SAND2017-1184C

Unclassified

W80-4 ANALYSIS PAIN POINTS (An Integrated Workflow Perspective)

Robert L. Clay

JOWOG-34 Keynote

Feb 6, 2017

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Acknowledgements

- *Randy Lober* – IWF PO lead
- W80-4: *Patty Hough*, Shawn English, Jay Dike, Terry Johnson, Nathan Spencer, Jeremy Templeton, Ryan Keedy, electrical folks!
- ASC: *Ed Hoffman*, Foss Friedman-Hill, Warren Hunt, Mike Glass, George Orient, Brian Adams, Martin Heinstein, Matt Staten, Brian Carnes, Trevor Hensley

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Outline

- Integrated Workflow Project Overview
- (W80-4) Use Cases & Requirements
- (W80-4) Analysis Pain Points
- Conclusions

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Sandia National Laboratories

Integrated Workflow (IWF) Project Overview

- ASC Integrated Project
  - Created by ASC program leads
  - Key goal is to deliver integrated [analysis workflow] capability to NW analysts (focus on W80-4)
  - Includes: SAW, Cubit, Sierra, Ramses, Dakota, Slycat, FOUS, …
  - ~$5M/yr (commitments, not new $)
  - RLC (PI); Randy Lober (PO); large team of PO's representing sub-projects
  - Management/administration as lightweight as possible
- History and Progress
  - Started late Q3 FY16
  - FY16 was info gathering
  - Q1 FY17 was planning, and getting teams aligned on integrated deliverables
    - Sounds easy; wasn't
    - Planning workshop was very effective (Oct 2016, documented)
  - FY17 planning still being refined, but *work is moving forward at pace*
  - FY18+ planning will happen in April to sync up with ASC cycle

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Outline

- Integrated Workflow (IWF) Project Overview
- (W80-4) Use Cases & Requirements
- (W80-4) Analysis Pain Points
- Conclusions

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

IWF Requirements Capture Process

- FY16
  - Initial W80-4 staff interviews and requirements capture - Hoffman & Hough
- FY17
  - October IWF planning workshop:
    - Analysts present their use cases & requirements
    - Synthesized rolled-up requirements and deliverables
    - Deliverable traceability developed, mapping to requirements
  - Detailed traceability worked Nov 2017
    - Follow-up interviews with all W80-4 stakeholder staff
    - Requirements->Deliverable gaps identified and partially closed
    - Graphical traceability developed (MindMap images)
  - Ongoing work
    - Simplifying top-level goals/deliverables (add clarity; improve structure)
    - Load balancing staff resources
    - Additional deliverable resolution

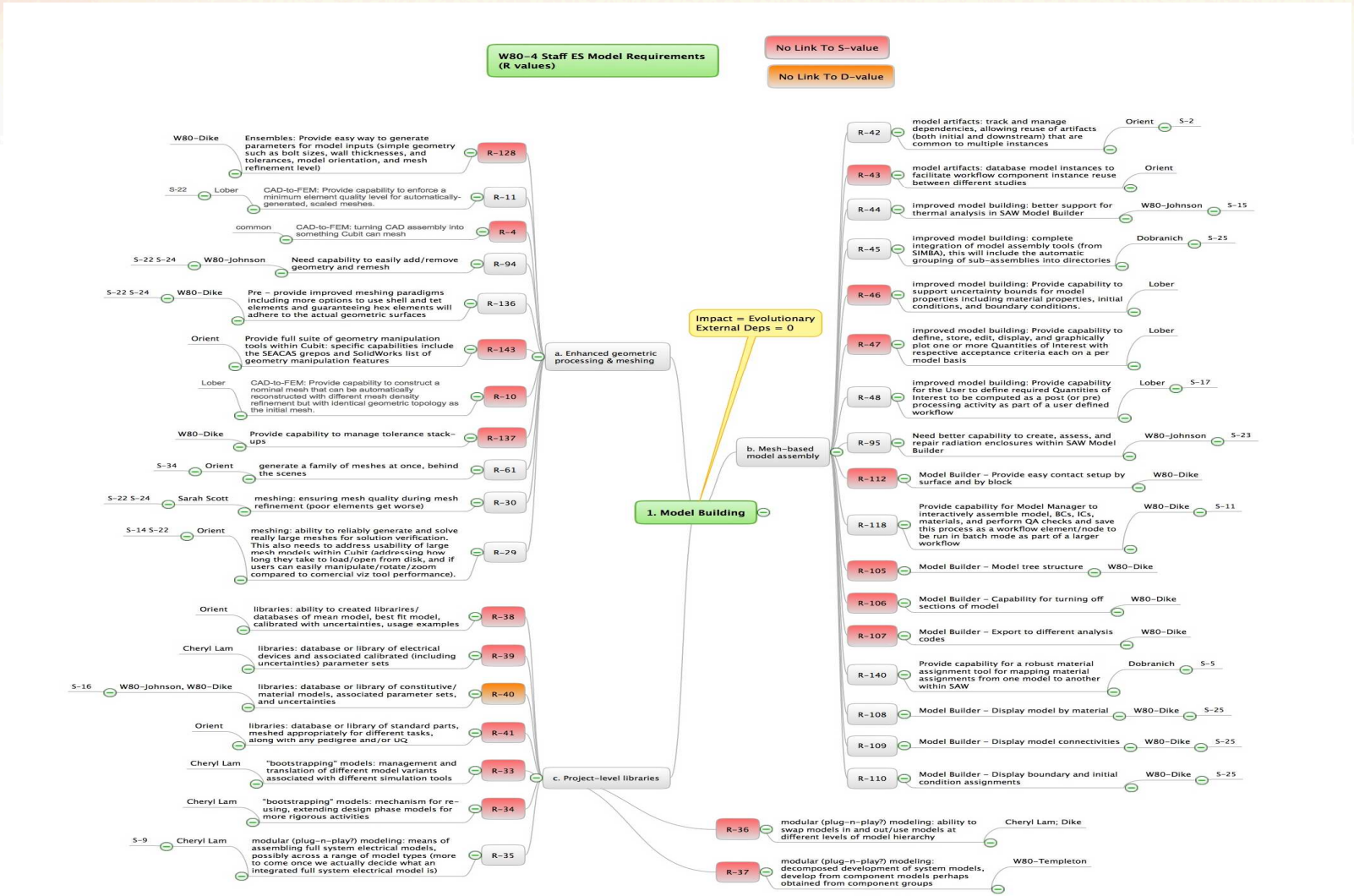
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

