Overdrive Controllers for Distributed Scientific Computation

Justin M. Wozniak

TCSC Doctoral Symposium

May 16, 2007

UNIVERSITY OF NOTRE DAME

Outline

- **Outline**
- Motivation
- Volunteer Storage
- Control Models
- Application to Scheduling
- Conclusion

Motivation

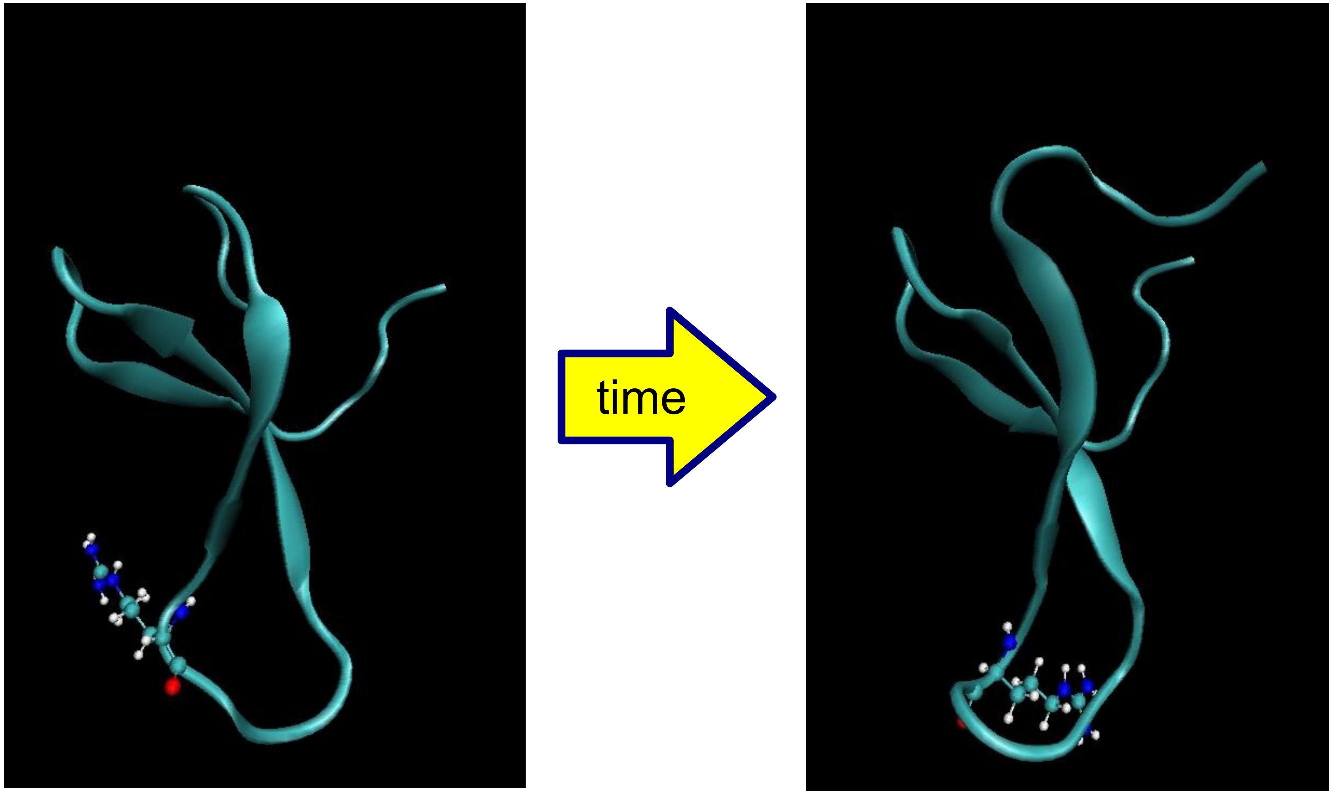
✗Outline
✗**Motivation**
✗Storage
✗Models
✗Scheduling
✗Conclusion

- Ordinary *naïve* scientific applications may be ported to modern, large distributed computer systems
  - Performance/utilization improvements
  - Hope to obtain benefits proportional to resources allocated
- Numerical scientific codes are a particularly simple case
  - Typically simple I/O
  - Well-understood underlying algorithms

Motivation

✗Outline
✗**Motivation**
✗Storage
✗Models
✗Scheduling
✗Conclusion

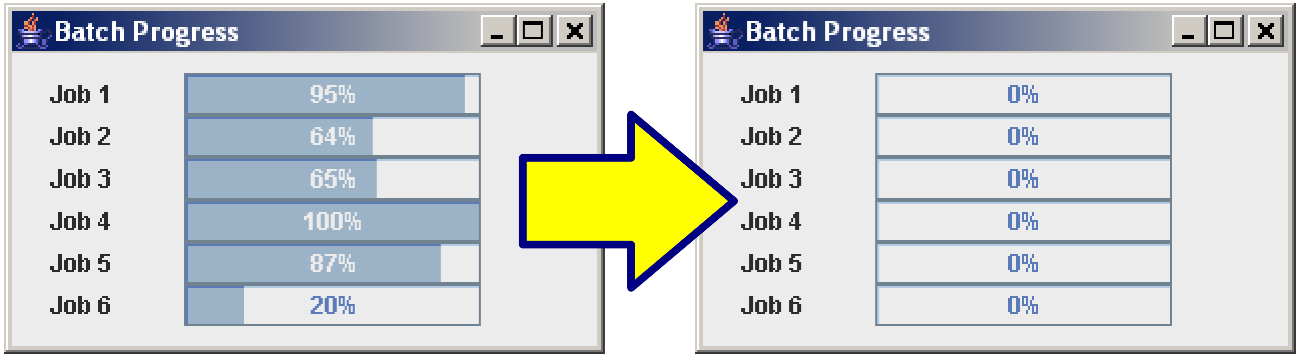
- Example: Transition Path Sampling (Chemistry)



