Characterization of Scientific Workflows

Shishir Bharathi   Ann Chervenak   Ewa Deelman
Gaurang Mehta   Mei-Hui Su   Karan Vahi

University of Southern California
Information Sciences Institute
Marina del Rey, CA

November 17, 2008

Motivation

- Extensive research in Workflow Systems over past few years.
  - New planners, workflow execution engines, scheduling algorithms
- Comparison of various systems requires good benchmarks.
- Dearth of freely available workflow applications.
- Montage is a popular example used by many researchers.
- Need to find out how a system/algorithm performs with other workflows.
- Some researchers have used randomly generated workflows.
  - Modeled workloads using "realistic data" from the Parallel Workload Archive, etc.

Motivation

- Extensive research in Workflow Systems over past few years.
  - New planners, workflow execution engines, scheduling algorithms
- Comparison of various systems requires good benchmarks.
- Dearth of freely available workflow applications.
- Montage is a popular example used by many researchers.
- Need to find out how a system/algorithm performs with other workflows.
- Some researchers have used randomly generated workflows.
  - Modeled workloads using "realistic data" from the Parallel Workload Archive, etc.

Not realistic enough for scientific workflows.

Objective

- Make available different types of scientific workflows to the community.
- Must cover a broad range of scale in terms of number of jobs, data processed.

Objective

- Make available different types of scientific workflows to the community.
- Must cover a broad range of scale in terms of number of jobs, data processed.
- Characterized workflows from different scientific communities.
- Identified basic structures common to many applications.
- Used information to generate synthetic but realistic workflows.

Outline

- Basic workflow structures.
- Characterization of five scientific workflows.
- Workflow generator.
- Workflow repository.
- Future directions.

Basic workflow structures

Process   Pipeline   Data Distribution   Data Aggregation   Data Redistribution

Job   Data
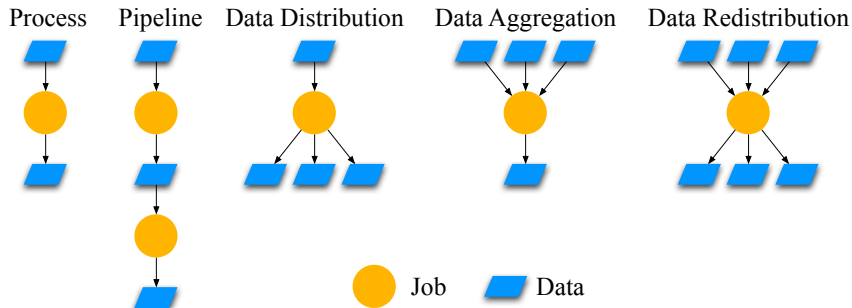
Montage - an astronomy application

- Created by NASA/IPAC Infrared Science Archive.
- Open source toolkit.
- Creates a mosaic of the sky from several input images.
- Inputs are re-projected to be of same spatial scale and rotation.
- Background emissions are corrected to be of the same level in all images.
- Re-projected, corrected images are co-added to form mosaic.

Montage Workflow
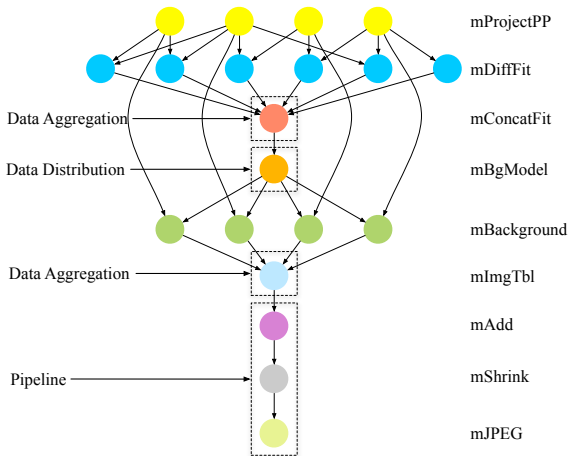
- Number of inputs may vary over time for the same final image.
- mBgModel more of a distribution job and not partitioning job.
- mAdd job most computationally intensive.

