Storage on the Grid – Work Breakdown Structure

Introduction

This document presents the work breakdown structure (WBS) of the "Storage on the Grid" project. This project was started as a response to the recorded incidents related to data intensive jobs running on FermiGrid and accessing data from the BlueArc storage [5]. The goal of this project is the evaluation of storage technologies for the use case of data intensive Grid jobs. The storage technologies considered are the Hadoop Distributed File System (HDFS) [1], Lustre [2], and Blue Arc (BA) [3]. The targeted infrastructures that will benefit from such evaluation are FermiGrid and the General Physics Computing Farm. This work is lead by the OSG group in the context of a loose collaboration with the FermiGrid, OSG Storage, DMS, and FEF groups at Fermilab.

Resources

This plan assumes the availability of the following resources

- Gabriele Garzoglio (marked as GG in the charts): Project Manager. 40% FTE until Mar 1, 2010 and 20% afterwards. As per the FY10 budget, once FermiCloud is deployed (assume Mar 1, 2010), Gabriele will help at 20% with possible development efforts for the FermiCloud infrastructure; therefore, the remaining effort for this storage project is 20%.
- New Hire (NH): Developer from the OSG Group. 20% FTE on the project, assuming a start date of Mar 1, 2010.
- Tanya Levshina (TL): OSG Storage Area Coordinator. We assume the help of her group in the installation of the storage technologies and data movement services (BeStMan, GridFTP, etc.).
- Steve Timm (ST): FermiCloud Project Manager. We assume the help of his group for the deployment of FermiCloud and the integration of the storage solutions with FermiGrid, our test environment.
- Extra Help (marked as Help! in the charts): Possible help assumed at 50% on limited specific tasks

This document describes two plans, assuming the availability of extra help for some tasks or not.

Commented WBS Items and Assumptions

1. Document Assessment Process

   1.1. Select relevant storage requirements/metrics

   See DMS' Lustre evaluation [2]

   1.2. Analyze selected storage metrics for Data Intensive jobs from RunII and IF

   This deliverable provides the baseline of the minimal expected storage performance, given the current *status quo*. Without some external help, this task should be abandoned as it would delay the project of 20 days.

   1.3. Document data access models for the technologies considered

   Examples of possible data access models: pre-staged scratch area; tape backed cache; external access mechanism (SRM, GridFTP, …); internal access mechanism (POSIX, SAM, Special API, …); …

2. Deploy the physical and virtual test infrastructure

   2.1. Design HW and VM layout to support the evaluation of storage technologies

   The deliverable of this task might influence the configuration and design of the FermiCloud infrastructure.

   2.2. Procure / commission Cloud infrastructure

   This tasks is assumed to be worked on by the FermiCloud team with no effort from this project, except for some initial consultation and design decisions (see previous item). We assume that the infrastructure will be available on Mar 1, 2010.

3. Prepare testing infrastructure

   This whole activity with its subtasks could benefit from external help. At the lab, several experts have the knowledge of setting up and using storage benchmarks. It is estimated that with their help at 50% FTE, this activity could take 15 calendar days, instead of 27.

   3.1. Commission storage benchmark for technology assessments

   See DMS' Lustre evaluation. Reuse DMS' storage benchmarks, if possible.

   3.2. Gather and learn to run real user jobs from RunII and IF

   We assume the availability of physicists from DZero, Minos, CDF, etc. Some people has given already tentatively their availability at the meeting on Nov 19, 2009 [4].

   3.3. Develop measurement suite for real user jobs

4. Assess Lustre

   4.1. Deploy storage service and related servers (BeStMan, GridFTP, …)

   As already tentatively agreed, we assume that the OSG Storage Group will be available for the basic deployment of the storage solutions and interfaces. We assume 3 days of full time effort: this estimate has a potentially large error.

   4.2. Integrate FermiGrid with storage service with Lustre

We assume that the FermiGrid team will be available to work at 0.5 FTE on this task.

4.3. Run benchmarks to compare performance with known results for Lustre. Optimize storage as appropriate

This task could benefit from external help. The expectation is that most of the work will consist in the study of the metrics and the tuning of the storage parameters. It is estimated that help from a storage expert at 50% FTE could cut down this time from 12 to 5 calendar days.

