Appendix A: Presentation Slides for “Beyond Mass Storage: How Do We Manage the Data of the Future?”
How do We Manage the Data of the Future?

Ideas for an International Collaboration

Ann Kerr

Chair, International Symposium
Global Data Interoperability - Challenges and Technologies

IEEE MSSTC
Problem/Current Situation

- Many Communities
- Many Problem Domains
- Many Solutions
- Many “Stovepipes”

Common Need is Secure Access to Global Data
Community Owned & Accessed Data

- Cost to collect and manage data high and growing
- Too much data for single PI or small research group to utilize effectively
- Need cross-discipline data sharing for many problems
- Modeling efforts need access to sensor and experimental data for validation and
- Modeling and Analysis drives insights into what data to collect
Many Communities Working on Parts of the Problem

- Application-specific ad hoc solutions with limited interaction
- Overlapping functionality
- Friendly Chaos
- Communications and Collaboration needs to be developed between and among Communities
What is Needed?

Communities to Start Dialogue

- Build A Common Vision and,
- Create A Global Architecture,
- Adopt Standards and Tools to effectively utilize and manage community-owned data
Anticipated Results

- Methodology/Technologies to facilitate ‘linkages’ among and between communities
- A ‘Community of Communities’ Solution
- Development of a ‘Common Software Bus’
- From ‘many to many’ to ‘many to One’
Possible Strategy: an International Collaborative Effort

that would...

- Foster Communication on Global Basis
- Develop Common Vision and Architecture (research issues)
- Support Federation of communities and sub architectures on a global basis
- Synthesize and Document results in a ‘White Paper’ that could drive R&D Efforts
- Multi-Agency Approach
Situation - Data Generators

- Sensor Systems and Networks (e.g., Smart Dust, Roadnet, Tsunami Warning)
- Experimental Devices (e.g., HEP)
- Computational Modeling (e.g., Tera/PetaFlop Systems)
Real Life: Current Situation
Goal for the Future: One to Many
An Approach

Step 1—Develop Constituents

- Standards (W3C, CMG)
- Academics developing capabilities (GRID, HPS, Digital Libraries, Storage, Persistent Archives)
- Applications - domains needing/using technologies to solve large scale problems
- Industry Partners, Vendors
- Stakeholders (Sponsors, Funding Agencies)
Step 2—Create Requirements Baseline

- Requirements - that are being met and those in need of solutions.
- Capabilities that are being provided by current and developing solutions.
- Reality Check - sample representative communities to verify reality of Requirements/Capabilities.
International Collaborative Effort

- Develop and Conduct International Workshops to accomplish Technical Goals
- Document Results of Each Workshop in formal white papers (e.g., Gridbook)
- Present White Papers and Proposals at International Symposium
- Obtain ‘buy-in’ from wide Communities of Users/Data for ‘Common Software Bus’
- Present Results to Stakeholders in form of R&D Plan
Workshop Scope/Format

- Position Paper Drafted in Advance
- Plenary Sessions to Set Goals/Objectives
- Breakout of Experts Sessions
- Writing Session to create Workshop Report
- Small with about 30 Experts
- Frequency - about every 6 months
- Length - 3 days to insure Reports are Reviewed, Endorsed and Documented
Workshop #1

- “Meet & Greet & Brag”
- Theme: What am I doing and What do I need?
- Develop Matrix of Communities, Needs and Capabilities and Holes to form baseline
- Document Results

Workshop #2

- Identify Applications
- Generate In-depth Application Profiles
- Document Results

Workshop #3

- Identify Technologies
- Profile of what’s being done and what architecture is needed
- Document Results
Possible Organization for Collaborative Effort

**Steering Committee**
- Coordinates Effort
- Endorses Recommendations
- Plans International Symposium

- Workshop 1
  - Develop Constituents
  - Program Committee

- Workshop 2
  - Identify Applications
  - Program Committee

- Workshop 3
  - Identify Technologies
  - Program Committee

**INTERNATIONAL SYMPOSIUM**
Forum for Exchange of Ideas and Adoption of Approach

Synthesized Results
- Integrated Set of Recommendations
- Blueprint for R&D Plan to Reach Goal