USQCD Call for Proposals
2018

This document describes the Call for Proposals for awards of time on the USQCD computer resources dedicated to lattice QCD and other lattice field theories. These are the CPU-, Skylake-, and KNL-clusters at BNL, the CPU and GPU-clusters at Fermilab, and the KNL-cluster at JLAB.

The awards will be for calculations that further the scientific goals of the collaboration, as laid out in the recent white papers and USQCD proposals, which can be found at http://www.usqcd.org/collaboration. An important reason for funding is relevance to the DOE experimental program.

In this allocation year, we expect to distribute about:

114 M Jpsi-core-hours on the pi0 cluster at FNAL
107.6 M Jpsi-core hours on the Skylake cluster at BNL
470 M Jpsi-core-hours on KNL clusters at JLAB and BNL
7.4 M GPU-hours on GPU clusters at FNAL and BNL

400 TB disk space at BNL
780 TB disk space and 500 TB tape at FNAL
800 TB disk space and 1000 TB tape at JLAB

on dedicated USQCD hardware. The resource numbers for BNL depend on the DOE FY2018 budget which is not yet finalized. They are somewhat optimistic estimates of what we expect to have available, but the final numbers may be different. Funding for new KNL cluster resources at JLab may allow us to allocate an additional 200 M Jpsi-core-hours (not included in the numbers listed above). Details are expected to be finalized before the end of March.

Important dates:

January 26: this Call for Proposals
March 2: Type A proposals due
April 6: reports to proponents sent out
April 20-21: All Hands' Meeting at Fermilab
May 31: allocations announced
July 1: new allocations start

The web site for the All Hands' Meeting is

http://www.usqcd.org/meetings/allHands2018/

It is mandatory to acknowledge USQCD in papers using calculations carried out on these resources. The preferred text can be found at the end of this Call.

Resources:
USQCD also has community resources through the DoE INCITE program at Argonne and Oak Ridge and on the NSF supercomputer Blue Waters at NCSA. These INCITE and NSF resources will be used as was described in the successful proposals by those groups making the proposals, following any modifications required as a result of the INCITE and NSF review processes. These projects and allocations will be listed together with the allocations of the SPC on the password-protected USQCD website. These INCITE and NSF allocations will be considered by the SPC when allocating other USQCD resources in order to achieve a balanced USQCD program that addresses the Collaboration's scientific priorities.

Because of the INCITE program described above, we expect to receive zero-priority time on the BG/Q at Argonne. Based on previous usage and availability, ANL zero-priority time is expected to be available for the second half of this calendar year and the first half of next year. As the total amount of zero-priority time cannot be reliably estimated, we will allocate this resource as percentages of total zero-priority usage. The SPC may readjust these percentage allocations based upon observed usage. The Oak Ridge facility does not provide a zero-priority queue.

I. USQCD computing resources

For the facilities at FNAL and JLAB, the Scientific Program Committee will allocate 7200 hours/year to Type A and Type B proposals. Of the 8760 hours in an average year the facilities at FNAL and JLAB intend to provide 8000 hours of uptime. We then reserve 400 hours (i.e., 5%) for each host laboratory's own use, and another 400 hours for Type C proposals and contingencies. The facility at BNL intends to provide 8760 hours of uptime on the GPU, KNL, and Skylake nodes allocated to USQCD.

At BNL:
on the BNL Institutional Cluster (IC) an estimated 438,000 node-hours, on up to 120 available nodes, corresponding to approx. 80 nodes on average through September 2018 and 40 nodes on average from October 2018 through June 2019

Dual-socket Broadwell CPU's
2 NVIDIA K80 GPU's per node
128 GBytes of memory per node
EDR Infiniband interconnect

total: 438,000 node-hours = 5.01 M GPU-hours

on the BNL KNL cluster an estimated 578,160 node-hours, on up to 144 available nodes, corresponding to approx. 66 nodes on average per year

