



DiPerF: automated Distributed PERformance testing Framework

Catalin Dumitrescu, Ioan Raicu,
Matei Ripeanu, Ian Foster

Distributed Systems Laboratory
Computer Science Department
University of Chicago

Introduction



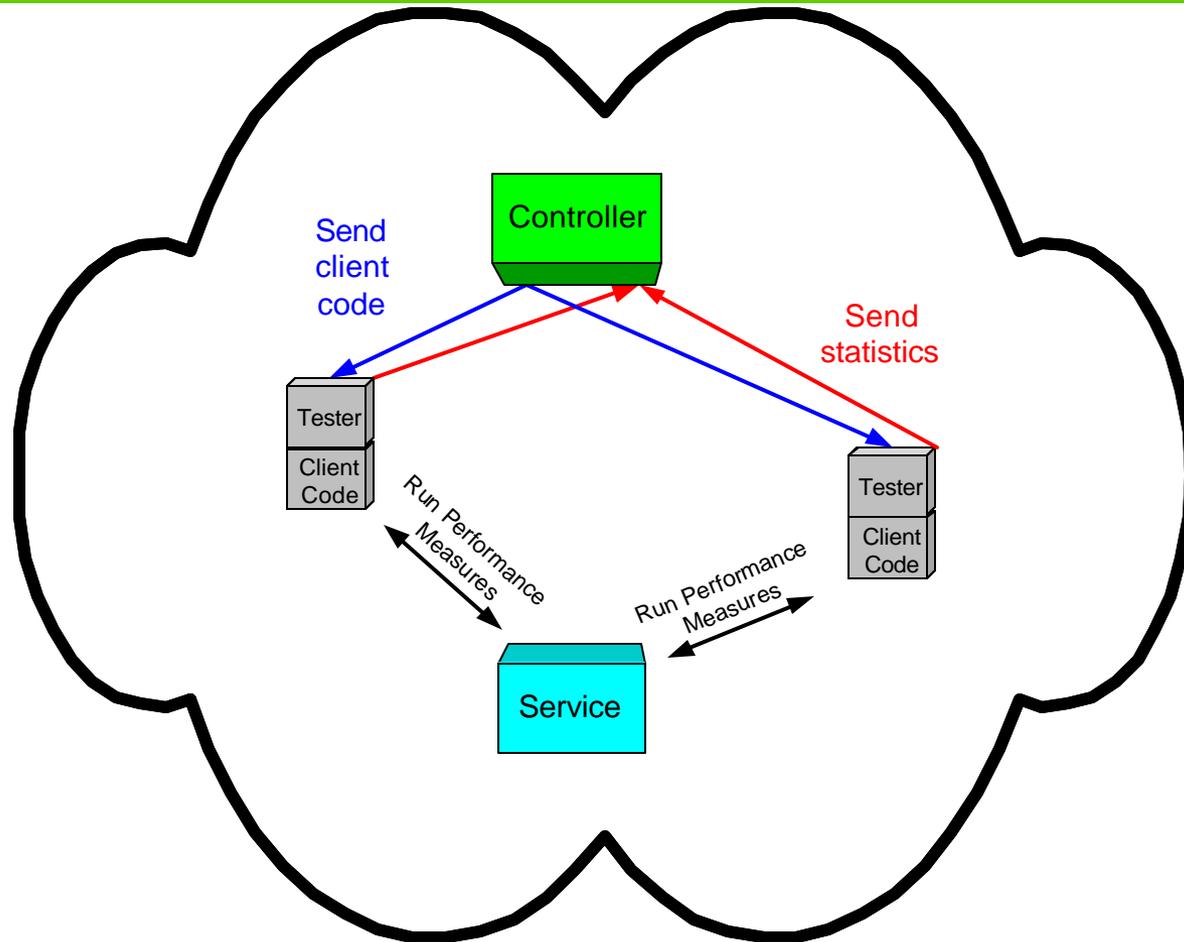
- Goals
 - Simplify and automate distributed performance testing
 - grid services
 - web service
 - network service
 - Define a comprehensive list of performance metrics
 - Produce accurate client views of service performance
 - Create analytical models of service performance
- Framework implementation
 - Grid3
 - PlanetLab
 - NFS style cluster (UChicago CS Cluster)