Motivation

✗Outline
✗**Motivation**
✗Storage
✗Models
✗Scheduling
✗Conclusion

- Biomolecular simulations may be managed as batches of similar jobs
- Whole (Monte Carlo) batches must complete to obtain a result
- Trajectory data is stored along the way


Batch Progress

Job 1	95%
Job 2	64%
Job 3	65%
Job 4	100%
Job 5	87%
Job 6	20%

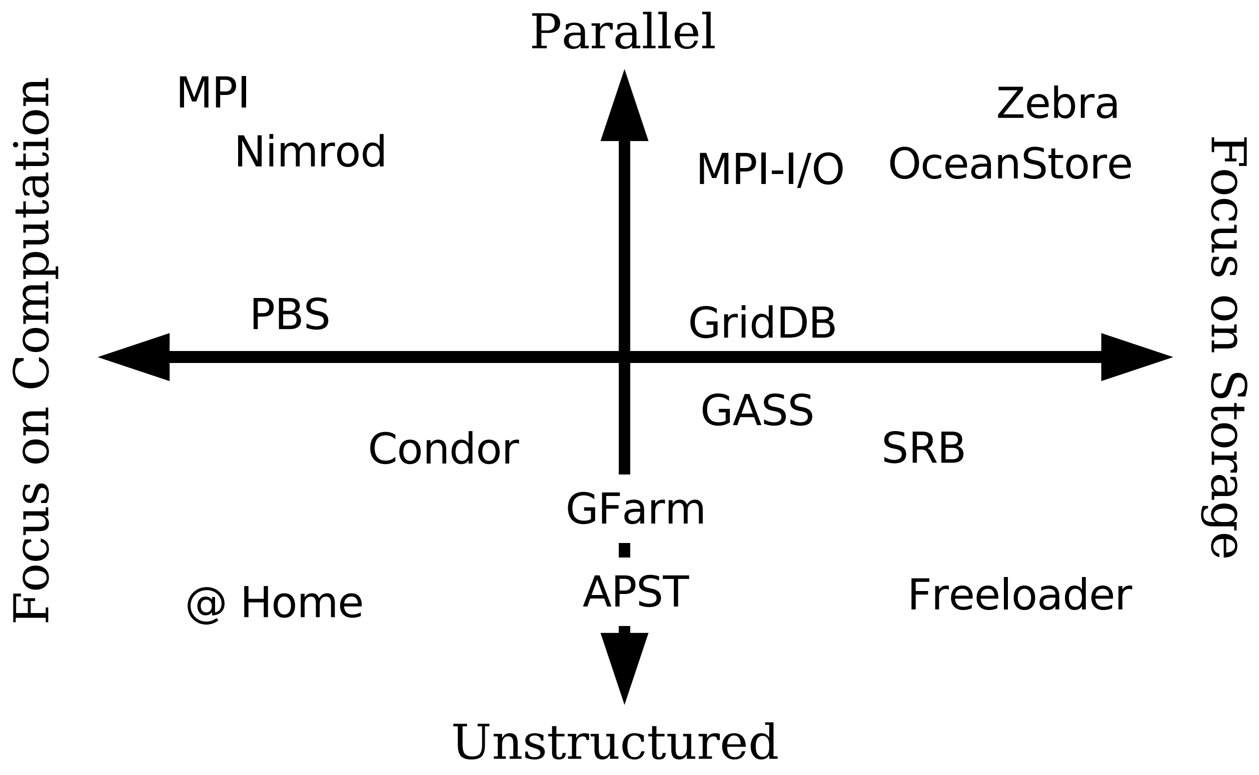
Batch Progress

Job 1	0%
Job 2	0%
Job 3	0%
Job 4	0%
Job 5	0%
Job 6	0%

Existing Software

✗Outline
✗**Motivation**
✗Storage
✗Models
✗Scheduling
✗Conclusion

- A variety of existing tools may be plugged together:



Parallel

Unstructured

Focus on Computation

Focus on Storage

MPI

Nimrod

Zebra

MPI-I/O

OceanStore

PBS

GridDB

GASS

Condor

SRB

GFarm

APST

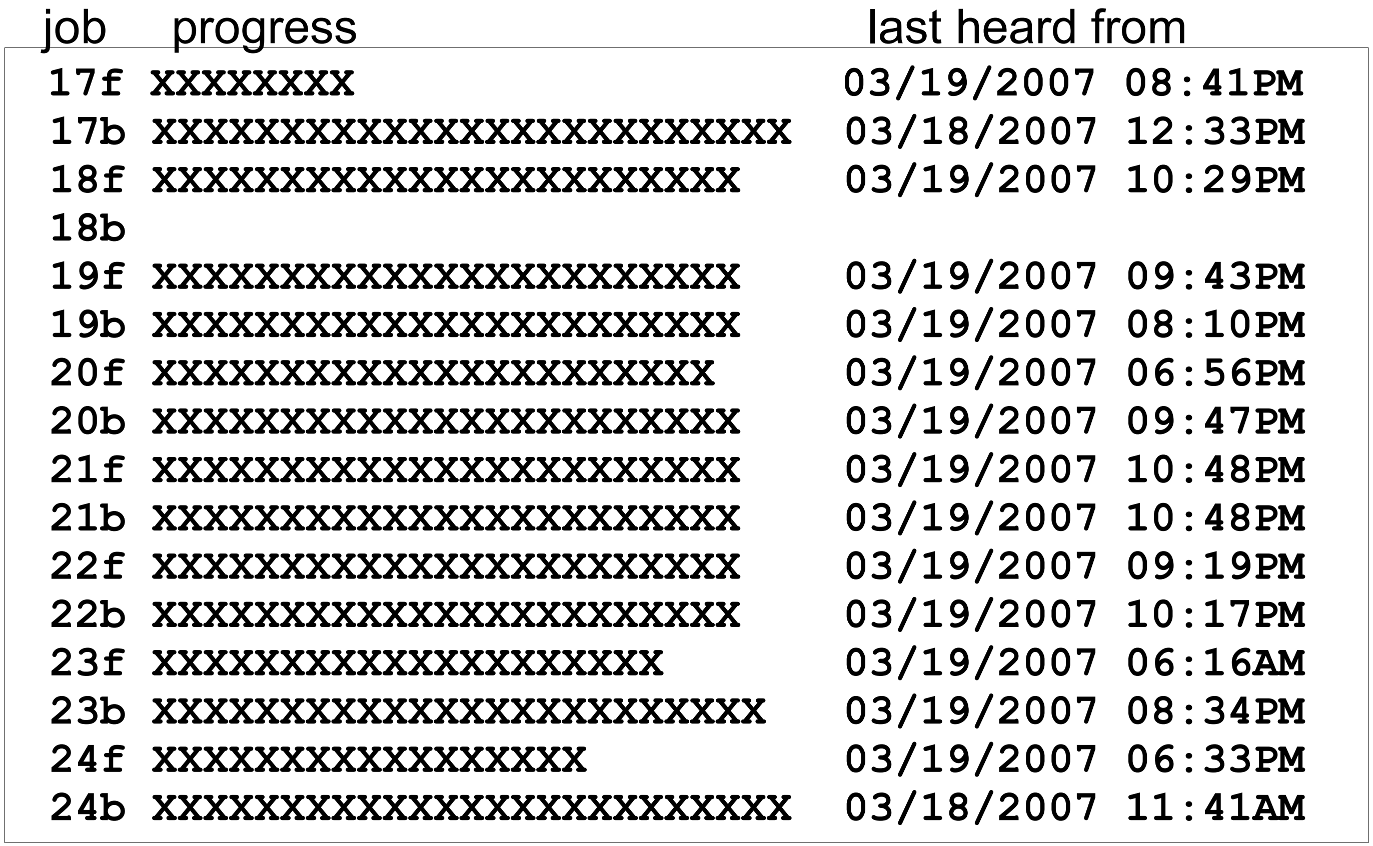
@ Home

Freeloader

Motivation

- ✗Outline
- ✗**Motivation**
- ✗Storage
- ✗Models
- ✗Scheduling
- ✗Conclusion

- Real-world job unpredictability:



```
job  progress                     last heard from
17f  XXXXXXXX                     03/19/2007 08:41PM
17b  XXXXXXXXXXXXXXXXXXXXXXXXX    03/18/2007 12:33PM
18f  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 10:29PM
18b
19f  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 09:43PM
19b  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 08:10PM
20f  XXXXXXXXXXXXXXXXXXXXXX       03/19/2007 06:56PM
20b  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 09:47PM
21f  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 10:48PM
21b  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 10:48PM
22f  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 09:19PM
22b  XXXXXXXXXXXXXXXXXXXXXXX      03/19/2007 10:17PM
23f  XXXXXXXXXXXXXXXXXXXX         03/19/2007 06:16AM
23b  XXXXXXXXXXXXXXXXXXXXXXXX     03/19/2007 08:34PM
24f  XXXXXXXXXXXXXXXXX            03/19/2007 06:33PM
24b  XXXXXXXXXXXXXXXXXXXXXXXXX    03/18/2007 11:41AM
```

Motivation

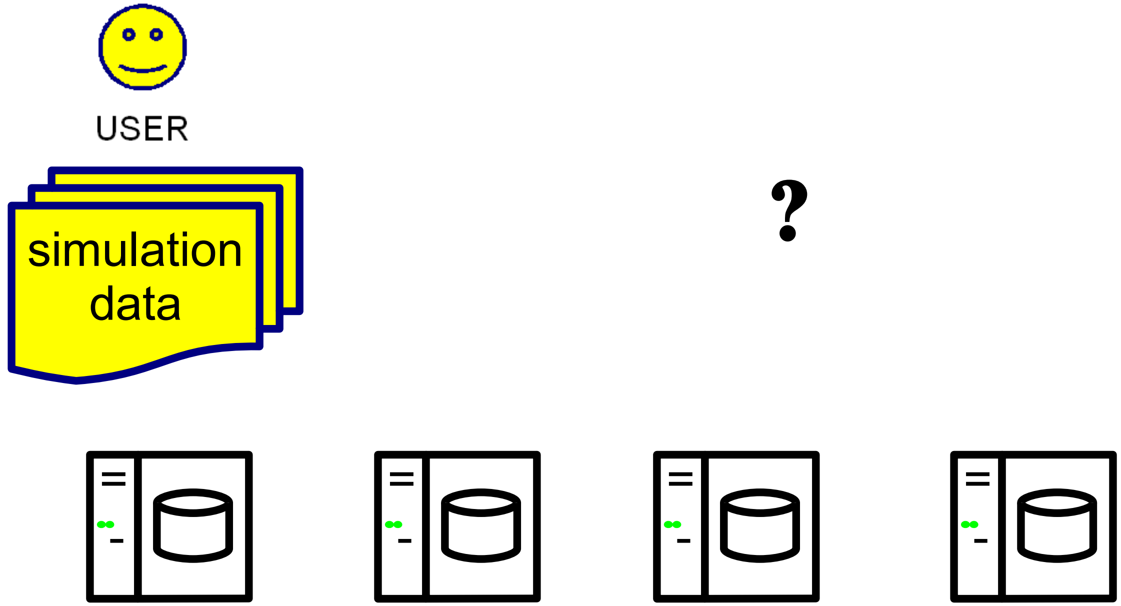
✗Outline
✗**Motivation**
✗Storage
✗Models
✗Scheduling
✗Conclusion

- Volunteer computing
- Volunteer storage
- Heterogeneous systems are averse to traditional parallel techniques
  - Variability in performance
  - Variability in reliability
- Users desire simple APIs, but unstructured systems require active controllers.

