Semi-Automatic Science Workflow Synthesis for High-End Computing on the NASA Earth Exchange

ESTF – October 28, 2014

Petr Votava, Andrew Michaelis, Ramakrishna Nemani, Hirofumi Hashimoto, Weile Wang

Background: NASA Earth Exchange

Vision: To engage and enable the Earth science community to address global Earth science challenges.

NEX is a collaborative compute platform that improves the availability of Earth science data, models, analysis tools and scientific results through a centralized environment that fosters knowledge sharing, collaboration, innovation and direct access to compute resources.

Engage:
Network, share and collaborate
Discuss and formulate new ideas
Portal, Virtual Institute

Enable:
Access to data
Access to computing
Access to knowledge

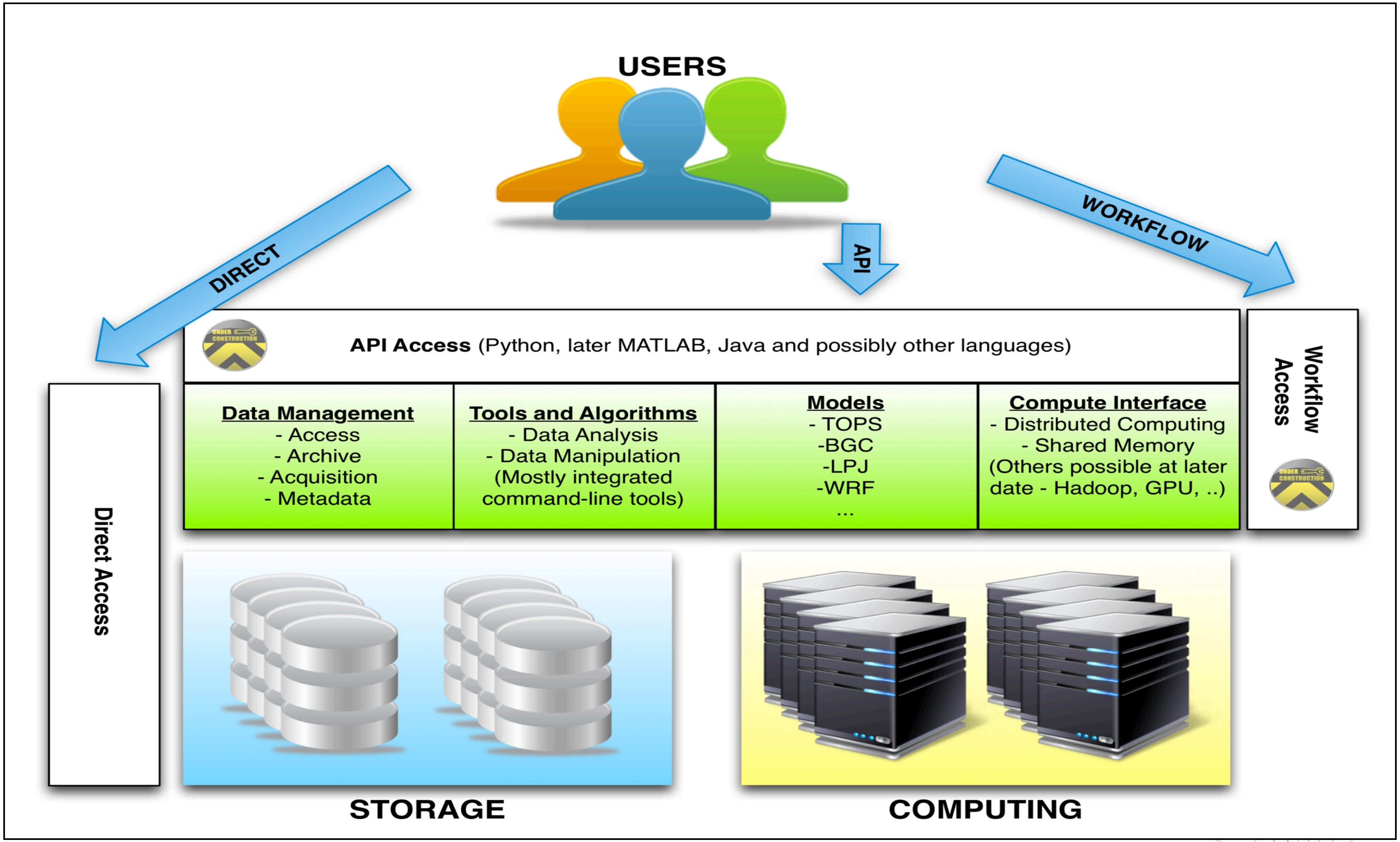
Role of workflows in NEX

NASA Earth Exchange

- (Currently using VisTrails Workflow Management system)
- Improve reproducibility and provenance
  - Results on NEX contain reference to the workflow, so it could be simply reproduced
  - Provenance can be tracked through data and code (through development and execution)
- Improve transparency
  - The process is visually depicted and each step can be examined.
- Improve collaboration
  - Multiple people can be editing and experimenting with the workflow and share results.
- Improve reuse

