Workflow Job Manager for NCCP

by Ivan Gibbs, Eric Fritzinger, Sergiu M. Dascalu, Frederick C. Harris, Jr., and Yantao Shen

May 20-24, 2013

Overview

1. Project Background
2. Research
3. Software Model
4. Prototype and Comparison to Other Tools
5. Conclusion

Project Background

Project Background

NCCP is a work by several people involved in the 2008-2013 NSF EPSCoR project.

Figure: NCCP homepage

Project Background

Modeling and Simulation–Use of models to assist in making decisions. Includes things such as :

- financial forcasting
- bridge models
- flight simulators

Project Background

Eric Fritzinger and Demeter–for data processing and coupling of scientific models.

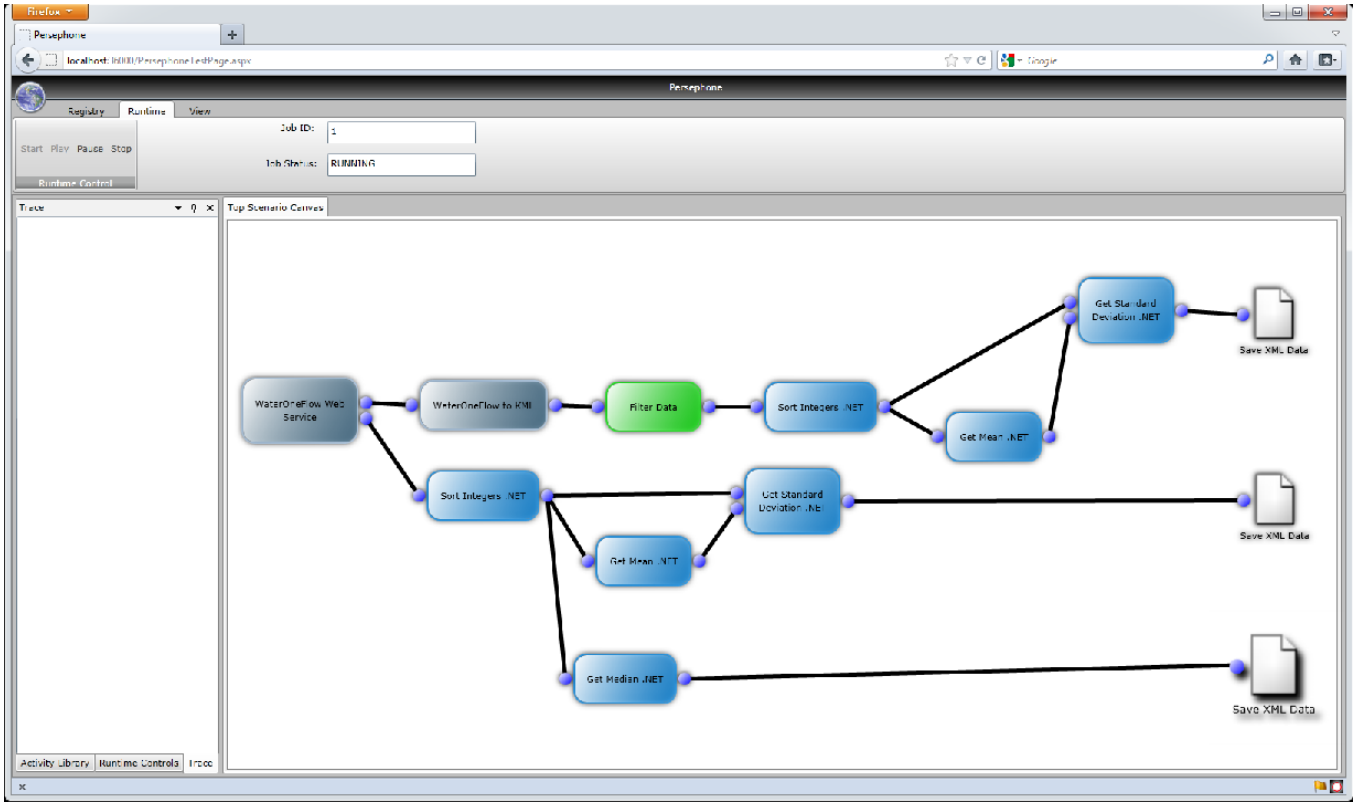


Figure: Persephone

Research

Service Oriented Architecture (SOA)

