Facilitating Distributed Algorithm Execution in a Grid Framework

ISP AI Forum Talk

Yuriy Sverchkov

Intelligent Systems Program

Adviser: Dr. Rich Tsui
RODS Lab
Department of Biomedical Informatics
University of Pittsburgh

November 6, 2009

Layout

Motivation

- Biomedical applications

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection
- Multiple organizations

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection
- Multiple organizations
  - Data collection

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection
- Multiple organizations
  - Data collection – hospitals, health-care providers

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection
- Multiple organizations
  - Data collection – hospitals, health-care providers
  - Data aggregation

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection
- Multiple organizations
  - Data collection – hospitals, health-care providers
  - Data aggregation – health departments

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection
- Multiple organizations
  - Data collection – hospitals, health-care providers
  - Data aggregation – health departments
  - Data analysis

Motivation

- Biomedical applications
  - Biosurveillance
  - Outbreak detection
- Multiple organizations
  - Data collection – hospitals, health-care providers
  - Data aggregation – health departments
  - Data analysis – researchers in health departments and universities

What is the Grid?

- Coordinated sharing of data and computational resources
- Collaboration between multiple institution
- Non-centralized (each institution maintains autonomous control over its own resources)

"The real and specific problem that underlies the Grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations."

I. Foster, C. Kesselman, S. Tuecke.

The Anatomy of the Grid: Enabling Scalable Virtual Organizations.

International J. Supercomputer Applications, 15(3), 2001.

Applications

- Computation-driven Grids
  - Teragrid
  - Enabling Grids for E-sciencE (EGEE)
  - Open Science Grid (OSG)
- Data-driven Grids
  - caGrid
  - MedGrid

Grid Middleware

Various solutions

- Globus
- gLite
- UNICORE
- VDT

Grid Middleware

Various solutions

- Globus
- gLite
- UNICORE
- VDT

The Globus Toolkit

- **WebServices**: Standardized interfaces for data queries and analytical services.
- **WS-GRAM**: Grid Resource Allocation Manager
- **MDS**: Monitoring and Discovery Service

Layout

Why distribute jobs?

Computational challenges

- Outbreak detection algorithms are required to process ever-increasing amounts of data
- Information is time-sensitive

Why distribute jobs?

Computational challenges

- Outbreak detection algorithms are required to process ever-increasing amounts of data
- Information is time-sensitive

- Decrease computation time by using distributed algorithms

Why distribute jobs?

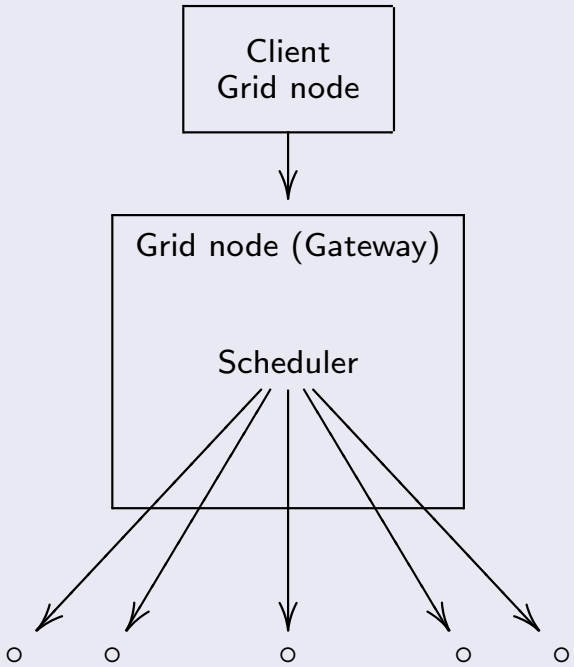
Computational challenges

- Outbreak detection algorithms are required to process ever-increasing amounts of data
- Information is time-sensitive

