



LHCb requirements for GridPP5

Andrew McNab
University of Manchester



Looking forward to 2015 and beyond

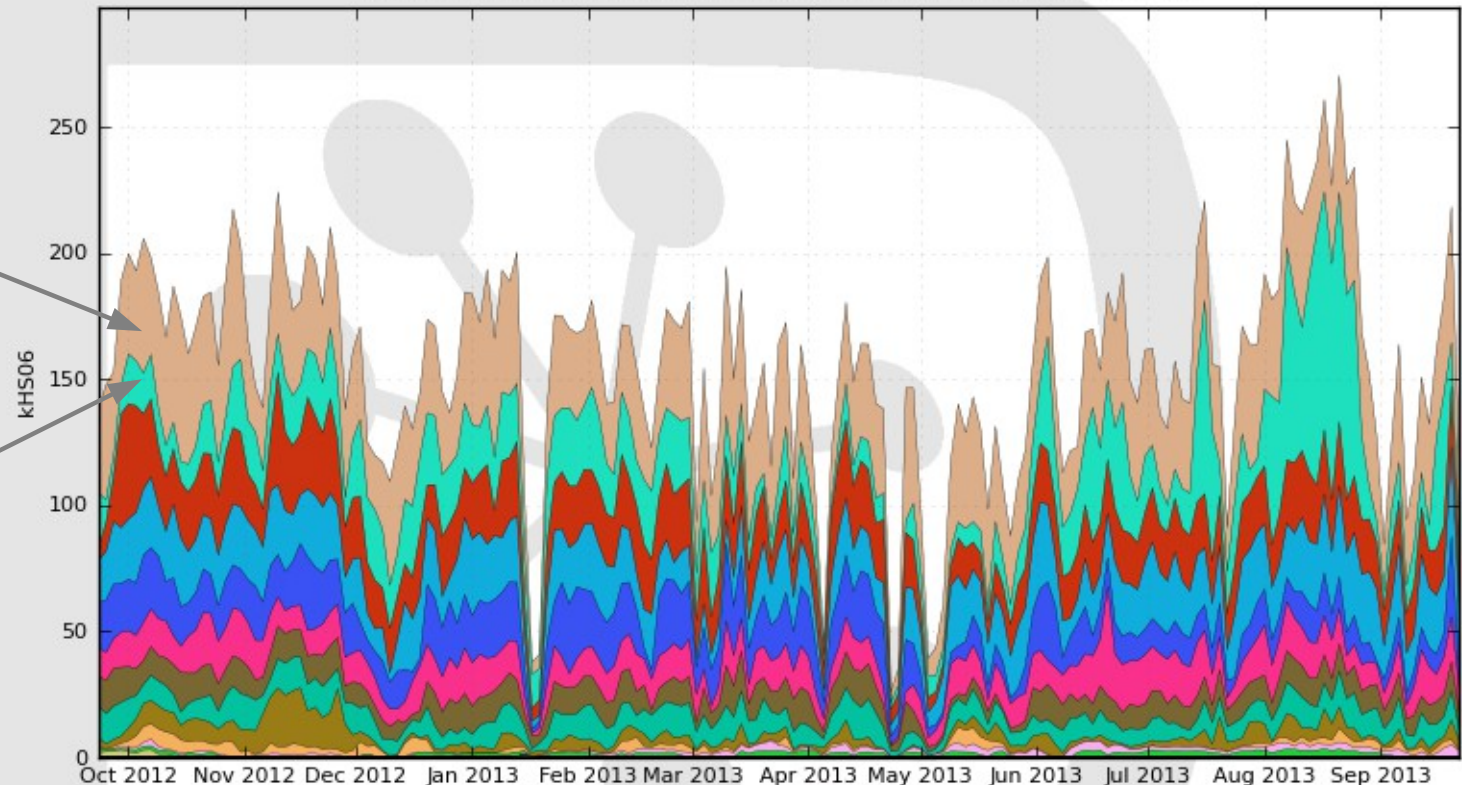
- Numbers/tables largely based on LHCb-PUB-2013-014
 - aimed at 2014/2015 but looks beyond that too
 - summarises changes to event sizes, going to 12kHz etc
- Also used internal LHCb computing meeting material
- And a review of DIRAC and Distributed Computing we did in the spring

- So start with some overall status and predictions about resource requirements, then talk about themes that are relevant to GridPP5.

