CERES DATA MANAGEMENT TEAM
& ASDC: TECHNICAL EVOLUTION

KATHLEEN D. MOORE & JEFF WALTER
MAY 16, 2017
CHALLENGES

• Science data processing paradigm: est’d 1990s
  • Different technologies, methodologies, operational paradigms, constraints...
  • Continuous modification
    • Code w/o refactoring: evolution → increased complexity
    • Environment changes
    • Increase in brittle dependencies
    • Insufficient modularity
  • No comprehensive review/refactoring
  • Highly complex – will increase with FM6 and RBI
CHALLENGES

• **Seven** science working groups
  • Quasi-independent work
  • Independent code bases

• Testing, Integration and Operations
  • Manually intensive SI&T and operations
  • Excessive human intervention, dependence on key individuals
  • CERES/ASDC testing and integration churn
  • Science working groups unable to run comprehensive test cases w/o ASDC operations staff

• Physical resource contention
HIGH-LEVEL APPROACH

• Rearchitect CERES code
• Modernize the existing
dev ➔ test ➔ integration ➔ operations paradigm
• Automate processing
OPERATIONS CONCEPT

Structural

Development

Testing

Systems Engineering/Architecting

Technical

• Code re-architecting
• Programming language
• Scripting framework
• Automation
  • Testing
  • Pipelining
Per-Subsystem Assessments:

1. Identify issues: structure, readability, maintainability
2. Update Operations Concept
3. Repeat
CODE RE-ARCHITECTURE

Atlassian

Bitbucket

Code Repository


CODE RE-ARCHITECTURE

Interviews w/ working groups, subsystems

Processing positives, negatives
Wishlists
AUTOMATED PROCESSING SYSTEM

• ASDC plan: hybrid cloud science data processing system for MAIA
  • Based on JPL’s HySDS system (currently evaluating)
  • Mission-specific production rules and archive system interfaces
• On-premise, heterogeneous science data production environments, combined:
AUTOMATED PROCESSING SYSTEM

Potential benefits:

• Reduction in hardware footprint

• Reduction in LoE (ops and testing)

• Elimination of stovepiping → economies of scale
NEXT STEPS

• Document Operations Concept
• Continue to actively solicit input from CERES stakeholders
• Schedule and plan
• Re-architect CERES codebase
• Build automated hybrid cloud processing system
  • MAIA, then CERES firmly in mind
  • Discuss cooperation/coordination with CLARREO-Pathfinder
  • Potential FLASHFlux and MISR integration