GridPP3 Financial Plan – June 2007

David Britton

1. Scope

This document contains the GridPP financial planning for:

a) All non-staff expenditure from the start of 1/Apr/07 to 31/Mar/11;
b) All SLA staff cost from 1/Sep/07 to 31/Mar/11;
c) All University Staff costs associated with Grants yet to be issued.

In addition to the sums shown in this document, the following spend from GridPP2 funds is expected in FY07:

1) Staff costs on the SLA at RAL between April-1st and Aug-31st 2007 – expected to be roughly £832k (to be updated).
2) All remaining expenditure on already-issued University Grants – Appendix-1 of this paper calculates unallocated GridPP2 funds assuming the University grants have been fully spent.

2. Funds

Details of the available funding are provided in Appendix-1. The total sum described in this plan is £28.95m consisting of £25.90m of new funding for the GridPP2+ and GridPP3 phases, together with £3.05m of previously awarded GridPP2 funding. The latter sum consists of (a) the balance of £1.27m (unallocated GridPP funds); (b) £1.51m budgeted for FY07 hardware; (c) £0.275m budgeted for FY07 travel.

3. Hardware

Summary: We have re-costed the hardware and aim to provide 85% of the requests for Tier-1 and Tier-2 hardware, as approved by the PPRP.

- Appendix-2 presents a summary of the current user requests for Tier-1 and Tier-2 hardware and the GridPP provision targets of 85% of the requests as approved by the PPRP.
- Appendix-3 presents details of our best-estimate of future CPU and Disk costs.
- Appendix-4 shows details of how the CPU will be provided.
- Appendix-5 shows details of how the Disk will be provided.
- Appendix-6 presents details of our current costing and provision of Tape storage at the Tier-1.

The total hardware-spend plan is shown in Table-1 below. The top three entries show the Tier-1 CPU, Disk, and Tape spend on capacity described in more details in the associated appendices. The forth line shows the proposed spend on tape infrastructure (drives, robot, M&O) which is also detailed in the Appendix-6. The fifth line shows the proposed spend on non-capacity machines at 10% of the CPU spend (see Appendix-7 for details) and the sixth line covers other miscellaneous items at the Tier-1. Due to the revised hardware costing and LHC schedule, the total Tier-1 proposed spend of £8.35m is slightly less than the sum of the £7.20m approved by the PPRP plus the remaining GridPP2 hardware budget of £1.5m. The last line of the table shows the Tier-2 proposed hardware spend as approved by the PPRP. Appendix-8 provides a justification for this sum in terms of the market-model adopted by GridPP. It is proposed to award the Tier-2 hardware money in two tranches, the first in FY07 and the second in FY09. The proposed method of division of the Tier-2 hardware between institutes is addressed in Appendix-9.

	WG	FY07	FY08	FY09	FY10	TOTAL PLAN
Tier-1 CPU		£493,550	£363,589	£342,729	£365,933	£1,565,802
Tier-1 Disk		£1,025,568	£970,255	£1,337,944	£760,069	£4,093,836
Tier-1 Tape		£241,611	£233,447	£0	£179,714	£654,772
Tape Inf.		£458,000	£200,000	£213,800	£521,400	£1,393,200
Non Cap. HW		£49,355	£36,359	£34,273	£36,593	£156,580
Tier-1 Misc		£113,000	£119,000	£125,000	£131,000	£488,000
Tier-1 HW	**A**	£2,381,085	£1,922,650	£2,053,746	£1,994,709	£8,352,190
Tier-2 HW	**B**	£2,176,500	£0	£2,176,500	£0	£4,353,000

Table-1: GridPP proposed hardware spending FY07-FY10.

4. Previously Approved Staff and Travel Costs

Tier-1 Staff (Work-package A)
The Tier-1 posts approved by the PPRP have been re-costed, resulting in a small net reduction from the total of £4.654m approved by the PPRP to £4.624m.

Tier-2 Staff (Work-package B)
The PPRP approved 13 out of 14.75 Tier-2 hardware support posts at a cost of £2.933m. Thirteen posts have been allocated as far as the Tier-2 level as described in Appendix-10 but the division at the institute level has not quite been finalised. The total (including the STFC inflation calculation) is now estimated to be £3.016m but this might vary slightly once the final institute division is known.

Support Staff (Work-package C)
There was a slight inconsistency in the feedback for GridPP2+ (Document "Grid PP 2+.doc" emailed to T.Doyle, 29/11/06) where 1-FTE was approved for R-GMA (Post D7) but the associated funding was only for 0.5-FTE. Assuming the 1-FTE is correct, the PPRP approved

sum should be £2.02m rather than £2.00m in the feedback summary table (see Appendix-1). After re-costing, the total is estimated to be £2.043m.

Operations Staff (Work-package D)
There was another inconsistency in the feedback to GridPP in Work-package-D where the total approved in the summary table (£1.802m) differed from the sum of the individual amounts approved in the GridPP2+ and GridPP3 feedback documents (£0.1397m + £1.7776m = £1.917m). We have re-costed the posts that were approved and the new total is estimated to be £1.941m.

Management (Work-package E)
The total approved by the PPRP was £0.987m. The posts have been re-costed and the Project Leader fraction adjusted to 90% (as discussed with STFC). The estimated cost is of this Work-package is now £1.096m.

Outreach (Work-package F)
The total approved by the PPRP was £0.122m. The posts have been re-costed and the new total is £0.130m. The plans for dissemination effort are contained in Appendix-11.

Travel and Other Costs (Work-package G)
The total approved by the PPRP for Travel and Other Costs was £0.63m. The total proposed here consists of £0.276m of GridPP2 funds plus £0.64m of GridPP3 funds for a total, including FY07 travel, of £0.91m. GridPP currently spends about £270k on travel per year with just over 80 FTE directly funded. Some travel funding supports non-funded FTE on GridPP related travel but it is hard to estimate a meaningful number for this. In GridPP3, the funded FTE count falls to about 60, a reduction of 25% but it is anticipated that there may be an increase in funding requests from non-funded FTE, particularly associated with Tier-2s. The GridPP3 budget reduces the travel funding by about 30% over present levels. One cost saving proposed is to reduce the number of GridPP collaboration meetings from three to two per year. The other costs included in WG-G are a budget line of £10k per annum for Miscellaneous items (licenses; test hardware; dissemination items; summer students; etc) and £5k per annum for project leader consumables (laptops, desktops, photocopies, personal printers, printed material etc).

5. New Costs

To complete the GridPP3 planning, three categories of previously un-costed posts have been included. The first category is posts that were defined in the GridPP3 submission but not included on the JeS forms because the hosting institutes had not been identified. The second category consists of transition funding to manage, in an orderly way, the large reduction in support effort over the critical LHC start-up period. The third category consists of new positions proposed to address the funding gap in experiment support for Grid middleware. The latter two categories address the following feedback from the March-07 GridPP oversight committee:

2. Plans for GridPP3
==========
The OC encouraged that we look flexibly at identified cost savings in GridPP2 that might be used in GridPP3 (subject to PPARC approval). They also encouraged us to identify the small number of things that can be done well (in the context of the overall plan proposed by Dave at the meeting).

3. Credibility Gap
==========
The OC recognised the potential derailment of the overall project due to lack of funding at the applications interfaces. A report was requested on where we are in terms of closing that credibility gap. (The background was of anticipated decreasing overall application efficiency due to poor implementations of the computing models).

Table-2 shows a summary of the proposed funding. More details on the contents of this table are given in Appendix-12. The Middleware Transition planning is described in Appendix-13 and the input from ATLAS, CMS and LHCb is provided in Appendices 14, 15 and 16.

		FTE	Start	Duration	Effort	Cost
Previously Identified Posts		**Frac**	**Date**	**(Months)**	**(Months)**	**Estimate**
NE9	UB Chair	25%	01-Sep-07	43.00	10.75	
NE8	Administrative Assist.	50%	01-Apr-08	36.00	18.00	
SubTotal				**79.00**	**28.75**	**£126,113**
Support Transition Posts						
NC28	VOMS Service	50%	01-Apr-08	24.00	12.00	
NC29	RTM Transition	50%	01-Nov-07	7.00	3.50	
NC30	Networking Completion	50%	01-Sep-07	7.00	3.50	
NC31	Metadata Transition	100%	01-Oct-07	6.60	6.60	
SubTotal				**44.60**	**25.60**	**£172,493**
Experiment Support Posts						
ND11	Atlas T1	50%	01-Apr-08	36.00	18.00	£107,500
ND12	LHCb T1	100%	01-Apr-08	36.00	36.00	£215,000
ND13	CMS T1	150%	01-Apr-08	36.00	54.00	£322,500
ND14	Atlas Ganga	100%	01-Apr-08	36.00	36.00	£215,000
ND15	LHC Ganga	50%	01-Apr-08	36.00	18.00	£107,500
SubTotal				**180.00**	**162.00**	**£967,500**
Grand Total				**303.60**		**£1,266,107**

Table-2: New costs proposed.

6. Full Plan Budget Summary

A full summary of the planned expenditure is presented in Table-3.