Framework



- Coordinates a distributed pool of machines
 - Tested with over 100 machines
 - Scalable to 1000s of machines
- Controller
 - Receives the address of the service and a client code
 - Distributes the client code across all machines in the pool
 - Gathers, aggregates, and summarizes performance statistics
- Tester
 - Receives client code
 - Runs the code and produce performance statistics
 - Sends back to “controller” statistic report

Architecture Overview

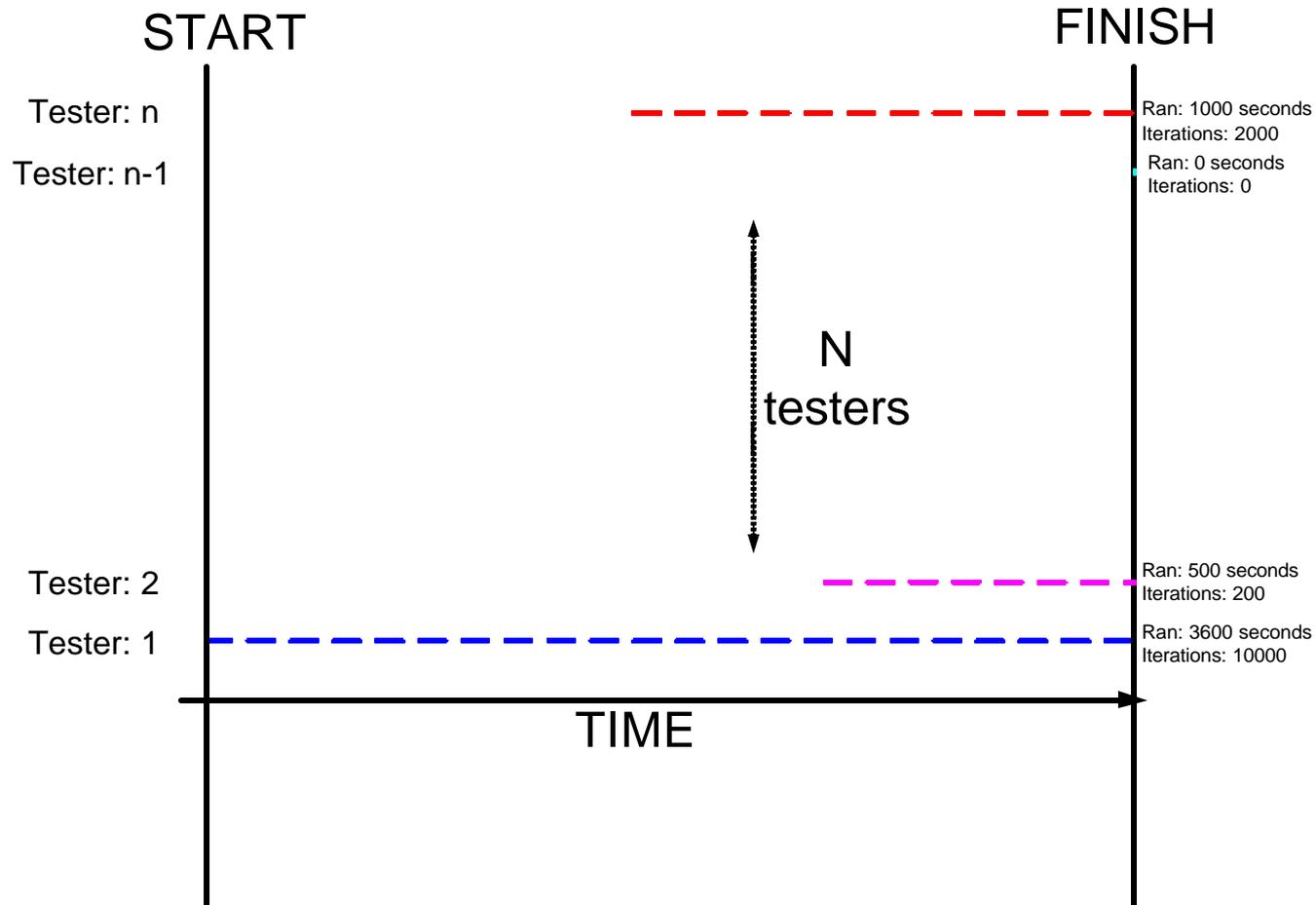


Time Synchronization



- Time synchronization needed at the testers for data aggregation at controller?
 - Distributed approach:
 - Tester uses Network Time Protocol (NTP) to synchronize time
 - Centralized approach:
 - Controller uses time translation to synchronize time
 - Uses network RTT to estimate offset between tester and controller

Metric Aggregation



Performance Metrics



- **service response time**
