

Επιμετρήσεις Επιδόσεων Υπολογιστικού Πλέγματος

«Test-driving the Grid»

Μάριος Δ. Δικαιάκος

Τμήμα Πληροφορικής
Πανεπιστήμιο Κύπρου



Acknowledgments



- In collaboration with:
 - *George Tsouloupas, UCY*
- With funding from:

crossgrid



egee
Enabling Grids for
E-science in Europe

CoreGRID

<http://grid.ucy.ac.cy>

Choosing an automobile



es techniques

Vehicle	Engine	Class	Year	Price
1990	1.8	Class A	1990	1.8
1991	1.8	Class B	1991	1.8
1992	1.8	Class C	1992	1.8
1993	1.8	Class D	1993	1.8
1994	1.8	Class E	1994	1.8
1995	1.8	Class F	1995	1.8
1996	1.8	Class G	1996	1.8
1997	1.8	Class H	1997	1.8
1998	1.8	Class I	1998	1.8
1999	1.8	Class J	1999	1.8
2000	1.8	Class K	2000	1.8
2001	1.8	Class L	2001	1.8
2002	1.8	Class M	2002	1.8
2003	1.8	Class N	2003	1.8
2004	1.8	Class O	2004	1.8
2005	1.8	Class P	2005	1.8
2006	1.8	Class Q	2006	1.8
2007	1.8	Class R	2007	1.8
2008	1.8	Class S	2008	1.8
2009	1.8	Class T	2009	1.8
2010	1.8	Class U	2010	1.8
2011	1.8	Class V	2011	1.8
2012	1.8	Class W	2012	1.8
2013	1.8	Class X	2013	1.8
2014	1.8	Class Y	2014	1.8
2015	1.8	Class Z	2015	1.8

Type of operator

For hire:		
Class A	98.1	1.3
Class B	94.8	1.5
Class C	73.7	2.6
Total for hire	93.9	1.5
Self-haulers	62.1	8.4
All operators	89.1	1.8



<http://grid.ucy.ac.cy>

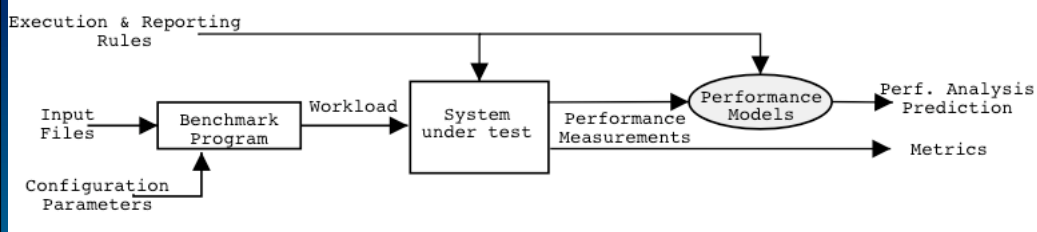
Benchmarks: definition



- Standardized programs (or detailed specifications) designed or chosen to investigate performance properties of computer systems:
 - Characterize performance capacity and behavior.
 - Compare different systems in a fair manner.
 - Guide the optimization and assessment of system designs and implementation.
 - Help researchers establish quantitative arguments in systems research.
- Complete applications, kernels, probes, or synthetic programs. E.g:
 - Whetstone, SPEC, Parkbench, NAS, TPC-C, Linpack, SPLASH, MediaBench...

<http://grid.ucy.ac.cy>

Benchmarking process



- Conditions of execution and measurement:
 - Well-defined and publicly available.
- Benchmarks are required to be:
 - Portable, fair, relevant, easy to measure, easy to explain.

<http://grid.ucy.ac.cy>

Benchmarking: not a panacea



- Questions of fairness and relevance.
 - Relatively easy to optimize systems for selected benchmarks; what about real workloads?
- Questions of cost.
 - Esp. when benchmarking network-centric systems.
- Provide no direct insights about derived metrics.
 - The more complex a system is, the more difficult it becomes to combine individual measurements into a meaningful system evaluation.
- What about performance prediction based on benchmarking?
 - Benchmark results need to be interpreted in the context of models for hardware, architecture, available resources.

