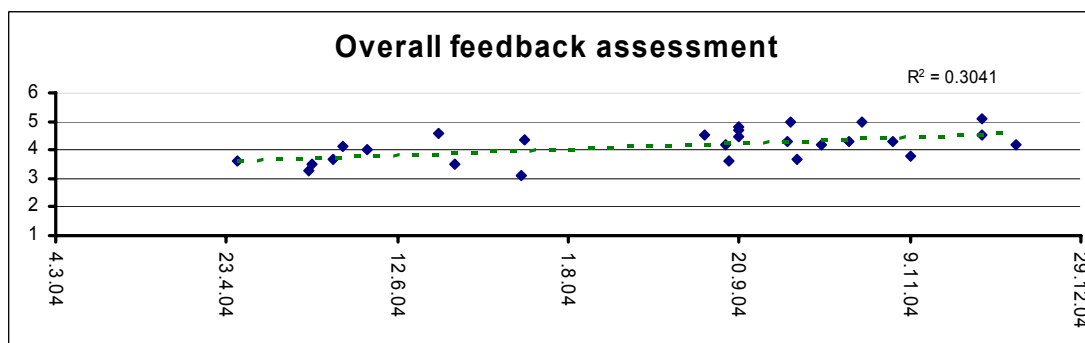


Figure 3: Graphical representation of the participant perception of the overall quality of EGEE training courses

The dashed straight line shows a regression of the quality data and the r^2 value is given.

As would be expected, in a project where training events are being presented across Europe by many different groups, the graph above shows that there is a degree of variation in the quality reported for courses. Encouragingly the general perception of the courses by attendees appears to be good and improving. It is likely that this reflects the re-use and iterative improvement of material as well as the improving focus of training and growing experience of trainers.

3.3. ACTIVE PARTNERS

Table 1: Delivery of courses by location (end of November 2004)

LOCATION	Number of Courses	Normalised for FTEs	Number of Days contact	Normalised for FTEs
Central Europe	10	1.9	19	3.6
United Kingdom	8	1.3	16	2.6
Italy	3	3	7	7
Northern Europe	2	1.5	4	3
Russia	6	1.4	11	2.7
Germany/Switzerland	8	8	17	17
South East Europe	2	0.6	3	0.8
Baltic	1	-	2	-
Spain	1	-	2	-
France	1	-	2	-
TOTAL	39	1.7	78	3.5

NB. The table above shows the courses and effort categorised by where the events were presented. This may not match the partner who provided the tutors (e.g. Spanish and Baltic events had tutors from the UK as did two of the events in Switzerland and one in Italy).

NB. The effort per partner in FTEs is shown in section 3.4 table 2.

As can be seen from table 1 (above) each of the NA3 federations has delivered at least two courses during the period under discussion. The first training plan (DNA3.1.1) calls for a total of 26 training events per year, whereas 39 have actually been delivered. The courses delivered have averaged almost exactly two days each across all of the federations. The total number of attendees so far reported is 1033.

The number of courses delivered is very roughly correlated to the effort, although it is clear that the German federation in particular has delivered more courses than might have been expected. Similarly this analysis does not show the effort contributed in support of courses in other areas. In this context it

is worth noting the efforts by INFN in Italy in providing support for many courses through the GENIUS training portal out of all proportion to their nominal effort. The UK has provided registration support for many courses and contributed in presenting courses in areas where NA3 is not represented (e.g. Spain and the Baltic countries). UEDIN provides EGEE wide (geographically and by activity) support for training course and provides the central coordination and management of the NA3 activity.

Both the Russian and German/Swiss federations have also made very strong contributions to the training activity, particularly in pioneering advanced courses on LGC2 administration and installation.

3.4. COOPERATION IN EGEE

EGEE activities outside NA3 have contributed expertise to the provision of training courses. For instance SA1 has contributed personnel and hardware support for some of the LCG2 installation and administration courses. Similarly many activities have contributed to the induction events. The expected model for much of this support is that experts in a particular domain will contribute material and training in the first instance which can be assimilated and built upon by the training team. On subsequent occasions this material may then be presented by others. A training materials archive has been created to assist with this process and dedicated hardware is being acquired in order to minimise the interference to the production facilities. However there remains a tension between the needs of training and development given that the knowledge of middleware developments can only reside with the development groups initially. Similarly 'hands-on' experience of middleware within the production environment will be concentrated initially within the operations activities.