The UK's place in LHCb computing

Normalized CPU usage by Country

52 Weeks from Week 39 of 2012 to Week 38 of 2013



Max: 270, Min: 26.1, Average: 156, Current: 120

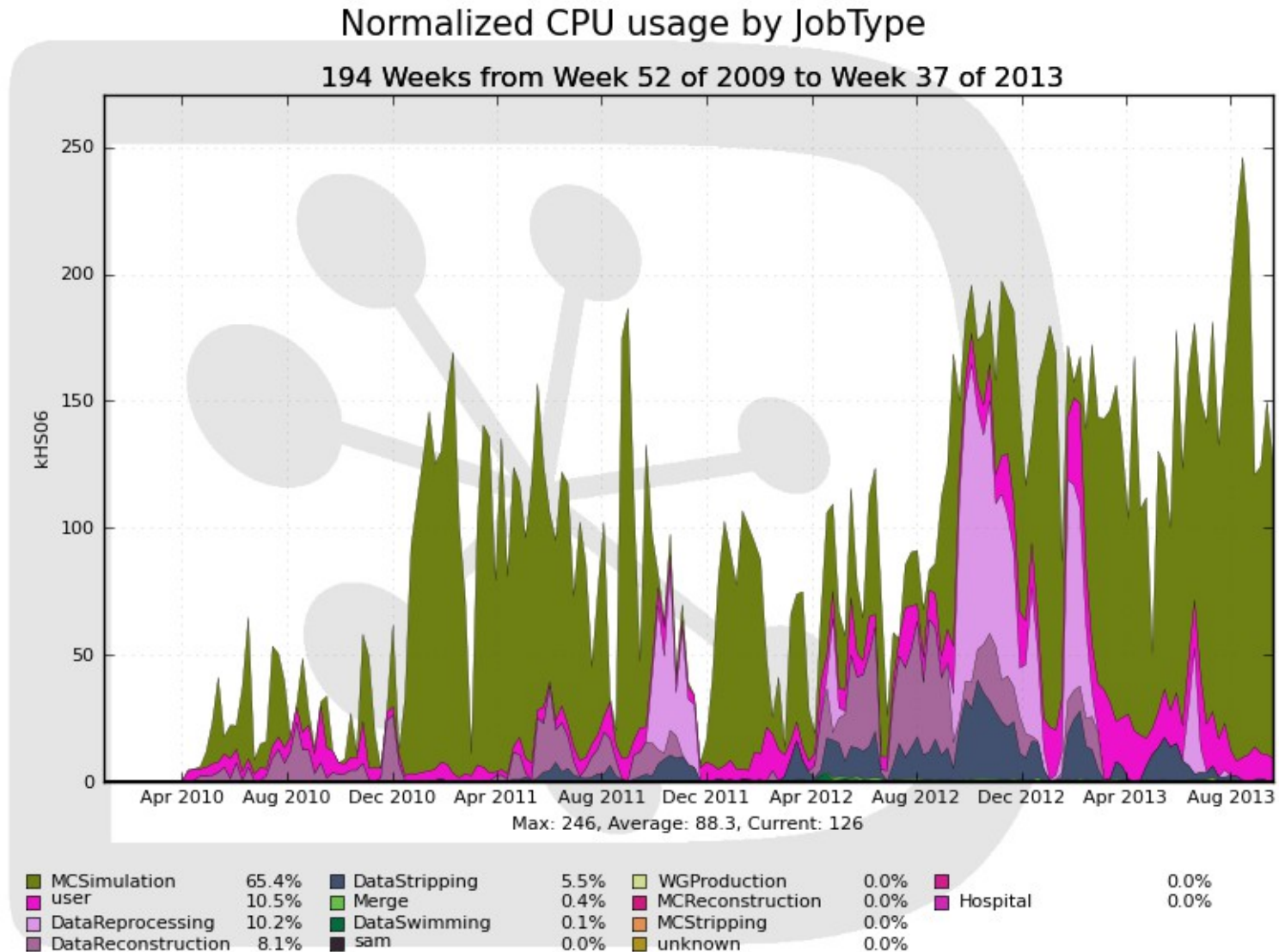
UK	22.8%	NL	6.0%	IE	0.1%	MULTIPLE	0.0%
CH	14.9%	ES	5.6%	BG	0.0%	ORG	0.0%
IT	13.5%	PL	3.0%	IL	0.0%	GR	0.0%
FR	13.1%	RO	1.1%	ANY	0.0%		
DE	9.8%	SU	0.9%	HU	0.0%		
RU	8.2%	BR	0.9%	CY	0.0%		

