CAF at Scale: Magnetic Fusion

BERKELEY LAB
Bringing Science Solutions to the World

Multithreaded Global Address Space Communication Techniques for Gyrokinetic Fusion Applications on Ultra-Scale Platforms

Robert Preissl
Lawrence Berkeley National Laboratory
Berkeley, CA, USA 94720
rpreissl@lbl.gov

Nathan Wichmann
CRAY Inc.
St. Paul, MN, USA, 55101
wichmann@cray.com

Bill Long
CRAY Inc.
St. Paul, MN, USA, 55101
longb@cray.com

John Shalf
Lawrence Berkeley National Laboratory
Berkeley, CA, USA 94720
jshalf@lbl.gov

Stephane Ethier
Princeton Plasma Physics Laboratory
Princeton, NJ, USA, 08543
ethier@pppl.gov

Alice Koniges
Lawrence Berkeley National Laboratory
Berkeley, CA, USA 94720
aekoniges@lbl.gov

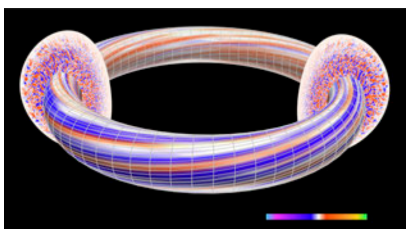


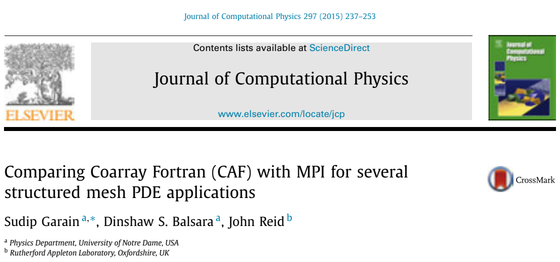
Figure 2: GTS field-line following grid & toroidal domain decomposition. Colors represent isocontours of the quasi-two-dimensional electrostatic potential

Preissl, R., Wichmann, N., Long, B., Shalf, J., Ethier, S., & Koniges, A. (2011, November). Multithreaded global address space communication techniques for gyrokinetic fusion applications on ultra-scale platforms. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis* (pp. 1-11).

- Application focus:
  - — The shift phase of charged particles in a tokamak simulation code
- Programming models studied:
  - — CAF + OpenMP or
  - — Two-sided MPI + OpenMP
- Highlights:
  - — Experiments on up to 130,560 processors
  - — 58% speed-up of the CAF implementation over the best multithreaded MPI shifter algorithm on largest scale
  - — “the complexity required to implement … MPI-2 one-sided, in addition to several other semantic limitations, is prohibitive.”

CAF at Scale: CFD, FFTs, Multigrid

BERKELEY LAB
Bringing Science Solutions to the World


Journal of Computational Physics 297 (2015) 237–253

Contents lists available at ScienceDirect

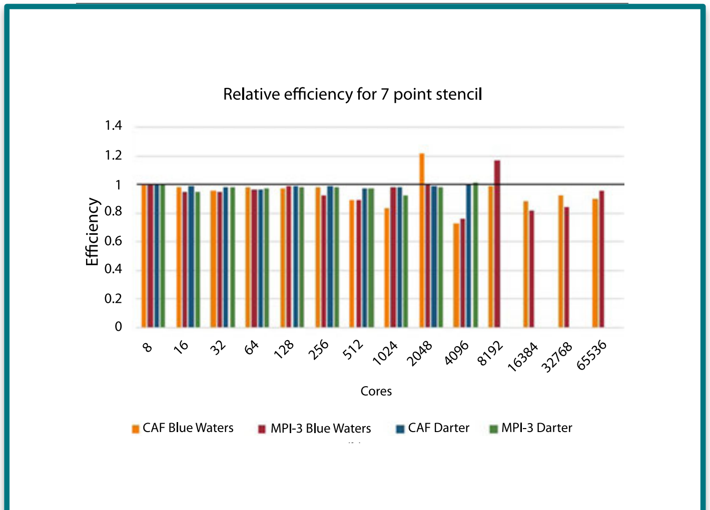
Journal of Computational Physics

www.elsevier.com/locate/jcp

Comparing Coarray Fortran (CAF) with MPI for several structured mesh PDE applications