Volunteer Storage

✗Outline
✗Motivation
✗**Storage**
✗Models
✗Scheduling
✗Conclusion

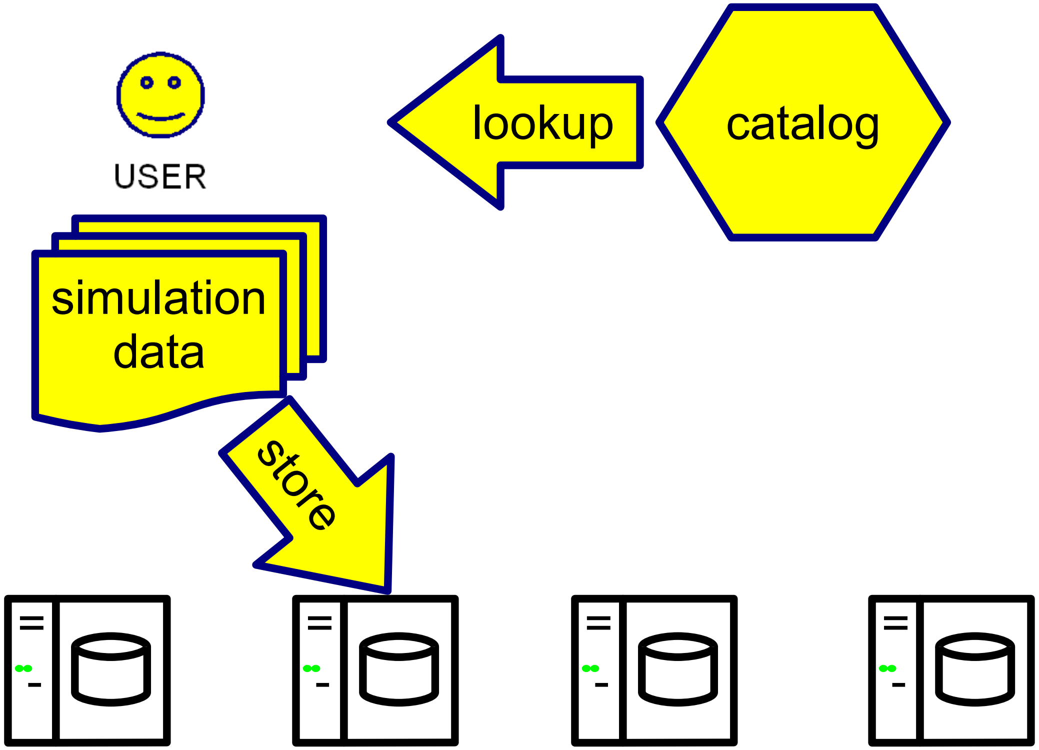
- Users volunteer storage space to an external system
- Resource revocation may occur at any time



Volunteer Storage

- ✗Outline
- ✗Motivation
- ✗**Storage**
- ✗Models
- ✗Scheduling
- ✗Conclusion

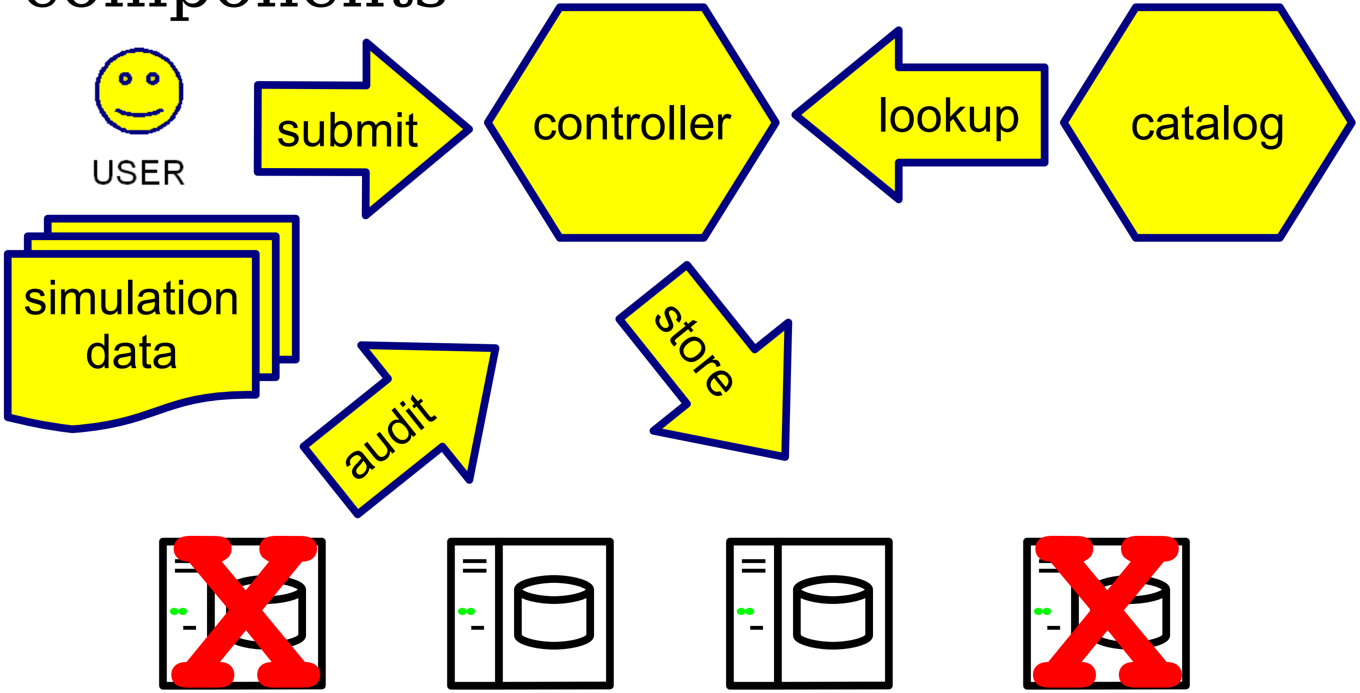
- Users have the ability to use *middleware* to locate and access remote resources



