Performance and Introspection at Exascale

1 For Form

- **Title:** Performance and Introspection at Exascale
- **Acronym:** PIE
- **Duration:** April 1st 2011, for three years
- **Keywords:** Introspection, automation, performance, component framework, Cactus, IPM, exascale applications, scalability

2 Summary (300 words)

At a time when simulation supported science is of increasing importance across many socially significant areas of research, the accessibility and utility of the ultrascale HPC systems that drive simulation science are in danger of eroding. The usefulness of petascale resources is already limited to science teams who can both deal with the complexities of expressing their science goals in HPC form and deal with complex performance optimization scenarios which vary across architectures and across runs. As we move toward exascale, without research and development effort applied to these two issues the utility of exascale systems will be of increasingly limited scientific and social relevance. We propose a program of research that will make it easier for scientists to express their problems in programmatic form suitable for HPC and also to make performance optimization more accessible and productive. These two efforts respectively form the HPC entry points and outputs that are growing steadily more complex. Below we describe a program of research to bring automatic code generation and scalable performance introspection to exascale.

3 Introduction

In the last decades, the emergence of simulation-based science and engineering has become dominant in many fields. One cannot imagine designing an aircraft, forecasting a hurricane storm surge, nor predicting the gravitational waveform produced when two black holes collide, without it. Already in such examples, we see a very diverse set of activities: discipline specific knowledge is critical, but one needs also to create mathematical descriptions of the phenomena, consider boundary conditions, develop appropriate numerical schemes to solve them, and so on. In doing so, one needs also to develop highly complex software environments from the computer system level to the application, and also understand the complex computer architectures in which they execute. One must also make sense of the output, requiring new and complex analysis and visualization tools. One also needs to bring together experts from each of these domains and enable them to work together to solve such complex problems, bringing in new challenges from software engineering to collaborative environments. These challenges are apparent in global scale issues, the target of the G8 solicitation where problems such as climate modeling, hurricane storm surge predictions, understanding nuclear fusion, predicting the flow of disease in global epidemics, require large collaborations, complex multiscale and multiphysics models, and efficient use of exascale machines including visualization, data management and analysis.

Scientific communities are currently developing their codes, and upgrading their tools, for *petascale* computing, for example in the NSF Blue Waters consortium in the USA, ■ [ADD EXAMPLES FROM EU sites and NL sites]. Already, it is clear that current paradigms and tools for simulation software development, deployment, analysis and visualization are severely strained. Although progress continues to be made to scale simulation codes to larger machines using current technologies such as hybrid MPI/OpenMP, scaling is only one of the issues faced. A core problem is that as the underlying architectures and machine size become more complex and grow, so do the application codes providing a quickly moving target that will be difficult to reach.

These complex codes are increasing hard to debug and verify, are increasingly specialized to particular architectures, and maybe of most concern, the only research groups able to take part at this level are long-lived teams of experts who have already assembled the critical mass of skills in large scale computing. Exascale machines, which may be deployed as early as XXXX will require scaling these tools by a further three orders of magnitude, and it is clear that our current model of software development will not best serve the scientific community using these machines.

4 Background

4.1 Cactus Framework

Cactus is an open-source framework designed for the collaborative development of large scale applications. Computational toolkits distributed with Cactus already provide a broad range of capabilities for solving initial value problems in a parallel environment. Cactus is used by applications in areas including relativistic astrophysics, computational fluid dynamics, reservoir simulations, quantum gravity, coastal science and computer science. Cactus has been used by over three dozen groups producing over 300 research publications and 50 student theses, and over 42,000,000 SUs of time are allocated to Cactus applications deployed by seven research groups running on the NSF TeraGrid. Two of the sixteen NSF PRAC awards for the Blue Waters petascale facility to be deployed in 2011 are for Cactus applications. In computer science, Cactus has pioneered the use of new technologies such as grid computing for scientific applications and is regularly used as a benchmark application and an application driver for research projects in computer science.

The Cactus group at LSU[[i],[ii]] is engaged in several projects that are working with application groups to address petascale challenges for scientific computing. The NSF PIF XiRel and CIGR projects (PI G. Allen)[[iii],[iv],[v]] are creating the basis for a modern, scalable cyberinfrastructure, providing enabling software for highly scalable codes able to leverage today's and tomorrow's petascale machines, preparing for next generation architectures and computing paradigms. Although XiRel and CIGR are focused in the computational relativity computing (for example generating a new community toolkit, see www.einsteintoolkit.org) the improvements in scaling, optimization and data management will advance all Cactus applications. The NSF SDCI Alpaca project (PI E. Schnetter)[[vi],[vii]] develops high-level tools to ensure correctness and improve performance of Petascale applications.

The Cactus team is also involved in the NSF Blue Waters project (LSU PI G. Allen)[[viii]], which will be the most powerful supercomputer in the world for open scientific research when it comes online in 2011. Cactus is part of the software development effort for Blue Waters, and in particular is working with NCSA to develop development interfaces and paradigms for large scale simulation development. Through the LSU HPCOPS award, the Cactus group are developing a Science Gateway which will include capabilities for simulation assembly, compilation and deployment across the NSF TeraGrid resources. Through the NSF EPSCOR CyberTools project, the Cactus team is involved in educational initiatives including summer REU programs, K-12 education and graduate mentoring.

