2021 USQCD Call for Proposals

February 2, 2021

1 Introduction

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

Awards will be made for calculations that further the scientific goals of the collaboration, as laid out in the recent white papers and USQCD hardware proposals, which can be found at https://www.usqcd.org/collaboration.html. It is important that USQCD sponsored research projects address the needs of the DOE funded experimental program, to ensure continued funding for USQCD-owned computational resources. In this allocation year, we expect to distribute computing and storage resources of about:

- Computational Resources

  79.04 M Sky-core hours on Skylake clusters at BNL and FNAL
  218.88 M KNL-core-hours on KNL clusters at JLAB and BNL
  1.13 M K80-GPU-hours on a GPU cluster at BNL
  1.84 M RTX2080-GPU-hours on the GeForce GPU cluster at JLAB

- Short-term Storage

  600 TB disk space and 600 TB tape at BNL
  600 TB disk space and 1000 TB tape at FNAL
  1000 TB disk space and 1000 TB tape at JLAB

on dedicated USQCD hardware. JLAB will install a new AMD MI100 GPU cluster by the beginning of July. We anticipate distributing around 0.4M MI100-GPU-hours this allocation year in a supplemental call following its installation. Members of USQCD who anticipate responding to that call are encouraged to obtain an account on the current MI50 evaluation system, as described below.

The resource numbers listed above depend on the next DOE fiscal year budget, and are therefore somewhat uncertain. In addition, the balance of available node-hours between the different platforms at BNL and FNAL may be adjusted, if necessary, to address the needs of the scientific program.

Long-term Storage: The USQCD has recently instituted a *Long-term Data-management Strategy and Plan* (USQCDSP) that can be found here, and on the USQCD Collaboration web page. As part of that strategy, each proposal must be accompanied by a *Data-Management Plan* (DMP) that must be maintained following any award; the USQCDSP provides a template, though collaborations are encouraged to use the online tool DMPtool to ensure straightforward maintenance. Short-term storage will be allocated as part of this allocation process, with tape storage allocated for six months beyond a project's end. The SPC will assess proposals for long-term storage according to the criteria laid out in the USQCDSP, and in particular will assess whether they are are of sufficiently broad interest that the cost need not be considered part of a project allocation, but borne by LQCD or NPPLC on behalf of the USQCD Community as a whole. We anticipate the SPC will perform subsequent assessments of long-term storage requests at six-month intervals.

Please note that it is mandatory to acknowledge USQCD in papers using calculations carried out on these resources. The preferred text can be found in section 5 at the end of this document.

Important dates:

03 Feb:	this Call for Proposals
08 Mar:	Type A proposals due
12 Apr:	reports to proponents sent out
30 Apr/1 May:	All Hands' Meeting at FNAL
31 May:	allocations announced
01 Jul:	new allocations start

The site managers at all three sites have agreed to provide small exploratory allocations to investigators in June, if needed, to enable all projects to get their codes ready on the allocated hardware prior to the start of the allocation year. Instructions for obtaining such early access will be included in the allocation notifications.

2 Resources

For the facility at JLAB, the Scientific Program Committee will allocate 7200 hours/year to Type A proposals. Of the 8760 hours in an average year, the facility at JLAB intends to provide 8000 hours of uptime. The facilities at BNL and FNAL intend to provide 8760 hours of uptime on the GPU, KNL, Skylake, and Cascade Lake nodes allocated to USQCD. About 10% of the node-hours available

at the three sites will be used for Type B and Type C proposals, for the incentives that are part of the jeopardy policy, and for other contingencies.