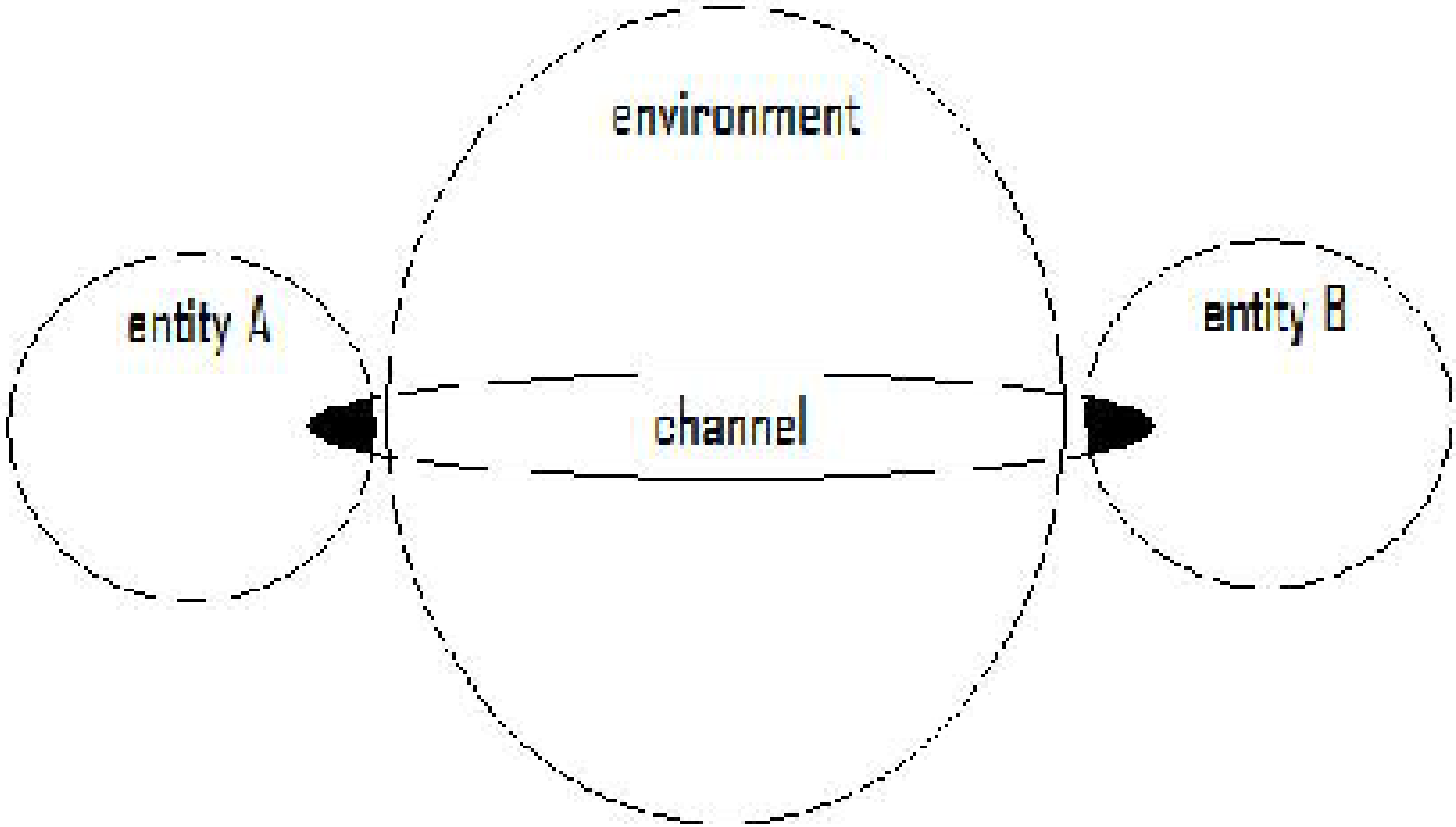


Figure: SOA diagram

Research

Web services

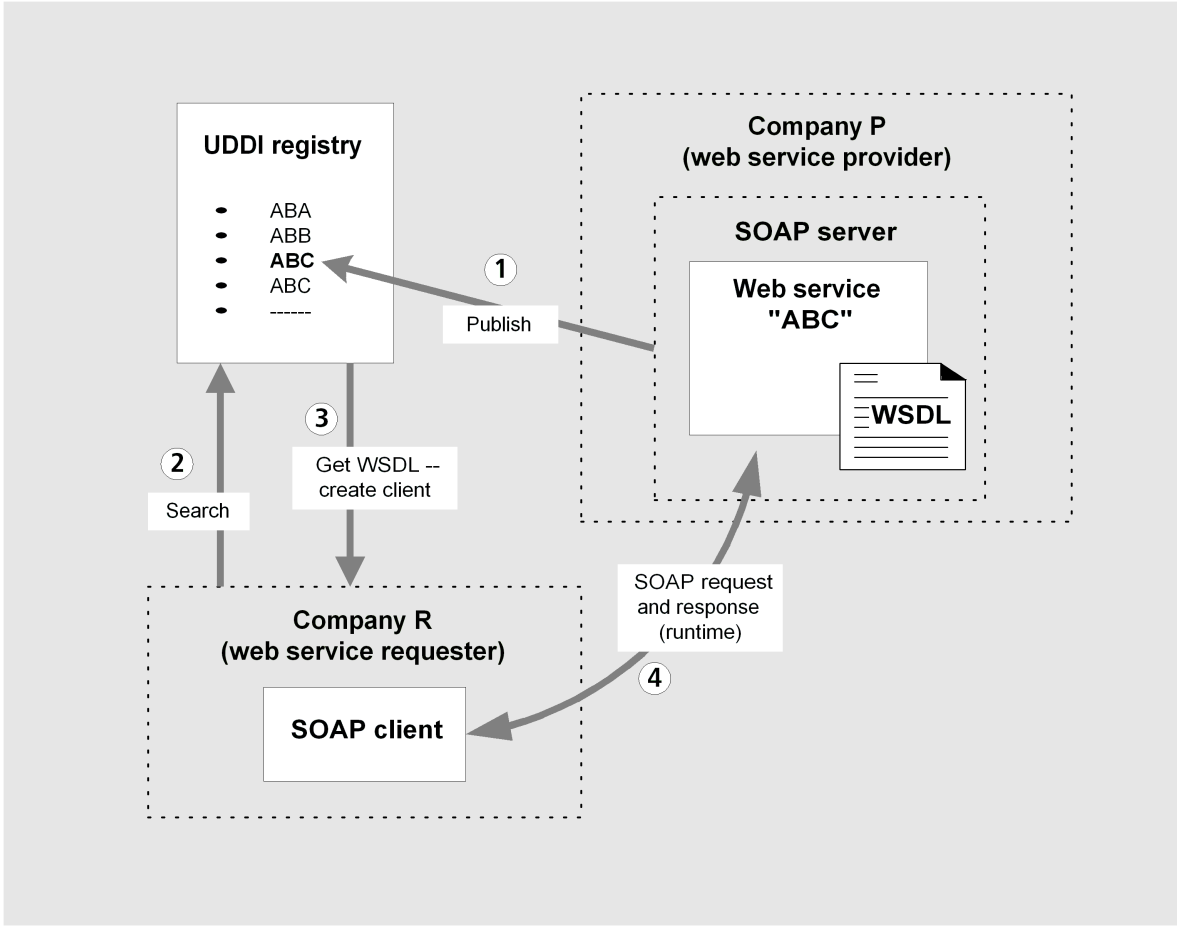


Figure: Web Services using WSDL and UDDI [1]

Research

Interoperability allows software to work together.
Some interoperability enabling standards for SOA are:

1. Common Object Request Broker Architecture (CORBA)
2. Java Platform Enterprise Edition (J2EE)
3. Distributed Component Object Model (DCOM)
4. Remote Method Invocation (RMI)
5. SOAP

Research

Scientific model interoperability through:

1. direct software implementation,
2. data interoperability standards.
   - Model Coupling Toolkit (MCT)
   - Earth System Modeling Framework (ESMF)
   - Common Component Architecture (CCA)
   - OASIS
   - and others

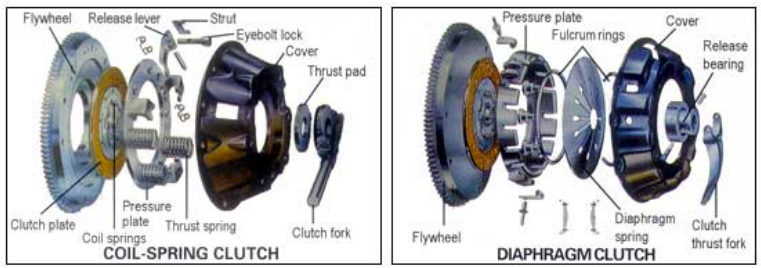


Figure: Clutch of a car[2]

Motivation

1. World Wide Web (WWW) enables more collaboration.
2. combine individual models for larger simulations
3. distributed computing offers low cost computer resources