 - time from when a client issues a request to when it is completed minus the network latency and minus the execution time of the client code
- **service throughput**
 - number of jobs issued per second and completed successfully by the end of the test time
- **service utilization**
 - percentage of service resource utilization throughout the entire test per client
- **service balance among clients**
 - ratio between number of jobs completed and service utilization per client
- **service load**
 - number of concurrent clients per second accessing the service
- **network latency to the service**
 - time taken for a minimum sized packet to traverse the network from the client to the service
- **time synchronization error**
 - real time difference between client and service measured as a function of network latency

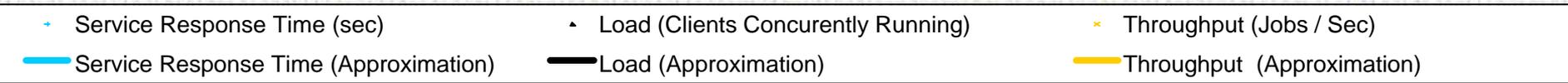
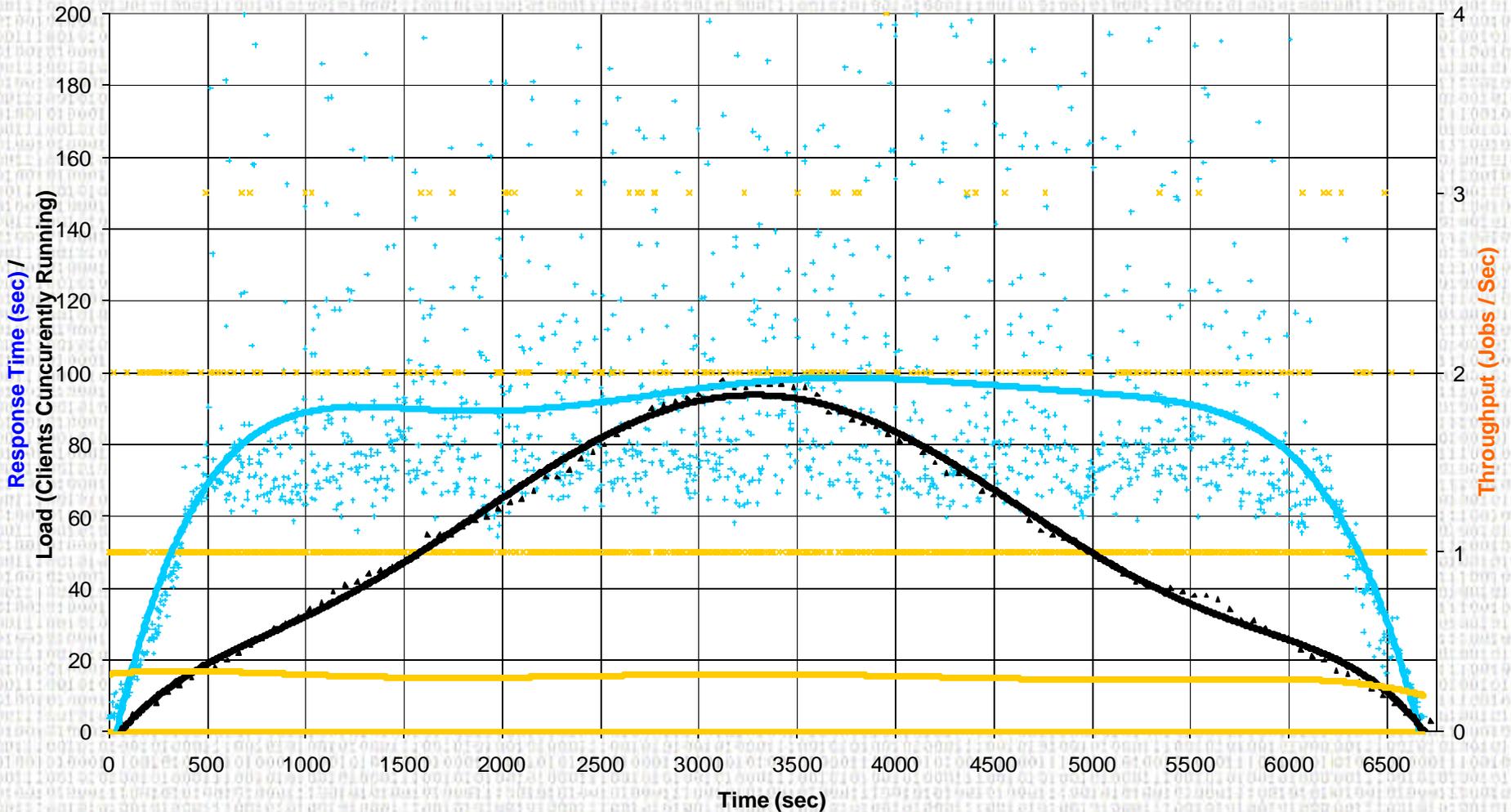
Services Tested



