

**CHSSI CCM-6**  
**Fast Parallel Methods for Multiple Length Scales**  
ITO&E Report  
Dimitrios A. Papaconstantopoulos, Principal Investigator

## Prologue

This is a revision of a report given to the external reviewers for the Operational Test Readiness Review (OTRR) of CHSSI project CCM-6. It incorporates changes noted during the external review. New text added to the document is in blue. Text to be deleted from the report is highlighted in red.

CCM-6 Project Team  
11 March 2003

## 1 Introduction

The Common High-performance Computing Software Support Initiative (CHSSI) project CCM-6, “Fast Parallel Methods for Multiple Length Scales,” seeks to develop a scalable software suite for calculating total energies and forces on atoms in systems ranging in size from small clusters of 10-100 atoms to large systems containing millions of atoms. A full discussion of the techniques used in these codes is contained in the CCM-6 Test and Evaluation Master Plan (TEMP) Addendum and the CCM-6 Software Development Plan (SDP). Briefly, CCM-6 consists of three modules:

- **Tight-Binding (TB):**

This module is a continuation of CHSSI project CCM-3, Tight-Binding Molecular Dynamics. It is designed to handle systems containing up to one thousand atoms with accuracy comparable to first-principles Density Functional Theory (DFT) codes. Assuming that the computational effort scales as the number of atoms cubed, TB calculations are approximately 1,000 times faster than comparable first-principles calculations. The CCM-6 project consists of:

1. a library of tight-binding parameters, including parameters for selected binary and ternary systems;
2. a fitting code for producing these parameters;

3. the *static* code, a tight-binding program for evaluating electronic structure and total energy;
4. the *TBMD* code for performing Tight-Binding Molecular Dynamics; and
5. SCTB, a Self-Consistent charge transfer Tight-Binding code.

Items 1-4 were developed as part of CCM-3 and have been expanded to include binary and ternary compound parameter sets, as well as spin-polarized parameters for the ferromagnetic elements. The SCTB code was developed as part of CCM-6, and fully implements a self-consistent charge transfer model, useful for studying the energetics and structures of molecules, surfaces, and defects. A complete description of all of the parts of the TB module, including publications, is available at

<http://cst-www.nrl.navy.mil/ccm6/tb/> .

- **Self-Consistent Atomic Deformation (SCAD):** This module implements a scalable parallel version of the SCAD method, which uses a fast implementation of Density Functional Theory to determine the energy of systems which can be described as overlapping atomic-like charge densities. It is therefore extremely useful in the study of ionic systems, including ferroelectrics, and can be used to predict Born effective charges and the splitting of the longitudinal and transverse optic phonon modes (“LO/TO splitting”) in ionic solids. The on-line manual

<http://cst-www.nrl.navy.mil/ccm6/scad/>

contains a complete description of the SCAD code, as well as a list of publications using the SCAD method. Because of its unique implementation of DFT, the computational effort in the SCAD code scales approximately linearly with the number of atoms in the system, and the method is highly scalable. The SCAD code has been tested on systems containing up to 128 atoms.

- **Atomistic Potentials (AP):** This module contains
  1. a library of atomistic potentials based on the Embedded Atom Method; and
  2. the Simulator Of Lattice Defects (SOLD) code, a scalable parallel EAM code.

The potentials are fit to first-principles, tight-binding, and experimental data in order to be able to handle a wide variety of structures, including interfaces, vacancies, surfaces, and defects. Potentials are being developed based on known DoD needs. The SOLD code is used to simulate the behavior of large-scale structures. The IOT&E test uses a system of 1372

atoms, using 500 Monte Carlo steps/atom to equilibrate the system and 4000 steps/atom to determine thermal expansion.

A complete description of the AP module and the SOLD code is given at

<http://cst-www.nrl.navy.mil/ccm6/ap/> .

Atomistic potentials for a variety of materials are also available at this site.