The Software Model

Table: Functional Requirements

R1	The job manager must allow the user to view the current status of a workflow.

The Software Model

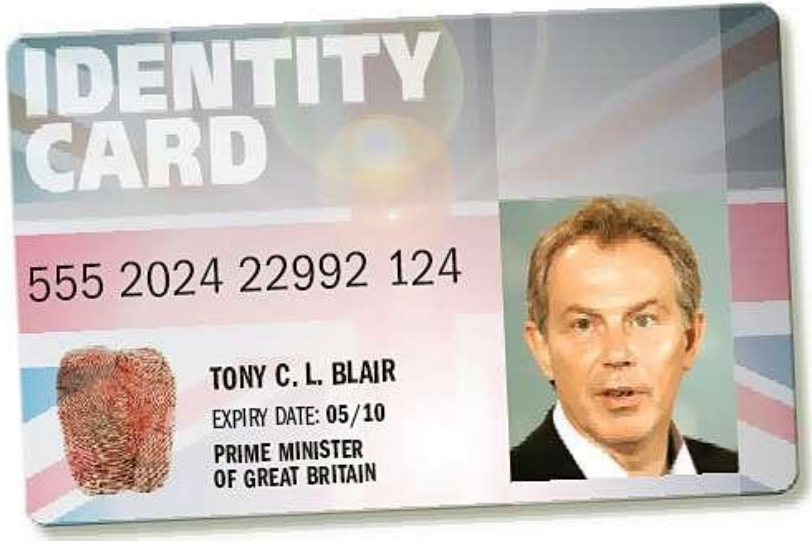
Table: Functional Requirements

R1	The job manager must allow the user to view the current status of a workflow.
R2	The job manager must allow the user to start a workflow.
R3	The job manager must allow the user to stop a workflow.
R4	The job manager must allow the user to pause a workflow.
R5	The job manager must allow the user to unpause a workflow.

The Software Model

Table: Functional Requirements

R1	The job manager must allow the user to view the current status of a workflow.
R2	The job manager must allow the user to start a workflow.
R3	The job manager must allow the user to stop a workflow.
R4	The job manager must allow the user to pause a workflow.
R5	The job manager must allow the user to unpause a workflow.
R6	The job manager must provide an ID for each job.