- Decrease computation time by using distributed algorithms
- Take advantage of the Grid framework
  - Secure data sharing
  - Secure sharing of computational resources

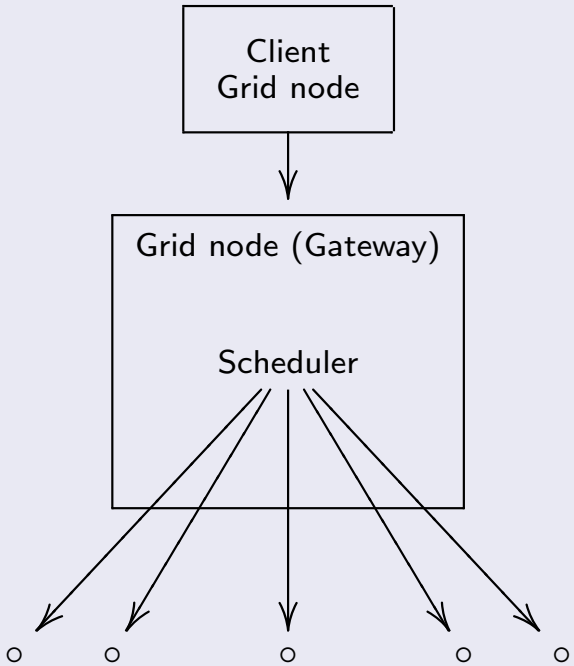
Job Distribution

Typical setup

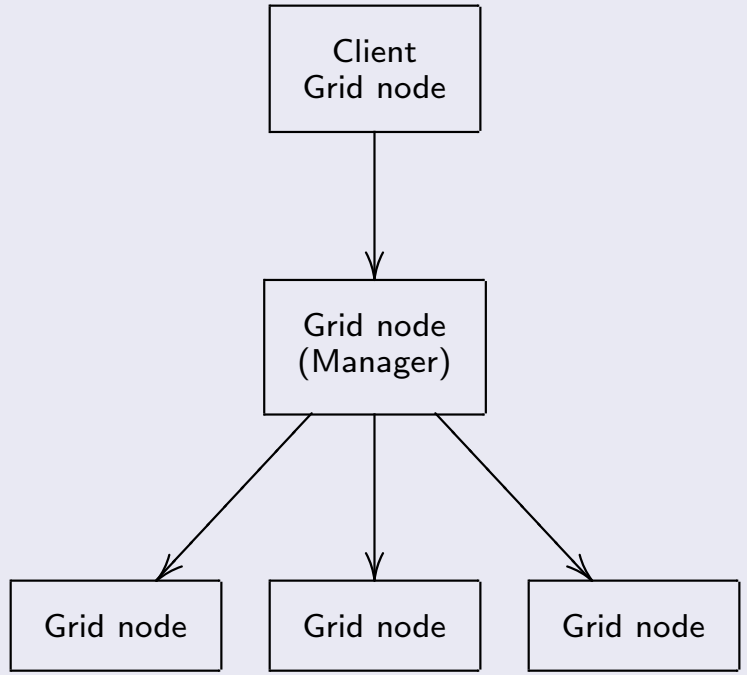


Job Distribution

Typical setup



Grid-driven setup



Previous Work

Previous approach by Tsai et. al.

- Developed ADMS
- Job division performed explicitly by the management service specified before jobs are sent for execution
- Simple and has little overhead
- Have to know information about available resources
- Algorithm must be set up on computation machines

Our approach

- Used existing software (Condor-Glidein)
- Queue-based
- Utilizes a full-blown job scheduler
- Flexible, allows adding and removing resources during execution
- Can take advantage of staging

Ming-Chi Tsai, Fu-Chiang Tsui, Michael M. Wagner

An Evaluation of Biosurveillance Grid – Dynamic Algorithm Distribution Across Multiple Computer Nodes

AMIA Annual Symposium Proceedings, 2007: 746-750.