## 2 Initial Operational Test and Evaluation (IOT&E) Review

The description of the CCM-6 Initial Operational Test and Evaluation (IOT&E) Review is contained in the CHSSI Operational Test and Readiness (OTRR) Plan of 6 January 2003. The object of the test is to determine the portability, correctness, and scalability of the modules using a set of Critical Testing Parameters and Objectives (CTPs). In addition, a set of Project Management Indicators (PMIs) are used to assess various issues concerning code development and usefulness to DoD researchers. The complete description of the CTPs and PMIs are in the CCM-6 CHSSI TEMP Addendum. The specific tests of the CTPs required for the IOT&E are given in the CCM-6 OTRR Plan.

### 2.1 Critical Testing Parameters and Objectives:

This section addresses the Critical Testing Parameters and Objectives of the IOT&E Test. Each parameter (portability, correctness, and scalability) is tested for each module. The module should meet the IOT&E Threshold for minimum acceptable behavior, but it is desirable for the module to meet the IOT&E Objective parameter, indicating full success of the code.

#### 2.1.1 The Tight-Binding (TB) Module:

The primary tests were conducted with the TBMD code, which is written in Fortran 90 and uses MPI to communicate between processors. The static and fitting codes (Fortran 77, MPI) were first tested, for monatomic systems, as part of CCM-3, and then retested for binary systems in CCM-6. The parallelization of the SCTB code (Fortran 90, MPI) is identical to these codes, i.e., by k-point.

1. *Portability:* The codes are required to compile and run on three HPC Platforms. This is both the IOT&E Threshold and Objective.

Outcome: The codes all compile and run on the

- IBM SP3 at ASC Wright-Patterson

- SGI Origin 2000 at ARL Aberdeen
  - Compaq SC40 at ERDC in Vicksburg, MS.
2. *Correctness*: The IOT&E Objective is to determine the mean-square displacement (*e.g.*, the Debye-Waller factor) of a binary or ternary compound at a temperature of 300K, using a system containing at least 500 atoms. Using this data and data collected at other temperatures for the same system, we will then calculate the mean-square displacement as a function of temperature, and extrapolate this result to determine the melting temperature of the solid using the Lindemann criterion.

The system chosen for study was the binary compound SiC. The Zincblende (ZnS) structure of this material is known to be metastable. For the initial test we used a 1024 atom supercell of SiC, at its equilibrium volume. The system was equilibrated at 300K by placing the atoms at their equilibrium positions in the supercell, giving them random velocities chosen to simulate a Maxwell-Boltzmann distribution at 600K, and run for 700 time steps (2 fs per time step). The positions and velocities of the final step of this simulation were used as a starting point for the next series of tests.

We next demonstrated that identical results were found on each of the target machines for any number of processors. To do this, we ran a 100 time-step simulation, starting from the above equilibrated parameters, using 8, 16, 32, 64, and 128 processors, on each of the target machines. We then looked at the output from the ENE file produced by the TBMD code, which lists the instantaneous temperature (a measure of the kinetic energy of the system), the kinetic energy, and the potential energy of the system. The ENE files from each of the 15 runs were identical up to round-off error. Figure 1 shows the instantaneous temperature of the system as a function of time step in these runs. The fluctuations in the temperature are in the expected range for a system of 1024 atoms.

Finally, using additional runs on a 432 atom unit cell, we determined the mean-square displacement of SiC as a function of temperature, as shown in Figure 2. We then did **at a** linear extrapolation of these results to determine the temperature at which the average displacement was 15% of the equilibrium nearest neighbor distance. This occurred at 3431K. By the Lindemann criterion this should be the melting temperature of the solid, and it is in good agreement with the experimental value of 3103K. **We conclude that the TBMD code and the parameters for SiC give correct results. The discrepancy between theory and experiment is**

$$T_{TB}/T_{exp} - 1 = 3431/3103 - 1 = 0.1057 = 10.57\% .$$

This is just outside the 10% value for the TB Module Correctness CTP. However, the value obtained is well within the expected accuracy of a physical theory, as can be verified in similar quantum mechanical calculations.