The Software Model

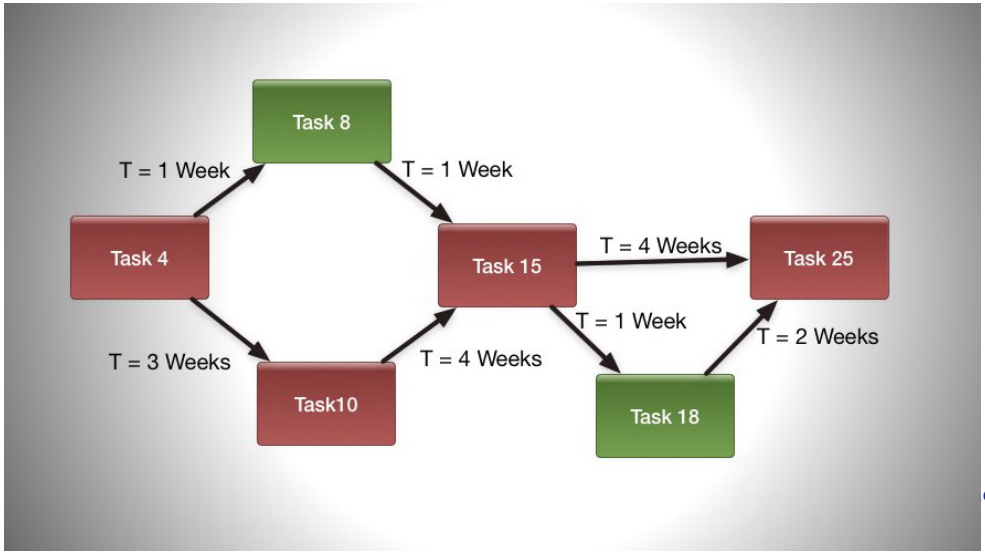
Table: Functional Requirements

R1	The job manager must allow the user to view the current status of a workflow.
R2	The job manager must allow the user to start a workflow.
R3	The job manager must allow the user to stop a workflow.
R4	The job manager must allow the user to pause a workflow.
R5	The job manager must allow the user to unpause a workflow.
R6	The job manager must provide an ID for each job.
R7	The job manager must provide the current state of a job.

The Software Model

Table: Functional Requirements

R1	The job manager must allow the user to view the current status of a workflow.
R2	The job manager must allow the user to start a workflow.
R3	The job manager must allow the user to stop a workflow.
R4	The job manager must allow the user to pause a workflow.
R5	The job manager must allow the user to unpause a workflow.
R6	The job manager must provide an ID for each job.
R7	The job manager must provide the current state of a job.
R8	The job manager must allow the user to visualize a job's execution.



The Software Model

Table: Functional Requirements

R1	The job manager must allow the user to view the current status of a workflow.
R2	The job manager must allow the user to start a workflow.
R3	The job manager must allow the user to stop a workflow.
R4	The job manager must allow the user to pause a workflow.
R5	The job manager must allow the user to unpause a workflow.
R6	The job manager must provide an ID for each job.
R7	The job manager must provide the current state of a job.
R8	The job manager must allow the user to visualize a job's execution.
R9	The job manager must allow the user to see all of their jobs.

The Software Model

Table: Functional Requirements

R1	The job manager must allow the user to view the current status of a workflow.
R2	The job manager must allow the user to start a workflow.
R3	The job manager must allow the user to stop a workflow.
R4	The job manager must allow the user to pause a workflow.
R5	The job manager must allow the user to unpause a workflow.
R6	The job manager must provide an ID for each job.
R7	The job manager must provide the current state of a job.
R8	The job manager must allow the user to visualize a job's execution.
R9	The job manager must allow the user to see all of their jobs.
R10	The job manager must allow multiple users to view a job concurrently.

The Software Model

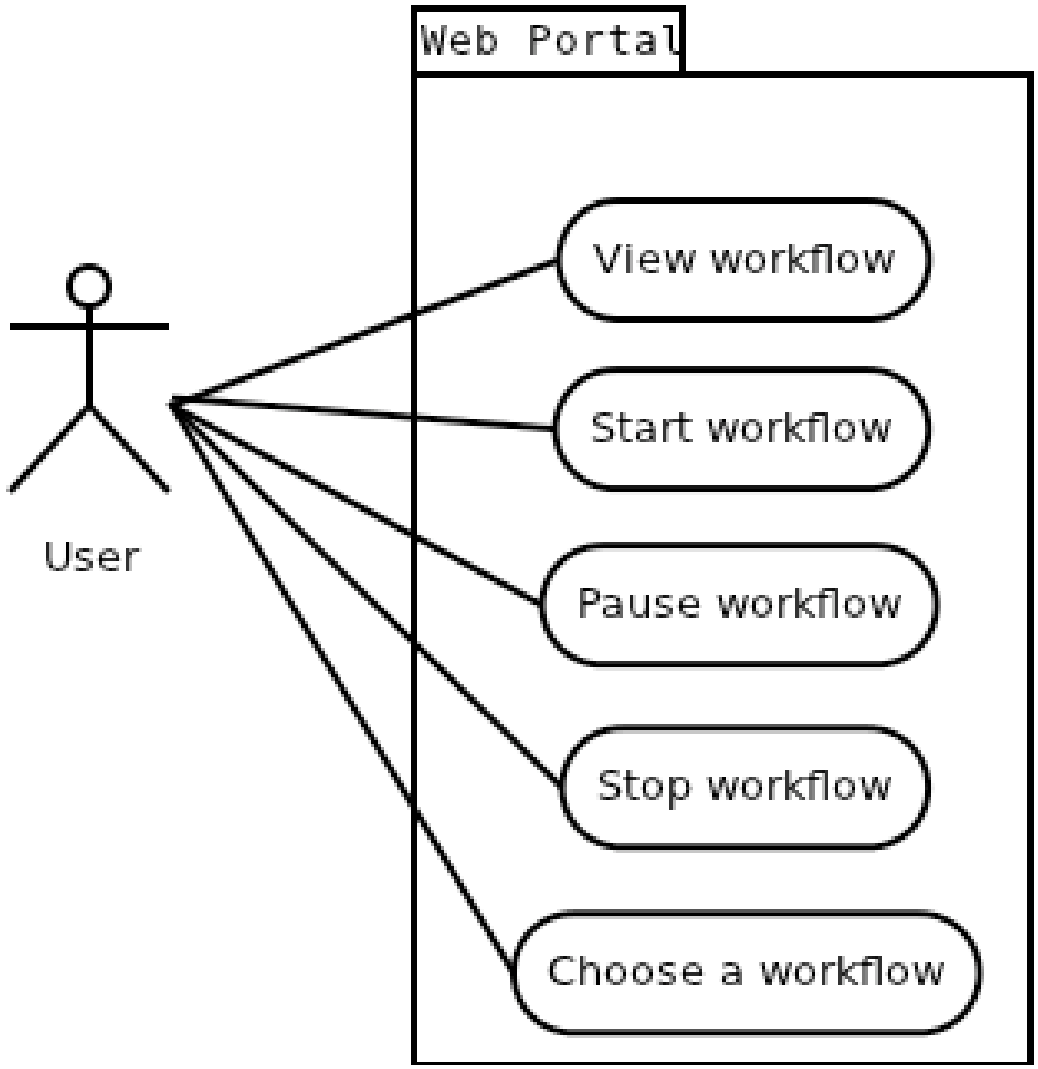
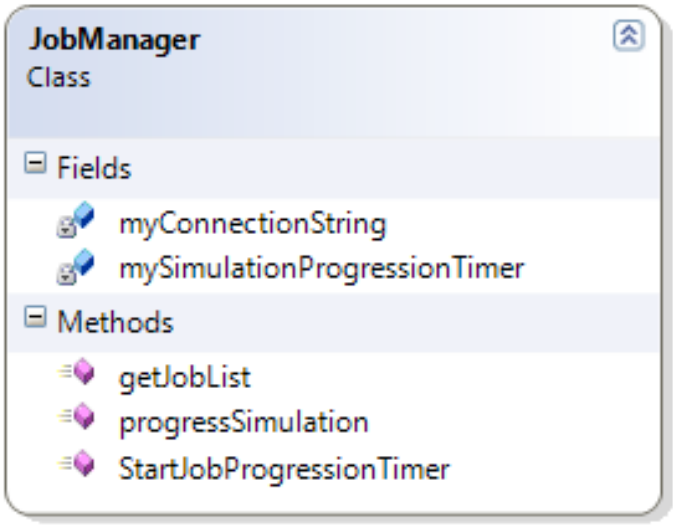