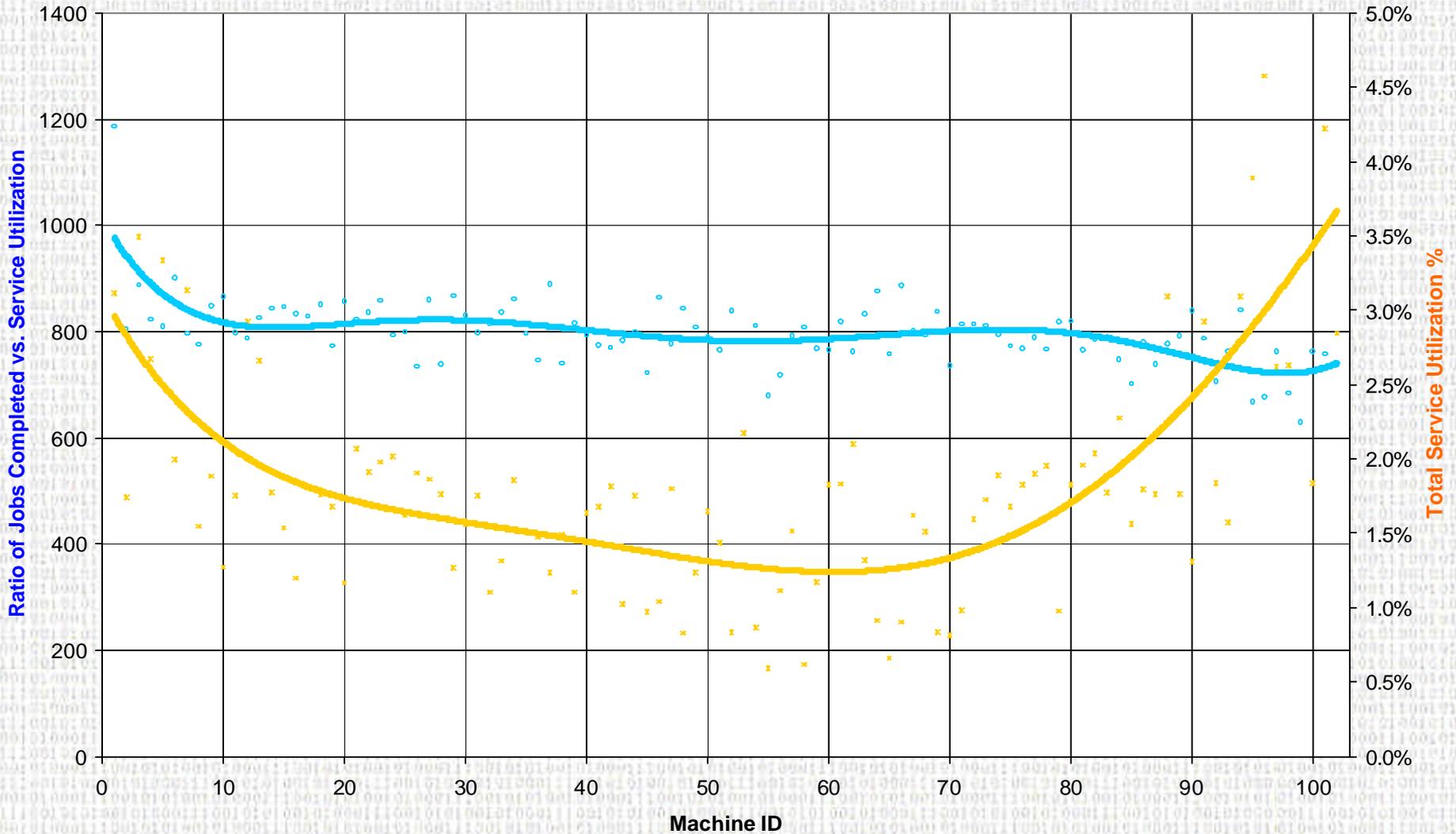
- HTTP
 - client used “wget” to invoke a program using CGI over HTTP on an Apache HTTP server
- GT2 GRAM
 - job submission via Globus Gatekeeper 2.4.3 using Globus Toolkit 2.4
- GT3.02
 - simple factory-based grid service; create a new service and invoke a method of the service without security features enabled
- GT3.2
 - identical to GT3.02 except that GT3.2 was an Alpha release
- MonaLisa
 - monitoring grid webservice

