JÜLICH
FORSCHUNGSZENTRUM

JSCNews
Jülich
Supercomputing
Centre

No. 164 • May 2008

## JUMP Upgrade to Power6

The largely well-proven IBM Power4 cluster JUMP is reaching the end of its life cycle: all 41 frames will be decommissioned in July 2008. Before the general purpose part of the Jülich dual supercomputer complex can be extended significantly, an interim solution based on IBM's Power6 technology will be established providing roughly the previous compute performance.

This solution will be implemented by installing one rack IBM p6-575 composed of 14 nodes running AIX, each containing

- 32 Power6 SMT processors with 4.7 GHz
- 128 GB main memory
- 10 Gbps Ethernet Adapter
- Infiniband Adapter

The overall peak performance of the new Power6 system will be 8.4 TFlop/s.

As before, access is granted to the Jülich storage server JUST where all user HOME, ARCH and WORK directories will be retained unchanged and will be accessible via GPFS. For JUMP users the transition to the new machine will be mostly transparent, although some work will be required to utilize the new architecture efficiently. Possible areas of modification are:

- Code re-compilation to gain more performance out of the Power6 processor
- Instead of the HPS Switch (Federation) Infiniband is now used for MPI communication. User environment variables will likely have to be changed.
- Adapted usage model: Instead of 1312 Power4 processors there are now 448 much higher-performing Power6 processors with the simultaneous multithreading (SMT) capability. Therefore fewer jobs than before can run simultaneously. The LoadLeveler job class will have to be adapted to this new situation.

For more detailed information, please watch the "Message-Of-Today". We will keep you informed about all updates of the new machine, which will go into operation at the beginning of July.
(Contact: Klaus Wolkersdorfer, ext. 6579)

## Blue Gene/P Porting and Scaling Workshop

From 22 to 24 April 2008, JSC organised the first Jülich Blue Gene/P Porting, Tuning, and Scaling Workshop to provide first-hand knowledge on how to port applications to the Blue Gene/P system, debug them on large numbers of nodes, and how to scale existing Blue Gene codes to the new full 16-rack system JUGENE of JSC. For this purpose, JSC provided CPU time in the order of one Blue Gene/P rack month for the 43 participants from France, Germany, Ireland, Spain, Sweden, Switzerland, the UK and USA.

Six advisors from the JSC parallel support team, as well as three experts from IBM and one debugging expert from Totalview Technologies, supported the participants with these tasks. On the first morning, IBM speakers introduced the Blue Gene system and architecture and gave practical tips and tricks on how to port and tune codes to the machine.

Forschungszentrum Jülich GmbH
in der Helmholtz-Gemeinschaft
Jülich Supercomputing Centre
52425 Jülich I Germany

Phone +49 2461 61-6402

jsc@fz-juelich.de
www.fz-juelich.de/jsc

The other two days of the workshop were dedicated to hands-on sessions where the participants tried to port and tune their application codes, which they brought to the workshop, with the help of the JSC, IBM, and Totalview experts. Almost all the participants succeeded in scaling their code to at least one Blue Gene/P rack (4,096 cores). A few participants were able to make use of four, eight, or even the full 16 rack (65,536 cores) system. Participants brought Fortran, C, and C++ codes parallelized with MPI and hybrid MPI/OpenMP from all kinds of application areas (DFT, electronic transport and structure, CFD, neuron simulation, discrete dipole approximation, molecular simulations, convective turbulence simulation, quantum chemistry, finite element tearing, astrophysics, exact diagonalization, and others).

A special success story of the workshop was the optimization of a piece of code in the CPMD programming package. Together with IBM CPMD expert Dr. Alessandro Curioni, members of Prof. Dominik Marx's group from Ruhr-Universität Bochum developed an alternative way of programming a certain communication pattern, which speeded up the scaling part of the ultrasoft pseudopotential by a factor of two, much more than expected.

At the same time as the practical experiments, more detailed tutorials on the parallel debugger Totalview and the parallel scalable performance analysis tool Scalasca, developed by JSC, were presented and prompted the participants to experiment with these tools. During the workshop, Totalview succeeded in attaching to 32,768 processes in one experiment. Participants effectively used Scalasca to make performance measurements of their application on up to 8,192 cores. For more information including the presentation slides see: *http://www.fz-juelich.de/jsc/bg-ws08/*.

## Usage Model of the Jülich Storage Server

Shortly before the new Blue Gene/P system was installed last year, a new storage server JUST was put into production as the GPFS server for the supercomputer complex. JUST has a capacity of 1 petabyte and an aggregated performance of 20 Gbytes/s. At the same time, a new usage model of the user file system was defined:

HOME: 3 Tbytes and 2 million datasets per project/group. Data is backed up daily and there is **no** migration to tape.

ARCH: 2 million datasets per project/group; no space limit. Data is backed up daily and **all** data will be migrated to tape.

WORK: 6 Tbytes and 4 million datasets per project/group. There is **no** backup of any data and data will be deleted 28 days after last usage. This file system is designed for maximum IO performance and thus is highly recommended for IO-bound jobs.

Please note that there is **no** automatic data migration from the user HOME directories any more. Instead, the user is supposed to move data manually to the ARCH directories where all files will be migrated to tape.

Currently, users tend to fill up the HOME directories without utilizing the ARCH directories to a sufficient extent. If this growth continues for the next few months, all HOME directories will be filled up and all users will suffer from insufficient space in HOME. For this reason we urge all users to move (not just copy!) big or not recently used datasets to ARCH. More detailed information can be found at:
*http://www.fz-juelich.de/jsc/just/userinfo.pdf*
(Contact: Lothar Wollschläger, ext. 6420)

## First PRACE Newsletter Issued

PRACE, the Partnership for Advanced Computing in Europe, published the first issue of the PRACE newsletter. The article "PRACE to create the most powerful computing environment in Europe" features an interview with the chairman of the PRACE Management Board, Prof. Achim Bachem, and describes the objectives of the PRACE project. The newsletter can be found at: *http://www.prace-project.eu/*.

## Events

**NIC Project Reports 2008**

Time: Thursday, 12 June 2008, 9:00 - 12:30
Venue: Hörsaal, JSC
09:00 S. Gottlöber, AIP: The small scale structure of the universe
09:30 Z. Fodor, NIC Zeuthen: Bericht der Forschungsgruppe "Elementarteilchenphysik"
10:00 R. Sommer, DESY Zeuthen: 2 Flavour QCD with non-perturbative precision
11:00 U. Hansmann, NIC: Bericht der Forschungsgruppe "Computergestützte Biologie und Biophysik"
11:30 R. Grauer, Ruhr-Univ. Bochum: Statistics of heavy particles in MHD turbulence
12:00 A. Muramatsu, Univ. Stuttgart: Numerical studies of correlated quantum systems
Details: *http://www.fz-juelich.de/nic/projectreports-2008*

**VSR Project Reports 2008**

Time: Monday, 16 June 2008, 13:30 - 15:00
Venue: Hörsaal, JSC
13:30 R. O. Jones, IFF: Structures of binary Ge/Te alloys: Why are they interesting? What have we learned?
14:15 H. Elbern, ICG-2: Combining models with measurements: Earth observation by data assimilation for atmospheric chemistry and pollution forecasting
14:35 M. Schultz, ICG-2: Chemie-Klima-Simulationen für den 5. Sachstandsbericht des IPCC
Details: *http://www.fz-juelich.de/vsr/seminar*

**Editor: Dr. Sabine Höfler-Thierfeldt, ext. 6765**