CESD – ESS CYBERINFRASTRUCTURE WORKING GROUPS
Annual Meeting Highlights

J. David Moulton (LANL)

Outline

Recap and Highlights from 2018

- Motivation, how we formed, our structure …
- Activities

Annual Meeting (yesterday)

- Recap of groups scope and activities
- Internal DOE, external highlights on Cyberinfrastructure
- Planning discussions

ESS Cyberinfrastructure Working Groups Structure

Following recommendations of the workshop report and feedback from 2015 AGU Town Hall the Working Groups Kickoff was in 2016

Established an Executive Committee

Deb Agarwal, Xingyuan Chen, Paul Hanson, Forrest Hoffman, David Moulton (Chair), Eric Pierce, Dean Williams, Roelof Versteeg, Stan Wullschleger

Established Three Working Groups

Data Management
Model-Data Integration
Software Engineering and Interoperability

Reporting on Working Group Activities

Informal reporting to S.C. (PMs) and E.C. (PIs)
Annual reporting at ESS PI Meeting

Model Development (MD)
Data Management (DM)
Software Engineering (SE)
Workflows for Model-Data Integration (WMDI)

SBR Project
TES Project
WMDI
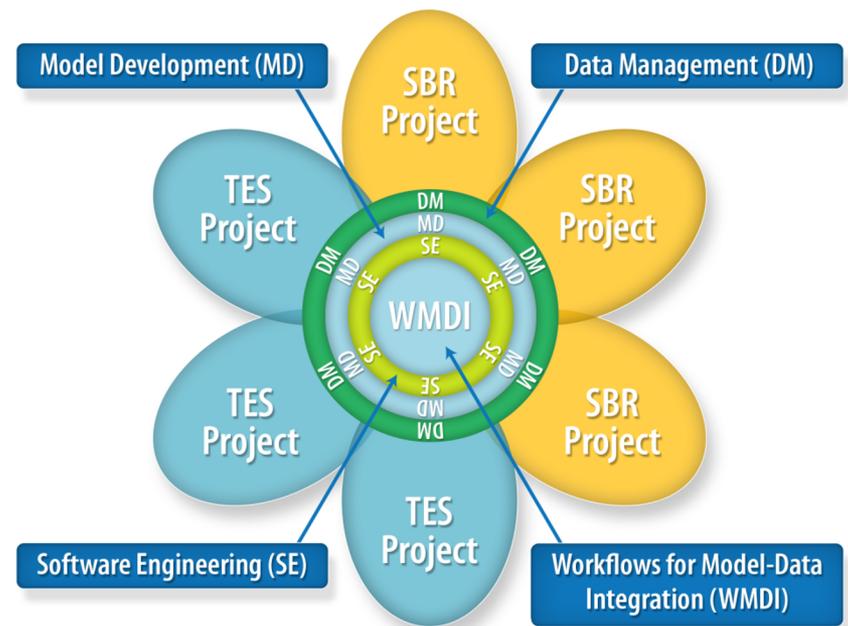
Environmental System Science

Highlights from the Working Groups

Model Intercomparison and Benchmarks participation

International Land Model Benchmarking (ILAMB) Workshop and Tools

Soil Carbon Dynamics Working Group for data synthesis

Community Outreach

Co-Led sessions for AGU Fall Meeting 2018:

- Computational Methods and Tools for Model-Data Integration
- Big Data in the Geosciences: New Approaches to Storage, Sharing, and Analysis

Breakout Sessions tonight

- Open Watersheds: Amplifying the Impact of SBR Science
- ILAMB Soil C Demonstration and Tutorial

Highlights from the Working Groups

Facilitate communication across projects

- MDI conducted a Community survey on workflows and model-data integration tools last year

- We had 30 respondents

- Suggestions for topics not covered and future activities will be incorporated for next year

Sample questions:

What software tools do you prefer for data analysis?

What software tools do you use or prefer for workflow management, provenance tracking and archiving?

Working Groups Annual Meeting

Annual Meeting to be held on:

Monday preceding the ESS PI Meeting

58 participants, 7 BER representatives

A quick recap of recent activities

Community oriented projects:

ESS DIVE

IDEAS-Watersheds

ILAMB/IOMB, RUBISCO Soil Carbon Dynamics Working Group

Cyberinfrastructure activities

Within DOE, and external to DOE

Brainstorming for the future …

DOE Flash Talks

Cyberinfrastructure in practice, progress and needs:

Danielle Christianson - Ameriflux processing pipelines

Chris Henry - KBase connectivity with ESS Cyberinfrastructure

Kjierstan Fagnan - National Microbiome Data Collaborative/JGI

Dave Millard - MyEMSL collaborations with ESS-DIVE, KBase

Ethan Coon and Dipankar Dwivedi - ExaSheds

Ben Sulman - NGEE Arctic - E3SM/PFLOTRAN Liaison talk

Roelof Versteeg - Phase 1 SBIR

John Tourtellot - Kitware Phase 1 SBIR.

Informal discussion (pop ups) included Joanna Martin from OSTI.