Volunteer Storage

✗Outline
✗Motivation
✗**Storage**
✗Models
✗Scheduling
✗Conclusion

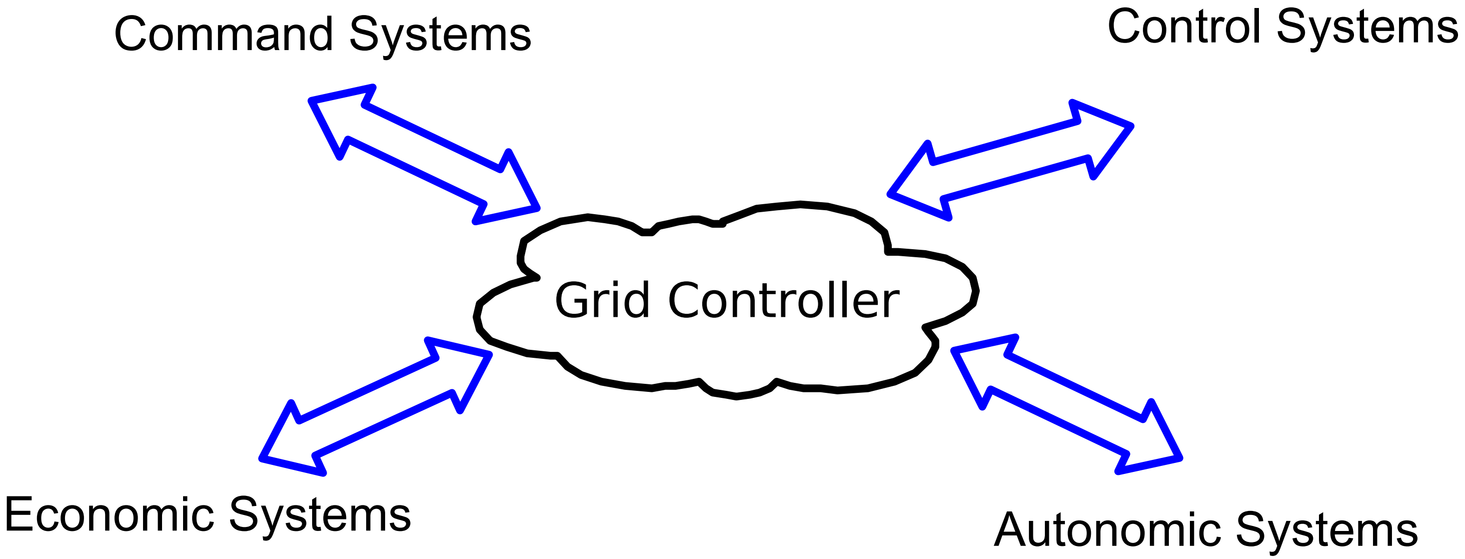
- *Overdrive Grid Controllers* are introduced to handle the complexity of large systems of unmanaged components



Overdrive Controller

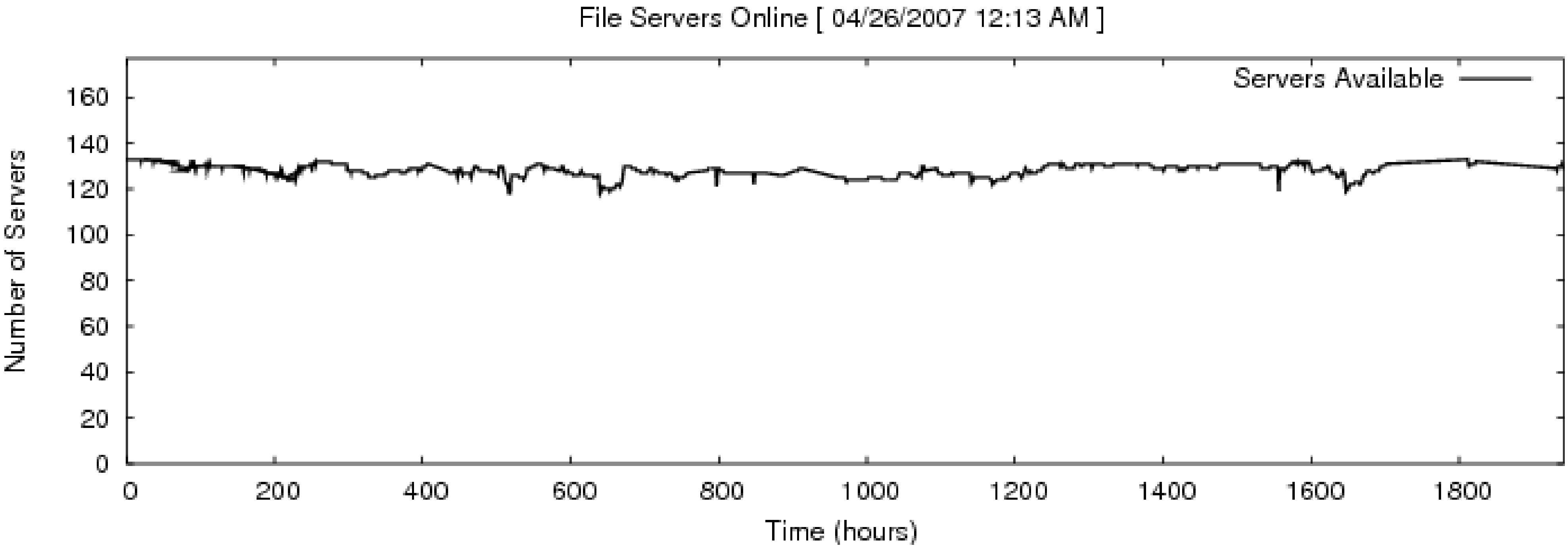
- ✗Outline
- ✗Motivation
- ✗Storage
- ✗**Models**
- ✗Scheduling
- ✗Conclusion

- What does this type of system borrow from other disciplines?