<http://grid.ucy.ac.cy>

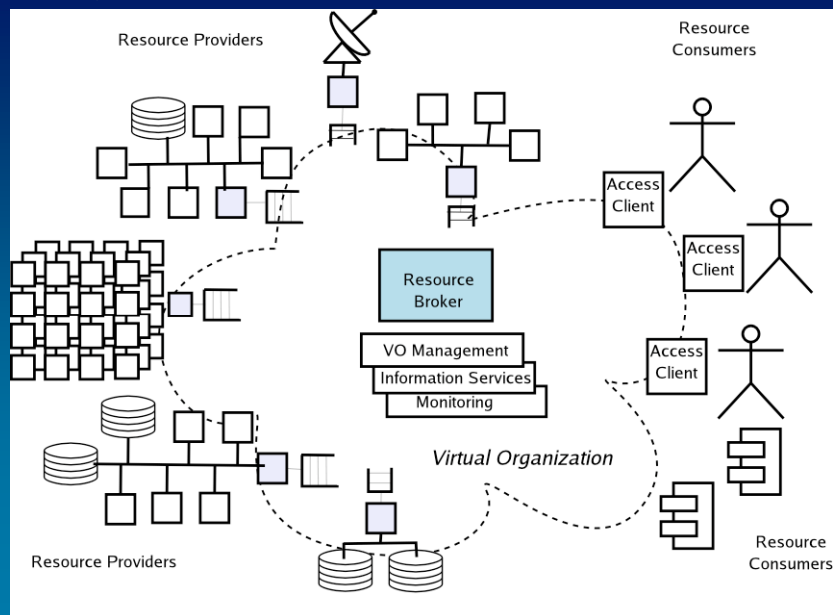
The Grid



- Middleware infrastructure that enables flexible, secure, coordinated resource sharing among dynamic collections of individuals and institutions (Foster, Kesselman, Tuecke).
- Enables communities ("Virtual Organizations") to share geographically distributed resources as they pursue common goals -- assuming the absence of...
 - Homogeneity
 - Central location
 - Central control
 - Existing trust relationships
- Enforces some level of resource virtualization.

<http://grid.ucy.ac.cy>

Grid operation



<http://grid.ucy.ac.cy>

A Grid (interactive) Application



Surgical Planning

- **Problem:** vascular diseases
- **Solution:** placement of a bypass by a surgeon
- Planning for intervention is based on 3D images obtained from MRI or CT scans.
- The attainable improvement in blood flow should determine which possibility is the best for a particular patient.
- A 3D arterial model is built on the basis of the images, and presented to the surgeon in an immersive environment.

*Source: Univ. of Amsterdam
CrossGrid project*

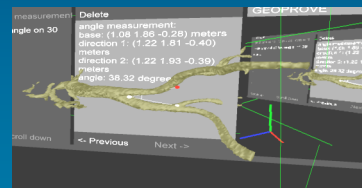


Observation



Stenosis

(narrowing of an artery)



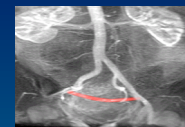
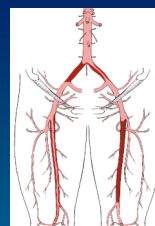
Viewing the arterial structure in an immersive 3D environment

<http://grid.ucy.ac.cy>

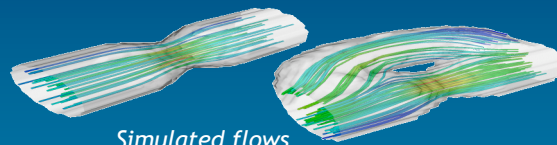
Surgical Planning



- **Goal:** Simulate vascular reconstruction
- **Method:**
 - Interactive Virtual Reality Environment to
 - View scanned data
 - Define proposed interventions
 - View simulation results
 - Advanced fluid code to simulate flows



Arterial structures from scans with proposed bypasses



Simulated flows

- Need Grid in interactive mode (the surgeon should not wait long...)
 - Access distributed computational resources for flow simulation and visualization, so get a high performance environment at low cost
 - Distribute simulations for different bypass configurations

<http://grid.ucy.ac.cy>

How about resource selection?



