National Science Foundation
4201 Wilson Boulevard
Arlington, Virginia 22230

NSF 06-05

Benchmarking Information Referenced in the NSF 11-511 "High Performance Computing System Acquisition: Towards a Petascale Computing Environment for Science and Engineering"

BENCHMARKING

Proposers are required to include, with each proposal, actual or estimated results of a set of benchmark runs for review and analysis. This benchmark data should include a core set of benchmarks described below and may, at the proposer's discretion, include data from additional benchmarks. All of the proposal contents, including actual or estimated benchmark data included with the proposal, will be provided to reviewers. Reviewers will also have access to a copy of the solicitation and to information about the benchmarks that proposers were asked to run. Reviewers will be asked to evaluate proposals based on consideration of both the qualitative and quantitative information supplied in the proposals. NSF will consider both the proposals themselves and the reviewers' evaluations of the proposals in selecting proposal(s) for award. NSF's decision-making will also take account of both the quantitative and qualitative information in the proposal. NSF views the benchmark data as information that is important but not the sole determinant in funding decisions.

As indicated in the solicitation, performance indicated by benchmark results may be used as the basis of performance measures included in award documents as acceptance criteria or other conditions of full funding.

The solicitation (NSF 11-511) asks proposers to:

"Provide a detailed analysis of the projected performance of the proposed system on a benchmark suite representative of science and engineering applications. This analysis should include actual results or estimated results for (a) the following benchmarks from the set that have been used in prior years under this solicitation and described in NSF 06-05: the High-Performance Computing Challenge benchmarks, updated version 1.4.1, and updated versions of WRF, PARATEC, MILC and HOMME application benchmarks; (b) an additional set of benchmarks identified by the proposing organization as best able to characterize the innovative capability of the resource being proposed. The system performance on an appropriate set of performance benchmarks will be a factor in the selection of the awards. Achievement of benchmark performance projections may be made an award condition. ***The actual results or estimated results of any benchmarks used must be submitted in the "Supplementary Documents" section of the proposal***."

The benchmarks provided by NSF should be run "as is." Minor changes in code in order to get the benchmarks to compile and/or run are permitted but should be described in the proposal. In addition, the modified version of the benchmark source code or execution scripts must be posted to a secure ftp site hosted by the proposing organization and accessible to NSF staff on the day following the proposal deadline date. In addition, at the discretion of the proposing organization, the benchmarks provided by NSF may also be run in a form in which the source code has been optimized by the proposer or vendor. If an optimized form of one or more of the NSF benchmarks is run, and/or if benchmarks other than those provided by NSF are used in addition to the NSF benchmarks, then detailed descriptions of the benchmark or code modifications, the results of the benchmark run, and copies of the version of the source code and execution scripts that were used in running the benchmark, must also be made available at the same secure ftp site on the day following the proposal deadline date. Any libraries with which the benchmarks were linked should be supplied to the HPC Resource Provider as part of the project requirements.

Benchmarks may be run on existing or prototype systems of the same design as proposed, or

estimated by well-justified extrapolation from analogous systems. In addition, proposers may choose to require vendors to demonstrate further the ability to support the research needs of the broad community of potential users by including performance data for a variety of specific applications. The choice of applications should be justified in terms of their scientific merit and their ability to characterize the potential of a system. Since optimizing system design for a particular set of applications can influence the architecture and "balance" of a system, the features of applications influencing the configuration of the proposed system should be fully explained.

If one of the benchmarks specified by NSF or by the proposing organization fails to run or cannot be run, a description of the reasons for this must be included. Benchmarks should be run on, or estimated for, a system that corresponds to what will be delivered if the proposal is successful. Any estimated benchmark performance results should be based on a well-justified extrapolation from analogous systems. "It is anticipated that demonstrated ability to achieve any benchmark results or other measures of performance provided in the proposal, whether actual or estimated, will be required as a performance metric for formal acceptance of the delivered system."

The benchmarks described below fall into two groups. Those in the first set, System Architecture Benchmarks, were selected to provide insight into the architectural features of the proposed system. Those in the second, Application Benchmarks, provide insight into how examples of applications that are of interest to groups of researchers supported by NSF.

