

Imsim On Demand

DESC Collaboration Meeting -- June 2014

**Jaroslava Schovancova, Torre Wenaus (BNL)
Debbie Bard, Jim Chiang, Seth Digel, Richard
Dubois, Warren Focke, *Tony Johnson*, Brian Van
Klaveren (SLAC)**

Contents

- Current goals of “Imsim on Demand”
- What currently exists (proof of principle)
- Technologies used
- Future plans?
- Request for input

Imsim On Demand goals

(Richard Dubois @ Dec '13 DESC meeting, Pitt)

- PhoSim is easy to set up and run on a small system
 - Much harder to manage production on thousands of cores and worse on big MPI machines
- Work towards a simple web interface for running ImSim on demand, supplying a catalogue and PhoSim parameters
 - Fire off jobs into the ether and notify the user when done (or if action is needed to rerun any failed jobs)
- Make use of existing workflow tools
 - Workflow engine to run graph of jobs
 - data catalogue to record where the data went
 - Interfaces to GRID, HPC, clouds etc to find the cycles

No need for user to obtain grid certificate or learn intricacies of installing tools or running on Grid or at SuperComputing Center

What currently exists

Parameter	Value
Job Name	%u-%t-%n Arbitrary name: %u=user name, %t=job type, %n=unique id
Catalog URL	https://dl.dropboxusercontent.com/u/12513001/bdgalb-fixed.dat
Sensor Regular Expression	R11_S_
Extra Commands	\$(PHOSIMDIR)/examples/quickbackground
Debug Mode	True
User Comment	R11_S%% with quickbackground
Expert Options	
Target Site	Panda (BNL) The site where the jobs will be run.

- <http://portal.lsstdesc.org/ImsimOnDemand>
 - Login with DESC confluence username/password
- Beta-level “Proof of principle”

What currently exists

Imsim on Demand

History

Imsim on Demand

Welcome to Imsim on Demand.

Parameter	Value
Job Name	<input type="text" value="%u-%t-%n"/> Arbitrary name: %u=user name, %t=job type, %n=unique id
Catalog URL	<input type="text" value="https://dl.dropboxusercontent.com/u/12513001/bdgals-fixed.dat"/>
Sensor Regular Expression	<input type="text" value="R11_S.."/>
Extra Commands	<input type="text" value="\${PHOSIMDIR}/examples/quickbackground"/>
Debug Mode	<input type="text" value="True"/>
User Comment	<input type="text" value="R11_S%% with quickbackground"/>
Expert Options	<input type="text" value="help"/>
Target Site	<input type="text" value="Panda (BNL)"/> The site where the jobs will be run.

What currently exists

Imsim on Demand

History

Confirm Imsim on Demand Settings

Messages

Your catalog URL has content length: 95,062,903

Your sensor regular expression matches 9 sensors

Parameter	Value
Job Name	%u-%t-%n
Catalog URL	https://dl.dropboxusercontent.com/u/12513001/bdgals-fixed.dat
Sensor Regular Expression	R11_S..
Extra Commands	\$(PHOSIMDIR)/examples/quickbackground
Debug	true
User Comment	R11_S%% with quickbackground
Expert Options	
Target Site	Panda (BNL)
<input type="button" value="Back"/>	<input type="button" value="Submit"/>

What currently exists

Imsim on Demand

History

Imsim on Demand job submitted

Your job tonyj-ImsimOnDemand-00019 has been submitted.

Your data will be available for download from <http://srs.slac.stanford.edu/DataCatalog/?experiment=LSST-DESC&folderPath=/LSST-DESC/ImsimOnDemand/1403121926840>

You will be sent an e-mail at tony_johnson@slac.stanford.edu when your job has completed.

You can monitor your job's progress using the [Pipeline](#)

Dear tonyj,
Job tonyj-ImsimOnDemand-00017 has finished. You can reach your data via:

<http://srs.slac.stanford.edu/DataCatalog/?experiment=LSST-DESC&folderPath=/LSST-DESC/ImsimOnDemand/1403113923444>

Please use: tonyj@slac.stanford.edu or reply to this e-mail with any concerns/suggestions.

