

MSCF Computational Grand Challenge Projects Electronic Proposal Guide

Deadline for receipt of complete proposals for CGCP projects is 5:00 PM PST, May 19, 2003

It is recommended that the Team Leader review the Call for Proposals of Computational Grand Challenge Projects (CGCP) in Environmental Molecular Science Research (<http://www.emsl.pnl.gov:2080/docs/tms/mscf/proposals/homepage.html>) and Frequently Asked Questions (<http://www.emsl.pnl.gov:2080/docs/tms/mscf/proposals/faq.html>) before submitting their team's proposal to the MSCF. Please contact the MSCF (mscf@emsl.pnl.gov or 509-376-3013) with any questions on the call for proposals or this form.

Proposal Submission Requirements. The Team Leader is required to submit the CGCP proposal for their team. All proposals are required to be submitted electronically via the EMSL User System (EUS). No paper proposals will be accepted. All parts of the electronic proposal form must be completed and submitted together at the same time for a proposal to be considered complete. Incomplete proposals will not be accepted or reviewed. All proposals must follow the format specified below. Proposals that do not follow the instructions, omit required information, or deviate from the proposal format will be considered as incomplete proposals, and will not be accepted or reviewed.

Before entering the EUS, information regarding team members names, affiliations and other information (described below) should be available for entry into the forms. Part of the EUS includes the capability to attach large sections of your proposal as Adobe PDF files (preferred) or as Microsoft Word files. **A User Proposal Number and User ID will be generated from your Pre-proposal submitted May 2nd via eMail. The Proposal Number and User ID will be supplied to the PI prior to final submission due May 19th. The complete proposal must be submitted via the EMSL User System using these IDs at <http://sos.emsl.pnl.gov:2080/EUS>.**

An acknowledgment for receipt of a complete proposal will be sent by email to the Team Leader. Each research participant entered into the EUS will receive an EMSL user ID number.

Questions about submission of the MSCF CGCP proposal form should be directed to the MSCF Contact (mscf@emsl.pnl.gov or 509-376-3013).

The EMSL User System is designed as a central utility for establishing and tracking collaborations with the EMSL. As you go through the EUS, the first several web pages will ask if you are an established or new EMSL user. The actual proposal section is divided into three parts: Participants, Details, and Logistics. The first section will need to be looped through for each member. The second section includes details of the research, a place to attach the bulk of the proposal, and anticipated computational needs. The third section contains questions about funding, materials to be used, and whether any parts of the research are considered proprietary.

A step-by-step description of what data needs to be entered into the EUS starts on the next page. Please review the requirements and have information and documents ready before you begin. The Adobe PDF or Microsoft Word files should be prepared ahead of time for quick attachment.

Before you enter the EMSL User Proposal system, you must agree to the Terms and Conditions for Using EMSL. After reading the terms and conditions, and you agree with the terms and conditions, select "I Agree" to enter the EUS. Select the appropriate option for past affiliation with the EMSL. Eventually you will get to the first section where the list of research collaborators for the proposal is to be entered. Note that almost all fields are required. If you or any of your team members are already affiliated with the EMSL, most of the information for participants will be filled in automatically. **Only one person can be the primary author.**

Participants

Research Team Leader and Members.

Include the following information for the Team Leader (listed first) and all of the Team Members. Postdoctoral and graduate students only need to be included if they will need computer accounts. You will also be asked if the team member plans on visiting the EMSL. Please respond accordingly.

Team Leader & All Members

Please have the following information ready for each team member.

EMSL User ID Number (if previous user)

First Name:

Full Middle Name:

Last Name:

Suffix: (if necessary)

Date of Birth:

Profession:

Telephone number:

Facsimile number:

Email address:

Country:

Type of Institution:

Institution Name:

Business Address:

Funding agencies

If previous EMSL User:

Enter your EMSL User ID Number

Check the information shown to sure our and make the necessary corrections.

If new EMSL User:

First Name:

Full Middle Name:

Last Name:

Suffix: (if necessary)

Date of Birth:

Profession:

Telephone number:

Facsimile number:

Email address:

Country:
Type of Institution:
Institution Name:
Business Address:

Once you have added yourself as the primary author, select the “Add Another Participant” to cycle through for the next team member. A separate window will open asking for this persons EMSL affiliation. Add the required information and continue with the next participant until all team members have been added. Then select the “Continue” button at the bottom of the page.

