Performance - Directed Resource Allocation

Seung-Hye Jang, Xingfu Wu, Valerie E. Taylor

Department of Computer Science
Texas A&M University

SIAM Conference on Computational Science & Engineering,
Orlando, Florida
February 12 – 15, 2005

Outline

- Motivation & Goals
- Performance Prediction: Prophesy
- Case Studies
  - Grid Physics Network (GriPhyN)
    - Grid2003 infrastructure
    - GEO LIGO pulsar search
  - Educational Application
    - Multiple servers
    - ADDAMLSS

Motivation

Galaxy Formation Simulation

LHC Simulations

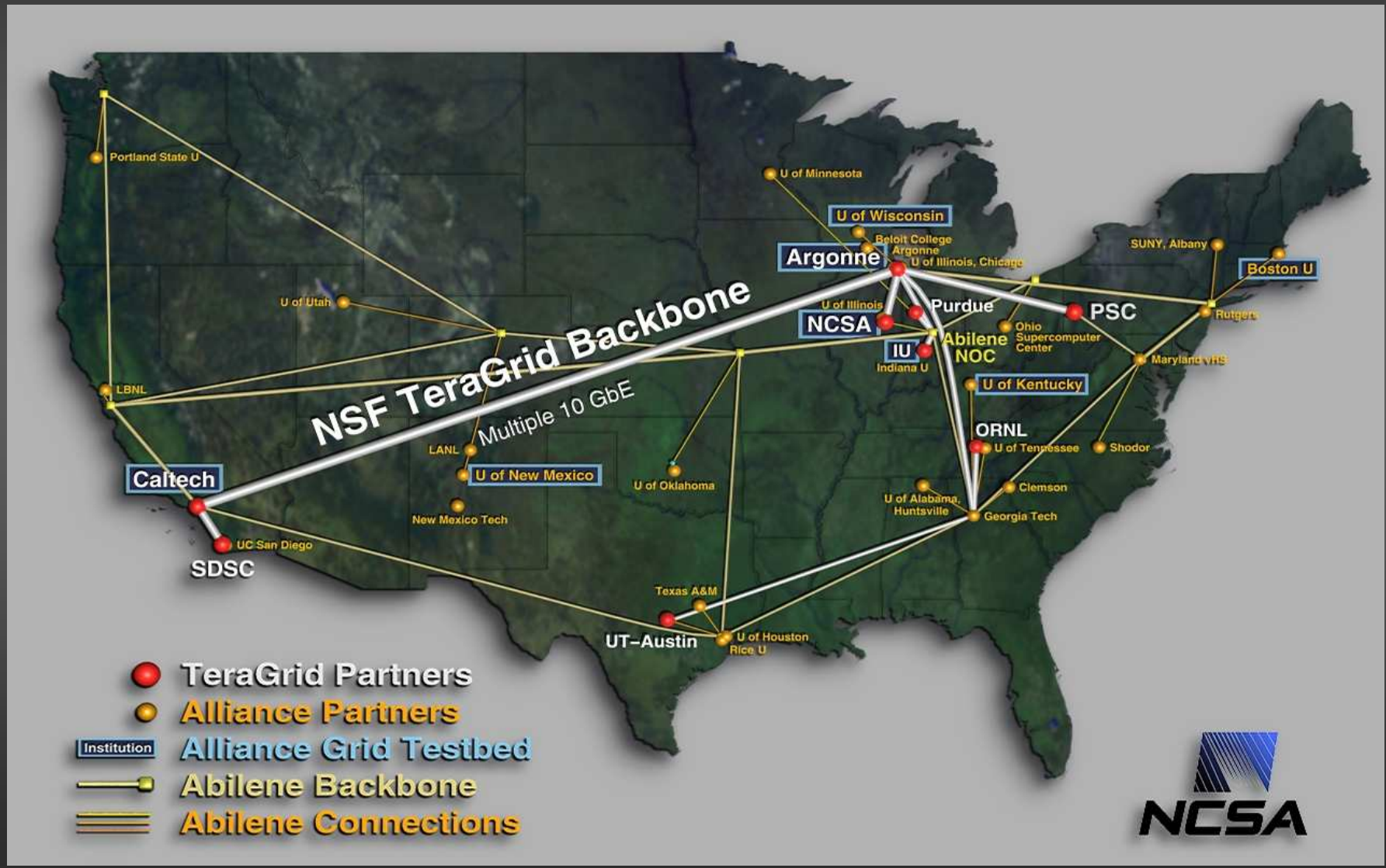
