

# Results of the 2007 User Survey for Lattice QCD (LQCD) Computing Facility

at  
Brookhaven National Laboratory (BNL), Brookhaven, New York  
Fermi National Accelerator Laboratory (FNAL), Batavia, Illinois  
Thomas Jefferson National Accelerator Facility (Jlab), Newport News, Virginia



For the U.S. Department of Energy  
Office of Science  
Offices of High Energy and Nuclear Physics

Prepared by: Bakul Banerjee

Date: November 07

Revision 1.0

Prepared by: Bakul Banerjee

Change Log

Revision No.	Pages Affected	Effective Date
Revision 1.0	Entire Document	11/15/07

## Table of Contents

Results of the 2007 User Survey for Lattice QCD (LQCD) Computing Facility.....	1
1 Objective.....	4
2 Summary of Results.....	4
3 Overview of the survey conducted .....	6
3.1 General requirements.....	6
3.2 Content of the survey.....	7
4 Test and deployment of the survey.....	7
5 Methodology for the analysis of the survey.....	8
6 Detailed Results .....	8
6.1 General.....	8
6.1.1 Employer.....	9
6.1.2 Employment level.....	9
6.1.3 Usage of LQCD computers.....	9
6.1.4 Job submission.....	9
6.1.5 Facility usage .....	9
6.2 User satisfaction.....	10
6.2.1 Overall user satisfaction with facilities (1-5 ratings).....	10
6.2.2 Documentation.....	11
6.2.3 User support.....	12
6.2.4 Reliability.....	13
6.2.5 Responsiveness .....	14
6.2.6 Ease of access .....	14
6.3 Support and communication .....	15
6.3.1 E-mail.....	15
6.3.2 Web support.....	16
6.3.3 Other tool support .....	16
6.4 User meeting needed (Y/N count) .....	16
6.5 Helpdesk (analysis of the last problem solved) .....	16
6.5.1 Knowledge of how to ask help (Y/N count).....	16
6.5.2 Help requested by facility (Y/N count).....	17
6.5.3 Time to initial response (working hours).....	17
6.5.4 Closeout from initial response (Y/N).....	17
6.5.5 Working days needed to solve the problem.....	17
6.5.6 Help provided was satisfactory (Y/N/NA) .....	17
6.6 Proposal and allocations .....	20
6.6.1 Satisfaction with the proposal process (Rating of 1-5).....	20
6.6.2 CFP clarification and time needed.....	20
6.6.3 Allocation process.....	20
6.7 Running jobs – Overall Job failure rate for the past year .....	21
6.8 Mass storage.....	22
6.8.1 Adequacy of disk storage.....	22
6.8.2 Disk storage needs in (GB).....	22
6.8.3 Disc storage backup needs (GB).....	23

	6.8.4	Tape library quality .....	23
	6.8.5	Tape library requirements in GB for FY08.....	23
	6.9	General comments .....	24
7		Lessons learned about conducting the survey.....	24
8		Acknowledgement .....	24

## 1 Objective

To fulfill the goal of the LQCD facility to continuously serve our user community in the best possible manner, a user survey was conducted during August and September of 2007. The objective of the survey was to assess the level of satisfaction experienced by the users of the Lattice QCD Computational Facility. Using the results of the survey, the Integrated Project Team (IPT) expects to understand how it can optimize the services using the limited resources available to the project. Results of the survey are presented in this document. An action plan for improvement will be developed and executed at a later date.

## 2 Summary of Results

Although the LQCD project team always strives to provide complete satisfaction, the user survey completed recently helps the team to fine-tune the services and improve satisfaction. A comprehensive set of questions for the survey was defined by the project team in collaboration with the LQCD Executive Committee and the Scientific Program Committee. The questions are designed to identify performances of individual facilities as well as the overall performance. The LQCD Facility consists of two clusters at Fermilab (FNAL) and Thomas Jefferson Lab (Jlab) and the QCDOC machine at Brookhaven National Laboratory (BNL). It is noted that the project management team that includes site managers at all three facilities works together closely toward an increasingly uniform process, in spite of the inherent differences in the machines and their setups.

The survey, targeted toward scientific users of the LQCD Computing Facility, is executed using a survey utility already in place at Fermilab. Survey questions are grouped in eight different categories. Total number of questions with alpha-numeric results is 29 and total number of questions with 1 to 5 satisfaction rating is 32. Details of the data collected and tabulated are given in the section titled "Detailed Results".

It is important to put forth a word of caution regarding the survey. Since the total population of users is relatively small, the outliers may affect the results of the survey significantly. For example, users of Jlab make up almost half of the total number of survey takers. A single unsatisfied customer may affect the satisfaction ranking of the facility. Descriptions of the questions and a summary of the survey results are given below:

1. General: Questions under this category are designed to collect demographic data of the user community.
  - a. 72% of the users are employed by a university or a college, the rest are mostly employed by the participating laboratories.
  - b. 46% of the users are faculty members. Post-docs make up a significant portion of the rest.
  - c. 48% of the users submit jobs daily. Only 5 users submit jobs occasionally.
  - d. About 57% of active users submit an average of less than 10 jobs per week. The maximum rate of submission per week can be as high as 100 jobs.
  - e. The number of users for the JLab is 25, whereas BNL and FNAL sites are used by 16 and 13 individuals.
2. User satisfaction: These questions assess the overall user satisfaction and satisfaction levels related to documentation, user support, reliability of the machines, responsiveness and accessibility. In this section, users who rated satisfaction level

to be either satisfied (rank=4) or very satisfied (rank=5) are considered to be satisfied in each subcategory.