Previous Work

Previous approach by Tsai et. al.

- Developed ADMS
- Job division performed explicitly by the management service specified before jobs are sent for execution
- Simple and has little overhead
- Have to know information about available resources
- Algorithm must be set up on computation machines

Our approach

- Used existing software (Condor-Glidein)
- Queue-based
- Utilizes a full-blown job scheduler
- Flexible, allows adding and removing resources during execution
- Can take advantage of staging

Ming-Chi Tsai, Fu-Chiang Tsui, Michael M. Wagner

An Evaluation of Biosurveillance Grid – Dynamic Algorithm Distribution Across Multiple Computer Nodes

AMIA Annual Symposium Proceedings, 2007: 746-750.

Previous Work

Previous approach by Tsai et. al.

- Developed ADMS
- Job division performed explicitly by the management service specified before jobs are sent for execution
- Simple and has little overhead
- Have to know information about available resources
- Algorithm must be set up on computation machines

Our approach

- Used existing software (Condor-Glidein)
- Queue-based
- Utilizes a full-blown job scheduler
- Flexible, allows adding and removing resources during execution
- Can take advantage of staging

Ming-Chi Tsai, Fu-Chiang Tsui, Michael M. Wagner

An Evaluation of Biosurveillance Grid – Dynamic Algorithm Distribution Across Multiple Computer Nodes

AMIA Annual Symposium Proceedings, 2007: 746-750.

Previous Work

Previous approach by Tsai et. al.

- Developed ADMS
- Job division performed explicitly by the management service specified before jobs are sent for execution
- Simple and has little overhead
- Have to know information about available resources
- Algorithm must be set up on computation machines

Our approach

- Used existing software (Condor-Glidein)
- Queue-based
- Utilizes a full-blown job scheduler
- Flexible, allows adding and removing resources during execution
- Can take advantage of staging

Ming-Chi Tsai, Fu-Chiang Tsui, Michael M. Wagner

An Evaluation of Biosurveillance Grid – Dynamic Algorithm Distribution Across Multiple Computer Nodes

AMIA Annual Symposium Proceedings, 2007: 746-750.

Previous Work

Previous approach by Tsai et. al.

- Developed ADMS
- Job division performed explicitly by the management service specified before jobs are sent for execution
- Simple and has little overhead
- Have to know information about available resources
- Algorithm must be set up on computation machines

Our approach

- Used existing software (Condor-Glidein)
- Queue-based
- Utilizes a full-blown job scheduler
- Flexible, allows adding and removing resources during execution
- Can take advantage of staging

Ming-Chi Tsai, Fu-Chiang Tsui, Michael M. Wagner

An Evaluation of Biosurveillance Grid – Dynamic Algorithm Distribution Across Multiple Computer Nodes
AMIA Annual Symposium Proceedings, 2007: 746-750.

Previous Work

Previous approach by Tsai et. al.

- Developed ADMS
- Job division performed explicitly by the management service specified before jobs are sent for execution
- Simple and has little overhead
- Have to know information about available resources
- Algorithm must be set up on computation machines

Our approach

- Used existing software (Condor-Glidein)
- Queue-based
- Utilizes a full-blown job scheduler
- Flexible, allows adding and removing resources during execution
- Can take advantage of staging

Ming-Chi Tsai, Fu-Chiang Tsui, Michael M. Wagner

An Evaluation of Biosurveillance Grid – Dynamic Algorithm Distribution Across Multiple Computer Nodes

AMIA Annual Symposium Proceedings, 2007: 746-750.

Job Scheduling

Job Schedulers

- Distribute jobs to computational units
- Monitor jobs
- Pause/resume jobs
- Cancel jobs
- Staging and cleanup
- Manage input and output

Various solutions that are officially supported by GRAM:

- Portable Batch System (PBS)
- Condor High-Throughput Computing System
- Platform LSF

Another option is the Sun Grid Engine

Job Scheduling

Job Schedulers

Various solutions that are

We use Condor

- Support for major platforms
- Supports heterogeneous pools of machines
- Highly configurable at the
  - Pool level
  - Machine level
  - Job level
- Can manage dedicated machines, cycle scavenge, or a mixture of both.
- Glidein – a means of adding Grid resources to a pool.

