Trident: Scalable Compute Archive and Analysis Systems

Arvind Gopu

Soichi Hayashi, Mike Young, Ralf Kotulla*, Robert Henschel, et al.

Indiana University – Pervasive Technology Institute/Research Technologies

*University of Wisconsin, Madison

March 11 2015

Big Data in Astronomy Workshop, Tucson, AZ

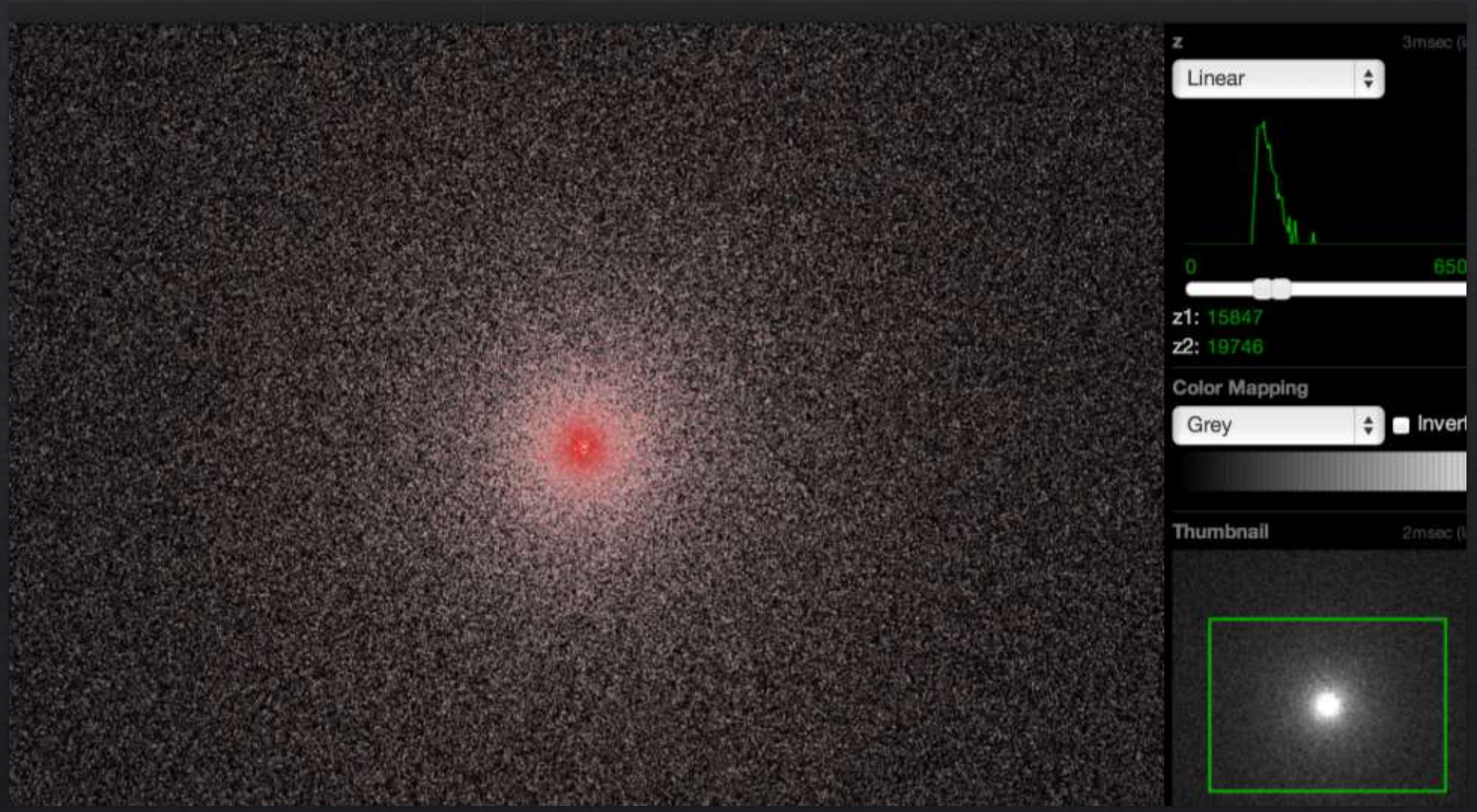
Trident in a Nutshell

- **Secure data archive**
- Metadata Search, View, Plotting Analysis
  - Build custom searches; Highly adaptive plotting interface
- Image Visualization and Basic Interactive Analysis
- Custom Integrated Data Processing Workflows
  - Cater to individual stakeholder, application, and audience
  - User/Operator execution; Datasets/collections
  - Built-in Data *Provenance* and Processing Logs → *Reproducibility*
- Scale up/down in terms of functionality, system complexity, etc.
  - **Adapts** to Big Data, and nominally large or small data
  - Develop component for each distinct functionality; then compose application instance ("Composition over Inheritance" design pattern)

DEMO!

- Video showing 4 features that demonstrate how Trident (using ODI-PPA and EMC-SCA as examples) serves as a Scalable Compute Archive with Visualization & Analysis Features
  - https://iu.box.com/ODI-PPA-Demo-2015-Short
- Plus a few screenshots
- Not Shown: Lots of other functions including
  - Administrative/operator/system health monitoring
  - GCS-SCA, etc.

Source Explorer: Overlay (Zoomed out)

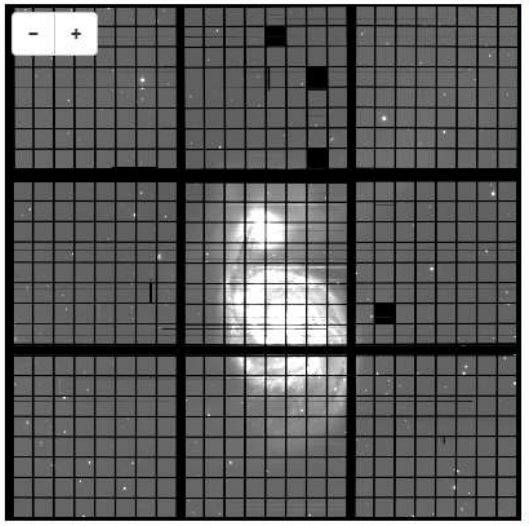


Source Explorer: Overlay (Zoomed In)