CMS

US ATLAS

LIGO

SDSS

Distributed TeraGrid Facility



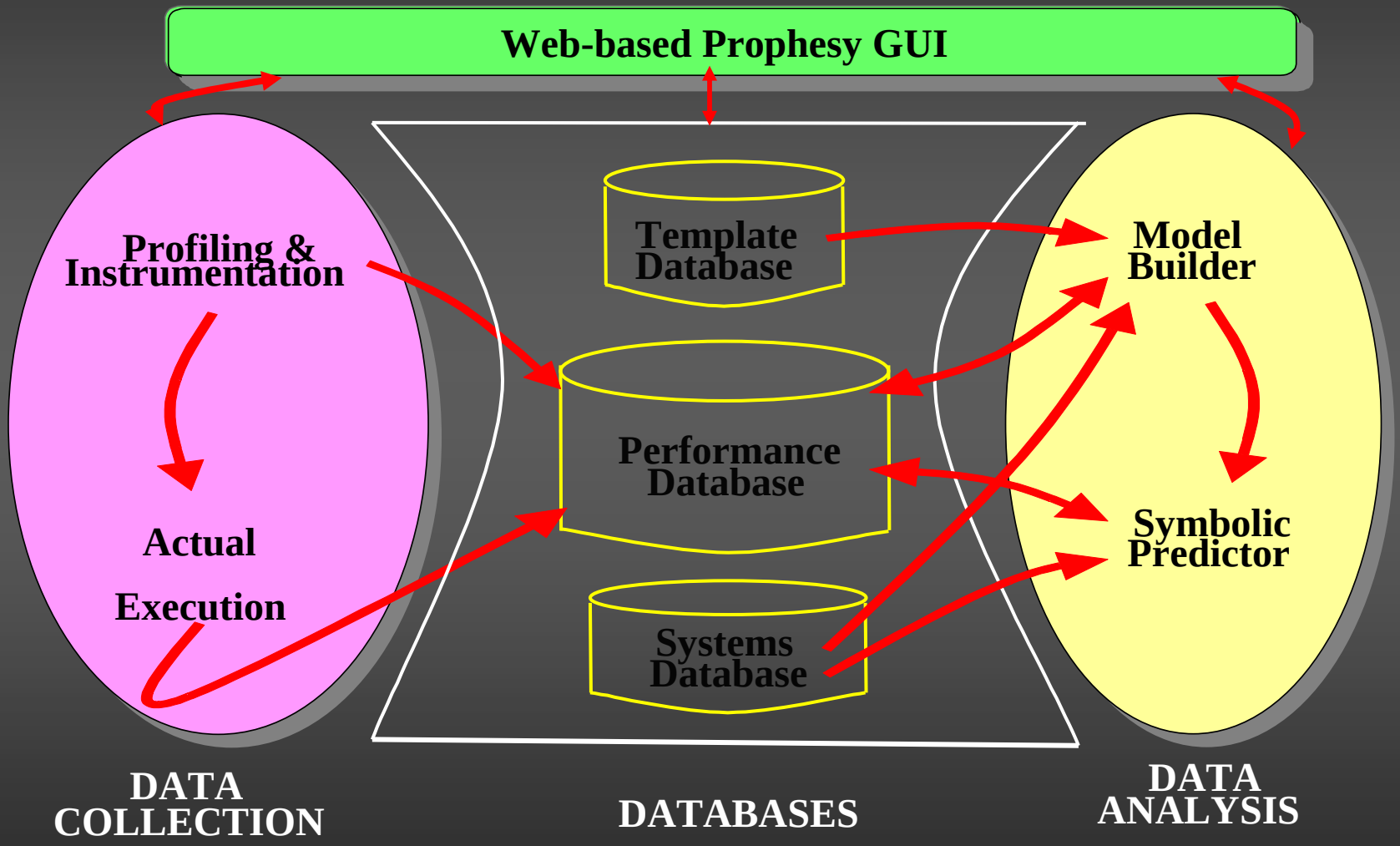
Goals

- To efficiently map jobs to appropriate resources
  - To present a resource planner that uses performance prediction based upon historical data
  - To select resources to reduce the execution time of the application

Performance Prediction

- Prophesy (http://prophesy.cs.tamu.edu)
  - Performance analysis and modeling of parallel and distributed applications
  - Three main components:
    - PAIDE System
    - Database
    - Model builder