NEX Software View

NASA Earth Exchange



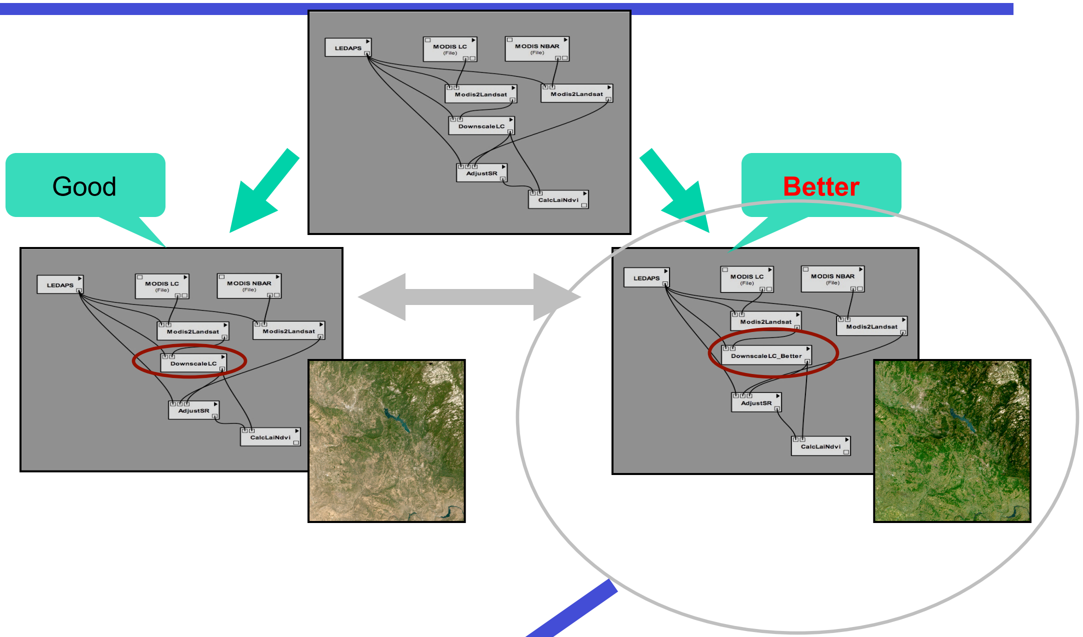
NEX Workflow Deployment Goal

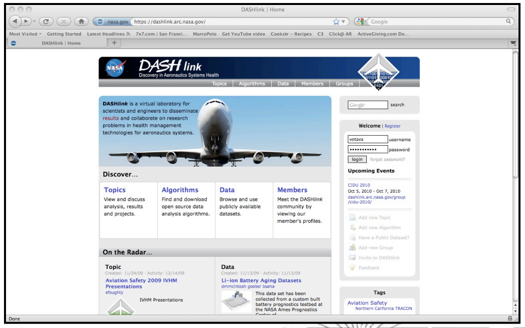
Process Design

Iterative Refinement/ Distributed Collaboration

Full-scale Execution

On-Line Result Sharing and Process Re-use





DASH*link* is a virtual laboratory for scientists and engineers to disseminate results and collaborate on research problems in health management technologies for aeronautics systems.

Discover...

Topics
View and discuss analysis, results and projects.

Algorithms
Find and download open source data analysis algorithms.

Data
Browse and use publicly available datasets.

Members
Meet the DASH*link* community by viewing our member's profiles.

votava

Upcoming Events

CIDU 2010
Oct 5, 2010 - Oct 7, 2010
dashlink.arc.nasa.gov/group/cidu-2010/

On the Radar...

Topic
Created: 11/24/09 · Activity: 12/14/09
Aviation Safety 2009 IVHM Presentations
efoughty
IVHM Presentations

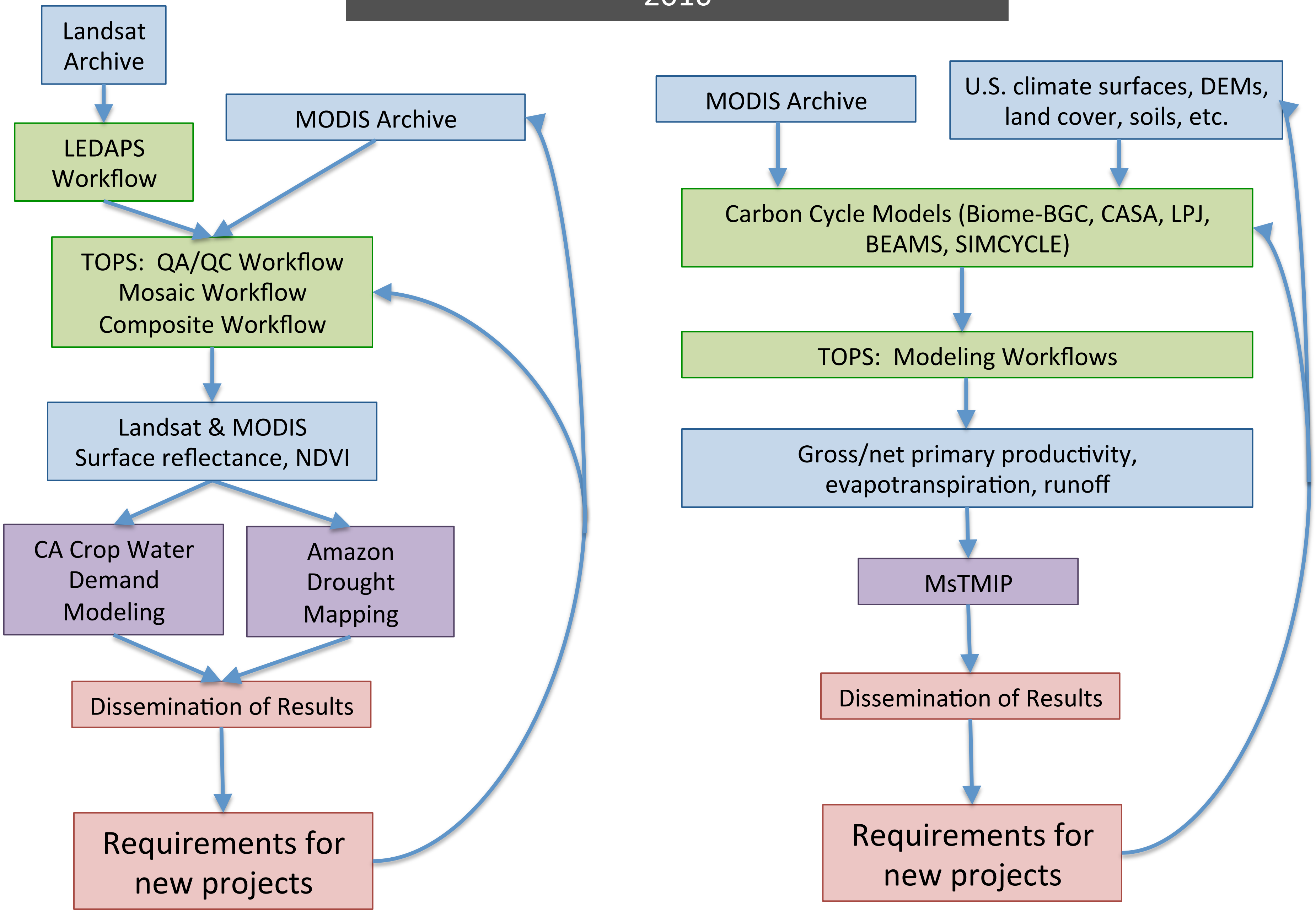
Data
Created: 11/13/09 · Activity: 11/13/09
Li-ion Battery Aging Datasets
dmmcintosh goebel bsaha
This data set has been collected from a custom built battery prognostics testbed at the NASA Ames Prognostics Center of