Training courses have also been delivered on request from other activities, for instance the Biomedical Application Developers course 6th – 8th October 2004 in Madrid.

3.5. TRAINING RESOURCES

The personnel resources available within the NA3 activity of EGEE are listed below:

Table 2: EGEE training FTEs per Federation, taken from the Training Plan (DNA3.1.1).

Federation	Effort in FTEs
South East Europe	3.44
Russia	4.13
Northern Europe	1.33
Italy	1
Germany/Switzerland	1
Central Europe	5.29
UK/Ireland	6
TOTAL	22.19

This equates to about 1 FTE per partner in NA3 (22 partners) or about 0.3 FTE per partner within EGEE (given 70 partners). As training is required over the whole geographical span of the project (and

beyond) and for all activities these figures give a measure of the effort compared to the size of the task.

It is becoming clear that the demand for training exceeds the original estimates and is likely to exceed the resources available within the NA3 activity. For instance the Biomedical Application Developers course held in Madrid (6th – 8th October 2004) had 20 places available but 60 applicants, and this has been repeated on other courses.

3.5.1. T-Infrastructure

The need for t-Infrastructure that is a parallel infrastructure dedicated to training has been clearly identified within EGEE. As part of the provision of LCG-2 installation/administration courses this has initially been partially met in an ad hoc, temporary manner by the removal of production infrastructure components (for instance in courses at Oxford and Karlsruhe). While this expedient has been highly successful it is not sustainable in the long term. In order to begin to address this particular requirement UEDIN is putting in place a 20 node grid dedicated to this type of course.

The relationship between t-Infrastructure and e-Infrastructure can be characterised as follows:

1. The t-Infrastructure emulates the e-Infrastructure as accurately as possible with respect to the technical, operational and management issues being covered by courses using the t-Infrastructure.
2. The version of t-Infrastructure may correspond to a future e-Infrastructure deployment to prepare developers and operations staff for that deployment.
3. The arrangements for authentication will permit rapid certification at or just before the start of the course.
4. The arrangements for authorisation may restrict imposed loads compared with normal use or not exist because the t-Infrastructure is operated in isolation.

Typical requirements on e-Infrastructure of an introductory course are that demonstrations must run at the moment they are needed in that course and exercises must run just as the students, often in concert, reach that exercise, therefore, reserved / pre-booked capacity must exist to give guaranteed response during the course. Typical requirements on t-Infrastructure for a course developing systems administration expertise include direct access to a group of isolated machines for each group of students, so that they can make systems-level mistakes and reboot with affecting other students or production services. Self-paced learners require “always-on” t-Infrastructure with a self-explanatory portal, such as that provided by GENIUS (<https://grid-demo.ct.infn.it>).

3.5.2. Web portals

Web portals provide important static training sites which nevertheless can reflect a fast moving subject area in ways that are impossible for more traditional forms of publishing. This makes them invaluable resources for training in an environment like EGEE. The portals allow for self-paced learning outside the traditional class-room scenario.

3.5.3. GENIUS/GILDA

The GENIUS portal provided by INFN using the GILDA test-bed grid (<https://gilda.ct.infn.it/>) provides easy access to a variety of grid applications in many of the EGEE subject domains coupled with a user friendly interface and a series of tutorials on getting started with grids. Coupled to this INFN have also produced a video tutorial for the GENIUS interface

(<http://real1.rm.cnr.it:8081/ramgen/Grid/Grid.rm>,
<http://real1.rm.cnr.it:8081/ramgen/Grid/Demo/DemoBarbera.rm>).

In order to simplify the initial experience of grid applications the GENIUS portal also has a demonstration mode where prior certification of users is not required. This allows prospective users to access the grid applications immediately.

A CD based instantiation of a grid User Interface client has also been produced by INFN to allow users to interact from their own LINUX machines with the GENIUS/GILDA grid. This is provided as an ISO image using the KNOPPIX LINUX distribution (<https://gilda.ct.infn.it/live-cd/gilda-live-ui.iso>).

Those wishing to use GILDA facilities to support their own training should complete a web-request at https://gilda.ct.infn.it/request_tutorial/.