At BNL

- **The BNL Institutional Cluster (IC)**
  An estimated 564 k node-hours, on up to 210 available nodes
  Dual-socket Broadwell CPUs
  2 NVIDIA K80 or P100 GPUs per node
  256 GBytes of memory per node
  EDR Infiniband interconnect
  TOTAL: 564 k node-hours = 1128 k K80-GPU-hours

- **The BNL KNL cluster**
  An estimated 252 k node-hours, on up to 144 available nodes
  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: 252 k node-hours = 16.13 M KNL-core-hours

- **BNL Skylake cluster**
  An estimated 490 k node-hours, on up to 64 available nodes
  Two Intel Xeon Gold 6150 CPU (36 total cores), 25 MB Cache, 2.7 GHz
  4 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: 490 k node-hours = 17.64 M Sky-core-hours

- **Short-Term Storage**
  The BNL systems will have access to 600 TBytes of disk storage and 600 TBytes of tape storage on the BNL GPFS system, with a peak bandwidth of 24 GBytes/second.

- **Long-Term Storage**
  BNL anticipates the provision of up to 6PB long-term tape storage.

On the BNL IC, users can request P100s, K80s or no preference in their job configuration. For further information see https://www.sdcc.bnl.gov/.

At FNAL

- **A 175-node cluster ("LQ1")**
  An estimated 1.46 M node-hours, on up to 175 available nodes
  Twenty-core, dual-socket, 2.5 GHz Intel Xeon 6248 (Cascade Lake) nodes
  40 cores per node
  196 GB memory/node
  Intel EDR Omni Path Network
  TOTAL: 1.46 M node-hours = 61.4 M Sky-core-hours, where the quoted Sky-core-hours reflects the slightly greater performance of Cascade Lake rather than Sky Lake cores.

- **Short-term Storage:**
  These clusters will share about 600 TBytes of disk space in a Lustre file systems. 1000 TBytes of tape access is also available.

- **Long-term Storage**
  FNAL anticipates the provision of up to 5PB long-term tape storage.

At JLAB

- **440 node Xeon Phi / KNL cluster ("16p/18p")**
  Single socket 64 core KNL (with AVX-512 8 double / 16 single precision)
  192 (98) GB main memory / node 16p (18p)
  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: 3.168 M node-hours = 202.75 M KNL-core-hours

- **32-node GeForce GPU cluster ("19g")**
  Eight-GPU RTX-2080 nodes
  8 GByte memory per GPU, 192 GByte memory per node.
  Each on 100g OmniPath Fabric
  Total: 230.4 k node-hours = 1.84 M RTX2080-GPU-hours

- **8-node GPU Cluster "21g"**
  Jefferson Lab intends to purchase an AMD GPU cluster in CY2021. The system will likely feature eight nodes each composed of two AMD EPYC 7502 CPUs and eight MI100 AMD GPUs with EDR Infiniband interconnect. Each GPU will have 32 GB of ECC memory and will be connected on each node through the AMD Infiniti fabric link. It is anticipated the system will be available on July 1. Jefferson Lab has an evaluation system, called qcd20g01, composed of four MI50 GPUs that is available to USQCD members for testing. Those wishing access to the evaluation system, and early access to the 21g cluster prior to the supplemental call, are requested

to email Robert Edwards (edwards@jlab.org). It has been announced that the OLCF Exascale system named Frontier will be composed of the next generation of AMD GPUs. As such, the Jefferson Lab system will benefit the USQCD science program as well as providing an on-ramp for users to the new Frontier system.
In a full allocation year, we anticipate awarding a total of 57.6 k node-hours = 0.46M MI100-GPU-hours.

- **Short-term Storage:**
  1.0 PB total shared disk space for LQCD, as detailed below, and 1.0 PB of tape storage:
  - **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.
  - **"work"**: a user-managed area of limited size that is not backed up.

  Both the cache and volatile are able to burst above managed quotas when needed. Note that requested disk space must include anything already present on disk that should be kept on disk.

- **Long-Term Storage:**
  JLAB anticipates the provision of several PB of long-term tape storage.

For further information see https://lqcd.jlab.org.