Tags
Aviation Safety
Northern California TRACON

NEX: Sample Workflows
2010

MODIS / Landsat

Climate Impacts



Landsat Archive

LEDAPS Workflow

MODIS Archive

TOPS: QA/QC Workflow
Mosaic Workflow
Composite Workflow

Landsat & MODIS
Surface reflectance, NDVI

CA Crop Water Demand Modeling

Amazon Drought Mapping

Dissemination of Results

Requirements for new projects

MODIS Archive

U.S. climate surfaces, DEMs, land cover, soils, etc.

Carbon Cycle Models (Biome-BGC, CASA, LPJ, BEAMS, SIMCYCLE)

TOPS: Modeling Workflows

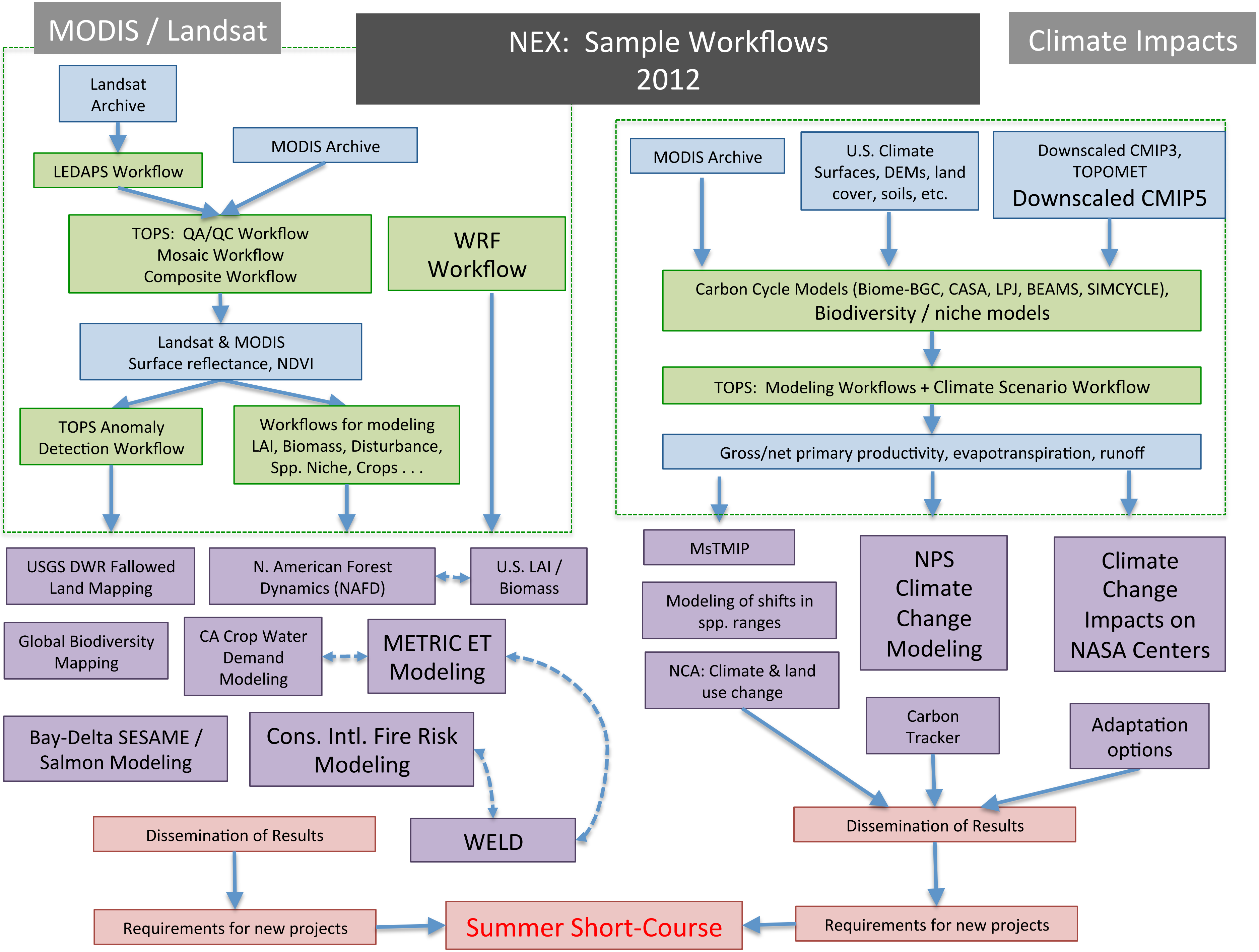
Gross/net primary productivity, evapotranspiration, runoff

MsTMIP

Dissemination of Results

Requirements for new projects

NEX: Sample Workflows 2012



MODIS / Landsat

Landsat Archive

LEDAPS Workflow

MODIS Archive

TOPS: QA/QC Workflow
Mosaic Workflow
Composite Workflow

WRF Workflow

Landsat & MODIS
Surface reflectance, NDVI

TOPS Anomaly Detection Workflow

Workflows for modeling LAI, Biomass, Disturbance, Spp. Niche, Crops . . .