Generated on 2013-09-23 16:10:40 UTC

.uk is the largest at almost 23%

.ch is mostly CERN Tier-0 and the LHCb HLT farm, which is enabled for offline work whenever possible

JobTypes since April 2010

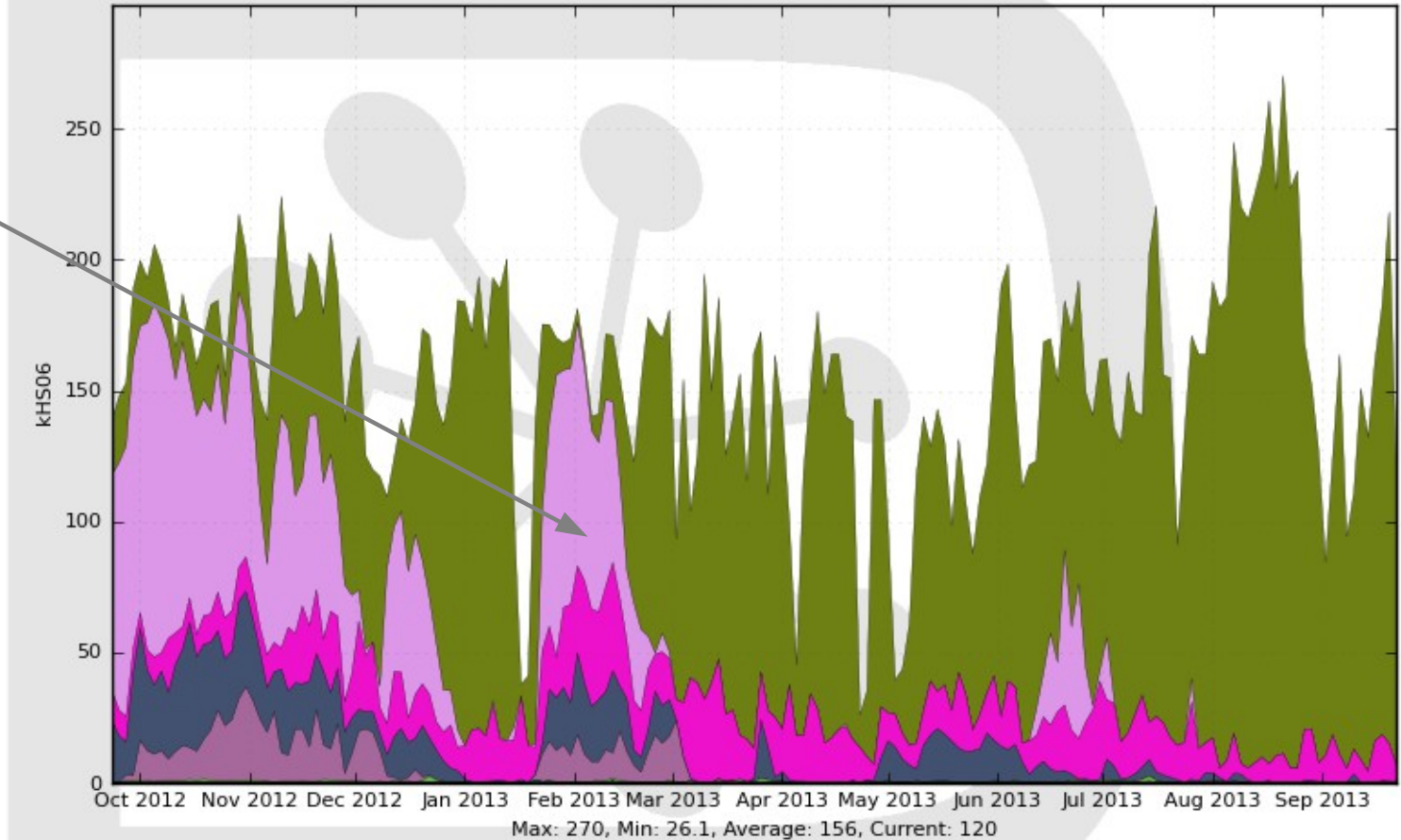


Generated on 2013-09-23 16:21:14 UTC

JobTypes in the last year

Normalized CPU usage by JobType