3.5.4. NA3 web portal

The UK NA3 partner (NeSC) has hosted a NA3 web site (<http://www.egee.nesc.ac.uk/>) which serves to advertise all of the EGEE courses, publish EGEE training statistics, provides facilities for requesting courses and makes EGEE course material available through a materials archive (see section 5).

The UK partner also makes available web registration for courses for those engaged in training throughout EGEE.

3.5.5. eLearning

AccessGrid (<http://www.accessgrid.org/>) has been used on an experimental basis for providing advanced course material for the Glasgow-based distributed computing MSc. This had the format of a traditional one hour lecture delivered over AccessGrid. The medium appeared to be satisfactory for this teaching format in the first instance.

The use of the SMIL web-based system for presenting video and slides (<http://www.w3.org/TR/XHTMLplusSMIL/>) was demonstrated at the NA3 Open Meeting in Karlsruhe by our GRNET partners.

4. TYPES OF COURSE

Four types of course were initially identified within the EGEE project. These are listed below with accompanying detail on the actually provision of courses, their distribution and the materials created to support them.

4.1. INDUCTION

4.1.1. Delivery

Table 3: Delivery of Induction courses

QUARTER	Number of courses	Number in Training Plan	Teaching days	Participants	Participant days	QA (average)
Q 1	9	5	14	296	4144	3.61
Q 2	8	0	11	190	2090	4.35
Q 3	9 (+ at least 2 more)	5	25	214	5350	4.48
Q4	-	0	-	-	-	-
1 st Year Totals	26	10	46	700	11584	4.14

NB. The Quality assessment is averaged for each quarter. The final value is reached by averaging over the values for the quarters.

4.1.2. Locations

Induction events have been staged in: UK, Austria, Switzerland, Italy, Greece, Slovakia, Netherlands, Russia, Germany, France, Lithuania, Czech Republic and Poland (including Isreal and Turkey within the period to end of December 2004).

4.1.3. Numbers

There have been 700 attendees at induction courses in the first 9 months. The Training plan expected 500 attendees per year.

Using the figures in the Training Plan the first year estimate of number of participant days for induction training is 1000 (10 events x 2 days x 50 participants). The actual delivery, in the first 9 months, has been 2.6 times the number of planned courses. The number of participants exceeded expectations by a factor of 1.4. There was approximately 10 times the expected number of participant days.

4.1.4. Material

A variety of induction material has been created and is discussed in deliverable DNA 3.2.1 (https://edms.cern.ch/file/504373/1/DNA3.2_document.doc). The topics covered include basic introduction to the concepts of grid, the EGEE project and it's context; EGEE operational organisation; overviews of EGEE middleware, applications domains and security; EGEE regional organisation; basic tutorials

introducing grids and taking users through joining a VO and running simple jobs (using the GENIUS portal).

The induction material has been re-used and updated in many courses. It has also been produced and delivered in languages other than English, e.g. German, Greek, French and Russian.

Materials will be made available through the Materials Archive (see section 5).

4.1.5. Current and projected demand

There has been a fairly constant demand for induction courses since the beginning of the project. Partly this has been fuelled by new regions starting to join the use of the LCG 2 middleware and the consequent training/dissemination activities. In this case it is difficult to disentangle training and dissemination with a major technical content.

Further spurs to this demand are the new application areas joining the EGEE project. Although initially this will tend to increase the demand for application developer and advanced courses as these are the segments most likely to be involved in the early phases of membership.

The expectation then should be that there will be a fairly constant demand for this type of course. A body of training material now exists to support the delivery of these courses and meeting the demand should be a matter of providing sufficient trainers. One caveat though is that there will still be some need to adapt the material for particular purposes.

4.2. APPLICATION DEVELOPER

4.2.1. Delivery

Table 4: Delivery of application developer courses

QUARTER	Number of courses	Number in Training plan	Teaching days	Participants	Participant days	QA (average)
Q 1	2	1	4	17	68	4.3
Q 2	1	1	1	8	8	3.6
Q 3	2 (+ at least 1 more)	4	6	35	210	4.33
Q4	-	2	-	-	-	-
1 st Year Totals	6	8	11	60	286	4.07

NB. The Quality assessment is averaged for each quarter. The final value is reached by averaging over the values for the quarters.

