

# DataGrid Project Status



Fabrizio Gagliardi – Project Manager  
Fabrizio.Gagliardi@cern.ch

# Outline



- ◆ Introduction to the review
- ◆ Achievements of the project
- ◆ Project deliverables
- ◆ Project budget
- ◆ Summary

# Introduction to the Review

- ◆ Review structure
  - **Today (AM):** 2003 Overview and current status
  - **Today (PM):** Application assessment, dissemination and quality, Application demos
  - **Tomorrow (AM):** Middleware development, security, summary and future exploitation
- ◆ Details are provided in the deliverables and periodic reports
- ◆ Need to follow the agenda firmly so please hold questions till the end of presentations

# DataGrid Summary



- ◆ Leading vehicle for Grid Research and deployment over the last three years
  - **Europe's flagship Grid Project**
- ◆ 2001
  - **Established organisational structures**
  - **Gathered/analysed requirements and surveyed software toolkits**
  - **Assembled large scale grid testbed and defined middleware architecture**
- ◆ 2002
  - **Improved services provided to application groups (quality issues in software, creation of documentation, tutorial programme)**
- ◆ 2003
  - **Built on solid technical base to provide advanced functionality to application groups and Grid related projects around the world.**

The results will form the basis of EGEE

# Themes of 3<sup>rd</sup> year of the project



- ◆ Complete the work as described in the technical annex
- ◆ Deliver and support the grid software for various national and international production grid programmes
- ◆ Define migration plan for future projects
- ◆ Elaborate exploitation plans

# Relationship with LCG

- ◆ LCG is a grid deployment project
  - Provides the computing infrastructure for the LHC experiments
  - World-wide coverage
  - Production requirements with strict deadlines
- ◆ LCG has worked closely with DataGrid over the last two years
  - Has adopted software products from all DataGrid middleware WPs
  - Has adopted many DataGrid procedures and templates
  - Established joint testing-team with software certification and scalability tests
  - Rapid feedback has helped improve the quality of DataGrid software



# DataGrid Achievements in Brief



- ◆ Complete the work as described in the technical annex
  - Major enhancements for functionality and stability in all middleware areas
    - VOMS based security model integrated in job-submission chain
    - Replica Location Service (RLS) deployed
    - Re-factorization of WMS, edg-replica-catalog
    - Support for MPI, interactive, and checkpointable jobs
    - Unified access to storage (SE)
    - R-GMA replaced MDS as basis for the information system
    - New automated fabric tools developed
    - Many more (details in the individual WP presentations)
  - Successful deployment of middleware for use by applications
    - Sept. 03: third major release of EDG middleware made available
    - Important improvements included in final EDG version 2.1 deployed end Nov. 03
    - Testbed focused on geographic distribution (more than 20 sites), not on computing power (many sites moved CPUs to production Grids)

# DataGrid Achievements II



- ◆ Applications assessed middleware on application testbed
  - Many scientific results achieved exploiting DataGrid testbed in all three application areas
  - New applications ported to Grid (WP10)
  - Earth Observation sites joined the testbed (WP9)
  - Grid environments (e.g. LCG) already in production use for HEP (WP8)
  
- ◆ Re-launched Application Working Group (AWG) has delivered important cross-application outputs
  - Re-launched March 2003 - Chaired by Vincent Breton (WP10)
  - Interactions with EU GridLab Project
  - Valuable recommendations and requirements
  - Excellent forum for discussions



# DataGrid Achievements III



- ◆ Delivery and support the grid software for various national and international production grid programmes
  - Successful delivery of middleware to the LCG production infrastructure:
    - EDG software installed on more than 40 sites throughout the world. Next generation LCG-2 deployed January 2004
  - Datagrid Middleware being exploited by other projects and production facilities:
    - LCG, EU CrossGrid Project, EU Grace Project, Italian INFN grid/grid.it, Dutch Dutchgrid, UK eScience programme, DataTAG, etc.
    - openlab: Involvement of industrial partners in Grid-related activities
    - EDG middleware being ported to different platforms such as IA64 as part of the CERN openlab project. (Enterasys Networks, HP, IBM, Intel Corporation, Oracle)
    - WP4 Quattor computing management tool now deployed at the CERN computing centre (approx. 2000 nodes). Other sites have expressed interest

# DataGrid Achievements IV



## ◆ Further important achievements

- Europe-wide network activity
  - Upgrade of numerous National research and Education networks
  - National networks inter-connected with GEANT
  - Grid data transport at GBit/s speed anticipated confidently
  - Close collaboration maintained with DataTAG and DANTE
- Active participation in international standards bodies
  - Providing contributions to different GGF Working and Research groups
- Coordination and cooperation with related Grid Projects
  - DataTAG, CrossGrid, GRIDSTART, GRACE, GridLab, etc.