Condor-Glidein

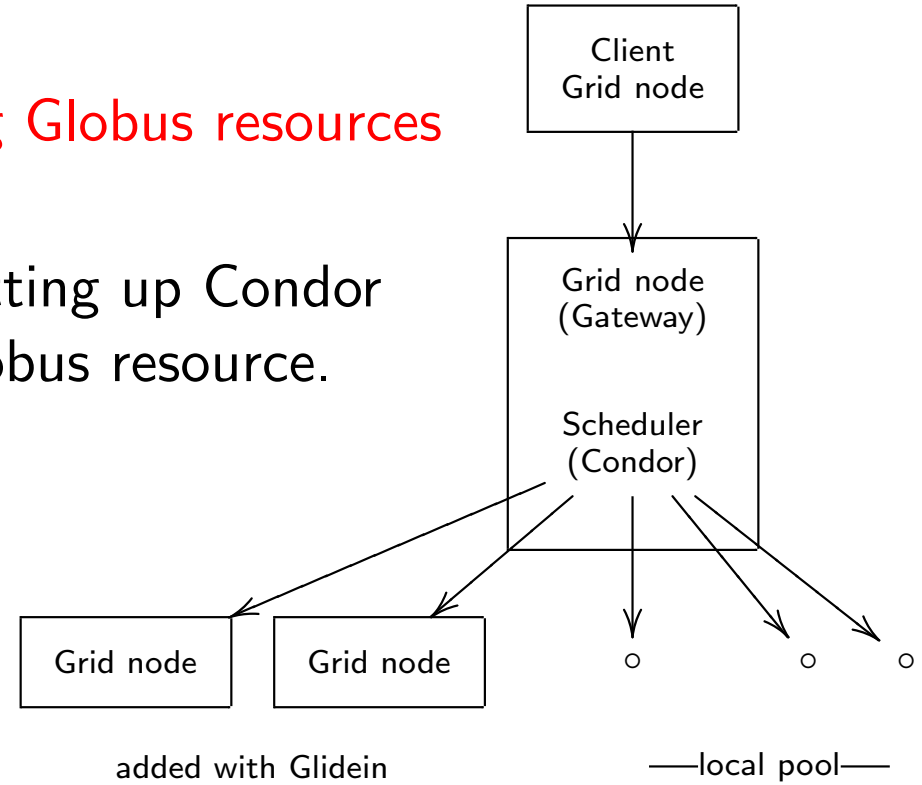
- A component of Condor

Condor-Glidein

- A component of Condor
- Allows (temporarily) adding Globus resources into the pool.