Sincerely Yours,
Imsim On Demand server



SRS Download Manager

File Edit Help

Add... Remove Details Clean Up Start downloading Pause downloading

File Name	File Size	Status
lsst_a_87393588_f2_R22_S11_C05_E001_lsstsimfits.fits.gz	818.5 kB	Done
lsst_a_87393588_f2_R22_S11_C06_E000_lsstsimfits.fits.gz	801.0 kB	Done
lsst_a_87393588_f2_R22_S11_C06_E001_lsstsimfits.fits.gz	801.2 kB	Done
lsst_a_87393588_f2_R22_S11_C07_E000_lsstsimfits.fits.gz	808.2 kB	Done
lsst_a_87393588_f2_R22_S11_C07_E001_lsstsimfits.fits.gz	808.5 kB	Done
lsst_a_87393588_f2_R22_S11_C10_E000_lsstsimfits.fits.gz	810.3 kB	Downloaded
lsst_a_87393588_f2_R22_S11_C10_E001_lsstsimfits.fits.gz	810.5 kB	Queued
lsst_a_87393588_f2_R22_S11_C11_E000_lsstsimfits.fits.gz	798.6 kB	Queued
lsst_a_87393588_f2_R22_S11_C11_E001_lsstsimfits.fits.gz	798.5 kB	Queued
lsst_a_87393588_f2_R22_S11_C12_E000_lsstsimfits.fits.gz	789.6 kB	Queued
lsst_a_87393588_f2_R22_S11_C12_E001_lsstsimfits.fits.gz	789.4 kB	Queued

Progress
Total Download Size: 27.0 MB Downloaded So Far: 13.5 MB
Download Rate: 386.1 kB/sec Time Remaining: 00:35



Folder [/LSST-DESC/ImsimOnDemand/1403115818663](#)

[Edit description](#)