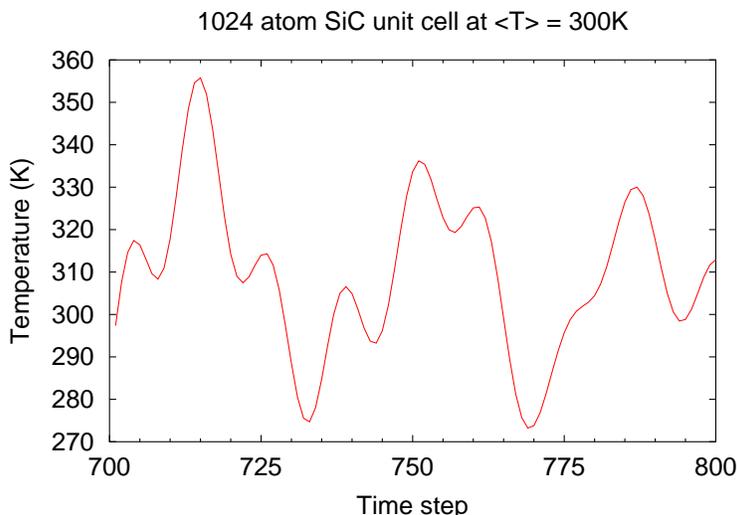


Figure 1: Instantaneous temperature of a system of 1024 atoms of the compound SiC, equilibrated at 300K.

3. *Scalability:* The IOT&E Objective is that a calculation run on 128 processors will be at least 4 times faster, measured by wall-clock times, as an otherwise identical run using 16 processors. The IOT&E Threshold is that a calculation run on 64 processors will be at least twice as fast as a calculation run on 8 processors.

The TBMD code uses the standard matrix diagonalization packages in the SCALAPACK libraries. As a result, on a scalar machine the CPU time involved in a calculation scales as the cube of the number of atoms ( $O[N^3]$  scaling), while the memory required for storage scales with the square of the number of atoms. This lead to scaling results which were highly dependent upon the number of atoms in the system. We ran calculations on the three machines using 1024 (the 100 time step calculations referred to above), 2000 (ten time steps), and 3456 atoms (3 or 5 time steps) in the supercell. Not all calculations were run on every machine. Because of memory limitations no machine was able to handle the 3456 atom unit cell with 8 processors, and the SGI Origin could not do this calculation with 16 processors.

The timings (wall clock time/time step) are shown in Table 1.

All platforms exceed the IOT&E Threshold ( $T(8)/T(64) > 2$ ) for all system sizes. The SGI exceeds the IOT&E Objective ( $T(16)/T(128) > 4$ ) even for relatively small systems (1024 atoms). The IBM SP3 and the Compaq SC40 satisfy the IOT&E Objective for a system size of 3456 atoms. As shown in the table, one could define an optimal number of processors to use on each machine for a given system size. This number increases with the number of atoms. It is

