**Overview**

The InSAR Scientific Computing Environment (ISCE) is a modern framework for efficient processing of interferometric and polynomic synthetic aperture radar (SAR) data. ISCE was developed at NASA/JPL and Stanford under an AGS Advanced Information Processing Technology program effort that ended in 2012. Integration with cloud computing, which is continuing through an effort called Cloud Enabled Scientific Computing Research Environment (CISCRE), under the NASA Advanced Information Processing Technology program, provides ISCE with powerful computing capabilities.

The image processing for ISCE sensors that will enable scientists to reduce measurements directly from radar satellites and aircraft.

**ISCE Concept and Processing Issues**

- **ISCE Component Architecture**
  - Facilitates Cloud VM Configuration
  - Supports AWS S3, EC2, SWF, and other cloud services
  - Relies on local storage and compute capability
  - Requires unnecessary data transfer before processing
  - Growing at a staggering rate
  - Dispersed around the world
  - Hard to discover
  - Requires search of potential solutions, which can be elastically provisioned in the cloud, where the stored data are available on the local network.

**Traditional Scientific Computing Environments**

- **Common issues**
  - Hard to discover
  - Distributed via site code that has to be:
    - Configured with proper versions for dependencies
    - Requires complex installation process
    - Lack infrastructure for community contributions
    - Documentation, comments, and guidance
  - Science data is:
    - Hard to discover
    - Dispersed around the world
    - Requiring constant data transfer before processing
    - Requires access to a running machine into a virtual machine image that may not be portable
    - Streamlined installation and configuration
    - Execution of software in the environment configured typically to developers
    - Slowing down research results
    - Easy migration between different versions of software and configurations during configuration,
    - Can be executed on cloud or local infrastructure

**Cloud Computing**

- **Levels of abstraction between hardware & processing infrastructure**
  - Virtual images to capture multiple run-time configurations
  - Streamlined and consistent environment
  - Execution of software in the environment configured typically to developers
  - Slowing down research results
- **Easy migration between different versions of software and configurations during configuration,**
  - Can be executed on cloud or local infrastructure

**Use of Cloud in Design of SAR Mission**

- **Parametric studies of DSHYD SAR antenna using the GRASP antenna analysis software were posted to cloud**
  - Execution of 62-element array was optimized from 4 hours to 10 minutes. This experiment demonstrates effective use of cloud computing to streamline data processing and analysis.

**Future Work**

- **Virtual Machine Image Catalog to discover algorithms**
  - Science data coming to discover raw and processed data as well as to creation parameters
- **Benchmarking cluster compute instances, SSDS, and high-I/O instances for ISCE processing**
- **Evaluate archived storage of ISCE data in cold storage**
- **Community infusion of ISCE and improved documentation**
  - This broad vision of interoperability to minimize data transfer to analyze results
- **Development of CISCRE**
  - Extended analysis for cloud computing and comparing with traditional environment
  - For cost of large scale processing (e.g. 50K ALOS scenes)
- **Large storage, backup, and archival costs**
- **Preliminary results demonstrate cloud computing is cost-effective**

**Interferometric Synthetic Aperture Radar (InSAR) Scientific Computing Environment on the Cloud**

Paul A. Rosen, Eric Gurdzo, Kajwaha Slams, Brett George, David Knight
Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109

**Interferogram Generation**

The ISCE workbench is configured so-called ALOS PALSAR images acquired over the Los Angeles area spanning 5 years. We formed all possible interferograms. 122 interferograms, containing 478 pairs of raw data, 53 was used as a distributed filesystem and subsequent storage of results. This task initially took 3 weeks on local infrastructure.

**Deployment of Cloud Computing to Streamline Data Processing.**

Deployed in production for:

- **MRO:** Data archival project
- **MER:** Data archival project
- **MSL:** every image is processed through Polyphony

**ISCE Component Architecture**

- **Facilitates Cloud VM Configuration**
  - Supports AWS S3, EC2, SWF, and other cloud services
  - Relies on local storage and compute capability
  - Requires unnecessary data transfer before processing
  - Growing at a staggering rate
  - Dispersed around the world
  - Hard to discover
  - Requires search of potential solutions, which can be elastically provisioned in the cloud, where the stored data are available on the local network.