External Flash Talks

Cyberinfrastructure in practice, progress and needs:

Dwayne Young - EPA Internet of Water

Tim Kern - USGS

Peter Arbuckle - USDA

Lejo Flores - CUAHSI

Matt Jones - DataONE

NMDC Vision expresses common theme:

Empower the research community to harness microbiome data exploration and discovery *through a collaborative integrative data science ecosystem.*

Looking Ahead

Explore adjustments to structure to help make us more effective during the rest of the year

Identify leads for small groups attacking ``hot topics''

Plan and hold a **workshop** on needs for model data

  Traceability of simulations (model input and model output)

 Include stages of data preparation (raw to product)

Work closely with community oriented projects:

  ESS-DIVE, IDEAS-Watersheds, RUBISCO, WHONDRS, etc.

Continue community outreach

# Building Community Cyberinfrastructure …
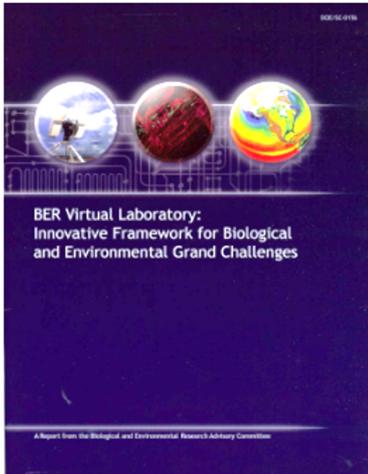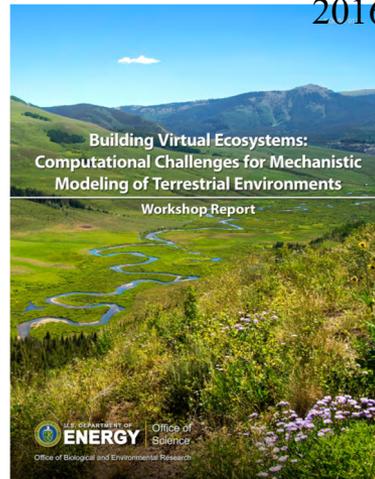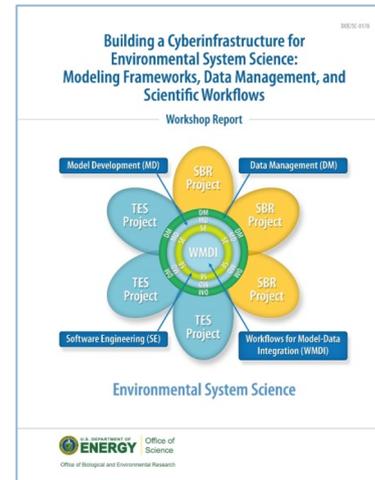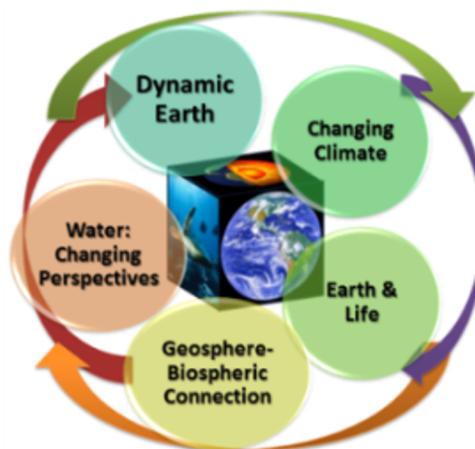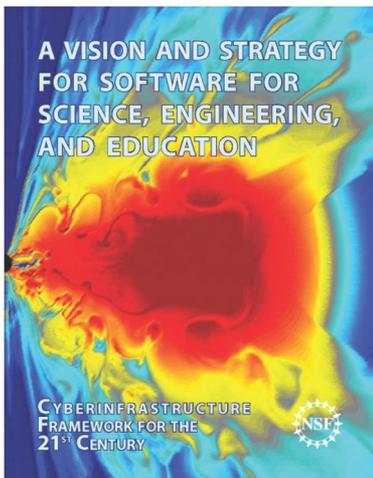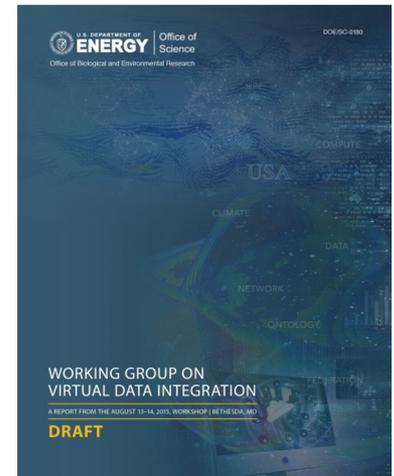
ESS Executive Committee Workshop on Data Infrastructure

Workshop:　　　　August 29-30, 2016; Germantown, MD
Organizers:　　　　Jay Hnilo (CESD) and David Moulton (LANL, EC – Chair)
Participants:　　　　10 participants representing Executive Committee, 7 DOE PMs