							Funding Source		PPRP
	WG	FY07	FY08	FY09	FY10	TOTAL PLAN	GridPP2	GridPP3	Approved
Tier-1 HW	A	£2,381,085	£1,922,650	£2,053,746	£1,994,709	£8,352,190	£1,508,318	£6,843,872	£7,204,000
Tier-2 HW	B	£2,176,500	£0	£2,176,500	£0	£4,353,000	£0	£4,353,000	£4,353,000
Tier-1 Staff	A	£593,295	£1,297,549	£1,342,963	£1,389,967	£4,623,774	£0	£4,623,774	£4,654,000
Tier-2 Staff	B	£166,429	£947,834	£949,971	£952,146	£3,016,380	£0	£3,016,380	£2,933,000
Support Staff	C	£528,191	£552,787	£509,621	£452,234	£2,042,833	£0	£2,042,833	£2,000,000
Operation etc	D	£137,634	£591,457	£601,126	£611,143	£1,941,361	£0	£1,941,361	£1,802,000
Management	E	£187,370	£301,187	£302,919	£304,719	£1,096,195	£0	£1,096,195	£987,000
Outreach	F	£23,046	£35,647	£35,647	£35,648	£129,988	£0	£129,988	£122,000
Group-N	N	£101,036	£388,006	£405,788	£371,276	£1,266,107	£1,266,107	£0	£0
Travel etc.	G	£276,558	£203,000	£212,400	£222,270	£914,228	£276,558	£637,670	£630,000
Total		£6,571,144	£6,240,117	£8,590,681	£6,334,112	£27,736,055	£3,050,983	£24,685,073	£24,685,000

Total Expenditure	£27,736,055
Working Allowance	£1,216,000
Total Funds Required	£28,952,055
GridPP2 Funds	£3,050,983
GridPP3 Funds	£25,901,000
Total Funds Available	£28,951,983
Project Balance	-£72

Table-3: Full Budget Summary.

Appendix-1: Available Funding

The "GridPP3" planning done here includes all hardware to be purchased from FY07 onwards and includes some GridPP2 planned hardware expenditure. For completeness, the planning also includes all funds not presently spent or encumbered within GridPP2. The total available funding is, therefore, the sum of the GridPP3 award approved by Science Committee in March 2007, plus the money remaining within the current project.

GridPP3 Funding

The approved GridPP3 funding is summarized in the Table -1 provided by STFC. The total is given in the middle of the table as £25901K of new funding, of which £1216K is reserved as Working Allowance.

Workpackage	Staff including indirects and estates (£k)	Hardware (£k)	Other costs	Total (£k)
A – Tier1	4,654	7,204		11,858
B – Tier 2	2,933	4,353		7,286
C – Grid Support	2,000			2,000
D – Grid Operations	1,802			1,802
E – Management	987			987
F – Outreach	122			122
G- Travel and other costs			630	630
Working Allowance				1,216
Total Project New Funding	**12,498**	**11,557**	**630**	**25901**
Contingency				1,250
Running costs (electricity)				2,500
Rolling Grant Funding				306
Grand Total				29,957

Table-1: The GridPP3 Award

GridPP2 Funding

Table-2 summarises the available funding within the current GridPP2 project. The top section shows the confirmed expenditure on the RAL SLA from FY01 to FY05 together with the predicted outturn for FY06. There is one non-SLA item which was a direct grant from PPARC. The second section shows the University Grants issued under the GridPP1, some of which was recovered when posts were terminated early, and the Grants issued under the GridPP2 project. The third section shows GridPP funding that went to CERN and for Globus support. The sum of the first three sections gives the expenditure which is then compared with the income from the GridPP1 and GridPP2 awards. To arrive at the available funding, the known encumbrances on the FY07 SLA are summed and the residual balance is £1.27m. The GridPP2 sum explicitly included in the "GridPP3" planning is £3.05m consisting of the non-staff sums of (a) the balance of £1.27m shown below (i.e. unallocated GridPP funds); (b) £1.51m for FY07 hardware; (c) £0.275m for FY07 travel.

RAL		
SLA FY01	£1,522,055	Includes pre-spend items in FY00
SLA FY02	£1,639,723	
SLA FY03	£1,602,243	
SLA FY04	£2,461,037	
SLA FY05	£2,357,724	
SLA FY06	£2,737,436	(Provisional outturn)
Non SLA	£26,667	*Direct grant from PPARC*
Sub Total	£12,346,885	
University Grants		
GridPP1 Issued	£4,603,425	*Includes Herriot-Watt extension and part of Dissemination Grant*
GridPP1 recovered	-£52,883	*Manchester and Bristol BaBar posts*
GridPP2 Issued	£6,439,676	*Includes part of the Disemination Grant which was re-issued*
Sub Total	£10,990,219	
Other Costs		
Globus Support	£12,658	
CERN	£5,666,835	
Sub Total	£5,679,493	
Expenditure		
Spend to date	£29,016,596	
Income		
GridPP1 Award	£17,000,000	
GridPP2 Award	£15,900,000	
Total Award	£32,900,000	
Encumberances		
SLA FY07 Staff	£832,420	
SLA FY07 Travel	£275,000	
SLA FY07 Hardware	£1,508,318	
Total Encumberance	£2,615,738	
Balance		
Balance	£1,267,665	

Table-2: Existing GridPP funding.

Appendix-2: Hardware Requirements

Tier-1 Hardware

Changes in the Experiment computing models and, more recently, changes in the LHC schedule have caused several iterations on the hardware requirements and we expect this process to continue. The current Tier-1 hardware requests are shown in Table-1 and the 85% provision approved by the PPRP is shown in Table-2. The units are in TB for Disk and Tape, and KSI2K for CPU.

Full Request	2008			2009			2010			2011		
	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape
ALICE	109	155	97	142	258	126	184	336	164	212	386	189
ATLAS	1242	2265	962	2461	3553	1869	4936	6197	3587	7024	8840	5616
BABAR	35	500	35	35	500	35	0	0	0	0	0	0
CMS	620	1330	1280	910	2120	2220	1340	3170	3160	1820	4750	4300
LHCB	191	364	161	514	908	572	606	1254	1093	700	1750	1550
OTHER	18	0	180	31	0	310	44	0	444	62	0	618
Total	2215	4614	2715	4093	7339	5132	7110	10957	8448	9818	15726	12273

Table-1: Tier-1 Hardware Requests (TB and KSI2K) May-2007.

"70% Scenario" (85% of hardware)	2008			2009			2010			2011		
	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape
ALICE	93	132	82	121	219	107	156	286	139	180	328	161
ATLAS	1056	1925	818	2092	3020	1589	4196	5267	3049	5970	7514	4774
BABAR	35	500	35	35	500	35	0	0	0	0	0	0
CMS	527	1131	1088	774	1802	1887	1139	2695	2686	1547	4038	3655
LHCB	162	309	137	437	772	486	515	1066	929	595	1488	1318
OTHER	15	0	153	26	0	264	38	0	377	53	0	525
Total	1888	3997	2313	3484	6313	4368	6044	9313	7181	8345	13367	10432

Table-2: Tier-1 Hardware Provision Targets (TB and KSI2K) May-2007.

Note: The BaBar numbers are identical in both tables and represent the Plan-B described in http://www.gridpp.ac.uk/tier1a/board/doc/BaBar2007/ .

Tier-2 Hardware

The current Tier-2 hardware requests are shown in Table-3 and the 85% provision approved by the PPRP is shown in Table-4. The units are in TB for Disk and Tape, and KSI2K for CPU.

Request	2008			2009			2010			2011		
	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape
ALICE	52	184	0	67	240	0	88	312	0	101	358	0
ATLAS	1033	2333	0	1748	3596	0	2951	6873	0	4146	9217	0
BABAR	40	1600	0	40	1600	0	0	0	0	0	0	0
CMS	340	1500	0	680	2200	0	1280	4050	0	1830	5200	0
LHCB	3	1092	0	6	2732	0	6	2732	0	6	2732	0
OTHER	107	197	0	176	214	0	250	252	0	345	253	0
Total	1575	6906	0	2717	10582	0	4575	14219	0	6428	17760	0

Table-3: Tier-2 Hardware Requests (TB and KSI2K) May-2007.

Provision Target	2008			2009			2010			2011		
	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape	Disk	CPU	Tape
	TB	KSI2K	TB	TB	KSI2K	TB	TB	KSI2K	TB	TB	KSI2K	TB
ALICE	44	156	0	57	204	0	75	265	0	86	304	0
ATLAS	878	1983	0	1486	3057	0	2508	5842	0	3524	7834	0
BABAR	40	1600	0	40	1600	0	0	0	0	0	0	0
CMS	289	1275	0	578	1870	0	1088	3443	0	1556	4420	0
LHCB	3	928	0	5	2322	0	5	2322	0	5	2322	0
OTHER	91	168	0	149	182	0	213	214	0	293	215	0
Total	1345	6110	0	2315	9235	0	3889	12086	0	5463	15096	0

Table-4: Tier-2 Hardware Provision Targets (TB and KSI2K) May-2007.