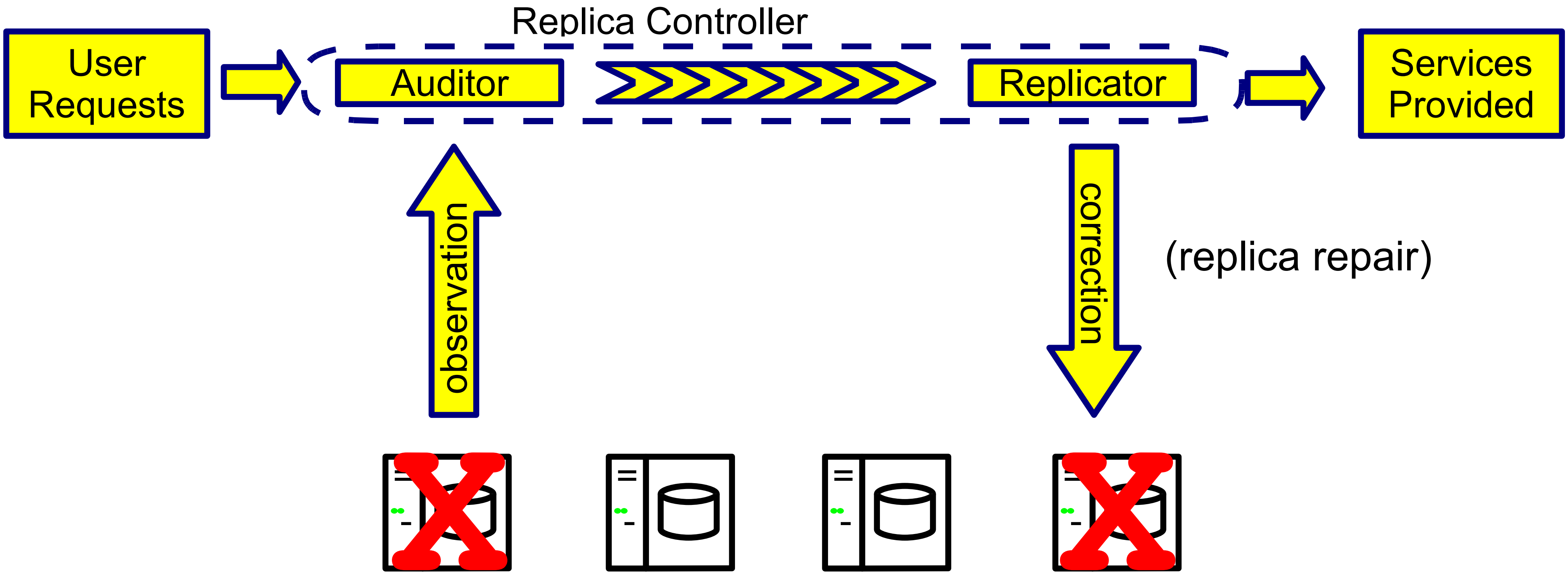
Autonomic System Management

- How do we control an unstable set of volunteer systems?



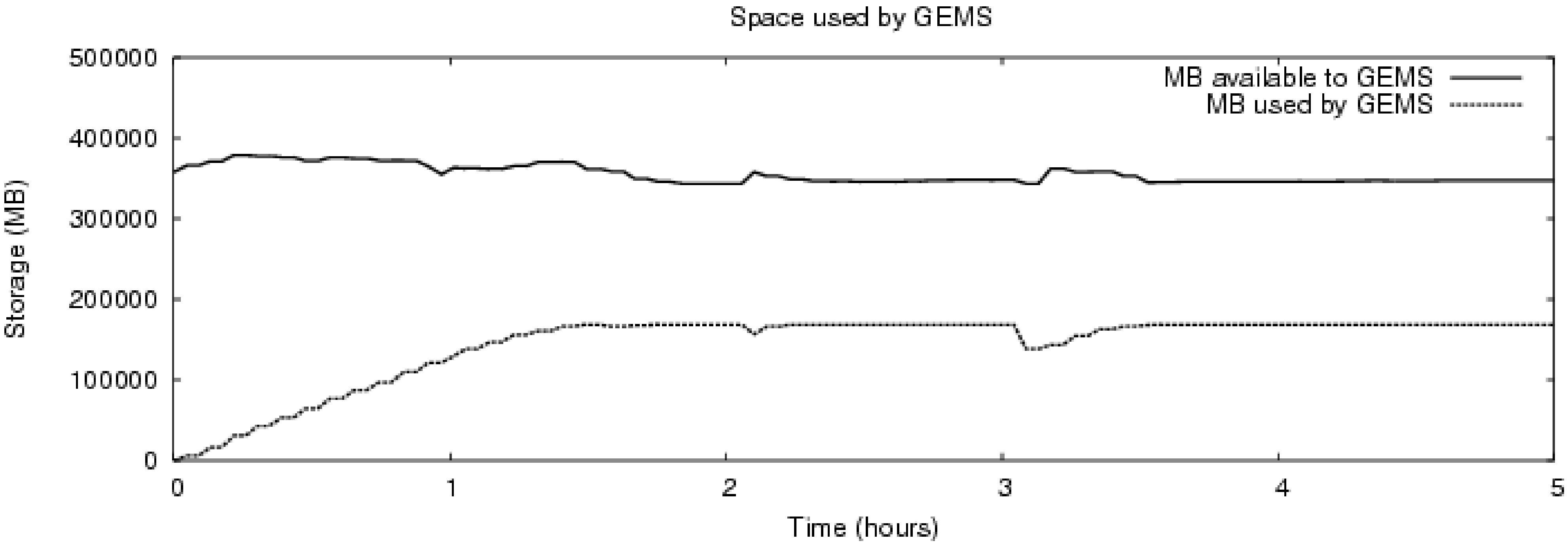
Decision Making

- How do we prioritize operations?