Image Preview, Metadata, QA, Provenance

m51 g Calibrated > OBSID: 20130511T235739.1

Add To Collection | Explore Image



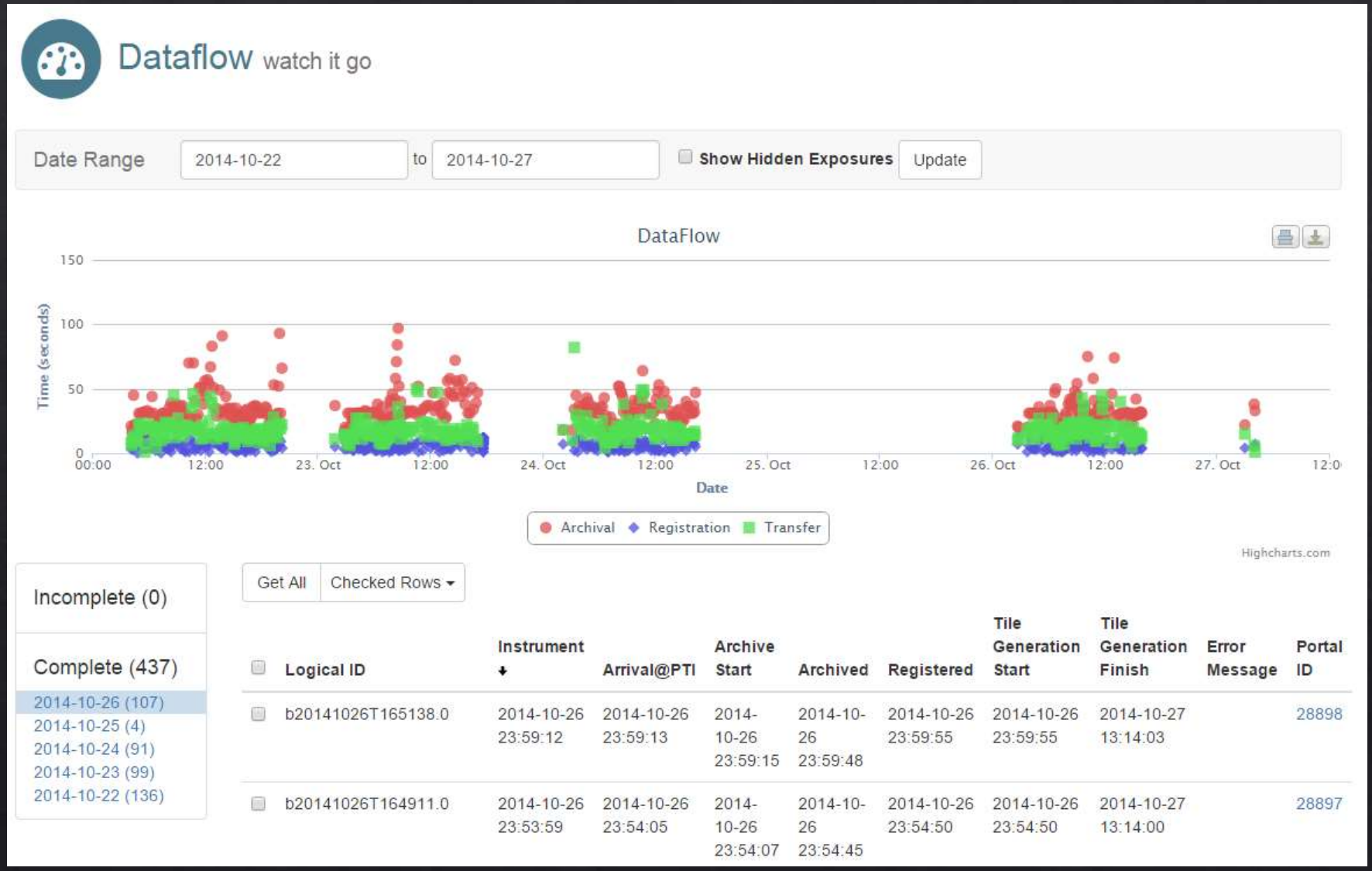
Fits Header | photZP | photZP_map | psfshape | seeing | wcs1 | wcs2

```
XTENSION  'IMAGE   '           / Image extension
BITPIX                  -32 / array data type
NAXIS                     2 / number of array dimensions
NAXIS1                 4096
NAXIS2                 4096
PCOUNT                    0 / number of parameters
GCOUNT                    1 / number of groups
QPIPESVN  'exported'        / QuickReduce Revision
EXTNAME   'OTA33.SCI'
OTA                      33 / OTA designation
FPPOS     'xy33    '        / position of detector in focal plane mosaic
PIXSIZE1               12.0 / pixel size of axis1 (microns)
PIXSIZE2               12.0 / pixel size of axis2 (microns)
```

Associations

		OBJECT	PROPID	EXPTIME	RA	DEC	PHOTDPTH	SEEING	WCSCAL	CAL-OBS
Parent images										
Raw Image		m51 g	2013A-0576	400	202.5851	47.2922	0	0	0	2013-05-11
Bias		master-bias	test	0.001	0	0	0	0	0	2013-05-06
Flat		master-flat odi_g	test	5	0	0	0	0	0	2013-05-08
Dark		master-dark	2013A-0468	1	0	0	0	0	0	2013-05-06
Child Images										
Stacks		m51 g		3200.01	202.4847	47.2374	0	0	0	2013-05-11

Instrument Data Flow Monitoring



DEMO!

- Briefly, let’s get back to the demo!