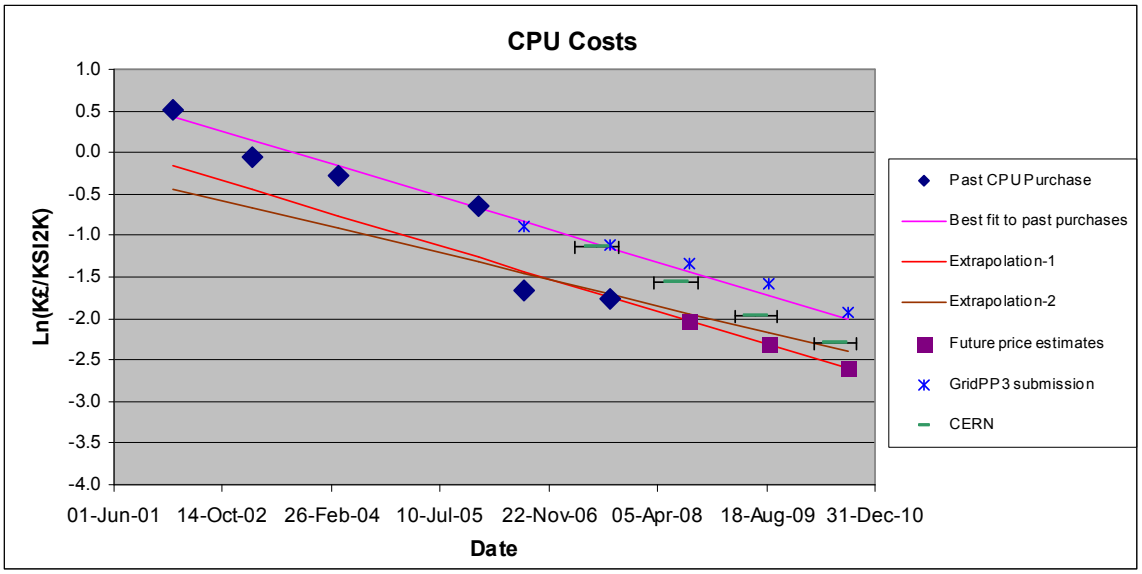
Note: The BaBar numbers are identical in both tables and represent the Plan-B described in http://www.gridpp.ac.uk/tier1a/board/doc/BaBar2007/ .

Appendix-3: Future CPU and Disk Costs

In order to prepare the final project plan for GridPP3, a new estimate of future hardware costs has been made, based on the most recent data available.

CPU Costs

Figure-1 shows the historical cost (logarithm) of CPU at the Tier-1 together with extrapolated costs for GridPP3. The first five blue diamonds from the left are Tier-1 CPU purchases between March-2002 and January 2006; the sixth point is an estimate of the cost for September 2007. The first four points suggest that prices were falling with a time constant of about 29-months, however, the fifth point shows a discontinuity associated with the introduction of the Woodcrest CPU. The purple squares are the future estimates used in this planning and are extrapolations using a 29-month time constant from the current pricing point.

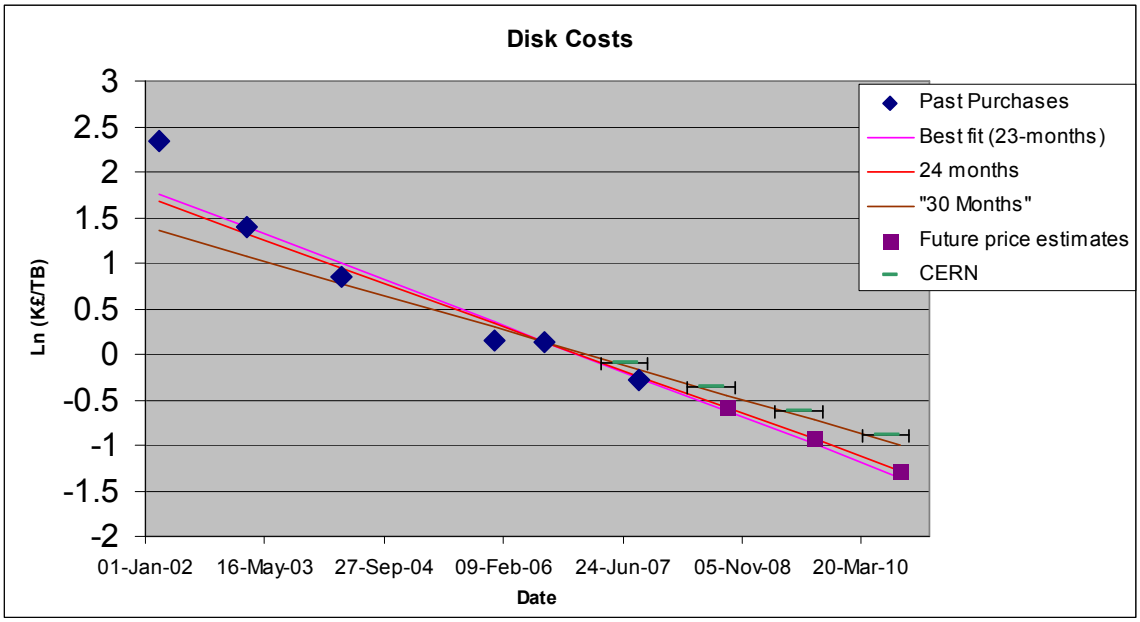


Tender Date	1-Sep-06	01-Sep-07	01-Sep-08	01-Sep-09	01-Sep-10
K£/KSI2K	0.19	0.17	0.13	0.10	0.07
Uncertainty	0%	5%	10%	15%	20%

The uncertainties are estimated by considering a more pessimistic 36-month time constant together with a Sep-07 cost of 0.18 K£/KSI2K. Note that the uncertainties do not account for possible discrete steps in the prices, in either direction!

Disk Costs

Figure-2 shows the historical cost (logarithm) of Disk at the Tier-1 together with extrapolated costs for GridPP3. The first five blue diamonds from the left are Tier-1 Disk purchases between March-2002 and January 2006; the sixth point is an estimate of the cost for September 2007. The second to sixth blue point suggest that Disk costs are falling with a time-constant of 23-months. The purple squares are the future estimates used in this planning and are extrapolations using a slightly more conservative 24-month time constant from the current pricing point.



Tender Date	1-Sep-06	01-Sep-07	01-Sep-08	01-Sep-09	01-Sep-10
K£/TB	1.16	0.76	0.56	0.39	0.28
Uncertainty	0%	10%	15%	20%	30%

The uncertainties are estimated by considering a more pessimistic 30-month time constant together with a Sep-07 cost of 0.84 K£/TB. Note that the uncertainties do not account for possible discrete steps in the prices, in either direction!

Appendix-4: Tier-1 CPU Planning

Table-1 shows the current planned CPU purchase schedule. This may change if the LHC schedule is delayed.

Planned Purchases			
Date	KSI2K	Budget	K£/KSI2K
01-Sep-07	2903	£493,550	0.170
01-Apr-08	0	£0	0.149
01-Sep-08	2766	£363,589	0.131
01-Sep-09	3475	£342,729	0.099
01-Sep-10	4945	£365,933	0.074
TOTAL		£1,565,802	

Table-1: CPU purchase schedule.

Table-2 shows the net CPU available at the Tier-1. The spend is tuned to provide a minimum of 102% of the hardware request within a particular year. The additional 2% is to account for out-of-service machines but not non-capacity machines. The provision by quarter is shown in Table-2 which takes into account existing CPU and assumes (a) that CPU is available 3 months after the purchase date; (b) that 10% of the CPU is phased out at the start of year-4 of its lifetime; (c) that an additional 10% is phased out at the start of year-5; (d) that the remaining 80% is phased out at the end of year-5.

The columns in Table-2 show the hardware target; the existing (May-2007) hardware which gradually phases-out; the new hardware purchased according to the schedule in Table-1; and the net total. The final column shows the current planned-to-be-pledged wLCG numbers: The miss-match primarily reflects the slippage in the LHC schedule and the subsequent re-tuning of the GridPP profile. It should be noted that the total GridPP hardware provision shown, includes resources for the non-LHC experiments.

Capacity Table						
Date	Target	Exists	Planned	Total	Prov.	Pledges
Date	KSI2K	KSI2K	KSI2K	KSI2K	Frac.	KSI2K
01-Jan-05		787	0	787		
01-Jan-06		830	0	830		
01-May-06		830	0	830		
01-Jan-07	1684	1039	0	1039	62%	
01-Apr-07	1684	1543	0	1543	92%	1300
01-Jul-07	1684	1543	0	1543	92%	
01-Oct-07	1684	1483	0	1483	88%	
01-Jan-08	1684	1324	2903	4227	251%	
01-Apr-08	3997	1324	2903	4227	106%	5220
01-Jul-08	3997	1324	2903	4227	106%	
01-Oct-08	3997	1174	2903	4077	102%	
01-Jan-09	3997	1167	5669	6836	171%	
01-Apr-09	6313	1167	5669	6836	108%	7380
01-Jul-09	6313	1143	5669	6812	108%	
01-Oct-09	6313	771	5669	6439	102%	
01-Jan-10	6313	671	9144	9814	155%	
01-Apr-10	9313	671	9144	9814	105%	10490
01-Jul-10	9313	646	9144	9790	105%	
01-Oct-10	9313	646	8853	9500	102%	
01-Jan-11	9313	596	13798	14394	155%	
01-Apr-11	13367	596	13798	14394	108%	

Table-2: CPU provision schedule.

Appendix-5: Tier-1 Disk Planning

Table-1 shows the current planned Disk purchase schedule. This may change if the LHC schedule is delayed.

Planned Purchases			
Date	TB	Cost	K£/TB
01-Sep-07	1349	£1,025,568	0.760
01-Apr-08	0	£0	0.651
01-Sep-08	1739	£970,255	0.558
01-Sep-09	3392	£1,337,944	0.394
01-Sep-10	2725	£760,069	0.279
TOTAL		£4,093,836	

Table-1: Disk purchase schedule.

