HPC Funding and Usage Proposal

Date: February 2, 2023

1 Introduction

This document describes the proposed path forward for usage and funding strategies for high-performance computing (HPC) resources at Colorado School of Mines (Mines). The intent of this guide is to facilitate researchers performing modeling and simulation on Mines' HPC resources to the best of support staff's ability, while adhering to federal grant and administration requirements and frameworks in place. This proposal is guided by three years of historical usage data (from March 01, 2019 through December 31, 2021) for evaluation and validation of the proposed approach.

The proposal incorporates discussions and subsequent resolutions arising from three meetings of an ad-hoc HPC committee tasked with working through said challenges. The committee was formed to ensure broad and equitable representation; its composition is based upon the perceived interest, investment and experience of members of the Mines HPC community, and the participation and guidance of accounting and Mines Research Council. The proposal represents the next step toward the creation of a new usage model that meets researcher needs while satisfying all compliance constraints.

Below are some points in summary form to establish the context in which the committee began addressing their charge. In particular, the following items are highlighted: (1) committee membership; (2) researcher needs and concerns; (3) relevant guidelines for compliance; (4) research administration needs and concerns; and (5) the proposed plan of action.

1.1 Committee Members

The committee members include Mines community members drawn from faculty from seven Departments as well as from Computing and Infrastructure (ITS/HPC) and Administration (ORA/A&O).

Table 1: Mines HPC Compliance Ad Hoc Committee Members.

Member	Dept	Title
Jeffrey Shragge (Chair)	GP	Associate Professor
Mahadevan Ganesh	AMS	Professor
Zhexuan Gong	PH	Assistant Professor
Mehmet Belviranli	CS	Assistant Professor
Mark Deinert	ME	Associate Professor
Vladan Stevanovic	MME	Associate Professor
Johanna Eagan	ORA	Director of Research Administration
Noelle Sanchez	A&O	Controller
Matt Ketterling	ITS	Sr Director of Infrastructure & RC Solutions
Torey Battelle	ITS/HPC	Asst Director of Research Computing

1.2 Researcher Needs and Concerns

Listed below are topics voiced by the Mines HPC community during the three ad-hoc HPC Committee meetings (with specifics inferred by HPC group):

- Compute Administration:

    - Reasonable queue times
    - Equitable job scheduling within those queues
    - Centralized support services
    - Equitable costs, both among research groups and with respect to cloud and other compute sourcing options
    - Transparent and timely invoicing
    - Reporting functions for usage

- Hardware/Software Infrastructure:

    - Provided and supported centrally by Mines, including:
        * Compute nodes
        * Management nodes
        * Data transfer nodes
        * File system
        * Networking infrastructure
        * File transfer software
        * Scheduler
        * Software to administer HPC services

- End User Services:

    - Provided and supported centrally by Mines, including:
        * Installations:
            · libraries to run software
            · user-requested applications
            · coding platforms
            · scientific languages
            · integrated languages
        * Training and outreach
        * Ticket response
        * Data transfer
        * Data and storage assistance
        * Research computing (RC) consultation services

1.3 Relevant Guidelines for Compliance:

The Office of Research Administration (ORA), represented by Johanna Eagan, emphasized the necessity of complying with Federal requirements as stipulated by grants awarded to Mines researchers. The document 'eCFR 2 CFR Part 200' was cited as the source for these requirements; Sections 200.403, 200.404, 200.405 and 200.468 were referenced as those most likely to affect how Mines researchers conduct their research. Summarized below are imperatives arising from compliance with Federal grant mandates,[1] with examples specific to Mines' HPC environment:

Expenses charged to sponsored funds must satisfy the following three principles:

- Consistent

   Must adhere to all relevant policies and procedures that in turn must apply uniformly to institution activities for both federally funded and those funded by other mechanisms.
   * Federally funded research must be charged the same rate as Mines funded research.
   * Unused compute cycles free-of-charge (i.e., as a free tier), where others are required to pay for usage, are not to be offered.
   * Priority access to compute cycles based on financial wherewithal is not considered a viable interpretation.

- Reasonable

   Expenses should not exceed those that would be incurred by a prudent person under like circumstances.
   * Services provided cannot exceed what would normally be paid in the general market.