Name	Type	Files	Events	Size	Created (UTC)	Links
lsst_a_87393588_f2_R22_S11_C00_E000.fits.gz	Dataset	1	0	815.9 kB	18-Jun-2014 18:30:11	Download
lsst_a_87393588_f2_R22_S11_C00_E001.fits.gz	Dataset	1	0	815.5 kB	18-Jun-2014 18:35:28	Download
lsst_a_87393588_f2_R22_S11_C01_E000.fits.gz	Dataset	1	0	808.2 kB	18-Jun-2014 18:30:14	Download
lsst_a_87393588_f2_R22_S11_C01_E001.fits.gz	Dataset	1	0	808.4 kB	18-Jun-2014 18:35:31	Download
lsst_a_87393588_f2_R22_S11_C02_E000.fits.gz	Dataset	1	0	804.1 kB	18-Jun-2014 18:30:17	Download
lsst_a_87393588_f2_R22_S11_C02_E001.fits.gz	Dataset	1	0	804.5 kB	18-Jun-2014 18:35:34	Download
lsst_a_87393588_f2_R22_S11_C03_E000.fits.gz	Dataset	1	0	831.5 kB	18-Jun-2014 18:30:20	Download
lsst_a_87393588_f2_R22_S11_C03_E001.fits.gz	Dataset	1	0	831.3 kB	18-Jun-2014 18:35:37	Download
lsst_a_87393588_f2_R22_S11_C04_E000.fits.gz	Dataset	1	0	811.4 kB	18-Jun-2014 18:30:22	Download
lsst_a_87393588_f2_R22_S11_C04_E001.fits.gz	Dataset	1	0	811.4 kB	18-Jun-2014 18:35:40	Download
lsst_a_87393588_f2_R22_S11_C05_E000.fits.gz	Dataset	1	0	818.4 kB	18-Jun-2014 18:30:25	Download
lsst_a_87393588_f2_R22_S11_C05_E001.fits.gz	Dataset	1	0	818.5 kB	18-Jun-2014 18:35:43	Download
lsst_a_87393588_f2_R22_S11_C06_E000.fits.gz	Dataset	1	0	801.0 kB	18-Jun-2014 18:30:28	Download
lsst_a_87393588_f2_R22_S11_C06_E001.fits.gz	Dataset	1	0	801.2 kB	18-Jun-2014 18:35:46	Download
lsst_a_87393588_f2_R22_S11_C07_E000.fits.gz	Dataset	1	0	808.2 kB	18-Jun-2014 18:30:31	Download
lsst_a_87393588_f2_R22_S11_C07_E001.fits.gz	Dataset	1	0	808.5 kB	18-Jun-2014 18:35:49	Download
lsst_a_87393588_f2_R22_S11_C10_E000.fits.gz	Dataset	1	0	810.3 kB	18-Jun-2014 18:30:34	Download
lsst_a_87393588_f2_R22_S11_C10_E001.fits.gz	Dataset	1	0	810.5 kB	18-Jun-2014 18:35:51	Download
lsst_a_87393588_f2_R22_S11_C11_E000.fits.gz	Dataset	1	0	798.6 kB	18-Jun-2014 18:30:37	Download
lsst_a_87393588_f2_R22_S11_C11_E001.fits.gz	Dataset	1	0	798.5 kB	18-Jun-2014 18:35:54	Download
lsst_a_87393588_f2_R22_S11_C12_E000.fits.gz	Dataset	1	0	789.6 kB	18-Jun-2014 18:30:40	Download
lsst_a_87393588_f2_R22_S11_C12_E001.fits.gz	Dataset	1	0	789.4 kB	18-Jun-2014 18:35:57	Download
lsst_a_87393588_f2_R22_S11_C13_E000.fits.gz	Dataset	1	0	789.2 kB	18-Jun-2014 18:30:43	Download
lsst_a_87393588_f2_R22_S11_C13_E001.fits.gz	Dataset	1	0	788.9 kB	18-Jun-2014 18:36:00	Download
lsst_a_87393588_f2_R22_S11_C14_E000.fits.gz	Dataset	1	0	810.8 kB	18-Jun-2014 18:30:46	Download
lsst_a_87393588_f2_R22_S11_C14_E001.fits.gz	Dataset	1	0	810.3 kB	18-Jun-2014 18:36:03	Download
lsst_a_87393588_f2_R22_S11_C15_E000.fits.gz	Dataset	1	0	795.3 kB	18-Jun-2014 18:30:49	Download
lsst_a_87393588_f2_R22_S11_C15_E001.fits.gz	Dataset	1	0	794.9 kB	18-Jun-2014 18:36:06	Download
lsst_a_87393588_f2_R22_S11_C16_E000.fits.gz	Dataset	1	0	799.7 kB	18-Jun-2014 18:30:51	Download
lsst_a_87393588_f2_R22_S11_C16_E001.fits.gz	Dataset	1	0	800.1 kB	18-Jun-2014 18:36:09	Download
lsst_a_87393588_f2_R22_S11_C17_E000.fits.gz	Dataset	1	0	802.1 kB	18-Jun-2014 18:30:54	Download
lsst_a_87393588_f2_R22_S11_C17_E001.fits.gz	Dataset	1	0	802.3 kB	18-Jun-2014 18:36:12	Download
lsst_e_87393588_f2_R22_S11_E000.fits.gz	Dataset	1	0	953.7 kB	18-Jun-2014 18:30:05	Download
lsst_e_87393588_f2_R22_S11_E001.fits.gz	Dataset	1	0	952.6 kB	18-Jun-2014 18:35:23	Download

[List Files](#) · [Download Files](#)

What currently exists

Imsim on Demand

History

Full History

Note: Clicking on the Status column will take you to the pipeline task that ran the job. Clicking on the Job column will allow you to rerun this task, or a similar one. Clicking on the Output Directory column will take you to the output.