Sudip Garain [a,*], Dinshaw S. Balsara [a], John Reid [b]

[a] Physics Department, University of Notre Dame, USA
[b] Rutherford Appleton Laboratory, Oxfordshire, UK



- Applications studied:
  - — Magnetohydrodynamics (MHD)
  - — 3D Fast Fourier Transforms (FFTs) used in infinite-order accurate spectral methods
  - — Multigrid methods with point-wise smoothers requiring fine-grained messaging
- Programming models studied:
  - — CAF or
  - — One-sided MPI-3
- Highlights:
  - — Simulations on up to 65,536 cores
  - — "… CAF either draws level with MPI-3 or shows a slight advantage over MPI-3."
  - — "CAF and MPI-3 are shown to provide substantial advantages over MPI-2.
  - — "CAF code is of course much easier to write and maintain…"

Garain, S., Balsara, D. S., & Reid, J. (2015). Comparing Coarray Fortran (CAF) with MPI for several structured mesh PDE applications. *Journal of Computational Physics*, *297*, 237-253.

CAF at Scale: Weather

Article

International Journal of HIGH PERFORMANCE COMPUTING APPLICATIONS

The International Journal of High Performance Computing Applications
2015, Vol. 29(3) 261–273

DOI: 10.1177/1094342015576773
hpc.sagepub.com
SAGE

A Partitioned Global Address Space implementation of the European Centre for Medium Range Weather Forecasts Integrated Forecasting System

George Mozdzynski, Mats Hamrud and Nils Wedi

Figure 7. EQ_REGIONS partitioning of grid-point space, showing a partition at the poles and then an increasing number of partitions as we approach the equator.

- Application:
  - European Centre for Medium Range Weather Forecasts (ECMWF) operational weather forecast model
- Programming models studied:
  - CAF or
  - Two-sided MPI
- Highlights:
  - Simulations on > 60K cores
  - performance improvement from switching to CAF peaks at 21% around 40K cores

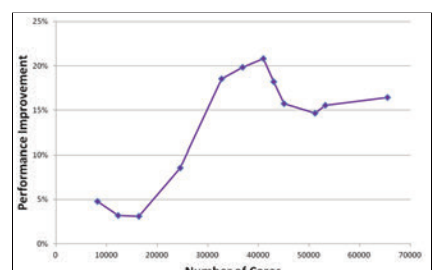
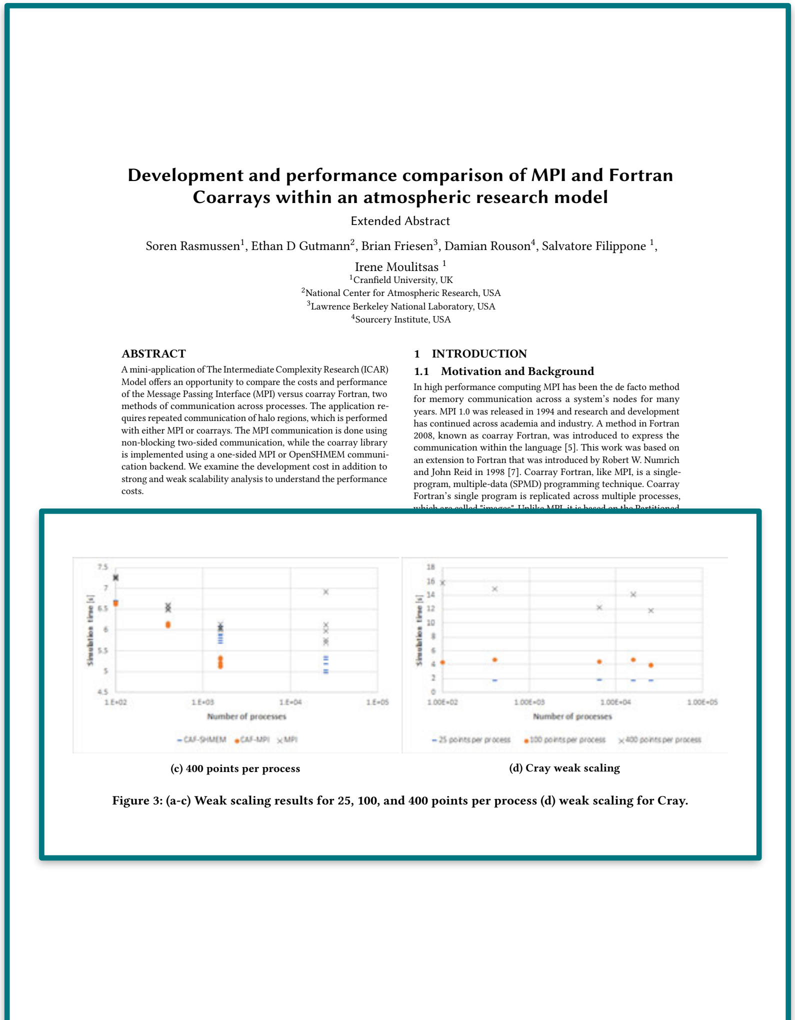