**Polyphony**

- **Pipelined-on-streamline parallel workflow development based on simple松 workflow framework**
  - Modular framework supporting:
    - Highly optimized data transfer
    - Dynamic provisioning of machines based on load
    - Distributed life system to scale clusters
  - Automatically process with nearly 30k across ALOS DICE instances
  - Workflows in Foster, C. C., Python, Jero, & Nido de
  - Deployed in production for:
    - MRO images reprocessed through Polyphony
    - CARIS (Carbon in the Arctic Reserve Vulnerability Experiment), La Thuile, Austria
    - Multiple other workflows and applications
  - Numerous staff scale executions across NASA

**Supertech A86, IX2, SWF, Domestic, and Supertech**

**CISCRE**

Collection of Storage and Cost-Effective Compute Capacity: Cloud computing enables us to collect large-scale storage with elastic compute capacity. This obviates local storage or discovering the data into the local infrastructure. Compute capacity can be elastically provisioned in the cloud, where the stored data are available on the local network.

**Paraview-based Processing**

- **CISCRE utilizes modules and can be customized with pre-configured machines with all the software and dependencies required by their environments.**
  - Pre-configured modules with some necessary features and dependencies for provisioning a machine and sharing the processing.
  - Paraview-based modules and cloud orchestration: CISCRE integrates capabilities with Polyphony, to enable scientists to obtain their results faster by concurrently employing large number of machines.
  - This enables scientists to cost-effectively maximize output productivity.

**Collaboration**

- **Collaboration of data enables sharing of raw data as well as higher-level products.**
  - The pre-configured machine images enable software developers and research scientists to share their implementations and compute results.
  - This significantly reduces the overhead currently associated with sharing and evaluating new algorithms.
  - The parallelization of algorithms in a cloud environment will enable distributed processing and lead to effective processing of large datasets. It will further facilitate efficient sharing of computational resources across the NASA community.

**Cloud or Local Infrastructure?**

- **In the conventional model data is transferred from DAAC to local storage and subsequently transferred to a running machine.**
  - The cloud computing allows the ability to store the data with the data in the cloud.
  - This reduces the latency and allows for faster processing, which can be paid as much as 30%.

**Management of Cloud Resources**

- **Cloud resources are charged in 1 hour increments**
  - Embarrassingly parallel problems can be processed faster by provisioning large capacity.
  - Costs can be optimized by utilizing up to 1 hour slots and using bid-based pricing.
  - Spot instances on AWS offer bid-based computing:
    - Users bid for how much they are willing to pay for compute
    - Price can vary as low as 0.01 per hour
  - Top 500 supercomputers can be provisioned for 100hrs (i.e. 120,000 cores, 10TB of storage)
  - Future research includes cost optimization through use of spot
  - Cloud offers long term archival storage options for as low as 1 GB/month.

**Results**

- **$256 – 4 hours**
  - This nets performed over a year ago using the AWS
  - Never cloud resources with faster processors, solid state disks, and on- and off-based budgets would provide greater speed and cost savings

**Citation**

Shams, Paul A. Technology Transfer. 2011.

**Acknowledgement**

The authors would like to thank the Earth Science Technology Office and High End Computing Program at NASA for support. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology. NASA, Copyright © 2012 California Institute of Technology. Government sponsored acknowledgement. All Rights Reserved.

**Table**

<table>
<thead>
<tr>
<th>Processing Steps</th>
<th>2</th>
<th>35</th>
<th>20 Gb</th>
<th>0.25</th>
<th>110 Gb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenes</td>
<td>325</td>
<td>735</td>
<td>Tb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nodes</td>
<td>62</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Execution Parameters</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resources</td>
<td>466,667</td>
<td>21,875</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>scenes = 250 Tb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure**

- **Cloud Computing Flow**
  - The ISCE framework is an open-source scientific computing platform designed to facilitate the processing of space-borne SAR data.
  - It supports the entire SAR data processing workflow, from data acquisition to final product delivery.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.
  - The ISCE framework is designed to be flexible, scalable, and extensible, allowing scientists to easily integrate new processing steps and algorithms.

**Alongside Cloud Computing Models**

- **Virtualized Cloud Compute Nodes**
  - **Cloud Pervasive Storage**
  - **Virtualized Cloud Compute Nodes**
  - **Cloud Computing Services**

**References**

Guzella, Eric, Gian Franco Saucio, Paul A. Rosen, and Howard Ed...