- a. General: 83% of users rate the satisfaction level to be satisfied or very satisfied. BNL, FNAL, and JLAB received similar ratings 83%, 93%, and 70% of users.
  - b. Document: Users of BNL, FNAL, Jlab are satisfied with documentation 88%, 69%, and 78% of times respectively.
  - c. User support: The ratings for user support for BNL, FNAL, and Jlab are 94%, 85%, and 80%.
  - d. Reliability: 81%, 85%, and 56% of users find BNL, FNAL, and Jlab reliable
  - e. Responsiveness: 100%, 96% and 72% of users find BNL, FNAL, and Jlab to be responsive.
  - f. Ease of access: 82%, 69%, and 68% of users find BNL, FNAL, and Jlab facilities easily accessible.
3. Communication: The topics covered were various modes of communications including e-mails, web communications, and other tools.
    - a. Email: 88%, 81%, and 72% of BNL, FNAL, and Jlab users find email related communications to be satisfactory
    - b. Web support: BNL, FNAL, and Jlab have satisfactory ratings from 73%, 74%, and 81% of users respectively.
    - c. Other tool support: Users were asked about their satisfaction level regarding various general purpose user tools, for example, various possible command line tools. 81%, 77%, and 72% of users are satisfied.
  4. User meetings: The need for user meetings was assessed in this question. 45 users indicated that additional user meetings are not needed, whereas 7 users think that user meetings might be helpful.
  5. Helpdesk: All three LQCD facilities operate site specific helpdesks, sometimes called problem reporting system. An extensive set of questions were posed to determine the usage and efficacy of the helpdesk at all three sites. After determining the awareness of the existence of the helpdesk, users were asked to rate their satisfaction regarding the last helpdesk request they submitted in terms of time to initial response and close out of the helpdesk ticket, and the level of satisfaction with the helpdesk request. 93% of users knew how to ask help. The response to the evaluation of the last problem report is given below:
    - a. The spread of the helpdesk request submission by users among BNL, FNAL, and Jlab is 23%, 52%, and 25%
    - b. Time to initial response: 31 out of 37 helpdesk requestors received initial response within 6 hours.
    - c. Problem solved by initial response: 68% of problems were solved using the initial response. About 23% of the problems required more than one day to solve. It is likely that a small fraction of problems may require modification of the system and may not be solved for months. It was difficult to characterize these problems from the survey. However, site managers indicate that some of these issues are related to software features that may not be available at the time of helpdesk request submission.

- d. Satisfaction with the helpdesk: 93% of users find the help received satisfactory.
6. Proposal and allocation: These questions are designed to understand the satisfaction level related to the proposal and allocation process.
  - a. Maximizing scientific output: 70% of users believe that the allocation process helps maximizing the scientific output
  - b. Satisfaction with the process: 69% of users find the allocation process satisfactory
  - c. Clarity of the Call for Proposal(CFP): 79% of users think that the CFP was clear
  - d. Transparency: 61% of users find the proposal process adequately transparent
  - e. Fairness: 63% of users find the allocation process to be fair.
7. Running jobs: The objective was to assess the success of job submissions. Approximately 83% of users had less than 10 job failures over the past year. A successful running of jobs on the LQCD machines depends on multiple factors including hardware, software, user configuration, experience and preferential styles. Although a set of questions were posed to determine the types of failures, the data obtained seemed to be uneven. This may be attributed to the limitations of the survey tool. Also, it is also likely that with the current set of software, users could not determine why a job failed. This issue will be referred to the Lattice Gauge software team.
8. Mass storage: Since mass storage is a continuously evolving issue for the LQCD project, these questions are design to understand the disk and tape storage requirements and satisfaction of the users.
  - a. Satisfaction with the disk space: 87% of users are satisfied with the storage solutions provided
  - b. 2007/2008 disk space requirements: The projected disk storage space requirement is 587 TB with approximately 500 TB backed up
  - c. Tape storage satisfaction at Jlab and FNAL: Approximately 32% and 19% users need tape storage at Jlab and FNAL respectively. Users find the instructions for tape storage adequate.
  - d. Estimated tape storage requirements: Approximately 320 TB of permanent tape storage is required with an additional 100 TB of temporary tape storage.
9. General comments: Users provided an extensive set of comments, both general and specific. These comments are very helpful.

### **3 Overview of the survey conducted**

#### **3.1 General requirements**

The survey was designed based on the following general requirements:

- The on-line survey should be easily accessible from various outside organizations for a limited time. The survey is anonymous.
- The tentative opening date for the survey is July 21 and it will remain open for one month.

- Results should be easily retrievable from the database repository in Excel format.
- The FNAL database group is not responsible for the analysis

### *3.2 Content of the survey*

#### **INVITATION LETTER**

The goal of the Lattice QCD Computational Facility, referred as LQCD, is to continuously serve our user community in the best possible manner. To fulfill this goal, we would like to know your level of satisfaction with the facility and to get feedback from you about how we can improve your satisfaction.

Please take few minutes to fill out the following on-line survey as best as you can.

[http://www-esh.fnal.gov/pls/default/qaire.show\\_questions?this\\_qaire\\_id=81](http://www-esh.fnal.gov/pls/default/qaire.show_questions?this_qaire_id=81)

Please answer every question and complete it by August 31, 2007. Please use tabs to navigate because using "Enter" will initiate a submission. Every effort has been made to keep the survey anonymous. While filling out the survey, if you come across a question related to a service that you or your group provide, please select the choice "Not applicable" or "Don't know". This will prevent self-evaluation.

