



THE UNIVERSITY OF
CHICAGO



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

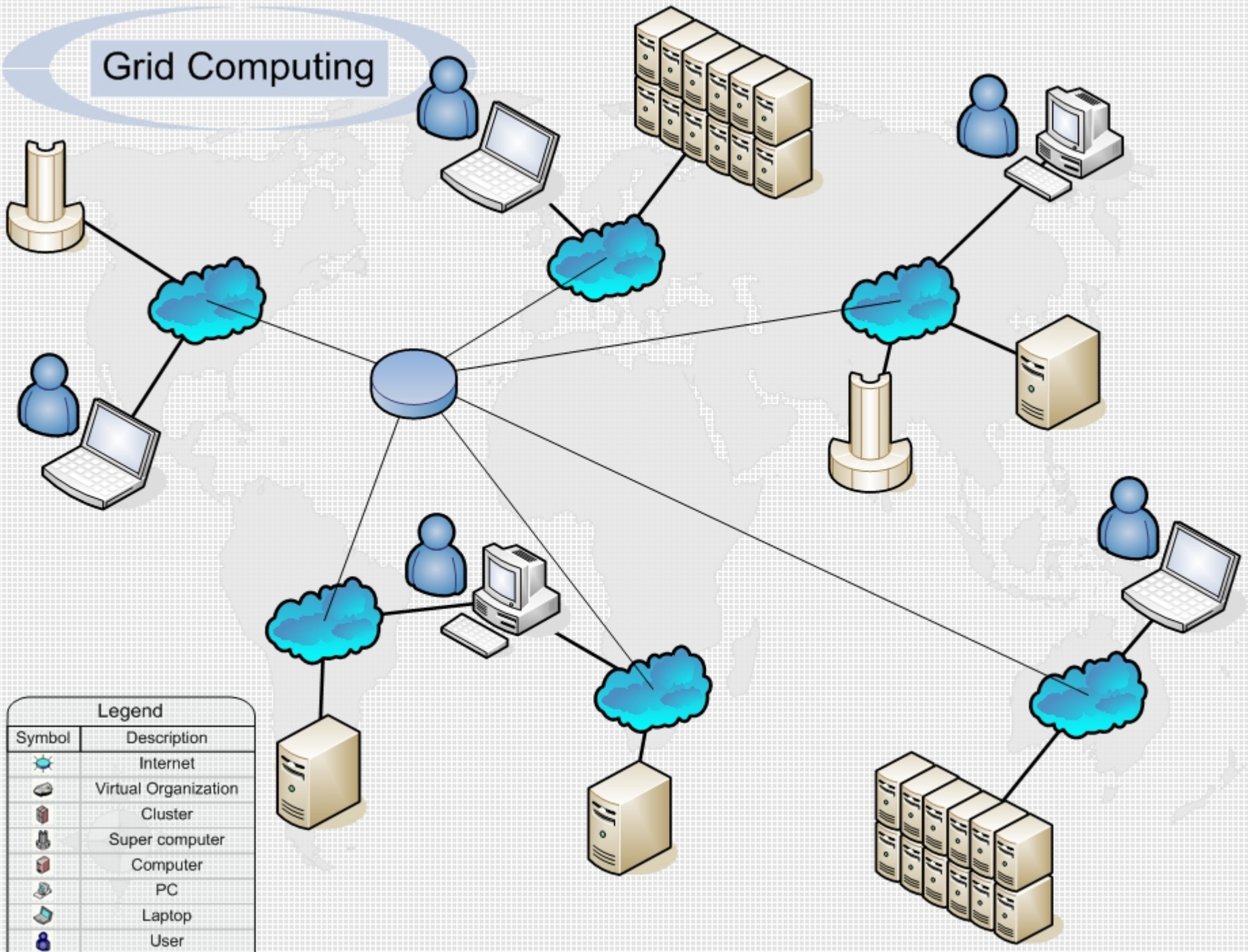
Ioan Raicu

Distributed Systems Laboratory
Computer Science Department
University of Chicago



February 22nd, 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 parallelization
- Applicable fields:
 - Astronomy
 - Medicine
 - Others