Table-2 shows the net Disk available at the Tier-1. The spend is tuned to provide a minimum of 105% of the hardware request within a particular year. The additional 5% is to account for out-of-service machines (hot-spares/broken units) and is based on experience. The provision by quarter is shown in Table-2 which takes into account existing disk and assumes (a) that disk is available 5 months after the purchase date; (b) that 10% of the disk is phased out at the start of year-4 of its lifetime; (c) that the remaining 90% is phased out at the end of year-4.

The columns in Table-2 show the hardware target; the existing (May-2007) hardware which gradually phases-out; the new hardware purchased according to the schedule in Table-1; and the net total. The final column shows the current planned-to-be-pledged wLCG numbers: The miss-match primarily reflects the slippage in the LHC schedule and the subsequent re-tuning of the GridPP profile. It should be noted that the total GridPP hardware provision shown, includes resources for the non-LHC experiments.

Capacity Table						
Date	Target	Exists	Planned	Total	Prov.	Pledges
Date	TB	TB	TB	TB	Frac	TB
01-Jan-05		200	0	200		
01-Jan-06		197	0	197		
01-May-06		166	0	166		
01-Jan-07	759	346	0	346	46%	
01-Apr-07	759	748	0	748	99%	640
01-Jul-07	759	748	0	748	99%	
01-Oct-07	759	748	0	748	99%	
01-Jan-08	759	748	0	748	99%	
01-Apr-08	1888	633	1349	1982	105%	2790
01-Jul-08	1888	633	1349	1982	105%	
01-Oct-08	1888	633	1349	1982	105%	
01-Jan-09	1888	633	1349	1982	105%	
01-Apr-09	3484	633	3089	3722	107%	3720
01-Jul-09	3484	633	3089	3722	107%	
01-Oct-09	3484	615	3089	3704	106%	
01-Jan-10	3484	570	3089	3659	105%	
01-Apr-10	6044	570	6481	7051	117%	5510
01-Jul-10	6044	570	6481	7051	117%	
01-Oct-10	6044	408	6346	6754	112%	
01-Jan-11	6044	0	6346	6346	105%	
01-Apr-11	8345	0	9071	9071	109%	
01-Jul-11	8345	0	9071	9071	109%	
01-Oct-11	8345	0	8762	8762	105%	
01-Jan-12	8345	0	8762	8762	105%	
01-Apr-12	8345	0	8762	8762	105%	

Table-2: Disk provision schedule.

Appendix-6: Tier-1 Tape Planning

A detailed tape plan has been prepared by Andrew Sansum and Dave Corney. Table-1 shows a summary of the key elements.

	2007	2008	2009	2010
Capacity [TB] Required in April	2000	2313	4368	7181
9940 Capacity [TB]	324	0	0	0
TK Capacity [TB] at start of FY	527	1676	2313	8736
TK Capacity [TB] purchased during year	1149	637	2055	0
Capacity [TB] delivered by December	2000	2313	4368	8736
Bandwidth Required in December [MB/s]	239	276	638	874
Bandwidth Provided by December [MB/s]	825	900	900	2180
Number of 9940B Bricks	6	0	0	0
Number of TK10 Bricks for bandwidth	15	20	20	20
Number of TK10 Bricks for repacking	3	3	3	3
Number of TK20 Bricks for bandwidth	0	0	0	16
Number of TK20 Bricks for repacking	0	0	0	3
Effective Bandwidth Increment per unit [MB/s]	45	45	45	80
Bandwidth Increment unit Cost [k£]	20	20	20	20
Bandwidth increment proposed [# of units]	0	0	0	0
Cost of increased bandwidth [K£]	0	0	0	0
Cumulative increase in bandwidth [MB/s]	0	0	0	0
Total bandwidth provided by December [MB/s]	825	900	900	2180
Capacity Cost [K£]	242	233	0	180
Robotic Costs [K£]	200	100	200	100
Bandwidth Costs[K£]	258	100	14	421
Bandwidth Increment Costs [K£]	0	0	0	0
Total Cost [K£]	700	433	214	701

Table-1: Summary of Tape planning.

Table-2 shows the net tape capacity available at the Tier-1 over the lifetime of GridPP3 in the same format as the CPU and Disk appendices. In this instance, the information is mostly already presented in Table-1 but is included again for completeness. The apparent over-provision in 2011 is due to the move from TK10 to TK20 technology in 2010 where the existing tapes are expected to be re-used with double the capacity. The last column of Table-2 shows the current planned-to-be-pledged numbers which are now very out of date (they were not modified in 2006 along with the CPU and Disk).

Capacity Table						
Date	Target	Exists	Planned	Total	Prov.	Pledges
	TB	TB	TB	TB	Frac	TB
01-Jan-05						
01-Jan-06						
01-May-06						
01-Jan-07	622			851	137%	
01-Apr-07	696					640
01-Jul-07	820					
01-Oct-07	1000					
01-Jan-08	1368			2000	146%	
01-Apr-08	2313					2790
01-Jul-08	2313					
01-Oct-08	2313					
01-Jan-09	2313			2313	100%	
01-Apr-09	4368					3720
01-Jul-09	4368					
01-Oct-09	4368					
01-Jan-10	4368			4368	100%	
01-Apr-10	7181					5510
01-Jul-10	7181					
01-Oct-10	7181					
01-Jan-11	7181			8736	122%	
01-Apr-11	10432					

Table-2: Tape storage provision.

Appendix – 7: Non capacity Machines

M.J.Bly
R.A.Sansum
16 March 2007

Circulation

Public

Introduction

For long range planning purposes during GRIDPP2 we have assumed that 10% of purchased capacity CPU would be unavailable for use in the batch system because:

- It may be flagged as faulty hardware and awaiting/undergoing diagnosis/repair
- It may be permanently broken awaiting cannibalisation (for out of warranty hardware)
- It may be in an automatically drained state by the automation system following the identification of an error condition (for example out of memory). Systems are auto-drained in order to minimise black holing of jobs following a fault
- It may be manually drained by the batch system – awaiting software upgrade (for example a new kernel or bios release.
- It may be allocated to non capacity duties.

Until recently the Tier-1 has succeeded in remaining within its target 10% non capacity limit, however the increasing pressure of non capacity requirements has now pushed non production capacity to over 11% (by KSI2K) and we expect it to rise to 13% by the middle of the year.

Overview

Choice of hardware to be deployed to non-production use is made appropriate to requirements of the service to be deployed, tensioned against the need to minimise loss of capacity to the experiments. Wherever possible older hardware is used, however:

- Compusys 2002 has inadequate networking (100Mb), memory (1GB) or cpu performance (1.4Ghz PIII) to meet the needs of many services.
- Clustervision 2003 Usually used for non-critical services as it is not always sufficiently reliable to run mission critical services, performance may be inadequate for some tasks. It is not on maintenance.

- Compusys 2004 The workhorse for important services. 10 systems are fitted with mirror disks to be used for resilient services – although hot swap is unavailable. It will shortly be off maintenance.

- Clustervision 2005 has only been used for 1 CPU intensive service (CE) but we are in the process of obtaining disks and sleds to enable 8 systems to be used for higher availability operations.

- Streamline 2006 – Protecting this for capacity use at present. Unsuitable for expansion with hot-swap disks.

Overall the hardware deployment is summarised as follows.

Generation	Total servers	Permanently broken	Non-capacity	Available KSI2K
Compusys 2002	156	1	16	189
Clustervision 2003	80	1	19	124
Compusys 2004	256	0	44	453
Clustervision 2006	52	0	1	262
Streamline 2006	64	0	0	545

Drained Production Hardware

There are only two systems that have had to be retired from service because they have not been worth repairing. Other hardware failures have been repaired, usually using spares purchased as and when required.
We do not have a reliable measure of what fraction of capacity is unavailable for production due to automatic or manual drain from the batch system, however we estimate that at any given time on average this is less than 10 systems (approx 2%)

Non Capacity Use

Area	Service	Hosts
WLCG CORE	UIs	6
	RB	2
	CE	2
	UK BDII	3
	LFC	1
	Site BDII	1
	Proxy	1
	FTS	1
dCache	PNFS	1
	Postgress	2
	GridFTP Doors	8
	Head Nodes	3
CASTOR	SRMS	6

Data Movers	NFS Movers	3
FabricCore	Oracle (FTS)	1
	Oracle (CASTOR)	2
	MYSQL	2
	Torque/Maui Scheduler	1
	Ganglia	1
	CACTI	1
	NIS	3
	MAIL	1
	Sysloggers	2
	Web Server	1
	System Front ends	2
VO Services	VO Boxes	4
	Objectivity DB	2
	Xrootd redirectors	2
	Xrootd test	2
	H1 centipeed	1
PPS		10
Temporary Development/Test	Ganglia	2
	CE/RB	2

There are also a further 15-20 services that use hardware no longer in the production service or other legacy systems that are not suitable for capacity use. We have not counted these up, but they include some key systems such as several systems to host the GOC, our NAGIOS master and slaves, the installation system server, and the password cracker. There will probably always be hardware too old for capacity use that remains useful for some low load and non-critical tasks.

Most of the CASTOR core infrastructure does not appear in this table as it is running on hardware purchased by CCLRC. There is no guarantee that CCLRC will continue indefinitely to fund the CASTOR core.

Future demand

We expect some modest increase in demand for hardware in the WLCG Core over the coming months as we further work on hardening the Grid infrastructure. Development work on gLite, SL4 and 64bit will also require additional hardware.
We expect to have to move some of the GOC equipment onto production hardware (for performance reasons).
We are also working to understand both the role and the requirements of the PPS service. It seems increasingly likely that the PPS service will become increasingly bound to the main production service and strategically we believe we need to increase our use of this service (once it is available) in order to understand deployment issues and test new releases. We expect the PPS will need to grow to 15-20 systems over the next 6 months.
Overall, growth in demand is expected to be modest over the remainder of 2007 and may even reduce once dCache systems are released.

