Matthieu Dorier, PhD

Research and development of I/O, storage, and data management solutions, in support for the efficient use of high-performance computing platforms.

6520 Double Eagle Dr, Apt 722
Woodridge, IL 60517, USA
+1 (630) 589-9383
matthieu.dorier@gmail.com
http://dorier.github.io

EXPERIENCE

Argonne National Laboratory, USA — *Postdoctoral Appointee*

FEBRUARY 2015 - PRESENT

- Conducted research on HPC I/O, in situ analysis, parallel storage, and HPC communication algorithms for MPI.
- Developed parallel software in C/C++ for HPC simulations and MPI trace processing/replaying. Conducted large-scale experiments.
- Established collaborations with several research groups in US (HDF Group, UC-Davis, RPI) and France (Inria Grenoble).
- Published numerous papers in top-ranked journals and conferences.

IRISA, Inria Rennes Bretagne–Atlantique, France — *PhD Student*

SEPTEMBER 2011 - DECEMBER 2014

- Conducted research on HPC I/O and in situ visualization/analysis.
- Developed parallel software in C/C++ for improving the I/O of HPC simulations and coupling them with visualization software.
- Contributed to several international collaborations.
- Published numerous papers in top-ranked journals and conferences.

National Center for Supercomputing Applications, USA — *Intern*

MAY 2010 - AUGUST 2010

- Investigated I/O performance issues in the CM1 atmospheric simulation, proposed/implemented new solutions to overcome them.

EDUCATION

Ecole Normal Supérieure de Rennes, France, *Normalien*

2011-2014 - PhD in Computer Science

- "Addressing the challenges of I/O variability in post-petascale HPC simulations." PhD thesis supervised by Gabriel Antoniu and Luc Bougé

2009-2011 - Master of Computer Science

- Specialization on parallel and distributed systems
- Obtained with honors, ranked 1st/70 students

2008-2012 - Magistère of Computer Science and Telecommunications

2008-2009 - Bachelor of Computer Science

EXPERTISE

HPC simulations, storage and I/O libraries, parallel file systems

Parallel and distributed data management

Software development in C, C++, with MPI, for extreme-scale supercomputers

AWARDS

2nd Gilles Kahn best PhD award of the French Computer Science Society and the French Academy of Science, 2015

C3I label (Certificat de Compétences en Calcul Intensif) of GENCI

2nd best PhD award of the Fondation Rennes 1, 2014

2nd prize of the ACM Student Research Competition at the ICS 2011 conference

LANGUAGES

French (native)

English (fluent)

German (limited)

SELECTED COLLABORATIONS AND PROJECTS

Joint Laboratory for Extreme–Scale Computing

Since 2009, I am an active participant of the Joint Laboratory for Extreme-Scale Computing (JLESC), which gathers researchers from Inria, ANL, UIUC, BSC, JSC, and RIKEN AICS. I participated to every biannual workshops since 2011. I was actively involved in building several collaboration projects between the KerData team (Inria), ANL and UIUC: FACCTS (France and Chicago Collaborating in the Sciences), PUF (Partner University Fund), and the Data@Exascale associated team. These collaborations allowed me to complete a 3-month internship at UIUC in 2009, a 1-month visit to UIUC and ANL in 2012, and a 3-month internship at ANL in 2013. Most of my research publications were done in the context of these collaborations.

Damaris — *middleware for efficient I/O and in situ visualization*

During my Master's and PhD, I developed Damaris, a middleware for efficient I/O and in situ visualization for HPC simulations. Damaris proposes to dedicate resources (single cores per multicore nodes, or entire nodes) to data management tasks. It consists of about 20000 lines of C++ and can be used in any C, C++ and Fortran-based HPC application. It provides a direct connection to the VisIt visualization software through the LibSim interface. Damaris was evaluated on some of the most powerful supercomputers in the world, including Blue Waters (NCSA), Titan (ORNL) and Kraken (NICS). In 2015, the VisIt developers integrated it in their package as an alternative to LibSim, starting from version 2.10 of VisIt. In 2015 also, Inria granted funding to the KerData research team to hire an engineer for 2 years dedicated to maintaining and supporting Damaris, and developing it further.

RESEARCH IMPACT

Publications

(publication list available at http://dorier.github.io and on Google Scholar)

I have published papers in top (CORE rank A) conferences including Cluster, IPDPS and SC. I have also published in recognized journals, including IEEE TPDS, Elsevier FGCS, and ACM ToPC.

Responsibilities

PC member of ACM/IEEE CCgrid 2016, EuroPar 2016, IEEE BigData 2016, ISC 2017, IEEE Cluster 2017, IEEE/ACM Supercomputing 2017.

Panelist at IEEE LDAV 2016. Session chair at IEEE Cluster 2016.

I regularly review papers for the widely referenced conferences and journals of my area: IEEE/ACM SC 2012; EuroPar 2011, 2014 and 2015; IEEE Cluster 2013 and 2015; ACM HPDC 2015; ACM/IEEE CCgrid 2013 and 2015; IEEE ICPADS 2015; ICPP 2013; ISPDC 2014; NPC 2015; IEEE HPCC 2011; IEEE BigData 2013 and 2016; Elsevier PARCO; IEEE TPDS.

SKILLS

Languages: main expertise in C, C++, and Ruby scripting. Also use Fortran and Python

Libraries: MPI, Boost, I/O libraries (HDF5, PnetCDF)

Software: ParaView (including Catalyst), VisIt (including LibSim), OrangeFS

Development tools: CMake, CTest, CppUnit, GDB, Valgrind, GIT, SVN

Platforms: I experimented on Blue Waters (supercomputer at NCSA), Titan (ORNL), Kraken (NICS), Intrepid (ANL), Grid'5000 (French grid testbed)

OPEN-SOURCE SOFTWARE

Links at http://dorier.github.io

Damaris
HPC I/O and data management library for large-scale simulations

Darshan-Ruby
Library for reading and processing Darshan (ANL) traces in an OO manner with Ruby

TEACHING AND SUPERVISION

During my Master's and PhD, I taught more than 250 hours at Lycée Chateaubriand (Rennes), the ENS Rennes and the INSA Rennes engineering schools, for Bachelor and Master students. Courses included programming in OCaml, C, C++, Ruby, and Java, object oriented programming, and FPGA programming.

During my PhD and postdoc, I supervised 1 bachelor student, 4 master's students and 2 PhD students for their internships.

HOBBIES

Playing piano, guitar, and banjo

Biking, hiking, and kayaking