Control Theory

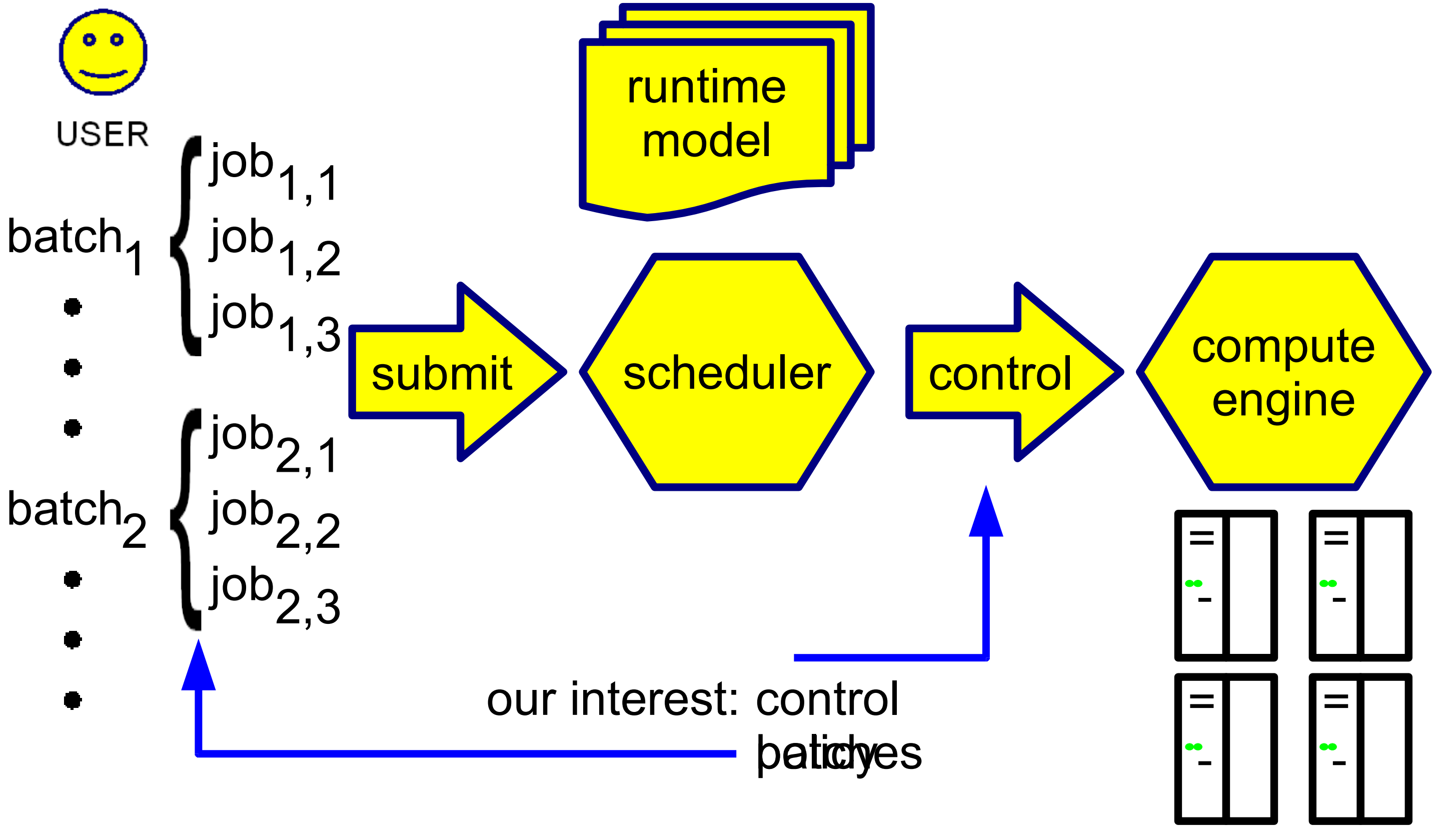
- How do we tell if the system is working?