It might be a good idea to review the survey before you start. Please contact Bakul Banerjee (phone 630-840-5251 or email [bakulb@fnal.gov](mailto:bakulb@fnal.gov)) for questions. Thank you in advance for your participation in this online survey.

#### **INSTRUCTIONS**

The LQCD facility is a distributed facility with dedicated clusters at FNAL and Jlab and the QCDOC machine at BNL. Some of the questions are site specific and others are generic. This survey covers all three sites, namely, BNL QCDOC, FNAL Clusters, and Jlab Clusters. The survey aims at capturing feedback for all three components of the LQCD facility. Although it may not directly impact you, it may ask some questions related to Cyber Security, Storage, and Network issues. The time period you should consider is January 2006 to June 2007 or past 18 months. The term Lattice QCD project means the approved project that your experiment is a part of.

**Please answer every question and complete it by August 31, 2007. Please use tabs to navigate because using "Enter" will initiate a submission.** Every effort has been made to keep the survey anonymous. While filling out the survey, if you come across a question related to a service that you or your group provide, please select the choice "Not applicable" or "Don't know". This will prevent self-evaluation. Please contact Bakul Banerjee at 630 840-5251 or [bakulb@fnal.gov](mailto:bakulb@fnal.gov) with your questions regarding the survey.

### **4 Test and deployment of the survey**

The request to participate in the survey was sent to the USQCD mailing list, with a follow-up to everybody who has an account to the facility. General users had 15 working days or three calendar weeks to complete the survey. Before releasing the survey to the general users, the survey was sent

to a selected group of expert users. Every effort had been made to keep the survey anonymous for those who chose to be so. Initial duration was Monday, July 30 to August 31, 2007. Due to the lack of adequate responses, additional requests were sent to key persons and the period was extended to September 15. Initial testers were site managers, project manager, Balian Joo, Bob Sugar, Andreas Kronfeld, and Paul McKenzie.

LQCD Associate Project Manager is responsible for the preparation of the analysis of the survey from the data collected. The survey report will be published on the USQCD web page.

## **5 Methodology for the analysis of the survey**

After the closing, the survey data was retrieved in the html format from the Fermilab's website: <http://www-esh.fnal.gov/pls/default/qaire.stats> and was translated into excel workbook by hand.

Although no accurate data is available for the total number of individual users of the facility, it is estimated that there are approximately 60 users with varied level of activities. The usable number of records was 54. The raw data needed only minimal cleaning, except for the data collected for categorization of the job failure types. This accounts for 90% survey response rate, which may be considered by general survey standards excellent.

The user survey data collected are available in three different categories. The analysis is done as follows:

1. Comments: Comments are inserted "as is" with minor formatting changes, but no analysis is done. Comments are deemed to be useful for general planning purposes
2. Non numeric data: Analysis is done by categories. Sometimes, the frequency distributions are done using appropriate binning.
3. Numeric data: These data is derived from the satisfaction level survey questions. To obtain normalized values for satisfaction, values for 0 and -1 scores were discarded. The rest of the counts were normalized by the total count at each column. The explanations for numerical values are given below.

- 5 – very satisfied
- 4 – satisfied
- 3 – neutral
- 2 – dissatisfied
- 1 – very dissatisfied
- 0 – no opinion
- 1 – not applicable

## **6 Detailed Results**

### **6.1 General**

**6.1.1 Employer**

Employed by	Count
BNL	6
FNAL	3
JLab	3
University/college	39
Other	2

**6.1.2 Employment level**

Type	Count
Student	8
Postdoc	17
Faculty	25
Other university staff	0
Lab scientist	4
Lab computing profs.	0
Other university staff	0

**6.1.3 Usage of LQCD computers**

Usage	Freq.
Daily	26
Weekly	15
Monthly	8
Occasionally	5

**6.1.4 Job submission**

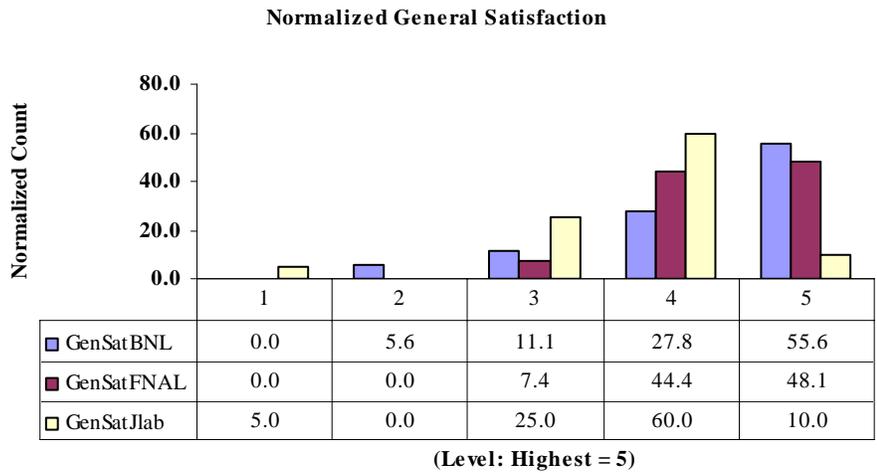
Avg. jobs submitted	Freq.
10	30
20	2
50	8
100	8
200	1
500	2
1000	2
5000	0