GT2.4

Response Time, Load, and Throughput (in time)

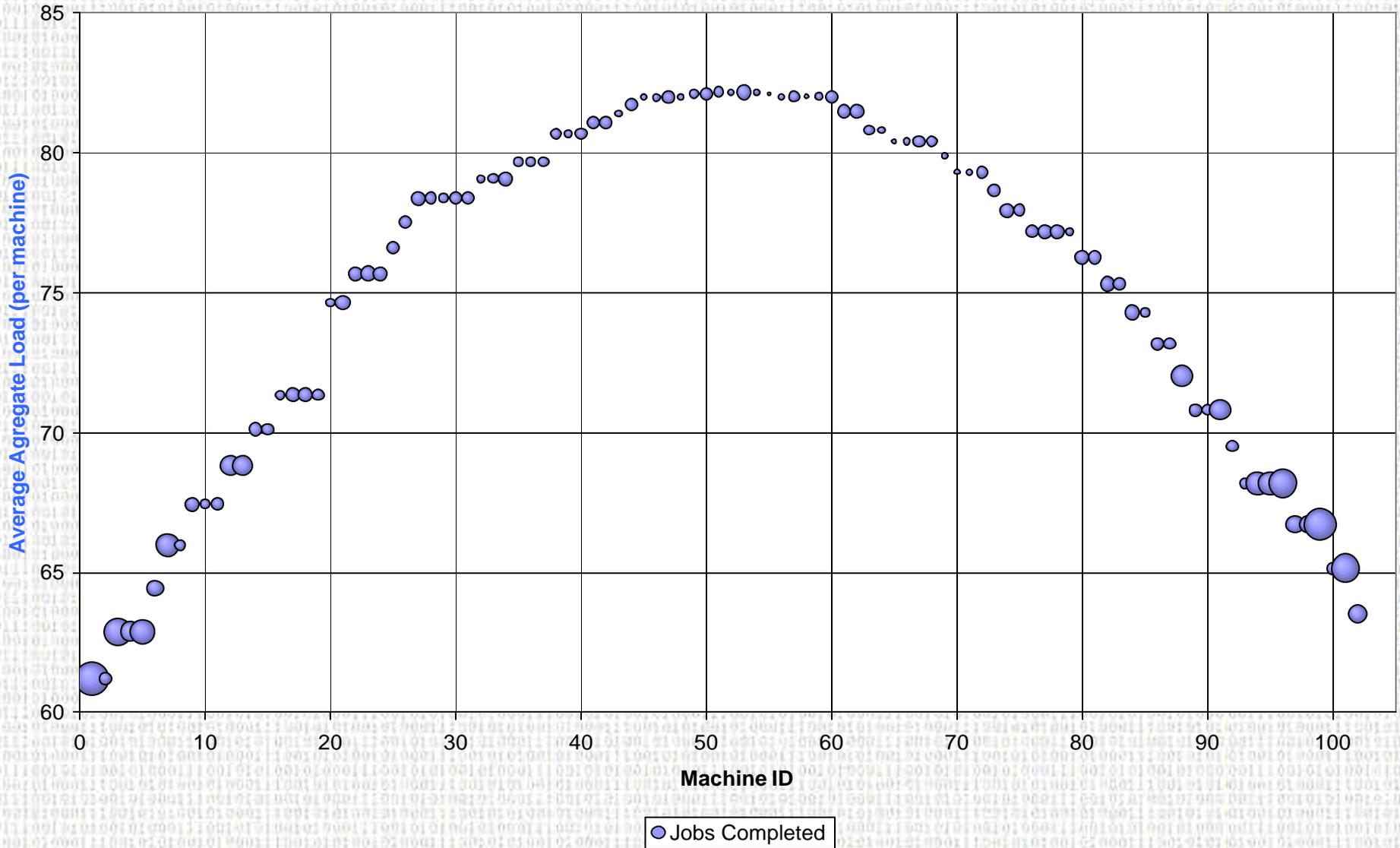


GT2.4 Service Utilization (per Machine)



GT2.4

Average Load and Jobs Completed (per Machine)



Analytical Model



- Model performance characteristics
 - used to estimate a service's performance based on the service load
 - Throughput
 - Service response time
- Dynamic resource allocator
 - Maintain QoS while maximizing resources utilization
- Polynomial approximations
 - Throughput: $y = -3 \cdot 10^{22} x^6 + 6 \cdot 10^{18} x^5 - 5 \cdot 10^{14} x^4 + 2 \cdot 10^{10} x^3 - 2 \cdot 10^7 x^2 + 0.0001x + 0.3212$
 - Service response time:
$$y = -2 \cdot 10^{19} x^6 + 4 \cdot 10^{15} x^5 - 3 \cdot 10^{11} x^4 + 10^7 x^3 - 0.0003x^2 + 0.2545x + 9.1001$$
- Neural networks

Contributions & Future Work



- Contributions
 - Analytical models: resource managers
 - Service capacity
 - Scalability study
 - Resource distribution among clients
 - Accurate client views of service performance
 - How network latency affects service performance
- Future Work
 - Verify analytical models
 - Polynomial Approximations
 - Neural Networks
 - Test more services

References



- Presentation Slides

- http://people.cs.uchicago.edu/~iraicu/research/documents/uchicago/cs33340/diperf_presentation.pdf

- Report

- http://people.cs.uchicago.edu/~iraicu/research/documents/uchicago/cs33340/diperf_report.pdf

- References

- <http://www.globus.org>
- <http://www.ivdgl.org>
- <http://www.planet-lab.org>
- <http://www.ntp.org>
- <http://xenia.media.mit.edu/~nelson/research/ntp-survey99/html/>