1.0 General Benchmarking Guidelines

All actual benchmark results reported in the proposal shall be executed on exactly the same system configuration and that system configuration shall be documented. Any hardware and software that is used in the benchmarking shall be provided as part of the acquired system, unless this requirement is waived in award negotiations. The documentation shall include, but not be limited to:

1.1 Hardware

- Description of the system topology used in the benchmarks
- Memory boards, Sections, and/or Banks
- Memory Size
- CPU Manufacturer Model and Speed
- Speed of the memory and memory bus (if applicable)
- I/O Boards and Bus Interfaces
- HBAs, Network Interface Cards and TCO Offload Engine (TOE) cards including firmware
- Network adapters, including firmware
- All communications hardware, including private channels
- RAID hardware including disks, cache, firmware, channels, GBICS and interfaces
- Fibre Channel switches, if used
- Any other hardware used as part of the benchmark configuration

1.2 Software

The entire computer system software shall be identical for each benchmark run and all tests must be run with that same system software configuration (as well as hardware configuration described above). This includes, but is not limited to, the values of variables such as I/O tuning parameters and system page size settings.

Any and all software used for the benchmark execution shall be included in the final system configuration and shall be described in the benchmark documentation. This includes:

- Operating system and all tunable parameters
- Network drivers
- Network stacks, include TOEs
- I/O Drivers
- File system software and/or Volume manager
- Compiler and libraries, including I/O and MPI libraries
- All patches and bug fixes
- Any additional software used as part of the benchmark configuration

1.3 Changes

1.3.1 Source Code Changes. For the primary benchmark data, vendors or proposing

organizations may change the source code to successfully execute the application and provide correct output but only to the minimal extent needed. If desired, the proposer may submit additional runs with source code vendor optimizations. The optimized performance will be accepted if the evaluation shows the improvement can be implemented in the actual code. The proposer must provide timings for both the modified source code and the original source code.

All source code changes, including allowed changes, must be fully documented. All software changes become the property of the NSF and the United States Government and may be incorporated into and used within existing codes without restriction.

1.3.2 Makefile Changes. Makefiles may be changed in the following circumstances:

1. Proposing organizations must include makefiles and documented rationale of all make file changes as part of the submission requirements.
2. Proposers must specify the appropriate libraries used during the build process.
3. Proposers may modify the set of compiler option(s) for each code, but only one (1) version of each compiler (e.g. C, C++, and FORTRAN) may be used for all benchmark executions. For each benchmark results based on IEEE floating point arithmetic should be submitted. If desired, additional results based on non-IEEE floating-point arithmetic may also be supplied.
4. Proposers are allowed to change the definition and location of the compiler that will be used.
5. The rules one (1), two (2), and three (3) above also apply to linker flags and libraries. Only one (1) version of a library may be used; however, it is understood that within a library's release there may be 32-bit and 64-bit versions. Note that allowed changes are described in some of the application sections.

1.3.3 Run Script Changes. Where provided, run scripts may not be changed except for those changes necessary to execute the code. Examples of such permissible changes include modifying the path names of variables, changing the number of CPUs, and setting environment variables to improve I/O performance.

The vendor must provide detailed documentation on any changes to the run scripts, and state why each of the changes was made.

1.3.4 Benchmark Operational Instructions Any deviation from the benchmarking instructions, questions of interpretation, and/or proposed changes must be formally submitted and approved by NSF, in writing (email) prior to the execution of the benchmarks and the submission of results. Any results submitted which do not follow the operational instructions and without prior approval of deviations **may not be evaluated.**

- Proposers must include makefiles and documented rationale of all makefile changes as part of the submission requirements.
- All benchmark files must be written to and read from a shared/clustered file system, as would be done on a production system.
- All temporary files must be written to and read from a shared/clustered file system, as would be done on a production system.
- Proposers should try to fully utilize all CPUs per node across all nodes. If less than the full number of CPUs per node are used, the reasons for doing so should be described.

2. 0 System Architecture Benchmarks:

Each proposal should include results of executing the HPC Challenge Benchmarks, Version 1.4.1. Descriptions of the benchmarks may be found at: http://icl.cs.utk.edu/hpcc/

The benchmarks themselves may be downloaded from: http://icl.cs.utk.edu/hpcc/software/index.html

These benchmarks are comprised of 7 tests:

- HPL - the Linpack TPP benchmark which measures the floating point rate of execution for solving a linear system of equations.
- DGEMM - measures the floating point rate of execution of double precision real matrix-matrix multiplication.
- STREAM - a simple synthetic benchmark program that measures sustainable

memory bandwidth (in GB/s) and the corresponding computation rate for simple vector kernel.
- PTRANS (parallel matrix transpose) - exercises the communications where pairs of processors communicate with each other simultaneously. It is a useful test of the total communications capacity of the network.
- RandomAccess - measures the rate of integer random updates of memory (GUPS).
- FFTE - measures the floating-point rate of execution of double precision complex one-dimensional Discrete Fourier Transform (DFT).
- Communication bandwidth and latency - a set of tests to measure latency and bandwidth of a number of simultaneous communications patterns; based on b_eff (effective bandwidth benchmark).

The site contains standard rules for the HPCC benchmarks that must be followed.