Motivation:

"the innovation most needed is a framework that allows seamless integration of multiscale observations, experiments, theory, and process understanding into predictive models for knowledge discovery" (BERAC 2013)

Exponential growth in the amount, variety and complexity of scientific data

Significant fragmentation across projects and disciplines still remains

A need for a data center that would be a foundational part of a community cyberinfrastructure

Towards a Shared ESS Cyberinfrastructure:
Vision and First Steps

Draft

Report from the ESS Executive Committee Workshop on Data Infrastructure
August 29-30, 2016.
DOE Headquarters, Germantown, MD

http://science.energy.gov/~/media/ber/pdf/workshop%20reports/Towards_a_Shared_ESS_Cyberinfrastructure.pdf

ESS Community Cyberinfrastructure and the Virtual Laboratory
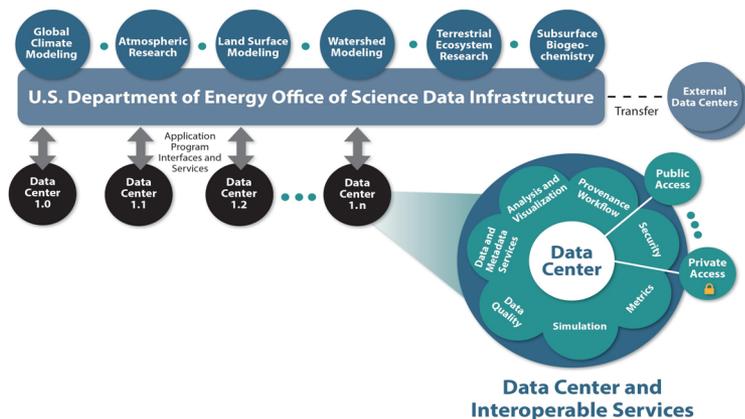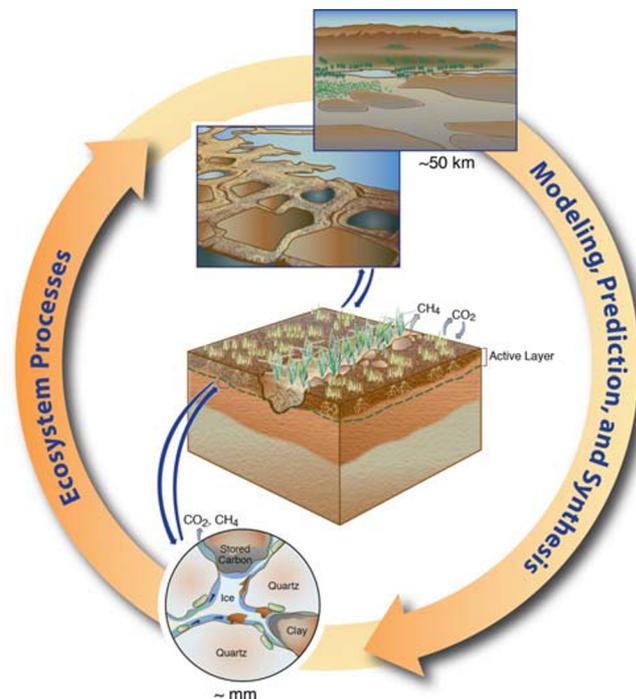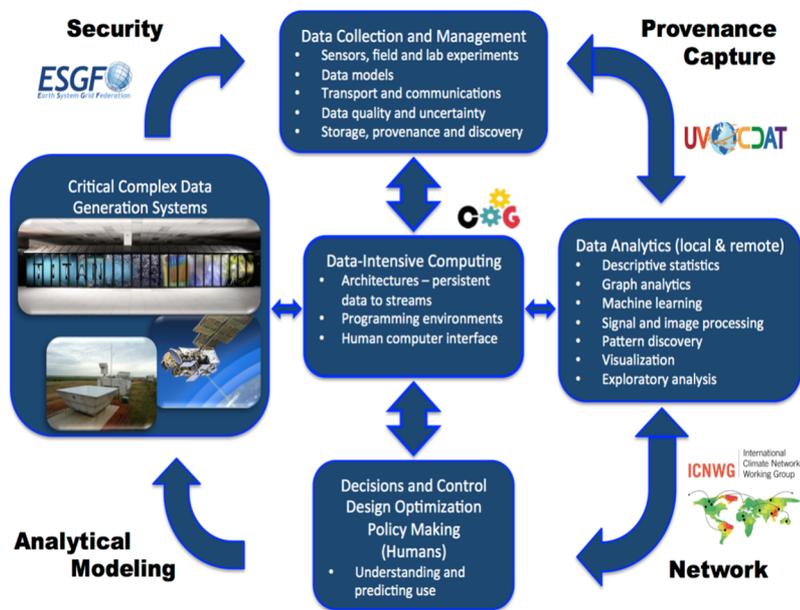
Reconciling *different view points and emphasis into requirements* for a new data center and a phased approach to community-based cyberinfrastructure development.