USGS DWR Fallowed Land Mapping

N. American Forest Dynamics (NAFD)

U.S. LAI / Biomass

Global Biodiversity Mapping

CA Crop Water Demand Modeling

METRIC ET Modeling

Bay-Delta SESAME / Salmon Modeling

Cons. Intl. Fire Risk Modeling

WELD

Dissemination of Results

Requirements for new projects

Summer Short-Course

Climate Impacts

MODIS Archive

U.S. Climate Surfaces, DEMs, land cover, soils, etc.

Downscaled CMIP3, TOPOMET
Downscaled CMIP5

Carbon Cycle Models (Biome-BGC, CASA, LPJ, BEAMS, SIMCYCLE),
Biodiversity / niche models

TOPS: Modeling Workflows + Climate Scenario Workflow

Gross/net primary productivity, evapotranspiration, runoff

MsTMIP

Modeling of shifts in spp. ranges

NCA: Climate & land use change

NPS Climate Change Modeling

Climate Change Impacts on NASA Centers

Carbon Tracker

Adaptation options

Dissemination of Results

Requirements for new projects

Focus of this project – barriers removal

- Learning curve
  - Fairly easy to develop new workflows in VisTrails from existing components, but it is harder to develop new components from existing/legacy code
  - This project aims to lower these barrier (not remove it – that would be enormous undertaking)
- Lack of repository accessible through standard protocols outside VisTrails (SQL, SPARQL, …)
  - Infrastructure that enables to find, match and manage individual components and entire workflows outside VisTrails. This is needed to integrate with the rest of the NEX development.
- Minimal support for HPC, which is important for NEX users
  - By the end of the project we aim to have our worklfow management system (VisTrails) running in distributed memory environments.

Project Drivers

- Improved usability of VisTrails workflow management system especially in the context of high-performance computing (HPC)
- This project provides a framework that enhances users ability to capture their research and scale it to HPC
  - But gives them the ability to do as little or as much detail as they are willing to capture
    - Provenance capture, workflows, individual workflow components, documentation, annotations
- Researchers/groups on NEX still work in relative isolation
- Limited research knowledge capture
  - Even people doing the research have hard time figuring out what they did 6 months ago
- Fits well with NEX vision
  - Tool for community contributions (with minimal effort)
  - A form of documentation

Use Case Application: WELD

- Many components of the project are tested within the context of Web Enabled Landsat (WELD) project
  - Global Landsat processing running on NEX
  - 1PB of data
  - Complex production pipeline
    - 30 separate steps
    - Operationalized by NEX team
- “Ideal” test case
  - Many separate processes delivered as a number of scripts
  - Provenance (process capture) is extremely important
  - Originally sequential process that had to be parallelized with non-trivial dependencies in order to scale it up to HPC.
  - Will take several iterations, upstream products are involved, so the automation and process capture is very important.

Project Overview

Capture

Capture processes as they are being executed together with their command line options and files that are being manipulated. Match process signature against NEX knowledge base.

Convert

Convert new processes into workflow components for VisTrails or potentially other workflow management system.

Manage

Store and retrieve workflow components, together with links to the knowledge base. Perform matching or partial matching of discovered processes.

Scale

Prototype on subsets of data and then seamlessly deploy workflows in the HEC environment on global problems.