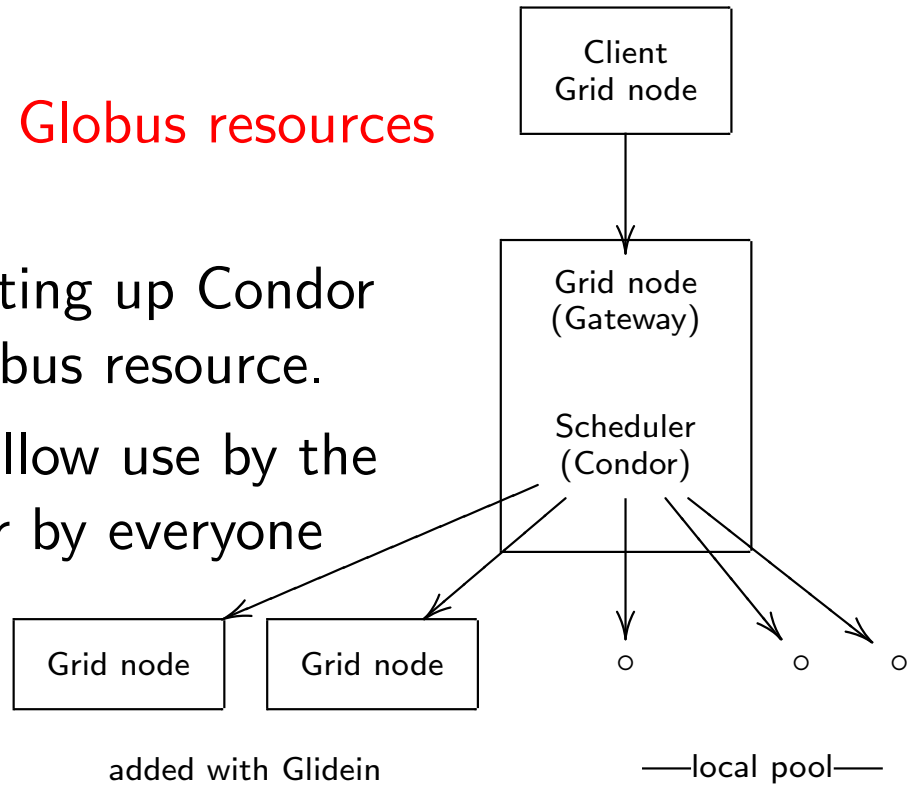
Condor-Glidein

- A component of Condor
- Allows (temporarily) adding Globus resources into the pool.
- Operates by temporarily setting up Condor daemons on the remote Globus resource.



Condor-Glidein

- A component of Condor
- Allows (temporarily) adding Globus resources into the pool.
- Operates by temporarily setting up Condor daemons on the remote Globus resource.
- Can be configured to only allow use by the user who invoked Glidein, or by everyone using the pool.



Layout

Software Distribution (Motivation)

There are many existing algorithms for various biomedical applications.
What is the current practice for biomedical researchers?

Software Distribution (Motivation)

There are many existing algorithms for various biomedical applications.
What is the current practice for biomedical researchers?

- Often people need to do one of the following:
  - implement their own version of an algorithm

Software Distribution (Motivation)

There are many existing algorithms for various biomedical applications.
What is the current practice for biomedical researchers?

- Often people need to do one of the following:
  - implement their own version of an algorithm
  - acquire their own copies to be able to use them

Software Distribution (Motivation)

There are many existing algorithms for various biomedical applications.
What is the current practice for biomedical researchers?

- Often people need to do one of the following:
  - implement their own version of an algorithm
  - acquire their own copies to be able to use them
  - send data to the owner of the algorithm

Software Distribution (Motivation)

There are many existing algorithms for various biomedical applications.
What is the current practice for biomedical researchers?

- Often people need to do one of the following:
  - implement their own version of an algorithm
  - acquire their own copies to be able to use them
  - send data to the owner of the algorithm

- What if the people holding the algorithms lack the computational resources to process your data?

Software Distribution (Motivation)

There are many existing algorithms for various biomedical applications.
What is the current practice for biomedical researchers?

- Often people need to do one of the following:
  - implement their own version of an algorithm
  - acquire their own copies to be able to use them
  - send data to the owner of the algorithm
  - the owner needs to set up a web-service
- What if the people holding the algorithms lack the computational resources to process your data?

Software Distribution (Motivation)

There are many existing algorithms for various biomedical applications.
What is the current practice for biomedical researchers?

- Often people need to do one of the following:
  - implement their own version of an algorithm
  - acquire their own copies to be able to use them
  - send data to the owner of the algorithm
  - the owner needs to set up a web-service
- What if the people holding the algorithms lack the computational resources to process your data?
- Information about where to get algorithms is external to the Grid.

Issues

Information

The algorithm to process the data exists, but researchers don't know where to get it.