4.4. Run measurement suite with real jobs for Lustre

This task could benefit from external help for the same reason as the task above. It is estimated that help from a storage expert at 50% FTE could cut down this time from 12 to 5 calendar days.

4.5. Document Lustre results

5. Assess HDFS

5.1. Deploy storage service and related servers (BeStMan, GridFTP, …)

As already tentatively agreed, we assume that the OSG Storage Group will be available for the basic deployment of the storage solutions and interfaces. We assume 3 days of full time effort: this estimate has a potentially large error.

5.2. Integrate FG with storage service with HDFS

For example, integrate HDFS VMs as WN. We assume that the FermiGrid team will be available to work at 0.5 FTE on this task.

5.3. Run benchmarks to compare performance with known results for HDFS. Optimize storage as appropriate

This task could benefit from external help for the same reason as task 4.3. It is estimated that help from a storage expert at 50% FTE could cut down this time from 12 to 5 calendar days.

5.4. Run measurement suite with real jobs for HDFS

This task could benefit from external help for the same reason as task 4.3. It is estimated that help from a storage expert at 50% FTE could cut down this time from 12 to 5 calendar days.

5.5. Document HDFS results

6. Assess BA for comparison

It should be noted that the WN of the GP Farm on FermiGrid and FermiCloud are in GCC and have about a 2 GBit network connection to the BA, which is in FCC. To stress the BA from FermiGrid or FermiCloud, we need to generate high IOPS test (e.g. moving a lot of small files), rather than high bandwidth tests (the network might saturate before the BA, in this case). These same tests should also used for Lustre and HDFS, to allow us a direct comparison with BA results.

6.1. Devise minimally disruptive testing technique

For the estimate on the amount of effort (1 FTE week), we assumed the availability of BlueArc experts to help with this task. This plan will define high IOPS tests appropriate for BA.

6.2. Run benchmarks to compare performance with known results for BA. Optimize storage as appropriate

This task could benefit from external help for the same reason as task 4.3. It is estimated that help from a storage expert at 50% FTE could cut down this time from 12 to 5 calendar days.

6.3. Run measurement suite with real jobs for BA

This task could benefit from external help for the same reason as task 4.3. It is estimated that help from a storage expert at 50% FTE could cut down this time from 12 to 5 calendar days.

6.4. Document BA results

7. Reports and documentation

7.1. Relate Data Intensiveness requirements and technology assessment

In both plans we assume that we had external help to deliver the analysis of the data intensiveness requirements for IF and RunII (see task 1.2)

7.2. Document relevance of the study for GPCF

7.3. Study references in literature to assess operational properties, long-term resilience, etc.

7.4. Gather all documentation

Plan

In this section we show two execution plans for the project. One assumes no external help, except for the “analysis of selected storage metrics for Data Intensive jobs from RunII and IF” (see comments on item 1.2 above). The other assumes help from one additional resource at 50% FTE on a few selected items.

The estimated termination of the plan without help is late Nov 2010. The estimated termination for the plan with help is early Sep 2010. Both timelines may be possibly shortened by reducing the scope of the storage performance study.