- Allocatable

   The charges incurred are specific for that project and that project receives the 'benefits' of that charge.
   * Federally funded computing nodes cannot be shared among researchers of other projects, unless usage for that node is re-reallocated.
   * Unused core-hours from federally funded sources cannot be gifted to other projects in a share tier.

1.4 Research Administration Needs and Concerns

The following items are meant to provide additional direction while formulating the business plan but which have not been explicitly stated elsewhere:

- Priority access

   If offered, are tiers determined by:
   * core-hour charge tiers (i.e., a higher core-hour rate allows for less wait time in queue)?

---

[1]An * indicates an example of how the requirement may impact the Mines computing model

* upfront 'investment' by research group in exchange for less wait time in queue?
        * a merit-based proposal system?
        * number of users in given research group?
        * usage-based weighting?

- How to translate 'equitable access' to actual use cases?

- How to ensure adequate personnel to provide necessary support levels?

- How to fund centralized resources such as network, storage, and administration?

- How to determine, monitor and evaluate direction of HPC@Mines (governance)?

- How to administer invoicing and remittance?

1.5 Proposal Plan of Action

The proposal identifies three phases comprising a successful plan of action:

1. The model;

2. The implementation of the model; and

3. The transition from current practices to the model.

Outstanding items of discussion will also be included to ensure that all viewpoints and situations are considered and addressed.

2 Model

2.1 Background

The ad hoc HPC committee met in late Fall 2021 and early Spring 2022 semesters for the purposes of: (1) more fully understanding the HPC compliance issues raised by ORA; (2) proposing a range of possible solutions that would ensure compliance; and (3) evaluating and ranking each raised proposal to generate a recommended pathway forward for the current and future Mines HPC facilities. As part of the discussions involving an open call for potentially workable models, three candidate approaches were analysed:

1. A mixed contribution core-hour model;

2. A solely university funded platform (i.e., no cost to users); and

3. A limited condo model (details provided in Section 2.3).

After examining and evaluation the strengths and shortcomings of the three proposed models listed above, the ad-hoc HPC committee reached a consensus to recommend adopting a mixed contribution core-hour model, the structure of which is described below. The selection of the mixed contribution core-hour model represents a majority preference that emerged after discussion among committee members over three meetings. A more detailed description of this process of elimination and selection of the model being proposed is given below.

HPC models presently in use at Mines exemplify the need for HPC model revision from the federal compliance angle. The condo model, currently represented by the Mines HPC cluster Mio, in which infrastructure is centrally owned and in which researchers were able to purchase nodes with 'pre-emptive access' to add to the core cluster, violates the 'consistent' clause of federal grant requirements (see Section 1.3). The aspect of the Mio model by which users have access to the entire pool of nodes when nodes owned by other research groups are not in use by said group members, fails to observe the 'allocatable' condition of Section 1.3; grant funds cannot be shared across projects; grant monies spent on hardware for a specific grant must be used for that specific grant only.

Mandates internal to Mines and independent of federal grant requirements disqualify another possible HPC model; guidance provided by Mines leadership that users of HPC resources share some of the fiscal burden precludes the university-as-sole-funder approach.

With the above considerations in mind, the ad-hoc HPC committee agreed to focus on the mixed contribution core-hour model, while entertaining the possibility of creating auxiliary models to meet the needs of pre-existing agreements (see below) and of research groups looking for alternative queue and/or job submission options.

2.2 Core-hour Model Description

The mixed contribution core-hour model supplies an HPC environment in which all infrastructure is centrally owned and all users of Mines HPC resources are subject to a usage fee, based on core-hour usage assessments. The university will provide the funds for the platform; therefore usage will fall under 'Other Direct Cost' for researchers. A standard core-hour rate (currently set at a ceiling of $0.02 per core hour) will be applied to researcher compute cycles and invoiced independently of IDC, which will be charged separately.

Federal grant compliance of the mixed contribution core-hour model rests to some degree on the interpretation of model guidelines and subsequent implementation. Listed here are model some attributes that, depending upon one's viewpoint, could be a pro or a con:

- Inventory and support contracts are centrally managed.

- Job scheduling is equitable, meaning that there is no pre-emptive access.

- All costs are shared among all users; there is no free usage (subsidies likely are available; however, details are not yet fully worked out).

- Highly homogeneous nodes help to streamline user support and onboarding.