Future Hardware Profile

Pressure on the hardware is increasing as recent purchases have been for relatively few systems (only 116 from a combination of FY05 and FY06). In principle we might be able to reduce demand for hardware by the use of some form of virtualisation, however:

1. Recently we have had to split several services over multiple instances of common hardware to improve performance.

2. Reliability may be impaired if unstable services share common hardware.

3. Virtualisation will increase the demand for high performance systems (but less of them).

Plans for purchases in 2007 and 2008 are not yet firm, however very rough estimates based on a guess of 60% of the original GRIDPP3 bid suggest that CPU system count may double by the end of 2007 alleviating pressure on hardware once more.

Hardware Purchasing Strategy for Service Systems

In the past we have almost entirely relied upon tasking batch workers to be service systems and to some extent we will continue to need to do this. However there is increasingly a divergence between the batch worker hardware specification:

- Features are being dropped from batch worker nodes (e.g., the Woodcrest servers cannot accommodate a second disk drive for resilience)

- There are increased requirements for resilience on some of the service nodes where we may wish to specify hot swap drives (which would have avoided the RB downtime) and redundant power supplies.

Some service systems that would have been purchased as batch worker nodes may need to be purchased as more specialised hardware raising the cost of these systems slightly.
We have also considered the possibility of buying significantly cheaper hardware for less critical low end service systems. However it is likely that it is usually cheaper to make use of old commodity hardware than to buy new cheap rack mount systems.

Conclusion

Non productive capacity hardware now amounts to approximately 11% of capacity (including 2% drained hosts). We expect it to climb a little to 13% by the summer but decline to 12% once dCache releases capacity. We do not expect it to fall below 10% until the next procured capacity comes online.
During the next procurement, we should consider purchasing some specialised hardware optimised for resilient running rather than relying entirely on batch worker nodes.

Appendix-8: Tier-2 Hardware Costing

Cost/Value of the Tier-2 Hardware

As a starting point, the cost of purchasing additional hardware to meet the Tier-2 requirements can be calculated within the Tier-1 costing model. The result is £2.05m for CPU and £2.75m for Disk, giving a total of £4.80m. However, this calculation ignores the fact that at the end of GridPP3 (March 31st 2011) much of this equipment would still have a significant lifetime remaining and thus have some residual value. The model of Tier-2 hardware is one where GridPP buys a resources service (an amount of resource and a level of service) from Tier-2 resource suppliers. A better estimate of the value of this hardware to GridPP, therefore, is the depreciation value of the hardware over the lifetime of the project. There is a tacit assumption here that the Tier-2 agreement to provide resources to GridPP terminates at the end of the GridPP3 project with no further obligations to provide access to resources which may have been funded with GridPP money.

Depreciation Value of Tier-2 Hardware

A simplistic linear depreciate of the equipment costs over a fixed period is not applicable here. Although GridPP has a reasonable understanding of the lifetime of CPU (a little under 5-years) and Disk (a little under 4 years), a linear depreciation model would produce anomalies because the underlying hardware prices are falling exponentially in terms of costs per TB or KSI2K. The value calculated from a linear depreciation model would underestimate the value in the early years and over-estimate the value in later years and, if GridPP were to base Tier-2 hardware funding on such a model, much of the value of hardware would not be funded in the lifetime of the GridPP project. Therefore, a model has been built which depreciates the hardware value exponentially with the expected 24-month halving period over the expected lifetime of the hardware. This model has been used to look at the Tier-2 resources required and to calculate a depreciation value per TB and per KSI2K by year. The result of this is an estimation of the value of Tier-2 CPU of £1.3m and Tier-2 Disk of £1.6m, giving a total "depreciation value" of £2.9m. A linear depreciation model gives a total of £2.1m.

GridPP Tier-2 Funding

Science Committee has approved the PPRP recommended figure of £4.35m for Tier-2 hardware, which is 53% above the exponentially depreciated value calculated above, and is 90% of the estimated full purchase cost. Therefore, it is reasonable to conclude that GridPP offers a premium (~50%) to the depreciation value in recognition of running costs, but less than the purchase cost (~90%) in the understanding that no obligations remain beyond the end of GridPP3. GridPP also offers the Tier-2s staff funding of £2.93m. This is more than the cost of providing the additional staff that would be necessary to operate the equipment were it co-located at the Tier-1. Running costs at the Tier-1 are 30-50% of the purchase costs (depending on whether spending on Tape infrastructure is included in the calculation).

In summary, GridPP views the package of Tier-2 Staff and Hardware funding as a reasonable offer for the provision of the resources specified, including running and operating costs.

Appendix-9: Tier-2 Hardware Allocation

In the GridPP3 Proposal the Tier-2 hardware was allocated to institutes using a number of metrics such as the number of LHC physicists, existing CPU and disk resources, CPU delivery to date and disk currently used. It is proposed to allocate the final Tier-2 hardware, using a formula to be agreed, in two tranches, one in September 2007 and one in September 2009 using metrics from the previous 6 months. The metric is likely to be based on

SAM Test Efficiency x (CPU Delivered + Disk Available)

The exact formula is still under discussion.

Appendix-10: Tier-2 Staff Allocation

In the GridPP3 Proposal we asked for 14.75 FTE of Tier-2 Hardware Support staff, and were awarded 13 FTE, compared with 9.5 FTE currently. In the Proposal the staff were allocated to institutes using a number of metrics such as the number of LHC physicists, existing CPU and disk resources, CPU delivery to date and disk currently used. This was done over a year ago so the whole process was recently repeated using similar metrics but with up to date values. The Tier-2 Board met on 5 June 2007 to discuss this. There were strong reservations from SouthGrid about basing the allocations purely on past performance and current capacity but the methodology was agreed by the Tier-2 Board and the following allocations to the Tier-2s were made:

London	4.00 FTE
NorthGrid	4.50 FTE
ScotGrid	1.75 FTE
SouthGrid	2.75 FTE

The Tier-2 Board also recommended a breakdown by institute within each Tier-2 but it was left to each Tier-2 to re-negotiate this if they want to within their Tier-2. The PMB endorsed this procedure at the Face to Face meeting on 8 June 2007.

The four individual Tier-2 Management Committees subsequently confirmed the internal breakdown of these posts as:

Brunel	0.50 FTE
Imperial	1.25 FTE
QMUL	1.25 FTE
RHUL	0.50 FTE
UCL	0.50 FTE
Lancaster	1.25 FTE
Liverpool	0.75 FTE
Manchester	2.00 FTE
Sheffield	0.50 FTE
Durham	0.25 FTE
Edinburgh	0.50 FTE
Glasgow	1.00 FTE
Birmingham	1.00 FTE
Bristol	0.45 FTE
Cambridge	0.45 FTE
Oxford	0.20 FTE
RAL PPD	0.65 FTE

Appendix-11: Outreach Plans

GridPP3 – Dissemination

Dissemination funding for GridPP2+ and GridPP3 has been cut substantially. GridPP currently employs a Dissemination Officer and an Events Officer, totaling 1.5 FTE. For GridPP3 and GridPP2+, this has been reduced to 0.5FTE in total.

GridPP2+

Sarah Pearce, the GridPP Dissemination Officer, currently works half time. GridPP's grant for GridPP2+ includes 0.5FTE to continue this post. In order to fund the Events Officer post (1FTE), it is proposed to use the remaining money from the GridPP2 dissemination budget, which will cover 1 FTE salary costs from September 07 – April 08.

GridPP3

At the start of GridPP3, the current Dissemination Officer will move to become Project Manager. We aim to employ a full-time Dissemination Officer for GridPP3, to work on both dissemination and events. The GridPP3 grant provides funding for 0,5FTE towards this. Our plans for GridPP3 are:

Working with NGS

NGS have recently appointed an Outreach Officer at Manchester University. As GridPP and NGS become more aligned, we will work closely with NGS outreach to disseminate Grid computing in the UK, including through press releases, events and literature.

EGEE

Bids are currently being submitted for EGEE-III, which starts in April 08. The outreach activity in EGEE-III (NA2) will be run by six 'competence clusters'. UKI has been encouraged to bid for one of these, covering 'Press and PR'. This involves building relationships with the press across the EGEE regions and sending out press releases, which we have extensive experience of during EGEE-II and GridPP. With both GridPP and NGS integrated into EGEE, there will be substantial overlap in the three bodies' press work.

The press and publicity cluster requires 2FTE. We therefore plan to use the GridPP Dissemination Officer 0.5FTE as matching funding for 0.5FTE of EGEE money. NGS will use their Outreach Officer funding in a similar way, resulting in a total of 2FTE.

Funding to support the Real Time Monitor was also cut in GridPP3. UKI has also bid for 0.5 FTE at Imperial College from EGEE to support the RTM, with the encouragement of EGEE management.

Contingency plans

Should UKI's bid to lead press work for EGEE be unsuccessful, we plan to fund a full-time Dissemination Officer for 1.5 years using GridPP3 funding. We will then look at options for funding during the final 1.5 years, including obtaining sponsorship from industry, bidding for STFC science and society funding or bidding jointly with NGS for outreach activities.