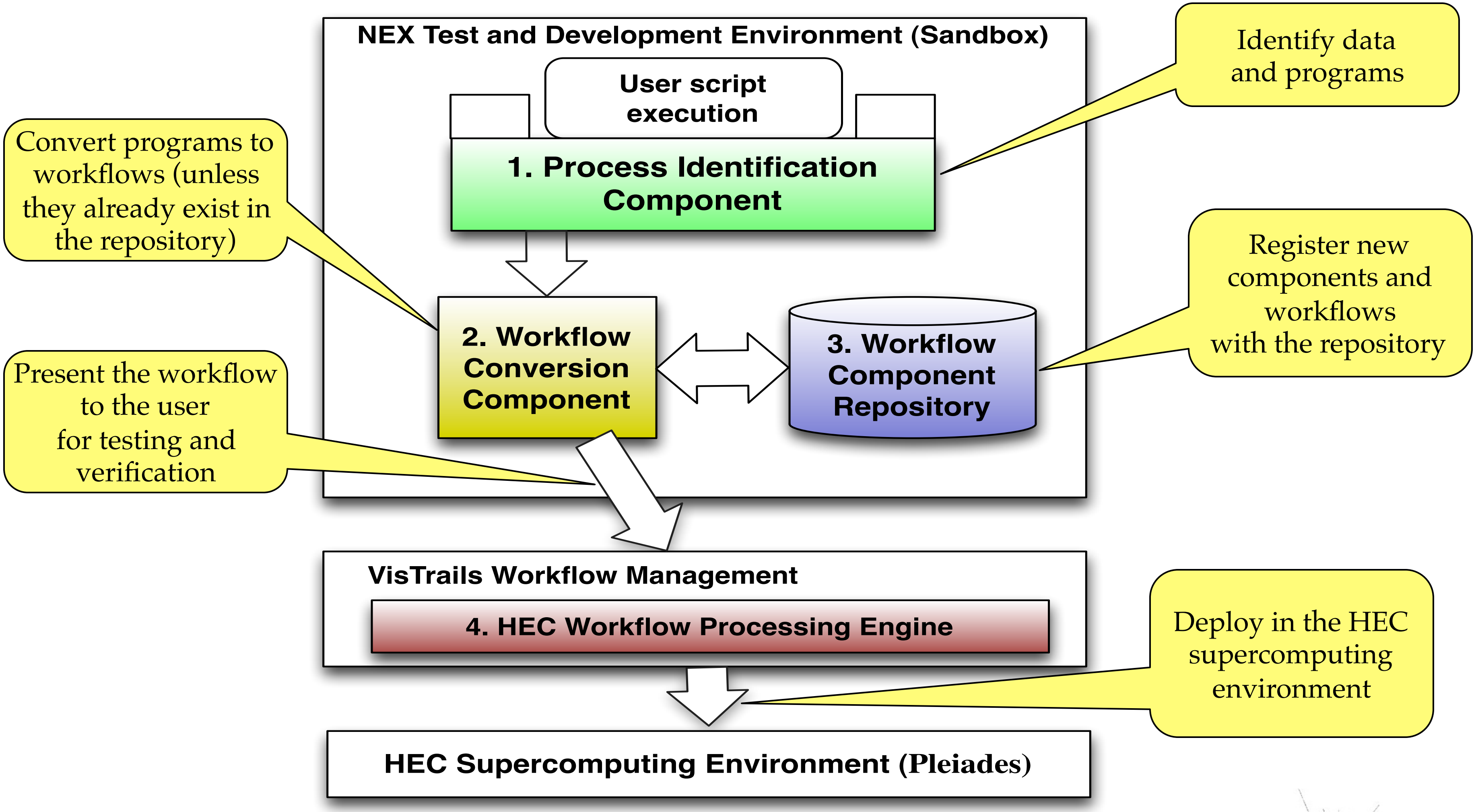
Key Components

- **Process Identification Component**
  - Identifies executable process and their sequencing during the runtime of user scripts
  - Identifies datasets used and/or generated during execution of user scripts
- **Workflow Conversion Component**
  - Converts each captured ore described process into a workflow component or fetches existing component from the repository
- **Workflow Component Repository**
  - Manages and stores workflow components
  - Provide means for user to add and retrieve annotations/tags/documentation etc.
- **HEC Workflow Engine**
  - Facilitates workflow execution at the HEC or other distributed environment

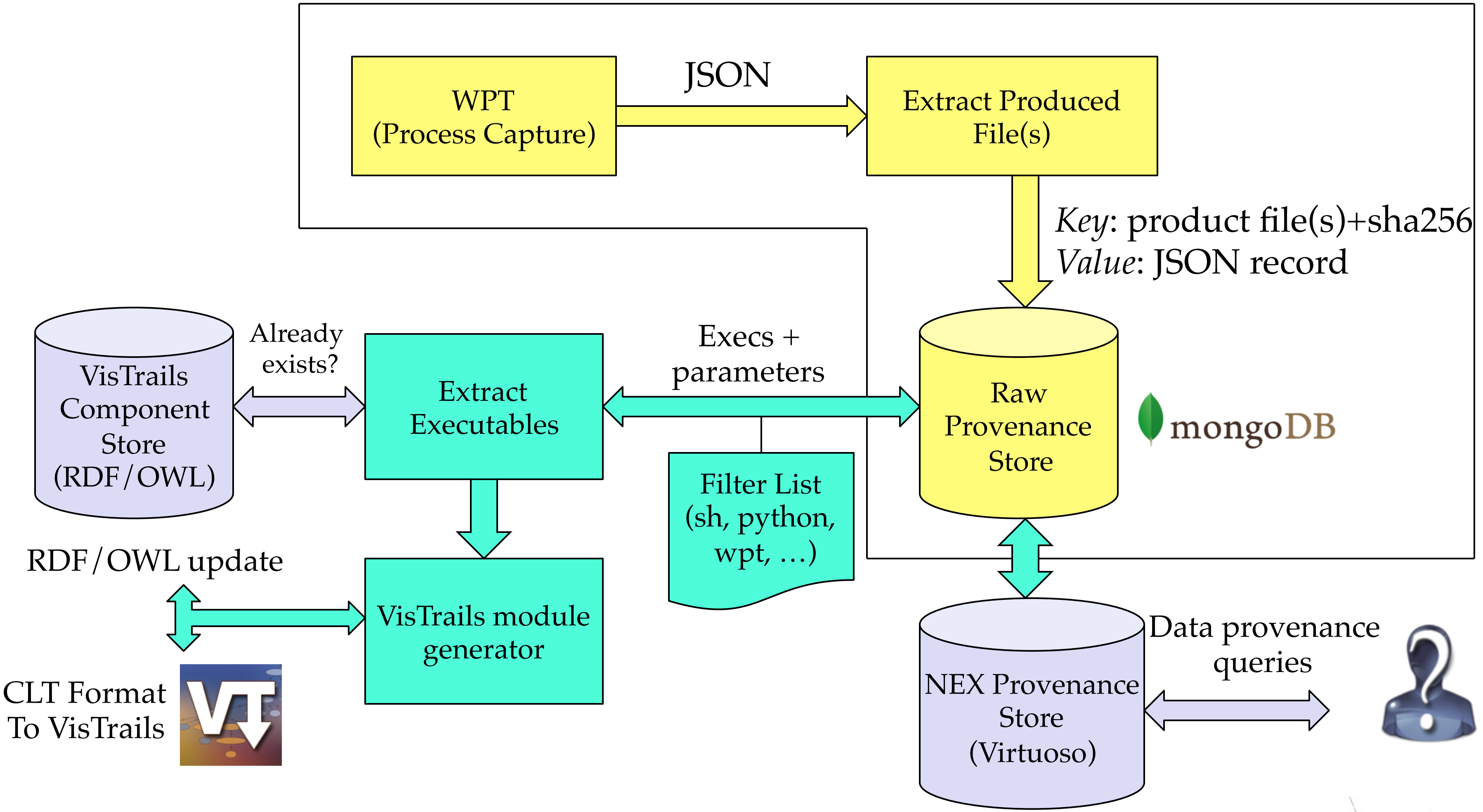
Putting It All Together (Global Landsat WELD Use Case)

1. Researcher comes to NEX to try out a new Landsat processing algorithm with their existing C codes executed through python/perl scripts.
   - The pipeline consists of 30 different programs that are being run in certain order
2. The researcher runs the scripts on the NEX sandbox within the workflow capture environment and the system captures the programs and their command line options as well as data and libraries that are involved in the processing.
3. The system tries to figure out the correct sequence of the calls and creates an approximate runtime call-graph of the entire process.
4. Next the system examines the graph and tries to match the components (executable programs) with existing NEX repository.
5. All the nodes in the graph are either converted to workflow components or are matched with existing components.
6. User does a few test runs of the workflow on single Landsat scenes and then deploys it to Pleiades to be executed in parallel on 15,000 global Landsat scenes.

