THE UNIVERSITY OF CHICAGO

Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets

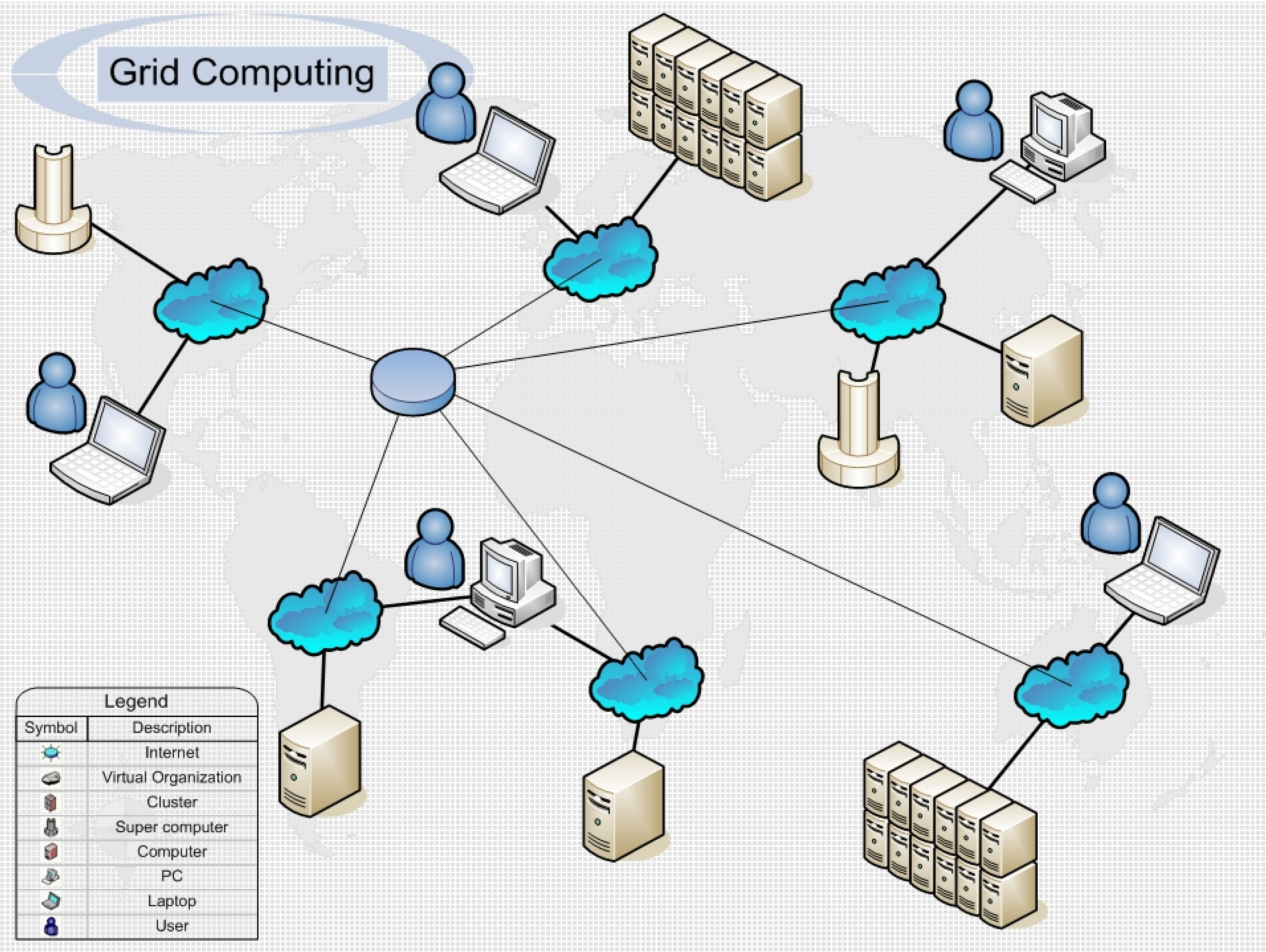
Ioan Raicu

Distributed Systems Laboratory
Computer Science Department
University of Chicago

ARGONNE NATIONAL LABORATORY

DSL Workshop 2006

June 2nd, 2006



Grid Computing

Legend	
Symbol	Description
	Internet
	Virtual Organization
	Cluster
	Super computer
	Computer
	PC
	Laptop
	User

Grid Computing

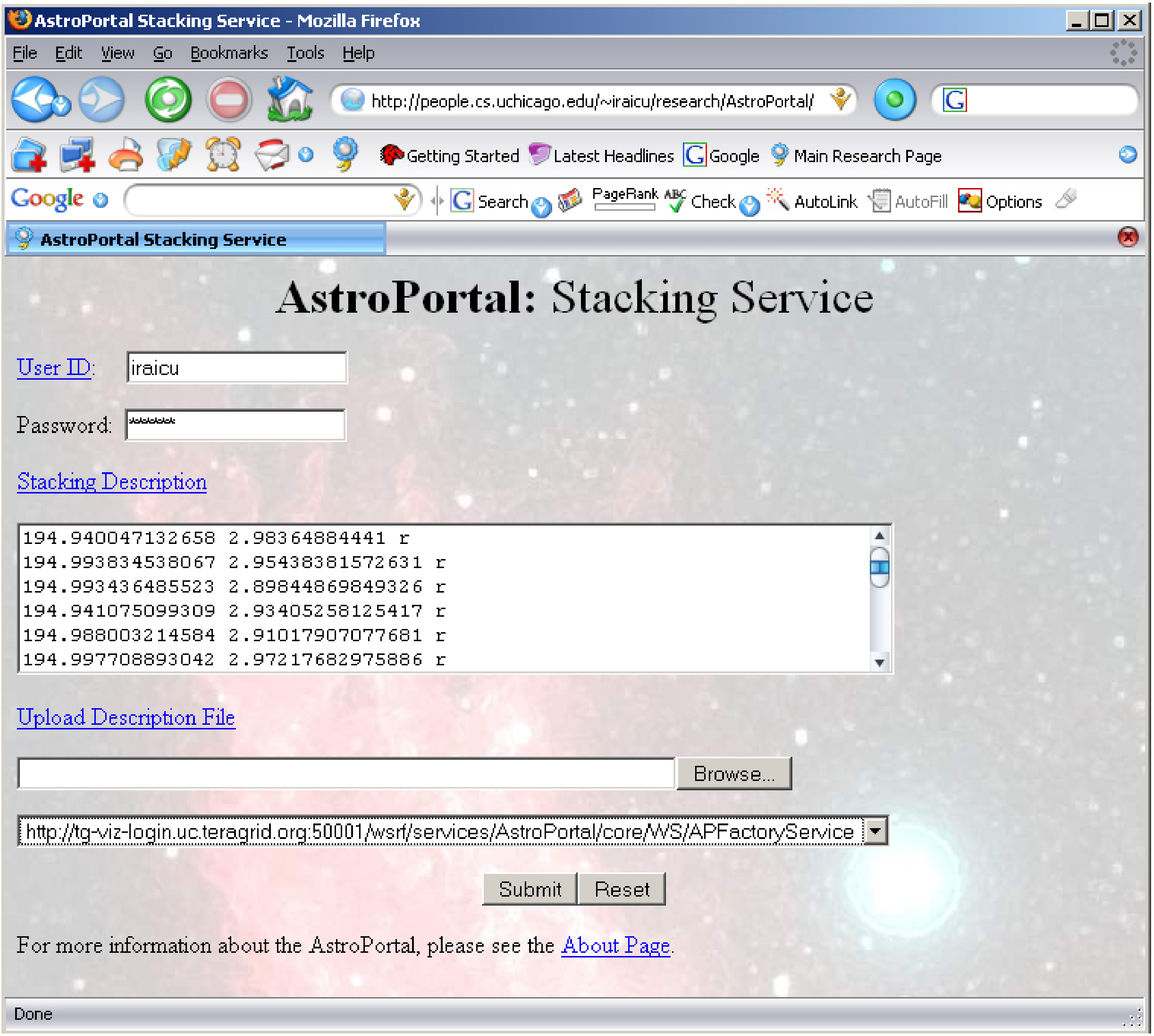
- Grid Computing's focus:
  - **large-scale resource sharing:** direct access to computers, software, data
  - innovative applications
  - high-performance orientation
- The 'Grid problem':
  - **Definition:** flexible, secure, and ***coordinated resource sharing among dynamic collections of individuals, institutions, and resources***
  - **Challenges:** Security (Authentication, Authorization), ***resource management (resource access, resource discovery, scheduling, data management)***

Introduction

- Science Portals: gateway to Grid resources
- Potential Applications Characteristics
  - Large data sets
  - Large number of users
  - Easy (but not necessarily trivial) parallelization
- Applicable fields:
  - Astronomy
  - Medicine
  - Others

Astronomy Field

- Astronomy datasets (i.e. SDSS) are the crown-jewels
  - SDSS DR4
    - 1.3M images
      - 300M+ objects
      - 3TB compressed images (2MB x 1.3M)
      - 8TB raw images (6.1MB x 1.3M)
    - 100K worldwide potential users
- Applications:
  - Stacking
  - Montage


AstroPortal Stacking Service - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://people.cs.uchicago.edu/~iraicu/research/AstroPortal/
Getting Started
Latest Headlines
Google
Main Research Page
Google
Search
PageRank
Check
AutoLink
AutoFill
Options
AstroPortal Stacking Service