Trident: Background

- **Team**: Combined 0.75 – 2.5 FTE at various times at IU + 0.5 – 2 External Pipeline Developers over several years (~2 years core)
- **First Trident Project**: WIYN One Degree Imager – Pipeline, Portal, Archive
  - Larger datasets → Paradigm Shift
  - Prototype, User Engagement Workshops, Design Review, Dev & Testing, Ops
  - Scope *decrease* early on, then Scope *increase* (owing to detector expansion)

Trident Projects and Web URLs

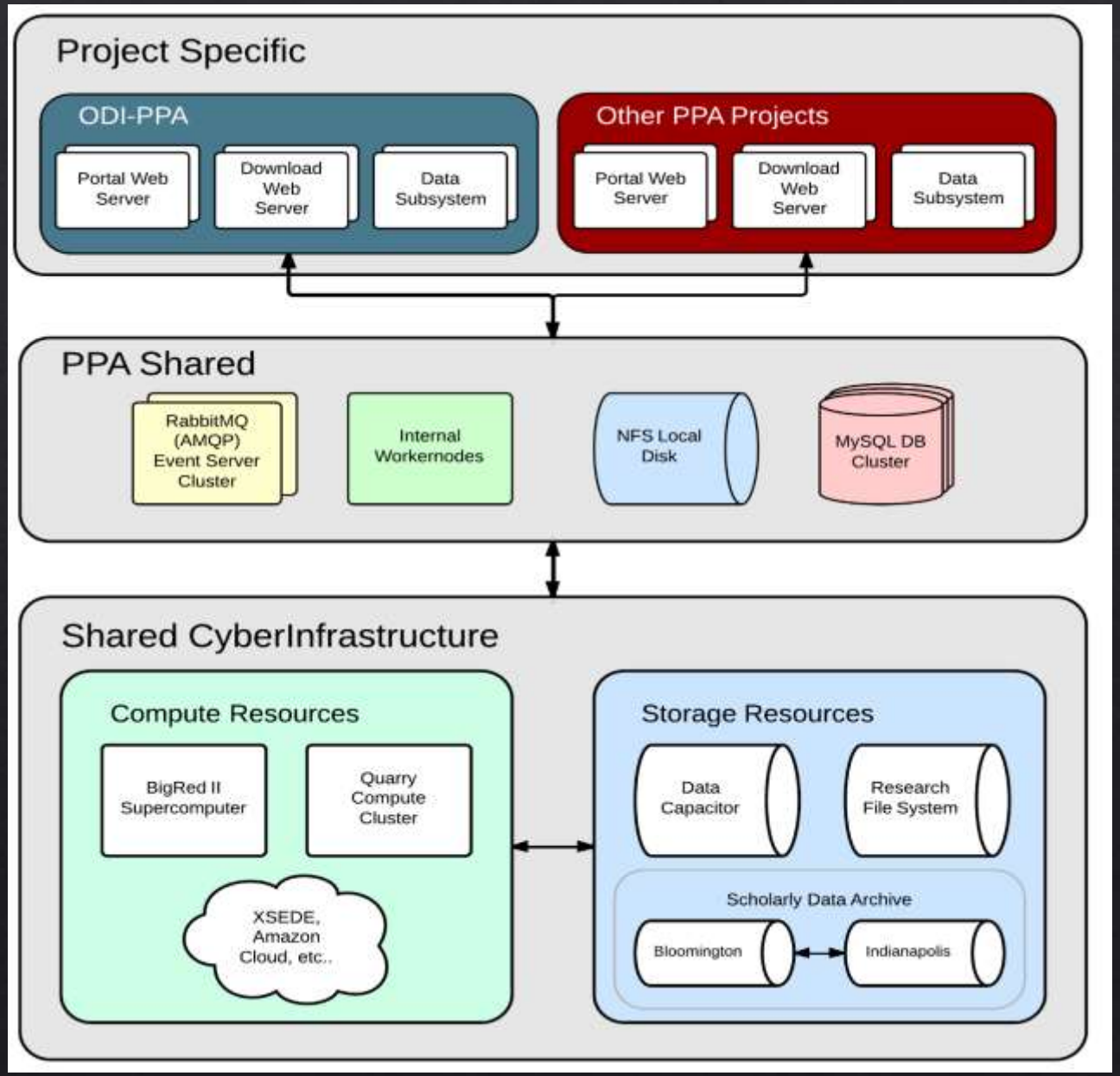
- ODI-PPA http://portal.odi.iu.edu
  - Full Fledged Portal with Metadata & Visual Image Analysis, Archival Data Access, Integrated Pipelines
- Offshoots
  - Electron Microscopy: http://portal.emcenter.iu.edu
    - Archival Data Access
  - Another Astronomy Project http://gcs.ppa.iu.edu
    - Metadata Analysis and Data Publication
- Prototype
  - LMU-PPA: Wendelstein Wide Field Imager (||| to ODI)
  - Infrared Astronomical Imaging
- Papers: http://ppa.iu.edu/publications