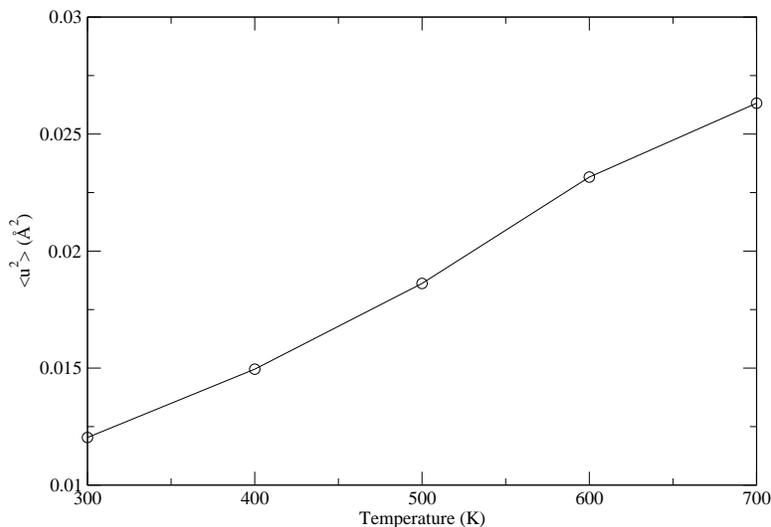


Figure 2: Mean-square displacement of atoms in a 432 atom unit cell of SiC as a function of temperature.

clear that a system containing more than 3456 atoms, running on the Compaq, produce even better scaling with 128 processors, however, ERDC has a ten hour time limit on all jobs. For larger cells, even if memory considerations would allow a 16 processor job to run, the 10 hour time limit might be exceeded before a meaningful number of iterations were computed. We therefore halted the calculations at 3456 atoms.

### 2.1.2 The Self-Consistent Atomic Deformation (SCAD) module:

The tests were conducted using the scalable parallel SCAD code, which is written in Fortran 77 and uses MPI for inter-processor communication.

1. *Portability*: The code is required to compile and run on three HPC Platforms. This is both the IOT&E Threshold and Objective.

Outcome: The code compiles and runs on the

- IBM SP3 at ASC Wright-Patterson
- SGI Origin 2000 at ARL Aberdeen
- Compaq SC40 at ERDC in Vicksburg, MS.

Table 1: Timings (wall clock time/time step, in seconds) for TBMD runs. The ratios  $T(8)/T(64)$  and  $T(16)/T(128)$  are to be used to determine the success of the code in achieving the IOT&E scalability Threshold and Objective, respectively. The time for the 3456 atom run on the IBM SP3 was mistakenly given as 4920 seconds in the initial report. Examination of the `timings` file for this run shows that the total time take was 9841 seconds for three iterations, or 3280 seconds per iteration. The basic conclusions of this report are not changed. The time for the 64 processor, 3456 atom run on the ARL Origin was added at the request of the review panel.

Platform	Atoms in Cell	Time Steps	Number of Processors					$T(8)/T(64)$	$T(16)/T(128)$
			8	16	32	64	128		
ARL	1024	100	616	556	370	86	78	7.2	7.1
SGI O2K	3456	3 or 5			10605	4309	1552		
ERDC	1000	100	263	194	103	89	86	3.0	2.3
Compaq	2000	10	2622	1026	491	635	317	4.1	3.2
SC40	3456	5		4594	2367	1917	1151		4.0
ASC	1024	100	394	268	159	157	193	2.5	1.4
IBM SP3	3456	3 or 5		11233	5353	3280	2326		4.8

## 2. Correctness:

The IOT&E objective is to use the SCAD code to compute the  $\Delta_1$  and/or  $\Delta_5$  phonon frequencies of NaCl at a point along the  $\Delta$  line  $(0, 0, x)$  midway between the origin ( $\Gamma$ ) and the high-symmetry  $X$  point  $(0, 0, 2\pi/a)$ . The phonon frequency will be computed in the frozen-phonon approximation. The point is to be chosen so that the supercell needed to perform the calculation will include at least 128 atoms.

The calculation chosen was the  $\Delta_1$  mode at the point  $(0, 0, 13\pi/(16a))$ . The unit cells required were determined using the FROZSL program, as described in the SCAD manual at

<http://cst-www.nrl.navy.mil/ccm6/scad/frozsl.html> .

This mode has two eigenstates, with frequencies of 110 and 230  $\text{cm}^{-1}$ . As shown in Figure 3, this is in excellent agreement with previous calculations using smaller unit cells. In addition, the acoustic mode frequency is in excellent agreement with experiment, while the optic mode frequency is within 10% of experiment.