Comparison of the Plans with and without Help

WBS	Name	Start	Finish	Work	Duration	Slack	Cost	Assigned to
1	**Document Assessment Process**	**Jan 11**	**Feb 17**	**16d**	**27d 4h**	**202d 5h**	**0**	
1.1	Select relevant storage requirements/metrics	Jan 11	Jan 20	3d	7d 4h	7d 4h	0	GG
1.2	Analyze selected storage metrics for Data Intensive jobs from RunII and IF	Jan 20	Feb 17	10d	20d	202d 5h	0	Help!
1.3	Document data access models for the technologies considered	Jan 20	Jan 29	3d	7d 4h	5d	0	GG
2	**Deploy the physical and virtual test infrastructure**	**Feb 1**	**Mar 1**	**4d**	**21d**	**14d**	**0**	
2.1	Design HW and VM layout to support the evaluation of storage technologies	Feb 1	Feb 10	3d	7d 4h	7d 4h	0	GG
2.2	Procure / commission Cloud infrastructure	Mar 1	Mar 1	1d	1d	14d	0	ST
3	**Prepare testing infrastructure**	**Feb 10**	**Mar 19**	**11d**	**27d 4h**		**0**	
3.1	Commission storage benchmark for technology assessments	Mar 1	Mar 19	3d	15d		0	NH
3.2	Gather and learn to run real user jobs from RunII and IF	Feb 10	Feb 19	3d	7d 4h	5d	0	GG
3.3	Develop measurement suite for real user jobs	Feb 22	Mar 10	5d	12d 4h	7d 4h	0	GG
4	**Assess Lustre**	**Mar 22**	**May 26**	**20d**	**48d**		**0**	
4.1	Deploy storage service and related servers (BeStMan, GridFTP, ...)	Mar 22	Mar 24	3d	3d		0	TL
4.2	Integrate FG with storage service with Lustre	Mar 25	Apr 7	5d	10d		0	ST
4.3	Run benchmarks to compare performance with known results for Lustre. Optimize storage as appropriate	Apr 8	Apr 26	5d	12d 4h		0	GG, NH
4.4	Run measurement suite with real jobs for Lustre	Apr 26	May 12	5d	12d 4h		0	GG, NH
4.5	Document Lustre results	May 13	May 26	2d	10d		0	NH
5	**Assess HDFS**	**May 27**	**Jul 26**	**20d**	**42d 1h**		**0**	
5.1	Deploy storage service and related servers (BeStMan, GridFTP, ...)	May 27	May 31	3d	3d	39d 1h	0	TL
5.2	Integrate FG with storage service with HDFS	May 27	Jun 7	5d	7d 1h		0	NH, ST
5.3	Run benchmarks to compare performance with known results for HDFS. Optimize storage as appropriate	Jun 7	Jun 23	5d	12d 4h		0	GG, NH
5.4	Run measurement suite with real jobs for HDFS	Jun 23	Jul 12	5d	12d 4h		0	GG, NH
5.5	Document HDFS results	Jul 12	Jul 26	2d	10d		0	NH
6	**Assess BA for comparison**	**Jul 26**	**Sep 29**	**17d**	**47d 4h**		**0**	
6.1	Devise minimally disruptive testing technique	Jul 26	Aug 11	5d	12d 4h		0	GG, NH
6.2	Run benchmarks to compare performance with known results for BA. Optimize storage as appropriate	Aug 11	Aug 30	5d	12d 4h		0	GG, NH
6.3	Run measurement suite with real jobs for BA	Aug 30	Sep 15	5d	12d 4h		0	GG, NH
6.4	Document BA results	Sep 15	Sep 29	2d	10d		0	NH
7	**Reports and documentation**	**Sep 29**	**Nov 29**	**14d**	**42d 4h**		**0**	
7.1	Relate Data Intensiveness requirements and Technology assessment	Sep 29	Oct 20	3d	15d		0	GG
7.2	Document relevance of the study for GPCF	Oct 20	Nov 10	3d	15d		0	GG
7.3	Study references in literature to assess operational properties, long-term resilience, etc.	Sep 29	Oct 20	3d	15d	15d	0	NH
7.4	Gather all documentation	Nov 10	Nov 29	5d	12d 4h		0	GG, NH