AstroPortal: Stacking Service

User ID: iraicu

Password: ******

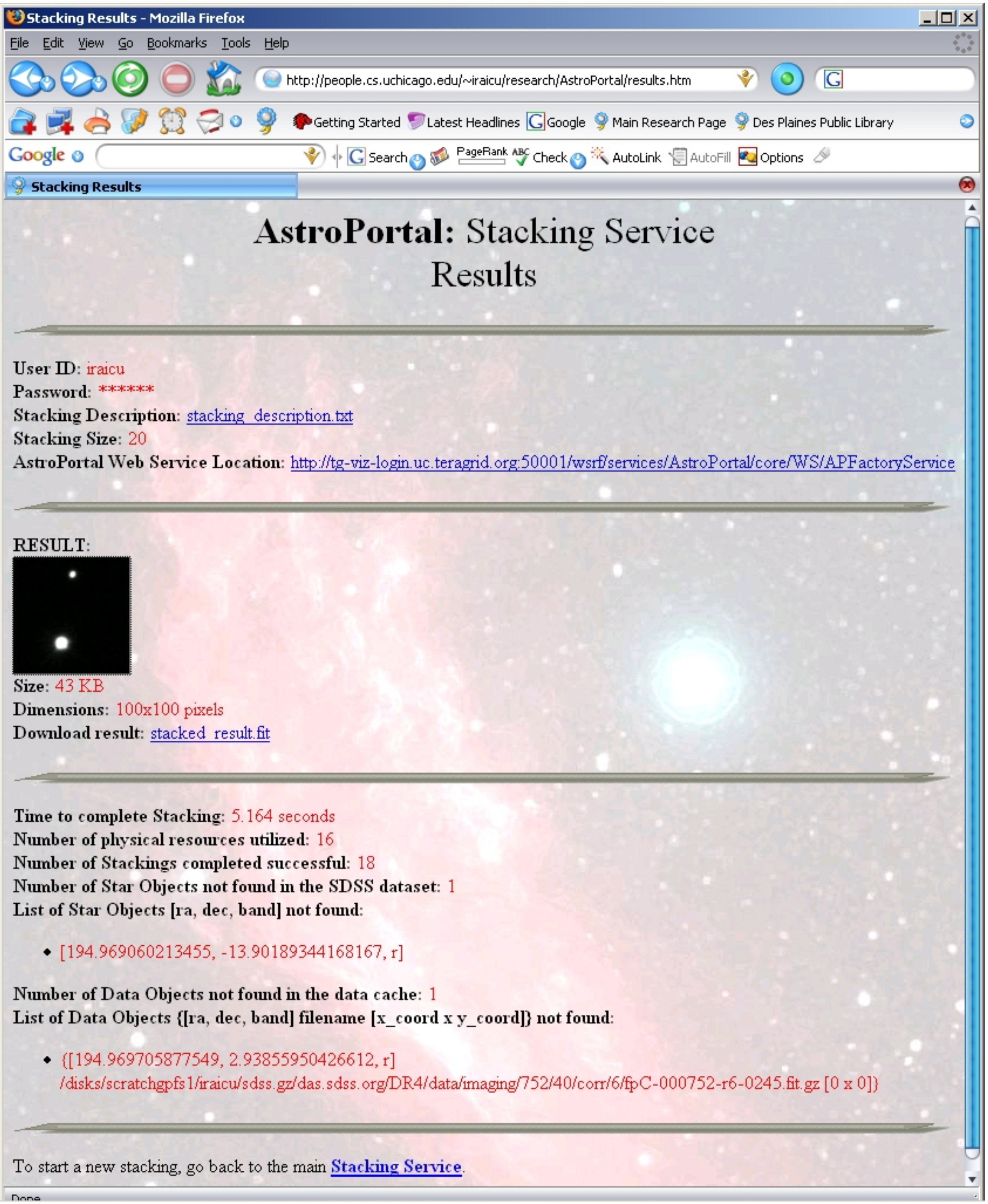
Stacking Description

```
194.940047132658 2.98364884441 r
194.993834538067 2.95438381572631 r
194.993436485523 2.89844869849326 r
194.941075099309 2.93405258125417 r
194.988003214584 2.91017907077681 r
194.997708893042 2.97217682975886 r
```

Upload Description File

http://tg-viz-login.uc.teragrid.org:50001/wsrf/services/AstroPortal/core/WS/APFactoryService

For more information about the AstroPortal, please see the About Page.


Stacking Results - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://people.cs.uchicago.edu/~iraicu/research/AstroPortal/results.htm
Getting Started
Latest Headlines
Google
Main Research Page
Des Plaines Public Library
Google
Search
PageRank
Check
AutoLink
AutoFill
Options
Stacking Results
Done

AstroPortal: Stacking Service Results

User ID: iraicu
Password: ******
Stacking Description: stacking_description.txt
Stacking Size: 20
AstroPortal Web Service Location: http://tg-viz-login.uc.teragrid.org:50001/wsrf/services/AstroPortal/core/WS/APFactoryService

RESULT:

Size: 43 KB
Dimensions: 100x100 pixels
Download result: stacked_result.fit

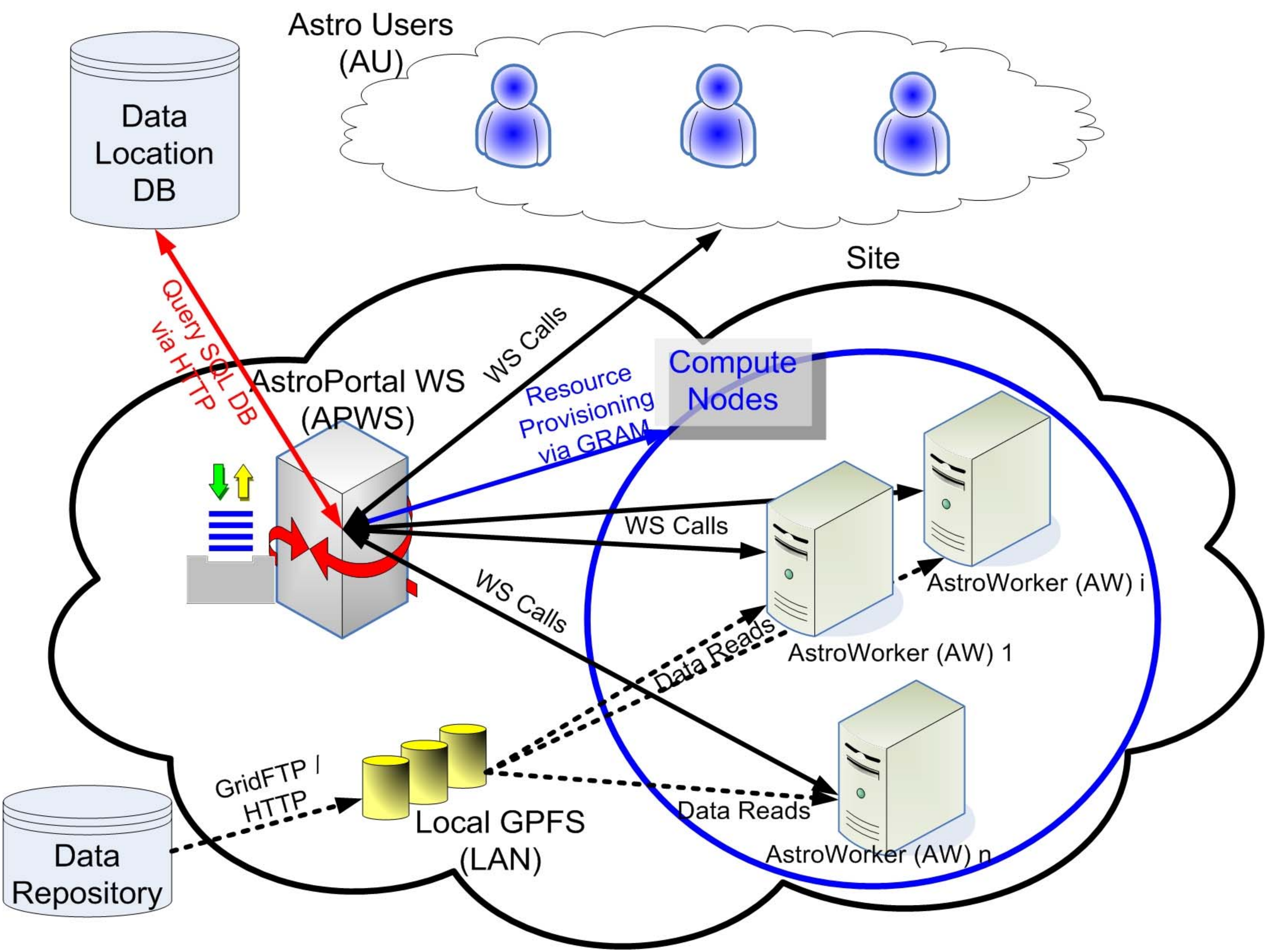
Time to complete Stacking: 5.164 seconds
Number of physical resources utilized: 16
Number of Stackings completed successful: 18
Number of Star Objects not found in the SDSS dataset: 1
List of Star Objects [ra, dec, band] not found:

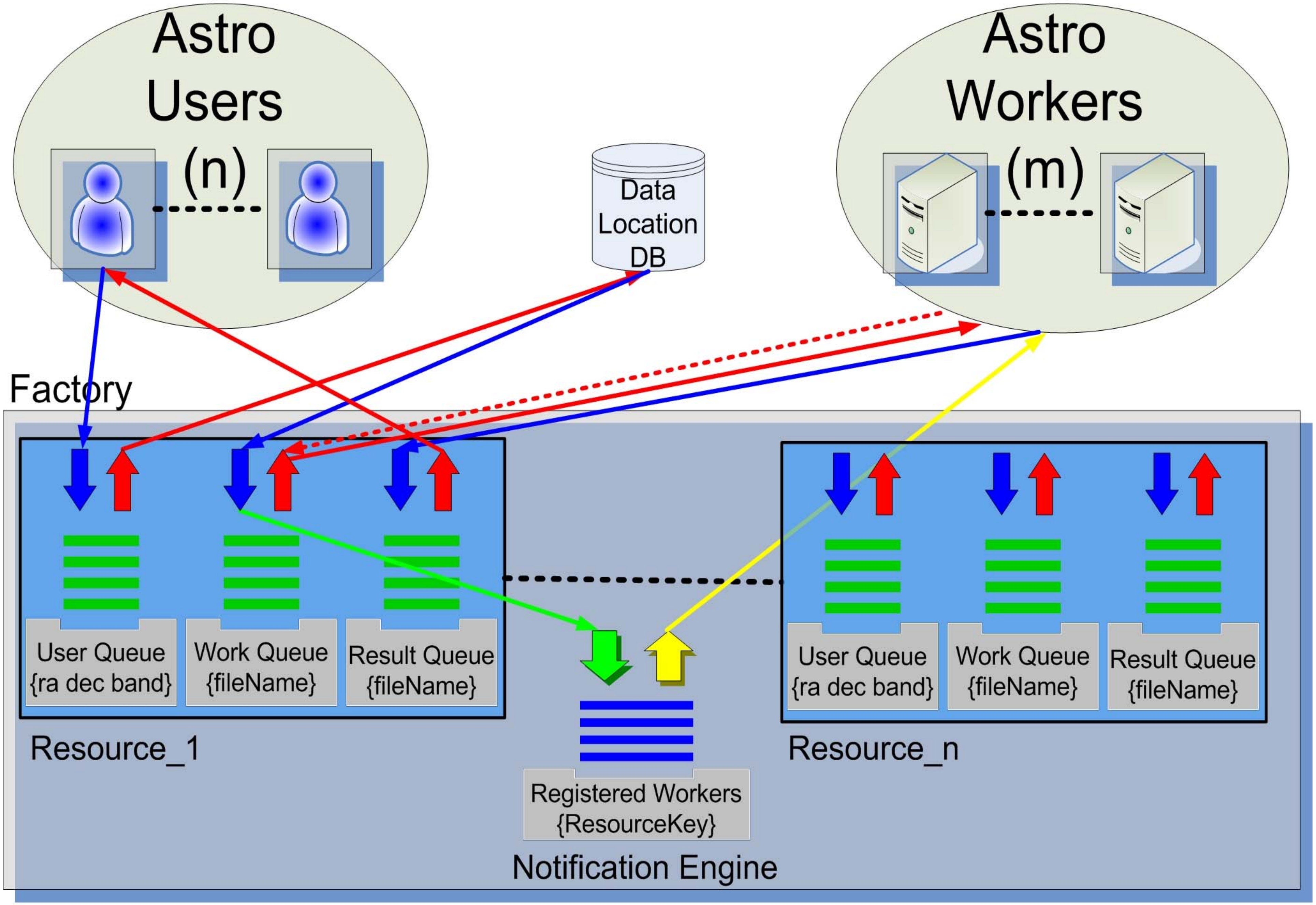
- [194.969060213455, -13.90189344168167, r]

Number of Data Objects not found in the data cache: 1
List of Data Objects {[ra, dec, band] filename [x_coord x y_coord]} not found:

- {[194.969705877549, 2.93855950426612, r] /disks/scratchgpfs1/iraicu/sdss.gz/das.sdss.org/DR4/data/imaging/752/40/corr/6/fpC-000752-r6-0245.fit.gz [0 x 0]}

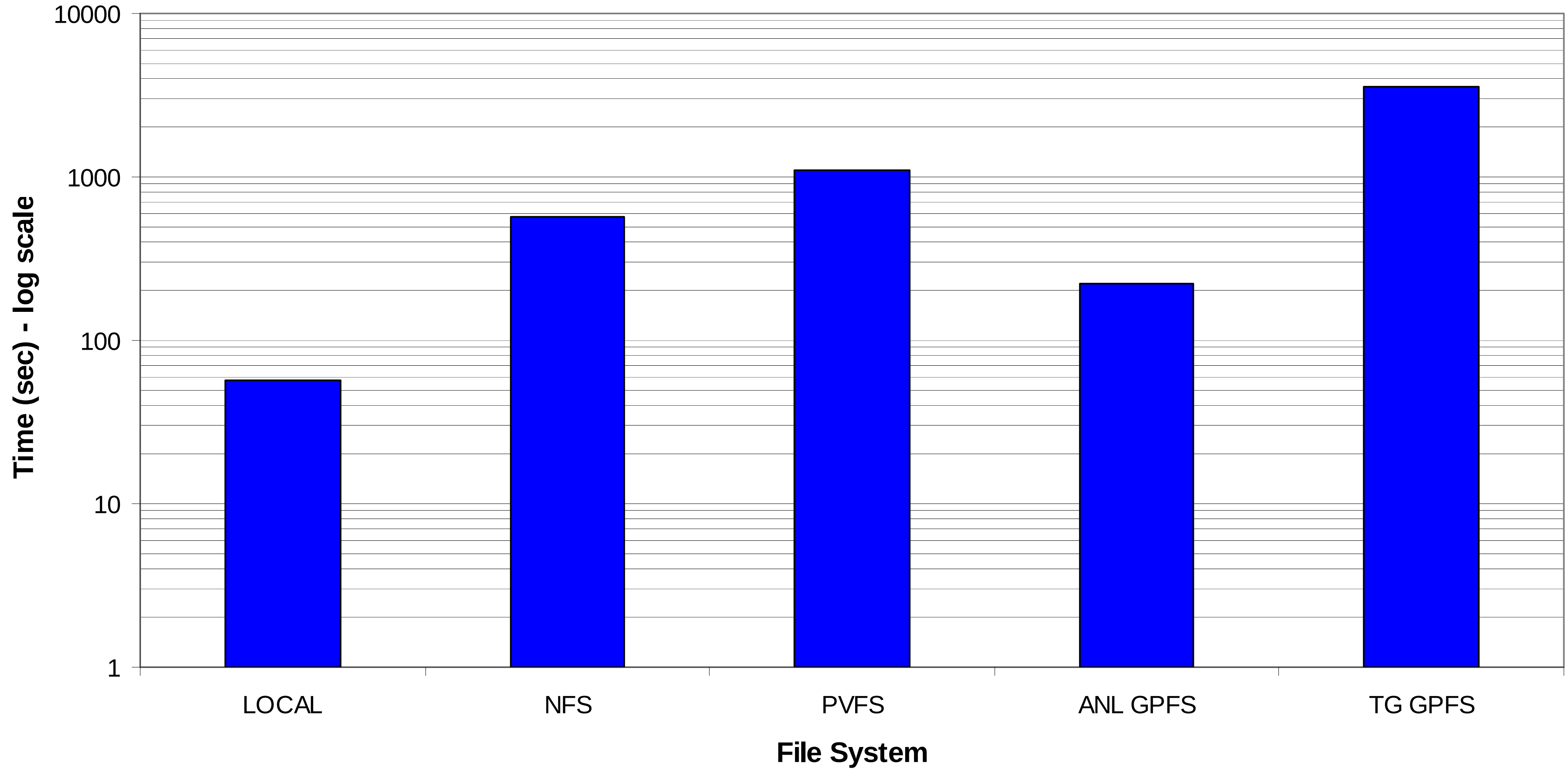
To start a new stacking, go back to the main **Stacking Service**.


Astro Users
(AU)
Data
Location
DB
Site
Query SQL DB
via HTTP
WS Calls
AstroPortal WS
(APWS)
Resource
Provisioning
via GRAM
Compute
Nodes
WS Calls
WS Calls
AstroWorker (AW) i
AstroWorker (AW) 1
Data Reads
GridFTP /
HTTP
Local GPFS
(LAN)
Data Reads
AstroWorker (AW) n
Data
Repository