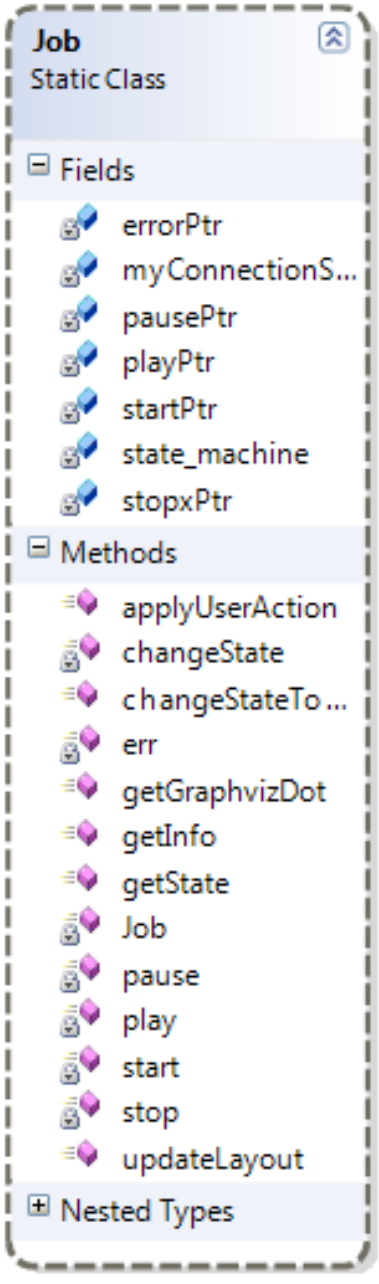
Figure: Use case diagram

The Software Model

Figure: Silverlight web application diagram

The Software Model





The Software Model

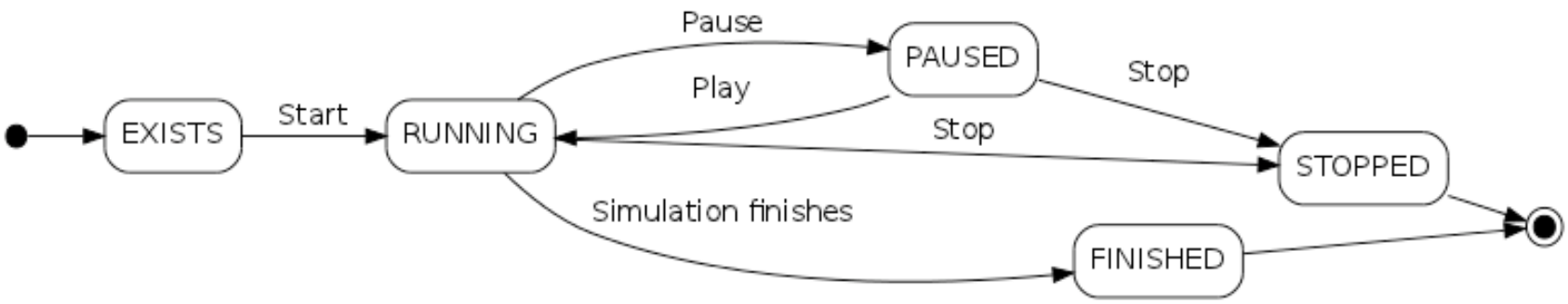


Figure: Workflow State Machine diagram

Prototype

- JobManager.swf
  - illustration of Job Manager client app
- Demeter_Stop.swf
  - Job Manager running in full size browser client: simulation is stopped
- Demeter_Finish.swf
  - Job Manager running in full size browser client: finishes simulation