Status for Resource Broker CERN_lxn1188: Thu Oct 7 07:46:29 BST 2004

TERAGRID



Ask a Grid Information Service!



dn: GlueSubClusterUniqueID-cgce.ifca.org.es,
GlueClusterUniqueID-cgce.ifca.org.es,
Mail-Forwarding-Service=CG2_0_4

dn: GlueCEUniqueID-cgce.ifca.org.es:2119/jobmanager-pbs-short,
Mds-Vo-name-ifcapro,mds-vo-name-local,o-grid

GlueChunkKey: GlueClusterUniqueID=cgce.ifca.org.es

GlueHostApplicationSoftwareRunTimeEnvironment: CG2_0_4

GlueHostApplicationSoftwareRunTimeEnvironment: CG2_0_4

GlueCEUniqueID: cgce.ifca.org.es:2119/jobmanager-pbs-short

GlueCEInfoGatekeeperPort: 2119

GlueCEInfoHostName: cgce.ifca.org.es

GlueCEInfoLRMSType: pbs

GlueCEInfoLRMSVersion: OpenPBS_2.4

GlueCEInfoTotalCPUs: 20

GlueCEStateEstimatedResponseTime: 0

GlueCEStateFreeCPUs: 20

GlueCEStateRunningJobs: 0

GlueCEStateMemoryVirtualSize: 1144

GlueForeignKey: GlueClusterUniqueID=cgce.ifca.org.es

<http://grid.ucy.ac.cy>

Motivation and Focus



- How can we characterize the performance of Grid resources?
 - ⇒ Support more advanced criteria for choosing resources: **performance**, cost, functionality, reliability, robustness...
 - ⇒ Drive the design and configuration of Grid infrastructures.
 - ⇒ Open marketplaces based on performance negotiation.
 - ⇒ Developing models for performance prediction.
- *"We have no real idea how the Grid and Grid applications could be characterized from the point of view of performance" (APART Working Group on Automatic Performance Analysis, Rackeve Workshop, 11/2003)*

<http://grid.ucy.ac.cy>

GridBench



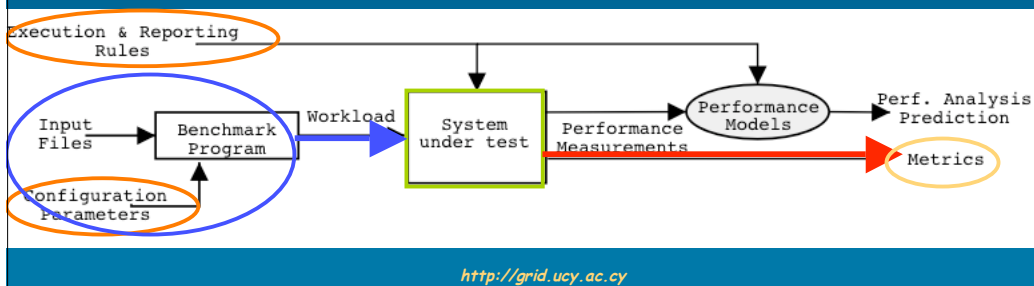
- The GridBench Framework:
 - A software tool (workbench) for **characterizing the performance** of Grids and Grid resources **quantitatively**, using **benchmarks**.
- GridBench Suite of Benchmarks:
 - A **hierarchical suite of benchmarks** deployed on a Grid testbed.
 - Geared towards high-performance and high-throughput computing needs.