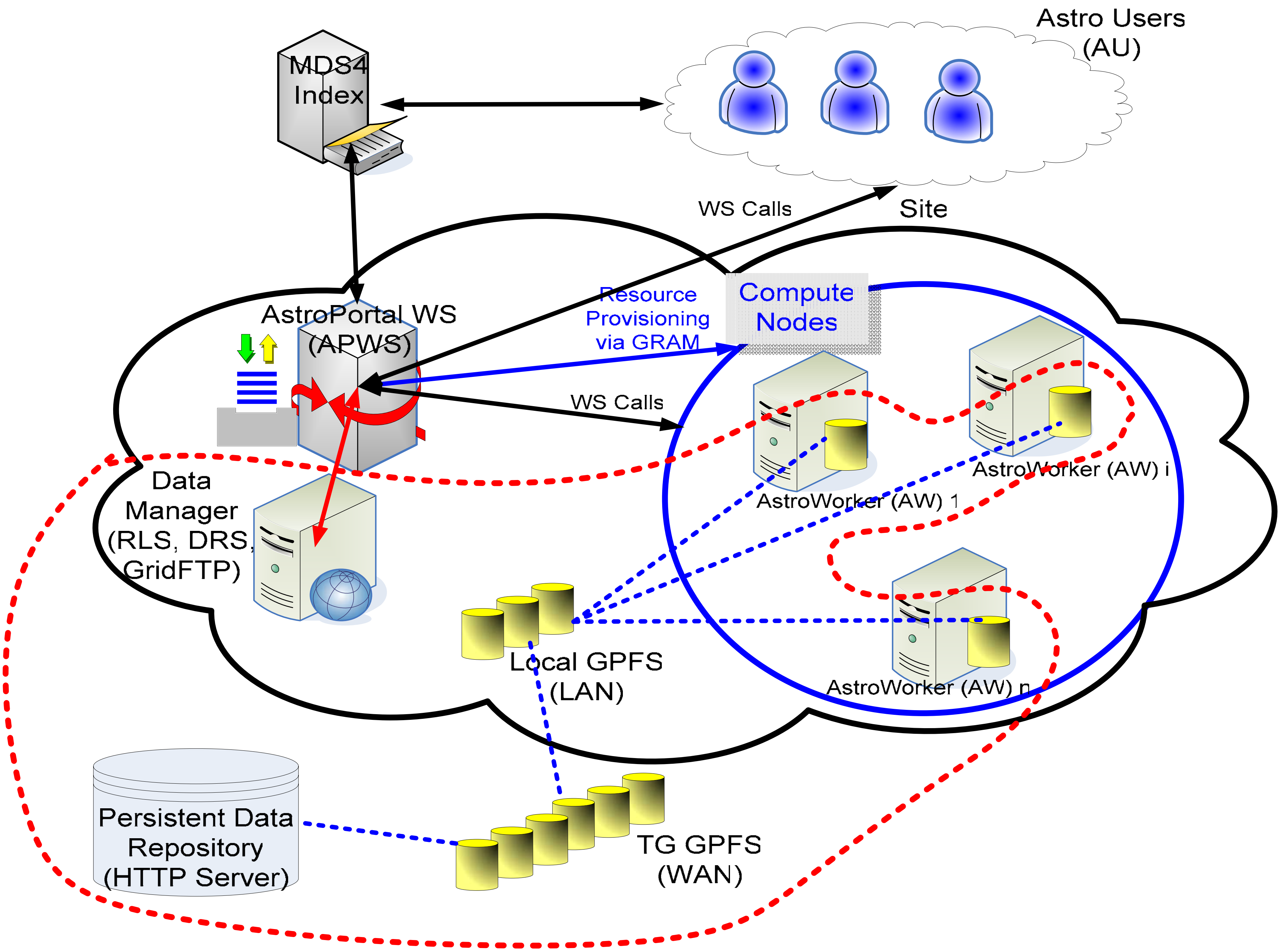


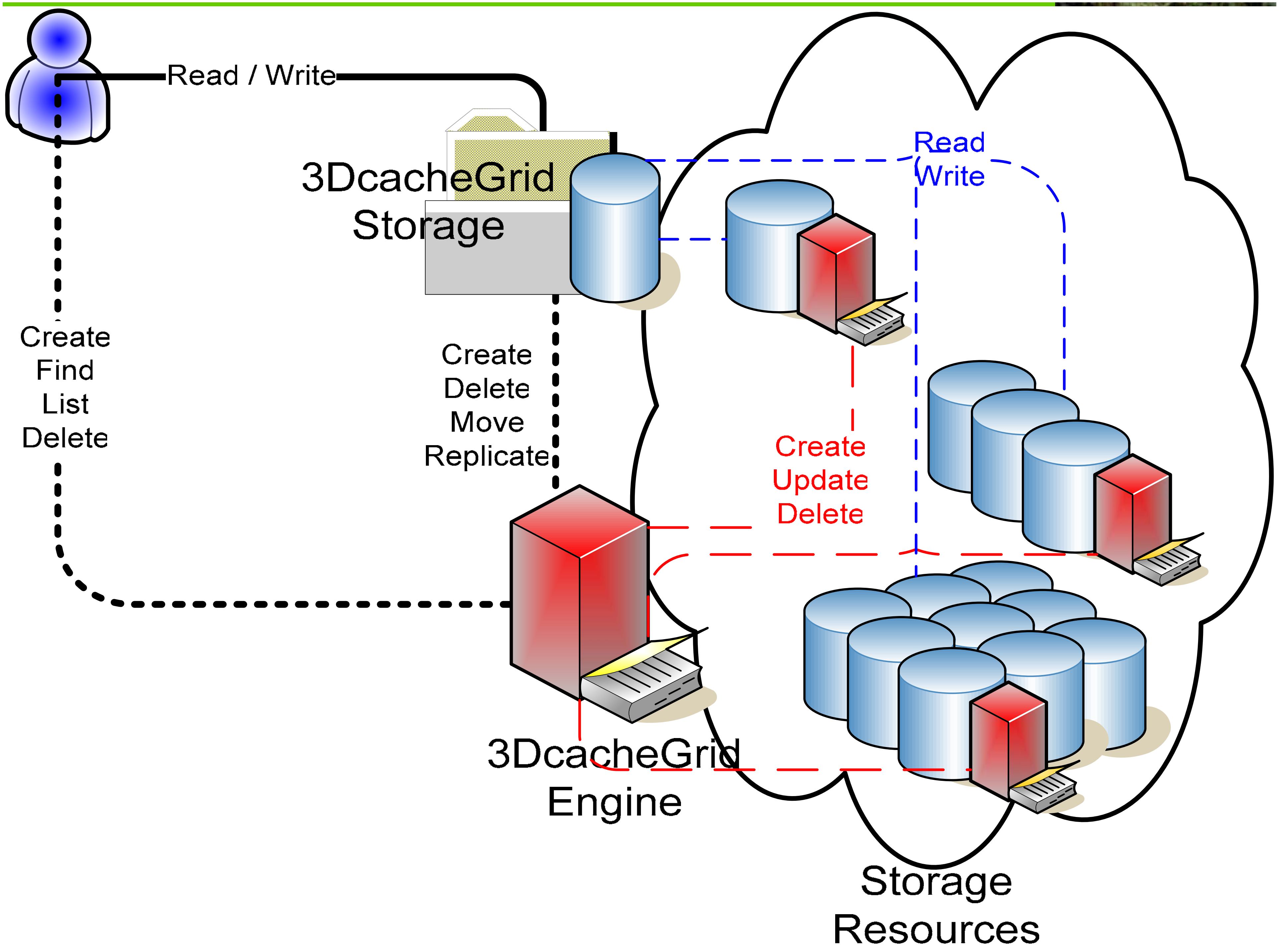
AstroPortal Web Service

O(100K)



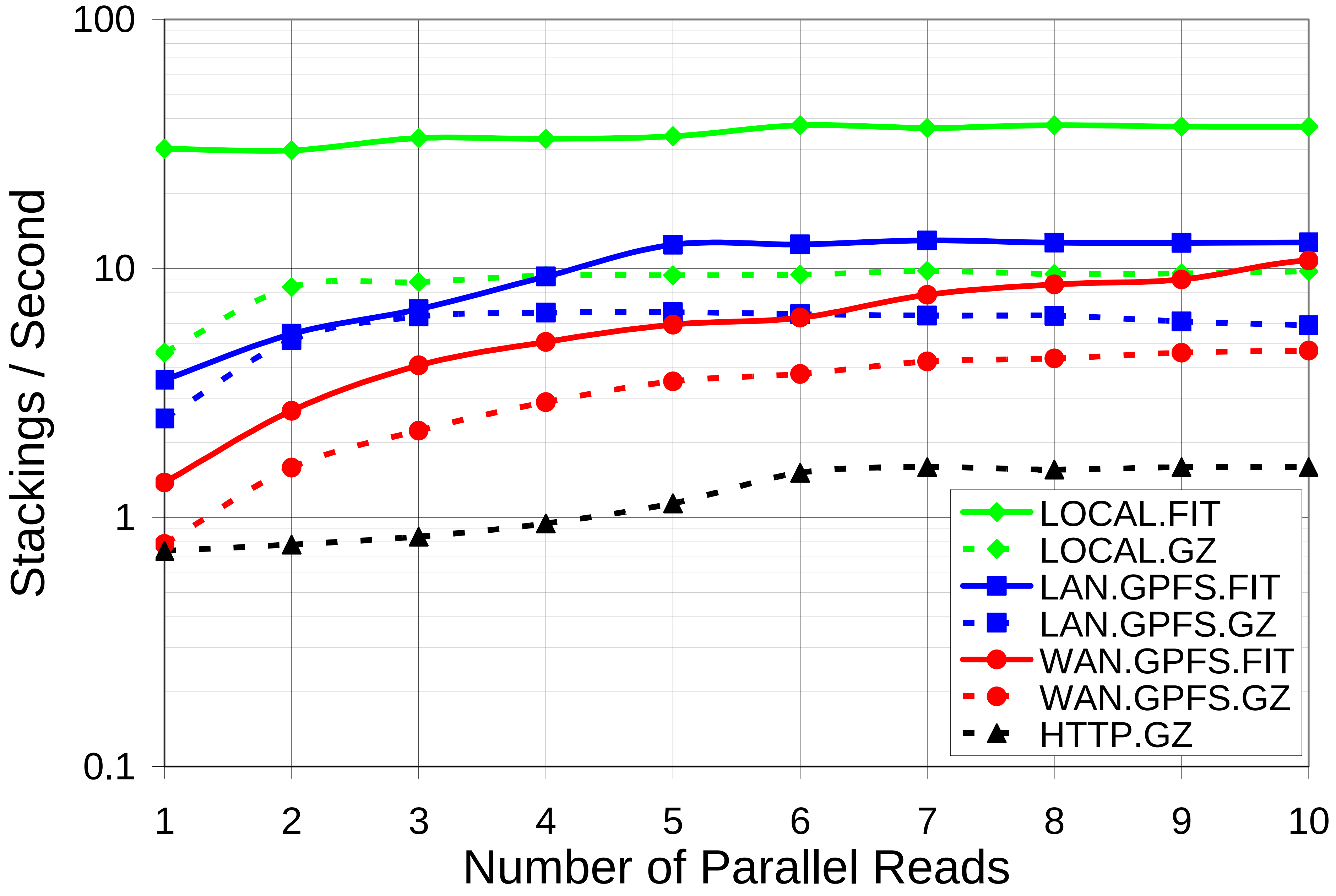
10000
1000
100
10
1
Time (sec) - log scale
LOCAL
NFS
PVFS
ANL GPFS
TG GPFS
File System


