Cloud Computing @ JPL Science Data Systems

Emily Law
Jet Propulsion Laboratory, California Institute of Technology
ESIP 2012 Winter Meeting

Outline

• Science Data Systems (SDS)
• Space & Earth SDSs
• SDS Common Components
• Components using Cloud Computing
• Use Case 1: LMMP
• Use Case 2: ACCE
• Strategy
Science Data Systems (SDS)

- Cover a wide variety of domain disciplines
  - Solar system exploration, Astrophysics, Earth science, Biomedicine, etc,…
- Each has its own communities, standards and systems
- But, there is a set of common components
- Some can greatly benefit from proven cloud computing technologies
Components using Cloud Computing

Use Case 1: Lunar Mapping & Modeling Project (LMMP)

- Provide science and exploration community a suite of lunar mapping and modeling tools and products that support the lunar exploration activities
- The tools and products are made available through a common, intuitive NASA portal
- Publicly available since March 2012
- http://lmmp.nasa.gov

Copyright 2012 California Institute of Technology. Government sponsorship acknowledged.
Challenge

- The image files LMMP manages range from a few gigabytes to hundreds of gigabytes in size with new data arriving every day.
- Lunar surface images are too large to efficiently load and manipulate in memory.
- LMMP must make the data readily available in a timely manner for users to view and analyze.
- LMMP needs to accommodate large numbers of users with minimal latency.

Cloud Computing Solutions

- Slice a large image into many small images and merge and resize until the last merge and reduce yields a reasonably sized image that depicts the entire image.
- Amazon E2C/S3
- Used distributed approach with Elastic MapReduce to tile images.
- Developed a hybrid solution (multi-tiered data access approach) to serve images to users by cloud storage.

Copyright 2012 California Institute of Technology. Government sponsorship acknowledged.
Findings

• Computing performance
  • Comparable especially for the new machines with significant processing capability
  • EC2’s “rental” model offers better performance per dollar than having to purchase and maintain local servers

• Storage
  • Pay for just the bandwidth consumed
  • Eliminate the need to purchase extra hardware and bandwidth to handle the occasional spikes in usage

• Cloud Deployment
  • Increase latency
  • Enable fault tolerance

Use Case 2: Airborne Cloud Computing Environment (ACCE)

• Multi-mission capability providing distributed SDS services applicable to space-borne missions
  • File Management
  • Workflow Management
  • Resource Management

• Extend the existing services to utilize cloud services, commercial, community and private
  • Storage
  • Compute Resources
Approach

- Explore the benefits of performing science data processing for airborne missions in the cloud
- Evaluate different cloud technologies
  - Amazon EC2/S3
    - Elastic compute resources and on-demand storage
  - Eucalyptus
    - Infrastructure software for establishing a private cloud
  - Hadoop – Distributed File System (DFS) and MapReduce
    - Increased processing performance on large data sets

Cloud deployment

Copyright 2012 California Institute of Technology. Government sponsorship acknowledged.
Findings

- Processing cost reduction
  - No investment in capital required (upfront or refresh costs)
  - Pay only for what you use

- Challenge
  - Host Environment
  - Support for ITAR-sensitive data
  - Data transfer rates between JPL and commercial cloud
  - JPL Firewall

Take Away

- Many benefits
  - Accessible from anywhere
  - Increase/decrease number of machines based on user defined parameters
  -Resizable compute capacity for unlimited growth
  - Utility Computing, pay by the drink, rapidly provisioned

- But…
  - Complexity
  - Security
  - Reliability
  - Feasibility
  - Standards
• Cloud computing is an approach and tool to support “Big Data”

• Cloud Computing Working Groups
  • Common architecture
  • Common cloud framework
  • Lessons Learned
  • Best Practices

• Further benchmarking, optimization, research

• Collaboration