CCA and SciDAC2

David Bernholdt, *CCA2 Lead PI*
with
Rob Armstrong, Randy Bramley, Gary
Kumfert, Lois McInnes, Jarek Nieplocha
CCA2 Elves

Background: CCA and SciDAC1

- CCA Forum started in 1998
  - Outgrowth of DOE2000 initiative

- Created an unsolicited proposal in 2000

- SciDAC1 was announced in 2000, and original proposal was revised to fit new program as CCTTSS

- CCTTSS focused on understanding requirements, defining, and implementing a component architecture specifically designed for HPC scientific computing
  - i.e. risky and research

The SciDAC2 Environment

- SciDAC1 is viewed as extremely successful in advancing computational science

- More emphasis on distinction between base research program and SciDAC (more applied/developmental research) than we saw in SciDAC1

- After five years of work, continuing ISICs are expected to have (significantly) more impact on apps than in SciDAC1

- Therefore, ISIC proposals must be strongly applications-driven
  - "Research" should be motivated by application needs
  - Strong ties to, and clear impact on specific applications will be expected

The SciDAC2 Environment (2)

- Program managers (PMs) are trying to grow the SciDAC budget
- But even if they are successful, we expect more *numerous* ISICs rather than more $ per ISIC
  - ISICs > $3M/yr will be a stretch (CCTTSS is ~$3M/yr)
- A major risk is that PMs will reduce or eliminate funding for existing ISICs in order to add new ones
  - There is a strong desire among many associated with SciDAC to see new visualization and/or data analysis ISICs
- Our competition is *not* that another group will come up with a better proposal for component technology, but that another group will persuade decision makers that their technology is more important to applications than components and take away our $
  - CCA2 must clearly illustrate past & future benefits to applications
  - Show that users are demanding CCA technologies

Advice from Fred

- "Research" is not a "dirty word", but…
  - Applied/developmental research clearly connected to applications fits better in SciDAC context

- All research ideas (we think should be in the proposal) should be included in pre-RFP whitepaper, and Fred will say if he considers any too much "basic research" for the SciDAC proposal. These would be better candidates for separate base-program proposals

- Produce the *right* proposal, with the *right* team, and the *right* budget. If that budget can't be met, negotiations will ensue
  - Caveat: anything > $3M/yr is unlikely

- He will interact as much as possible *before* the RFP is issued. After that he is constrained in dealing with us.

Thoughts on Applications

- ITER (fusion) is #1 priority for Office of Science

- Climate is an ongoing priority

- Accelerator design, astrophysics have been viewed as very successful, also some quantum chemistry projects

- Biology is a new addition to SciDAC

Thoughts on CCA Research

- SciDAC1 has been a period of unprecedented stability in DOE HPC platforms. This is changing. How can CCA help users achieve portability?

- Users need a spectrum of solutions. In IDEs, for example, Eclipse is very popular, but is a heavyweight thing to get started with. There should be a spectrum of approaches so that users can choose trade-offs. Similarly for CCA. Is there a "CCA lite" concept that would simplify adoption by more users?

CCA2 Research Challenges

- Using the component environment to enhance and expand the capabilities of/available to scientific applications
  - CQoS, PCI, other initiatives

- Supporting a growing user base
  - CCA tools infrastructure -- maintenance, porting, basic improvements
  - Tutorials, documentation (everyone's responsibility)
  - User outreach and application support
  - Users must be able to bet their applications on CCA being available forever
    - Critical mass for open source project immortality?
    - Injecting CCA ideas into OMG CORBA (or other) immortal environment?

4

CCA2 Research Challenges (2)

- Building the component ecosystem
  - Component toolkit
  - Interoperability w/ other HPC frameworks Component "automation"

- Usability of the CCA environment
  - Build environments
  - Component Repository/Component Deployment Methodology
  - Comfortable look & feel for scientific community
  - A comprehensive "CCA lite" approach to provide users with a spectrum of cost/functionality trade-offs

CCA2 Project Structure

Four focus areas:
- similar to CCTTSS proposal

- User and applications outreach and support

- Component technology initiatives

- CCA environment and tools (incl. infrastructure, usability)

- Component toolkit

- Tight linkages between focus areas
- Think of as a 4d matrix with entries connecting many activities
  - Various applications will motivate different technology initiatives, tool development, and toolkit additions
  - Technology initiatives will also motivate tool development
  - …

5

Applications Hierarchy

1. SciDAC applications
2. Non-SciDAC DOE applications

   Primary motivators of CCA2 R&D should be from these two categories

3. Non-DOE applications

   These should be cast as *beneficiaries* of work motivated by apps in categories 1 & 2, Too much effort into non-DOE apps likely to raise questions

---

Target Applications (SciDAC)

- Fusion (especially new FSP Prototypes)
  – Center for Simulation of Wave Interaction with Magnetohydrodynamics (CSWIM), Randy Bramley, CS lead
  – Center for Plasma Edge Simulation (CPES), Scott Klasky, CS lead
- Climate
  – ESMF dominates the agenda for climate community. What is best role for CCA in this context?
  – Rob Armstrong & Jay Larson, POCs
- Combustion (CFRFS)
  – Jaideep Ray, POC
- Chemistry
  – Internal to CCTTSS in SciDAC1
  – Anticipate major call for large, multi-institution "center" to develop and deploy next-generation HPC chemistry tools (i.e. "endstation"). FY07 $ but call may be on similar schedule to SciDAC2
  – Theresa Windus, POC