In terms of the Training Plan the expectation was that Application Developer courses would generally be 4 days long. In practice it has been found that developers are unwilling to commit this length of time to training courses, which reflects their work load. This has meant that these courses have in reality been run over two days (from experience a 2 – 3 day course also appears to be better in terms of maintaining participant focus).

Similarly the average expected attendance of 25 has not been met. This does not necessarily reflect a lack of demand but the attendance at courses has to be limited by the trainer /participant ratios that are required for these courses to be delivered effectively.

4.2.2. Locations

Application developer's courses have been staged in: Spain, Switzerland, Venezuela and UK.

4.2.3. Numbers

There have been 60 attendees at Application Developer courses. The Training plan expected 200 attendees per year. The Technical Annex calls for 800 participant days/year of this type of course. This will not be met in the first year (first 9 months – 286 participant/days).

4.2.4. Material

In response to demand from the NA4 (Applications) activity the Application Developer courses have concentrated on the LCG-2 APIs, Globus Toolkit and Web Services, in anticipation of the gLite middleware. Topics covered have included: Web Services architecture and protocols; WSDL; JAX-RPC; Ant; deployment using JWSDP and Tomcat and practicals on creating servers and clients from WSDL definitions.

Material on using UML modelling for good software design practice in conjunction with web services (WSDL) has also been produced.

Materials will be made available through the Materials Archive (see section 5).

4.2.5. Current and projected demand

There is clearly a great demand for these courses as many of them have been heavily over subscribed. It is clear from the figures shown above that while the expected number of courses have been presented, they have generally been on a smaller scale than was forecast. One reason for this is that the nature of this material requires a lower ratio of students to teachers than was initially expected. This is because this type of course demands a high proportion of practical work and involvement from demonstrators. Similarly a greater proportion of practical work requires considerably more preparation.

The demand for web services training is expected to remain high up to and beyond the introduction of the gLite middleware. Once gLite is introduced this training will be linked with gLite specific APIs.

As LCG-2 will continue to be supported some demand for training of the type described will remain but it is expected to decrease over the second half of the project.

The demand for developer's courses will also increase as:

- A – new infrastructure makes EGEE more attractive to developers
- B – new disciplines join.

Consequently we expect that a substantial increase in this category of course in the coming year.

4.3. ADVANCED

4.3.1. Delivery

Table 5: Delivery of advanced courses

QUARTER	Number of courses	Number in Training plan	Teaching days	Participants	Participant days	QA (average)
Q 1	0	0	0	0	0	-
Q 2	4	0	22	168	3696	3.7
Q 3	1 (+1 to be delivered)	0	2	18	36	4.2
Q4	-	0	-	-	-	-
1 st Year Totals	5	0	24	186	3732	3.95

NB. The Quality assessment is averaged for each quarter. The final value is reached by averaging over the values for the quarters.

4.3.2. Locations

Advanced courses have been staged in: Italy, Russia, UK and Germany.

4.3.3. Numbers

There have been 186 attendees at advanced courses. The Training plan expected 50 attendees per year. The Technical Annex calls for 250 participant days/year for this type of course, in this case it has been exceeded greater than 100 fold.

4.3.4. Material

Materials produced have so far dealt with the installation and administration of LCG-2.

4.3.5. Current and projected demand

This type of course includes the LCG-2 installation and administration courses (and gLite in turn). Initially there was expected to be a relatively low demand for these courses as much of the training was expected to occur within operational centres and to have a degree of 'on the job training' character.

However it is becoming clear that middleware installation, and particularly configuration, requires more training at more sites (for instance university departments) than was initially expected. This is intimately tied with the provision of user support and these need to be coordinated.

One limitation on the provision of the courses outlined above is the heavy requirement for hardware for installation and the need to keep this separate from the production infrastructure.

4.4. ADVANCED RETREATS (WORKSHOPS)

4.4.1. Delivery

Table 6: Delivery of advanced retreats

QUARTER	Number of courses	Number in Training plan	Teaching days	Participants	Participant days	QA (average)
Q 1	1	0	1	16	16	-
Q 2	2	0	2	20	40	-
Q 3	1	0	5	85	425	-
Q4	-	0	-	-	-	-
1 st Year Totals	4	0	8	121	481	-

NB. The Quality assessment is not collected for workshops as this was not deemed necessary.