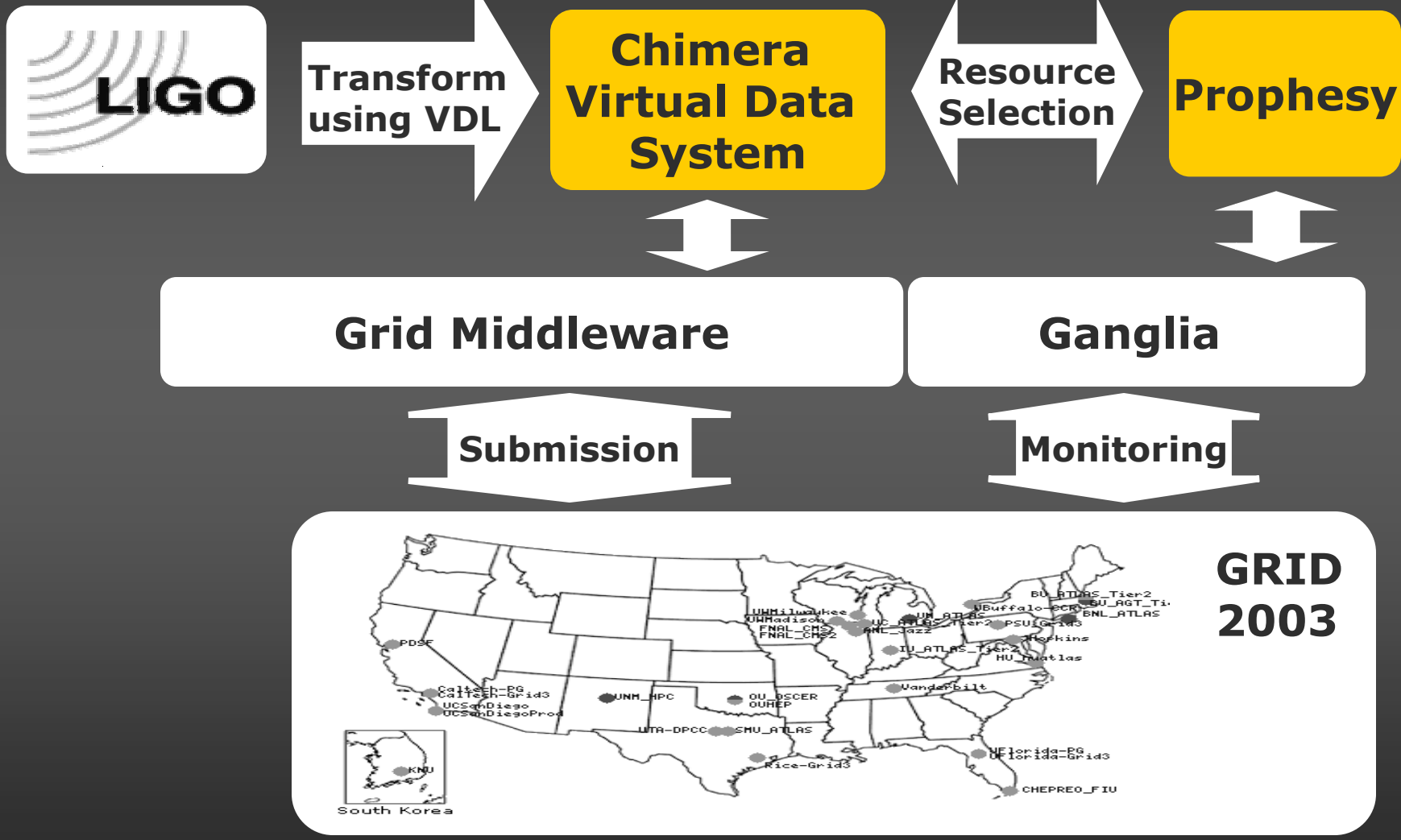
Prophesy System



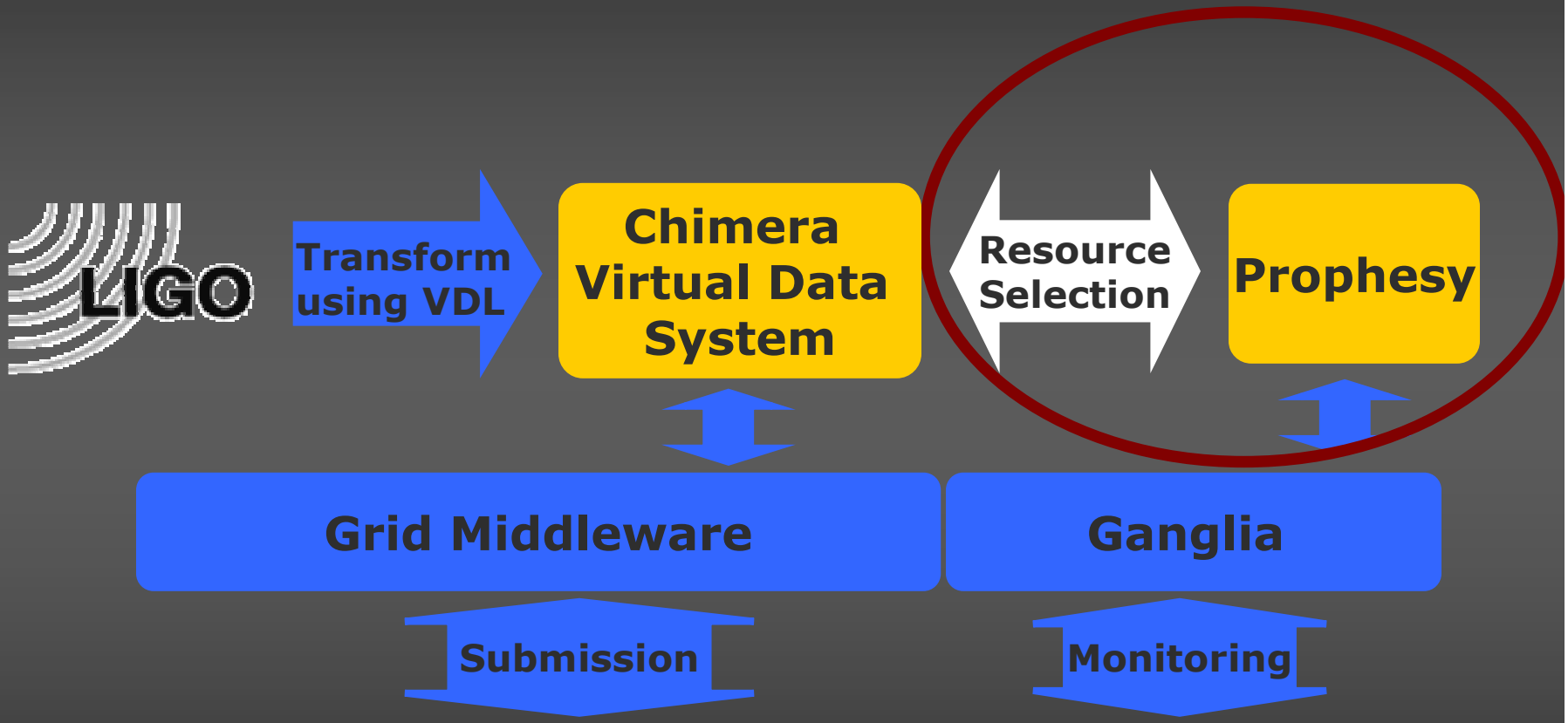
Prophesy Model Builder

- Utilize information in the Prophesy databases
  - Performance database
  - Template database
  - System database
- Three techniques
  - Curve Fitting
  - Parameterization
  - Kernel Coupling