Astro Users
(AU)
MDS4
Index
WS Calls
Site
Resource
Provisioning
via GRAM
Compute
Nodes
AstroPortal WS
(APWS)
WS Calls
AstroWorker (AW) i
AstroWorker (AW) 1
Data
Manager
(RLS, DRS,
GridFTP)
Local GPFS
(LAN)
AstroWorker (AW) n
Persistent Data
Repository
(HTTP Server)
TG GPFS
(WAN)

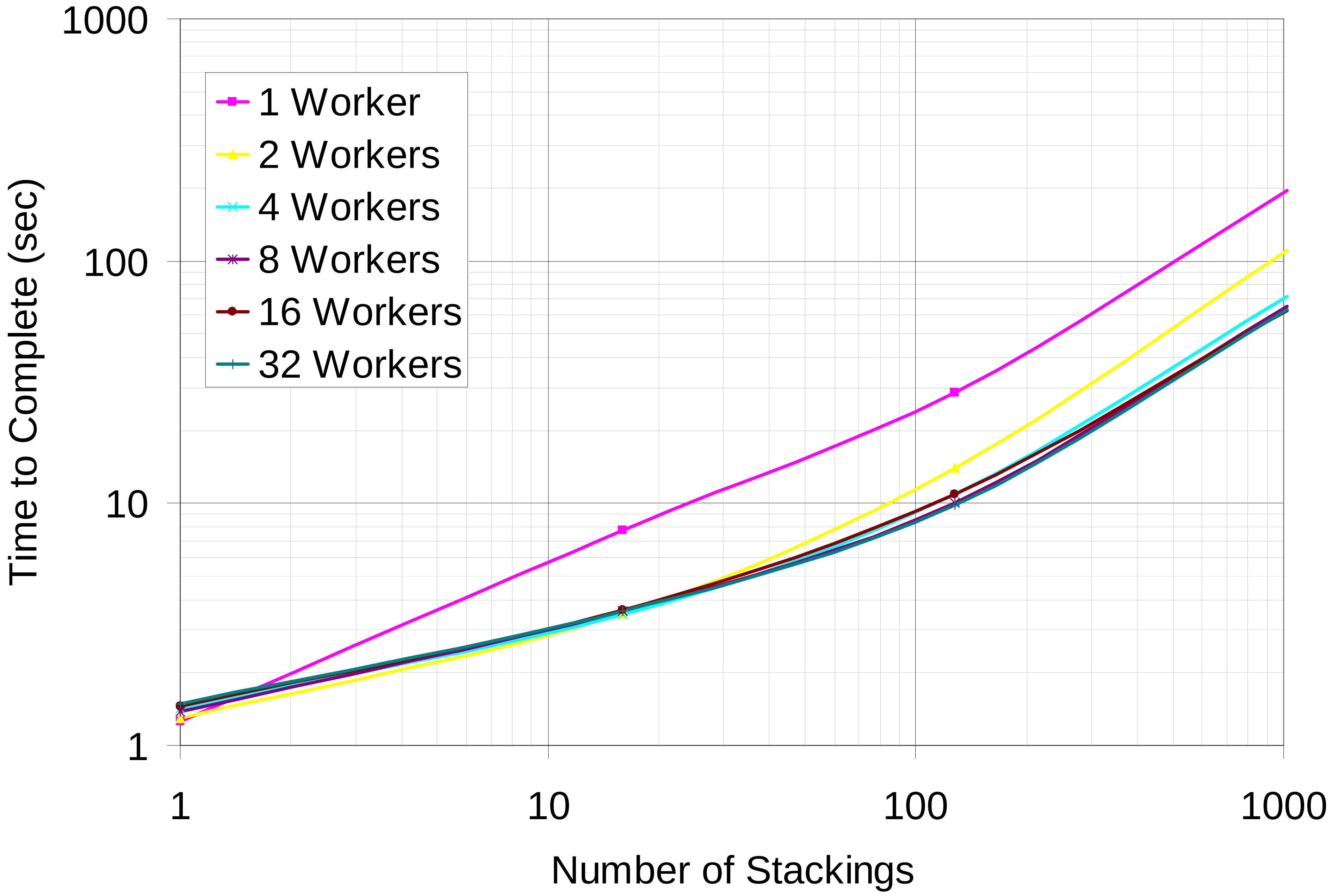


Read / Write
3DcacheGrid
Storage
Read
Write
Create
Find
List
Delete
Create
Delete
Move
Replicate
Create
Update
Delete
3DcacheGrid
Engine
Storage
Resources