Sponsorship
GridPP's Oversight Committee has suggested funding some of our dissemination work with sponsorship from industry. There are a number of options for this, including sponsoring GridPP as a whole; sponsoring a project within GridPP, such as LHC@home; paying costs for a person; sponsoring single events; or funding dissemination awards. We plan to identify one or two business with which we have good relations and suggest a range of possibilities.

Appendix-12: Previously Un-Costed Posts

Posts defined in the GridPP3 proposal

The following two posts were defined in the GridPP3 proposal but were not included on any JeS form because the institute was not then known:

Post-NE8: Administrative Assistant (18 project months):
It is proposed to continue the present administrative assistant (50% FTE) from 1/Apr/08 for a period of 36 months. Continuation from 1/Sep/07 to 31/Mar/08 was previously approved by the PPRP and is included in Work-package E.

Post-NE9: UB-Chair (9 project months):
The GridPP3 submission included a proposal to fund 25% of an FTE for 3 years to fill the crucial role of chair of the User Board. We believe the Oversight Committee recognizes the importance of this role (Paragraph 6.1 of the minutes of the 9th meeting on June-30th, 2006). It is proposed to fund this role at RAL from 1/Apr/08 for a period of 3 years

Transition funding identified for Middleware Support

Post-NC28: VOMS-Service (12 project months):
The VOMS service is essential to the function of the GridPP infrastructure. The removal of the post a Manchester has been addressed by negotiating a phased transition of responsibility from GridPP to NGS to provide VOMS service for the UK (but is subject to funding for the NGS beyond 2009, which is not yet established). GridPP proposes to fund 50% of an FTE for 24 months from 1/Apr/08 with a clear plan to converge on a common service with the NGS.

Post-NC29: Real Time Monitor (RTM) (3.5 project months)
The RTM has become the public face of the Grid and has been the show piece of GridPP demonstrations for several years. Strong pressure has been received from EGEE and LCG to continue this work. It is now viewed as highly likely that funding will be received from EGEE-III starting in April-08. GridPP proposes to bridge the gap from Sep-07 with 7 months of 50% FTE funding.

Post-NC30: Networking Transition (3.5 project months):
The reduction in Networking support by the PPRP both in GridPP3 and the transitional GridPP2+ period endangers the deployment of an operational tool in time for LHC startup.
The requested additional effort will be used to tailor the capability of GridPP Gridmon interface to access additional sources of network performance data. In so doing Gridmon will provide the GridPP community access to a far greater granularity of detail on
the end-to-end path with respect to network performance monitoring and allow individual sites to understand better network performance issues that affect them directly.

Post-NC31: Metadata Transition (6.6 project months):
ScotGrid funding from April-2008 has been identified to help address the reduction of Metadata support at Glasgow. Some of the transition period can be funded from existing GridPP2 funds but it is estimated that there is a shortfall of 6.6 FTE-months.

Posts identified for Experiment Support

The reduction of Grid Support both on the GridPP side (PPRP cuts to the GridPP3 proposal) and on the Experiment side (reduction in the Rolling Grant posts) has lead to a support gap in the crucial area of integrating Experiment applications and the Grid. GridPP has consulted with the three major LHC experiments as to how to best address this problem. The experiment input is contained in Appendices 14, 15, and 16. In response GridPP proposes to fund the following posts[1]:

Post-ND11: ATLAS Tier-1 Operation Support (50% FTE for 3 years; 18 project months)
This post will support reprocessing at the Tier-1; support data movement within the UK Tier1 - Tier2 cloud; and support ATLAS Grid-based analysis.

Post-ND14: ATLAS GANGA Support (100% FTE for 3 years; 36 project months):
This post will maintain, develop and support Ganga. The tasks to be taken on by the post are: new developments on the Ganga plug-ins specific to ATLAS as the ATLAS computing model and data management tools adjust to the life of real data taking, user support and training of the support shift helpers.

Post-ND13: CMS Tier-1 Support (150% FTE for 3 years; 54 project months):
This effort will be expended as (a) a 0.5FTE contribution to a whole post shared with the RAL Tier-1 centre; this post will focus on technical operations at the Tier-1 required to support CMS services at the centre, including operation of phedex and specialised storage management and workload management services; (b) a 1.0FTE post at RAL PPD managed within the CMS collaboration, taking on the Tier-1 Data Manager role and carrying out data operations at the centre in liaison with Tier-1 staff, include event reprocessing, skimming and bulk analysis on behalf of the Uk and wider collaboration.

Post-ND12: LHCb Tier-1 Support (100% FTE for 3 years; 36 project months):
The focus of this post will be to provide full technical support for reconstruction, stripping and analysis activities at the UK Tier 1 centre. Tasks will include the optimisation of the mass storage systems during data taking and reprocessing, continued evolution of the DIRAC workload management system, development of LHCb-specific services (e.g. those currently covered by VO boxes), monitoring of processes/sites and user support. Support will also be given to Monte-Carlo production at UK Tier 2s (which use the Tier 1 for storage) to ensure high efficiency.

[1] Note Added: ATLAS are engaged in a complimentary internal effort to address this shortfall and may request to modify the details of the posts within the funding envelope proposed. CMS have indicated that they would like to start 1 FTE in Sep-07 rather than Apr-08 as currently proposed; LHCb have confirmed that they are happy with the proposal as it stands.

Post-ND15: LHCb GANGA Support (50% FTE for 3 years; 18 project months):
This post will provide Ganga maintenance, development and support. The tasks to be taken on by the post are: new developments on the Ganga plug-ins specific to LHCb as the LHCb computing model adjusts to the life of real data taking, development of the Ganga core to reflect new usage patterns by an expanding user base, and user support and training. Another component will also be to provide initial support to Ganga uptake outside HEP, such as the current use within the Cambridge Ontology project with support from the STFC KITE club.

Post Costings

		FTE	Start	Duration	Effort	Cost
Previously Identified Posts		**Frac**	**Date**	**(Months)**	**(Months)**	**Estimate**
NE9	UB Chair	25%	01-Sep-07	43.00	10.75	
NE8	Administrative Assist.	50%	01-Apr-08	36.00	18.00	
	SubTotal			**79.00**	**28.75**	**£126,113**
Support Transition Posts						
NC28	VOMS Service	50%	01-Apr-08	24.00	12.00	
NC29	RTM Transition	50%	01-Nov-07	7.00	3.50	
NC30	Networking Completion	50%	01-Sep-07	7.00	3.50	
NC31	Metadata Transition	100%	01-Oct-07	6.60	6.60	
	SubTotal			**44.60**	**25.60**	**£172,493**
Experiment Support Posts						
ND11	Atlas T1	50%	01-Apr-08	36.00	18.00	£107,500
ND12	LHCb T1	100%	01-Apr-08	36.00	36.00	£215,000
ND13	CMS T1	150%	01-Apr-08	36.00	54.00	£322,500
ND14	Atlas Ganga	100%	01-Apr-08	36.00	36.00	£215,000
ND15	LHC Ganga	50%	01-Apr-08	36.00	18.00	£107,500
	SubTotal			**180.00**	**162.00**	**£967,500**
	Grand Total			**303.60**		**£1,266,107**

Appendix-13: Middleware Transition Planning

[Awaiting Draft – RM]

Appendix-14: ATLAS Input

2 May 2007

ATLAS UK Applications Support in GridPP3 and the Rolling Grants

In response to the invitation from GridPP and a request from STFC that the situation be outlined to the ATLAS Oversight Committee, this document outlines the situation regarding ATLAS eScience applications and computing support in the GridPP3 era.

Outline

In the last Rolling Grants round, the GridPP applications posts and the posts in the ATLAS eScience programme were moved into the core programme and considered by the PPGP rather than the PPRP as before. The cases for the continuation of each post came from the experiments, as requested, and were repeated in the group submissions. The Core Computing is divided into two components; Category A M&O, which is an STFC commitment, based on the fraction of qualifying authors in the UK; and Category B, which is based on expertise and the 'service' element required of collaborators, but is not calculated on a formulaic basis. ATLAS estimated that between 2008 and September 2011, 45 staff-years of effort are required from the UK in Category B, and the Category A requirement will remain of order 5.5 staff-years.

In total, almost exactly 50% of the 21.5 e-Science and GridPP Applications posts requested have been funded. The continuing contribution to computing and computing operations will be lower, however, as the posts have varying and increasing fractions of effort on exploitation, as might be expected.

This outturn has had a particularly negative effect on core computing activities such as the validation framework and the distributed analysis interface, GANGA. The loss is both in the effort just as it is most needed (given the delays in the LHC schedule) and in the skills of the personnel concerned. It has had a particularly hard impact on the Category A M&O in-kind contribution to the ATLAS core computing, but also on the more technical aspects of the M&O B contributions.

Core Computing Category A M&O Activities

The M&O A contributions to computing (agreed by PPARC in the October 2005 RRB). These are calculated on the basis of the fraction of qualified authors under each funding agency, and can be delivered either in cash or in kind. Until the Summer of 2007, the UK contribution is covered completely in kind using ATLAS eScience and GridPP effort. By construction, the activities are those suited to computing specialists and not general physicists, which is why the levy is required.

It is not entirely surprising, given its remit, that in almost all cases the posts lost are more 'hard core' computing, and especially those providing the Category A. The yearly requirement from the UK in Category A is almost exactly 1.56 FTEs; after the grants round, only 0.35FTEs survive. Without other action this would lead to a direct annual charge to PPARC of about 100kCHF; it also means that important UK products such as the Run Time Testing framework will be under-resourced without new effort being found.