<http://grid.ucy.ac.cy>

Key challenges for benchmarking Grids



- Measuring a moving and fuzzy target
- Trusting our measurements
- Describing performance (metrics)
- Managing cost
- In search of relevance



Measuring a moving target



- Grid jobs run upon heterogeneous resources which are:
 - Assembled dynamically and subject to change.
 - Described inaccurately or inadequately by GIS.
 - Often not properly operable, due to configuration errors, operator faults etc.
- Hence:
 - We need to capture the real set of resources we try to characterize.
 - Functionality benchmarking is equally important.

Trusting measurements



- A benchmarking job submitted to the Grid will run upon resources for which:
 - Exclusivity is **not guaranteed**.
 - Often, a benchmark **runs in co-location** with other jobs.
- Hence:
 - We **cannot trust** all of our measurements.
 - Need to **filter out polluted measurements**.

<http://grid.ucy.ac.cy>

Describing performance (metrics)



- Performance capacity of Grid infrastructure is defined by:
 - The performance of a **hierarchical collection** of measurable entities (CPUs, computers, clusters, collections of clusters..)
- Thus:
 - Small sets of metrics **not adequate** for Grids.
 - Definition, organization, storage, and interpretation requires **advanced, open data models**, amenable to post-processing (statistical, data mining, AI).
- ...what about **interpretation of metrics?**

<http://grid.ucy.ac.cy>

Managing cost



- Grid benchmarking seeks to capture a **complete** and **valid** "performance snapshot" of a Grid infrastructure. But:
 - Not all resources are available and running properly at a given time.
 - Jobs are susceptible to partial failure and degraded performance.
 - Hence, the derivation of a complete snapshot requires:
 - **Series** of experiments, measurements, analyses.
 - **Integration of metrics** from multiple sites and runs.
 - **Job submission** to the Grid takes **effort!**
- ⇒ **The cost of benchmarking increases substantially.**

<http://grid.ucy.ac.cy>

In search of relevance



- Relevant benchmarks are the ones **producing realistic workloads**, i.e., workloads representative of:
 - The **prevalent programming model**.
 - "Killer" Grid applications.
- However, the field:
 - Is still not mature; a prevalent programming model has yet to arise.
 - Production-quality Grid infrastructures are just beginning to emerge; no clear clues about typical workloads and applications.

<http://grid.ucy.ac.cy>

Elements of GridBench



- A **hierarchical model** for the Grid architecture.
- A **hierarchical suite of benchmarks** characterizing the performance of abstract-model elements [HPC, HTC, MPI].
- A **platform-independent language** (GBDL) for specifying the configuration and for representing the conditions and results of benchmarking experiments.
- **GridBench**: a virtual workbench for administering Grid benchmarks, archiving, publishing and browsing results.

<http://grid.ucy.ac.cy>

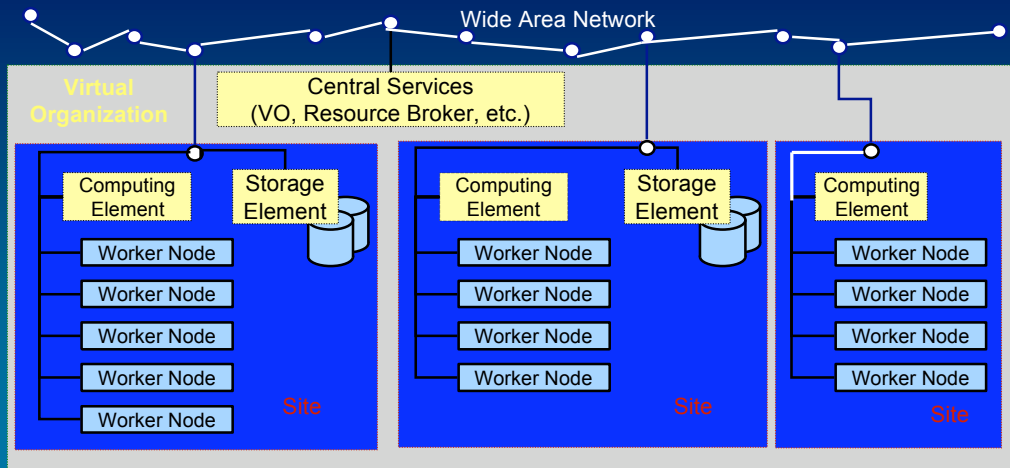
Outline



- **A simple reference architecture for the Grid.**
- The GridBench suite of benchmarks.
- GridBench description language.
- Filtering polluted measurements.
- The GridBench software.
- Using GridBench to characterize Grid sites.
- Conclusions and Future Work.