Trident: Architecture / Design Choices / Thought Process

- Web based Science Gateway model
  - Balance legacy UI features and modern UI design; Consistent look and feel
  - Provide desktop application features (and related backend processing)
  - Integrate pipelines on existing CI accessible via portal
    - Custom UI for each of these
- Often times, UI tends to take a backseat – usability becomes an afterthought
- Typical preference for generalized large scale web and middleware solutions: Pros and Cons
- Framework of Frameworks
  - Not *one* package/tarball/RPM model – instead, customized to each stakeholder's requirements.

Trident: Architecture / Design Choices / Thought Process

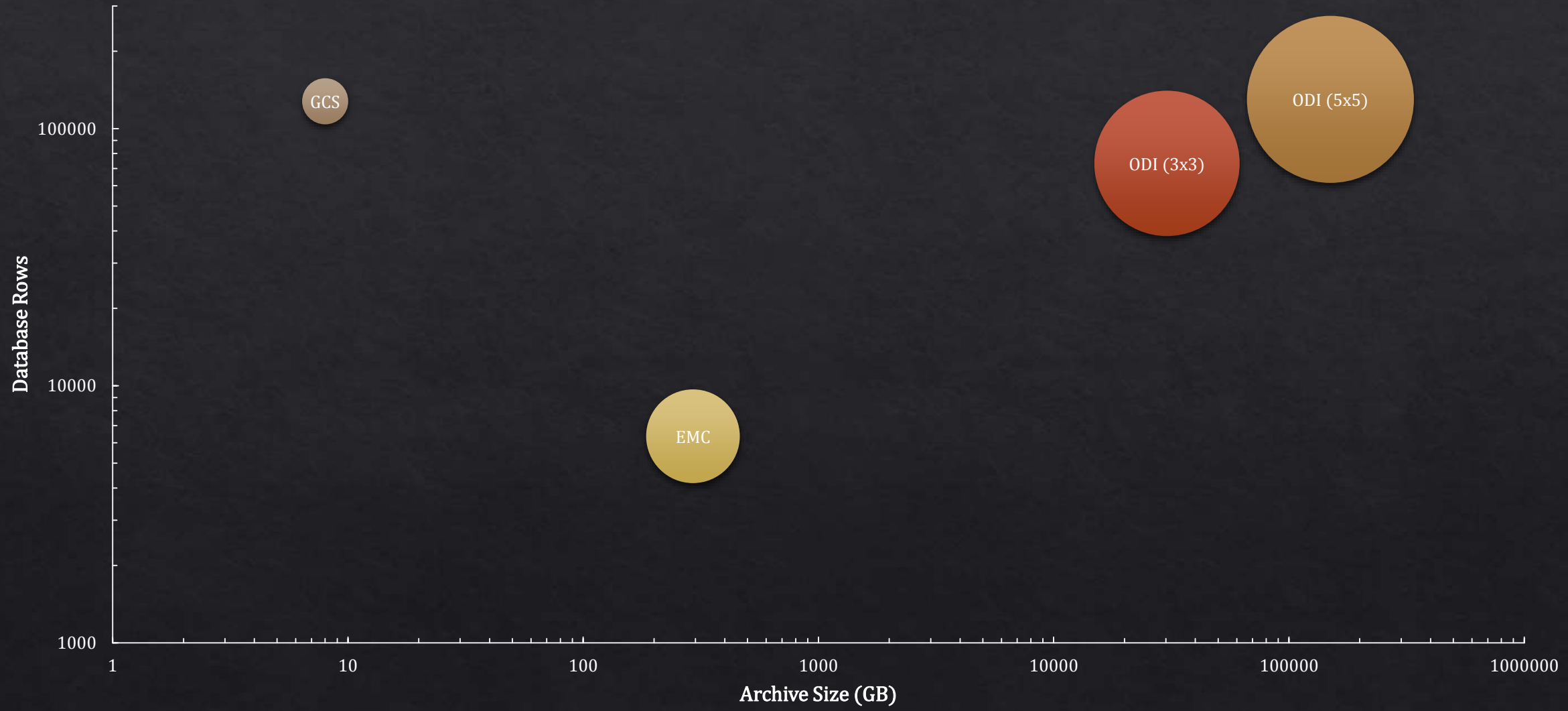
- MVC Design Pattern
- Light Weight Services + AMQP
  - + Portability, Ease of Use, Scalability
  - - Harder to track/monitor system health
- RHEL 6.5/Centos 7.x KVM
  - Separate production/test/dev stacks
  - Leveraged Project and Institutional CyberInfrastructure