Process Flow


NEX Test and Development Environment (Sandbox)
User script execution
1. Process Identification Component
2. Workflow Conversion Component
3. Workflow Component Repository
VisTrails Workflow Management
4. HEC Workflow Processing Engine
HEC Supercomputing Environment (Pleiades)
Identify data and programs
Convert programs to workflows (unless they already exist in the repository)
Register new components and workflows with the repository
Present the workflow to the user for testing and verification
Deploy in the HEC supercomputing environment

Capture/Convert/Store -Architecture Overview

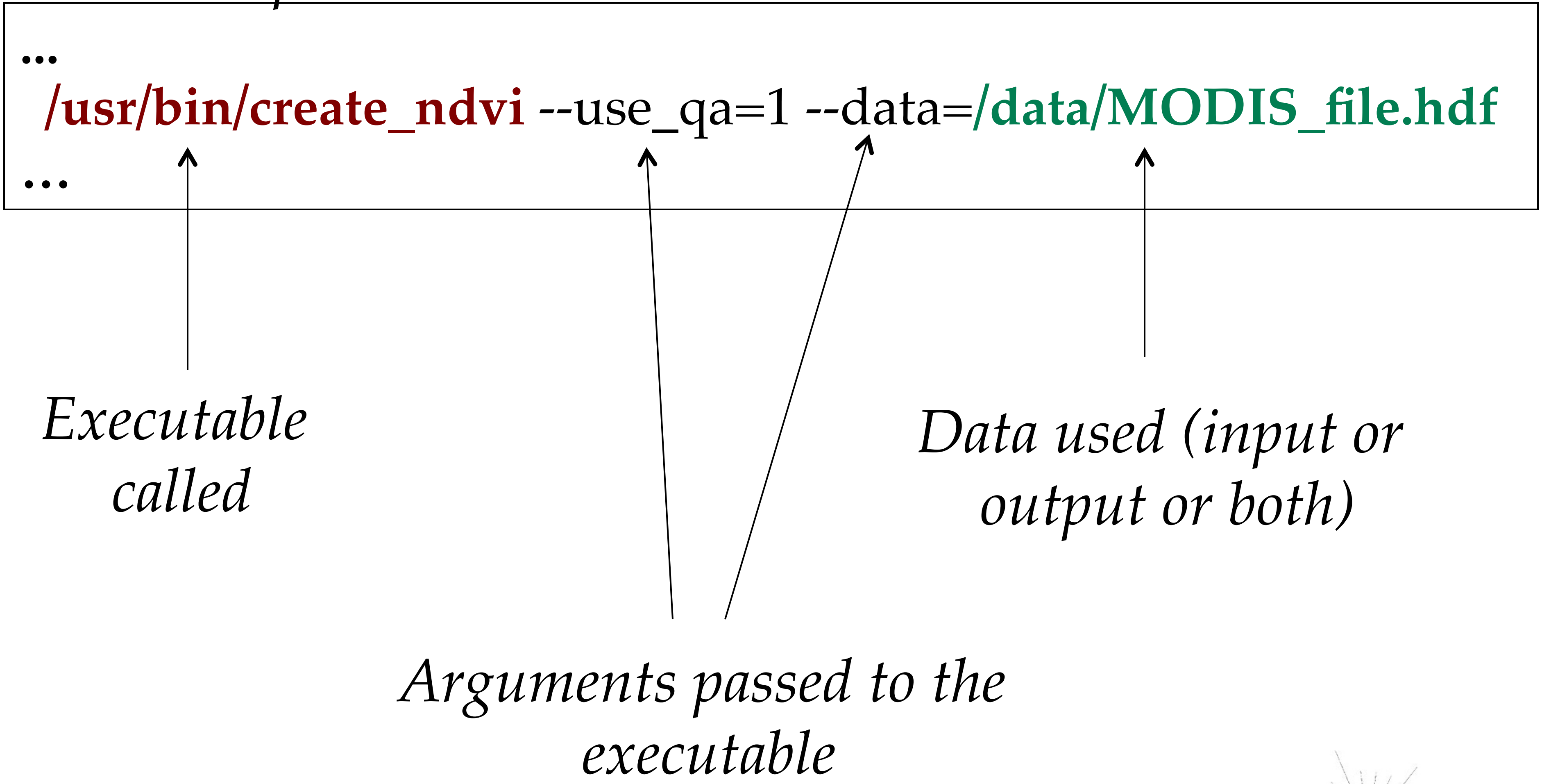


Process Identification

- Developed using Ptrace API
- Captures information from the execution of a user code (C, MATLAB, Python, Java, …)
  - Capture low-level underlying process file access, process chains in real-time, libraries used, …
  - Minimal or no burden to user, **transparent**
  - Low overhead w.r.t. process **performance**
    - Still have some room for improvement
  - Language or tool **independent**
  - Report in JSON format

For Example

In the script:



```
...
/usr/bin/create_ndvi --use_qa=1 --data=/data/MODIS_file.hdf
...
```

Executable called

Data used (input or output or both)

