Climate Science Performance, Data and Productivity on Titan

Ben Mayer\textsuperscript{1}
Pat Worley\textsuperscript{1}
Rafael Ferreira da Silva\textsuperscript{2}
Abigail Gaddis\textsuperscript{1}

1 Oak Ridge National Laboratory
2 USC Information Sciences Institute
ACKNOWLEDGEMENTS

• Department of Energy for providing support through the Panorama and ACME projects

• Ashley Barker, Marcia Branstetter, Christopher Fuson, Judy Hill, Gideon Juve, Eric Lingerfelt, Vickie Lynch, Ross Miller, Matt Norman, Suzanne Parete-Koon, Duane Rosenberg and Brian Smith
Introduction

• Run and Data Management of Climate simulations

• Experiential phase is 6-9 months for high resolution

• Long period of time, likely can optimize
  – Shortening cycle time can have large impact on cost and ability to advance model capability

• Need to make measurements to know where to optimize
Productivity

• Looking at Current and Better Practices
• How Model and Machine are changing
• What the impact on science looks like
Terms

• ACME is the DOE branch of the CESM climate model.

• ACME is made up of component models (Land, Sea Ice, Ocean, Atmosphere, etc)

• Case – specific configuration of climate model (resolution, active models, parameters, etc)

• High Resolution – ~0.25° The T341 and NE120 resolutions are examples

• Low Resolution – 1° The T85 and NE30 resolutions are examples
Impact on Science

• Faster turn around time on experiment execution
  – Allows faster model development
  – Reduces cost of experiment

• Automation further reduces burden on scientists
  – Allows for higher resolution (spatial/temporal) models to be processed with ease
  – Allows analysis of the details – Hurricanes – monthly vs daily to hourly output
Complete ACME Workflow

DOE Earth System Model Testbed

Science Input
- Initialization Files
- Name List Files

Input Data Sets

Machine Config

Configuration UI + Rule engine to guide valid configs

Configure ESM Case or Ensemble

Build ESM

Run ESM

Output Data

Archive to Storage

Analysis (UV-CDAT)
- Exploratory Analysis
- Explanatory Analysis
- Diagnostics Generation

UV-CDAT & Dakota
- Uncertainty Quantification

Legend
- Monitoring & Provenance Dataflow (Simulation Manager)
- Dataset Dataflow ESGF
- User Driven Interaction
- Automated Workflow Process Control
- Process level Dataflow

Manually Provided File(s)

Model Source (svn/git)

Configuration Status

- Retrieve required Datasets
- Store manually provided files

- Build status

- ESM run status
- Store history files

- Configuration Information (Store and/or Retrieve)

Rapid, reliable, secure data transport and synchronization: Globus Online

Enterprise E2E Database
Enables Search/Discovery, Automated Reproducibility, Workflow Status, Monitoring Dashboard, Data Archive and Sharing

Simulation Manager & Provenance AKUNA + ProvEn

Data Archive ESGF

Single sign on and group management: Globus Nexus

Climate Change Science Institute
Oak Ridge National Laboratory
Progress Timing
Data Source

• Collecting many times of events and performance parameters automatically as the ACME model is run
• Data used from 2012 and 2014 to sample Jaguar and Titan
• Many different cases run, many of them for one to a few trips through the system
• Data presented here is from three cases that have many trips through the machine
  – Case will do similar amount of work each time (+/- I/O)
• Sample bias – Only capturing data from successful runs
Variability in Initialization time

![Graph showing variability in initialization time over dates from 4/1/12 to 3/17/13. The graph displays data points for t341 FAMIP, t85 FAMIP, and t341 F1850.]
Variability in Main Loop time

Variability in Main Loop time

Variability in Main Loop time

Variability in Main Loop time

Variability in Main Loop time

Variability in Main Loop time
Request Size vs Time in Queue
Date Submitted vs Queue wait time

Date of Submission

Queue Wait time (Sec)

- under 1k
- 1k to 10k
- 10k-20k
- 20k-100k+
Utilization of Requested time

DOE Earth System Model
Typical Queue Sequence

Create Job
Wait in Queue
Run Job

Allocated time

Unused Allocation time

Create Job
Ratio of Compute to Archive data

- Three projects with allocations: 33.5 Million hour (Mhr), 116Mhr and 50Mhr.
- Number of TB to Mhr of compute is: 4.8, 0.5, 5.0
- High ratios are from projects that were focused on production simulations
(Data) Rates of Post Processing Tasks

• Several post processing steps that are not in the time critical chain until end of simulation

• Processes are very human time intense

• Analysis – Two types for atmosphere
  – AMWG 6.5 hours
  – CDAT based 58.7

• Data Transfer between sites – Average 591Mbps with standard deviation of 621Mbps. N=38

• HPSS ingest (anecdote) about 200MB/s

• ESGF publication – Little data movement other than directory structure shuffle, but can take order day

• Data time/space inversion – not running yet
(Data) Rates of Post Processing Tasks

- Several post processing steps that are not in the time critical chain until end of simulation
- Processes are very human time intense
- Analysis – Two types for atmosphere
  - AMWG 6.45 hours
  - CDAT based 58.74
- Data Transfer between sites – Average 591Mbps with standard deviation of 621Mbps. N=38
  - Further testing shows high variability {2200, 220, 2150} Mbps
- HPSS ingest (anecdote) about 200MB/s
- ESGF publication – Little data movement other than directory structure shuffle, but can take order day
Model Changes

• New dynamics CAM4 to CAM5
  – CAM5 is 2-4x more computationally expensive

• Move from moderate resolution to High resolution (T85 or NE30 to NE120)
  – Should be 4x more resolution in each of the horizontal direction (16x)
  – Seeing 27x change in performance
  – Vertical levels are held constant
  – Output data increases by 16x, but is adjustable
Machine Changes

• With each generation moving to be 5-10x more capable

• Summit has a significant increase in theoretical compute performance, but a static file system performance.
  – Changes in technology likely will change ratio of delivered performance to magic fairy flops
Automation

• Data management is using 3-6 FTE in the project
• That is 7-14% of budget
• Using Pegasus Workflow Manager
• Early stages
Future Work

• Continue automation
• Adding more post processing tasks
• Adding more timing – easier with automation
• Build simple analytic model relating queue time to “rate of progress” and take measurements
• Capture failed instances