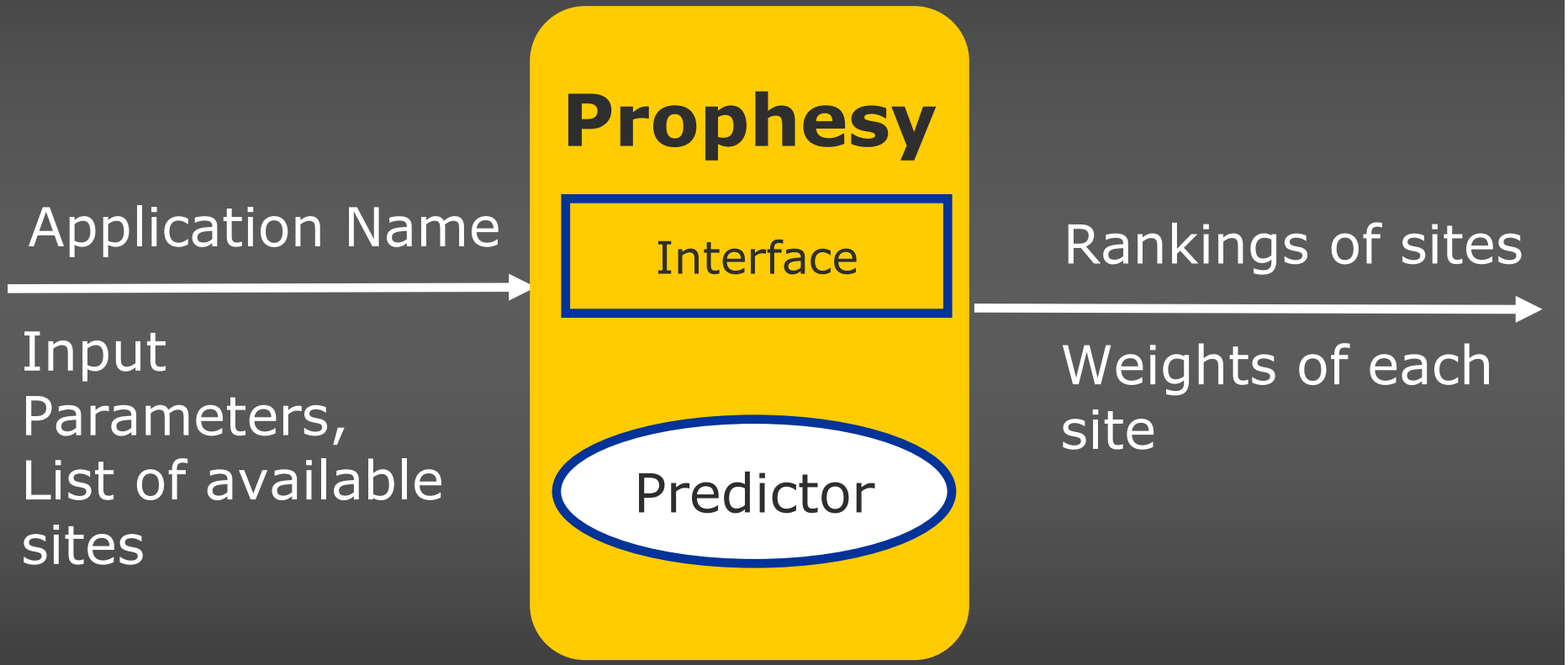
Case Study 1: GriPhyN



Resource Selector

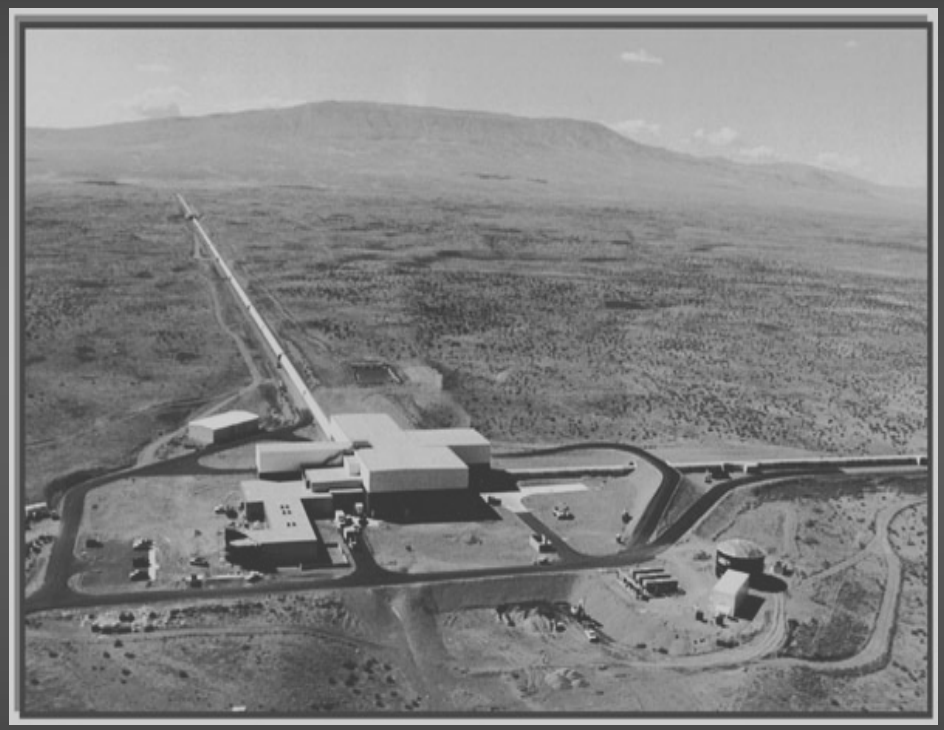


Resource Selector



Application
- GEO LIGO Pulsar Search -

- The pulsar search is a process of finding celestial objects that may emit gravitational waves
  - GEO (German-English Observatory) LIGO (Laser Interferometer Gravitational-wave Observatory) pulsar search is the most frequent coherent search method that generates F-statistic for known pulsars



Grid2003 Testbed

Execution Environment

Site Name	CPUs	Batch	VO	Compute Nodes		
				Processors	Cache	Memory
tier2b.cacr.caltech.edu	66	CONDOR	USCMS	2 X Intel Xeon 2.4 – 2.8 GHz	512 KB	1 GB
atlas.iu.edu	64	PBS	USAtlas	2 X Intel Xeon 2.4 GHz	512 KB	2.5 GB
pdsfgrid3.nersc.gov	400	LSF	USAtlas	2 X PIII 650-1.8 GHz, 2 X AMD 2100+ - 2600+	256 KB	2 GB
atlas.dpcc.uta.edu	162	PBS	USAtlas	2 X Intel Xeon 2.4 – 2.6 GHz	512 KB	2 GB
nest.phys.uwm.edu	296	CONDOR	LIGO	1 X PIII 1GHz	256 KB	0.5 GB