Details

Proposal Title:

The first field is the title for the proposal. This should be short and concise.

Proposal Abstract [Maximum 2 paragraphs]:

This should be about 500 words in length briefly describing what will be done, include anticipated results. This can be cut & pasted from a previously prepared Word document.

Primary Research Area:

Choose the research area most closely associated with your proposal. These areas are most in line with the DOE’s research directives. If none of these listed are appropriate, select “Other” and specify it in the next field.

- Biological and Life Sciences (excludes medical applications)
- Chemistry (excludes materials chemistry)
- Earth Sciences
- Engineering
- Environmental Sciences
- Materials Sciences (incl. condensed matter phys. & materials chem..)
- Medical Applications
- Optics
- Physics (excludes condensed matter physics)
- Polymers
- Other - Specify

Attachments:

This is where the bulk of the proposal will be entered. This can be either a single document or several smaller documents. They will be attached to the EUS proposal. A description of each section follows. The maximum length for each text section is shown in brackets. The length of the instructions for a section will not be counted against the length of that section. Requested benchmark information on scalable performance for parallel codes must be reported in terms of speedup factor versus number of processors and be based upon real elapsed wall clock times. Specifications of requested MSCF computer resources must be given in terms of processor wall-hours.

1. Project Definition [Maximum 3 pages]:

Definition of project and objectives including background and relevance to the environmental problems and research needs facing the U.S. Department of Energy and the Nation. What is the “Grand Challenge” aspect of the project?

2. Proposed Work [Maximum 10 pages]:

Describe the proposed work in enough detail to justify resource allocation. Describe the computational methods/approach that will be used. Give expected outcomes/results from the project and impacts on DOE environmental mission goals and/or national environmental research needs. Include a project work plan (in six month periods) for the duration of the proposed project (3 years) with expectation of more detail at the beginning and less at the end.

3. Software [Maximum 3 pages]:

Describe the software to be used in enough detail to demonstrate efficient scalable utilization of the MSCF’s resources for parallel calculations starting at 16 processors and going higher. If the code is not parallel (or does not efficiently scale), how much computer time will be required to develop an efficient, scalable parallel implementation of the code? If the code is parallel, provide benchmark information in enough detail to demonstrate the performance and scalability. Benchmark information must be reported in terms of speedup factor versus number of nodes and be based upon real elapsed wall clock times. If software development/porting efforts are planned, identify the length of time projected for that work and the amount of MSCF computer resources needed. Identify any MSCF available software (e.g., NWChem, Ecce, parallel tools and libraries) that your project proposes to use. The MSCF does not guarantee the ability to provide any external third-party software for a project or user. Users will be required to provide a valid executed software license agreement (or provide suitable public domain documentation) before being able to install software on any MSCF computer system.

4. Proposed MSCF Computational Resources [Maximum 3 pages]:

Estimate the MSCF computational resources (in processor wall-hours) required for the proposed work for each six month period in the work plan. Be as specific as possible (e.g., 64 processors x 500 wall clock hours = 32,000 processor wall-hours) and state the basis for your estimate (e.g., runs of previous jobs on an SMP or other parallel computer system). If the basis of your estimate is from results on a computer system that differs from the MSCF resources, use the ratio of the theoretical MFLOPS of the processors to convert to processor wall-hours. If development/porting of software is planned, also include the computational resources required for it. The requested computational resources specified here must sum to match those given in the Logistics section of the proposal.

5. Proposed MSCF Storage Resources [Maximum 2 pages]:

Describe the type and amount of storage needed for the proposed project on a yearly and cumulative basis. Distinguish between short-term and long-term disk storage. Identify what, if any, types of data in the NWfs database your project will make freely accessible to other EMSL researchers and/or publicly available. The storage resources specified here must be consistent with those given in the Logistics section of the proposal.

6. Access to Other Computational Resources [Maximum 2 pages]:

Identify and quantify the other computational resources the team will have for accomplishing the proposed project. Note that calculations that do not scale efficiently with the number of processors or

can be done effectively with less than 16 processors are expected to be run primarily on smaller computers not in the MSCF. Teams that are planning to acquire and use the Ecce software from the MSCF should identify the types of desktop workstations used (including versions of the operating system).