Intel Xeon Phi 7230 CPU (64 cores), 16 GB RAM on chip, 1.3 GHz
2 x 512 GB SSD (with 512 MB internal buffer) for local storage
192 GB DDR4 dual-rank RAM
Dual-rail (2x) Intel Omni-Path Host Fabric Interface Adapter 100 series
Intel TOR Omni-Path switches
1) dual-rail, non-blocking
2) 400 Gbps peak aggregate bi-directional bandwidth

total: 578,160 K node-hours = 111.0 M Jpsi-core-hours

approx. 64 nodes on the BNL Skylake cluster

Two Intel Xeon Gold 6150 CPU (36 total cores), 25 MB Cache, 2.7 GHz
1 x 4 TB SATA (6 Gbps) disk drive for local storage
192 GB DDR4 dual-rank RAM
Infiniband EDR Host Fabric Interface Adapter VPI QSFP28
Mellanox non-blocking Infiniband EDR switches

total: 8760 * 64 = 560.64 K node-hours = 107.6 M Jpsi-core-hours

The BNL systems will have access to 400 TBytes of storage on the BNL GPFS system, with a peak bandwidth of 24 GBytes/second.

For further information see https://www.sdcc.bnl.gov/.

At FNAL:

314 node cluster ("Pi0")

Eight-core, dual-socket 2.6 GHz Intel Xeon (Ivy Bridge) nodes
16 cores per node
128 GB memory/node

total: 7200*314*16*3.14 = 113.6 M Jpsi-core-hours

32 node cluster ("Pi0g")

Eight-core, dual-socket 2.6 GHz Intel Xeon (Ivy Bridge) nodes
128 GB memory/node
4 GPUs NVIDIA K40m (Kepler Tesla) per node, GPU rating 2.6
(128 total GPUs available)
GPU memory (ECC on) 11.5 GB/GPU

total = 7200*128*2.6 = 2396 K GPU-hours

These clusters will share about 780 TBytes of disk space in Lustre file systems. 500 TBytes of tape access is also available.

For further information see http://www.usqcd.org/fnal/.

At JLAB:
260 node Xeon Phi / KNL cluster ("16p")
Single socket 64 core KNL (with AVX-512 8 double / 16 single precision)
192 GB main memory / node
32 GB high bandwidth on package memory (6x higher bandwidth)
100 Gbps bi-directional Omnipath network fabric (total 25 GB/s/node)
32 nodes / switch, 16 up-links to core / switch
total: 7200*260 = 1.87 M node-hours = 359.4 M Jpsi core hours

Shared Disk
0.8 PB total for LQCD, as follows:
- Write-through cache: this is never full and data are auto-migrated to tape, with the disk copy automatically deleted as needed. Thus this consumes tape resources.
- Volatile: this is never full, with least recently used data auto deleted.

Both are able to burst above managed quotas when needed.
Finally, there is "/work", a user-managed area of limited size that is not backed up.
Note that requested disk space must include anything already present on disk that should be kept on disk. Tape is for time scales of 1-18 months and is not to be considered permanent.

For further information see also http://lqcd.jlab.org.

The following table is used to convert the different platforms to Jpsi-core-hours:

1 J/psi core-hour	=	1 Jpsi core-hour
1 BG/Q core-hour	=	1.64 Jpsi core-hour
1 pi0 core-hour	=	3.14 Jpsi core-hour
1 C2050 hour	=	1 GPU hour = 31 Jpsi equivalent core-hour
1 K20 hour	=	2.2 GPU hour
1 K40 hour	=	2.6 GPU hour
1 K80 hour	=	5.72 GPU hour
1 BNL IC node-hour	=	11.44 GPU hour
1 KNL node-hour	=	192 Jpsi core-hour
1 Skylake node-hour	=	192 Jpsi core-hour

The above numbers are based on appropriate averages of asqtad, DWF fermion, and Clover inverters. The conversion of GPU to Jpsi is based on the average of application performance on user jobs across all GPU systems at FNAL and JLab. Note that the GPU allocations are handled in GPU-hours, but the conversion to Jpsi-equivalent core hour shown above is used for overall accounting.

II. USQCD allocation procedures

The remainder of this document describes USQCD allocation procedures that are not expected to vary much from year to year. All members of the USQCD Collaboration are eligible to submit proposals. Those interested in joining the Collaboration should contact Paul Mackenzie (mackenzie@fnal.gov).