**6.1.5 Facility usage**

Facility	Users
BNL	16
FNAL	13
Jlab	25

## 6.2 User satisfaction

### 6.2.1 Overall user satisfaction with facilities (1-5 ratings)

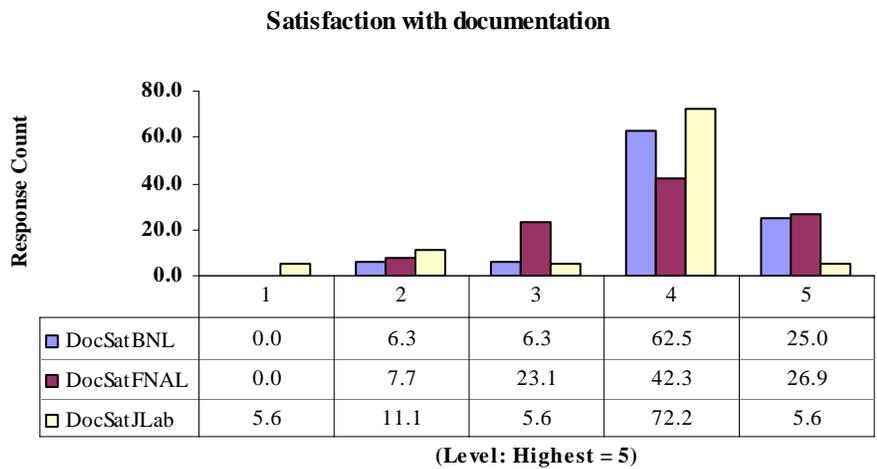


### Comments

1. I just hope the IO speed can be faster. I know it's not easy for a large supercomputer.
2. Very effective.
3. Long queues. High failure rates.
4. The low IO rate on the qcdoc makes it unsuitable for some valence calculations.
5. No stable environment, unclear responsibilities of staff. Eventually, I gave up.
6. Getting the Kerberos and (obsolete version of) SSH working for FNAL was a bit of trouble; but now that it's over I am very happy. On QCDOC at BNL, all of the setup scripts (at least one for compiling, another for running, some incompatibility between the two if I remember correctly) were also a bit of trouble.
7. Outstanding service at FNAL in every respect of cluster management
8. I will discuss this survey with my students and postdocs who use these facilities extensively, and have them complete the remaining parts in detail

9. Some problems ( instability, file server issues, ... ) but administrators always ready to help quickly
10. Having three different security systems is quite annoying
11. Increased support staff at JLab has resulted in increased level of satisfaction
12. My only reservation is that I frequently need access to a few nodes for several short, independent runs. A few nodes have been set aside for this purpose, but more is needed.

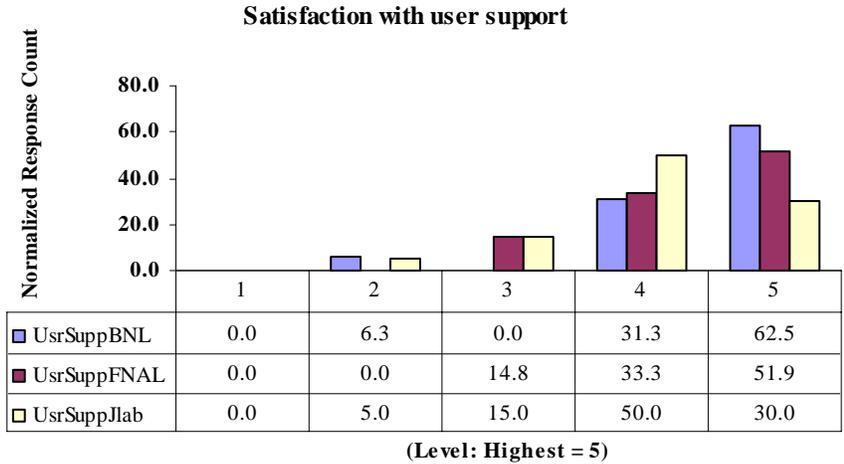
**6.2.2 Documentation**



**Comments**

1. All sites do a reasonable job at site level documentation.
2. Difficult to find information needed. Best if user experience doesn't change over long periods of time. That would avoid having to reread everything again and again for changes.
3. I needed documentation on the PBS queuing system. The documentation was great, but I had to be told where to find it, (it's location on the web site was not obvious).
4. for some details, it an email to the administrators helps
5. I find that it's much easier to get information by word-of-mouth than looking online.

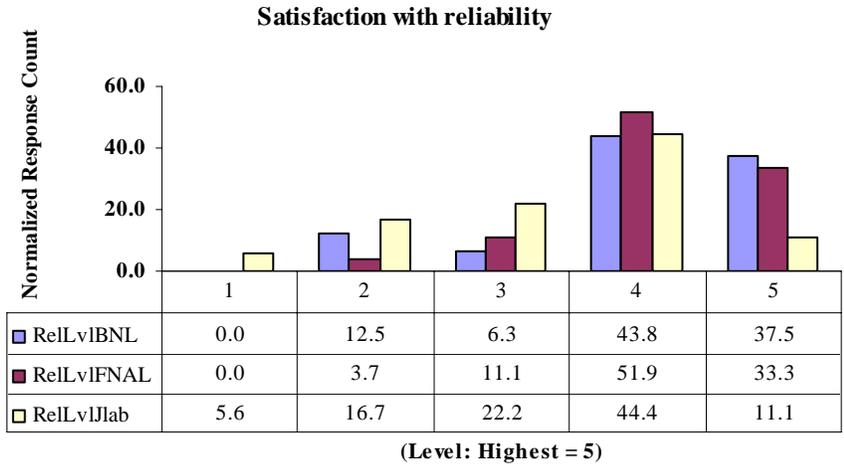
### 6.2.3 User support



#### Comments

1. A good and responsive service. Queries are resolved by email and generally solves the problem.
2. The user support staff at all sites are helpful and quite motivated.
3. Great job!!
4. It took weeks to get my QCDOC account at BNL. On the other hand, Don Holmgren and Amitoj Singh (that I know of) at FNAL deserve medals.
5. very fast response from FNAL administration, excellent!
6. I really appreciate the prompt responses from lqcd-admin, and, in particular, that they will respond on weekends if they can.
7. Outstanding user support at FNAL
8. Always ready to provide uncomplicated, immediate solutions
9. Has been quick, mostly knowledgeable. biggest problem has been communicating what the problem is and if it is my misunderstanding or a limitation of the equipment.