Figure 14. Performance improvement of the T2047 (~10 km) model with 137 levels by using Fortran2008 coarrays on HECToR (Cray XE6).

Mozdzynski, G., Hamrud, M., & Wedi, N. (2015). A partitioned global address space implementation of the European centre for medium range weather forecasts integrated forecasting system. *The International Journal of High Performance Computing Applications*, *29*(3), 261-273.

CAF at Scale: Climate



Development and performance comparison of MPI and Fortran Coarrays within an atmospheric research model

Extended Abstract

Soren Rasmussen[1], Ethan D Gutmann[2], Brian Friesen[3], Damian Rouson[4], Salvatore Filippone [1], Irene Moulitsas [1]
[1]Cranfield University, UK
[2]National Center for Atmospheric Research, USA
[3]Lawrence Berkeley National Laboratory, USA
[4]Sourcery Institute, USA

ABSTRACT

A mini-application of The Intermediate Complexity Research (ICAR) Model offers an opportunity to compare the costs and performance of the Message Passing Interface (MPI) versus coarray Fortran, two methods of communication across processes. The application requires repeated communication of halo regions, which is performed with either MPI or coarrays. The MPI communication is done using non-blocking two-sided communication, while the coarray library is implemented using a one-sided MPI or OpenSHMEM communication backend. We examine the development cost in addition to strong and weak scalability analysis to understand the performance costs.

1 INTRODUCTION

1.1 Motivation and Background

In high performance computing MPI has been the de facto method for memory communication across a system's nodes for many years. MPI 1.0 was released in 1994 and research and development has continued across academia and industry. A method in Fortran 2008, known as coarray Fortran, was introduced to express the communication within the language [5]. This work was based on an extension to Fortran that was introduced by Robert W. Numrich and John Reid in 1998 [7]. Coarray Fortran, like MPI, is a single-program, multiple-data (SPMD) programming technique. Coarray Fortran's single program is replicated across multiple processes,

(c) 400 points per process

(d) Cray weak scaling

Figure 3: (a-c) Weak scaling results for 25, 100, and 400 points per process (d) weak scaling for Cray.

Application:

- Intermediate Complexity Atmospheric Research (ICAR) model
- Regional impacts of global climate change

Programming models studied:

- CAF over one-sided MPI
- CAF over OpenSHMEM
- Two-sided MPI
- Cray CAF

Highlights:

- "… we used up to 25,600 processes and found that at every data point OpenSHMEM was outperforming MPI."
- "The coarray Fortran with MPI backend stopped being usable as we went over 2,000 processes… the initialization time started to increase exponentially."

Rasmussen, S., Gutmann, E. D., Friesen, B., Rouson, D., Filippone, S., & Moulitsas, I. (2018). Development and performance comparison of MPI and Fortran Coarrays within an atmospheric research model. *Parallel Applications Workshop - Alternatives to MPI+x (PAW-ATM)*, Dallas, Texas, USA.