<http://grid.ucy.ac.cy>

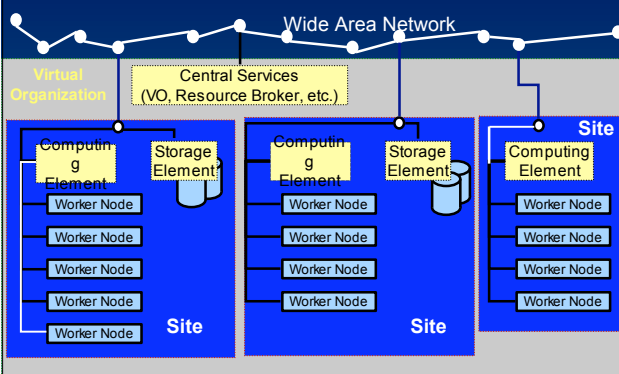
A simple reference architecture



- ✓ Inspired by the DataGrid/CrossGrid/LCG architecture (Globus 2-based)
- ✓ Represented by the GLUE Schema.
- ✓ Necessary for defining benchmarking-targets and interpreting metrics.

<http://grid.ucy.ac.cy>

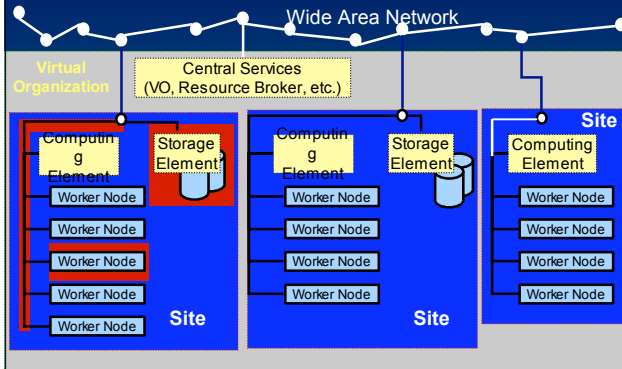
GridBench: a hierarchical approach



- Performance measurements at the different levels of the Grid architecture.

<http://grid.ucy.ac.cy>

A hierarchical approach

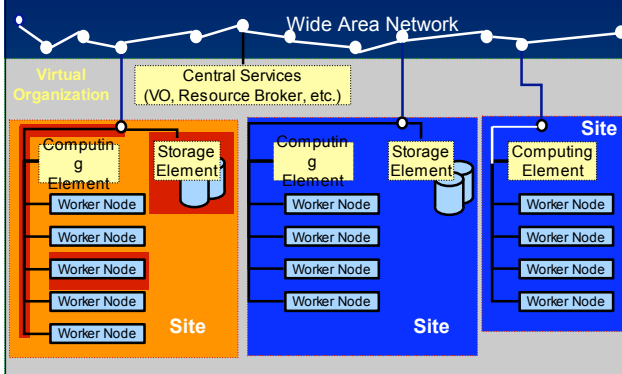


Individual Resources
(cluster nodes, mass storage)

- Performance measurements at the different levels of the Grid architecture.