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

1 KNL-core-hour = 0.563 Sky-core-hours
1 K80-GPU-hour = P100-GPU-hour = 33.25 Sky-core-hours
1 RTX2080-GPU-hour = 0.65 K80-GPU-hour = 21.6 Sky-core-hours

We expect
1 MI100-GPU-hour ≈ 130 Sky-core-hours

The above numbers are based on appropriate averages of asqtad, DWF fermion, and Clover inverters, whilst the conversion of GPU to Sky-core-hours is based on the average of application performance on user jobs across all GPU systems at BNL, FNAL and JLab. The KNL and GPU allocations use KNL-core-hours, and K80-GPU-hours and RTX2080-GPU-hours as appropriate; the conversion to Sky-core-hours is shown for reference, but is application dependent.

3 Procedures

This section describes the USQCD allocation procedures. All members of the USQCD Collaboration are eligible to submit proposals[1]. Requests can be of three types:

A) Requests for potentially large amounts of time for calculations which address the scientific needs of the collaboration or support calculations of benefit for the whole USQCD Collaboration. There is no minimum size to the request. Allocations are for one year.

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 500 k Sky-core-hours or less on clusters, or 25 K K80-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 20 K Sky-core-hours on clusters 2 K GPU-hours on the K80 or RTX 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. Requests of Type A should be submitted via the **2021 USQCD Proposal Submission Form** available at the All-Hands Meeting Web Site and must be received by the deadline listed in section 1.

Requests of Type B can be made anytime of the year, and will start in the nearest month. A PDF file of the proposal should be sent by email to the chair of the SPC, currently David Richards (dgr@jlab.org).

Requests of Type C should be made in an e-mail message to
James Simone (simone@fnal.gov) for computing at FNAL,
Peter Boyle (pboyle@bnl.gov) for computing at BNL,
Robert Edwards (edwards@jlab.org) for computing at JLAB.

For requests and awards on Skylake clusters, the Scientific Program Committee will use "Sky-core-hours" as a common unit. Requests and awards for KNL clusters are in terms of "KNL-core-hours", and "K80-GPU-hours", and "RTX2010-GPU-hours" are used for requests and awards on GPU clusters, reflecting the

[1]Those interested in joining the Collaboration should contact Andreas Kronfeld (ask@fnal.gov)

quite different properties of each GPU cluster. Conversion factors for the various platforms are given in section 2.

USQCD has adopted a policy to encourage even use of allocations throughout the year. The policy encourages early use of resources in the first calendar quarter and requires projects with Type A allocations to use some of their allocation in the first calendar quarter. A detailed statement of the rules for the 2019-2020 allocations is posted on the USQCD website and available from the following link: https://www.usqcd.org/jeopardy.pdf.

3.1 Policy directives

USQCD has adopted the following policies for proposals and awards of Type A and B:

1) Proposals of Type A are for large scale investigations, 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 500 k Sky-core-hours on Skylake or KNL clusters, or more than 25 k GPU-hours on the K80 and RTX GPU clusters will receive more scrutiny.

2) 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 accompanying DMP 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 USQCDSP referenced earlier.

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

3) 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 strategic importance in the narrative, and their data-sharing strategy in their DMP, as in the case of Type A proposals.

4) 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 USQCDSP. 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.

5) 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 in their DMP 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 the DMP 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.

6) 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. In general, the SPC can be reasonably confident that projects which employ USQCD software are able to use USQCD resources efficiently. Projects that employ software developed outside the collaboration must include information in the proposal to show that it performs efficiently on its target platform(s). Information on portability is welcome, but not mandatory.

7) 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.

4 Proposal format

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. For proposals of Type A, the body, including bibliography and embedded figures, should not exceed 12 pages in length, with a 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. Proposals of Type B should not be longer than about 6 pages. For proposals of both Type A and Type B, the option of an abbreviated form for *continuation proposals* is available, as described later in this section. Resumes, Curricula Vitae, publication lists and similar personal information are not requested and should not be included in the submission of any proposal for USQCD resources. The title page, proposal narrative and optional appended figures should be combined into a single PDF file. All proposals must be accompanied by a DMP as a single, separate PDF file.