5 Joint Research Plan

Our research plan will provide a new paradigm for the development of robust, scalable simulation codes at the petascale and exascale level that will support a new generation of scientists using computation for scientific inquiry as well as traditional computational scientists. Through the enhancement of an already well-used component framework for the development of large scale simulation codes this work will have broad impact across many scientific disciplines, including astrophysics, coastal science, and computational fluid dynamics. To provide focus to the research and development, and leverage an already strong collaboration of scientists experienced in the challenges of large scale computation, we concentrate on an application of great societal impact — understanding the origins of the universe — which includes scientific challenges such as nuclear fusion and radiation transport that are important for a range of other important applications.

Our team brings all the necessary skills and experience to deliver the above high level goal. Below we describe the specific work and in the next section we describ how the expertise of consortium member is leveraged to achieve those goals.

5.1 Research Plan

- *Robust Problem Formulation:* Develop procedures to translate governing equations into a robust formalism for further, automated processing and implementation as simulation framework components.
- *Automated Code Generation:* Generate Cactus components from equations, enable automated optimisations, allow tuning for different hardware architectures and different parallel paradigms (MPI, OpenMP, hybrid, UPC, accelerators, etc.)
- *Science Introspection:* Add intelligence to component interfaces at the scientific level (interfaces between models, robustness of user parameters).
- *Performance Introspection:* Integrate run time performance information for highly concurrent, multicomponent simulations into the framework to support dynamic decision making, dynamic optimisations, automated application level performance problem detection and reporting.
- *Metadata and Workflow:* Support complete simulation pipelines including data import, pre- and post-processing, analysis, visualisation with the above model, tracking sufficient metadata to ensure reproducible results.

6 The Consortium

To address this solicitation, we have assembled a uniquely experienced team of researchers from across the G8 countries who have proven skills in large scale scientific computing and have well established projects including the collaborative development of core software essential for this project.

In physics, LSU, Southampton University, Albert Einstein Institute and collaborating site UIB already collaborate on scientific studies in computational relativity and are core partners in the community Einstein Toolkit. In computer science, LSU, Munich, and collaborating sites LBL and BSC provide the experience and research in performance modeling and tool development. Collaborating partner ANL will provide new collaboration tools levering Web 2.0 technologies that will enhance the integration of the project.

6.1 Louisiana State University (Gabrielle Allen, Erik Schnetter)

Located at the Center for Computation & Technology (CCT) at Louisiana State University, the Computational Frameworks and Numerical Relativity groups work closely together on developing both the underlying technologies and complex physical models for large scale simulations of real world processes such as relativistic astrophysics. The groups together form the core support and development for the Cactus Computational Toolkit, and lead the development of community domain-specific toolkits for astrophysics (the Einstein Toolkit, http://www.einsteintoolkit.org), and coastal modeling and computational fluid dynamics (the CFD Toolkit).

The team is funded through awards primarily from NSF, including a PetaApps award to develop algorithms for modeling Gamma-Ray Bursts at large scale and a physics at the information frontier award to develop the open, community Einstein Toolkit. The group is part of the Blue Waters project, contributing both to the software development through Cactus, and through a PRAC award to deploy astrophysics simulations on the petascale system when it is in operation in 2011.

The CCT team is deeply involved in the education of next generation researchers in computational science. Allen leads the CyberTools project in Louisiana and is co-PI for a three year NSF REU in computational science which this summer will train 16 undergraduates in research projects at CCT.

LSU already has a long standing collaboration with the sites and investigators in Germany, United Kingdom and Spain, reaching back to the EU funded Astrophysics Network (www.eu-network.org) that joined together ■ [XXXX] countries in developing both computational models (Whisky) and infrastructure (Cactus). LSU also has a close connection to LBL who were involved, via John Shalf, in contributing to the development of the I/O layers of Cactus and in large scale demonstrations that led e.g. to together winning the Bandwidth Challenge at Supercomputing in ■ [XXXX].

LSU's role in the project will be

- Act as the Lead site for the project
- Provide the infrastructure development needed in Cactus
- Provide the broad science cases: coastal modeling, reservoir simulations
- Contribute towards the relativistic astrophysics use case
- Provide educational opportunities related to the project through programs at CCT such as the REU in Computational Science.

6.2 Lawrence Berkeley Laboratory (David Skinner, Karl Fuerlinger)

A core component of our research program is connecting appplication level and performance engineering spaces. This need is aligned with work already underway at Lawrence Berkeley Lab (LBL) where the monitoring and analysis of production HPC workflows is an important concern. We propose to deliver performance introspection to applications through a direct extension techniques in wide use today by a diverse variety of HPC codes. The Integrated Performance Monitoring (IPM) framework in use at DOE, NSF, and commerical computing cemters provides a low-overhead production-ready interface for job and task level performance analysis. AT NERSC alone 300,000 application performance profiles have been generated in the last six years. This profiling effort has led to great improvements in the understanding of how application and architecture meet to deliver scientific progress. Currently users, staff, and all stakeholders in the performance of scientific applciations are able to regularly and easily assess tasks, jobs, and workloads.