Astronomy Field



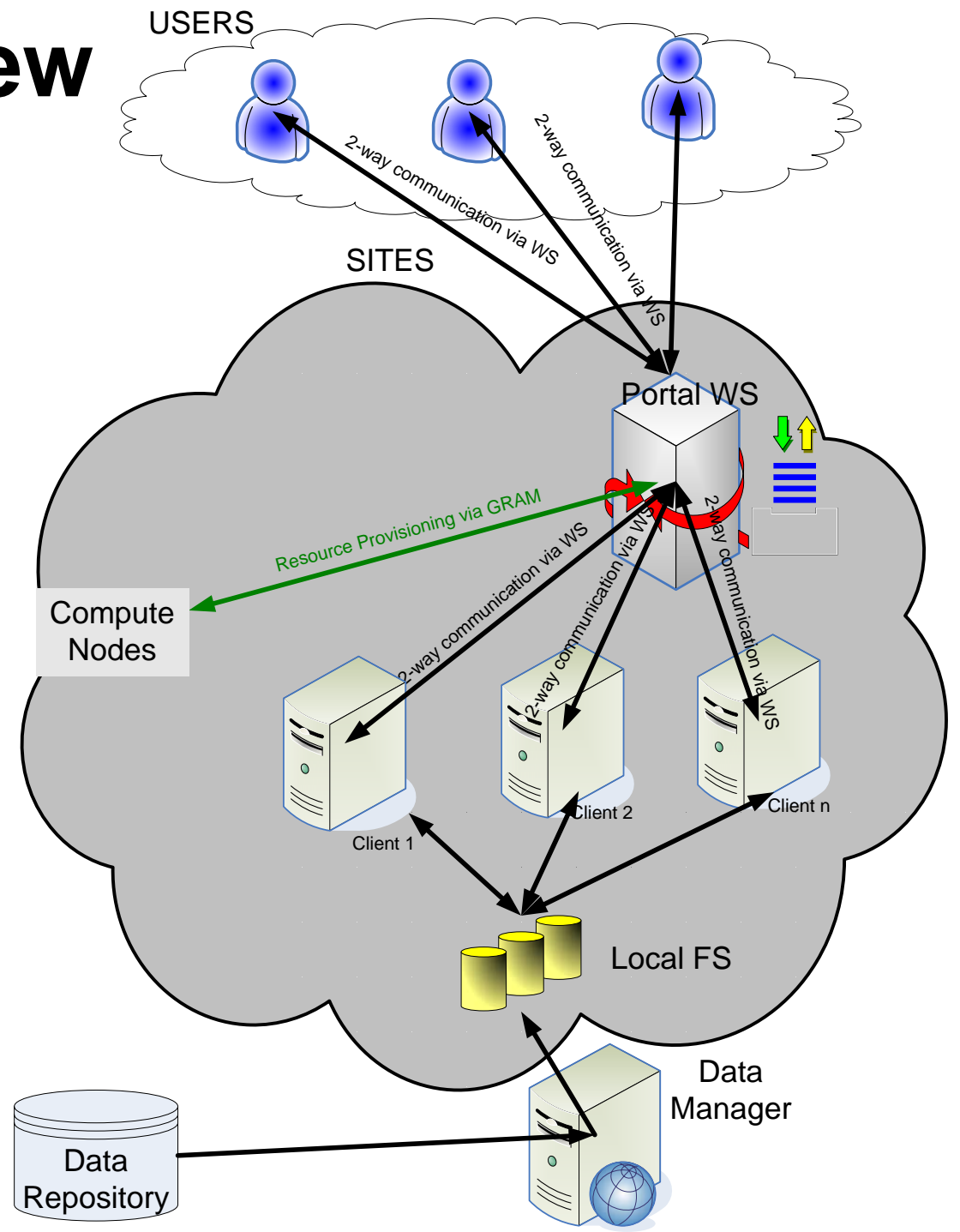
- Astronomy datasets (i.e. SDSS) are the crown-jewels
 - SDSS DR4
 - 500K images
 - 300M+ objects
 - 1TB+ compressed images (2MB x 500K)
 - 3TB+ raw images (6.1MB x 500K)
 - 100K worldwide potential users
- Applications:
 - Stacking
 - Montage