Selected Grid2003 Sites (Batch: local scheduler, VO: Virtual Organization)

Training Set

Parameters			UTA-DPCC		UWMilwaukee		IU_Atlas_Tier2		PDSF		Caltech-PG	
Alpha	Freq	Delta	Load	Execution Time (sec)	Load	Execution Time (sec)	Load	Execution Time (sec)	Load	Execution Time (sec)	Load	Execution Time (sec)
51	20001	2	0.58	516.33	0.00	2314.01	0.86	330.85	0.81	298.29	0.92	679.92
51	30001	2	0.03	623.96	0.01	3073.89	0.84	493.85	0.73	441.96	0.75	902.14
51	40001	2	0.55	848.64	0.90	4649.19	0.00	655.76	1.31	596.23	0.75	1245.50
51	50001	2	0.96	920.96	0.67	5485.29	0.00	842.04	1.37	747.34	0.70	1527.59
51	60001	2	0.54	1227.83	0.75	6339.79	0.98	1006.24	0.87	2590.73	0.75	1903.95
51	70001	2	0.56	1393.55	0.39	7645.19	0.94	1173.29	0.67	1046.95	0.50	2013.06
51	80001	2	1.01	1685.02	0.53	7865.04	0.86	1338.41	0.74	1191.15	1.23	2697.18
51	90001	2	0.46	1760.32	1.00	8718.98	0.88	1503.90	0.71	1356.91	0.99	2880.33
151	10001	2	0.50	569.54	0.00	3556.01	0.00	506.81	0.86	558.23	0.80	746.41
251	10001	2	0.51	1074.80	0.85	7421.57	0.00	842.76	0.94	2151.70	0.75	1534.51
351	10001	2	0.51	1375.05	0.85	8850.15	0.86	1184.85	0.85	1043.60	0.69	2523.76
451	10001	2	0.54	1898.35	0.22	10220.80	0.91	1495.69	0.66	1340.08	0.50	2243.62
551	10001	2	0.50	2397.77	0.47	10893.11	0.00	1820.16	0.77	1632.12	0.43	3725.79
651	10001	2	0.53	3390.52	0.25	11727.07	0.96	2149.54	1.03	1925.37	0.44	3867.02
751	10001	2	0.59	3975.04	0.52	12681.72	0.00	2481.77	0.81	2721.29	0.96	5017.09
851	10001	2	0.84	4316.29	0.19	13467.01	0.90	2809.56	0.73	2512.31	0.80	5533.36
951	10001	2	1.03	4364.20	0.81	14398.22	0.00	3142.17	1.22	2828.59	0.75	5607.35
1051	10001	2	0.52	5035.14	0.18	15266.22	0.00	3470.28	0.78	3093.50	0.72	5929.08
51	10001	11	0.51	1480.62	3.56	2684.43	0.00	931.36	0.35	1028.01	0.88	1872.93
51	10001	101	0.01	19098.76	0.34	37183.61	0.04	8520.45	0.46	9440.83	0.95	17270.18
151	20001	11	0.51	16659.95	0.25	20414.31	0.00	5380.85	0.90	6001.33	0.75	8530.43
251	20001	11	0.07	9510.51	0.06	35462.17	0.50	8950.66	0.90	8005.11	0.75	15375.46
451	20001	2	0.61	2175.48	0.50	12644.95	0.00	2984.89	0.93	3289.23	0.41	4576.72
151	10001	11	0.52	2586.53	0.91	10573.75	0.16	2758.50	0.66	3002.53	0.50	3411.83
251	10001	11	0.05	4218.16	1.68	18508.70	0.16	4571.52	0.86	11808.97	0.66	6687.94

Comparison

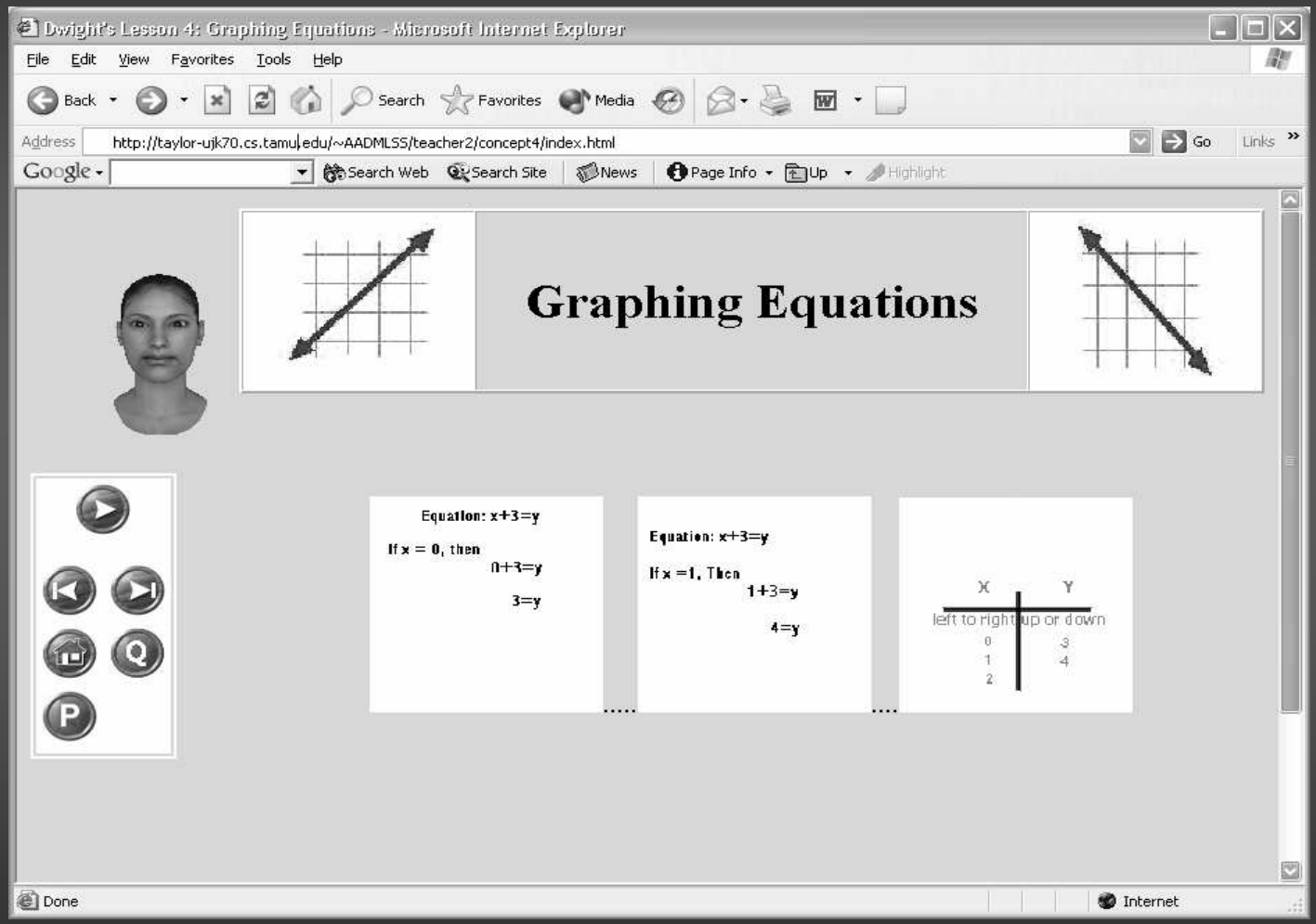
- Load-based selection method
  - Ganglia [1] to monitor the Grid2003
  - Selects the least loaded site
- Random selection method
- Static method based on a Linpack [2] benchmark