4.4.2. Locations

Workshops have been staged in: UK, Slovakia and Italy

4.4.3. Numbers

The Technical annex calls for at least 360 participant days/year of this type of event. This has been comfortably exceeded in the first 9 months of the project (441 participant days). We have been informed that many of these meetings have taken place within the period but details are not available.

4.4.4. Material

The CONDOR week workshop in the UK is currently working on producing a document describing experiences and solutions from the spread of users and administrators present at the meeting.

4.4.5. Current and projected demand

There appears to be a steady but moderate demand for this type of meeting.

5. MATERIALS ARCHIVE

5.1. DESIGN

A new repository of training materials is currently being beta-tested.

This repository is based on a relational database with a web based front end. Design documentation is available at: <http://www.egee.nesc.ac.uk/trgmat/index.html>. During testing it is not generally available. On completion it will be open for anybody to download material and use it, in compliance with EGEE's open source policy.

A Unified Modelling Language (UML) class Diagram for the current database is shown below.

As the resources available for this project are strictly limited and there is a need for a functional system to be created as quickly as possible, it is important to limit the design to a minimal set of core functionality. This will allow a usable prototype to be created rapidly which can be modified to cope with future requirements, as long as extensibility is built into the system.

There are many on-going projects aimed at producing generalised materials repositories for education. These tend to have a number of features that are not required for the current EGEE repository.

For example:

1. The ability for many users to make depositions to the archive.
2. Flexible access controls to facilitate copyright control.
3. Support for many different user communities.
4. Generalised support to cope with all possible forms of material.

(See UK JORUM project, http://www.jorum.ac.uk/docs/Vol1_Fin.pdf)

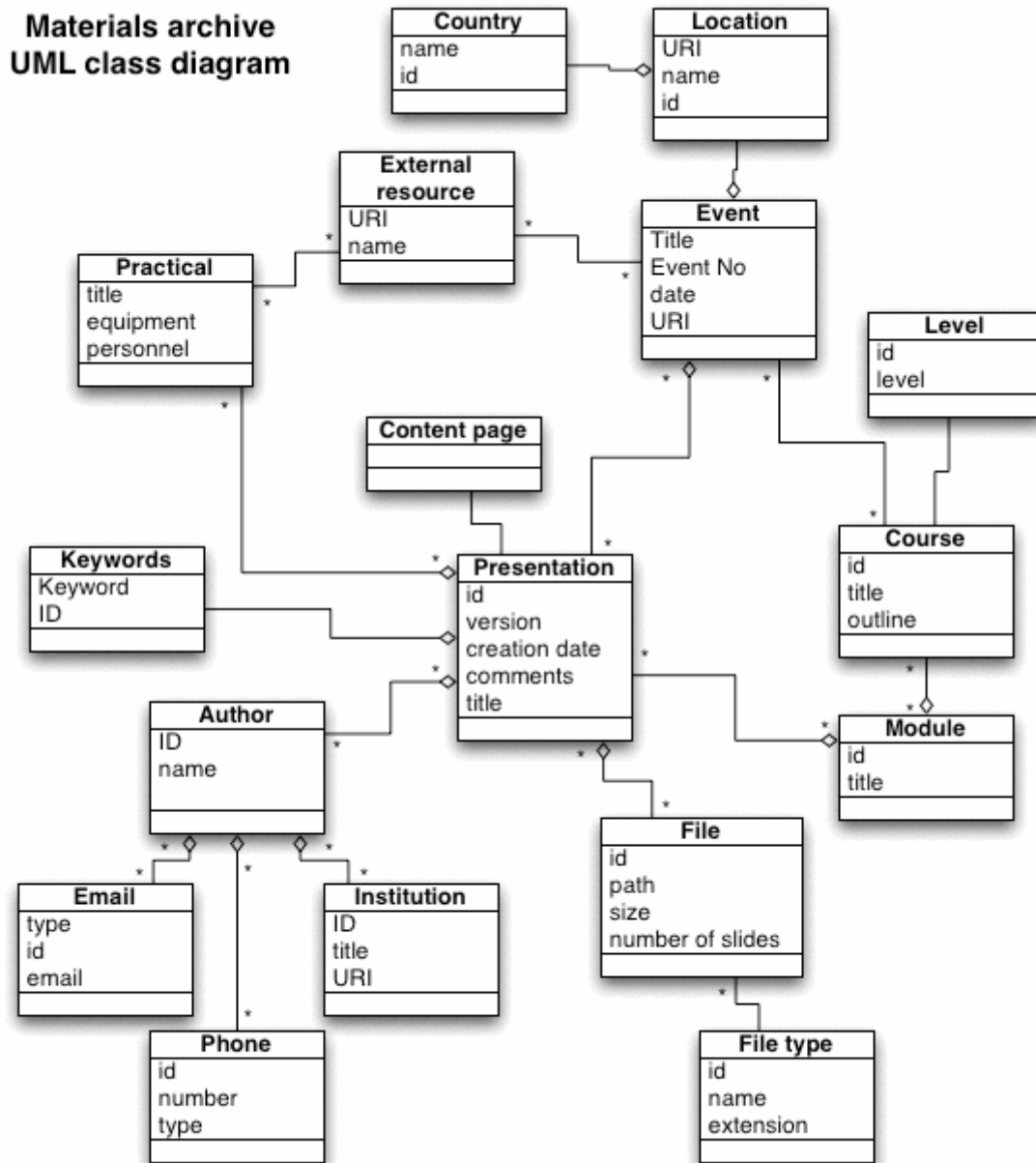
Many of these projects are relatively long term whereas, given the time constraints of the EGEE project, the NA3 materials archive will have to be established very rapidly.