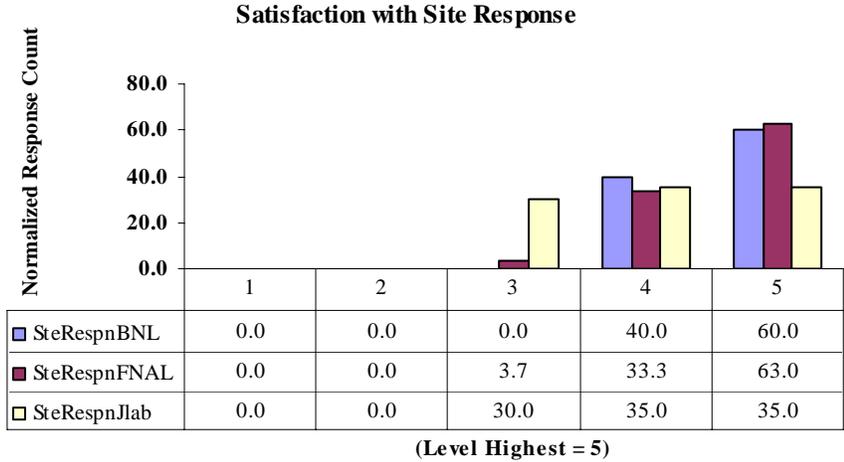
### 6.2.4 Reliability



#### Comments

1. Getting jobs to successfully start can be a chore on the qcdoc. However, once running it is quite stable. The FNAL clusters have overcome some initial problems scaling their systems to large size. Initial failure rates on kaon were high. JLab is experiencing scaling problems ramping their systems to large sizes - failure rates can be high on 7n. The JLab 6n and 4g clusters, and FNAL pion and kaon clusters are fairly stable.
2. I have never had a problem with job failure or unreasonable downtime.
3. significant number of sporadic, inexplicable job failures (MPI ?) ==> wasted CPU time
4. failure rate has decrease to the point where only a few percent of time is lost.
5. Irregular service interruptions are a bit of an inconvenience.
6. Some problems with jobs hanging and thereby allowing time to effectively be wasted. Somewhat peculiar to my mode of submitting many small programs within an overall large job.

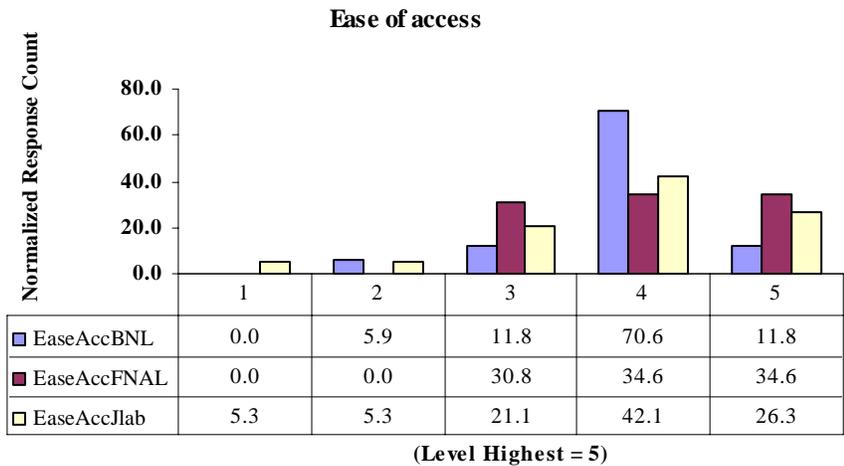
### 6.2.5 Responsiveness



#### Comments

1. See comment on support. Staff could not be better in dealing with our group.
2. The staff is very helpful and quick to respond.
3. The responsiveness of the FNAL team is outstanding