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

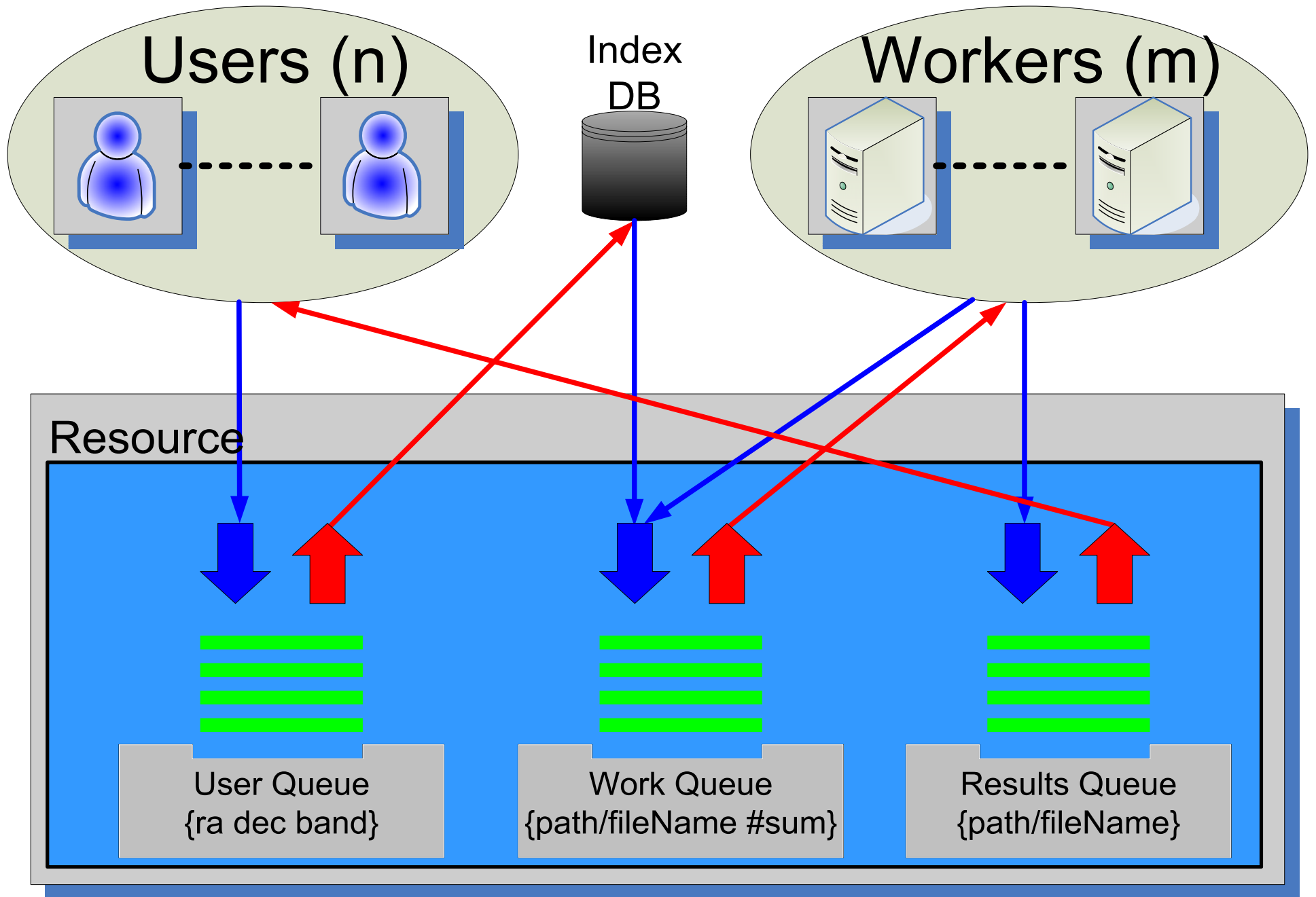
Generic Overview



Functionality Overview



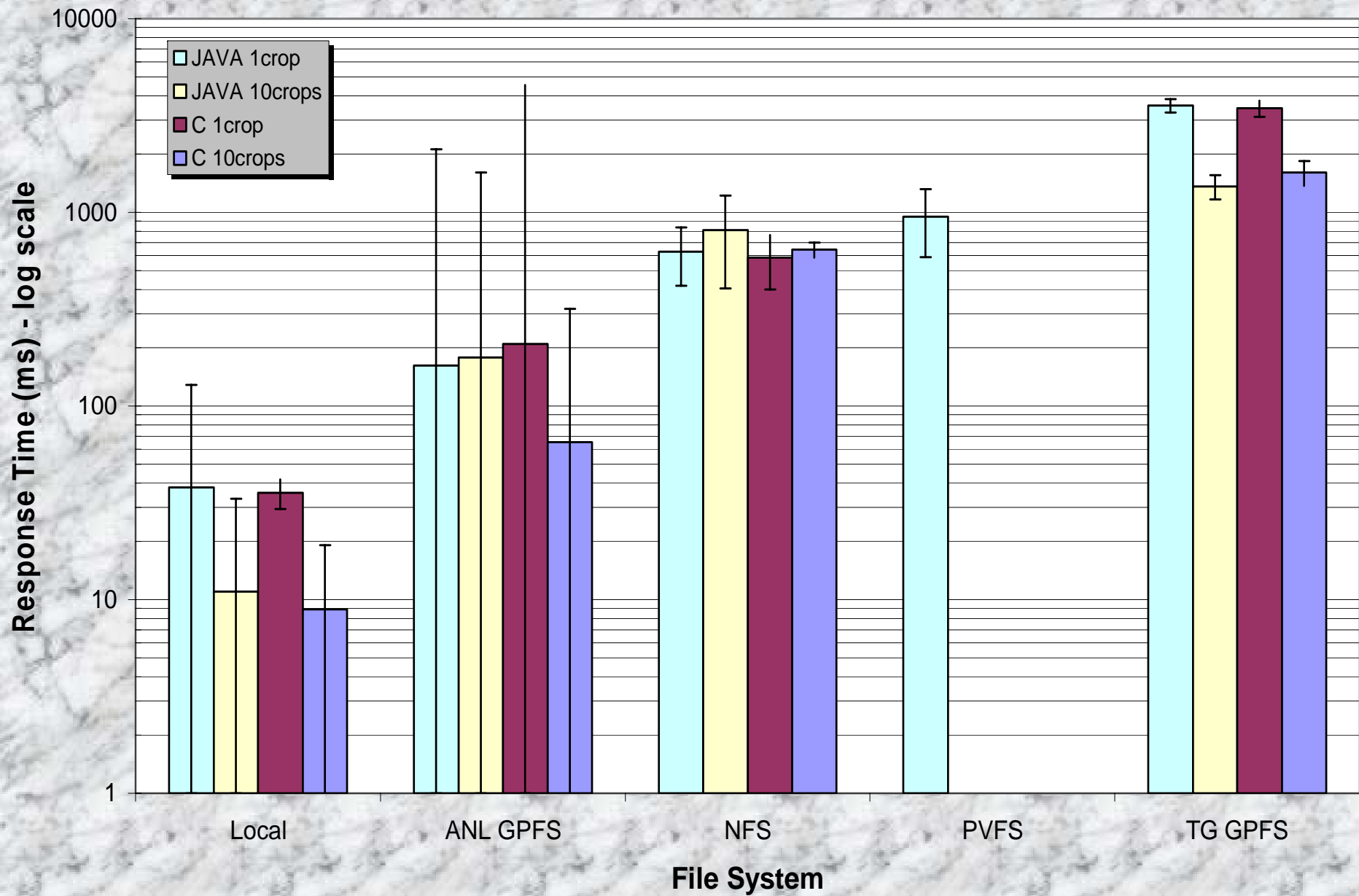
- Input
 - A set of {band ra dec} tuples plus operation to be performed (GetAll, SumAll, etc...)
- Work
 - GetAll: crop ROIs
 - SumAll: crop ROIs and stack them
- Output
 - GetAll: A set of images corresponding to the above tuples
 - SumAll: 1 image corresponding to the summation of the above tuples