As can be seen from the diagram below the proposed EGEE repository can be viewed as a minimal set which corresponds with the definition of a Learning Object {quoted by JORUM as : "an aggregation of one or more digital assets, incorporating metadata, which represents an educationally meaningful stand alone unit"}. In this case the 'stand alone unit' is considered to be the presentation. The metadata represented in the diagram is the same as the initial examples given by JORUM for useful metadata, e.g. author, description, title, educational level.

As the time is not available to make a systematic survey of all the current and proposed teaching archives of all of the potential stakeholders in EGEE, it has been decided to produce a simple structure for the current archive which will facilitate translation into the other schemas used for individual archives. For instance the current EGEE archive schema can be seen as a subset of the mandatory fields defined in the LOM core (<http://www.cetis.ac.uk/profiles/uklomcore>) and translation would mainly require the addition of catalogue and metadata – metadata information.

It is expected that the most likely route will be through a XML representation which will allow instantiation into either schema. Creation of the translation applications is beyond the scope of the current activity.

Figure 4: EGEE NA3 Training materials archive UML class diagram



The model is built around the central concept of a ‘presentation’. This is expected to be the minimal unit of material which will be included in the archive. Usually this will be a file containing a number of slides that correspond to a single talk. However it could represent a single slide if required. Storing the material as a collection of single slides was considered but rejected as being too cumbersome in

the circumstances. Additionally, even if a user is looking for individual slides in order to create a new talk, it is often critical to have access to the context for individual slides when trying to interpret them.

Each presentation will have a set of simple metadata associated with it as shown in the diagram. It is expected that in some cases explanatory notes to accompany a talk may be available. Further comments may be added about the talk in the curation process to add to the information available to users.

Provenance information, e.g. the authors and modifiers, will be attached to each presentation.

A presentation may be part of a course. A course may represent a definitive set of presentations which form the recommended content for a particular subject or it may represent a specific instance of linked presentations which were linked at an event.

An event represents an occasion at which one or more instance of a presentation was given. Its information should be compatible with that stored in the NeSC events registration database to facilitate interoperability.

The actual location of the physical file representing a presentation will be held in the file entity. This has an attribute 'type' which should be used to indicate the format of the file.

One or more presentations may be linked to practical materials, which are generally expected to be stored in an external site (e.g. a tar archive on a FTP server).

It is widely recognised that as the volume and scope of such a collection grows it is well worth investing in well-structured curation processes such as those developed in digital libraries. The combined EGEE material, including documentation produced by all activities has clearly crossed the threshold where this would be worthwhile. As yet the resources and skills have not been found to develop an EGEE digital library.

5.2. IMPLEMENTATION

The database for the materials archive is currently implemented using SQLServer (<http://www.microsoft.com/sql/default.msp>) with a MS Access (<http://office.microsoft.com/en-gb/FX010857911033.aspx>) front end for generating forms and maintenance. The SQL used in creating the database schema is shown in Appendix 4.