Issues

Information

The algorithm to process the data exists, but researchers don't know where to get it.

Security

- Sensitive Data: Data provider can't send the data.

Issues

Information

The algorithm to process the data exists, but researchers don't know where to get it.

Security

- Sensitive Data: Data provider can't send the data.
- Proprietary/Patented Algorithms: Algorithm provider can't send algorithm.

Issues

Information

The algorithm to process the data exists, but researchers don't know where to get it.

Security

- Sensitive Data: Data provider can't send the data.
- Proprietary/Patented Algorithms: Algorithm provider can't send algorithm.

More generally, issues of colocation.

Approach to the information issue

We present an information service that allows Grid users to publish and discover information about the location of algorithms and software using a Grid webservice.

Approach to the information issue

We present an information service that allows Grid users to publish and discover information about the location of algorithms and software using a Grid webservice.

A Grid WebService to

- Present a list of available algorithms
  - Identifying information
  - Description of algorithms
  - Execution information
  - URI(s) of the software

Approach to the information issue

We present an information service that allows Grid users to publish and discover information about the location of algorithms and software using a Grid webservice.

A Grid WebService to

- Present a list of available algorithms
  - Identifying information
  - Description of algorithms
  - Execution information
  - URI(s) of the software
- Allow for a hierarchy or network of information services

Downloadable Algorithm Information Service

Data queries

- List all algorithms
- List algorithms with a set of attributes matching a set of values
- List algorithms with a keyword appearing in a set of attributes

Downloadable Algorithm Information Service

Data queries

- List all algorithms
- List algorithms with a set of attributes matching a set of values
- List algorithms with a keyword appearing in a set of attributes

Administrative operations

- Add an algorithm
- Remove an algorithm
- Add this information node as a subnode of another node
- Remove this information subnode

Layout

Downloadable Algorithm Information Service: Future Applications

Due to the standards that Grid services follow, they can easily be made to interact with other webservices and client applications.

Downloadable Algorithm Information Service: Future Applications

Due to the standards that Grid services follow, they can easily be made to interact with other webservices and client applications.

- Easily transform XML output to human-readable form to be presented to users of the information service.

Downloadable Algorithm Information Service: Future Applications

Due to the standards that Grid services follow, they can easily be made to interact with other webservices and client applications.

- Easily transform XML output to human-readable form to be presented to users of the information service.
- Build a service that utilizes this service, GridFTP, and GRAM to download and execute the algorithm.

Downloadable Algorithm Information Service: Future Applications

Due to the standards that Grid services follow, they can easily be made to interact with other webservices and client applications.

- Easily transform XML output to human-readable form to be presented to users of the information service.
- Build a service that utilizes this service, GridFTP, and GRAM to download and execute the algorithm.
- Integrate with workflow services to include algorithm download and execution in a workflow.

Implications for the security issue

Recall the security concerns

- Sensitive Data: Data provider can't send the data.
- Proprietary/Patented Algorithms: Algorithm provider can't send algorithm.

Directly enabled by the information service:

- Get the algorithms to the party with the data.
- Use Grid security to only allow certain parties to see certain information.
- Only allow secure downloads, using GridFTP or other secure protocols.

Future work: certificate-protected exectution.

Implications for the security issue

Recall the security concerns

- Sensitive Data: Data provider can't send the data.
- Proprietary/Patented Algorithms: Algorithm provider can't send algorithm.

Directly enabled by the information service:

- Get the algorithms to the party with the data.
- Use Grid security to only allow certain parties to see certain information.
- Only allow secure downloads, using GridFTP or other secure protocols.

Future work: certificate-protected exectution.

Implications for the security issue

Recall the security concerns

- Sensitive Data: Data provider can't send the data.
- Proprietary/Patented Algorithms: Algorithm provider can't send algorithm.

Directly enabled by the information service:

- Get the algorithms to the party with the data.
- Use Grid security to only allow certain parties to see certain information.
- Only allow secure downloads, using GridFTP or other secure protocols.