Submit Time	Job	User	Task Type	Status	Output Directory	User Comment
18-Jun-2014 13:09	tonyj-ImsimOnDemand-00019	tonyj	ImsimOnDemand	Running	View dir	R11_S%% with quickbackground
18-Jun-2014 11:52	tonyj-ImsimOnDemand-00018	tonyj	ImsimOnDemand	Terminated	View dir	R11_S%% with quickbackground
18-Jun-2014 11:24	djbard-test-1	djbard	ImsimOnDemand	Success	View dir	does it work at BNL?
18-Jun-2014 11:23	djbard-test-0	djbard	ImsimOnDemand	Success	View dir	does it work at slac?
18-Jun-2014 10:52	tonyj-ImsimOnDemand-00017	tonyj	ImsimOnDemand	Success	View dir	R11_S01 at BNL (Test manual copy)
18-Jun-2014 09:33	tonyj-ImsimOnDemand-00016	tonyj	ImsimOnDemand	Failed	View dir	R11_S01 at BNL (Test manual copy)
18-Jun-2014 04:46	tonyj-ImsimOnDemand-00015	tonyj	ImsimOnDemand	Success	View dir	R11_S01 at BNL (Test file transformation)
17-Jun-2014 17:25	tonyj-ImsimOnDemand-00014	tonyj	ImsimOnDemand	Success	View dir	R12_S01 at BNL
17-Jun-2014 17:22	tonyj-ImsimOnDemand-00013	tonyj	ImsimOnDemand	Success	View dir	R12_S01 at SLAC
17-Jun-2014 16:41	tonyj-ImsimOnDemand-00012	tonyj	ImsimOnDemand	Success	View dir	R12_S01 at SLAC
16-Jun-2014 15:48	jschovan-ImsimOnDemand-00003	jschovan	ImsimOnDemand	Success	View dir	test.jschovan
16-Jun-2014 14:30	jschovan-ImsimOnDemand-00002	jschovan	ImsimOnDemand	Success	View dir	test.jschovan
16-Jun-2014 14:14	tonyj-ImsimOnDemand-00011	tonyj	ImsimOnDemand	Success	View dir	R12_S01 at SLAC

Tools currently used

- SLAC workflow engine
- Web interface and Data Catalog
 - Originally developed for Fermi Gamma-Ray telescope
- Panda distributed analysis system
 - Grid interface (originally used by Atlas)
- Xrootd globally accessible data storage
 - At BNL and SLAC

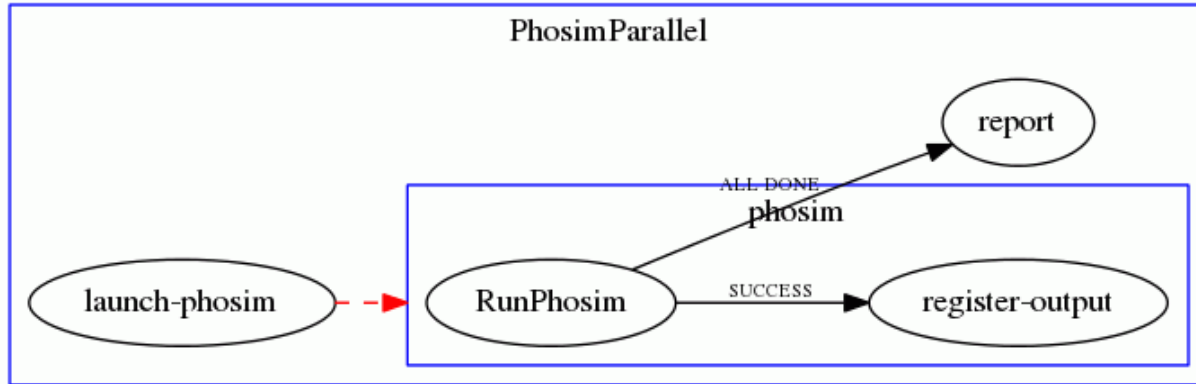
Panda and Grid Tools

- Panda -- Atlas Workload Management and Analysis System
 - Being generalized under ASCR + DOE HEP funded multi-year effort
 - LSST DESC project is early “pilot” of Panda’s use beyond Atlas
 - LSST VO set up -- currently @ BNL, next @ UT Arlington, ...
- CERNVM File System (CVMFS)
 - Allows one central installation of the software to be propagated to Grid sites
 - /cvmfs/oasis.opensciencegrid.org/lstt/test.lsst/bnl/phosim/3.3.2/phosim-3.3.2
- Data Access: xrootd
 - HEP standard distributed data access tool called xrootd (SLAC developed)
 - Exposes data stores to distributed (as well as local) clients, soon via http clients as well as via xrootd tools
 - Also supported by SLAC data catalog and download manager

More Details: https://docs.google.com/presentation/d/1zRez2sUZYTdqFpUaSGPyJviXfO331qXBam2S1_O6k7s/edit?usp=sharing

Workflow Engine

Allows arbitrary workflows to be defined. Can include embedded jython scripts to control flow (for example optional steps) and supports parallelism



Flexible job submission architecture allows jobs to be submitted to various sites, including SLAC, IN2P3, Grid (Panda, Dirac), and NERSC.