<http://grid.ucy.ac.cy>

A hierarchical approach



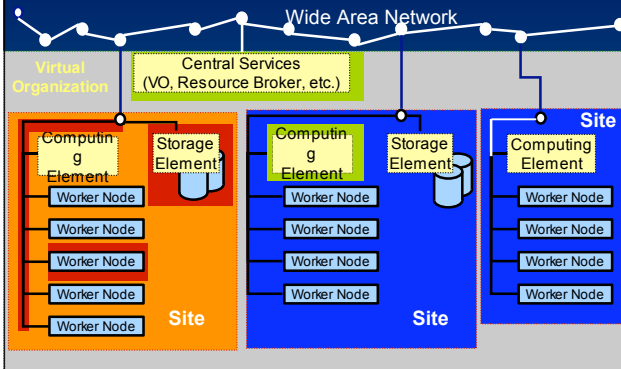
Individual Resources
(cluster nodes, mass storage)

Sites
(clusters, SMPs)

- Performance measurements at the different levels of the Grid architecture.

<http://grid.ucy.ac.cy>

A hierarchical approach



Individual Resources
(cluster nodes, mass storage)

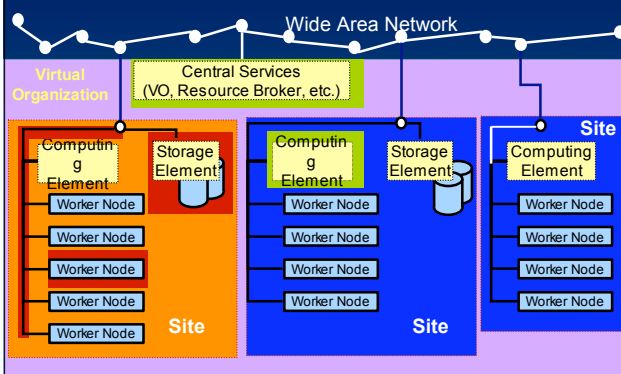
Sites
(clusters, SMPs)

Middleware
(middleware layer providing access to shared resources)

- Performance measurements at the different levels of the Grid architecture.

<http://grid.ucy.ac.cy>

A hierarchical approach



Individual Resources
(cluster nodes, mass storage)

Sites
(clusters, SMPs)

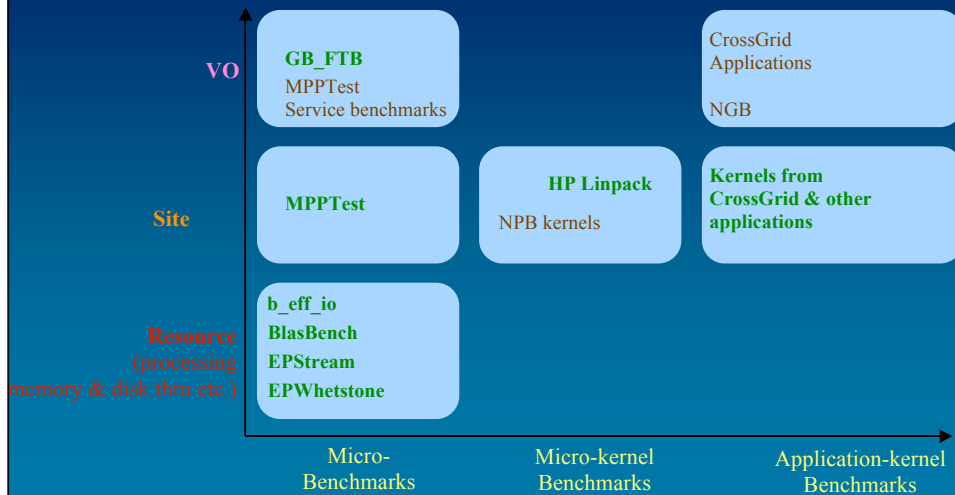
Middleware
(middleware layer providing access to shared resources)

Virtual Organization
(multiple sites, VO)

- Performance measurements at the different levels of the Grid architecture.