ATLAS UK has re-examined the programme, and attempted to mitigate the effect of this shortfall. In 07/08, flexibilities in budgets and gap savings have allowed some posts to be extended. After discussion with the PIs and the ATLAS central management, some funded RA effort is being moved to Category A (with a loss of Category B contributions). *It is to be hoped that this flexibility is recognised by future grant assessors and not penalised; while a ring-fence is not appropriate, we feel some explicit guidance would be fitting for future Grants Panel members.* Such adjustments will have some negative effect on other exploitation tasks, but are required for the greater good.

However, the scope for flexibility is limited, and the skills needed for various tasks are not present in the community in all cases. As a result of this process, we believe that the full Category A contribution can be met in-kind in 07/08, but that there will be a shortfall beyond that point of about 0.4 FTEs, falling in the area of the validation framework (RTT) and installation support. RTT will be reduced to ~0.4 from 1.4FTE per year. The software installation and environment effort will fall from 1.2 to 0.2 FTE. (The reason the losses are greater than 0.4FTE is that we are supplying additional effort in other Category A areas.) An option that would allow some of the skilled effort to be retained would be to negotiate PDAS positions with CERN and to fund the UK employer costs with additional monies from STFC. This would equate well with the cash sum that would be required to be paid to CERN otherwise, and the ATLAS management would prefer the skilled person rather than cash.

Core Computing Category B M&O Activities

In the category B area, the areas of the core operational effort and the distributed analysis system were particularly badly hit. Other areas were also reduced, some, such as ATLANTIS, with UK leadership, but are less mission critical.

Core Operations

Core operations have two components: the effort required 'centrally' for the operation of the ATLAS simulation and reconstruction; and the effort required for running the same operations (essentially at the Tier 1, but the staff are often elsewhere), along with the data placement within the UK Tier 1 and Tier 2s, and user support for the analysis system. The central operations component counts towards Category A, and this is an activity we have used to mitigate the problems in that area. The UK has made a modest contribution to the operations at CERN in the past; this contribution is vital, as it allows direct knowledge transfer into the vital UK 'cloud' ATLAS operations effort.
Both components are facing a rising requirement in terms of effort, and the underlying production system still requires additional development. The UK planned to contribute to the

development effort with expertise freed from running experiments, but this has not proven possible given the Grants Round results.

UK Operations

UK operations for ATLAS itself involves several components:

- Support for the reprocessing at the Tier 1.
- Support for the data movement within the UK Tier 1-Tier 2 cloud.
- User support for ATLAS Grid-based analysis.

In addition, the ATLAS UK relies on technical support effort at the Tier 1 and Tier 2s, funded by GridPP.

The technical support at the Tier 1 has been at the level of 0.5FTE from GridPP. There has been no Tier-1 based support for the ATLAS Tier 1 activities, but we have had about 0.25FTEs elsewhere from ATLAS eScience. This is clearly inadequate. Experience indicates that ~1FTE of expert and dedicated effort is required associated with the Tier 1 operations. Approximately 0.25FTE of operations effort will be gained from proposed shift system, which will count towards the service element for authorship. However, the co-ordinating 0.25 FTEs is increasingly required for the data movement activity below, and is not best sited for the operations work. There still remains a shortfall of ~0.5FTEs of ATLAS skilled effort at or associated with the UK Tier 1; this should come from a single individual, who should have a significant level of expertise.

The data movement operations will also be covered by a shift activity, with a similar level of effort. This is increasingly being co-ordinated by effort previously managing the production activity in the UK. This will, in time, be complemented by a third shift to organise UK computing user support.

Distributed Analysis/GANGA

The UK provided the initial funding for the GANGA distributed analysis system, GANGA. The core of the system is shared with LHCb, but there a considerable fraction of the project that must be ATLAS-specific. The UK investment has been used to obtain 'matching effort' from elsewhere. ATLAS requested 3.5 staff-years of ATLAS-specific development effort in this area, backed by 3.5 staff-years from LHCb; this was reduced to no ATLAS-specific effort and 1.7staff-years of core effort. ATALS also requested 3.5 staff-years of effort for the analysis system user support. This requires both highly-expert effort and less exert effort. The less-expert component can be realised through the user support shifts (0.75 staff-years), but a significant shortfall exists in the expert effort. We estimate that a minimum of 3.5 staff-years of dedicated effort is still required to cover the development, maintenance and user support short fall.

The sudden collapse in GANGA support in mid-2007 is particularly ill-timed. GANGA has in the last 6 months taken-off in the target user community, which has generated both new user requirements and a large support load. The middleware on which it sits is still evolving, and requires changes in the GANGA layer. Most importantly, we are now expecting real data in 2008, and this will place new challenges.

Priorities

The two main priorities of ATLAS UK concerning GridPP and the Grid are:

1. ATLAS operational support for the UK Tier 1 centre, ~0.5FTE.
2. Ganga maintenance, development and support. The tasks to be taken on by the post are: new developments on the Ganga plug-ins specific to ATLAS as the ATLAS computing model and data management tools adjust to the life of real data taking, user support and training of the support shift helpers, ~1FTE.

Concerning out broader core computing needs, our priorities are:

1. Validation framework development and maintenance (RTT), ~0.4FTE.
2. Software installation and environment effort, ~0.5FTE.

Roger Jones (Lancaster) - on behalf of ATLAS (UK)
2nd May 2007

Appendix-15: CMS Input

CMS Computing Effort 2007 – 11: Input to GridPP3 planning

DMN 11/5/07 – Draft 1 - Not for circulation

Summary

We identify an urgent need for 1.0 – 1.5FTE of expert CMS data operations effort at the Rutherford Tier-1, which is currently unfunded. We describe the associated work programme, and identify key responsibilities in this area. We do not propose here routes by which GridPP may be able to steer and prioritise resources to meet this need, though we welcome the opportunity to supply such input if required. We note that points made in this document have been fully discussed within both CMSUK and the CMS computing management, and have been presented to the CMSUK Oversight Committee.

CMSUK computing programme

CMSUK has contributed to the software and computing programme of CMS in several areas during the construction phase. Key amongst these have been:

- **Data management software**: The UK is the originator of PhEDEx, the tool handling all higher-level data transfer and tracking. This is a vital product for CMS operations, which the UK leads and must support.

- **Workload management software**: The UK has contributed to the development of BOSS, the main interface between the various Grid WMS systems and higher level job submission tools, again a vital component.

- **Tier-1 centres**: The UK provides Tier-1 services to CMS via the centre at RAL; CMSUK has been active for many years in attaining a close working relationship with the centre staff, and in testing and exploiting the centre services.

- **Tier-2 centres**: CMSUK institutes provide reliable Tier-2 centres with growing resources, each of which is in use for both Grid-based analysis and simulation.

- **Management**: The UK has contributed to the management and coordination of CMS software and computing for many years, currently providing the chair of the Computing Resource Board, one of the Integration Coordinators, and the leader of PhEDEx and the web tools group.

- **User support and documentation**: CMSUK has been proactive in supplying support, documentation and tutorials for CMS, with the original ‘CMS workbook’ concept coming from the UK.

This programme is substantial compared to the size of the UK collaboration, and is driven by the strong expertise in computing at each CMSUK institute, typically shared between CMS and other

experiments. The UK has been one of the most active contributors to CMS computing, and it has always been fully recognised within the UK collaboration, and within UK funding agencies, that the computing programme forms a vital part of the overall CMSUK enterprise, underpinning both the UK physics programme and that of the collaboration as a whole.

Funding outlook

The UK computing project has averaged around 6FTE of effort over the last three years. Even at this level, there has often been tension between development and operational tasks, e.g. during data challenges. The effort is currently drawn from a variety of sources, including: GridPP2 applications area funding; e-science funding; and group rolling grants.

Given the strong history of computing in CMSUK, and the critical nature of this area in the exploitation phase, it was our expectation that the majority of e-science and GridPP2 posts would be renewed in the last rolling grant round. The grant applications from institutes reflected this, with around 7FTE of non-academic effort requested in total. In fact, only 2.5FTE were renewed, with 1.5FTE of this in the core programme of groups, in recognition of expertise present within the collaboration. Of the 3.0FTE funded through GridPP2, no posts were renewed; it is therefore possible that the 'funding gap' is due in part to a misunderstanding of the future of such posts within GridPP3.

We are already experiencing a large shortfall in effort that will become significantly worse in April 2008, beyond which there is no possibility to reprofile either e-science of GridPP2 funds. The UK collaboration has attempted to deal with the 2007/8 shortfall in several ways:

- We will focus upon the core deliverables of the UK project: PhEDEx / data management, and centre operations. The UK contribution to central documentation and support has already been dropped, and is sorely missed. We will not renew our contribution to BOSS.

- Remaining e-science and GridPP2 funds have been diverted to support key posts until April 2008, to prevent the loss of essential staff.

- Additional resources at the 1.0FTE level have been found within rolling grants (diverting resources that would otherwise have been expended on physics exploitation, with the support of the CMS OsC)

- 0.5FTE of effort will be steered within the Tier-1 centre to more experiment-focussed activities.

However, there is still an absence of RAL-resident effort required to run the Tier-1 centre for CMS-specific operations. As of April 2008, there will be only 0.5FTE available to carry out data operations at the Tier-1, a factor of 3-to-4 too low.