52 Weeks from Week 39 of 2012 to Week 38 of 2013



Note the reprocessing campaigns in light purple

Now expect that our completed "Reco14" reprocessing of 2011+2012 is the final version of the Run I data

Generated on 2013-09-23 16:21:09 UTC

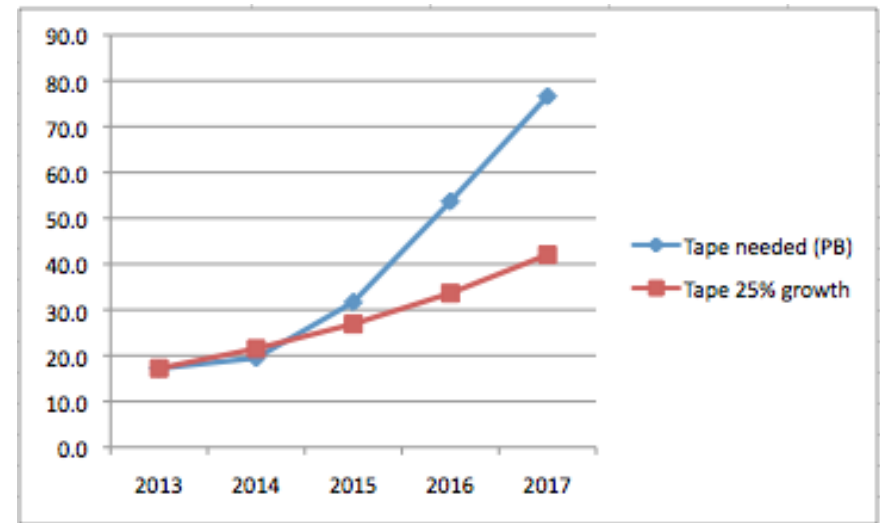
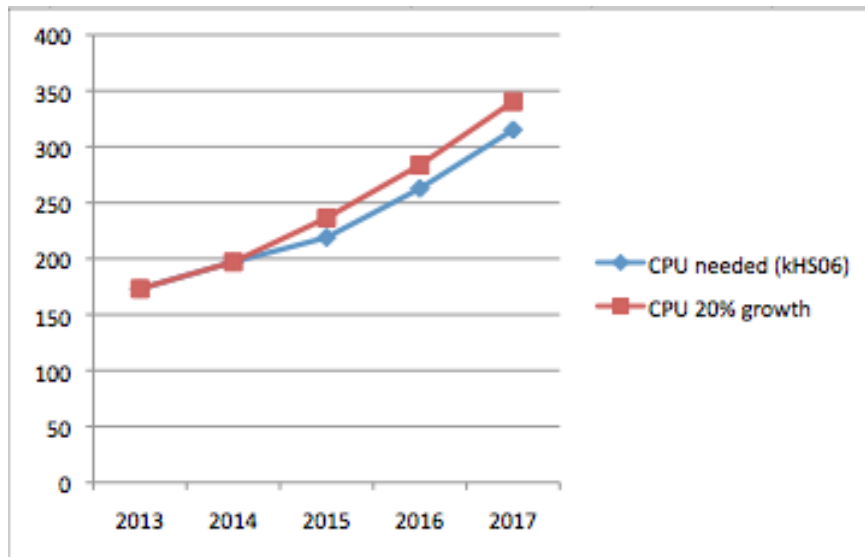
Reco and reprocessing in Run 2