Montage execution profile

Executed on viz-cluster at ISI - 8 nodes with dual Intel Xeon 2.4 GHz processors, 2 GB Memory

Job	Count	Runtime		Inputs		Outputs	
		Mean(s)	Variance	Mean(Bytes)	Variance	Mean(Bytes)	Variance
mProject	45	13.59	0.06	4222384.00	0	8324854.00	4.6e+08
mDiffFit	107	10.59	0.01	16646919.33	1.5e+09	568468.13	8.7e+10
mConcatFit	1	13.60	0.00	32207.00	0	22465.00	0
mBgModel	1	10.88	0.00	30005.00	0	2393.00	0
mBackground	45	10.74	0.03	8334797.00	4.6e+08	8324864.00	4.6e+08
mImgtbl	1	10.69	0.00	374626420.00	0	12851.00	0
mAdd	1	30.34	0.00	374632035.00	0	346930560.00	0
mShrink	1	12.26	0.00	173465280.00	0	6940800.00	0
mJPEG	1	10.96	0.00	6940800.00	0	339858.00	0

- Number of jobs of each type.
- Runtime statistics for each job type.
- Total sizes of inputs consumed and outputs produced.

CyberShake - Earthquake science workflow

- Created by the Southern California Earthquake Center (SCEC).
- Characterizes earthquake hazards using Probabilistic Seismic Hazard Analysis (PSHA) technique.
- MPI based simulation generates Strain Green Tensors for region of interest.
- Synthetic seismograms are calculated for each rupture that was predicted.
- Spectral acceleration and probabilistic hazard curves are then generated.

CyberShake Workflow
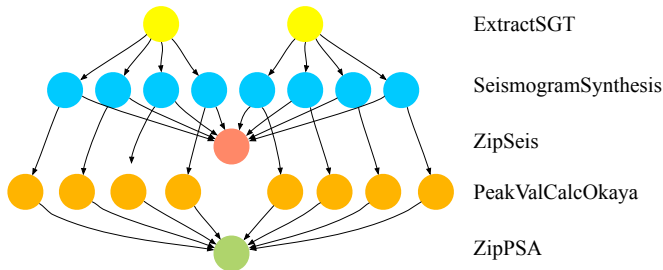
- Highly parallel in structure.
- SeismogramSynthesis jobs are most computationally intensive.
- If a large number of $\langle source, rupture \rangle$ pairs are analyzed, ExtractSGT and the Zip jobs may also need more time on compute resources.

Genome sequencing - Epigenomics application

- Modeled as a data processing pipeline by USC Epigenomics center.
- DNA sequence data generated by Illumina-Solexa Genetic Analyzer.
- Split into several chunks operated on in parallel.
- Converted into format used by the Maq sequencing system.
- Noisy and contaminating sequences filtered out.
- Remaining sequences mapped in a reference genome.

Epigenomics Workflow

fastQSplit

filterContams

sol2sanger

fastq2bfq

map

mapMerge

maqIndex

pileup
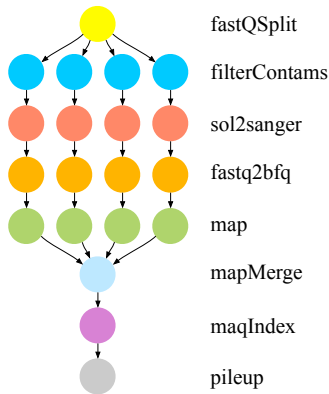
- Largely pipelined application.
- Sequence data obtained from multiple "lanes" are split into several chunks.
- Filtered, formatted data is aligned by the "map" job which are the most computationally intensive.

Inspiral Analysis Workflow - Gravitational Physics

- Created by the Laser Interferometer Gravitational Wave Observatory (LIGO).
- Detection of gravitational waves produced by various events in the universe.
- Analysis of data obtained from the coalescing of compact binary systems.
- Time-frequency data split into smaller blocks.
- Waveforms belonging to the parameter space generated for each block.
- Matched filter output computed and trigger generated if inspiral detected.

Inspiral Analysis Workflow
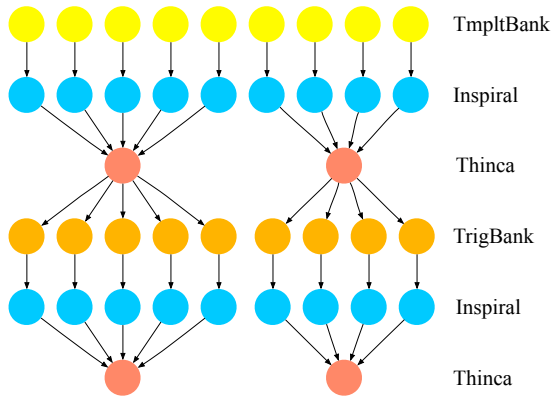
- Simple structure allowing for greater parallelism.
- Inspiral jobs are the most computationally intensive.
- Triggers produced by the Inspiral analysis jobs are tested for consistency by the Thinca jobs.
- Template banks can be generated from the trigger outputs.