Volunteer CPUs

✗Outline
✗Motivation
✗Storage
✗Models
✗**Scheduling**
✗Conclusion

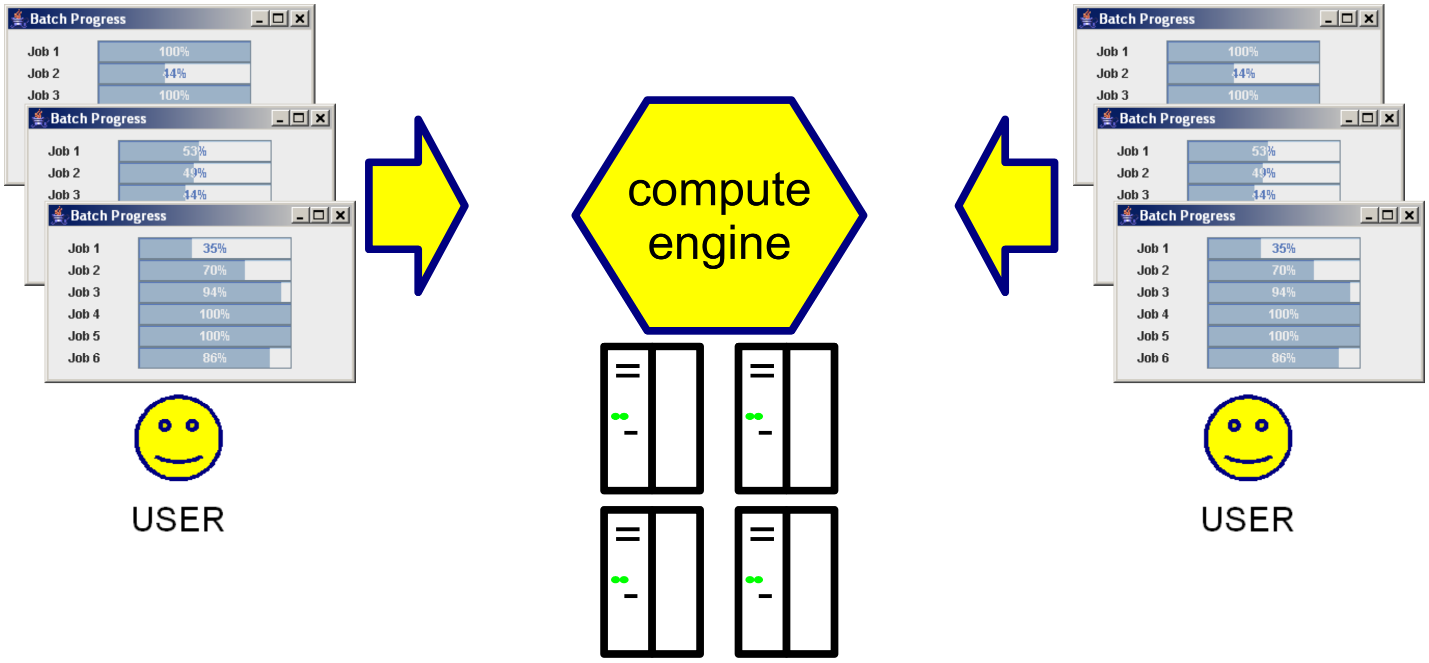
East simulated framework



Economic Systems

- ✗Outline
- ✗Motivation
- ✗**Definitions**
- ✗Estimation
- ✗Simulation
- ✗Conclusion

- Contention



“Overdrive Controllers…” Annual Simulation Symposium 17

Estimation

- ✗Outline
- ✗Motivation
- ✗Storage
- ✗Models
- ✗**Scheduling**
- ✗Conclusion

- Inherent Unpredictability
  - Grid resources are volatile in availability and variable in performance
  - User estimates are pretty bad
- The Fundamental Enforcement Policy Question:

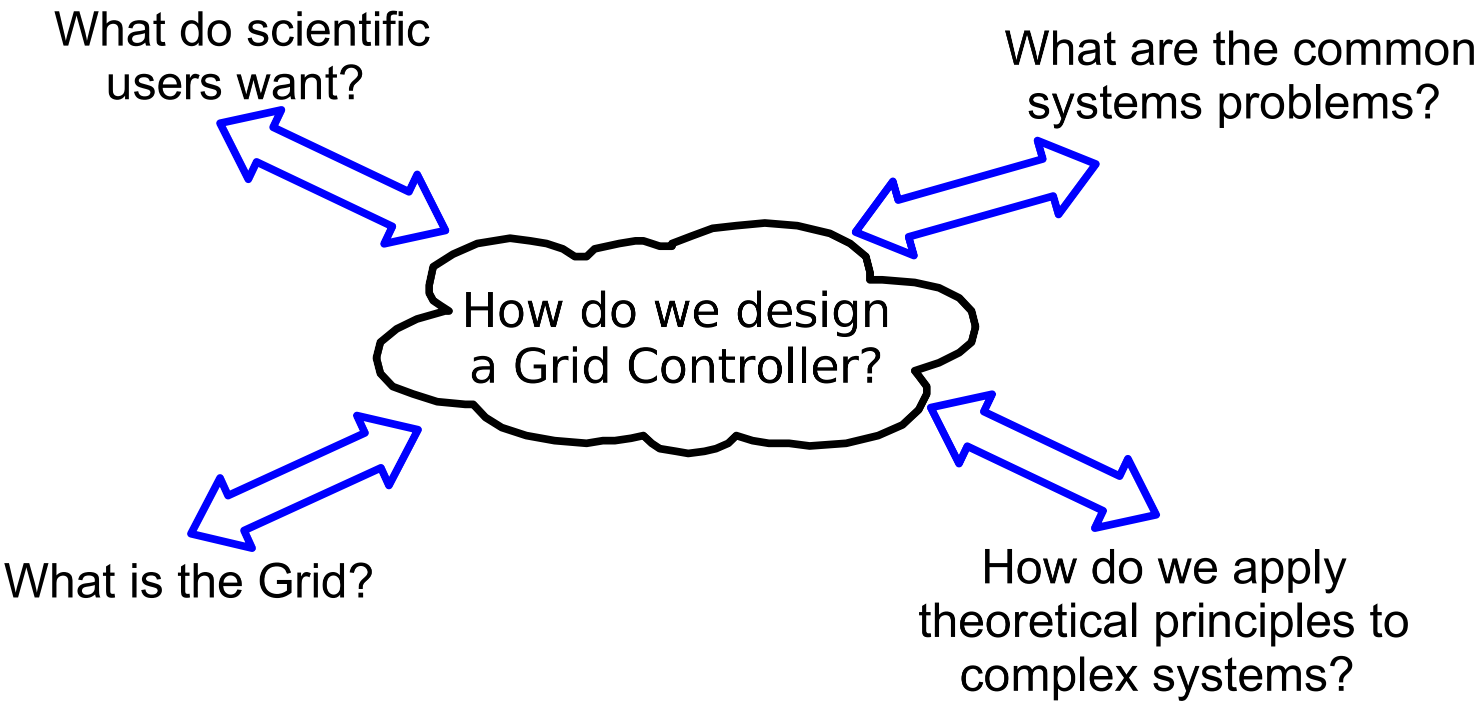
When should the system kill jobs

to satisfy guarantees?

Integration

✗Outline
✗Motivation
✗Storage
✗Models
✗Scheduling
✗**Conclusion**

- How do we tie these controllers together?



How do we design a Grid Controller?

What do scientific users want?

What are the common systems problems?

What is the Grid?

How do we apply theoretical principles to complex systems?

Summary

✗Outline
✗Motivation
✗Storage
✗Models
✗Scheduling
✗**Conclusion**

- Develop a model for software controllers that simplify user experience and manage complex computer systems
- Apply model to:
  - Replica storage systems
  - Deadline-driven job scheduling
- Support ongoing projects in computational chemistry