- LHCb plans a radical change to its model
- Data will be split into 3 streams in the High Level Trigger farm
- 20% (“TURBO”) will be processed within the HLT immediately, rather than reconstructed offline, using rapid calibrations
 - A lot of subdetector work done to provide these calibrations
- Some of the rest (“PARKED”) will be not be processed, perhaps not until near or during LS2. Maybe 0-50% of total? Depends on offline resources.
- The remainder (“FULL”) will be stored as RAW for ~2 weeks and only then reconstructed - in case of any problems in prompt calibrations. Not reprocessed for the remainder of Run 2.
- **So much flatter CPU profile, modulated by the accelerator.**

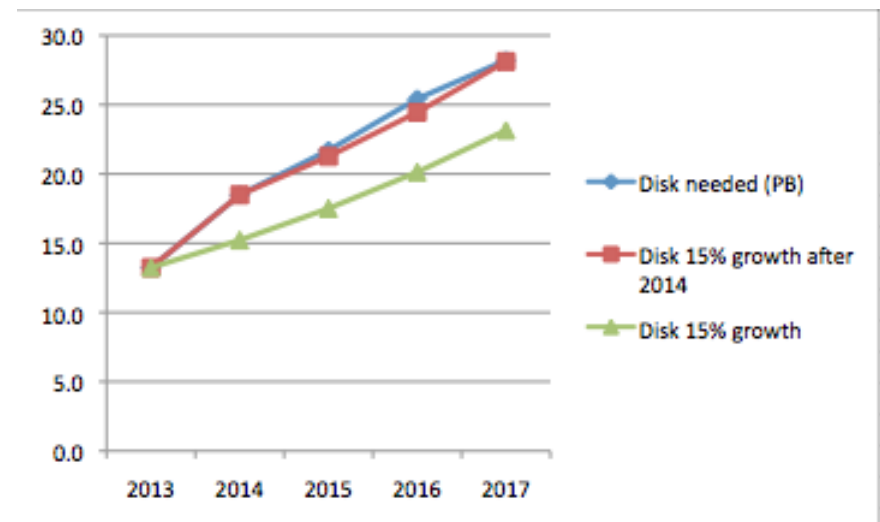
Computing assumptions: 1

- Planning on a constant budget, which means:
 - 20% per year increase in CPU resources
 - 15% per year increase in Disk resources
 - 25% per year increase in Tape resources
 - Assume that pledges to LHCb will increase by this amount in each of 2015-2017
 - Historically not true so far!
 - Starting from 2014 baseline that is not yet secured
- Total CPU for simulation increases more slowly than accumulated luminosity
 - So fewer simulated events per pb^{-1} ? or faster simulation?

Growth assumptions as graphs



- CPU looks ok
- Tape is worrying: mostly RAW.
- Disk assumes we get the pulse in 2014 then 15% growth after that



Growth assumptions as tables

Power (kHS06)	Request 2015	Forecast 2016	Forecast 2017
Tier 0	44	53	63
Tier 1	123	148	177
Tier 2	52	62	74
Total WLCG	219	263	315
HLT farm	10	10	10
Yandex	10	10	10
Total non-WLCG	20	20	20

LHCb Disk (PB)	2015 Request	2016 Forecast	2017 Forecast
Tier0	6.7	8.3	9.5
Tier1	12.5	14.2	15.4
Tier2	2.5	2.9	3.3
Total	21.7	25.5	28.3



Computing assumptions: 2

- 20% slope in CPU increase assumes software can take full advantage of how Moore's law is playing out in practice
 - No longer through increased cycles: do more work per cycle
 - Requires major development in software
 - To make it more vectorisable
 - To take full advantage of multicore architectures
 - Memory limitations can affect performance
- This development will need cycles of testing and optimisation on a range of configurations of real sites
 - Is this something GridPP can help with?
 - Some kind of experiment testing co-ordination to help sites?