[1] Ganglia, http://ganglia.sourceforge.net/.
[2] Linpack software library, http://www.netlib.org/benchmark/Linpackc/.

Experimental Results

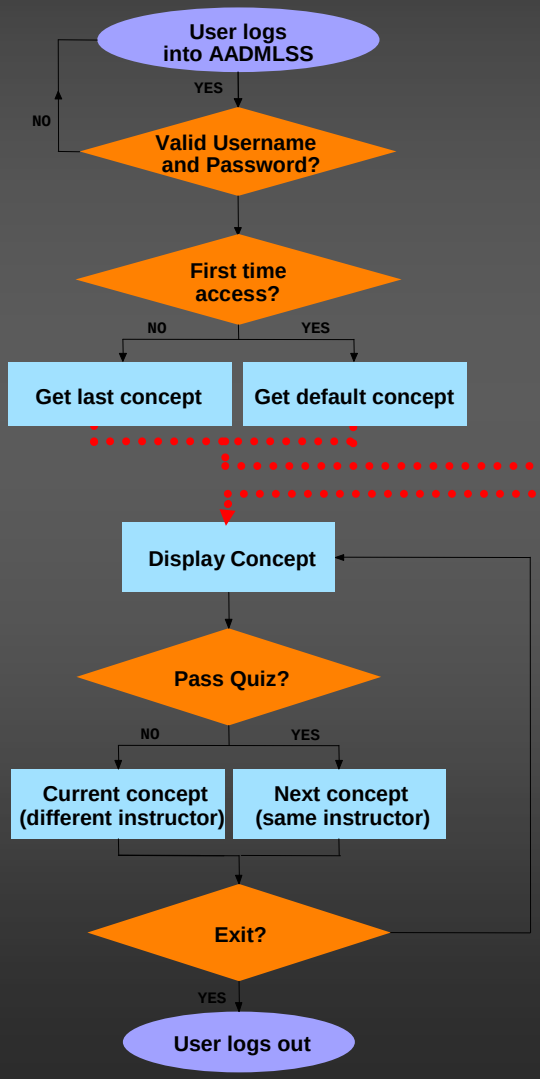
Parameters			Prediction		Load-based			Random			UTA-DPCC (Linpack)	
Alpha	Freq	Delta	Site	Time	Site	Time	Error	Site	Time	Error	Time	Error
51	10001	2	IU	171.43	IU	171.43	0.00%	Caltech	215.49	20.45%	187.31	8.48%
51	110001	2	PDSF	2006.90	UTA	1922.62	-4.38%	IU	1835.39	-9.34%	1922.62	-4.38%
351	20001	11	IU	12625.87	UTA	22209.95	43.15%	IU	12625.87	0.00%	22209.95	43.15%
451	10001	11	IU	8329.12	IU	8329.12	0.00%	UTA	8569.79	2.81%	8569.79	2.81%
451	20001	2	IU	2924.13	UTA	5404.14	45.89%	IU	2924.13	0.00%	5404.14	45.89%
551	40001	2	PDSF	6466.63	IU	7242.46	10.71%	Caltech	9501.03	31.94%	11053.77	41.50%
651	20001	2	IU	4322.18	UTA	8581.00	49.63%	IU	4322.18	0.00%	8581.00	49.63%
751	20001	11	PDSF	24467.19	UTA	38318.12	36.15%	IU	27402.86	10.71%	38318.12	36.15%
851	20001	2	PDSF	6204.16	IU	5623.11	-10.33%	UTA	7909.41	21.56%	7909.41	21.56%
1351	10001	2	PDSF	5007.49	UTA	11697.98	57.19%	PDSF	5007.49	0.00%	11697.98	57.19%
Average Error Rate							**22.80%**			**7.81%**		**30.20%**