Comparison to Other Tools

The Trident Workbench job manager.
http://www.youtube.com/watch?v=BZMjqVgJiIo (1:00 and 2:00 - 3:00)

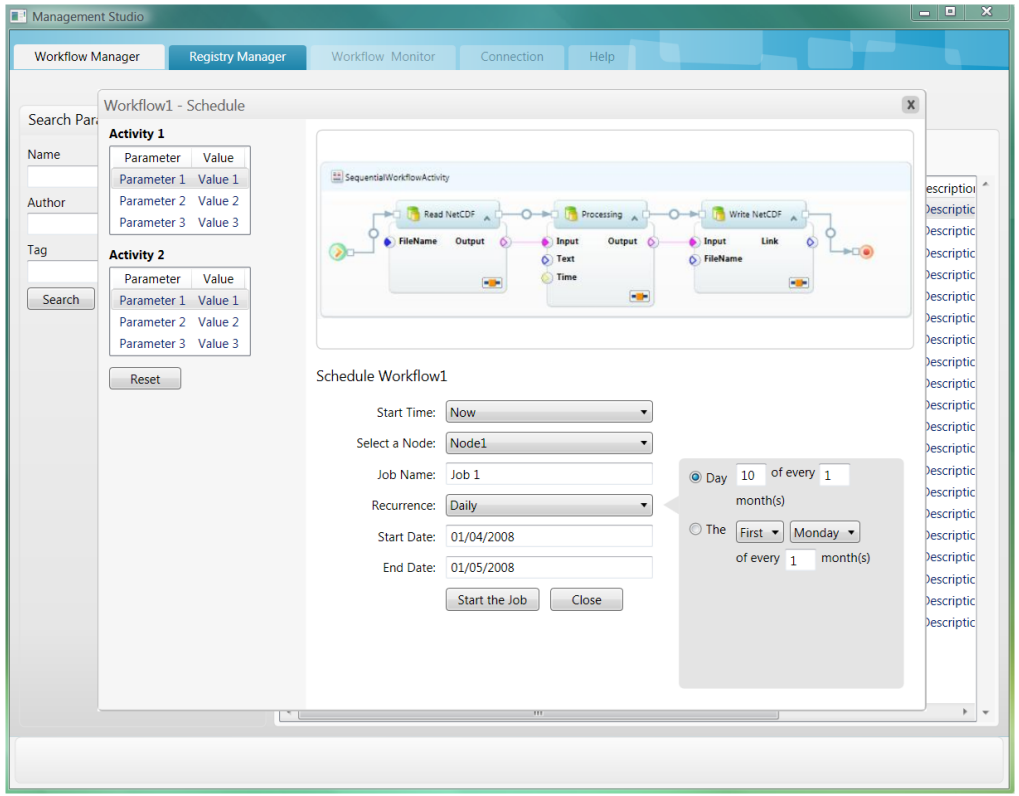


Figure: Trident workflow manager[3]

Comparison to Other Tools

The RealFlow job manager.

- uses esoteric language (job note, IDOC, etc.)
- IDOC - Independent Domain Of Computation
- http://www.youtube.com/watch?v=-1pnD5w7kz8

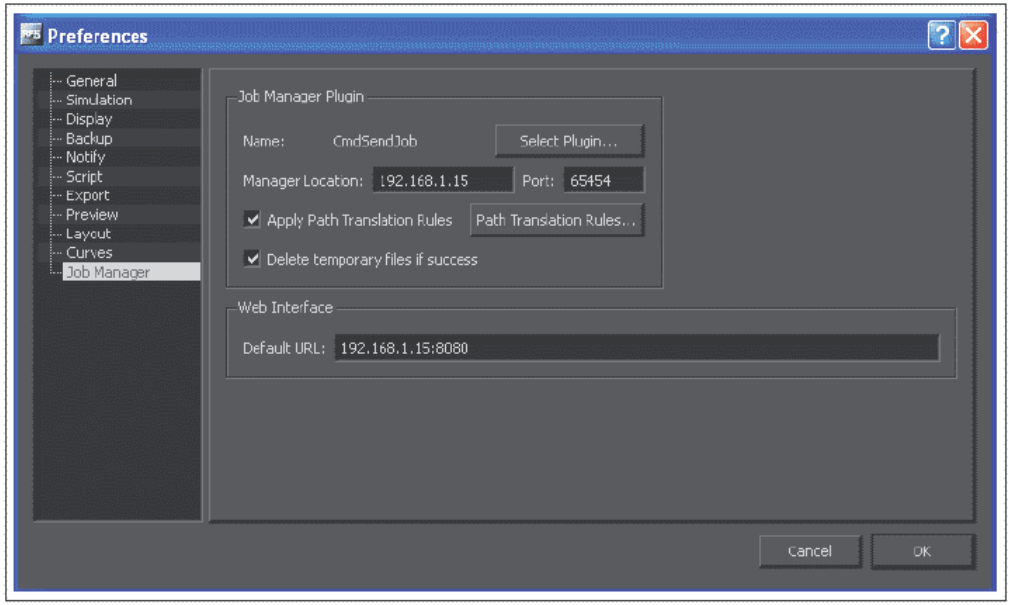


Figure: RealFlow Job Manager Preferences[4]

Comparison to Other Tools

The HPC job manager.

