MxN Working Group Meeting Report!

(This is your brain…)

At The Radisson, Knoxville, TN
January 22, 2004

James “Jeeembo” Kohl, M.C.
Oak Ridge National Laboratory
2nd Biennial, Official, 100% Pure, CCA Approved…

MxN Working Group Meeting

Report!

(This is OUR brain…! 😊)

(This is our brain!)
The “Basic” Problem:
Parallel Data Exchange

“M”

“N”
Cooperating Parallel Components...?!

- CCA-specific capability (non-ad-hoc…)
- Requirement of emerging scientific apps
  ⇒ Generalized model coupling (multi-parallel)
  ⇒ Efficient daisy-chaining of numerical libraries
  ⇒ Visualization (parallel rendering / serial gathering)
- Fundamental issue ~ data decompositions
  ⇒ Different distribution for each parallel component
  ⇒ Need to describe/reconcile disparate data organizations
- And, more than just parallel data…
  …Parallel method invocations, too!
MxN Internal Structure

- High-Level MxN Run-Time Library
- MxN Data Exchange and Synchronization Interface
- Parallel Data Mapping and Communication
- Data Decomposition Specification
- Data Distribution System
- Local Data
MxN Progress to Date…

- Generalized MxN specification
  ⇒ Based on CUMULVS (ORNL) and PAWS (LANL)
- SCMD peer MxN component solution
  ⇒ Visualization; combustion and climate apps…
- Parallel Remote Method Invocation (PRMI)
  ⇒ Preliminary semantics and specifications
- Distributed MxN solution
  ⇒ MPI I/O approach
MxN “Explicit” Component Solution

• Port-based direct invocation of MxN methods
  ⇒ Most general solution, but…
  ⇒ Most challenging to the end-user scientist.
    • “Assembly language” level interface…
    • Preliminary platform for experimentation, higher-level functs

• Several demonstrations (SC00, SC01, SC02)
  ⇒ Increasingly elaborate visualization solutions
    • Incorporated generalized data description since SC01 (DADF)
  ⇒ Two main prototypes (reused several apps):
    • CumulvsMxN ~ first inter-framework component capabilities!
    • PawsMxN ~ ping-pong coupling experiments
MxN Meeting Topics

- Overview of MxN Specification ~ Jeeembo
- Data Reorganization Interface (DRI) ~ David
- Distributed Data Broker (DDB) ~ Dan
- Distributed CCA Framework (DCA) ~ Felipe
- ECho and MxN ~ Hasan
- Open Discussion
Data Reorganization Interface (DRI)

• DARPA Effort for Distributed Array Support
  ⇒ No academic involvement (as opposed to CCA? 😊)
  ⇒ Related to (absorbed by) HPSI effort…

• Simple series of Puts/Gets for redistribution
  ⇒ Through user-managed buffers
  ⇒ Uses MPI as basis for the communication design

• Evaluation: Different domain (real-time)
  ⇒ But we can steal some of their ideas…!
Distributed Data Broker (DDB)

• Designed for Earth System Model  
  ⇒ Coupling 2 Atmosphere and 2 Ocean models

• Works with MPI or Multi-MPI w/PVM  
  ⇒ One large integrated parallel application (Fortran)

• DDB library hierarchy on each processor  
  ⇒ Handles generic rectangular data decompositions  
  ⇒ Registration phase, then SendData / GetData…

• Now Implemented as a CCA/Babel Component!  
  ⇒ Nice Interface, Gruesome Internals  
  ⇒ Steal ideas, simplicity of interface (for simple data)
Distributed CCA Framework (DCA)

• MPI-Based MxN-Capable Distributed FW
  ⇒ Most Parallel Cases Don’t Map to Distributed!

• 3 Levels of Parallelism:
  ⇒ “Component” ~ Distributed Computing
  ⇒ “Process” ~ SPMD/Parallel Computing
  ⇒ “Thread” ~ Node-localized Thread Computing

• MPI Communicator Groups for Collective Calls
  ⇒ Caller dictates data layout at provider…
    • Point of controversy, how to specify provider decomp…?
DCA Decomposition Issues

Static (SIDL) versus Dynamic (Run-Time) Meta-Data?
**ECho and MxN**

- Dynamic Publish/Subscribe (Event) Environment
  - Source-side filtering for data for redistribution
    - A la “one-sided” communication, but not quite… 😊
    - ECL (C-subset) dynamic code generation
    - PBIO efficient binary data format
    - PDS Proactive Directory Service for discovery
  - Simple application demo of SmartPointers for Viz
  - Future work ~ use XML for describing data distributions and custom filter generation…
MxN Action Items / Tasks:

- Parallel MxN Specification Changes:
  ⇒ Eliminate the requestTransfer() suite of methods…
    • Seemed like a good idea at the time… but no.
  ⇒ Add simplified interface alternatives
    • One-stop-shopping method, for all-in-one MxN-ing…

- Create Combined Distributed/Parallel FW Demo
  ⇒ Co-located MxN and Application Components…
    • Allow data wrangling in shared address space (cheating!)
  ⇒ Combine multiple FWs as Distributed Components
    • Theoretically possible, need practical proof-of-concept.
MxN Action Items / Tasks (cont.)

- TSTT Mesh Integration with DADF
  ⇒ Use “General Map” Interface…
- Distributed MxN and CCA Issues
  ⇒ Continue Education and Exploration of Issues
- Parallel Rendering / Viz using MxN
  ⇒ Not just Mx1, but real MxN*…
MxN Action Items / Tasks (cont.)

• Killer Integration Application Demo!
  ⇒ TSTT + TOPS + CCA/Babel + DADF + MxN
  • Demonstrate full integration capabilities… (if any!)
  ⇒ Choose a simple contrived or synthetic appl…
  • Else perhaps a real application from Climate/Fusion:
    → Jay’s “Big Lake” dream code…
    → Ongoing Fusion integration opportunities…
MxN Summary

• Stable specification and component solutions
  ⇒ Update and Streamline, Provide Simple Alternatives
• Distributed/Parallel Framework MxN experiments
• TSTT and DADF Integration
• Continue Distributed MxN and CCA Front
• Parallel Rendering / Viz Demo
• Killer Integration Application Demo!
  ⇒ “Can’t we all get along…?”

http://www.csm.ornl.gov/cca/mxn/