- Usage charges will post as 'Other Direct Cost' to projects; IDC will be assessed separately.

Past HPC compute usage across the Mines HPC community, on Wendian only, is presented in Table 2 The table presents total CPU core-hours used by Mines departments, with the associated percentage of total usage, between March 2019 and December 2021. The purpose of presenting this data is to provide an historical sense of how the proposed model's core-hour rate may have translated to past usage.

Table 2: Wendian CPU core-hour usage by department between March 2019 and December 2021.

Department	Usage (CPU core hrs)	Percentage (%)
Metallurgical & Materials Eng.	16,846,394.17	33.20
Chemistry	14,475,342.63	28.52
Chemical & Biological Eng.	8,891,339.63	17.52
Mechanical Eng.	3,499,328.23	6.90
Geophysics	2,481,916.93	4.89
Applied Math & Statistics	1,287,339.17	2.54
Physics	895,944.47	1.77
Admin	870,605.38	1.74
Chemical Eng.	724,415.32	1.43
EOG	385,111.90	0.76
Civil & Environmental Eng.	339,933.42	0.67
Geology & Geological Eng.	49,254.07	0.10
Computer Science	2,154.63	0.0042
Total	50,749,079.95	100.00

A second important way of examining usage is to consider user type based on research contributions to the Mines HPC system (specifically Wendian). For this purpose, we introduce and use the acronym PEPA (pre-emptable priority access). Table 3 shows the usage by two different groups: (1) non-PEPA or those who use the facility without making any direct financial contribution to the system; and (2) PEPA or those who have provided a direct financial contribution (via a 'PEPA value purchase').

Table 3: Wendian CPU core-hour totals by PEPA status between March 2019 and December 2021.

Usage Status	Usage (CPU core hrs)	Percentage (%)
Non-PEPA	41,860,695.88	82.49
PEPA	8,888,384.07	17.51
Total	50,749,079.95	100.00

Table 4 presents data comparing actual usage among Wendian PEPA research groups, anonymized for reasons of privacy. Fees assessed were based on per-node costs, designated at $8500 for 192 GB

nodes and $10,427 for 384 GB nodes (fees reflect a $3000 per node subsidy by the university). Note that factors other than compute rate may have driven research groups' monetary decisions: for example, 'priority access' status may have outweighed dollars for computations per hour. A core-hour rate analysis for groups historically running in the 'full' QoS (meaning those users not invested directly in Wendian) is potentially complicated by skewed pre-emption and possible cancellation occurrences; should such analysis be deemed of value, a reasonable estimate can be calculated and included in future reporting.

Table 4: Example of Wendian usage between March 2019 and December 2021 and equivalent cost for six Mines research groups based on a $0.02/core hour model.

Research Group	Usage CPU hrs	Cost $
CH 2	2,151,925	43,039
ME 2	1,469,318	29,386
PH 1	837,319	16,746
ME 1	193,054	3861
CH 1	104,988	2100
AMS 1	30,666	613
Total	475,961	95,745

2.3 Limited Condo Model Description

This description is included as a possible future option to the selected core-hour model; in the event that researcher desires and needs change, it presents a potential basis for discussion and alternative model direction. We refer to this concept as the limited condo model approach; it is based on the Mio version outlined above, with a constraint that precludes sharing of nodes not owned by a given research group with any other users. This framework would allow centralized management of resources (i.e., network, storage, infrastructure, support, maintenance) to the research groups' advantage while eliminating pre-emption and queue wait time (except intra-group). One downside for users might be that a research group would need to ensure purchase of sufficient compute capacity to meet all of their computational needs. On balance, it may be a workable enhancement to the mixed contribution core-hour model.

3 Implementation

The implementation of the mixed contribution core-hour model will at minimum center around the following themes, described in detail below:

1. Governance for HPC services

2. Resource Allocation

3. Resource Monitoring

4. Resource Accounting

5. Administrative Support

6. HPC services technical support

3.1 Governance of HPC Services

Governance of HPC at Mines will be the purview of a Mines' recognized official HPC Strategic Advisory Committee, to be created concurrent with the implementation of the Plan. Members will be drawn from the Mines HPC community, with a committee charge to provide oversight and strategic direction to Mines HPC facility administrators. Still to be established are determining the constituency and scope of the committee, and obtaining official committee recognition by Mines Leadership. Once founded, the committee will be responsible for completing the following near-term items:

- A communication plan will be developed and executed, with oversight provided by the HPC Strategic Advisory Committee.