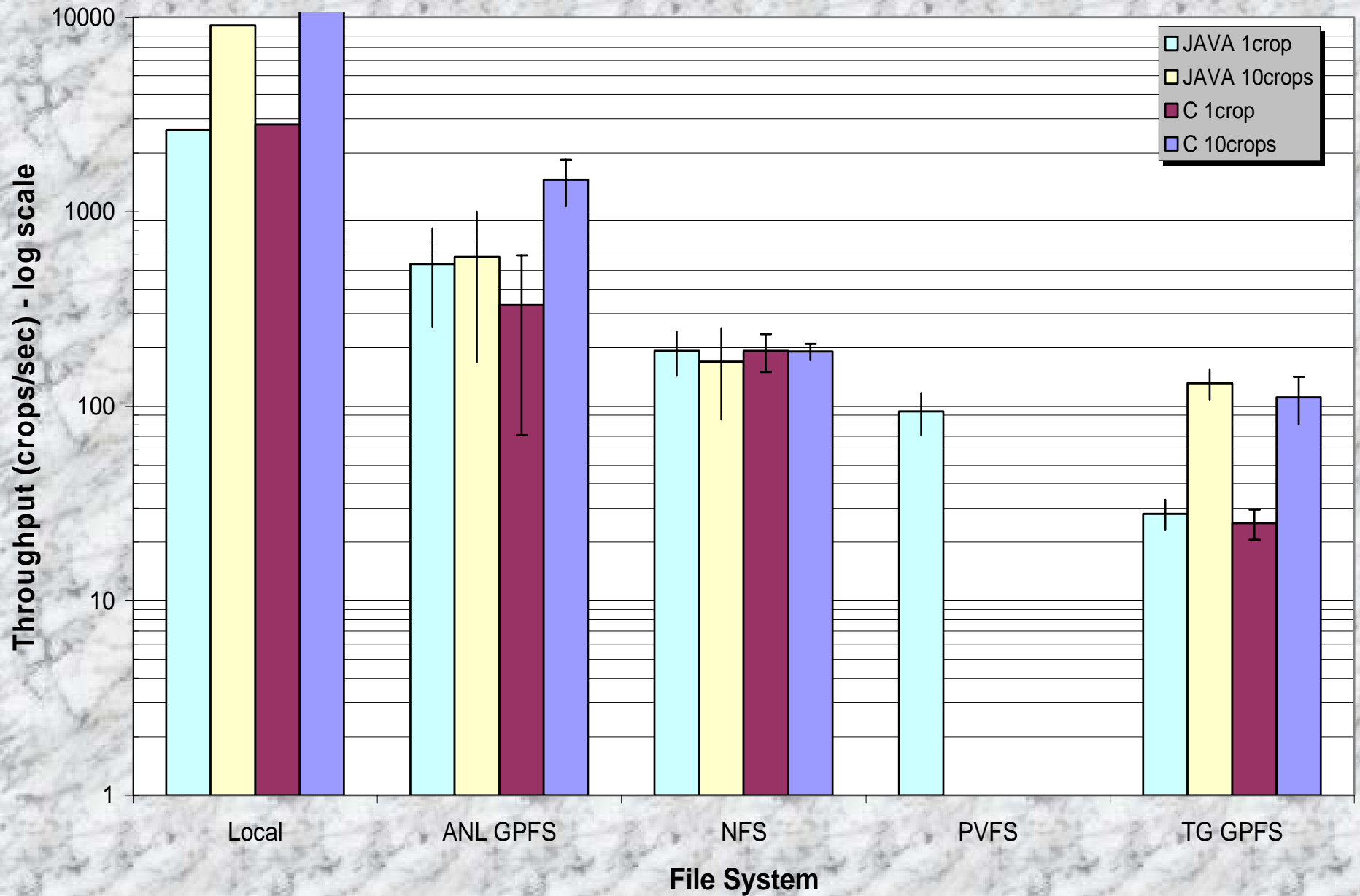
AstroPortal WS

Summary FIT Client Performance

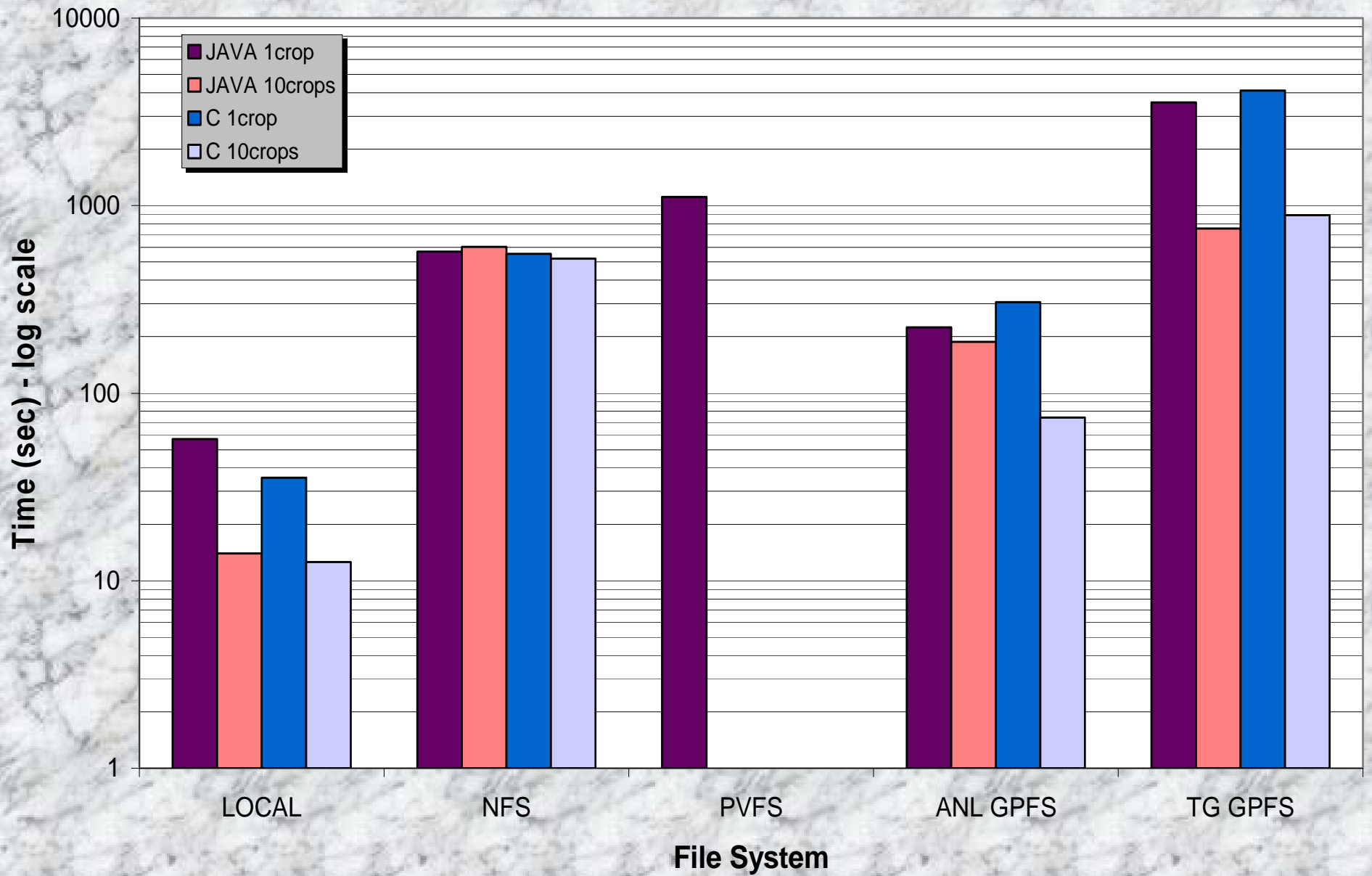
Response Time



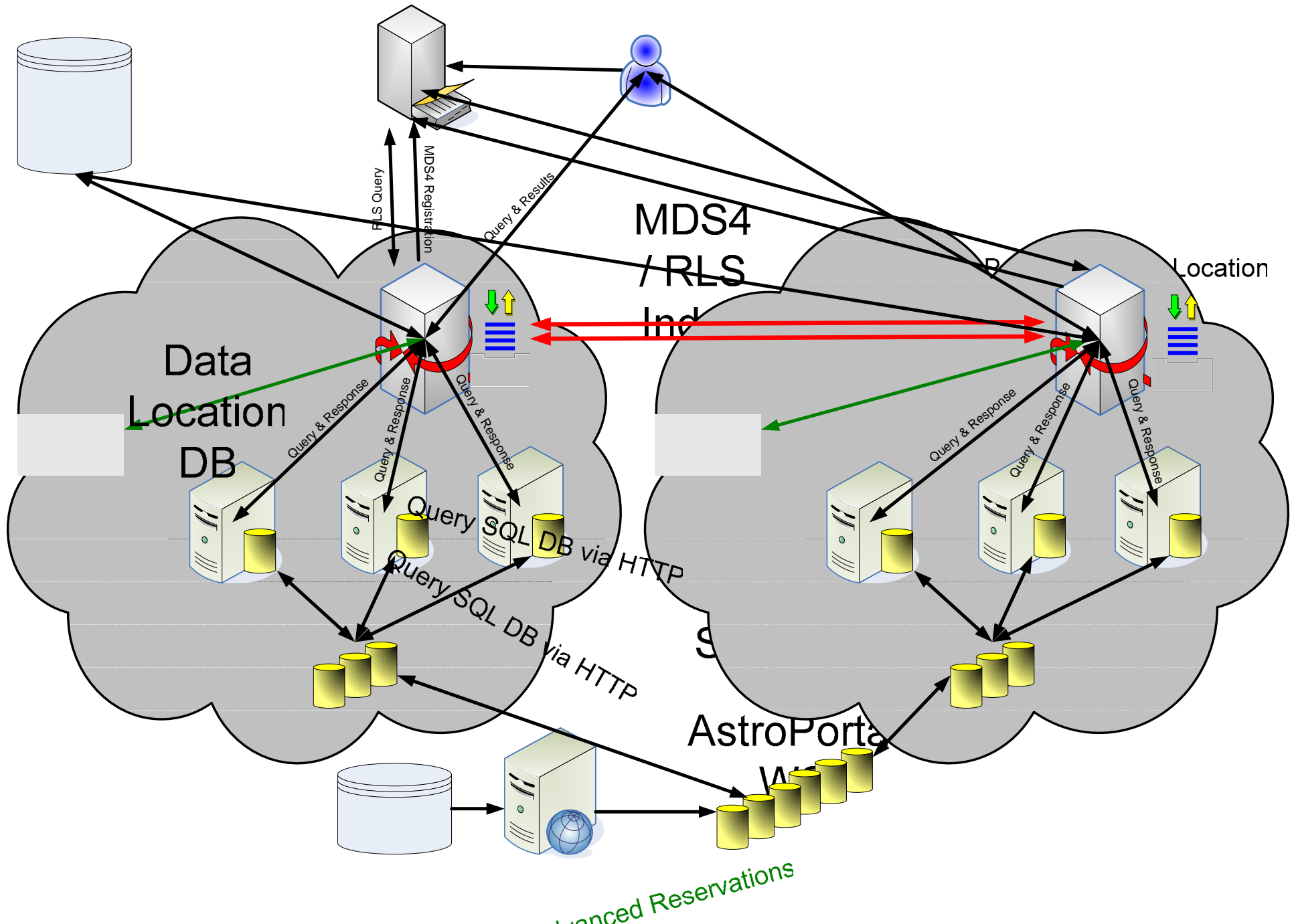
Summary FIT Client Performance Throughput



Time to complete O(100K) Crops



Target Implementation



Open Research Questions



- Cluster 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: Cluster 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



- The inter-site communication among the AP WS and its effects on the overall system performance is very interesting
- Workload management, moving the work vs. moving the data
- Algorithms, the amount of state information, and the frequency of state information exchanges will affect the performance of the overall system

Questions?