Trident: Open Source Libraries and Applications Used

- Common Tools/Applications
  - Apache, MySQL, Cobbler, Puppet, Munin; Lustre; Confluence, JIRA
- Open Source *Libraries*
  - PHP-Zend, Bootstrap, AngularJS, jQuery, HighCharts, TileViewer
  - NumPy, SciPy, PyFITS, etc.
- Integrated Pipelines – incl. custom input and output UIs
  - QuickReduce (WIYN), SWarp, SourceExtractor, IRAF
  - ViZier for catalog lookup and overlay

Trident: Scalability

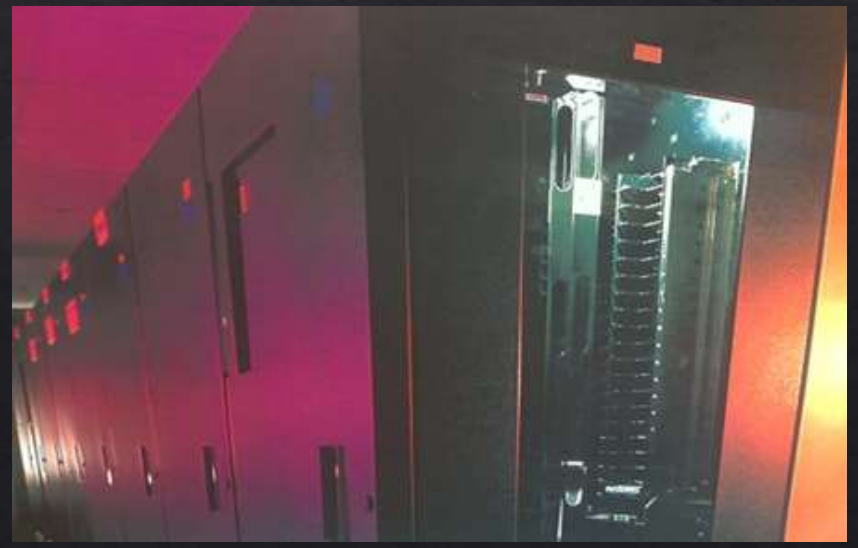


Keys to Scaling Up: DESIGN

- AMQP: No ports to open, no files to create/cleanup on shared file systems → Clean and Secure
- Custom UI and Integrated Applications
  - Plus Search Form Builder (for example)
- Any Unix Command Line Appl. can be integrated with custom UIs
  - Initial version: Integrated QR in 2 weeks ; SWarp in 1 week
  - Highly Customizable & Polished version: 4-8 weeks for development and testing
- Scaling Horizontally / Resource Scaling
  - A user workflow job could go to one of N backend instances
  - Add more resources as required