- Governance guidelines to be developed by the HPC Strategic Advisory Committee.

3.2 Resource Allocation

Procedures to obtain resource allocations for research groups:

- Computation

  A proposal-based system, as is currently in place, will serve as the mechanism for access to HPC systems; once accepted, the group's computational environment will be created in accordance with the contents of the proposal. HPC will request an estimated number of core-hours for the given allocation period (likely one year), and will provide command-line information for checking usage by project account number. The form is fluid; its present incarnation is here: Project Allocation Request (PI Proposals).

- Storage: see Storage Policy below

3.3 Resource Monitoring

Procedures to monitor resource allocations (HPC to develop scripts and instructions to users):

- Computation

  - Command line access:
    * Run by date and project acct number with username, usage, allocation, CPU/GPU
    * Create script to track above, and also to produce allocation totals
  - Ganglia, XDMoD, other GUI:
    * Develop GUI representation of command line output
    * Format: text, graphs, charts
    * Determine access permissions/qualifications

- Storage

  - Develop script to monitor storage: residency/duration, quantity, location
  - Possibly devise GUI representation of above

3.4   Resource Accounting

3.4.1   Computation

In accordance with the mixed-contribution core-hour model (Section 2.2), research groups will be charged for HPC usage at the core-hour level on a monthly basis. Prior to being granted access, and as part of the proposal process, each research group will provide the accounting information needed to invoice for monthly compute usage, including a responsible individual and corresponding index. For non-research user categories, such as academic classes and exploratory requests, an oversight entity will be required to perform such management functions. The billing mechanisms will be automated, yet transparent, with details available upon request.

Reviewing the numbers upon which other sections of this proposal are based in an attempt to forecast use percentages and dollar values for such categories suggests that tracking mechanisms will need to be adjusted to provide useful accounting information. Presently, usage is tracked by project; an academic class can be set up as a project, but strict enforcement is not prioritized. For example, if a professor has other project accounts or if a student has alternative access, the PI may not prioritize the students' inclusion of the class project when running class assignments. Similarly, Mines Tech Fee subsidies are indicated, but not necessarily required by individual job submission. A revamp of parameters required to execute a job is one way to tailor the output of information to inform future subsidy projections, and will be implemented with the new business plan.

Examples of charge group categories, funding sources and other related items we may expect to see are bulleted below:

- Categories of charge groups for computation:

  - Research groups (projects)
  - Academic classes
  - Tech Fee/other student uses
  - Production systems: APIs and science gateways
  - Exploratory projects: viability of HPC for use case
  - Development

- Funding sources and requirements for account reconciliation:

  - Grants and PI invoicing:
    * Grants, external: e.g., NSF, NASA, NIH
    * Grants, internal: start-up, other awards
    * Consortia

        ∗ Other funding mechanisms
- Responsible party contact
- Frequency of invoicing
- Expectations for payment and reconciliation of accounts

3.4.2 Storage

Storage is available in several forms, including high-availability (HA) and replicated (Orebits), fast-access on Wendian (short-term: scratch and sets directories; and long-term: projects directory), along with other more conventional options available external to HPC resources. See Appendix A for a detailed description of CSM Storage policy.

3.5 Administrative Support

Administrative support for HPC resources will be executed primarily by Mines' HPC support group (CIARC). Policies and procedures governing this support will be enforced by Mines' HPC Strategic Funding Committee, with development and consultation provided by CIARC.

3.6 Technical Support

Technical support for HPC resources will be executed primarily by Mines' HPC support group (CIARC). Users of HPC platforms may request assistance with HPC-related issues by submitting a support form via Mines' ticketing system (currently TDX). Service categories include applications for allocation of HPC cycles, new user access, software installation, ease-of-access webapps, technical and professional consultation and a general support option, available for issues not included under other service categories. Depending on the scope of support required, time-to-resolution will vary. At the very least, an acknowledgement of ticket submission will be made within 24 hours of receipt by a member of CIARC.

4 Transition

4.1 Evaluation and Plan

Several research groups have invested grant monies (or other sources of funding) in Wendian, in accordance with agreements in place and approved at the time. These agreements will be addressed individually, on a case-by-case basis, adhering as closely and consistently as possible to a common plan developed to best meet each party's circumstances.