The energies computed by the SCAD code are the same, within round-off error, on all platforms for each number of processors.

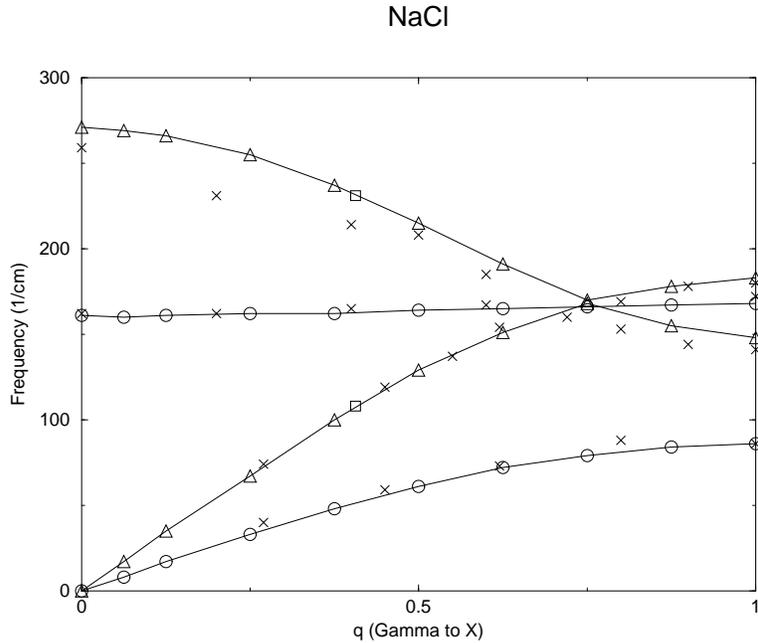


Figure 3: Phonons as a function of position along the  $\Delta$  line in NaCl. The circles and diamonds represent the SCAD calculations, and the crosses represent experiment. The boxes are the computed  $\Delta_1$  phonons at  $(0, 0, 13\pi/(16a))$ , and require a 128 atom supercell in order to do the computation.

3. *Scalability*: The IOT&E Objective is that a calculation run on 128 processors will be at least 4 times faster, measured by wall-clock times, as an otherwise identical run using 16 processors. The IOT&E Threshold is that a calculation run on 64 processors will be at least twice as fast as a calculation run on 8 processors.

In the course of the calculations it was noted that computational time and scalability were extremely sensitive to the number of atoms on each processor. Accordingly, we tabulate timings for 32, 64, and 128 atom unit cells, using a k-point appropriate to each cell. These results are shown in Table 2.

Problems with scaling are caused by two factors: (1) communication between processors, and (2) the calculation of the Madelung energy, which is not parallelized. Factor (1) is minimized on the shared-memory Origin 2000.

The SCAD code exceeds the IOT&E Threshold ( $T(8)/T(64) > 2$ ) on all platforms. The code exceeds the IOT&E Objective ( $T(16)/T(128) > 4$ ) on the Origin 2000, and nearly meets the objective on the IBM SP3 (3.76).

Table 2: Timings (wall clock time/time step, in seconds) for SCAD runs. The ratios  $T(8)/T(64)$  and  $T(16)/T(128)$  for the 128 atom unit cells are to be used to determine the success of the code in achieving the IOT&E scalability Threshold and Objective, respectively. **Note that the time for the 64 atom with 16 processors on the ARL SGI O2K was originally listed as 293 seconds. The correct value is 2933 seconds.**

Platform	Atoms in Cell	Number of Processors					$T(8)/T(64)$	$T(16)/T(128)$
		8	16	32	64	128		
ARL Origin 2000	32	2379	1352	873				
	64		2933	2254	1137			
	128	14290	7825	4593	3367	1669	4.2	4.7
ERDC Compaq SC40	32	2052	1094	781				
	64		4264	3052	2137			
	128	18779	12472	8270	6763	4749	2.8	2.6
ASC IBM SP3	32	2478	1333	765				
	64		3871	2133	1368			
	128	21803	11127	5920	3722	2957	5.9	3.8

### 2.1.3 The Atomistic (AP) Module

The tests were conducted using the scalable parallel SOLD code, which is written in Fortran 77 and uses MPI for inter-processor communication.