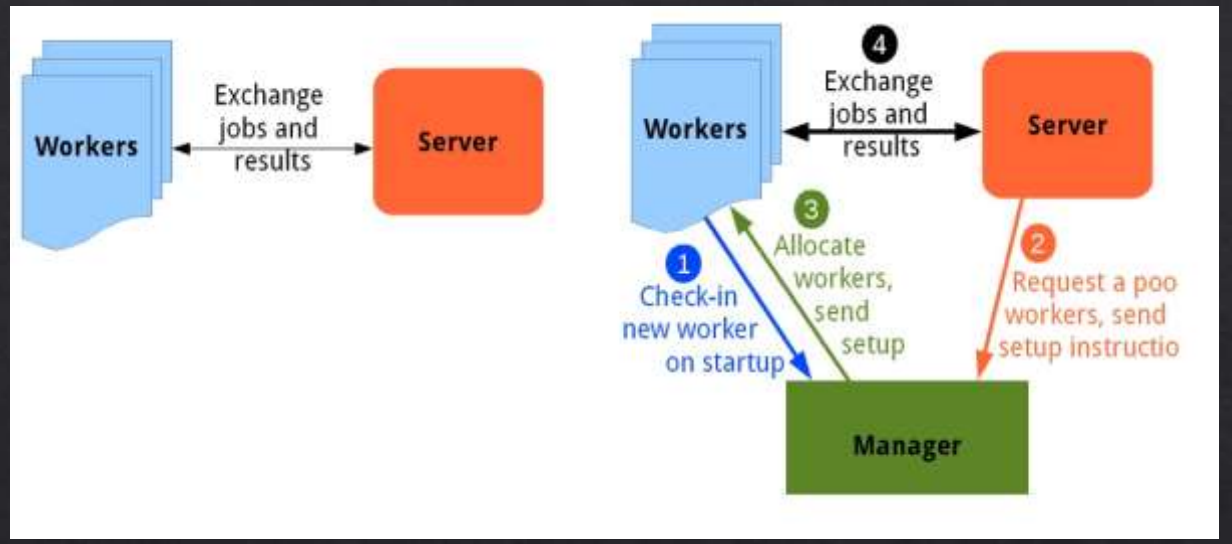
Keys to Scaling Up: LEVERAGED IU CYBERINFRASTRUCTURE

- Petaflop scale Big Red II Supercomputer
  - 32-core and 16-core nodes (latter with GPUs)
- 42 PB Scholarly Data Archive – geographically replicated secure tape archival system
- 5 PB Data Capacitor - high performance disk cache
- Multiple Virtual Machine options
- State of the Art Data Center with a ton of expansion space