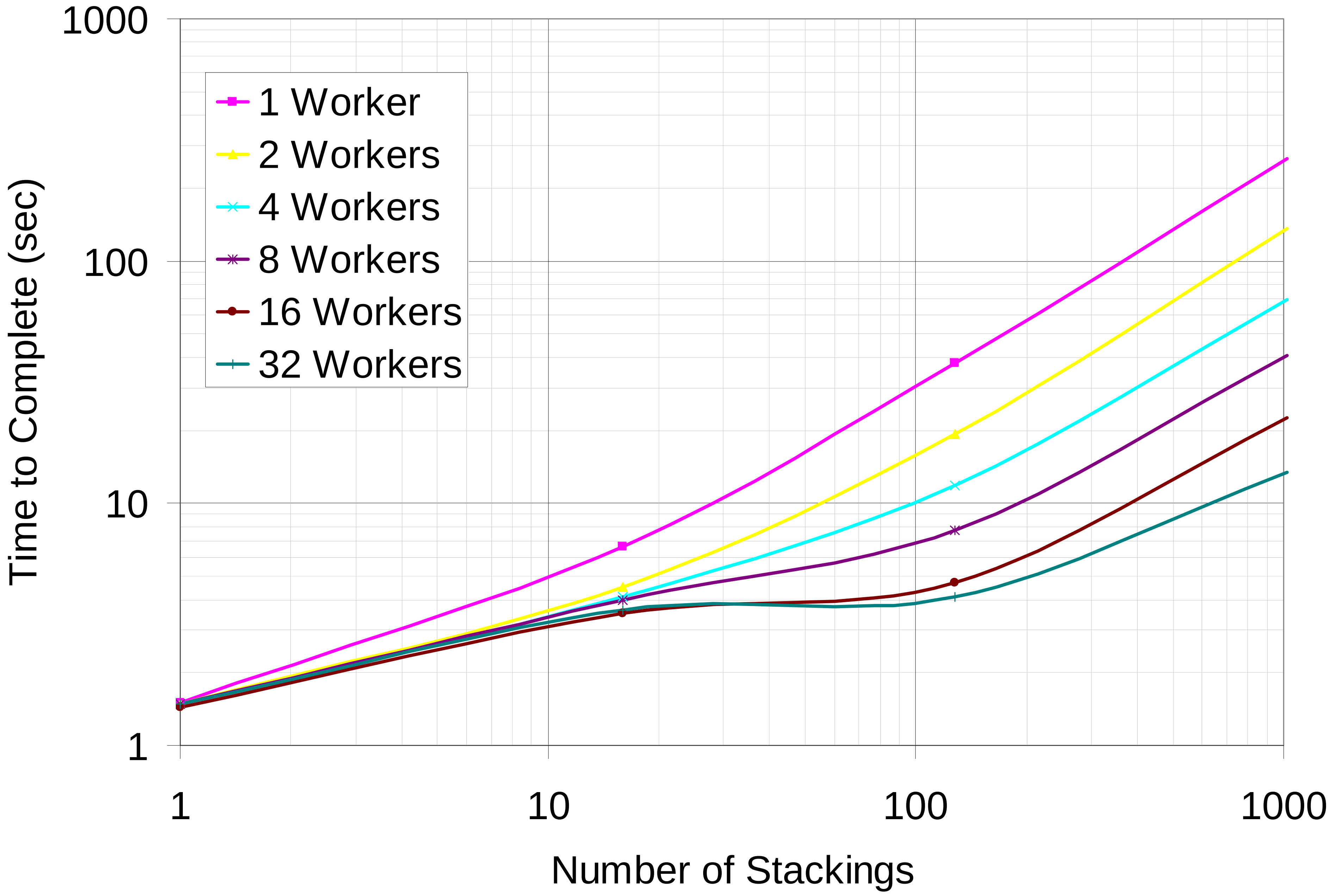
1 Worker – Multiple Threads



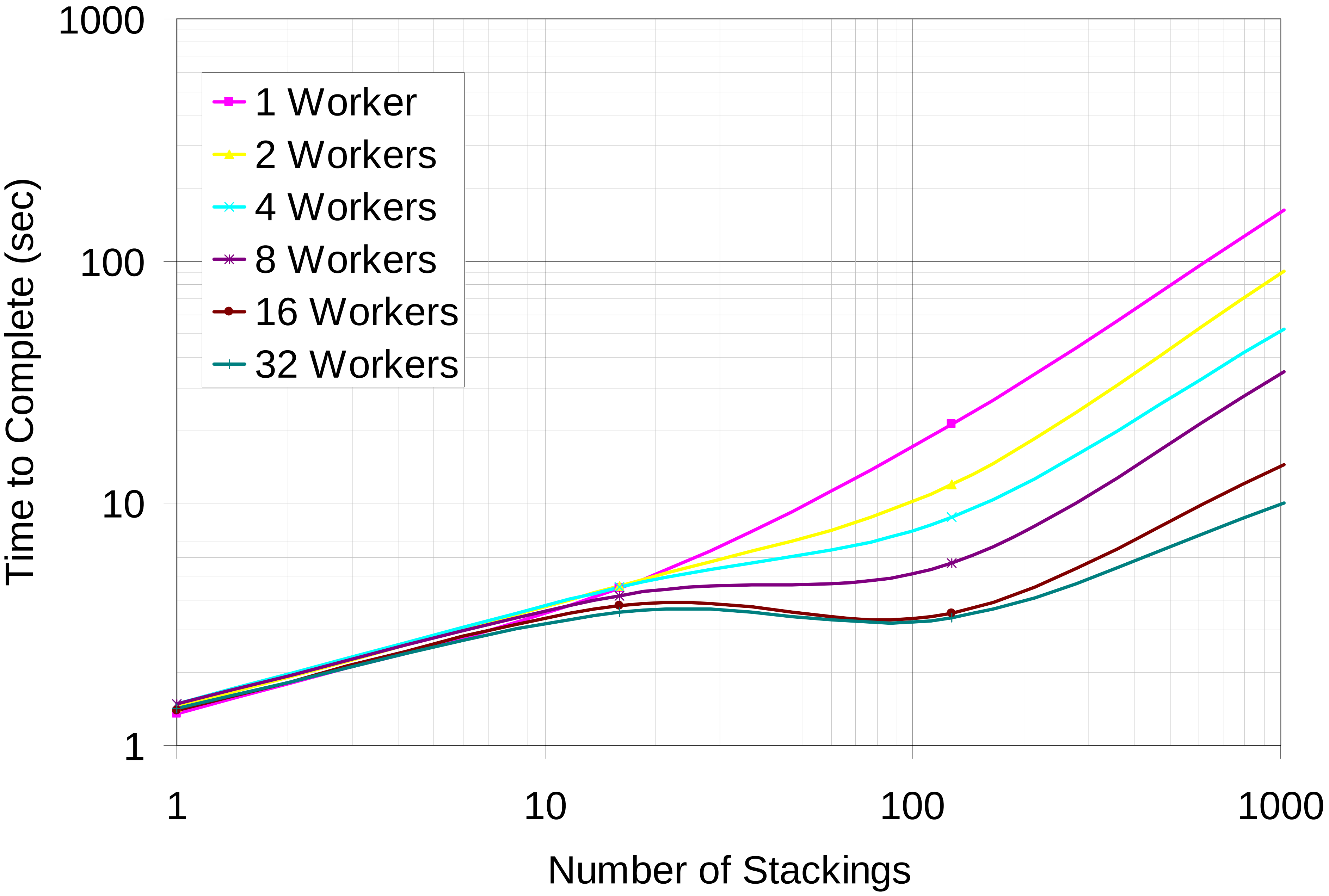
HTTP.GZ


1000
100
10
1
Time to Complete (sec)
1 Worker
2 Workers
4 Workers
8 Workers
16 Workers
32 Workers
1
10
100
1000
Number of Stackings

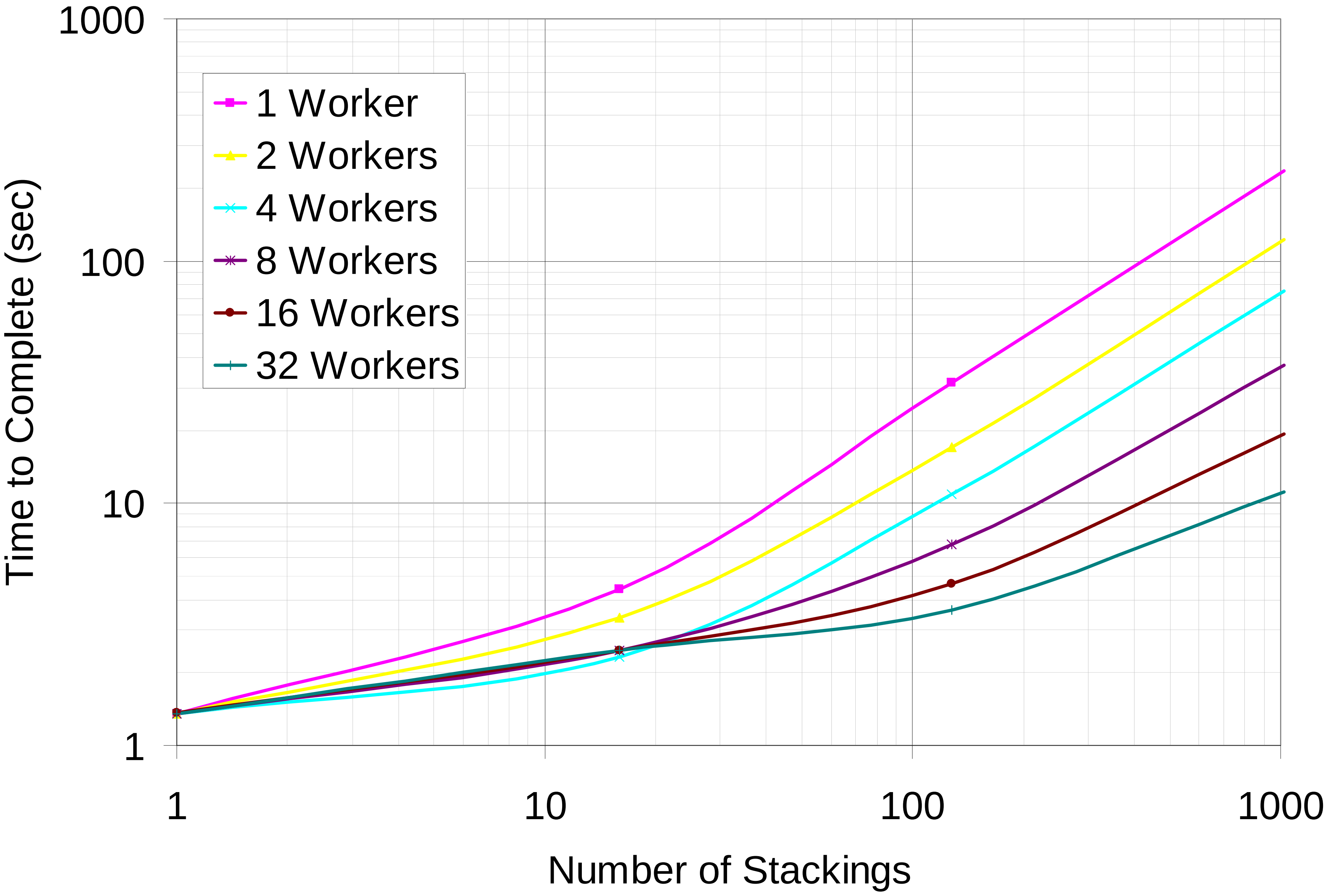
WAN.GZ


1000
100
10
1
Time to Complete (sec)
1 Worker
2 Workers
4 Workers
8 Workers
16 Workers
32 Workers
1
10
100
1000
Number of Stackings

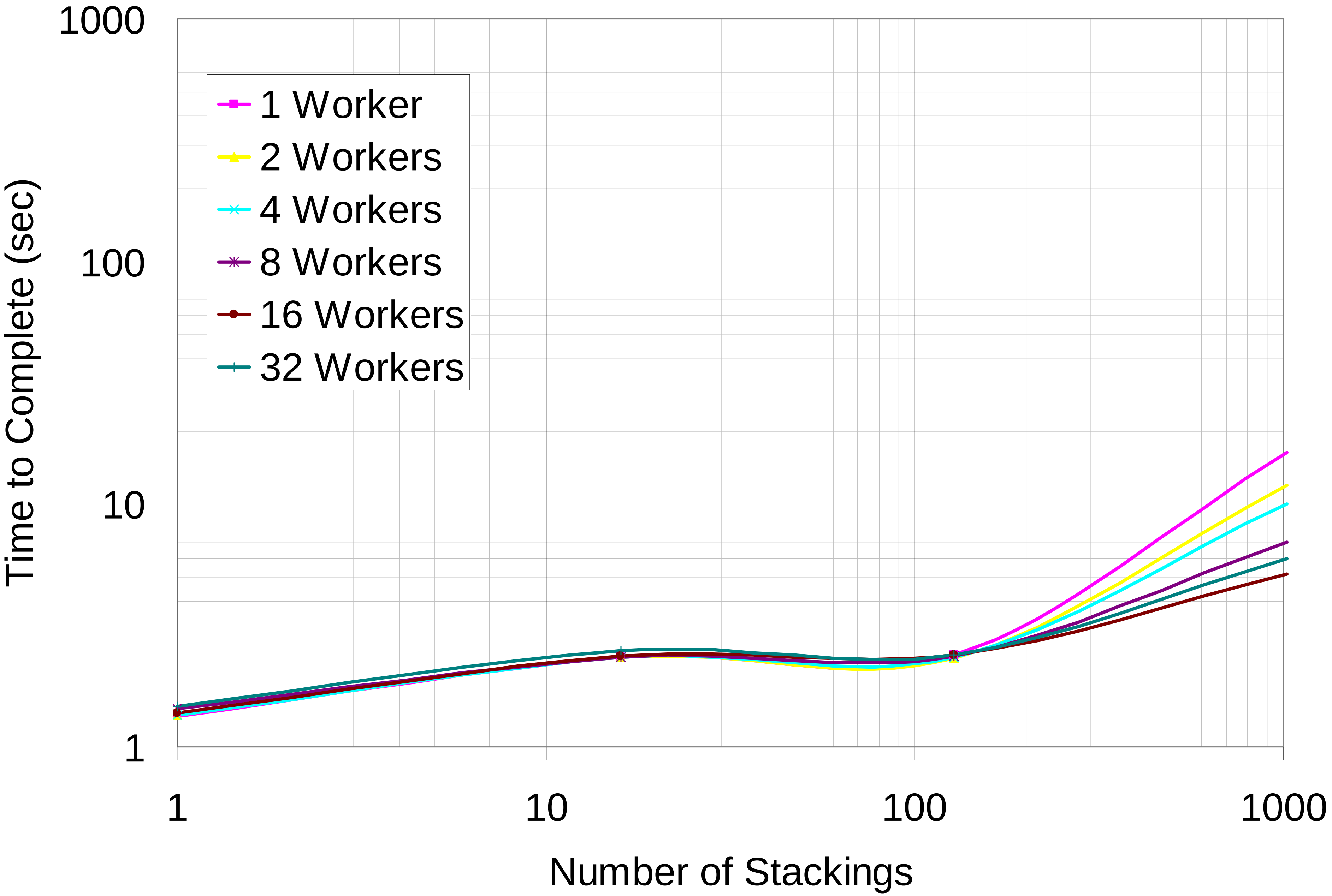
WAN.FIT


1000
100
10
1
Time to Complete (sec)
1 Worker
2 Workers
4 Workers
8 Workers
16 Workers
32 Workers
1
10
100
1000
Number of Stackings