As a guiding principle, existing agreements will be honored until the date at which the new business plan is executed, at which point three principles, targeting the existing tiers, will describe the terms for HPC usage moving forward:

1. Research groups whose investment is paid in full will have pre-emption and priority access privileges for five years from the date of payment[2]. At that time, their privileges will expire, and the same provisions governing usage by non-invested users will apply;

2. Research groups with partial payments will be subject to the above on a pro rata basis, determined by the percentage of payment made. This percentage will translate as a percentage of the total core under agreement;

3. Research groups that have not made any payments and have been accommodated during this period of transition will have all privileges expire upon initiation of the new plan. Research groups to which refunds have been issued will fall under the "partial payment" category and be treated accordingly.

4.2 Outstanding Topics of Consideration

There are a number of outstanding issues that have not yet been discussed in this planning document. We list them here for completeness:

- GPU usage 'core-hour' rate;

- Accommodation for use cases other than established research;

- Possible tiered options for accelerated research;

- Other specific situations not addressed by document.

---

[2]Past experience dictates that a three-year "repair or replacement" warranty on node hardware is standard, generally accompanied by a 24-hour on-site response time (M-F 8-5) provision for Mines HPC. CIARC deems it reasonable to provide (internal) support for two more years (for a total of five years) on node investments made by the Mines HPC community; any extended warranty costs are anticipated to be covered by the university.

Appendix A - Proposed Wendian Storage Policy

Author: Mines HPC Group
Date: February 2, 2023

This appendix concerns the organization and management of data on Wendian. The overarching goal is to produce a coherent, efficient and sensible file system policy that will meet the majority of user, facilitator and system support needs for the reasonably foreseeable future. The intent of these storage policies is to provide clearly defined spaces and guidelines for specific categories of data. The content is a distillation of conversations among members of the HPC Steering Committee Policy Working Group and the Mines CIARC group, and incorporates some feedback from its initial iteration.

The proposed policy is based on the following two assumptions that: (1) we will charge for long-term storage; and (2) purging policies and quotas will be enforced. In the event that neither proves necessary, the policy can be easily adjusted. Disruption to user behavior and expectations should be minimal, compared to that incurred by introducing these mandates at a later time.

Summary descriptions of Wendian storage options:

/home:
A long-term directory assigned to each individual user for storing small files. Not subject to purge. Size limit of 20 GB. No charge to research group or user.

/scratch:
A short-term directory assigned to each individual user for storing data currently necessary for active research projects. Subject to purge on a six-month cycle. No limits (within reason) to amount of data. No charge to research group or user.

/beegfs/projects:
A directory assigned to each research group for storing data that:

- is currently necessary for active research projects, and;

- requires shared access to more than one member of the research group.

The /beegfs/projects directory is owned by the research group. Subdirectories are managed by the research group, including user access permissions. This directory is subject to purge on a six-month cycle, by last active date of file. There are no limits (within reason) to the amount of data that can be stored. There is no charge to the research group or the user. A 'projects' directory must be explicitly requested by the PI.

/beegfs/sets:
A directory assigned to each research group for long-term (greater than six months) storage of files. This directory is also owned by the research group and allows shared access within the research group, with permissions managed by the PI. The /beegfs/sets directory is not subject to purge.

Although no size restrictions (within reason) are imposed, a fee of $5.66 per TB per month will be assessed on more than 10TB of data in a project's sets directory. Enforcement of the 10TB limit will consist of weekly monitoring of directory size by CIARC, with notification sent to any PIs whose content exceeds 10TB. Individual discussions with PIs about options for storing the data will ensue, and once all parties are in agreement with a path forward, this plan will be implemented. Billing arrangements are not yet established, and will be agreed upon by the PI and CIARC.

Table 5: Proposed Wendian Storage Organization

Directory:	Purpose:	Purge By Time:	Limit by Quota:	Cost:
/home	small files	n/a	20 GB	inc w account
/scratch	individual user files active temporary current research	6 mos	n/a	inc w account
/beegfs/projects	active temporary shared group datasets	6 mos	n/a	none
/beegfs/sets	permanent group files shared group datasets long-term storage	n/a	n/a	<10TB: no charge >10TB: $5.66 /TB/mo