SIPHT - search for sRNAs

- Created by the bioinformatics project at Harvard.
- Kingdom-wide search for small untranslated RNAs that regualate several processes in bacteria.
- Search for sRNA encoding genes for all of the bacterial replicons in the NCBI database.
- Prediction of Rho-independent transcriptional terminators.
- BLAST (Basic Local Alignment Search Tools) comparison of inter-genetic regions.
- Annotations of sRNAs that are found.

SIPHT Workflow
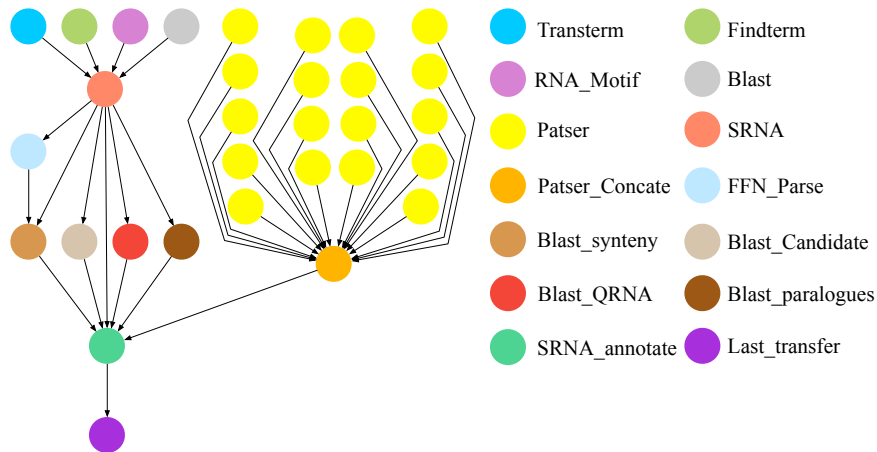
- Relatively fixed structure.
- BLAST jobs and sRNA prediction is also compute intensive.

Workflow generator

- Workflows generated using information from actual executions of scientific workflows.
- Supports various parameters for each workflow type allowing for the generation of workflows of different scale.
- Output in DAX (Directed Acyclic Graph in XML) format.
- Includes annotations for job runtimes and sizes of data items.

```
<job id="ID00000" namespace="Montage" name="mProjectPP" version="1.0" runtime="8.98">
<uses file="region.hdr" link="input" size="306"/>
<uses file="2mass-atlas-ID00000s-jID00000.fits" link="input" size="2111040"/>
...
</job>
```

Workflow generator

- Workflows generated using information from actual executions of scientific workflows.
- Supports various parameters for each workflow type allowing for the generation of workflows of different scale.
- Output in DAX (Directed Acyclic Graph in XML) format.
- Includes annotations for job runtimes and sizes of data items.

```
<job id="ID00000" namespace="Montage" name="mProjectPP" version="1.0" runtime="8.98">
<uses file="region.hdr" link="input" size="306"/>
<uses file="2mass-atlas-ID00000s-jID00000.fits" link="input" size="2111040"/>
...
</job>
```

- Workflow can be simulated or executed on real Grids.
  - We plan to provide a client program that can simulate the execution of each job in the workflow.
  - Standard UNIX utilities can be used to create synthetic data files that can then be transferred to and from Grid sites.

Workflow generation steps - Montage example

- Identification of individual jobs and their compositions.
  - If degree is specified, estimate number of input images and choose the number of images that overlap.
  - Generate one mProject job for each input image and one mDiffFit job for each pair of input images that overlap.
  - Include other jobs as necessary.

Workflow generation steps - Montage example

- Identification of individual jobs and their compositions.
  - If degree is specified, estimate number of input images and choose the number of images that overlap.
  - Generate one mProject job for each input image and one mDiffFit job for each pair of input images that overlap.
  - Include other jobs as necessary.
- Annotation of workflow with job runtimes and data sizes.
  - Often input images have constant size. Therefore, processing time is also picked from a single distribution.
  - Sizes of intermediate data products are estimated from the number and types of inputs being processed.
  - Scale of the image produced is also taken into account.
  - Runtimes are scaled appropriately using estimates from actual executions.

Workflow repository

- Created an initial library of synthetic scientific workflows.
- Provides benchmarks to research communities to evaluate tools and systems.
- Also includes statistics from executions of workflows of each type on the Grid.

http://vtcpc.isi.edu/pegasus/index.php/WorkflowGenerator

Future directions

- Expand on workflows in the repository.
- Collect more statistics about current workflows and feed them back to the generator.
- We invite scientific communities to contribute information about their applications.
  - Workflow structure.
  - Data sizes and job runtimes.
  - Algorithms/codes do not need to be contributed.

> Contacts: shishir@isi.edu, Pegasus team.