*Fig 1: The timeline and resource assignments of the plan **without help***

WBS	Name	Start	Finish	Work	Duration	Slack	Cost	Assigned to
1	**Document Assessment Process**	**Jan 11**	**Feb 17**	**16d**	**27d 4h**	**146d 3h**	**0**	
1.1	Select relevant storage requirements/metrics	Jan 11	Jan 20	3d	7d 4h	9d 1h	0	GG
1.2	Analyze selected storage metrics for Data Intensive jobs from RunII and IF	Jan 20	Feb 17	10d	20d	146d 3h	0	Help!
1.3	Document data access models for the technologies considered	Jan 20	Jan 29	3d	7d 4h	9d	0	GG
2	**Deploy the physical and virtual test infrastructure**	**Feb 1**	**Mar 1**	**4d**	**21d**	**1d 4h**	**0**	
2.1	Design HW and VM layout to support the evaluation of storage technologies	Feb 1	Feb 10	3d	7d 4h	9d 1h	0	GG
2.2	Procure / commission Cloud infrastructure	Mar 1	Mar 1	1d	1d	1d 4h	0	ST
3	**Prepare testing infrastructure**	**Feb 10**	**Mar 3**	**11d**	**15d**		**0**	
3.1	Commission storage benchmark for technology assessments	Mar 1	Mar 3	3d	2d 4h		0	Help!, NH
3.2	Gather and learn to run real user jobs from RunII and IF	Feb 10	Feb 12	3d	2d 1h	9d 2h	0	Help!, GG
3.3	Develop measurement suite for real user jobs	Feb 12	Feb 18	5d	3d 4h	9d 2h	0	Help!, GG
4	**Assess Lustre**	**Mar 3**	**Apr 20**	**20d**	**34d**		**0**	
4.1	Deploy storage service and related servers (BeStMan, GridFTP, ...)	Mar 3	Mar 8	3d	3d		0	TL
4.2	Integrate FG with storage service with Lustre	Mar 8	Mar 22	5d	10d		0	ST
4.3	Run benchmarks to compare performance with known results for Lustre. Optimize storage as appropriate	Mar 22	Mar 30	5d	5d 4h		0	Help!, GG, NH
4.4	Run measurement suite with real jobs for Lustre	Mar 30	Apr 6	5d	5d 4h		0	Help!, GG, NH
4.5	Document Lustre results	Apr 6	Apr 20	2d	10d		0	NH
5	**Assess HDFS**	**Apr 20**	**May 28**	**20d**	**28d 2h**		**0**	
5.1	Deploy storage service and related servers (BeStMan, GridFTP, ...)	Apr 20	Apr 23	3d	3d	25d 2h	0	TL
5.2	Integrate FG with storage service with HDFS	Apr 20	Apr 29	5d	7d 1h		0	NH, ST
5.3	Run benchmarks to compare performance with known results for HDFS. Optimize storage as appropriate	Apr 29	May 7	5d	5d 4h		0	Help!, GG, NH
5.4	Run measurement suite with real jobs for HDFS	May 7	May 14	5d	5d 4h		0	Help!, GG, NH
5.5	Document HDFS results	May 14	May 28	2d	10d		0	NH
6	**Assess BA for comparison**	**May 28**	**Jul 13**	**17d**	**31d 5h**		**0**	
6.1	Devise minimally disruptive testing technique	May 28	Jun 16	5d	12d 4h		0	GG, NH
6.2	Run benchmarks to compare performance with known results for BA. Optimize storage as appropriate	Jun 16	Jun 21	5d	3d 4h		0	Help!, GG, NH
6.3	Run measurement suite with real jobs for BA	Jun 21	Jun 29	5d	5d 4h		0	Help!, GG, NH
6.4	Document BA results	Jun 29	Jul 13	2d	10d		0	NH
7	**Reports and documentation**	**Jul 13**	**Sep 9**	**14d**	**42d 4h**		**0**	
7.1	Relate Data Intensiveness requirements and Technology assessment	Jul 13	Aug 3	3d	15d		0	GG
7.2	Document relevance of the study for GPCF	Aug 3	Aug 24	3d	15d		0	GG
7.3	Study references in literature to assess operational properties, long-term resilience, etc.	Jul 13	Aug 3	3d	15d	15d	0	NH
7.4	Gather all documentation	Aug 24	Sep 9	5d	12d 4h		0	GG, NH