T2-D sites

- T2-D proposal: request some T2 sites to add LHCb storage and pass qualification tests to become “T2-D” sites
- Provides more storage and moves closer to GPD Tier-2 models
 - Tier-2 sites spend a lot of time learning how to provide storage but traditionally we've not exploited that
 - Originally we didn't run reconstruction and reprocessing at Tier-2 either. Now have “Attached Tier-2” (or T2-R) sites processing data from a nominated Tier-1.
- T2-D is seen as a way of plugging LHCb's Tier-1 storage gap
 - 33% of LHCb sites are in countries without a Tier1. Pro-rata Tier-1 pledges never equalled total REBUS Tier-1 request.
 - The hope is that those countries will provide some T2-Ds...



T2-D obligations

- 300TB available for LHCb space tokens
 - Start with 100TB, until next year? end of LS1?
 - (How many UK sites will GridPP funding model encourage? 2? If too many, sites lose out by providing LHCb storage.)
- Named site contact (and deputy) within local ops team
 - Does not have to be an LHCb member
 - Must not be an LHCb member who isn't in local ops team
- Respond promptly to LHCb tickets during working hours
 - Should do this already in principle ...
- Site contact should participate in infrequent status meetings
- Achieve quality requirements over some timescale (year?)



Supporting T2-D etc sites

- Storage to be used for data so can run analysis jobs at T2-Ds
 - LHCb is working on improved data placement so can make efficient use of many T2-D sites if they become available
- Setting up monitoring of existing quality criteria (availability, reliability, perfSonar stats, ...) to give feedback about metrics we would like sites to optimise.
- Have to address concerns about support burden of Tier-2 sites
 - Is there a way that existing GridPP support structures can support the WLCG supported experiments' more?
 - As the experiments' rather than ops SAM tests are now WLCG's focus, could the ROD team get involved in the experiments' SAM test monitoring/ticketing for GridPP sites?



Multicore etc

- Ongoing work to support more parallelism within event processing
 - GaudiMP and GaudiHive
 - Optimising event model and applications for vectorisation
 - May be applicable to GPGPUs too
- This is largely hidden from the Grid apart from the need to match jobs/tasks to appropriate machines
 - And even then, it's our overlay grid, DIRAC, that does it
- This is one area that LHCb UK is hoping to work on as part of the LS2 upgrade preparations
 - So there may be opportunities for joint work if GridPP identifies it as an area to work on / support



VMs / Clouds / Vac / BOINC

- Have already produced a generic LHCb VM contextualisation
- This has been used to run production jobs on CERN Ibex OpenStack site, Vac at 3 UK sites, and BOINC on assorted laptops.
- LHCb DIRAC's philosophy is to run job agents anywhere and then provide an overlay grid which hides this from the users
 - eg with the pilot jobs and central task queues that most other experiments have now adopted
 - or with the HLT farm which is consequently one of our largest offline computing sites
- So we should be able to exploit any reasonable flavour of site
 - We don't have a strong bias for or against VMs and Clouds.



Summary

- The UK sites make the largest national contribution to LHCb
- LHCb is changing its model for Run 2 for a smoother profile during the year
- Request continued growth in CPU, Disk and Tape resources
- Working on software changes to exploit:
 - Multicore etc architectures
 - VMs / Clouds / Vac / BOINC etc
 - T2-D sites
- Specific GridPP5 overlaps are: continued provision of essential resources; support of UK T2-D sites; new architecture testing; enhanced site/experiment monitoring and support.