Requests can be of three types:

A) requests for potentially large amounts of time on USQCD dedicated resources and/or leadership class computers, to support calculations of benefit for the whole USQCD Collaboration and/or addressing critical scientific needs. There is no minimum size to the request. However, small requests will be not considered suitable for leadership-class resources. Allocations are for one year on USQCD resources.

B) requests for medium amounts of time on USQCD dedicated resources intended to support calculations in an early stage of development which address, or have the potential to address the scientific needs of the collaboration. There is no maximum, but the request is encouraged to be below 2.5 M Jpsi-equivalent core-hours or less on clusters, or 100 K GPU hours or less on GPU clusters. Allocations are for up to 6 months.

C) requests for exploratory calculations, such as those needed to develop and/or benchmark code, acquire expertise on the use of the machines, or to perform investigations of limited scope. The amount of time used by such projects should not exceed 100 K Jpsi core-hours on clusters or 10 K GPU-hours on the GPU-clusters.

Requests of Type A and B must be made in writing to the Scientific Program Committee and are subject to the policies spelled out below. These proposals must also specify the amount of disk and tape storage they will carry forward and the amount of each that will be created in the coming year. Projects will be charged for new disks and tapes as well as existing disk usage. How this will be implemented is discussed in section (iii).

Requests of Type B can be made anytime of the year, and will start in the nearest month. Requests should be sent in an e-mail message to the chair of the SPC, currently Aida El-Khadra (axk@illinois.edu).

Requests of Type C should be made in an e-mail message to

Paul Mackenzie (mackenzie@fnal.gov) for computing at FNAL,
Robert Mawhinney (rdm10@columbia.edu) for computing at BNL,
Chip Watson (Chip.Watson@jlab.org) for computing at JLAB.

Collaboration members who wish to perform calculations on USQCD hardware or use USQCD zero-priority time at Argonne can present requests according to the procedures specified below. The Scientific Program Committee will handle requests and awards on CPU and GPU clusters in their respective units, Jpsi core-hours and GPU hours. Requests for zero-priority time at Argonne should be made as a percentage of zero-priority usage. Conversion factors for the various platforms are given in section I. In addition, since the various GPU clusters have quite different properties, it may be useful if proposals asking for GPU time included a preference, if any, for a particular USQCD GPU system.

USQCD has adopted a policy to encourage even use of allocations throughout the year, similar to policies in use at supercomputer centers such as NERSC. Our policy requires projects to use some of their allocation in each calendar quarter. Projects that fail to do this will forfeit some of their allocation for the quarter, which will be added to the allocations of projects that are ready to run. A detailed statement of the rules for the 2018-2019 allocations will be posted on the USQCD website and announced in a separate email later this spring.

The remainder of this document deals with requests of Types A and B. It is organized as follows:

i) policy directives regarding the usage of awarded resources;

ii) guidelines for the format of the proposals;

iii) procedures that will be followed to reach a consensus on the research programs and the allocations;

i) Policy directives

1) This Call for Proposals is for calculations that will further the physics goals of the USQCD Collaboration, as stated in the proposals for funding submitted to the DOE (see http://www.usqcd.org/), and have the potential of benefiting additional research projects by members of the Collaboration. In particular, the scientific goals are described in the science sections of the recent hardware, SciDAC, and INCITE proposals and in the recent white papers, which are placed on the same web-site. It is important to our success in continued funding that we demonstrate continued importance in helping DoE

experiments to succeed.

2) Proposals of Type A are for investigations of very large scale, which may require a substantial fraction of the available resources. Proposals of Type B are intended for investigations at an early stage of development, and are smaller in size. There is no strict lower limit for the resources requested in Type A proposals, and no strict upper limit on Type B Proposals. However, Type B requests for significantly more than 2.5 M Jpsi-equivalent core-hours on clusters or more than 100 K hours on GPU-clusters, will receive significant scrutiny.

Proposals that request zero-priority time on the leadership-class computers at Argonne should be of Type A and should demonstrate that they can efficiently make use of leadership class computers.