Arguments passed to the executable

Capture Report

- **JSON**

```
“nametag” {
      “pid” : “<pid>”,
      “executable” : “/usr/bin/create_ndvi ”,
      “args” : “[‘--use_qa=1’, ‘--data=/data/MODIS_file.hdf’]”,
      “Children” : [“c1,…]”,
      “fileAccess” : [[“/data/MODIS_file.hdf”, “<mode>”],
      …..]”,
      “ctime” : “datetime”
},
“nametag2” {
      “pid” : “<pid>”,
      .
}
```

Goal: Intelligent History and improved provenance

- Expanding our current provenance capture to an entire session rather than just a single process
  - Produced data can be tagged with the session, so that we can see other related work that was done during the same session
  - An expanded way of documenting scientific process
  - Also – not all data are usually generated by a single script
    - This way we can really trace data across sessions and days/months.
- We can then derive the entire sequence of calls that produced any data file (and possibly rerun the sequence and reproduce the result)

(Executable) Provenance Capture

- We are able to capture and store all of data provenance from users experiments
- We are close to having infrastructure, where we can reproduce data produced by user runs directly from command line
  - For example if user runs set of complex scripts to produce a dataset, we are capturing the entire environment and will be able to reproduce it on the same system
  - With some additional effort, we may be able to package this in a way where we could reproduce the results on different system
    - Perhaps enhance ReproZip package to run as a non-root user

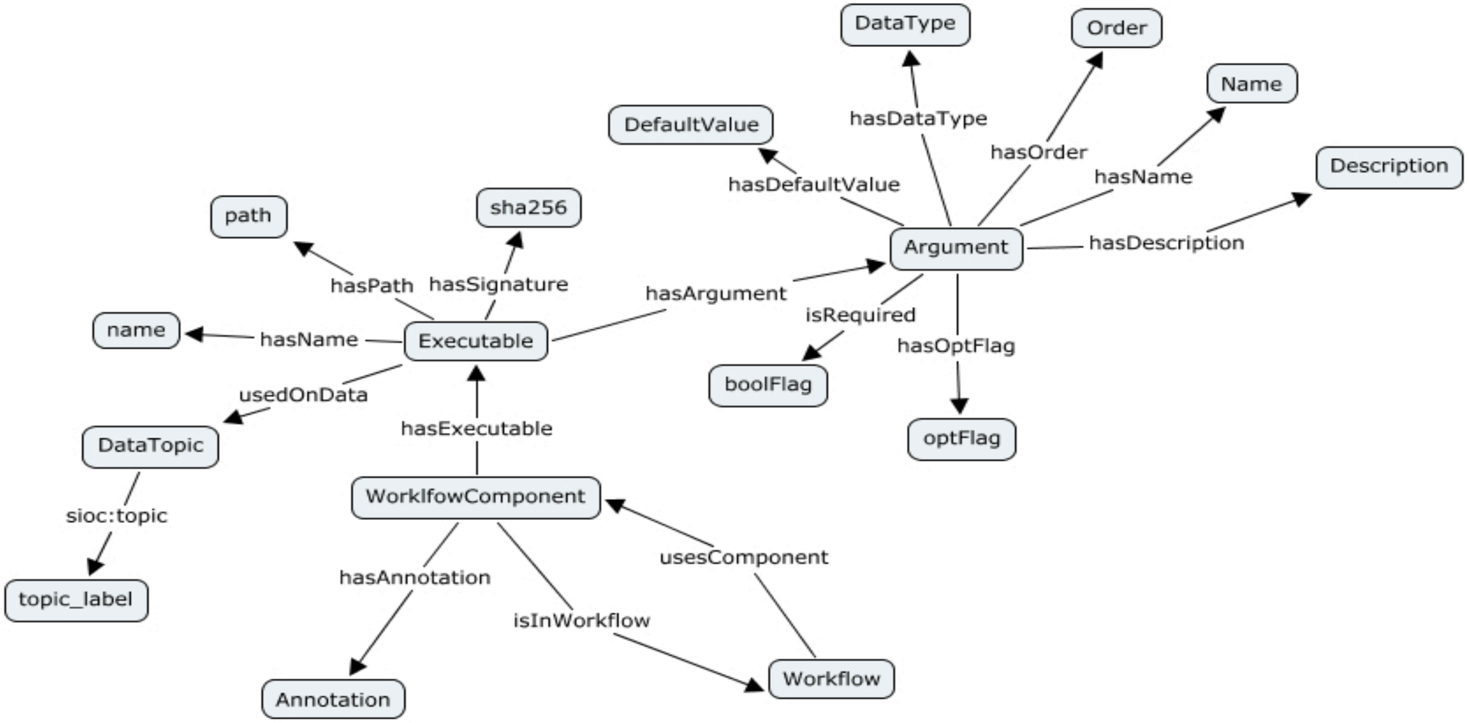
Immediate Practical Benefits of the Workflow Capture

- The current implementation provides a great way to do a fairly comprehensive provenance tracking of any process on NEX
  - Run process within the WPT program
    - Captures executables, all involved files + all involved libraries = great way to capture the environment settings and the dependencies
  - Stream results directly to a database/datastore
  - For every file in the database we can then search for:
    a) What exactly was involved in the production of the data?
    b) Compare if two files were generated with the same inputs, processes AND libraries (environments)
    c) If there are problems with execution, we can see what the difference is in terms of libraries and versions = great for debugging and moving code between different platforms (one of the goals of NEX)
- The tool is now accessible to all NAS supercomputing users through the loadable modules

Conversion Process

- Extract the executables for particular process from the provenance store
- Filter based on what is required
  - Some of the executables that are usually implied are also reported (sh, python, wpt, …)
- Look for a match in the *Component Repository*
  - So that we don't have to re-do the conversion
  - Interface is at works
- Convert to CLT format and insert into VisTrails path