We propose an extension to that approach which publishes performance profiles not just to the stakeholders listed above but also to the application itself, in the present case to the CACTUS framework. . Ex post facto performance analysis has sufficed for many decades but must give way to truly runtime analysis when concurrencies and costs are driven suffceintly by Moore's Law. Unless we move towards such performacne introspection at runtime, and architect them sensibly into application frameworks, performance becomes more mysterious and eventually limits researchers access to HPC resources as they have a limited number of "shots" to get the algorithm, settings, and domain decomposition correct.

Through the currently proposed research program, LBL will be a source of understanding of HPC application's runtime needs. We will build an interrface for communicating performance information on an as-needed basis to the CACTUS runtime. Instead of making performance engineering decisions on a after-the-fact or per job basis, our approach offers the ability to hust-in-time decisions about algorithm and library level strategies. This perspective can provide great value to end users in that the are not required to wager about performance choices prior to issuing their workflows.

Synergistic Funding: ■ [SDCI,BACATEC]

6.3 University of Southampton (Ian Hawke)

The General Relativity group at the University of Southampton is the largest classical relativistic astrophysics group in Europe with a focus on computing gravitational wave signatures. As part of this effort the group has been involved in long term collaborations with the AEI and LSU groups, particularly around the Whisky code for relativistic hydrodynamics.

The relevant work of the group is on high accuracy algorithms for complex multiphysics problems, such as the interaction between fluid and elastic matter interfaces coupled to radiation transport in binary neutron star mergers. To better leverage cutting-edge approaches from outside numerical relativity the group is closely involved with the Computational Modelling Group (cmg.soton.ac.uk) and Iridis3 computing resources here at Southampton.

Southampton's role in the project will be

- Provide multiphysics coupling mechanism for Cactus
- Provide high accuracy algorithms for Whisky
- Outreach/dissemination through CMG?

Synergistic Funding: ■ [XXXX]

6.4 Albert Einstein Institute (Ian Hinder, Luciano Rezzolla)

The Numerical Relativity group at the Albert Einstein Institute focuses on the simulation of astrophysical compact object coalescences in Einstein's full theory of General Relativity and is among the world-leaders in this field, having produced over 100 publications over the last 6 years. We use the `Cactus Computational Toolkit` and `Carpet` mesh-refinement driver extensively to simulate the inspiral and merger of black holes and neutron stars, expected to be among the most promising candidates for the direct detection of gravitational waves by current and planned experiments. These simulations are necessary to accurately predict the astrophysical waveform; a necessary ingredient in currently used search techniques. We use supercomputing resources throughout Germany, as well as the TeraGrid through our partership with LSU.

We also develop the `Kranc` automated code generation package which has been used to construct the `McLachlan` Einstein evolution code, which has been made freely available as part of the Einstein Toolkit. Our role in the project will be to enhance `Kranc` to generate codes which support detailed performance introspection and feedback to enable dynamical load balancing on large numbers of cores. *What about making Kranc support hydro? Is that too much? no, Kranc hydro is fine, but it may be too specific.*

- Automated code generation
- Automatic optimisations for efficient implementations on new architectures
- (Something about automatically generating component interfaces, or automatically deriving implementations from such component descriptions)

Synergistic Funding: ■ [XXXX]

6.5 University of Munich (Dieter Kranzlmueller)

6.6 University of the Balearic Islands (Carles Bona, Sascha Husa, Denis Pollney, Toni Arbona, Joan Masso)

The IAC3 is a research institute of the Balearic Islands University. Its focus is on simulation science, particularly addressing socially relevant aspects. An example is its participation in the 30M Euro Spanish project cvREMOD, which intends to significantly cut national and European-level health-care costs through the implementation of preventive medicine and accurate evaluation of the need and benefit of intervention. The technical core of such strategy is based on patient specific physiological simulation.

The complexity of models that must be addressed and combined in such scenarios is daunting (multi-scale models at organ, tissue, cellular, and biochemical levels), and certainly out of the scope of traditional techniques for writing high performance simulation software. The IAC3 is developing a model management tool (Simflowny) based on a formal representation of physical models based on either hyperbolic or parabolic PDE (for instance, Maxwell's equations, Navier-Stokes, MHD, heat transfer...), physical problems, and even discretization techniques. This formal representation is actually independent of Simflowny, and may be used as a specification- as it is or improved - in other software environments. Simflowny is essentially a management tool for such models and problems, and it also generates code automatically for Cactus. The role of the IAC3 in the project will be to adapt the formal representations to the requirements of petascale computing in cooperation with the rest of partners, and tune the code generation to meet the objectives of the project in terms of performance.

7 Points

- Highlighting and mentoring of young researchers.
- Long standing collaboration on physics and infrastructure (EU Network and beyond), LBL connection via Shalf and scaling work, bandwidth challenges with LBL.
- Two week long working meetings each year?