1. *Portability*: The code is required to compile and run on three HPC Platforms. This is both the IOT&E Threshold and Objective.

Outcome: The code compiles and runs on the

- IBM SP3 at ASC Wright-Patterson
- SGI Origin 2000 at ARL Aberdeen
- Compaq SC40 at ERDC in Vicksburg, MS.

2. *Correctness*: The IOT&E objective is to apply the SOLD code to determine the coefficients of thermal expansion of fcc Cu at several temperatures in the range 50-1200 K.

The thermal expansion coefficients were calculated by zero-pressure Monte Carlo simulations using a  $7 \times 7 \times 7$  cubic block with 1372 atoms. The simulations were performed at 13 temperatures. At each temperature, the system was brought to equilibrium by 500 Monte Carlo steps

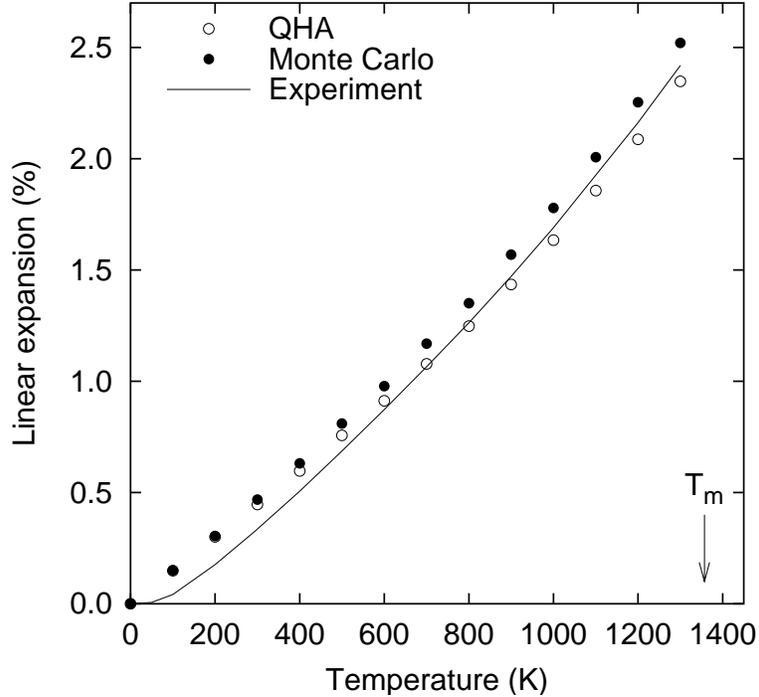


Figure 4: Thermal expansion of copper as a function of temperature as determined by the SOLD code in Monte Carlo mode (solid circles), compared to the Quasi-Harmonic Method (QHA) (open circles) and experiment (solid line).

per atom and the thermal expansion factor was determined after at least 4000 Monte Carlo steps per atom during the thermal equilibrium stage.

The results are shown in Figure 4. All platforms give the same results.

3. *Scalability:* The IOT&E Objective is that a calculation run on 128 processors will be at least 4 times faster, measured by wall-clock times, as an otherwise identical run using 16 processors. The IOT&E Threshold is that a calculation run on 64 processors will be at least twice as fast as a calculation run on 8 processors. To test this, the 1000 K calculation was run on 16, 32, 64, and 128 processors. The wall-clock times are listed in Table 3.

The SOLD code exceeds the IOT&E Objective ( $T(16)/T(128) > 4$ ) on all platforms.