### 6.2.6 Ease of access

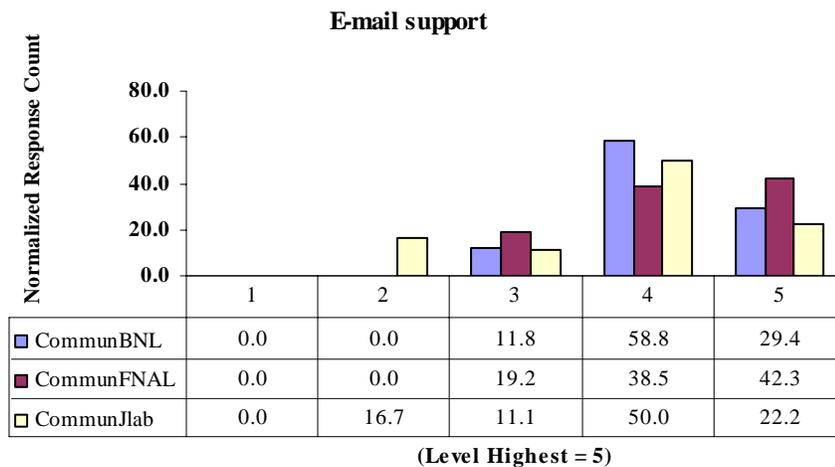


### Comments

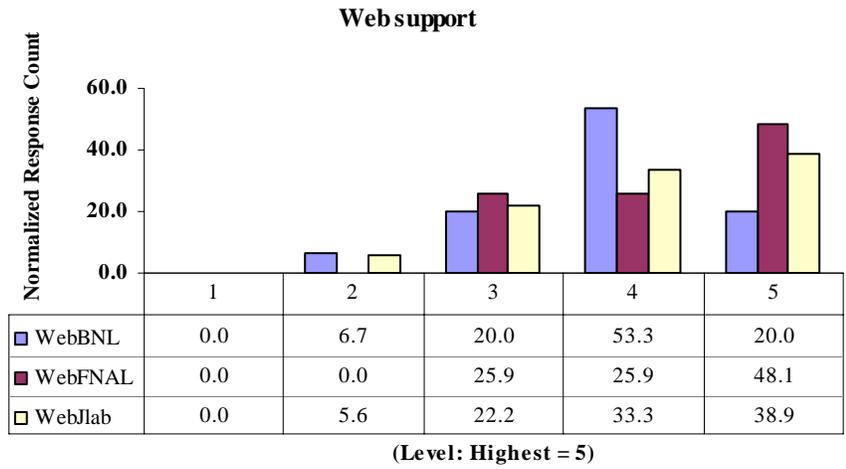
1. The FNAL Kerberos system is arcane (at least to me), and doesn't mesh well with other sites. Thus, pushing to that system has been a pain. Pulling into it is fine. Fortunately, that's straightforward pulling from BNL and JLab (ssh tunnels).
2. Very difficult to get through firewall, impossible with large file transfers
3. For some reason I can't ssh directly from my notebook (Mac), I have to ssh from a Unix server. It is also inconvenient that I can't access qcdoc host directly (have to go through ssh.bnl.gov).
4. Once I figured out how to tunnel through quark and ssh.qcdoc, I can get to host/b in one shot.
5. login via login server complicates file transfers; file transfer very slow.
6. Tight cyber-security restrictions make it more cumbersome to transfer files.
7. It's easy to transfer things to and from tape, but getting files off of tape is extremely slow and often produces a long time delay between starting to set up jobs and being able to run. This results in more downtime for the nodes.
8. This is one of the reasons that I prefer to use the FNAL clusters.
9. Local to where I use machines.

## 6.3 Support and communication

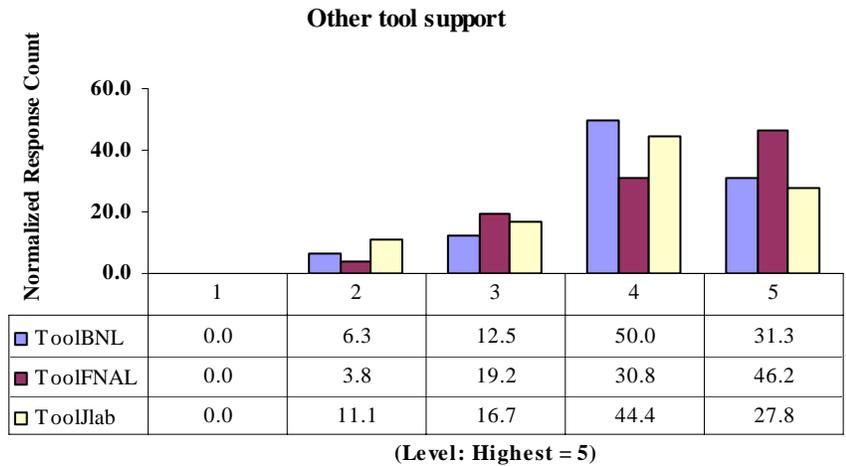
### 6.3.1 E-mail



**6.3.2 Web support**



**6.3.3 Other tool support**



**6.4 User meeting needed (Y/N count)**

Needed?	Count
Y	7
N	45

**6.5 Helpdesk (analysis of the last problem solved)**

**6.5.1 Knowledge of how to ask help (Y/N count)**

Knows	Count
Y	50
N	4

### 6.5.2 Help requested by facility (Y/N count)

Help asked?	BNL	FNAL	Jlab
Y	10	23	11
N	34	21	33

### 6.5.3 Time to initial response (working hours)

<=Hours	Freq.
6	31
12	2

### 6.5.4 Closeout from initial response (Y/N)

Closed?	Count
Y	28
N	13

### 6.5.5 Working days needed to solve the problem

<=Days	Freq.
1	29
3	7
5	2

### 6.5.6 Help provided was satisfactory (Y/N/NA)

Satisfied?	Count
Y	38
N	1
NA	2

### Comments: Providing help - Briefly describe the nature of your last resolved problem report or helpdesk ticket

1. Had missed an announcement of downtime and job wouldn't start. Support staff kindly pointed out the announcement they had sent earlier.