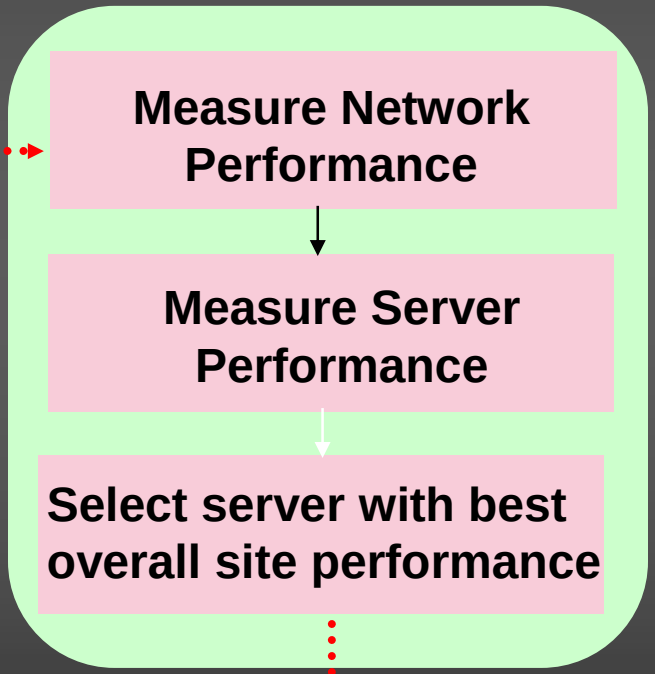
Case Study 2: AADMLSS



African American Distributed Multiple Learning Styles System (AADMLSS) developed by Dr. Juan E. Gilbert

Site Selection Process



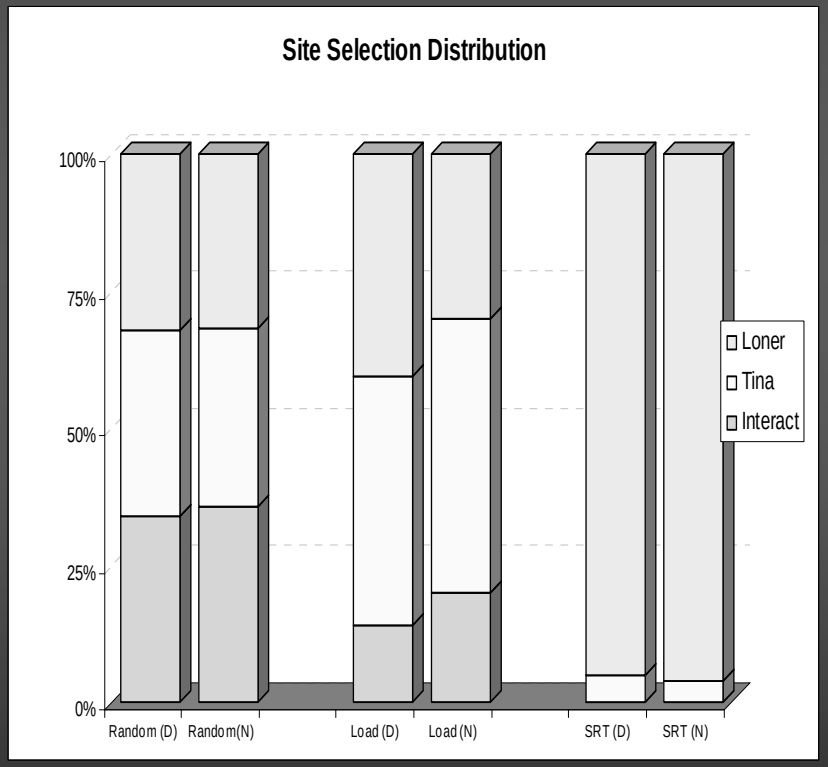


Testbed Overview

CATEGORY	SPECS	Loner (TX)	Tina (MA)	Interact (AL)
Hardware	CPU Speed (MHz)	997.62	1993.56	697.87
	Bus Speed (MB/s)	205	638	214
	Memory (MB)	256	256	256
	Hard Disk (GB)	30	40	10
Software	O/S	Redhat Linux 9.0	Redhat Linux 9.0	Redhat Linux 9.0
	Web Server	Apache 2.0	Apache 2.0	Apache 2.0
	Web Application	PHP 4.2	PHP 4.2	PHP 4.1