Table 3: Wall-clock times for SOLD code Monte Carlo simulations of copper at 1000 K. The ratio  $T(16)/T(128)$  is to be used to determine the success of the code in achieving the IOT&E scalability Objective.

Platform	Processors					T(8)/T(64)	T(16)/T(128)
	8	16	32	64	128		
ARL SGI O2K		39670	23610	14700	9374		4.2
ERDC Compaq SC40		17940	10730	5586	2956		6.1
ASC IBM SP3		32420	17420	9534	6722		4.8

## 2.2 Project Management Indicators

CCM-6 is required to meet several program management indicators, which we outline here. More comments and notes on the PMIs are in the CCM-6 CHSSI TEMP Addendum.

- Project Planning, Tracking and Oversight

1. *Description and objectives of project developed, approved and on file with CTA Leader and CHSSI PM showing innovative use of HPC technology to solve critical military application.*

The modules in CCM-6 are used to study the structure of materials at the atomic level, which can result in increased reliability and lower lifetime operating costs for DoD systems.

2. *Comprehensive Software Development Plan emphasizes best practices from standard software engineering principles developed, approved and on file with CTA Leader and CHSSI PM.*

All programs are written in standard Fortran 77 or Fortran 90 with interprocessor communication controlled by the Message Passing Interface (MPI) libraries.

3. *Software Development Plan (SDP) reviewed at least annually with CTA Leader and CHSSI PM for currency/update.*

SDP last updated August, 2002.

4. *Project milestones, financial data and metrics approved, tracked and reported regularly to CTA Leader.*

Financial data is reported monthly, performance metrics quarterly, as required.

5. *Principle project team members and organization (responsibilities) determined and provided to CTA Leader and CHSSI PM.*

The members of the team and their responsibilities are listed in the TEMP and SDP

6. *Project progress reviewed by CTA Leader at least semi-annually and reviewed by independent team and CHSSI PM annually.*

Reports are filed quarterly with the CTA Leader and the CHSSI PM.

- Requirements Management/Configuration Control

1. *Comprehensive Technical Reference Manual and Software Development Manual developed and maintained current for use by development team and approved users.*

All manuals are online and available to users at

<http://cst-www.nrl.navy.mil/ccm6/>

2. *Requirements, input, review, approval/rejection and feedback process developed and known by team and user community.*

Complete contact information is available at the above website.

3. *Error identification, fix and testing procedures developed and made known to team and user community.*

The website informs users on the mechanism for reporting errors.

4. *Project team members maintain regular contact with one another and Project Leader reviews progress and lessons learned routinely.*

Team members and the team leader are in contact on a regular basis.

- Quality Assurance, Tech Transfer and User Support

1. *Comprehensive, easy to follow User Manual developed and maintained current with latest software release.*

Manuals are updated with each software release.

2. *Comprehensive, easy to follow Installation Guide, Lessons Learned and Training Program developed and maintained current.*

These items are available on the website. We note that Installation procedures necessarily change with HPC center, as no two centers can seem to agree on the location of common software, e.g., BLAS, LAPACK, and SCALAPACK.

3. *“Helpdesk” and on-line assistance provided as required and available.*

Help is available to registered users from the developers. Registration information is kept on file by the project leader.

4. *Multi-level software testing, error fixes, lessons learned and validation/verification methods, schedules and results documented and made available to user community, CTA Leader and CHSSI PM.*

Web pages are regularly updated to indicate changes in the software.

5. *User access, responsibilities and project sensitivities and security policies, procedures, and requirements.*

User access is controlled by an online registration process.

6. *Software release, export control, licensing and/or data exchange policies, procedures and requirements documented, maintained and advertised to user community and made available to CTA Leader and CHSSI PM.*

Software release policy is on web-site registration page.

7. *Project files are kept current with all required and historical data and information and are reviewed regularly by CTA Leader.*

The CTA Leader regularly reviews all aspects of CCM-6.