*Fig 2: The timeline and resource assignments of the plan **with help***

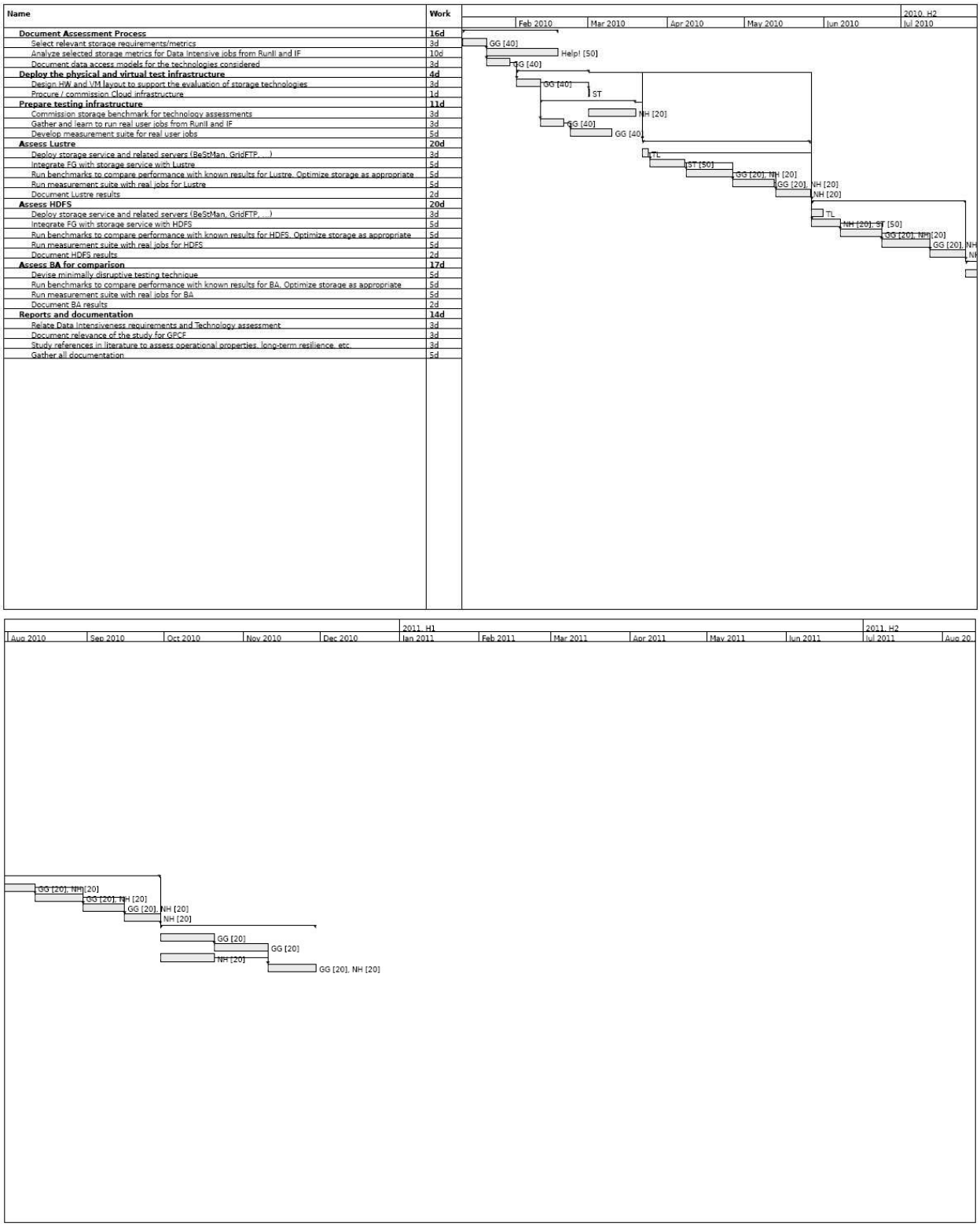

*Fig 3: The Gantt chart for the plan **without help***

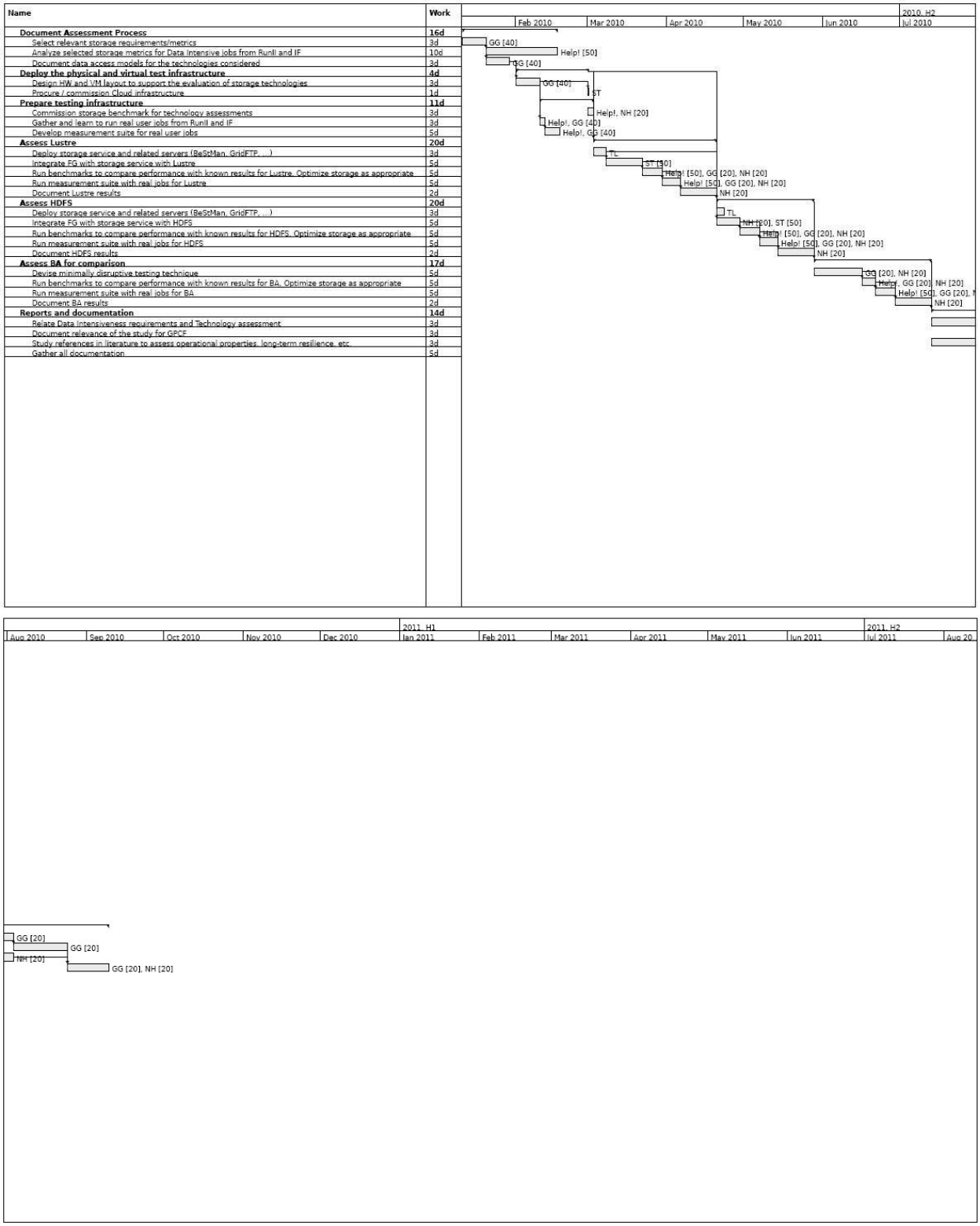

*Fig 4: The Gantt chart for the plan **with help***

References

[1] G. Attebury, A. Baranovskiy, K. Bloom, B. Bockelman, D. Kcira, J. Letts, T. Levshina, C. Lundestedt, T. Martin, W. Maier, H. Pi, A. Rana, I. Sfiligoi, A. Sim, M. Thomas, F. Wuerthwein, “Hadoop Distributed File System for the Grid”, to be published in the proceedings of the IEEE Nuclear Science Symposium, Oct 2009.

[2] G. Oleynik, J. Bakken, R. Rechenmacher, J. Simone, D. Holmgren, N. Seenu, D. Petravick, M. Crawford, S. Fuess, A. Kulyavtsev, D. Litvintsev, A. Moibenko, T. Perelmutov, V. Podstavkov, S. Naymola, S. Wolbers, “Lustre Evaluation as part of 2008 Storage Evaluation”, CD-docdb 2817

[3] R. Pasetes, Central NAS Major Upgrade Plan 2009 – doc db 3229

[4] Slides of the Nov 19, 2009 meeting on Storage on FermiGrid and GPCF: docdb 3279

[5] G. Garzoglio, “Storage Services for the Fermilab Grid Facility” – doc db 3279