# DataGrid Achievements V



## ◆ Further important achievements

- Open source software license established
  - Used throughout 2003 and submitted for approval as an OSI compliant license; model for CrossGrid and Globus license.
- Tutorials
  - “Education & Outreach Manager” assigned to coordinate and lead tutorial activities.
  - More than 600 people trained in over 25 events
- Two Project conferences organized
  - May 2003 in Barcelona, September 2003 in Heidelberg

# DataGrid Third Year Deliverables



**D1.6, D2.5, D3.5, D4.5, D5.5, D6.7:** Components and documentation for the third project release

**D1.7, D2.6, D3.6, D4.6, D5.6:** Final Evaluation report

**D6.8:** Final evaluation of testbed operation

**D7.4:** Final report on network infrastructure and services

**D7.7:** Final security report

**D8.4:** Report on the results of HEP application migration #2 and final application report

**D9.4:** EO application platform installation

**D9.5:** EO application processing demonstration and final report

**D10.4:** Final report including the 2nd bio-testbed release

**D11.6:** Second annual conference and grid forum

**D11.7:** Final conference

**D11.9:** Contribution to international standards

**D12.13-16:** Third year quarterly reports

**D12.19:** Third year annual report

**COMPLETED AND DELIVERED**

*D6.6, D7.6, D8.3, D9.3  
and D10.3 which had been  
rescheduled for the first  
quarter of the third year  
were also delivered*

# Project Major Issues

- ◆ Build on fast moving technology
  - GT2 -> GT3 (OGSI) -> GT4 (WSRF)
  - Decision at the beginning of year 3 to stay with GT2 and only prepare a few components for OGSI by building on plain web-services was a sensible choice
- ◆ Relations with other Grid projects, LCG in particular
  - True production use with LCG required the adoption of more strict procedures
- ◆ Globus and other tools support
  - Now based on VDT; together with LCG an excellent relationship and support structure could be established
- ◆ Persistency of project results
  - Exploitation of project results (in particular through EGEE)
  - More details will be given at the presentation at the end of the review

# Project Budget



Partners	Cost basis	Original Plan Y1	Reimbursed	Deviation Y1	Original Plan Y2	Reimbursed	Deviation Y2	Original Plan Y3	Claimed	Deviation Y3	EU ceiling	Total planned	Overall Deviation
CERII		494 597	393 547	-20%	588 221	640 957	9,0%	463 761	541 404	17%	1 546 579	1 575 908	1,90%
CERII coord.		130 666	88 682	-32%	120 055	126 922	5,7%	169 372	157 269	-7%	420 093	372 873	-11,24%
<b>Total CERII</b>	AC	625 263	482 229	-23%	708 276	767 879	8,4%	633 133	698 673	10%	1 966 672	1 948 781	-0,91%
ITC-irst	FC	102 170	78 033	-24%	88 920	66 511	-25,2%	83 120	127 660	54%	274 210	272 204	-0,73%
UH	AC	90 624	13 775	-85%	91 820	71 206	-22,5%	93 394	159 600	71%	275 838	244 581	-11,33%
VR	AC	124 844	0		123 466	132 999	7,7%	123 636	197 890	60%	371 946	330 889	-11,04%
ZIB	AC	90 156	76 487	-15%	93 156	105 785	13,6%	93 156	94 193	1%	276 468	276 465	0,00%
EVG HEI UHH	AC	90 156	59 436	-34%	93 156	79 657	-14,5%	93 156	121 748	31%	276 468	260 841	-5,65%
CHRS	FF	278 547	251 686	-10%	314 111	338 368	7,7%	318 732	343 791	8%	911 390	933 845	2,46%
CSSI	FF	165 100	177 828	8%	165 100	181 371	9,9%	165 200	138 092	-16%	495 400	497 291	0,38%
CEA	FC	60 264	62 281	3%	61 377	62 700	2,2%	62 604	92 963	48%	184 245	217 944	18,29%
IFAE	AC	64 392	70 740	10%	64 392	38 983	-39,5%	0	19 061		128 784	128 784	0,00%
ESA	AC	206 452	206 042	0%	254 873	274 437	7,7%	183 675	164 522	-10%	645 000	645 001	0,00%
IIFIH	AC	319 189	256 610	-20%	300 592	455 860	51,7%	300 592	219 127	-27%	920 373	931 597	1,22%
DATAMAT	FF	151 629	141 968	-6%	175 974	156 218	-11,2%	142 406	159 360	12%	470 009	457 546	-2,65%
CHIR	FF	89 836	101 040	12%	89 836	91 382	1,7%	90 836	78 218	-14%	270 508	270 640	0,05%
CESNET	FF	45 656	51 071	12%	45 656	50 101	9,7%	45 706	47 766	5%	137 018	148 938	8,70%
FOM	AC	77 136	68 349	-11%	69 919	121 221	73,4%	71 125	81 868	15%	218 180	271 438	24,41%
KIMI	FC	59 655	66 580	12%	59 778	49 613	-17,0%	39 504	42 425	7%	158 937	158 618	-0,20%
SARA	FC	71 300	66 346	-7%	55 696	52 775	-5,2%	46 622	48 727	5%	173 618	167 848	-3,32%
PPARC	AC	396 916	97 504	-75%	386 404	386 310	0,0%	414 005	431 160	4%	1 197 325	914 974	-23,58%
MTA SZTAKI	FF	88 850	95 629	8%	94 367	94 417	0,1%	98 543	95 163	-3%	281 760	285 209	1,22%
IBM	FC	77 218	92 911	20%	79 429	85 604	7,8%	81 707	79 170	-3%	238 354	257 685	8,11%
<b>Total</b>		<b>3 275 353</b>	<b>2 516 543</b>	<b>-23%</b>	<b>3 416 298</b>	<b>3 663 397</b>	<b>7,2%</b>	<b>3 180 852</b>	<b>3 441 177</b>	<b>8%</b>	<b>9 872 503</b>	<b>9 621 117</b>	<b>-2,55%</b>