A coldfusion (<http://www.houseoffusion.com/>) based dynamic web front end has now been implemented and allows limited searching of the database. This was presented at the NA3 Open meeting in Karlsruhe (24/9/04). The representatives of the NA3 training community suggested some enhancements that will be incorporated as soon as resources allow (for instance metadata on language has been added to the 'presentation' and 'event' tables).

5.3. ISSUES

5.3.1. Copyright

The ownership of copyright is in general one of the most important issues in the implementation of publicly available archives. It differs from the copyright of software, in that not only is re-use encouraged, but the modification of training material to support the students on a particular course is absolutely essential. Sometimes this re-crafts examples and exercises to match a discipline, other times it adapts to language or background technical experience. This is expressly forbidden by the EGEE software licences (<http://public.eu-egee.org/license/copyright.html>), and so a more relaxed public licence is required for training material, including programs used in examples and exercises. To quote:

“You are permitted to copy and distribute verbatim copies of this document containing this copyright notice, but modifying this document is not allowed.”

In order that optimum use of effort is possible and in order that EGEE prepares communities currently outside EGEE use of material and donation by external e-Infrastructure educators should be welcomed and external parties should be encouraged to use EGEE training material. At present, we rely on the Consortium Agreement to clarify the ownership of the IPR, so that EGEE contributions to the repositories can be made publicly available. This does not permit the external use nor does it indemnify EGEE if external donations are made, which contravene IPR ownership, e.g. many academics think of their training material as their own IPR, whereas it often is owned by their employing institution.

5.3.2. Maintenance

The issues discussed in the section above mean that it is not currently possible to allow free submission to the materials repository the material must be moderated before inclusion. Obviously this is a very resource intensive procedure. This leads to a trade off as moderation is helpful for quality maintenance but can inhibit subscription and volume growth.

6. CONCLUSION

The first 9 months of this training have been very successful, exceeding planned targets by substantial margins. Overall, 1033 people have attended 39 training events that have delivered a total of 15,797 participant days. This breaks down into 11,584 participant days at induction events, 286 participant days at developer courses, 3,732 at advanced courses and 481 at workshops. We measure the views of the participants using a standard QA form, and the results show we are always in the upper half of scores and also show a positive trend of improvement.

In every case except developer courses, this is well in excess of the planned targets for the first 9 months. This has been achieved because of the enthusiasm and hard work of members of EGEE drawn from many activities across all regions. The developer courses have taken longer to prepare, have required higher trainer to participant ratios and have more diversity than originally anticipated. The increasing functionality and resources of the operational EGEE systems and the growing diversity of user communities from more and more disciplines guarantee that demand for developer courses will significantly increase – it already exceeds supply. Therefore, in the next year increased effort and attention will be devoted to developer training. Our analysis of demand shows that all other aspects of training will continue to be in great demand – see the Revised Training Plan DNA3.1.2.

We have recognised the vital importance of providing sufficient good quality t-Infrastructure. T-Infrastructure emulates the production e-Infrastructure but has additional properties that meet particular training and student needs – such as pre-production features, short response times, light weight authorisation and convenient portals. The initial work, EGEE can boast that it has pioneered exceptionally good t-Infrastructure for introductory, developer and advanced courses requiring grids. This is a valuable collaboration between NA3 and NA4.

An unexpected requirement has been the large demand for training courses in installing and running LCG2 services. As more and more sites join in the provision of grid-mediated resources, there are more staff who have to set up and operate EGEE infrastructure. Many of the operational difficulties are due to configuration and operational errors at new sites. Consequently there is not only high demand for systems administration training, it is crucial for EGEE that this training be given. Effective collaboration between SA1 and NA3 has delivered early versions of these courses. They require exclusive use of well-provisioned t-Infrastructure as well as experts with experience of systems administration. This will require more investment as e-Infrastructure is rolled out to EGEE sites and the trained staff produced will be a vital underpinning of Europe's eventual e-Infrastructure.

To improve training efficiency and to support self-paced learning we are establishing a repository of training material at the NA3 web site and a collection of tutorials available with GILDA, the preferred t-Infrastructure for introductory courses. In conjunction with SA1, JRA1, NA2 and NA4 we have begun the development of integrated, user-oriented information about EGEE and its use.