2. Can't recall
3. Problems getting through the obstacle course of various security measures.
4. By email contact: the move to 64 bit libraries caused difficulty with out F90 code. Resolved by clear message within very short time.
5. f90 compiler needed to be updated after upgrade
6. Machine had to be power cycled
7. Account renewal
8. Using the CCPR system at JLab - works fine.
9. Failed jobs due to some queue glitch because of something the staff was working on. (I don't know the details.)
10. Too many to list
11. I was helping a colleague use my scripts. There were issues with the setup of his user account. (This is in no way a typical problem in my experience, but it is my last resolved problem.)
12. Trouble mounting scratch disk on compute node correctly.
13. I could not access a front-end machine due to NFS problem.
14. Queue down
15. I had trouble opening an x-term on lqcd. The version of ssh on my desktop had been upgraded by the computing division at FNAL, and was no longer compatible, but Amitoj and Nirmal tracked it down.
16. When first obtaining a BNL account, couldn't ssh into the bnl server. The private key had been truncated in the e-mail I sent. This problem was resolved fairly quickly over the phone.
17. None.
18. At FNAL, my only need for help was with setting up Kerberos and SSH. I was initially pointed to some pages that weren't very helpful (or weren't even there), but after a few tries, I was directed to a very satisfactory solution. The email response time was excellent.
19. Job could not be killed since one of the running nodes crashed
20. "Disappearance" of files in the silo: new srmLs tool listed files in cache only. bug was fixed by Ying Chen
21. Kerberos logins

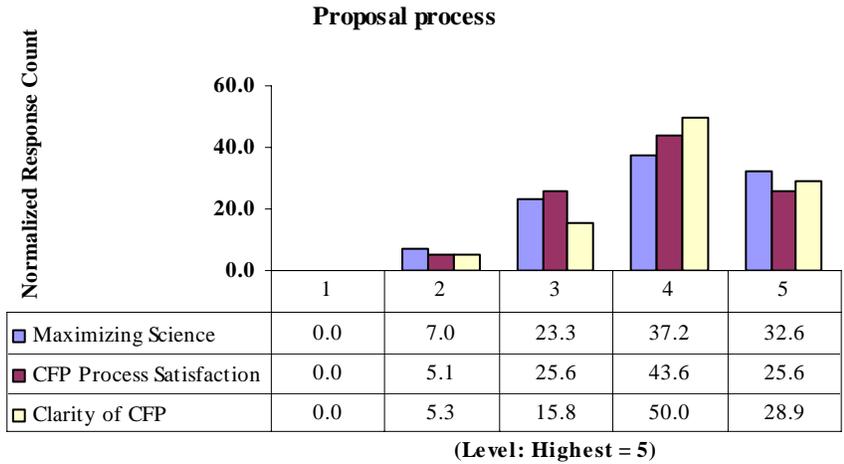
22. Kerberos problem
23. Job expired the walltime and cannot kill it.
24. Shell environment set up wrong
25. A zombie job could not be killed on the weekend.
26. I wish that more emails were sent out to lqcd-user when the cluster is temporally down, even if lqcd-admin is going to have it back up soon. This saves user's time trying to find why their jobs have been failing. I like the website that shows the cluster usage. I also appreciate "lqota", although it would be nice if we could see a breakdown of how we used time between the various clusters (qcd,pion,kaon).
27. Great help from FNAL team to set up our initial accounts and the logistics to start our Higgs project
28. Mysterious job completion errors. Distributed to appropriate staff member - gave sensible advice.
29. Kaon was down
30. Don't recall
31. Last ticket was regarding a crate being down, which was resolved in a day. Other times I have requested by e-mail help regarding access etc. which were also immediately attended to.
32. Problem reporting and resolution is now adequate at both FNAL and JLab.
33. Failure of compiler
34. The last interaction I recall was help in renewing my Kerberos principal. I can't recall my last computing problem.
35. Last resolved report was essentially solved by submitting jobs to larger numbers of nodes, which seemed to alleviate problem

**Comments: What needs should be better addressed?**

1. On-line instructions for individual PIs who wish to submit proposals for use of the facility. A more transparent, documented process.
2. NA
3. None that I can think of.
4. Website not always up to date
5. 24/7 coverage would be nice, but I realize that could be expensive unless part of lab wide support.

### 6.6 Proposal and allocations

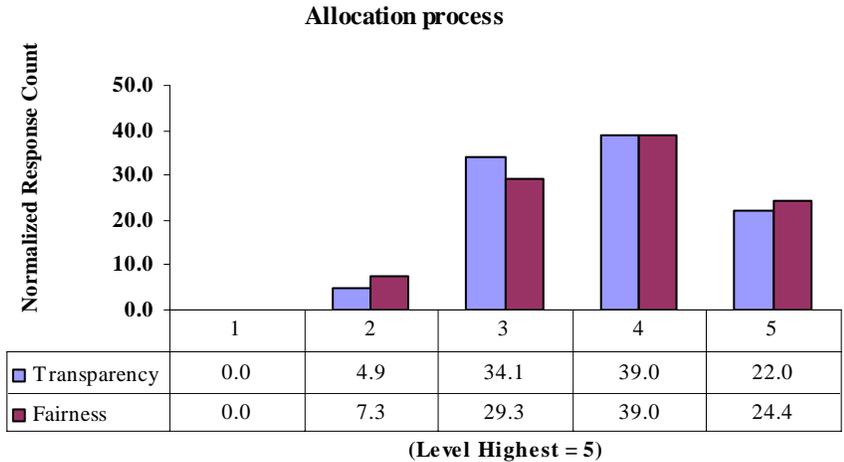
#### 6.6.1 Satisfaction with the proposal process (Rating of 1-5)



#### 6.6.2 CFP clarification and time needed

Rating	Clarification needed?	Time adequate?
Y	7	40
N	34	1

#### 6.6.3 Allocation process



**Comments**

1. Because I do not belong to a large collaboration, I have found myself largely outside of the communication loop and in the dark re. USQCD. I often find out about things after the fact. Perhaps that is also because I am a bit of a newcomer to USQCD.
2. I am part of HPQCD with Us scientists who run the allocation requests.
3. Not enough allocation for applications
4. I've never been the lead person on a proposal and so have no direct experience with the procedures.
5. Where are the proposals for 2007? I couldn't find them on the web.
6. I'm just a student, I wasn't involved in the proposal.
7. Since I am on the SPC, I recuse myself from answering these.
8. The process works best for established big user groups with computational programs that run for decades.
9. Different groups adopted different strategies in submitting proposals, e.g. several small projects encompassed in single large proposal, or several smaller proposals. Committee fairly recognized this in its allocations.