IWF Requirements Capture Process

- Prelim user input – summer 2016
- October 2016 workshop, followed by Randy doing a lot of requirements data analysis
- The following eye charts give a sense of the analysis of the relationship between rolled-up key deliverables and requirements
  - Looking for coverage (of requirements) and proper mapping (between deliverables to requirements)
  - 154 requirements doesn't make for 154 deliverables
  - We synthesized integrated deliverables that captured groups of requirements, and started organizing around 'themes'
  - First cut was team-based; redux accounted for more cross-cutting solutions

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

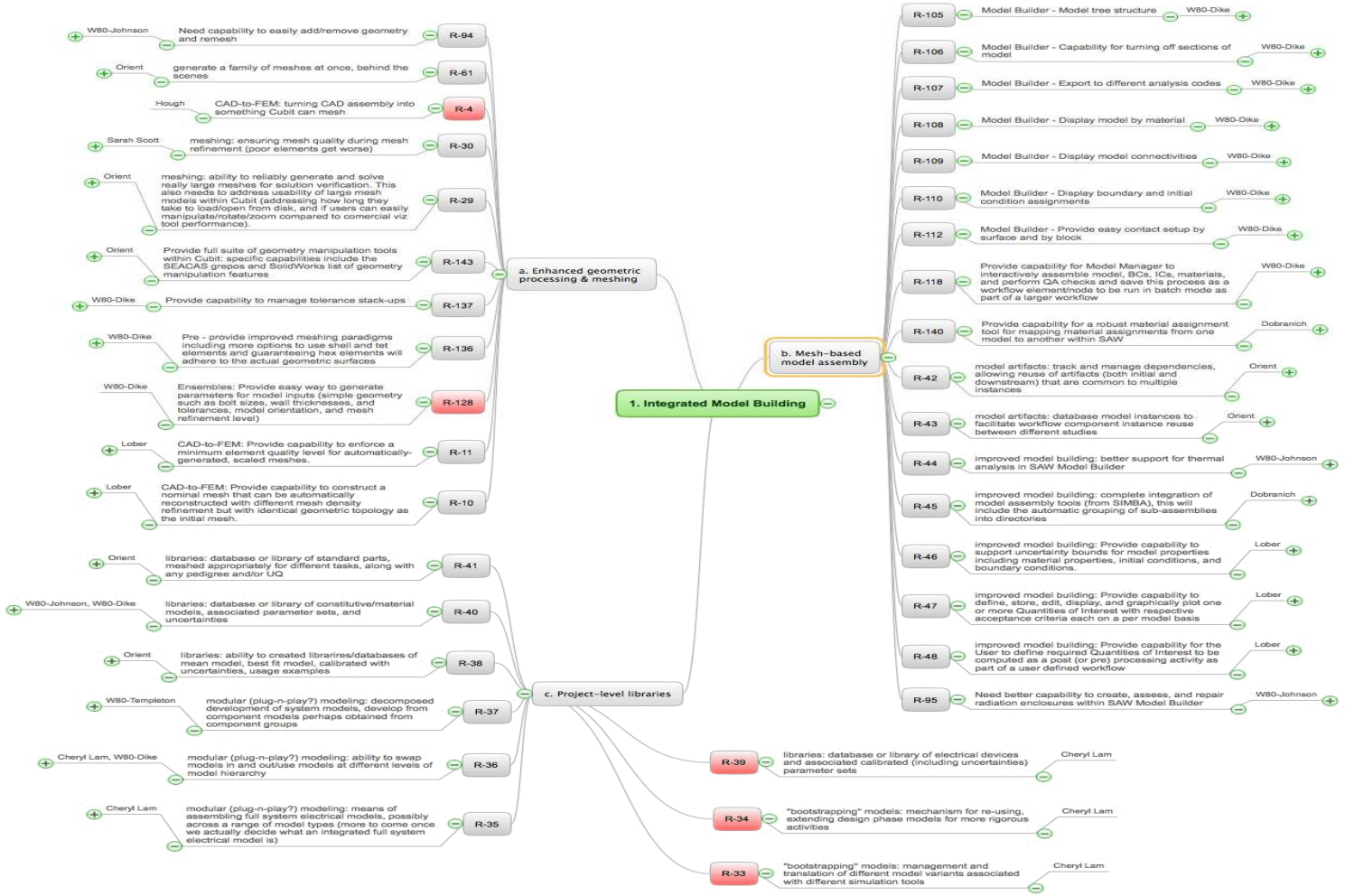
Model Building requirements, at Oct'17 IWF Workshop



Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Sandia National Laboratories

Model Building requirements, after analysis



Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Sandia National Laboratories

Deliverable Reference – part 1

D Value	Deliverable Title	Theme
D1	Provide full functional SAW Cubit	2.a.
D2	Cubit Pre-Analysis Verification Tools	1.b.
D3	Faster/Bigger Cubit Serial Tet Mesher	1.a.
D4	Parallel Tet Mesher	1.a.
D5	Tet Mesh Scaling for Solution Verification	1.a
D6	Next Generation Workflow (NGW)	3.a. & 3.c.
D7	SIMBA Features in Integrated Framework	1.b.
D8	Simulation Data Management (SDM) - Tagging & Branching	2.b.
D9	Slycat/SAW Integration	4.c.
D10	Visualization Tool/SAW Integration	4.a.
D11	In Situ Simulation Results and Automated Post Analysis	4.a.
D12	Provide Improved Sierra/SAW Integration	2.a. & 2.c.
D13	Catalyst integration into all Sierra modules	2.b.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Deliverable Reference – part 2

D Value	Deliverable Title	Theme
D14	Provide Extended output capabilities	TBD
D15	Metadata propagation/preservation within Sierra	5.c.
D16	Dakota Examples Repository	2.a.
D17	Dakota Visualization and Insight	2.a.
D18	Improve Dakota/SAW Integration	2.a.
D19	<removed>	
D20	Support Ensemble Parallel Modes	TBD
D21	V&V Specific New User Training	2.c.
D22	Provide Credibility of Workflows	TBD
D23	Provide Solution Verification Workflow	TBD
D24	Advance Ensemble Analysis Capabilities	4.c.
D25	Dakota Steering	TBD
D26	Report Generation Framework hardening	4.b.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Deliverable Reference – part 3

D Value	Deliverable Title	Theme
D27	Geometric Tolerance management	TBD
D28	Export Model to external analysis codes	TBD
D29	Uncertainty Management at Workflow level	TBD
D30	Model Assembly	1.b.
D31	Library Capability Across Projects	2.b.
D32	Improved Repository Searching	2.b.
D33	Code Deployment Plan	2.c.
D34	Validation Metric Library	TBD
D35	Metadata creation and retention	TBD
D36	Improved Job Submission Process	3.b.
D37	Improved Job Submission resource leveraging	3.b.
D38	Human in the loop Workflows	3.a.
D39	Workflow Customization	3.b.
D40	Universal Journaling	TBD

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

IWF Capability Development Themes

- Integrated Model Building
- Workflow Automation
- Ensemble Analysis
- Analysis Insight Assistance
- Infrastructure (“trail blazing”)

These “themes” provide an organizing structure for the IWF project

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Outline

- Integrated Workflow (IWF) Project Overview
- (W80-4) Use Cases & Requirements
- (W80-4) Analysis Pain Points
- Conclusions