Future work: certificate-protected exectution.

Implications for the security issue

Recall the security concerns

- Sensitive Data: Data provider can't send the data.
- Proprietary/Patented Algorithms: Algorithm provider can't send algorithm.

Directly enabled by the information service:

- Get the algorithms to the party with the data.
- Use Grid security to only allow certain parties to see certain information.
- Only allow secure downloads, using GridFTP or other secure protocols.

Future work: certificate-protected exectution.

Implications for the security issue

Recall the security concerns

- Sensitive Data: Data provider can't send the data.
- Proprietary/Patented Algorithms: Algorithm provider can't send algorithm.

Directly enabled by the information service:

- Get the algorithms to the party with the data.
- Use Grid security to only allow certain parties to see certain information.
- Only allow secure downloads, using GridFTP or other secure protocols.

Future work: certificate-protected exectution.

Layout

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool
  - Helps combine the computational resources of multiple organizations

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool
  - Helps combine the computational resources of multiple organizations
- Condor in general

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool
  - Helps combine the computational resources of multiple organizations
- Condor in general
  - Is flexible and configurable

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool
  - Helps combine the computational resources of multiple organizations
- Condor in general
  - Is flexible and configurable
  - Can do cycle scavenging

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool
  - Helps combine the computational resources of multiple organizations
- Condor in general
  - Is flexible and configurable
  - Can do cycle scavenging
- Advantageous to organizations that do not have dedicated computation servers, but lots of workstations

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool
  - Helps combine the computational resources of multiple organizations
- Condor in general
  - Is flexible and configurable
  - Can do cycle scavenging
- Advantageous to organizations that do not have dedicated computation servers, but lots of workstations
  - Hospital and health department workstations

Advantages of using Condor

- Condor with Glidein and the Globus Toolkit
  - Enables the addition of Grid resources to the pool
  - Helps combine the computational resources of multiple organizations
- Condor in general
  - Is flexible and configurable
  - Can do cycle scavenging
- Advantageous to organizations that do not have dedicated computation servers, but lots of workstations
  - Hospital and health department workstations
  - Instructional labs in universities

Questions?

Thank you for your attention.

This work is supported by CDC grant 1U38HK000063-01

Workflows

- What are workflows for?

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components
- Our use case:

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components
- Our use case:
  - Data needs to be split up into small chunks to be handled by each job

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components
- Our use case:
  - Data needs to be split up into small chunks to be handled by each job
  - Individual jobs are executed independently on chunks of data

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components
- Our use case:
  - Data needs to be split up into small chunks to be handled by each job
  - Individual jobs are executed independently on chunks of data
  - Results are consolidated (often more complex than a simple aggregation)

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components
- Our use case:
  - Data needs to be split up into small chunks to be handled by each job
  - Individual jobs are executed independently on chunks of data
  - Results are consolidated (often more complex than a simple aggregation)
  - May have some post-processing

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components
- Our use case:
  - Data needs to be split up into small chunks to be handled by each job
  - Individual jobs are executed independently on chunks of data
  - Results are consolidated (often more complex than a simple aggregation)
  - May have some post-processing
- Tools for designing and executing workflows:
  - Taverna workbench
  - Kepler scientific workflow system
  - Business Process Execution Language (BPEL)

Workflows

- What are workflows for?
  - A workflow consists of a series of interrelated operations
  - Workflow management systems can handle the execution of complex workflows consisting of many components
- Our use case:
  - Data needs to be split up into small chunks to be handled by each job
  - Individual jobs are executed independently on chunks of data
  - Results are consolidated (often more complex than a simple aggregation)
  - May have some post-processing
- Tools for designing and executing workflows:
  - Taverna workbench
    - Has been used with Globus, Condor, and caGrid
  - Kepler scientific workflow system
  - Business Process Execution Language (BPEL)