**6.7 Running jobs – Overall Job failure rate for the past year**

<= # of Failed job	Freq.
10	40
20	3
40	2
60	1
80	1
100	0

**Comments**

1. On the 7n, there are some initial problems that maybe hardware related, or possibly network driver related. 6n appears fairly clean, although there are still some linger MPI related errors on what I think are a large number of small messages.
2. Most failures seem to be due to a bad node(s) that may get assigned over and over when jobs are in a stream. Then there are glitches after changes to the system, or due to crashes. I have fewer file access problems and often those have cryptic error messages and are due to system changes as well.
3. Frequent, unpredictable, undocumented changes to the user environment made it difficult to figure out if a problem was due to a hardware issue or not

4. These numbers are for the bulk of my running of production-quality jobs. User error would be much higher for development jobs.
5. These numbers are off the top of my head, and could have large systematic errors.
6. Often, my first submission has an error in the qsub script or in my preparation of the run directory, so I need a second.
7. substantial number of failures during calculation due to node communication problems
8. Haven't run much yet
9. This is early for us to judge. We just started our project
10. most failures are from a node going down
11. usually a system down
12. failures due to occasional very slow performance that causes job to run over queue limit
13. Current systems seem to be pushing the limit of scalability for networked file systems as well as circumventing the use of networked file systems by using remote copies to move files between local file systems on different machines.
14. In my case, hanging jobs largely resolved by submitting to large number of nodes.

## 6.8 Mass storage

### 6.8.1 Adequacy of disk storage

Storage OK?	Count
Y	43
N	6

### 6.8.2 Disk storage needs in (GB)

<= Storage requested (GB)	Freq.
<b>100</b>	14
<b>1000</b>	12
<b>10000</b>	8
<b>100000</b>	1
<b>1000000</b>	1

Total disk storage requested = 587361 GB

### 6.8.3 *Disc storage backup needs (GB)*

<= Storage backup needed	Freq.
100	33
1000	0
10000	2
100000	0
1000000	1

### 6.8.4 *Tape library quality*

Tape library OK	Tape at Jlab	Instruction at Jlab	Tape at FNAL	Instruction at FNAL
Y	16	16	9	10
N	33	33	40	0

### 6.8.5 *Tape library requirements in GB for FY08*

<= Tape in GB	Permanent storage	Temporary storage
100	15	17
1000	4	6
10000	6	6
100000	8	1
1000000	0	0

### Comments

1. These figures reflect a lattice archive for USQCD community use.
2. May have to store propagators at BNL, not sure how much space they will take up yet.
3. For more accurate mass storage requirements, please refer to past and future communication with the LHPC collaboration!
4. I didn't answer the first question because the answer is "sometimes" there is adequate disk space for my work. Most of the local Fermilab users work out of /pnfs/volatile, which sometimes has enough space and sometimes doesn't. It depends upon how conscientious users are in clearing out their files when they are finished. As far as the storage estimate goes, that requires a real calculation, and shouldn't be done quickly in a survey or anonymously. But, in general, we need tens of terabytes of storage -- both on volatile and on tape. I don't understand what "online disk space means", so I can't answer that.
5. Others in the collaboration are using the disk space
6. Our disk storage need is limited yet but expected to grow with the development of the Higgs project
7. these estimates are very rough

8. Slightly confusing here, since I am only listing requirements for the SPECTRUM project, but there are associated projects that also need space.

## **6.9 General comments**

1. While my responses might be somewhat negative for the LQCD facilities; overall, they rate very well compared to other national centers. For example, I much prefer the LQCD facilities to ORNL (very negative experience so far), PSC, and NERSC. These national centers have good user support, but they are heavily overloaded with queue policies that can be highly restrictive thus forcing the user into suboptimal and unreasonable job running to get reasonable throughput. Also, some of these systems can be fairly unreliable (system errors). The LQCD facilities are highly (cost) effective and also human time effective. My criticisms merely reflect that the LQCD systems could be better.
2. Support and Communications: Overall the support is great and the communication is good (esp. web). Sometimes the emails to users are confusing (for the non-expert) and sometimes I wonder if users should know about events sooner (e.g. effects of crashes). Still, the emails have improved over the past year or two and I find them more helpful now.
3. The negative feedback applies ONLY to JLab and only to the period of 2-4 years ago when I tried to use the machines extensively. After that, I gave up.
4. I want to thank again the FNAL team all what they have done to help us
5. Fairshare system at JLab can be 'gamed' a little too easily - loopholes should be closed.
6. Still waiting for a common runtime environment across all USQCD sites. A little progress towards convergence of JLab and FNAL runtime environments is noticeable.

## **7 Lessons learned about conducting the survey**

After reviewing various options at three sites, Fermilab's ESH surveying tool was considered to be the quickest and easiest way to deploy the survey. Although the basic requirements for the survey was met by the tool and Fermilab's in-house expertise and support for tool was invaluable, this tool lacks significant capabilities for conducting a survey with a large number of questions and multiple data types. The surveying done by this tool before the LQCD survey contained less than ten simple questions. It might be worth while for Fermilab to invest in the deployment of better tools for handling various performance measures for the Laboratory.

## **8 Acknowledgement**

Matt Arena and Mary Sulek, the administrative intern of Fermilab's Computing Division, provided valuable assistance to this project.