3) All Type A and B proposals are expected to address the scientific needs of the USQCD Collaboration. Proposals of Type A are for investigations that benefit the whole USQCD Collaboration. Thus, it is expected that the calculations will either produce data, such as lattice gauge fields or quark propagators, that can be used by the entire Collaboration, or that the calculations produce physics results listed among the Collaboration's strategic goals.

Accordingly, proponents planning to generate multi-purpose data must describe in their proposal what data will be made available to the whole Collaboration, and how soon, and specify clearly what physics analyses they would like to perform in an "exclusive manner" on these data (see below), and the expected time to complete them, in accordance with the USQCD data management plan available on the website.

Similarly, proponents planning important physics analyses should explain how the proposed work meets our strategic goals and how its results would interest the broader physics community.

4) Proposals of Type B are not required to share data, although if they do so it is a plus. Type B proposals may also be scientifically valuable even if not closely aligned with USQCD goals. In that case the proposal should contain a clear discussion of the physics motivations. If appropriate, Type B proposals may discuss data-sharing and strategic importance as in the case of Type A proposals.

5) The data that will be made available to the whole Collaboration must be released promptly. "Promptly" should be interpreted with common sense and in accordance with the USQCD data management plan available on the

website. Lattice gauge fields and propagators do not have to be released as they are produced, especially if the group is still testing the production environment. On the other hand, it is not considered reasonable to delay release of, say, 444 files, just because the last 56 will not be available for a few months.
After a period during which such data will remain for the exclusive use of the members of the USQCD Collaboration, and possibly of members of other collaborations under reciprocal agreements, the data will be made available worldwide as decided by the Executive Committee.

6) The USQCD Collaboration recognizes that the production of shared data will generally entail a substantial amount of work by the investigators generating the data. They should therefore be given priority in analyzing the data, particularly for their principal physics interests. Thus, proponents are encouraged to outline a set of physics analyses that they would like to carry out with these data in an exclusive manner and the amount of time that they would like to reserve to themselves to complete such calculations.

When using the shared data, all other members of the USQCD collaboration agree to respect such exclusivity. Thus, they shall refrain from using the data to reproduce the reserved or closely similar analyses. In its evaluation of the proposals, the Scientific Program Committee will in particular examine the requests for exclusive use of the data and will ask the proposers to revise it in case the request was found too broad or excessive in any other form. Once an accepted proposal has been posted on the Collaboration website, it should be deemed by all parties that the request for exclusive use has been accepted by the Scientific Program Committee. Any dispute that may arise about the use of such data will have to be directed to the Scientific Program Committee for resolution and all members of the Collaboration should abide by the decisions of this Committee.

7) Usage of the USQCD software, developed under our SciDAC grants, is recommended, but not required. USQCD software is designed to be efficient and portable, and its development leverages efforts throughout the Collaboration. If you use this software, the SPC can be confident that your project can use USQCD resources efficiently. Software developed outside the collaboration must be documented to show that it performs efficiently on its target platform(s). Information on portability is welcome, but not mandatory.

8) The investigators whose proposals have been selected by the Scientific Program Committee for a possible award of USQCD resources shall agree to have their proposals posted on a password protected website, available only to our Collaboration, for consideration during the All Hands' Meeting.

9) The investigators receiving a Type A allocation of time following this Call for Proposals must maintain a public web page that reasonably documents their plans, progress, and the availability of data. These pages should contain information that funding agencies and review panels can use to determine whether USQCD is a well-run organization. The public web page need not contain unpublished scientific results, or other sensitive information.

ii) Format of the proposals

The proposals should contain a title page with the **title, abstract, and a complete author list, which should include all participating investigators and their affiliations.** The body, including bibliography and embedded figures, should not exceed 12 pages in length for requests of Type A, and 10 pages in length for requests of Type B, with font size of 11pt or larger. If necessary, further figures, with captions but without text, can be appended, for a maximum of 8 additional pages. CVs, publication lists and similar personal information are not requested and should not be submitted. The title page, proposal narrative and optional appended figures should be submitted as a single pdf file, in an attachment to an e-mail message sent to axk@illinois.edu.