SciDAC Application Areas w/o Established Collaborations

- Astrophysics
  - Strong intellectual influence (Doug Swesty @ SciDAC2005)
  - Haven't found a compelling need for actual CCA tools

- Accelerator Modeling

- Biology
  - New addition in SciDAC2
  - Not sure who the players are yet, nor how the RFP will be structured
  - Very different from other HPC communities
  - Haven't been proselytized for five years already about CCA
  - Jarek Nieplocha discussing with PNNL biologists
  - Need others to make contacts too!

ISIC Collaborations

SciDAC1 was viewed as having too little collaboration among ISICs

- Scalable Systems successor
  - Nothing yet

- PERC
  - Continued collaboration with Oregon group
  - Empirical optimization (selection among variants/CQoS)
  - Performance modeling (?)
  - Performance composition (?)
  - Performance testing frameworks (?)
  - Boyana Norris, Sameer Shende, POCs

- SDM
  - Integration of components into scientific workflow environments
  - Data analysis framework (?)
  - Steve Parker, POC

ISIC Collaborations (2)

- TOPS
  - ?
  - Lois McInnes, POC
- TSTT
  - Interface development (?)
  - ?
  - Lori Freitag Diachin, POC
- APDEC
  - ?
  - Jaideep Ray, POC
- Visualization
  - Data movement/management/processing to allow viz people to focus on viz tools for end-users (?)
  - Steve Parker, POC

Thoughts on CCA Usability

- Current experience with CCA in general is that people can use it after some effort, but that initial effort can be dauntingly high
  - High enough that some won't get over the barrier

- Users are interested in different benefits of CCA, often just one or two (not all)
  - Managing code complexity (component-ness)
  - Facilitating collaboration over software
  - Language interoperability
  - Language-neutral interface specification
  - Code coupling

We Need a Spectrum of Approaches for Usability

- Wizards & automatic code generation
  - Eclipse integration

- "CCA lite" approaches to lower build-based barriers
  - Build of CCA tools
  - Build of user code in CCA environment
  - Solutions may involve allowing users to select which CCA features they need at the moment
    - Dynamic linking
    - Language interoperability
    - Interactivity
  - But provide an evolutionary path to "CCA Complete"

- The ultimate build solution is, at best, a (very) long-term solution. We need immediate results!

Rough Schedule and Milestones

- 9 Oct   Draft outline of proposal
- 31 Oct   Draft preproposal submitted to Fred
- SC2005   Discussions with collaborators
- [Iterate at preproposal level until Fred likes]
- [Write draft of full proposal (won't be able to tune to RFP yet)]

- **15 Dec   RFP Issued by DOE**
- [Finalize preproposal and submit as early as possible]
- [Begin modifying full proposal to satisfy particulars of RFP]

- **24 Jan   Preproposal due to ORNL for internal review**
- **31 Jan   Preproposal due to DOE**

- **15 Feb   Proposal due to ORNL for "red team" review**
- **[Do other participants also have internal review requirements?]**
- [Revise proposal based on red team comments]
- **15 Mar   Proposal due to DOE**

All items in red are *estimated* dates   All items in blue depend on deadlines in red

Objectives for this Meeting

- Consider applications, research initiatives, etc. discussed to date

- Formulate a coherent and cohesive plan for an interesting CS research project with a strong impact on specific computational science applications
  - Remember the 4d matrix – everything should be extensively linked in all dimensions

- Capture this plan in a detailed outline for the proposal
  - Including who will participate in various activities

Organization of Discussions

Thursday
- Short whole-group discussion to insure everything we need to discuss is on the outline

Lunch
- Breakout sessions for discussion and writing

Dinner/Evening
- Informal discussions as desired

Friday
- Reports from discussion leaders
- Wrap-up discussion

CCA
Common Component Architecture

Breakout Sessions

	Location A	Location B
Session 1 (1:00-3:00pm)	**User and Applications Outreach and Support** *Lead: Randy* *Notes: Jaideep*	**CCA Tools & Environment** *Lead: Gary* *Notes: Boyana*
Session 2 (3:30-5:30pm)	**Technology Initiatives** *Lead: Lois* *Notes: Jim K*	**Component Toolkit** *Lead: Rob* *Notes: Manoj*

CCA
Common Component Architecture

We Need a Good Name!

- CCTTSS was not a good name
  - Center for Component Technology for Terascale Simulation Software
  - No one remembered it
  - No one used it
- To most people the CCA effort is identical to the SciDAC project, but…
- An increasing number of groups outside of the SciDAC project are contributing to the development of the CCA
- A good name might help make the distinction between the SciDAC project and the larger CCA world, collectively represented by the CCA Forum
- Wiki page established to collect suggestions
  - https://www.cca-forum.org/wiki/tiki-index.php?page=CCA2+Naming
  - Linked from top CCA2 wiki page