An Evaluation of Alternative Designs for a Grid Information Service

Warren Smith
Computer Sciences Corporation
NASA Ames Research Center

Abdul Waheed
MRJ Technology Solutions
NASA Ames Research Center

David Meyers
Directory Research L.L.C.
NASA Ames Research Center

Jerry Yan
NASA Ames Research Center

Motivation

- The Globus info. service wasn't working well
  - Many updates of data from Globus daemons
  - Saturated the single server
  - Users couldn't retrieve information
- Created a second server for NASA & Alliance
  - Things were great on that server, but a bit slow on the other server
- How exactly is the info. service being used?
- What are the best servers & configurations?

Outline

- Background
- Workload characterization
- Methodology
- Performance evaluation
- Conclusions

Globus Metacomputing Directory Service

- The Globus grid information service
- Information about organizations, people, computers, networks, software, applications, ...
- Accessed using the Lightweight Directory Access Protocol (LDAP)
- Specified data format and access protocol but not implementation

MDS History

- Originally contained in a single Netscape LDAP server that ended up at NCSA
- Then contained in 2 Netscape LDAP servers at NCSA
  - 1 server for NASA and the Alliance
  - 1 server for everything else
- Now:
  - NASA is currently using 4 Netscape LDAP servers at 3 sites
  - Globus v1.1.3 totally changed the MDS
    - No central servers
    - Many OpenLDAP servers

LDAP

- Data organized as entries in a tree
- Entries named using their position in the tree
- Entries can be added, deleted, modified
- Searches can be performed over attributes of entries
- Many providers of LDAP servers

Improving LDAP Performance

- Distribution of data
  - Sub-trees of data can be placed on different servers
  - Find the server by using referrals
  - Supports more updates but slows broad searches
- Replication of data
  - More servers for searches
  - More servers to update data
- Indexes
  - Lookup table based on an attribute value
  - Servers can construct and maintain
  - Improve search performance

Uses of MDS

- Host registration and periodic update
  - Host, network interfaces, networks, software
- Periodic GRAM reporter updates
  - Load, queues, users that have access, running applications
  - By default: Every 30 seconds and no user or application information
- Finding GRAM contacts for a host
- Determine status of applications
- Currently, complex queries not performed often

Workload Characterization

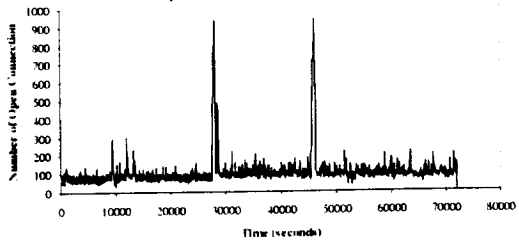
- 20 hours of trace data
- Recorded when there was a single server
- 86,695 connections
- 633,672 operations
- 8.8 operations per second

Operations in Workload

Operation	Number of Operations	Percent of Total Operations	Percent of Operations Resulting in Error
Add	1,044	0.73	90.33
Delete	81	0.06	7.41
Modify	13,4611	93.84	2.83
Search	7710	5.37	76.10
Total	143,446	100.00	7.41

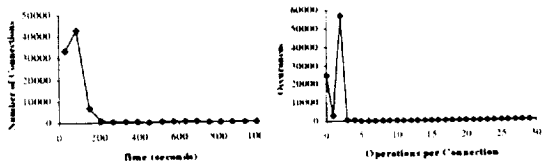
- Modifies are from GRAM reporter
- Adds are tried when modifies fail
  - Typically adds also fail

Open Connections



- 90 open connections on average
- Periodic spike: Netscape LDAP maintenance?

Connection Properties



- 88% less than 120 sec, 97% less than 240 sec
- 97% have 2 or less operations

Classification of MDS Connections

- Classified 99.96% of all connections
- 67.45% are connections from GRAM reporters
- 22.34% are connect, bind failure, unbind, close
- Use this knowledge in our experiments

Experimental Analysis I

- Perform trace-driven simulation of accesses to the MDS
- Trace data does not record changes to entries in the MDS
  - Use our knowledge of Globus to construct realistic entries off-line
  - For example, place a random number as the value in the free nodes attribute
- Simulator:
  - Written in Java
  - Runs on 1+ workstations
  - Simulate accesses in real time, faster, or slower

Experimental Analysis II

- Use 1 or 2 UltraSparc workstations to run MDS servers
- Start the server(s)
- Load initial MDS contents
- Run the simulation
- Record MDS access times in the simulator

Single Server Comparison

Load	LDAP Server	Add (ms)	Delete (ms)	Modify (ms)	Search (ms)	Weighted Average (ms)
1.0	OpenLDAP	101	108	117	[illegible]	[illegible]
	Vendor 1	121	106	159	2463	283
	Vendor 2	933	1230	903	1358	928
2.0	OpenLDAP	104	141	233	287	379
	Vendor 1	270	351	581	4031	764
	Vendor 2	31722	36920	26558	22945	26407

- OpenLDAP crashed unless we limited open connections to 50 and new connections per second to 20
- Vendor 2 optimized for search performance
- CPU load for Vendor 2 was 5.5, Vendor 1 was 0.1

Indexing

- Hashtable for a specific attribute keyed on the value
- Advantage: Improve search performance
- Disadvantage: Must be maintained
- We added an index to the Vendor 1 server
  - GlobalJobID attribute
  - Improve job status searches which are made over the entire tree
- Improved search performance by 79%
- Decreased update performance by 38%
- Slower updates may not be worth faster searches

Data Distribution

- Purpose is to
  - Support very large databases
  - Improve add, modify, and delete performance
  - Can improve search performance
- For our experiment, we:
  - Distribute data across 2 servers
  - NASA & Alliance on one, everything else on the other
  - Used the LDAP server from Vendor 1
  - Performed a simulation in $\frac{1}{4}$ real time

Data Distribution II

- In general, we found that:
  - Update time increased by 27%
  - Due to higher load on one machine
  - Search time decreased by 76%
- On the NASA/Alliance server:
  - XXX initial records
  - Update time ..
  - Search time ..
- On the other server:
  - XXX initial records
  - Update time...
  - Search time...

Globus 1.1.3

- Grid Resource Information Server (GRIS) on each machine
  - OpenLDAP front end over a GRAM reporter
  - Can be queried (almost) like any other LDAP server
- Organizational server
  - OpenLDAP
  - GRIS servers register
  - Chaining to get data from GRIS servers
  - Caches data from GRIS servers
- Could be higher-level servers or not

Globus 1.1.3

- "Pull" model: data pulled to the organizational server when user does a search
- No updates to "push"
  - Less load on the servers
- If data isn't cached, searches may be slow
- For broad searches, could be very slow
- Have not yet evaluated experimentally

Conclusions

- Analyzed accesses to the Globus information service
  - 90 connections open on average
  - 8.8 operations per second
  - 94% of operations are modifies
- Evaluated LDAP servers using trace-driven simulations
  - OpenLDAP not very robust
  - One commercial server has better performance
  - Other commercial server appears to be optimized for search performance and SMPs
  - Indexing reduces search time but increases update time
  - Distributing data increased our update time and decreased the search time

Future Work

- Evaluate more configurations of LDAP servers
- Evaluate Globus v1.1.3 MDS
- Develop a simulation environment with synthetic users, machines, daemons
  - More flexible experiments
  - Planning future needs