7. Requested MSCF Support Resources [Maximum 2 pages]:

Identify and describe any MSCF support resources that are needed for the proposed project (e.g., specific features/functionality not presently in NWChem or Ecce, guidance in using the parallel tools and libraries to develop efficiently scalable parallel software, etc.). Note that porting and parallelizing of purely sequential code is not likely to be supported.

8. References:

Include the appropriate reference citations for the proposal.

9. Curriculum Vitae for Team Leader and Team Members [Maximum 2 pages each].

Include curriculum vitae for the Team Leader plus up to **six** of the team members. Not all team members need to have a vitae. The vitae should include recent publications relevant to the proposed research.

Preferred Start Date:

Choose a date (month, day, year) when you expect to begin work. The earliest date is August 1, 2003 since this is when we plan on announcing the accepted proposals.

Preferred End Date:

Choose a date (month, day, year) when you expect to complete the work. For a 3-year proposal this would be July 31, 2006.

Will you desire the assistance of EMSL staff in obtaining and interpreting the results:

Yes No

Select yes if you anticipate the assistance of EMSL staff otherwise select no.

Do you anticipate this proposal resulting in a joint publication with EMSL staff members:

Yes No

Select yes if you have EMSL staff members on the team, otherwise select no.

PNNL Staff Contact:

If you have discussed this proposal with a PNNL Staff member, please enter his/her name.

Primary EMSL Facility:

Molecular Science Computing (MSCF): Jones, Donald R.

Since you will be primarily using the HP Linux cluster, select the MSCF.

EMSL Facilities List:

Resource List:

Select MSCF from the Facilities list that will provide a list of Resources from which to choose. First choose the "Computing: HP 256-processor Linux Cluster (opus)"

Estimated Hours of Usage:

Click the mouse on this field and enter the total number of processor wall-hours requested for the first year. (Allocation awards are expected to be on the order of 250,000-500,000 processor wall-hours for a one-year period). For example: "250,000" Click on the "Add Resource" button and the selected resource request will appear in the field below.

If you made a mistake, just click on the wrong entry followed by clicking on the "Remove Resource" button.

Next select the "Computing: Data Storage Archive (NWfs)" in the **Resource List**: Enter the estimated amount of Disk Storage space in Gbytes or Tbytes for each of the three years. The resource request will appear in the field below along with the computing hours.

When you have completed entering all the data, select the "Continue" button at the bottom of the page.

Logistics

Finally a few questions regarding the proprietary nature of the proposal and whether you will need additional equipment or if you intend to bring any equipment or materials to the EMSL as part of this proposal.

Will any results of this proposed research be considered proprietary?

Yes No

The proposed research will likely be considered non-proprietary, unless otherwise specified here. If the intellectual property provisions given in Appendix C of the EMSL Non-Proprietary Use Agreement (contact toni@pnl.gov for a copy of the Appendix C of the NPUA) are not acceptable for the proposed research, then the proposed research will be deemed "proprietary". Users engaged in proprietary research at EMSL are obligated to pay full cost recovery for their use of EMSL. In return, those users have the option to take title to any inventions made during the proposed research and to treat as proprietary all technical data resulting from the proposed research. Intellectual property developed in the course of non-proprietary research will be protected in accordance with DOE policies (as described in Appendix C of the EMSL Non-Proprietary Use Agreement - (contact toni@pnl.gov for a copy), which in general assign the intellectual property and any resulting commercial benefits to the discoverer.

Is this proposal associated with a peer reviewed funded project:

Yes No

If this Grand Challenge proposal is funded by a peer reviewed funded project select yes, otherwise, no.

Funding Agency (Agencies):

Use pick list provided in the EUS to select all the agencies from which each team member is funded. If "other", then please specify.

The following Materials and Equipment section probably will all be answered with a No response. But we need to ask for your response.

Do you plan on operating the NMR Remotely?

Yes No

Will your research involve the use of human blood, tissues, DNA, cells, cell lines, or human biological samples in any form?

Yes No

Will you be bringing or sending any chemicals to the EMSL facility?

Yes No

Will you be bringing or sending any samples to the EMSL facility?

Yes No

If you intend to bring any equipment to EMSL as part of this proposed research, please list it in the **User Equipment:** field. If you are bringing computers that will need to connect to the PNNL network, please list them as well.

If you have any additional needs or comments regarding the proposal or the process, please enter them in the **Comment:** field.

When you are finished, select the “Submit Proposal” button at the bottom of the page.

Congratulations! You are now finished.