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Analysis Pain Points

- Time (Speed)
- Credibility
- Data Management
- Process Management
- Scalability (Problem Size)
- Complexity (Human Factors)

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Analysis Pain Points – Time

- Time is *the* top-level driver – analysts need more time
- Time is entirely cross-cutting
- Every aspect of analysis impacts time
  - Problem setup (e.g., meshing, model building, workflow building, …)
  - Simulation execution (e.g., running ensembles, debugging failed jobs, …)
  - Results analysis (e.g., visual analytics, graphics generation, ...)
  - Report generation (e.g., automated assurance info, ...)
- It's a zero-sum game – time spent on one thing reduces time to use on another
- Bottom line on time
  - We need more time spent on analysis
  - And less time spent on using the analysis tools

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Analysis Pain Points – Credibility

- Can the result be reproduced?
  - Repeatability is mandatory (not just by the same analyst!)
- Can the result be verified?
  - Demonstrate solution convergence
- Can the result be validated?
- Can the workflow be shared with another analyst?
  - Workflow must be portable (user agnostic)
- Credibility of ensemble analysis is crucial as automated processes dominate
  - One example: auto-generation of QOI plots
- Bottom line for Credibility
  - Must be an integrated part of the workflow – not increased overhead/work
  - Reproducibility & sharing of workflows must be standard procedure

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Analysis Pain Points – Data Management

- Data management underpins assurance and credibility
- There are many aspects of (requirements for) data management
  - Tracking (and using!) meta data from design through analysis
  - Automated data capture as part of analysis workflows
  - Analysis data repository with versioning, dependency tracking, provenance, …
  - 'Archival' – this is used to certify NW system requirements
  - Full security model – kerberos, NTK, Sigma 15, ACLs, team based, ...
  - Distributed computing environment
    - jobs run on local and remote machines
    - 'central' database
  - Open-source, portable solution
  - Scalable solution (user's archiving 1 TB runs these days, and we expect that to grow)

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Analysis Pain Points – Process Management

- Process management essentially is workflow
- There are many aspects of (requirements for) process management
  - Automated analysis workflows – time saving, key to reproducibility, sharing, lowering complexity, …
  - Ensemble analysis workflows are the new normal
  - Simple to build, run, share, debug, save, modify, …
  - Intuitive – should be able to see a workflow and make sense of it without being a ‘workflow expert’
  - Open-source, portable solution
  - Scalable solution – works on ACS-x systems
  - Extensible – easily add [user-contributed] components
  - Full security model – kerberos, NTK, Sigma 15, ACLs, team based, ...
  - Distributed computing environment

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Analysis Pain Points – Scalability

- Growth of computing resources will continue
  - Support for ACS-x systems is a *key deliverable*
- Scalability is a cutting pain point – it all needs to scale
- Scalability requirements are embedded within:
  - Process and Data management of ensemble analysis workflows
  - Model creation tools for generating large scale models
  - Preprocessing tools capable of scaling up existing models for convergence work
  - In Situ visualization and automated post-processing of quantities of interest (user configured)
  - Advanced methods to harvest and aggregate ensemble data and quantities of interest (at scale)
  - Framework to support automated report generation (reduces analyst burden of manually creating and recreating large numbers of similar QOI plots/images and locating them into a report format)
- Bottom line – workflow capabilities need to operate smoothly at scale

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Analysis Pain Points – Complexity

- Reducing complexity improves everything, basically
- Simpler means easier
  - Easier to build, run, share, test, debug, understand, …
  - More intuitive – look at the workflow and understand it on sight
  - *Ease of use can be strategic*
- Simpler and easier means:
  - Faster (less time messing with tools; more time analyzing the problem)
  - Fewer mistakes
  - Easier to explain to others – simpler to share [workflows]
  - Promotes reusable components – user provided
  - Fewer ‘exotic’ solutions [that one person understands]
- Very sophisiticated workflows can be constructed from relatively simple components
  - Make it as complicated as it needs to be
  - But no more so

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Outline

- Integrated Workflow (IWF) Project Overview
- (W80-4) Use Cases & Requirements
- (W80-4) Analysis Pain Points
- Conclusions

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Conclusions

- Analysts want to spend less time on the tools, and more time on analysis.
- Analysts want to gain fidelity and credibility of their work without adding time and complex processes.
- Analysts [develop and] use workflows, although not necessarily with workflow management systems.
- Analysts are willing participants to help define better [ASC] solutions.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Questions?

rlclay@sandia.gov

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE

Sandia National Laboratories

Save that one for the panel!

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND2016-7247 PE