CMS Global View

CMS is currently defining M&O responsibilities in the software and computing area. Key in this process is the definition of the data operations plan; 'data operations' encompasses all steps in the movement and processing of data from the interface to the online system, to the users' desktop. This overall programme contains components that are initially the responsibility of central teams at CERN

and FNAL, and other aspects that are the responsibility of the remote Tier-1 and Tier-2 centres. The UK has full engagement in the definition of this plan.

The current estimate for the effort required in the UK for data operations (in addition to technical support and operations of services) is 2.0FTE at the Tier-1 centre, and 1.0FTE across the UK for the Tier-2 centres. This does not include the general contribution of the collaboration to offline shifts, etc, or the support of specialised software and services such as PhEDEx. This effort is entirely classed as Category B M&O; the CMSUK pro-rata contribution to M&O Cat A is below 1FTE, and a UK funding line must be identified through which to make the equivalent cash contribution. We note that the UK Cat B contribution is consistent with a pro-rata contribution to a project requiring over 100FTE for M&O.

It is therefore clear that we are missing around 1.5FTE from the absolute minimal team which could meet UK responsibilities, and that this problem particularly affects the Tier-1 centre operations.

Proposal to GridPP

The UK plan for LHC computing, which is intended to enable us as effective contributors to LHC analysis, can only succeed if a correct balance is found between technical resources (hardware, middleware, networks) and human resources. This is the central tension within the project plan, and we recognise that every effort is being made by GridPP management to optimise the balance.

Expert effort is required to operate both the 'generic infrastructure', and to perform tasks directly on behalf of the experiments. It is our belief that this latter effort must be embedded in the experimental programme to be fully effective. We require staff who are CMS collaborators, and under CMS management control, while following our established model of a closest possible working relationship with centre operations staff.

We therefore encourage GridPP to identify a route through which 1.0-1.5FTE may be funded at the Tier-1 centre for CMS operations. This effort would be directed towards the following work programme:

- Support of data operations for physics at RAL; monitoring of workflow; identification of job failure modes and data access problems, and their remedy; proactive investigation of problems, and optimisation of the centre's efficiency by adjustment of scheduling, caching and data-access algorithms.

- Management, and allocation between competing data-processing tasks, of resources allocated to CMS in pursuit of the physics goals of the experiment; maintenance and curation of the resident datasets and databases; recovery of data due to technical failures at the site.

- Operation, management and debugging of data transfer, data consistency checking, and data curation tools.

- Technical liaison with centre operations staff, and with the overall CMS data operations sub-project, to plan and execute the CMS reprocessing programme at the centre.

- Technical liaison with centre staff to specify, develop, test and deploy new hardware and systems, and to ensure that they meet CMS requirements as data volumes increase by more than an order of magnitude.

- Pragmatic support of local and remote CMS users who (exceptionally) use the Tier-1 centre for a large-scale or urgent analysis project.

- Liaison with, and support for, Tier-2 centres relying upon the Tier-1 for data access (NB: in the CMS model, this implies contact with a wide range of Tier-2 centres worldwide)

- Technical contribution to the ongoing development of CMS tools and facilities that will be required during the very steep ramp-up to design luminosity and beyond.

We emphasise three key points relating to this work plan. Firstly, that the programme of work is extremely close to that carried out in the operation of the BaBar Tier-A centre in the UK, albeit at a much larger scale. In fact, the Tier-1 programme is managed at several levels by ex-BaBar collaborators, exploiting their experience of providing a reliable service for a running experiment.

Secondly, we note that the programme of work is not simply an operations task aimed at providing a response to a relatively constant and well-anticipated demand. Instead, it will involve a very sharp ramp-up in performance and scale as the LHC dataset increases, bearing in mind that it is the entire accumulated dataset that must be made accessible to the collaboration. This will require a great deal of technical development and deployment of new resources, in addition to the day-to-day tasks. We expect the Tier-1 to be confronted with an overwhelming demand for data on a continuous basis, requiring absolutely robust hardware and software systems.

Thirdly, we recall that the hardware resources at the Tier-1 are sufficiently constrained (likely to be <70% of the CMS request in the GridPP3 proposal) that we are at the threshold for efficient Tier-1 operation. This implies, primarily, that resources must be used with extremely high efficiency; but also, that a more demanding data access model, based upon high speed tape IO, will be required compared to that at CMS centres with the ability to store entire large datasets on disk. These issues will be particularly important in the first year of CMS running, during which data access will necessarily be less well-coordinated than in later years as we find out feet and study intensively the performance of detector and machine. The particularly strong liaison between storage management experts and RAL-affiliated CMS staff must be continued if we are to meet this key challenge, which implies adequate funding on both sides.

Appendix-16: LHCb Input

27 April 2007

LHCb UK Applications Support in GridPP3

Universities of Bristol, Cambridge, Edinburgh, Glasgow, Imperial College London, Liverpool, Oxford and RAL

In response to the invitation from GridPP, this document outlines the situation regarding LHCb eScience applications and computing support in the GridPP3 era.

Outline

At the last grants round, LHCb UK estimated that, over the period April 2007 to October 2011, there was a requirement for 30 FTE to provide the UK share of the core computing for the experiment and to support UK computing. This was split between the areas of core computing (20.4 FTE) and UK computing (9.6 FTE). It should be noted:

- The LHCb Computing Management has estimated that approximately 37 FTE and 33 FTE are required to support the core computing activities of the experiment in 2007 and 2008 respectively. Taking the UK's authorship share (currently 18.6%), this indicates a UK contribution of 6-6.5 FTE/year for core computing alone, which is in-line with the request for a total of 20.4 FTE.
- In order to host a Tier 1 centre for LHCb, it is a condition that **at least** 0.5 FTE of support effort is continuously provided by the host country. However, experience with data challenges in the UK has shown that 1 FTE is a more realistic estimate, especially for the early exploitation of LHC.

Despite the above, the PPGP recommended the funding of just 12.1 FTE in total for LHCb until October 2011 to cover both core computing (6.1 FTE) and UK computing (6 FTE); averaging to just 2.7 FTE/year. Furthermore, the PPGP funding did not include support for the Tier 1 post. This potential lack of effort from the UK has caused serious concern within the experiment. If the UK has to provide all the expected level of computing effort from within current resources, then the physics exploitation and detector (RICH and VELO) related software, for which the UK is also expected to deliver, will be seriously comprised. This will increase the risk to the UK's deliverables and have a significant negative impact on UK influence and leadership.

In order to alleviate this problem, LHCb UK has reviewed the available UK staff effort in order to divert some into core computing and instigated a review of the Category A M&O for the experiment. Currently, LHCb does not have Category A M&O associated with computing and the LHCb management has agreed to review the situation in the context of other "service tasks" that will be required for the successful operation of the experiment. Even so, there is

still a major concern for the high priority UK effort outlined below which can not be provided for through the current UK effort or Category A M&O.

Priorities

In the LHCb computing model, the six external Tier 1 centres are used for data reconstruction, event stripping and analysis, with the Tier 2 centres dedicated largely to Monte-Carlo production. The Tier 1 centres will therefore play a vital and essential role in the exploitation of LHCb data and the collaboration views the provision of a technical support post for the UK Tier 1 centre as the **highest priority**. This post is currently located at RAL through an eScience grant that terminates in April 2008. This means that essential expertise, built up over a period of 30 months, will be suddenly lost just before the main LHC run in 2008.

In addition, the Ganga project to provide a Grid user interface for both the LHCb and ATLAS experiments has become a flagship application of the GridPP collaboration. This has been a largely UK led project, since the original concept in 2001, and the current project leader is a UK LHCb physicist. Due to the funding for two posts in GridPP1 and GridPP2, it has been possible to leverage additional effort from the ARDA group at CERN to work on Ganga. The PPGP, however, terminated one of the GridPP supported posts from September 2007 and extended the other (via the grants line) for only 20 months. This in turn has unfortunately placed the future of the CERN supported posts in jeopardy since these are only provided on a "matched" basis. Ganga is widely used and will play a central role in exploiting the Grid for physics analysis. The LHCb collaboration views the solution of continued UK support for Ganga with **high priority**.

In view of the above, the two main priorities of LHCb UK are:

3. LHCb technical support/coordinator for the Tier 1 centre. The focus of this post will be to provide full technical support for reconstruction, stripping and analysis activities at the UK Tier 1 centre. Tasks will include the optimisation of the mass storage systems during data taking and reprocessing, continued evolution of the DIRAC workload management system, development of LHCb-specific services (e.g. those currently covered by VO boxes), monitoring of processes/sites and user support. Support will also be given to Monte-Carlo production at UK Tier 2 centres (which use the Tier 1 for storage) to ensure high efficiency.
4. Ganga maintenance, development and support. The tasks to be taken on by the post are: new developments on the Ganga plug-ins specific to LHCb as the LHCb computing model adjusts to the life of real data taking, development of the Ganga core to reflect new usage patterns by an expanding user base, and user support and training. Another component will also be to provide initial support to Ganga uptake outside HEP, such as the current use within the Cambridge Ontology project with support from the STFC KITE club.

We believe that two separate posts are **strongly** justified for these activities and this would indeed be our preference. However, we recognise the limitations of current funding and would welcome discussion on other proposals that would provide funding for either, both or a combination of the above posts. We have resisted the temptation of submitting a long list of

posts and are presenting what we consider the absolute minimum requirement in the hope that with some flexibility a pragmatic solution can be found.

Finally, we should like to take this opportunity to thank GridPP for attempting to address and respond to the funding shortfalls in software and computing for the LHC experiments.

Glenn Patrick (STFC) - on behalf of LHCb (UK)
27th April 2007