THE UNIVERSITY OF
CHICAGO



ARGONNE
NATIONAL LABORATORY

Terminology



- **Site:** A TeraGrid site, such as UC/ANL, SDSC, NCSA, PSC, ORNL, TACC, etc...
- **User:** user from the astronomy domain who wants to query the data set with a 5-tuple (path & file name, x-coordinate, y-coordinate, height, and width)
- **AstroPortal Web Service (AP WS):** A WS that gives users an entry point into accessing TG resources to process the user's queries
- **MDS4 Index:** A standard MDS4 Index used for resource (AP WS) discovery by the users
- **Compute Nodes - AstroClient (AC):** dedicated nodes in TG that are reserved in advance to be used for processing queries from the AP WS
- **Data Repository:** the original data set in compressed format that can be accessed via GridFTP
- **AstroData (AD) Manager:** A data resource manager that keeps the data set up to date between the data repository, and the corresponding file systems (Local GPFS, TG GPFS, etc...); in the distributed version, the AD Manager could also use RLS to manage data replication; the AD Manager also communicates with the AP WS in order to keep the AP WS data set index updated with the latest data set location
- **Local GPFS:** Refers to site local GPFS accessed over a LAN
- **TG GPFS:** TeraGrid wide GPFS accessed over a WAN
- **RFT:** Used to update the working data set on GPFS from the data repository
- **GRAM:** Used to make advanced reservations of AC compute nodes by being scheduler independent
- **RLS:** used to keep track of the data replicas in the distributed AP architecture