The **last sentence** of the abstract must state the amount of computer time in Jpsi-equivalent core-hours for CPU, Skylake, and KNL clusters and in GPU-hours for GPU-clusters, in addition to the disk and tape storage needs in TBytes. Proposals lacking this information will be returned without review.

The body of the proposal should contain the following information, if possible in the order below:

1) The physics goals of the calculation.

2) The computational strategy, including such details as gauge and fermionic actions, parameters, computational methods.

3) The software used, including a description of the main algorithms and the code base employed. If you use USQCD software, it is not necessary to document performance in the proposal. If you use your own code base, then the proposal should provide enough information to show that it performs efficiently on its target platform(s). Information on portability is welcome, but not mandatory. As feedback for the software development team, proposals may include an explanation of deficiencies of the USQCD software for carrying out the proposed work.

4) The amount and type of resources requested. Here one should also state which machine is most desirable and why, and whether it is feasible or

desirable to run some parts of the proposed work on one machine, and other parts on another. If relevant, proposals of Type A should indicate longer-term computing needs here.

In addition to CPU time, proposals must specify how much mass storage is needed. The resources section of the proposal should state how much tape and disk storage is already in use, and how much new storage is needed, for disk and tape, in Tbytes. In addition, please also restate the storage request in Jpsi-equivalent core-hours, using the following conversion factor, which reflect the current replacement costs for disk storage and tapes:

1 Tbyte disk = 40 K Jpsi-equivalent core-hour
1 Tbyte tape = 6 K Jpsi-equivalent core-hour

Projects using disk storage will be charged 25% of these costs every three months. Projects will be charged for tape usage when a file is written at the full cost of tape storage; when tape files are deleted, they will receive a 40% refund of the charge.

Proposals should discuss whether these files will be used by one, a few, or several project(s). The cost for files (e.g., gauge configurations) that are available for use by all USQCD members will be borne by USQCD and not a specific physics project. The charge for files used by a single project will be deducted from the computing allocation: projects are thus encouraged to figure out whether it is more cost-effective to store or re-compute a file. If a few (2-3) projects share a file, they will share the charge.

Projects that expect to have large I/O requirements, such as those that use eigenvalue and deflation methods, are requested to note that in their proposal and to work with the site managers to handle these needs as painlessly as possible.

5) Readiness and anticipated run schedule: Are the codes and scripts ready for production running? If not, what is the anticipated timeframe for the start of the runs? Please provide a plan for a quarterly run schedule.

6) Data sharing and exclusive rights: If relevant, what data will be made available to the entire Collaboration, and the schedule for sharing it.
What calculations the investigators would like to perform in an "exclusive manner" (see above in the section on policy directives), and for how long they would like to reserve to themselves this exclusive right.

iii) Procedure for the awards

The Scientific Program Committee will receive proposals until the deadline. Proposals not stating the total request in the last sentence of the abstract will be returned without review.

Proposals that are considered meritorious and conforming to the goals of the Collaboration will be posted on the web at http://www.usqcd.org/, in the Collaboration's password-protected area. Proposals recommended for awards in previous years can be found there too.

The Scientific Program Committee (SPC) will make a preliminary assessment of the proposals. Before the All Hands Meeting, the SPC will send a report to the proponents which may seek further information about the proposal.

Following the All Hands' Meeting, the SPC will determine a set of recommendations on the awards. The quality of the initial proposal, the proponents' response to questions posed in the written report, and the views of the Collaboration expressed at the All Hands' Meeting will all influence the outcome. The SPC will send its recommendations to the Executive Committee after the All Hands' Meeting, and inform the proponents once the recommendations have been accepted by the Executive Committee. The successful proposals and the size of their awards will be posted on the web.

Scientific publications describing calculations carried out with these awards should acknowledge the use of USQCD resources, by including the following sentence in the Acknowledgments:

"Computations for this work were carried out in part on facilities of the USQCD Collaboration, which are funded by the Office of Science of the U.S. Department of Energy."

Projects whose sole source of computing is USQCD should omit the phrase "in part".