Panda + Workflow engine integration

Workflow engine and Panda support job monitoring, now integrated via links



summary / PhosimParallel / phosim

Task phosim Stream 12.0

Rollback Stream

Stream 12.0

Execution 1

Is Latest 1

Status Success

Submitted 17-Jun-2014 17:44:08.135

Started 17-Jun-2014 17:46:47.547

Ended 17-Jun-2014 18:15:17.675

Variables

Name	Type	Value
SENSOR	String	R12_S01

Stream Processes

Show only latest execution

Process	Status	Type	Created	Submitted	Started	Ended	Job id	CPU	Host	Links
register-output	Success	Script	17-Jun-2014 17:44:08		17-Jun-2014 18:15:15	17-Jun-2014 18:15:17				Messages
RunPhosim	Success	Batch	17-Jun-2014 17:44:08	17-Jun-2014 17:44:25	17-Jun-2014 17:46:45	17-Jun-2014 18:14:56	2569	1657.14	acas1508.usatlas.bnl.gov	Messages : Panda

Substreams

Nothing found to display.

All Substreams

Task	Process	▽	⤴	⤵	⤶	⤷	⤸	⤹	⤺	⤻	⤼	Total
phosim	RunPhosim	0	0	0	0	0	1	0	0	0	0	1
	register-output	0	0	0	0	0	1	0	0	0	0	1
Totals		0	0	0	0	0	2	0	0	0	0	2

Version 4.1-SNAPSHOT | Jira (Front-End) (Server) | Help

Page updated: 06/19/2014 04:40:47

Start refreshing page every secs [Start Refreshing](#)

User: tonyj (Switch/Logout) Mode: [Prod] [Dev] [Test] Preferences

[Task List](#) [Message Viewer](#) [Usage Plots](#) [Fair Share Plots](#) [Admin](#) [JMX](#)

PanDA monitor

Job details for PanDA job 2569 Generated June 19, 2014, 11:41 a.m. UTC

PanDA Attempt#	Owner / VO	TaskID	Status	Created	Start	End	Mod	Site	Priority
2569 Attempt 1	tonyj / lsst	1	finished	2014-06-18 00:44	06-18 00:46	06-18 01:20	06-18 01:25	ANALY_BNL-LSST	2000

Job name: job.30649200.phosim.1.tonyj.0 #1

Status **finished** indicates that the job has successfully completed.

LSST job info: Pipeline task: **phosim** process: **30649200**

View the job's [stdout](#), [job outputs](#)

View the job wrapper's (verbose) [stdout](#)

[Download the job cache tarball](#) containing the job execution scripts

View the pilot job's [stdout](#), [stderr](#), [batch log](#)

[Look for PanDA logger records for this job](#)

1 job files

Filename (Type)	Size (bytes)	Status	Dataset
job.30649200.phosim.1.tonyj.0.job.log.tgz (log)	49294717	ready	panda.lsst.user.30649200.phosim.tonyj

Other key job parameters

Processing type	phosim
Job type	panda
Payload script (transformation)	lsst-trf.sh
Attempt number	1 of a maximum 3
Output destination	local
CPU consumption time (s)	1643
Job parameters	--lsstJobParams="DP_USER=tonyj PIPELINE_TASK=phosim PIPELINE_PROCESSINSTANCE=30649200 PIPELINE_EXECUTIONNUMBER=1 PIPELINE"



(Possible) Future Plans

- Incorporate Catsim
- Allow appending multiple catalogs
- More Imsim options (what would be most useful?)
- More interactive web interface (for example select rafts from diagram)
- Ability to view images without downloading
 - For example by using JS9
- Ability to run DMStack tools on images?
 - Producing data as fits, database?
- Alternatives to full phosim?
 - Phenomenological simulation?
- Extend to more than a single image
- Extend to run at more sites
 - Other sites in DESC VO, IN2P3, NERSC
- What would be most useful to you?

Conclusions

- A “Proof of Principle” version of Imsim on Demand now exists
 - Please try it out and give feedback
 - Does it work, does it do what you would like?
- We need input from working groups, individuals on what features would be most useful to add
 - We have fairly regular meetings -- anyone welcome
 - Contact me (tonyj@slac.stanford.edu) to be kept in the loop