Workflow Repository

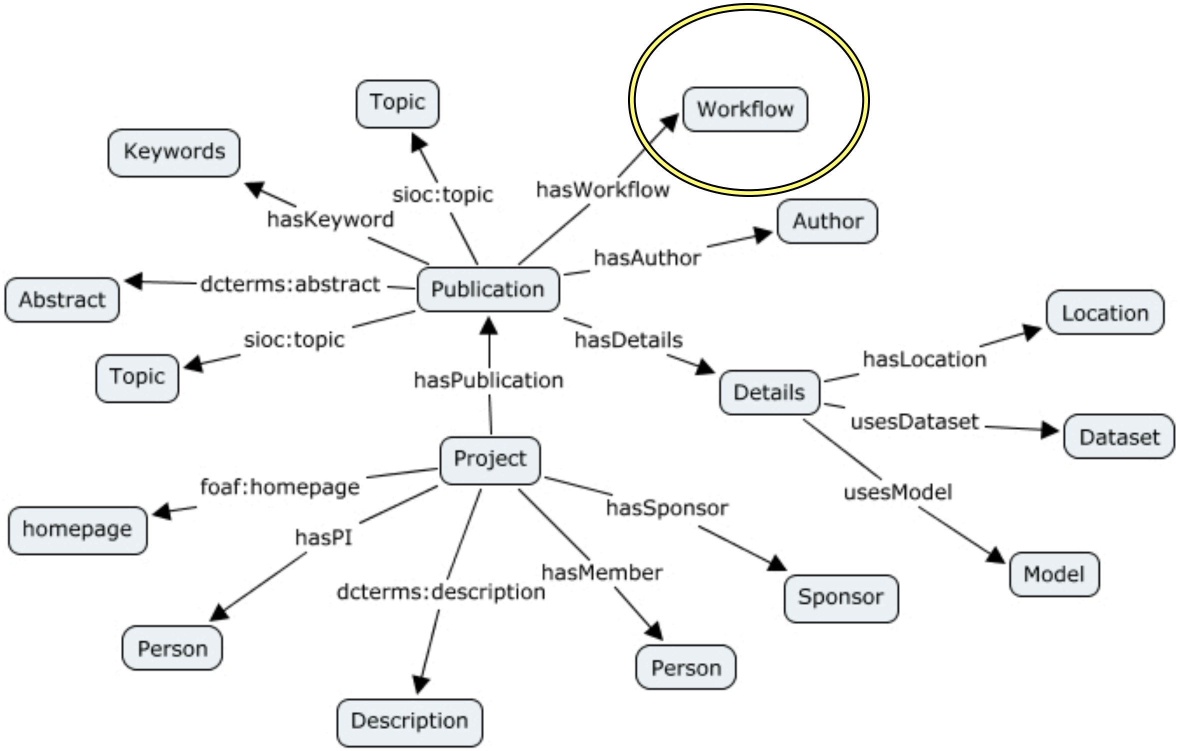


- RDF/OWL Ontology
  - Access through SPARQL and API interfaces
  - Currently using Virtuoso as a back-end

Workflow Repository - Integration

- Integrates with other parts of NEX knowledge base (projects, publications, data etc.)
- We can start asking more complex questions
  - What tools can be used with MODIS data?
    - Based on the knowledge that some projects were using particular tools to derive products from MODIS data.
  - What workflows were used to generate our Amazon 2010 publication?
  - What computational resources were needed to produce global Landsat WELD data for July 2010?

Example - Publication Integration



Publications can be traced back to workflows, projects, authors and even NASA programs

Status: Starting to populate the data store with publications and integrating with existing graphs – projects, GCMD, …

Workflow HPC Interface

- Working with VisTrails team in order to extend VisTrails interface to HPC
- Integrating “Suspended Module” with the NEX scheduler interface
  - Long running jobs (i.e. jobs submitted through the scheduler on Pleiades)
    - Implementation through a callback to PBS status
  - Upon restart, VisTrails keeps monitoring the current job and only proceeds when the job is completed
  - Gives us also ability to “chain” subsequent supercomputing jobs together
    - Perfect match for WELD – there are over 30 separate phases of the processing
- Further integration with the NEX supercomputing interface being developed under NASA CMAC project + VisTrails new parallel branch
  - Improves load balancing and ease of interaction with PBS scheduler

Test Case – WELD (Global Landsat Processing)

- NEX is the production facility for WELD Landsat processing
  - Products are developed and delivered to the community
  - Provenance (process capture) for WELD is really important
- Each month of WELD data required over 100,000 executions of various processing codes
  - That's 24,000,000 for single processing of 18 years
  - Plus another 200,000,000 log entries (performance metrics for analytics, error tracking, …)
  - 2.4 million Landsat scenes for a single processing of the 18-year record
- We must be able to tell where each of the piece of data came from and what algorithms were used
- Prototype testing with our workflow capture tool (we have tested it on other processes within HPC, but this is at very large scale)
  - Improved provenance retrieval and management
  - Improved performance estimates and analytics -> improved scheduling and more efficient processing – also feedback to the developers and scientists for code improvements
- The same process can be applied to other sensors and/or higher level products
  - Prototype large-scale data production pipeline

Relationship to NEX

- Thank to ESTO (AIST), which provides funding for low-level TRL development of technologies that are needed based on user feedback and requirements
- There is already established user community on NEX available for testing and rapid feedback
  - Since the requirements often come from the users, they have good reason to participate
- This leads to agile technology infusion
  - Test and integrate developed technologies rapidly as we go and try them out on full-scale problems in real-world environment
  - Fail fast – risk mitigation within the project itself
  - But also have a vehicle to integrate mature technologies upon exit
    - Code repositories that can be built on
    - Anomaly Detection toolkit (AIST 2008) was scaled up and is currently used in global studies (couple papers will come out in the next few months)
- Diverse architecture mix (HPC, single-core, GPU, large shared memory systems, quantum)
  - Can test, prototype and develop wide variety of technologies supporting diverse applications

Thank You

https://nex.nasa.gov