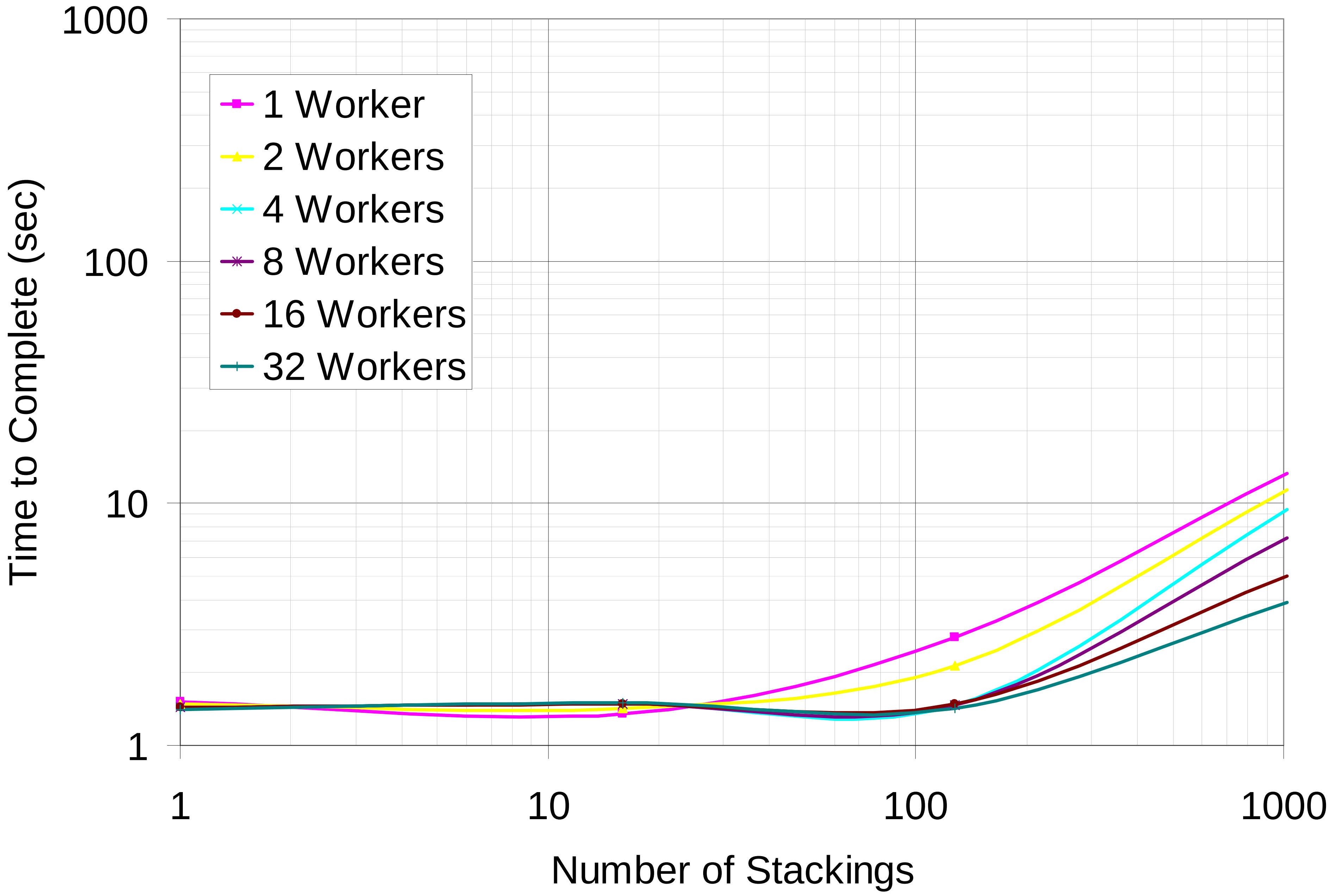
LAN.GZ


1000
100
10
1
Time to Complete (sec)
1 Worker
2 Workers
4 Workers
8 Workers
16 Workers
32 Workers
1
10
100
1000
Number of Stackings

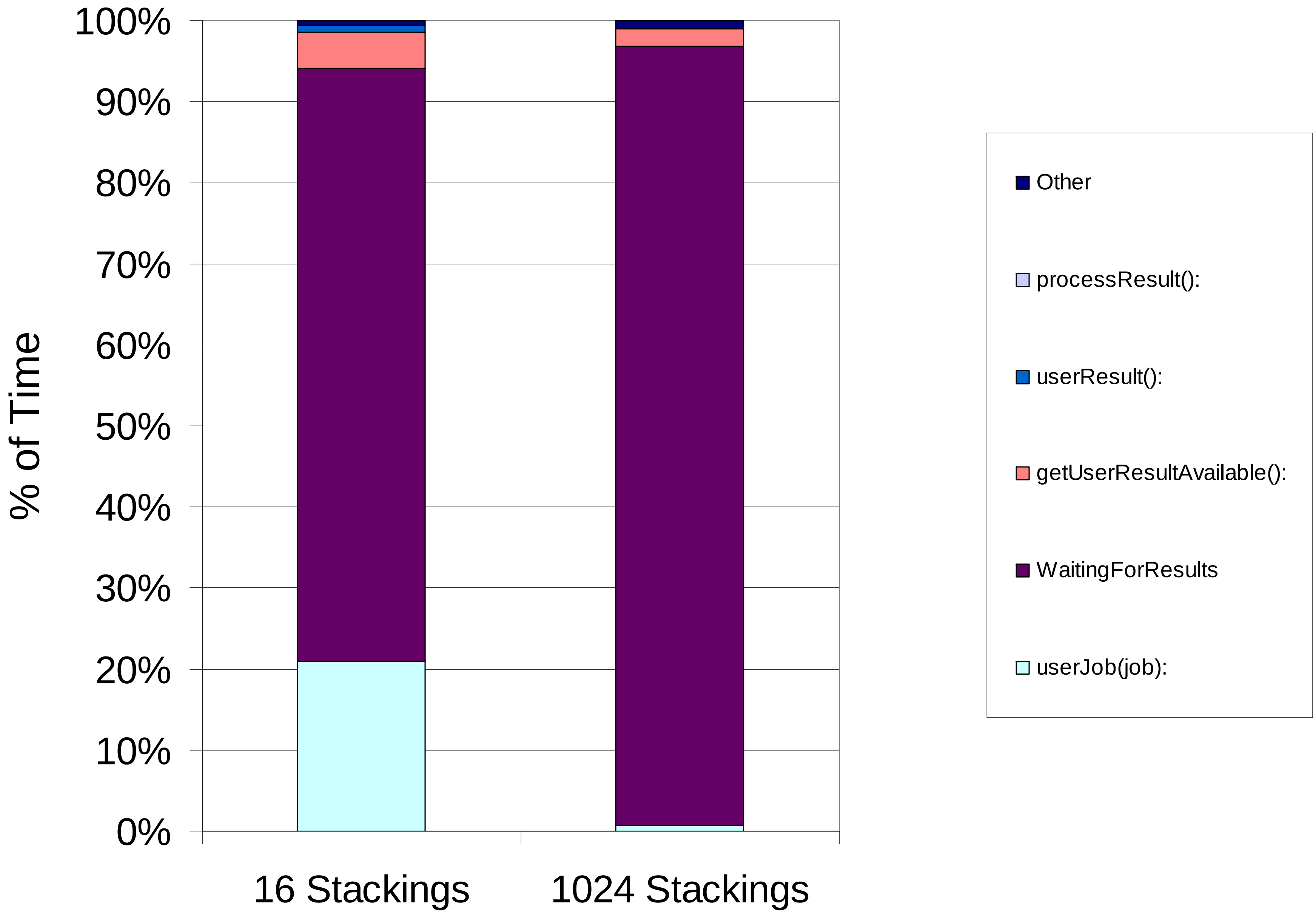
LAN.FIT


1000
100
10
1
Time to Complete (sec)
1
10
100
1000
Number of Stackings
1 Worker
2 Workers
4 Workers
8 Workers
16 Workers
32 Workers