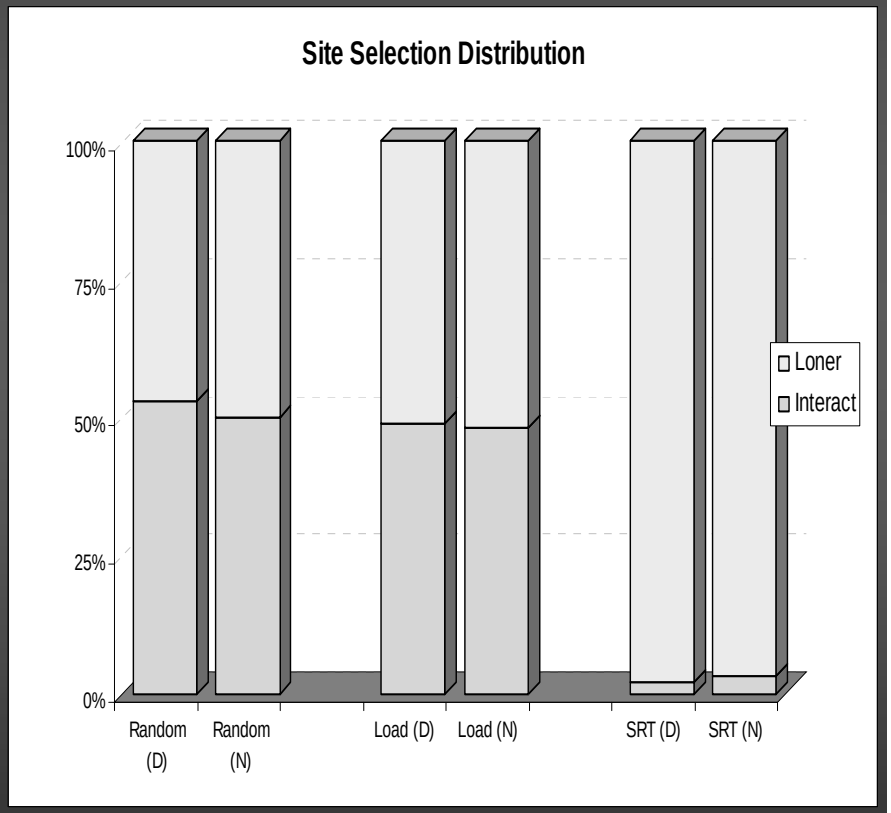
Experimental Results – 3 Servers -

Concept	SRT-LOAD (%)	SRT-RANDOM (%)
3/0/0 D	6.21	14.05
3/0/1 D	12.13	21.94
3/0/2 N	14.02	25.83
3/0/3 N	18.12	23.52
3/1/0 N	8.05	12.04
3/1/1 N	7.31	12.25
3/1/2 N	12.60	18.74
3/1/3 N	10.96	19.11
3/2/0 N	7.93	12.58
3/2/1 N	8.05	14.25
3/2/2 N	9.14	15.97
3/2/3 D	9.79	20.58
3/3/0 D	8.94	13.64
3/3/1 D	8.26	16.74
3/3/2 D	9.21	15.21
3/3/3 D	9.07	19.36
AVERAGE	**10.04**	**17.24**



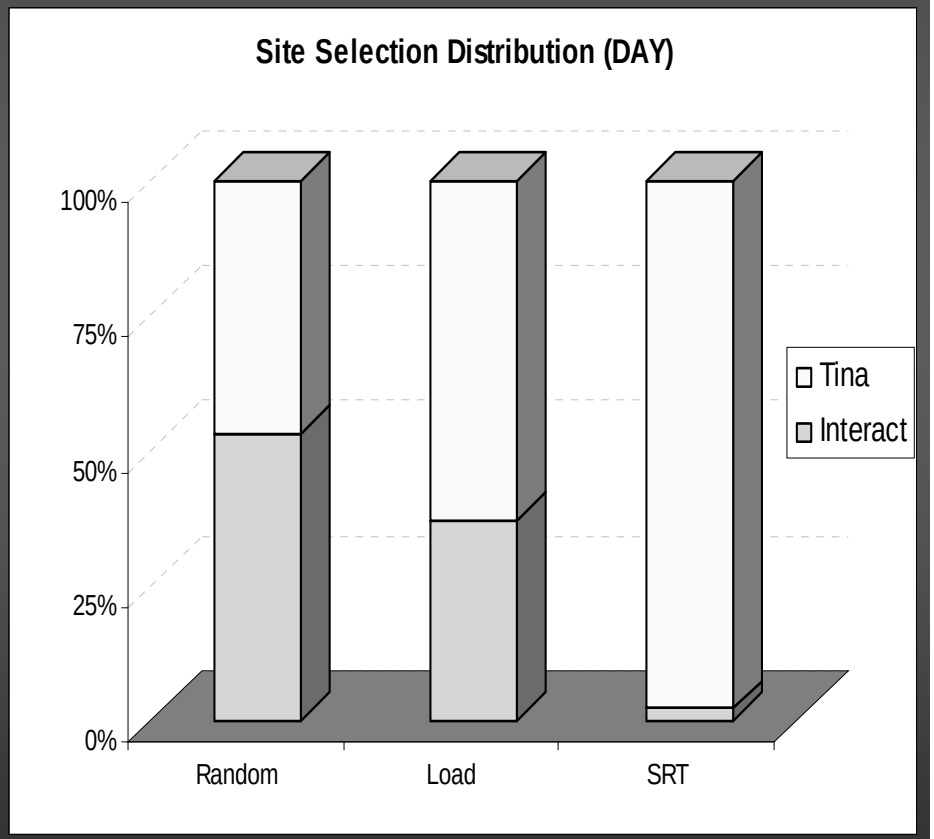
Experimental Results – 2 Servers (local and remote) -

Concept	SRT-LOAD (%)	SRT-RANDOM (%)
3/0/0 D	9.91	10.24
3/0/1 D	13.04	15.06
3/0/2 D	18.06	19.16
3/0/3 D	20.54	21.29
3/1/0 N	9.81	9.58
3/1/1 N	7.02	7.91
3/1/2 N	11.35	12.15
3/1/3 N	10.47	10.36
3/2/0 D	8.56	8.67
3/2/1 D	8.75	9.75
3/2/2 D	10.06	10.92
3/2/3 D	10.15	10.50
3/3/0 N	8.41	9.56
3/3/1 N	8.58	8.08
3/3/2 N	8.31	7.95
3/3/3 N	10.21	10.19
AVERAGE	**10.83**	**11.34**



Experimental Results – 2 Servers (remote and remote) -

Concept	SRT-LOAD (%)	SRT-RANDOM (%)
3/0/0 D	3.13	4.03
3/0/1 D	4.26	5.97
3/0/2 D	7.02	8.28
3/0/3 D	8.64	9.02
3/1/0 D	3.25	4.94
3/1/1 D	3.27	4.10
3/1/2 D	3.93	5.97
3/1/3 D	3.64	4.08
3/2/0 D	3.15	3.32
3/2/1 D	4.39	5.20
3/2/2 D	5.80	5.97
3/2/3 D	6.52	6.95
3/3/0 D	4.39	5.64
3/3/1 D	4.16	5.20
3/3/2 D	4.81	5.73
3/3/3 D	5.02	5.58
AVERAGE	**4.71**	**5.62**



Summary

- Presented Resource selection method based upon performance predictions
- Illustrated the advantages of using performance predictions using two case studies
  - Large scale scientific application GEO LIGO on Grid2003
    - An average 23% better than load-based selection
  - AADMLSS on 3 servers
    - An average 10% better than load-based selection

Thanks!

Questions?