ForHLR: a New Tier-2 High-Performance Computing System for Research

Robert Barthel and Simon Raffeiner

Steinbuch Centre for Computing, Karlsruhe Institute of Technology

ForHLR (short for German "Forschungshochleistungsrechner") is a new petaflop high-performance computing (HPC) system dedicated to multiparallel, i.e., 100 – 1000 core, applications and consists of two installation phases: ForHLR Phase I and ForHLR Phase II. The former has already been in operation since September 2014, while the latter was inaugurated on March 4, 2016. Both systems are available to the entire German scientific community. All compute proposals are to be peer-reviewed by the HLRS steering committee.

1 Introduction

ForHLR, short for German "Forschungshochleistungsrechner", is the current installation of the tier-2 high-performance computing (HPC) system in the state of Baden-Württemberg, Germany [1]. This HPC system gaps the currently fastest European supercomputer, i.e., Hazel Han, (tier-1 system) and the HPC enabling (tier-3) infrastructure, i.e., bwUniCluster and the bwForCluster JUSTUS, MLS&WISO, NEMO and BinAC, in the state of Baden-Württemberg, Germany. The ForHLR system consists of two cluster: ForHLR Phase I and ForHLR Phase II.

2 Hardware Architecture

2.1 ForHLR Phase I

The ForHLR I system marks the first phase of the ForHLR installation and has been cofunded by the Ministry of Science, Research and the Arts Baden-Württemberg, Germany and DFG (German Research Foundation).

The ForHLR Phase I infrastructure [2] consists of 512 compute nodes with two 10-core Intel Xeon E5-2670v2 ("Ivy Bridge") processors and 64 gigabytes of main memory each, plus 16 compute nodes with four 8-core Intel Xeon E5-4620v2 processors and 512 gigabytes of main memory each. Total peak performance for the system is 216 TeraFLOPS, the total amount of main memory is 41.1 Terabytes. All nodes are connected to a central InfiniBand fabric at 56 GBit/s per node (4x FDR).

On ForHLR Phase I two parallel file system based on Lustre are used for globally (i.e., on all nodes) accessible user data. Both file systems differ in usability lifetime, redundancy and total

capacity. While `$HOME` is permanent and under backup, its 469 Terabyte is not primarily for computing activities. Globally available data scratch is provided by the workspace file system with a total capacity of 938 Terabyte. However workspaces have a limited lifetime and no redundance.

2.2 ForHLR Phase II

The ForHLR Phase II system [3] marks the second phase of the ForHLR installation and has also been cofunded by the Ministry of Science, Research and the Arts Baden-Württemberg, Germany and DFG (German Research Foundation).

It consists of 1152 compute nodes with two 10-core Intel Xeon E5-2660v3 ("Haswell") processors and 64 gigabytes of main memory each, plus 21 compute nodes with four 12-core Intel Xeon E7-4830v3 processors and 1 Terabyte of main memory each. Total peak performance for the system is one PetaFLOPS, the total amount of main memory is 95 Terabytes. Similar to ForHLR Phase I, a permanent (`$HOME`) as well as a workspace parallel filesystems are available for data storage with 670 Terabyte and 4.8 Petabyte, respectively.

All nodes are connected to a central InfiniBand fabric at 56 GBit/s per node (4x FDR), while the InfiniBand Backbone uses 100 GBit/s connections (4x EDR). The ForHLR II system is located at KIT's Campus North and is coupled to the existing Lustre parallel filesystems through a 320 Gbps long-range InfiniBand link bridging the 11 kilometers to SCC's second data center at Campus South.

A novelty of this installation is the highly efficient hot-water cooling system, which reuses the waste heat for the heating of the nearby office building.

3 Access & Registration

The ForHLR system is available to all research areas and has been customized, in particular, to energy science, environmental science, materials science and engineering technology. Moreover both systems are dedicated to solving highly complex scientific scenarios using highly scalable and multiparallel application codes, i.e., utilizing 100 – 1000 CPU cores. Smaller compute projects ought to be for ForHLR Phase I while the most resource demanding ones are preferably assigned to ForHLR Phase II.

Although academic use is free of charge, preliminary accounts are given only for a short term to test ForHLR architecture and the software stack while accounts of long term compute projects are bound to approved CPU hours.

Compute projects are peer-reviewed in terms of application scalability/usability and CPU hour feasibility for the ForHLR systems. To be proper reviewed applicants must submit a 3 to 5 page long project description elsewhere [4].

4 Citation

As every funded work, the use of the ForHLR systems have to be cited in any work related publication as follows:

"This work was performed on the computational resource ForHLR Phase I (ForHLR Phase II) funded by the Ministry of Science, Research and the Arts Baden-Württemberg and DFG (German Research Foundation)."

References

[1] Hartenstein, H., T. Walter, and P. Castellaz. "Aktuelle Umsetzungskonzepte der Universitäten des Landes Baden-Württemberg für Hochleistungsrechnen und datenintensive Dienste." Praxis der Informationsverarbeitung und Kommunikation, Band 36, Heft 2 (2013): 99-108. http://dx.doi.org/10.1515/pik-2013-0007

[2] Häfner, H.: "ForHLR – ein Hochleistungsrechner für die Forschung". SCC-News, 02-2014, pp 4–5.

[3] Gernert, H., R. Lohner and R. Mayer: "ForHLR – Der neue Forschungshochleistungsrechner am KIT". SCC-News, 01-2016, pp 8–12.

[4] http://www.scc.kit.edu/dienste/proposals.php