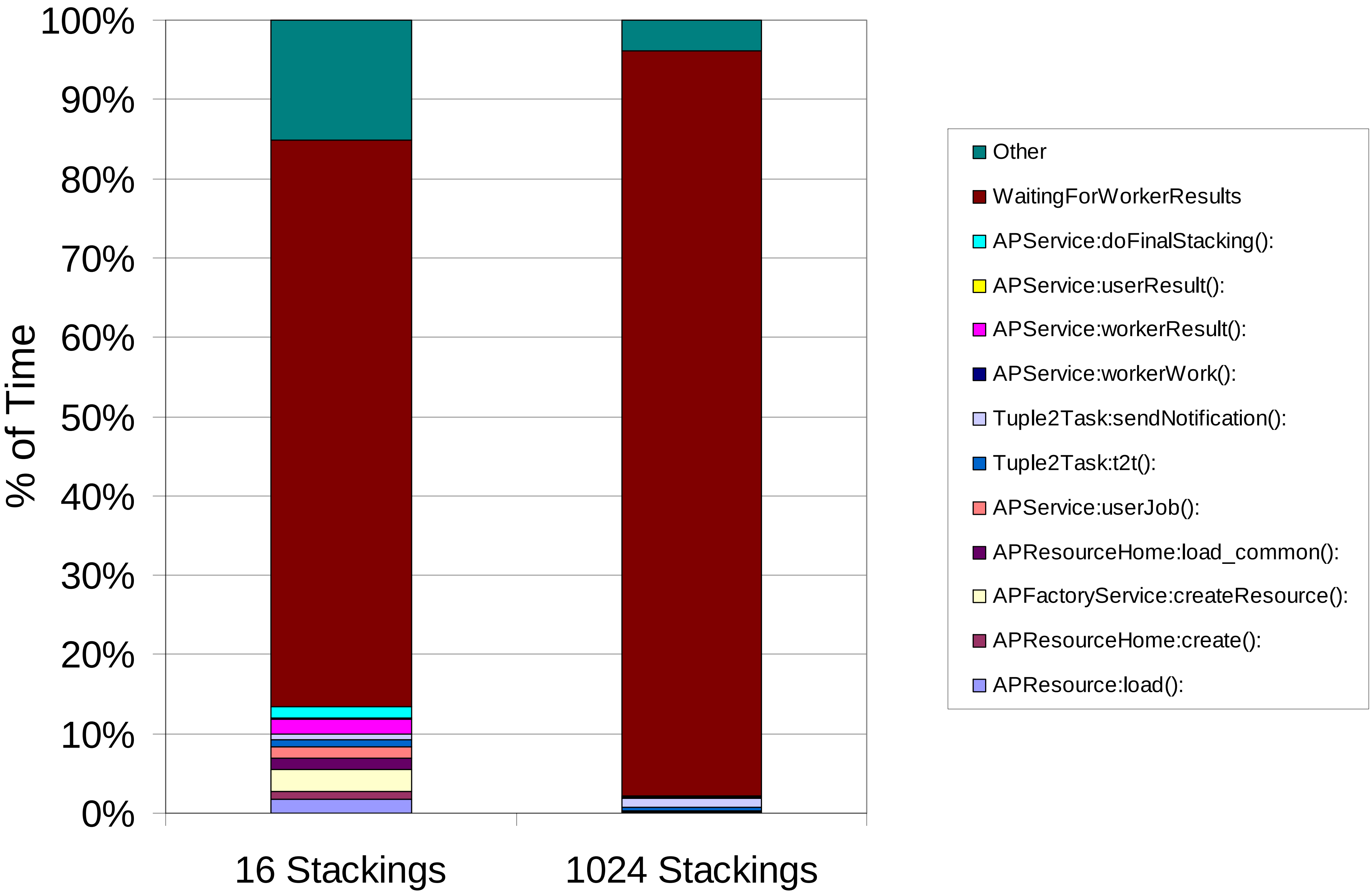
LOCAL.FIT


1000
100
10
1
Time to Complete (sec)
1
10
100
1000
Number of Stackings
1 Worker
2 Workers
4 Workers
8 Workers
16 Workers
32 Workers

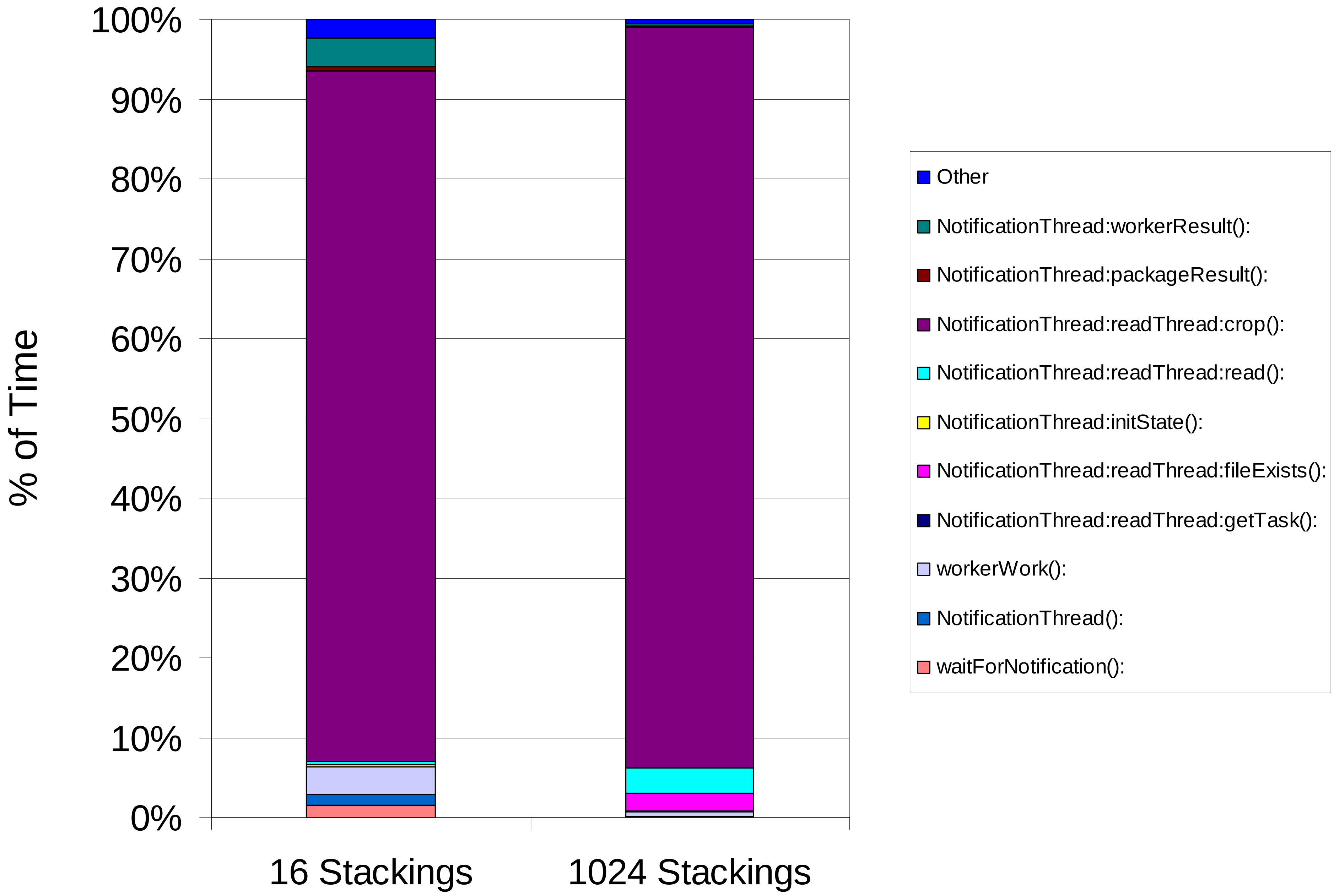
AstroUser


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% of Time
16 Stackings
1024 Stackings
■ Other
□ processResult():
■ userResult():
■ getUserResultAvailable():
■ WaitingForResults
□ userJob(job):

AstroPortal Web Service



AstroWorker



Open Research Questions

- Site level
  - advanced reservations
  - resource allocation
  - resource de-allocation
- Data management
  - Data location and replication
  - Data caching hierarchies
- Resource management
  - Distributed resource management between various sites

Open Research Questions: Site Level

- Leverage techniques used in large clusters
- Find heuristics will apply for managing efficiently the set of resources depending on the workload characteristics, number of users, data set size and distribution, etc…
- How to perform efficient state transfer among worker resources while maintaining a dynamic system

Open Research Questions: Data Management

- Very large data set distributed among various sites
- Replication strategies to meet the desired QoS
- Data placement based on past workloads and access patterns

Open Research Questions Resource Management

- Harness entire TeraGrid pool of resources (8 sites) rather than just 1 site
- Workload management, moving the work vs. moving the data

Other Domains: Medical Field

- Medium to large medical datasets are hard to acquire
  - Typical medium size data set (of CT images)
    - 1000 patient case studies
      - 100K images (1000 cases x 100 images)
        - » 1M+ objects (i.e. organs, tissues, abnormalities, etc…)
        - » 0.4TB+ raw images (4MB x 100K)
    - 10K+ potential users from 1K+ of different institutions (research labs, hospitals, etc…)
- Applications:
  - Making datasets available to trusted parties
  - Allowing image processing algorithms to be dynamically applied
  - Normal tissue classification in CT images
  - Lung cancer image databases

Questions?

- More information: http://people.cs.uchicago.edu/~iraicu/research/
- Related materials and further readings:
  - Ioan Raicu, Ian Foster, Alex Szalay, Gabriela Turcu. "***AstroPortal: A Science Gateway for Large-scale Astronomy Data Analysis***", to appear at TeraGrid Conference 2006, June 2006.
  - Alex Szalay, Julian Bunn, Jim Gray, Ian Foster, Ioan Raicu. "***The Importance of Data Locality in Distributed Computing Applications***", NSF Workflow Workshop 2006.
  - Ioan Raicu, Ian Foster, Alex Szalay. "***Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets***", under review at SuperComputing 2006.
  - Ioan Raicu, Ian Foster, Alex Szalay, Gabriela Turcu, Catalin Dumitrescu. "***Enabling Large-scale Astronomy Data Analysis with the AstroPortal***," under preparation for the HPC Analytics Challenge at SC06.
  - Ioan Raicu, Ian Foster, Elizeu Santos-Neto, John Bresnahan. "***3DcacheGrid: Dynamic Distributed Data Cache Grid Engine***," under preparation for the HPC Storage Challenge at SC06.