The last sentence of the abstract should state the amount of computer time in units of Sky-core-hours for Sky and Cascade Lake clusters, KNL-core-hours for KNL clusters, and in units of K80-GPU-hours and RTX2080-GPU-hours as appropriate for GPU clusters, in addition to the disk and tape storage needs in TBytes.

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 the project uses USQCD software, it is not necessary to document performance in the proposal. If the project uses another 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. If a Type A proposal requests time on more

than one platform, it should include a description of how the various parts of the proposed work will be split up among the different platforms. If relevant, proposals of Type A should indicate longer-term computing needs here.

In addition to computing time, proposals must specify their storage needs. In particular, Type A proposals should state in Tbytes how much tape and disk storage is already in use, and how much new disk and tape storage is needed. In addition, the storage request should also be restated in Sky-core-hours, using the following conversion factors, which reflect the current replacement costs for disk storage and tapes:

1 Tbyte disk = 7.5 K Sky-core-hour
1 Tbyte tape = 1.3 K Sky-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.

Type A proposals should discuss in their accompanying DMP 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 time frame for the start of the runs? Type A proposals need to provide a plan for a quarterly run schedule.

Continuation Proposals

Recognising the considerable burden in writing a proposal, projects that received USQCD resources last year may submit an abbreviated narrative that refers back to last year's proposal for the motivation, science case, goals, methods and codes, as appropriate in sections 1)-3) above. This should be prominently stated in both the abstract and introduction of the proposal. Continuation proposals that use this option must, however, include the following information:

1. accomplishments of the project in the past year(s)
2. description of changes in the methodology or codes used, if any, and the reasons for these changes
3. specific goals of the project for the new allocation year and how they fit into the overall goals of the project

The other sections of the proposal, including a detailed justification of the requested resources, should be the same as for new proposals.

Both new and continuing proposals should include in the DMP a discussion of Data sharing and exclusive rights (for Type A proposals). In particular, 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 section 3.1), and for how long they would like to reserve to themselves this exclusive right.

The Scientific Program Committee will use a web interface for Type A proposals:

2021 USQCD proposal submission form

The form includes PDF file uploads for the proposal file, and for the DMP file, and requests the following summary information:

i. Proposal PI (Last Name, First Name)
ii. PI Institution
iii. Computational resource request for:
    a. CPU time (in M Sky-core-hours)
    b. KNL time (in M KNL-core-hours)
    c. GPU time (in k K80-GPU-hours)
    d. GeForce GPU time (in k RTX2080-GPU-hours)
iv. Short-term disk and tape storage (in Tbytes)
    a. to be carried forward from a previous allocation
    b. to be created new this year
v. Quarterly run plan for each requested resource (CPU, GPU, KNL) as a percentage of the total request for that resource
vi. Site preferences

5 Awards procedure

The Scientific Program Committee will receive proposals until the deadline. Each proposal will be read in detail by two members of the SPC, chosen by the chair to minimize conflicts of interest, who will then lead the discussion of the proposal by the SPC as a whole. Following this preliminary assessment, and in accord with the time line outlined earlier, the SPC will send a report to the proponents which may include a request for additional information or clarification.

Following the All Hands' Meeting, the SPC will determine a set of recommendations on the awards. Each member of the SPC will independently draft a possible allocation of resources based on the total available, and the aggregate of these allocations will form the basis for the discussion of the final recommendation. Note that members of the SPC will exclude from their independent allocations proposals on which they are participants, and refrain from commenting on such proposals. The scientific quality of the proposal, the computational appropriateness, the proponents' response to questions posed in the written report, the need to maintain a balanced physics portfolio, alignment with the DOE HEP and NP programs, 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*".