- all text interface

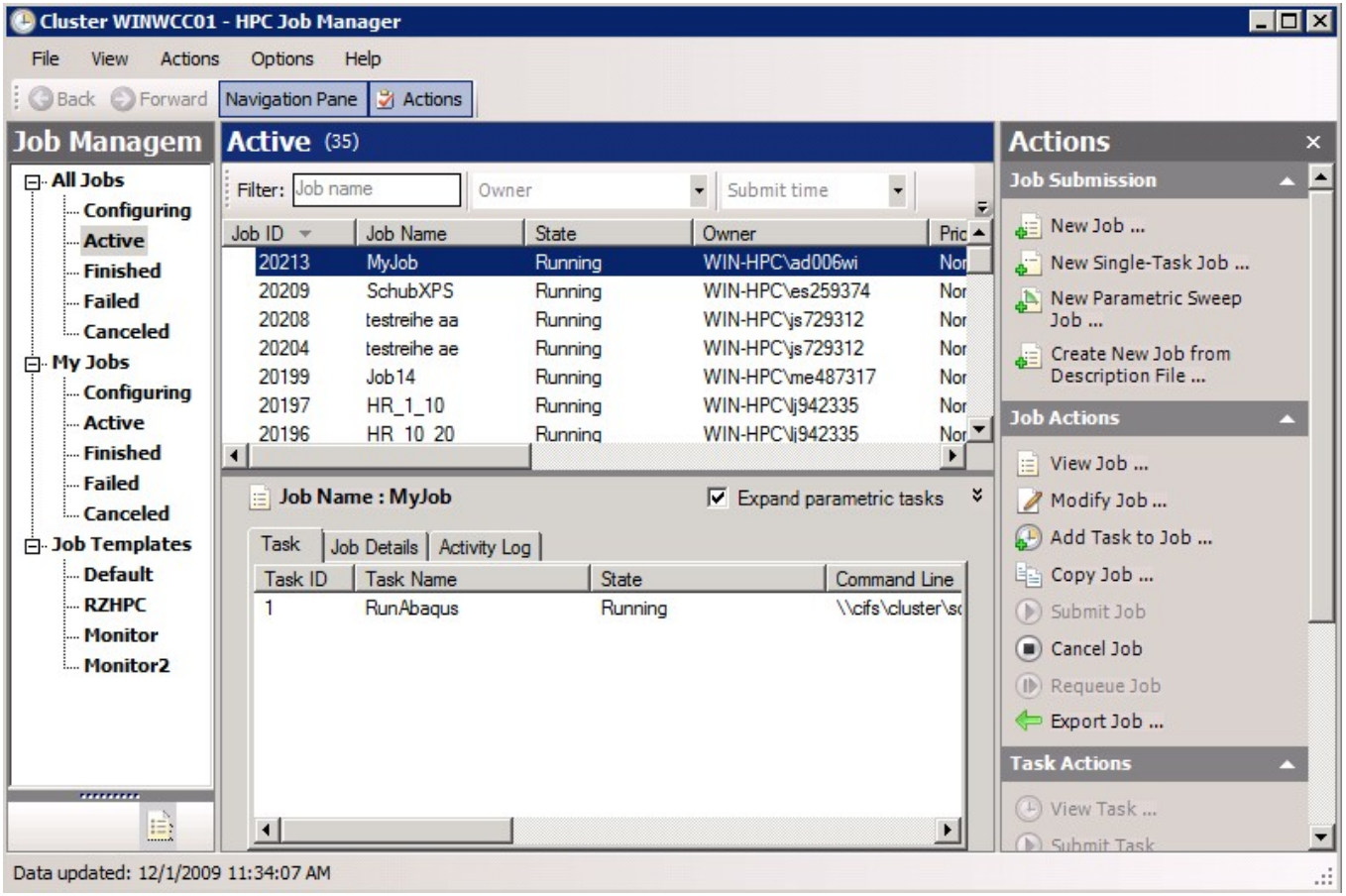


Figure: HPC Job Manager[5]

Comparison to Other Tools

The Taverna job manager.