  - Co-Location capabilities for collaborating group



Keys to Scaling Up: SCALABLE APPLICATIONS

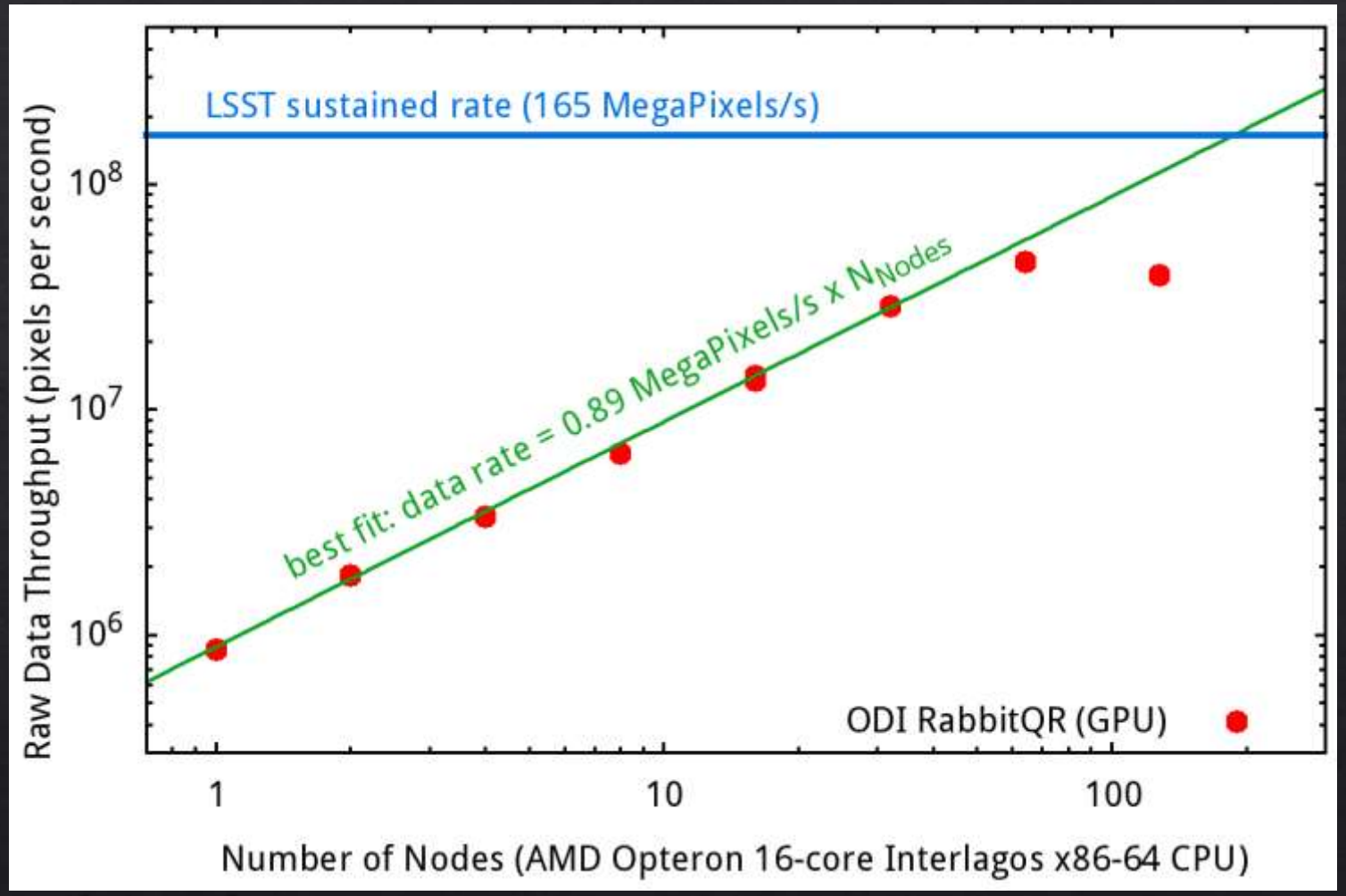
- Example: Parallelized Quick Reduce using AMQP
- Design: Server-Manager-Worker framework (vs. typical master/slave)
  - Advantage: Flexible # of nodes during execution, easy support for workflows
- Scaling: Most I/O work done on nodes, minimal communication between servers and workers → good scaling