An additional test in the System Architecture Benchmarks is:

- Scalable Parallel IO Benchmark Test1 (SPIOBENCH)-

SPIOBench must be run in its entirety.

The ratio of I/O processors/nodes to CPU processors/nodes may differ for the benchmark system and the full proposed systems, but a full disclosure of the number of I/O nodes and CPU nodes for both the benchmarked and proposed systems is required.

All files associated with SPIOBench must be located on a shared file system at run-time, and SPIOBench itself must be executed from that same shared file system. The hardware and software configuration for the shared file system must be explicitly stated in the vendor's submission.

All application temporary files must be written to and read from the shared/clustered file system, as would be done on a production system.

Following completion of the tests, type the command **make tar** in the *spiobench* directory to create the *spiobench_results.tar* file of the entire directory in the parent directory. This tar file must contain the results, the makefile with the tested compile and link settings, and the source files. Return the *spiobench_results.tar* file as the deliverable for SPIOBench.

For more details, please read the *README* file in the *spiobench* directory.

The Scalable Parallel I/O Benchmark measures the ability of the system to transfer data to/from the proposed shared file system. SPIOBench tests reading and writing to the shared file system across 16, 32, 48, 64, 128, 256, 384, and 512 processors.

Vendors are to configure the system using the same hardware and software being proposed.

Unless otherwise noted, the HPCC benchmarks shall be executed on actual hardware on at least processor counts of 1024 processors and 2048 processors as well as for the number of processors in the system being proposed. Proposers may provide estimated performance at the full system size level if there does not exist a system of that size at the time the proposal is submitted. However, upon delivery of the system, in the event the proposal is successful, it is expected that the delivered system shall perform at, or exceed, any estimated figures as these results will constitute a portion of any acceptance criteria.

CAUTION: Vendors are cautioned, particularly for estimated or extrapolated times, that the delivered systems will be required to demonstrate or exceed the reported levels of performance

Benchmark results must be provided in tabular form as provided below:

Procs	G-HPL	G-PTRANS	G-FFTE	G-Random Access	G-STREAM Triad	EP-STREAM Triad	EP-DGEMM	Random Ring Bandwidth	Random Ring Latency	HPL percent of peak
Count	TFlop/s	GB/s	Gup/s	GFlop/s	GB/s	GB/s	GFlop/s	GB/s	usec	percent
Base line	N										

Proposers are encouraged to submit their results for the HPCC benchmarks to the HPC Challenge upload site via:
http://icl.cs.utk.edu/hpcc/custom/index.html?lid=52&slid=77

3.0 Application Benchmarks

Four application benchmarks have been identified. They have been selected because of their ability to act as indicators of how a system will perform on the broad range of codes used by the NSF science and engineering communities.

a. WRF[1] - Multi-Agency mesoscale atmospheric modeling code: Part 1, Part 2(4 GB)
b. MILC[2] - Particle physics lattice QCD code (496 KB)
c. PARATEC[2] - Parallel Total Energy Code (592 KB)
d. HOMME[3] - High Order Methods Modeling Environment, tools to create a high-performance scalable global atmospheric model. (2.2 MB)

Each of the four Application Benchmarks above come packaged with README files, the necessary source codes, makefiles, scripts, input data sets, output datasets and mechanisms to be used to verify that correct results have been obtained. Also included are the processor counts required for each of the Application Benchmarks.

The SPIOBench and four application benchmarks are available for download. Please send an email request to Doug Baggett: dbaggett@nsf.gov. For additional information, Please include your name, organization and reference NSF Solicitation
NSF 11-511 in your request.

The principal metric collected for the Application Benchmarks is **both** wall time and CPU execution time at specified processor counts. In addition to reporting the execution times generated outputs, compiler switches and makefile modifications required to arrive at an executable shall also be provided. Benchmarks may be run on existing or prototype systems of the same design as proposed, or estimated by well-justified extrapolation from analogous systems. Benchmarks should be run on, or estimated for, a system that corresponds to what will be delivered if the proposal is successful. Any estimated benchmark performance results should be based on a well-justified extrapolation from analogous systems.

---

[1]Courtesy of the DoD High Performance Computing Modernization Program

[2]Courtesy of Department of Energy: NERSC

[3]Courtesy of NCAR

It is anticipated that demonstrated ability to achieve any benchmark results or other measures of performance provided in the proposal, whether actual or estimated, will be required as one of the performance metrics for formal acceptance of the delivered system.

Finally the results for the Application Benchmarks shall include system descriptive information as found in Section 1.0 above.

Any questions regarding benchmarks under this solicitation should be referred to Irene Qualters at iqualter@nsf.gov and Barry Schneider at bschneid@nsf.gov