Table shows status at 31<sup>st</sup> December 2003

# Project Effort



		FUNDED				UNFUNDED				F+UF			
		C	E	PM Dev.	% Dev.	C	E	PM Dev.	% Dev.	C	E	PM Dev.	% Dev.
<b>CERN</b>	AC	<b>230</b>	245	-15	-6%	<b>546</b>	424	122	29%	<b>776</b>	669	107	16%
<b>ITC</b>	FC	<b>103</b>	72	31	42%	<b>3</b>		3		<b>106</b>	72	34	47%
<b>UH</b>	AC	<b>84</b>	36	48	133%	<b>60</b>	36	24	67%	<b>144</b>	72	72	100%
<b>VR</b>	AC	<b>48</b>	54	-6	-11%	<b>21</b>	50	-29	-58%	<b>69</b>	104	-35	-33%
<b>ZIB</b>	AC	<b>35</b>	36	-1	-4%	<b>37</b>	36	1	4%	<b>72</b>	72	0	0%
<b>EVG-HEI</b>	AC	<b>37</b>	36	1	1%	<b>45</b>	36	9	24%	<b>81</b>	72	9	13%
<b>CNRS</b>	FF	<b>310</b>	305	5	2%	<b>418</b>	15	403		<b>728</b>	320	408	127%
<b>CS SI</b>	FF	<b>78</b>	81	-3	-4%	<b>0</b>				<b>78</b>	81	-3	-4%
<b>CEA</b>	FC	<b>37</b>	31	6	19%	<b>57</b>		57		<b>94</b>	31	63	203%
<b>IFAE</b>	AC	<b>36</b>	36	0	0%	<b>62</b>	42	20	47%	<b>98</b>	78	20	25%
<b>ESA</b>	AC	<b>86</b>	84	2	2%	<b>82</b>	13	69	532%	<b>168</b>	97	71	73%
<b>INFN</b>	AC	<b>269</b>	228	41	18%	<b>645</b>	733	-88	-12%	<b>914</b>	961	-47	-5%
<b>DATAMAT</b>	FF	<b>116</b>	108	8	7%	<b>0</b>				<b>116</b>	108	8	7%
<b>CNR</b>	FF	<b>65</b>	60	5	9%	<b>64</b>		64		<b>129</b>	60	69	115%
<b>CESnet</b>	FF	<b>132</b>	132	0	0%	<b>12</b>	12			<b>144</b>	144	0	0%
<b>FOM</b>	AC	<b>42</b>	36	6	17%	<b>81</b>	102	-21	-21%	<b>123</b>	138	-15	-11%
<b>KNMI</b>	FC	<b>27</b>	26	1	3%	<b>9</b>		9		<b>36</b>	26	10	37%
<b>SARA</b>	FC	<b>28</b>	30	-2	-7%	<b>5</b>		5		<b>33</b>	30	3	11%
<b>PPARC</b>	AC	<b>167</b>	156	11	7%	<b>820</b>	534	286	53%	<b>986</b>	690	296	43%
<b>SZTAKI</b>	FF	<b>76</b>	75	1	1%	<b>39</b>		39		<b>115</b>	75	40	53%
<b>IBM</b>	FC	<b>38</b>	36	2	5%	<b>0</b>				<b>38</b>	36	2	5%
<b>Others (*)</b>						<b>433</b>	610	-177	-29%	<b>433</b>	610	-177	-29%
<b>Total</b>		<b>2042</b>	<b>1903</b>	<b>139</b>	<b>7%</b>	<b>3438</b>	<b>2643</b>	<b>795</b>	<b>30%</b>	<b>5480</b>	<b>4546</b>	<b>934</b>	<b>21%</b>

**Table shows status at 31<sup>st</sup> December 2003**

# Summary

- ◆ Project completed within budget and on schedule
- ◆ All partners actively participated and made it to the end
- ◆ DataGrid has gone the extra mile to produce output beyond the original contracted programme of work
- ◆ Functionality and performance of software and testbed(s) completed according to the original plans
- ◆ The value of EDG technology has been confirmed through its adoption by many projects including LCG for one of the largest scientific enterprises to-date
- ◆ Follow-on project successfully proposed and being launched (EGEE)