- Caveat: Network-served data can be I/O bottleneck (even at 40GBit/s bandwidth)
- Benchmark Numbers: Performance scales linearly with number of nodes out to ~128 nodes40 MegaPixels/s raw data, limited by I/O

Trident + AMQP QR: Case Study of LSST Data Volumes

- LSST rate (165MPix/s) → would need 185 nodes
  - All internal I/O (inter-process comm, export for SourceExtractor) using memory or RAM-disk
- Investigating I/O bottleneck on shared cluster file system
- Network-I/O performance critical for future systems



- 50% of LSST needs today on shared supercomputer!

Thank You!

- Feel free to register on ODI-PPA: https://portal.odi.iu.edu
- Check out related papers: http://ppa.iu.edu/publications

Trident Project Acknowledgments

- We are grateful to several people/groups including (but not limited to):
  - The IU-PTI/RT Storage, HPS, HPFS teams that maintain the SDA tape system, the super computing clusters, and the Data Capacitor disk systems!
  - Thomas Lee, system administrator consultant
  - WIYN Collaborators including UW-Milwaukee/Madison and NOAO (ODI-PPA)
  - David Morgan, Operations Coordinator for IU EM Center (EMC-SCA)
  - Kathy Rhode, IU Astronomy Professor (GCS-SCA)
  - Claus Goessl, LMU, Munich (LMU-SCA)
  - Developers/maintainers of all the open source libraries and applications we use within the Trident!

ODI-PPA – Partner Organizations

ODI-PPA is a collaboration of the following organizations

- Pervasive Technology Institute (PTI) and UITS Research Technologies (RT) + IU Astronomy
  - Capitalize on the expertise of PTI members who have led the effort to provide scientists in many different fields with user-friendly access to super-computing facilities
  - IU Astronomy Expertise, Feedback
- University of Wisconsin (Astronomy)
  - Build on the experience of Python based pipeline development experience.
- WIYN
  - Experience running telescopes, and supporting Astronomy scientific community
- NOAO Science Data Management group
  - Build on the experience of SDM and the legacy of IRAF and NHPPS Pipeline system