- all text interface

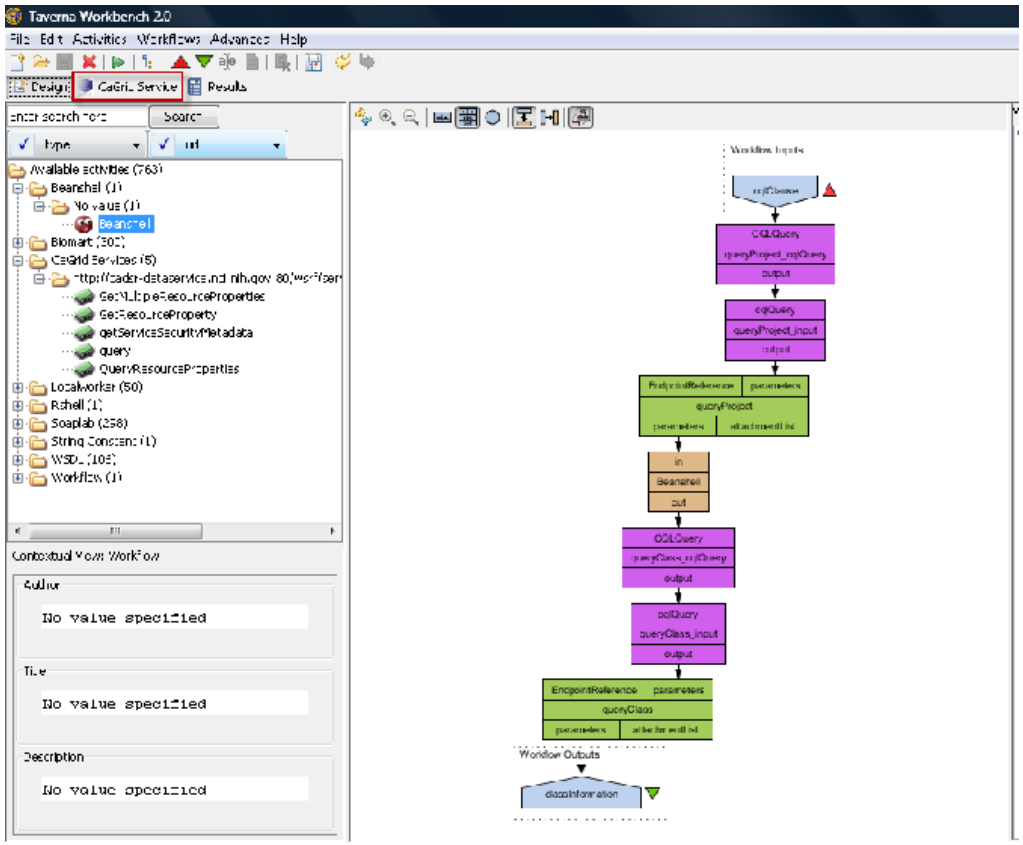


Figure: Taverna Job Manager[6]

Comparison to Other Tools

- Condor-G: gives access to multi-domain resources
- DAGMan: manages dependencies between jobs for Condor
- Globus: to build Grids for distributed computing

Table: Comparison of Technical Aspects

Software	GUI	Platforms	Condor-G	DAGMan	Globus	GraphViz
Trident Workbench	stand-alone web portal	W/M/L	N	N	N	N
RealFlow Job Mgmt		W/M/L	N	N	N	N
Natural Organic	web portal	W	N	N	N	N
HPC Job Mgr.		W/M/L	Y	Y	Y	N
Kepler	stand-alone	W/M/L	N	N	Y	N
Taverna Workbench	JRE client cmd line	JRE	N	N	Y	Y
Pegasus	stand-alone Java webstart	W/M/L Web	Y	Y	Y	N
NCCP	web portal	W/M	N	N	N	Y

Comparison to Other Tools

Table: Comparison of Functional Aspects

Software	pause/ resume/ stop	monitor execution via web	distribute execution
Trident Workbench	Y	Y	Y
RealFlow Job Mgmt	Y	Y	Y
Natural Organic	N	Y	Y
HPC Job Mgr.	Y	Y	Y
Kepler	Y	Y	Y
Taverna Workbench	Y	Y	Y
Pegasus	N	Y	Y
NCCP	Y	Y	Y

Conclusions

- offered a job management interface as a web service
- satisfied functional requirements
- implemented functions compare well with other applications

Contributions

- Job Manager presents a depiction of simulation progress with a dynamic graph

Contributions

- Job Manager presents a depiction of simulation progress with a dynamic graph
- uses separation of concerns through a state-transition matrix to guard for concurrency

Contributions

- Job Manager presents a depiction of simulation progress with a dynamic graph
- uses separation of concerns through a state-transition matrix to guard for concurrency
- user friendly interface for small devices

Contributions

- Job Manager presents a depiction of simulation progress with a dynamic graph
- uses separation of concerns through a state-transition matrix to guard for concurrency
- user friendly interface for small devices
- built an integrated interface for NCCP

Contributions

- Job Manager presents a depiction of simulation progress with a dynamic graph
- uses separation of concerns through a state-transition matrix to guard for concurrency
- user friendly interface for small devices
- built an integrated interface for NCCP
- extensible interface for Demeter

Acknowledgements

- Doctors Sergiu Dascalu, Fred Harris, and Yantao Shen for their support and encouragments
- Eric Fritzinger and Michael McMahon for the explanations of Demeter, Persephone, and the NCCP.
- National Science Foundation EPSCoR Grant for the Nevada Infrastructure for Climate Change Science, Education, & Outreach, No. 0814372

References

1 Oracle Corporation. Building Web Services–Chapter 1. http://docs.oracle.com/cd/E19792-01/816-7862/index.html, September 2012

2 indiacar.com. Transmission Systems: The Clutch. http://www.indiacar.com/infobank/theclutch.htm, November 2012

3 Microsoft Research. Project Trident: A Scientific Workflow Workbench. http://research.microsoft.com/en-us/collaboration/tools/trident.aspx, October 2012.

4 RealFlow. Job Manager. http://thevault.realflow.com/job_manager-83.html, October 2012.

References (continued)

5 Rachen-und Kommunikationszentrum. Using Abaqus with the Windows Batch System. http://www.rz.rwth-aachen.de/aw/cms/rz/Themen/hochleistungsrechnen/nutzung/nutzung_des_rechners_unter_windows/~sxn/Windows_BatchSystem_Examples_Abaqus/?lang=de, November 2012.

6 Taverna. Taverna Users Guide. http://www.cagrid.org/display/workflow13/Taverna+Users+Guide, October 2012.