<http://grid.ucy.ac.cy>

The GridBench suite of benchmarks



<http://grid.ucy.ac.cy>

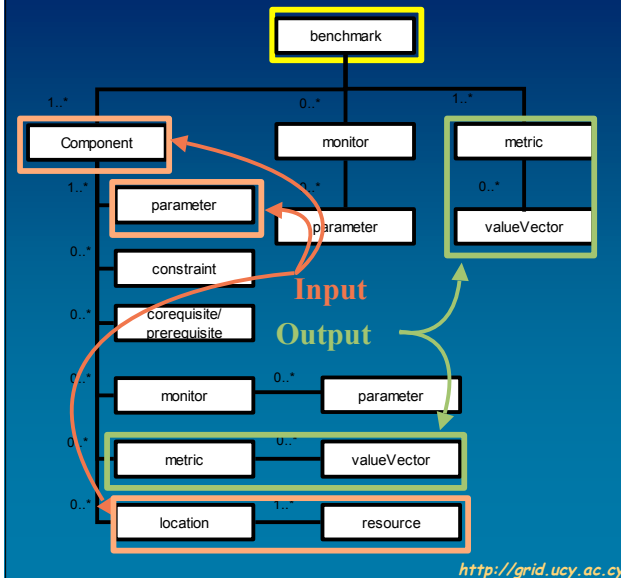
The GridBench Description Language



- An XML-schema specification representing benchmarking metadata.
- Allows easy transformation to different job description formats.
- Benchmark definition co-exists with benchmarking configuration and results.
- GBDL elements include metadata about benchmark components:
 - Parameters
 - Location
 - Co-requisites
 - Prerequisites
 - Constraints
 - Metrics
 - Metrics
 - Archive
 - Monitoring

<http://grid.ucy.ac.cy>

GBDL structure



Metrics representation



- **Benchmark-level metrics vs. component-level metrics:**
 - e.g. benchmark completion time vs. component completion time.

- **Single-value metrics**

```
<metric name="epwhetstone_IPS" type="value">
  <vector unit="MIPS">53</vector>
</metric>
```

- **Multi-value metrics**

```
<metric name="xfer-rate" type="list">
  <vector unit="bps">
    2039430,2083930,1909830,2184750,...
  </vector>
  <vector unit="second" toffset="1055327287">
    0,10,20,30,...
  </vector>
</metric>
```

<http://grid.ucy.ac.cy>

GBDL example



```
<benchmark date="031010121503" name="epwhetstone">
  <component id="A" name="epwhetstone" type="mpi">
    <location type="multi">
      <resource cpucount="2" name="apelatis.grid.ucy.ac.cy"/>
      <resource cpucount="2" name="ce010.fzk.de"/>
    </location>
    <parameter name="execpath" type="attribute">/opt/cg/gb/bin</parameter>
    <parameter name="stage_executable" type="attribute">manual</parameter>
    <parameter name="executable" type="attribute">epwhet</parameter>
    <parameter name="nloops" type="value">50000</parameter>
    <metric name="completion" unit="s">28.23</metric>
    <metric name="whetstone_mips" type="value">
      <vector name="hostname">wn001.grid.ucy.ac.cy
        wn001.grid.ucy.ac.cy
        n16.fzk.de
        n15.fzk.de
      </vector>
      <vector name="mips">684.023 671.813 784.234 777.564 </vector>
    </metric>
  </component> </benchmark> http://grid.ucy.ac.cy
```

Outline



- A simple reference architecture for the Grid.
- The GridBench suite of benchmarks.
- GridBench description language.
- **Filtering polluted measurements.**
- The GridBench software.
- Using GridBench to characterize Grid sites.
- Conclusions and Future Work.

Identifying polluted measurements



- Often a benchmark will run on a resource, in co-location with other jobs:
 - "Fellow passengers:" co-allocated by the Resource Broker.
 - "Free-riders:" unauthorized users, unknown to the VO.
 - "Runaways:" O/S processes, zombies, etc.
- These jobs may "pollute" our **measurements** and affect seriously the characterization accuracy.
- Their effect on metric accuracy can be:
 - Identified through **monitoring**.
 - Reduced through **access control** (for fellow passengers) and **remote healing** (for runaways).

<http://grid.ucy.ac.cy>

Filtering polluted measurements

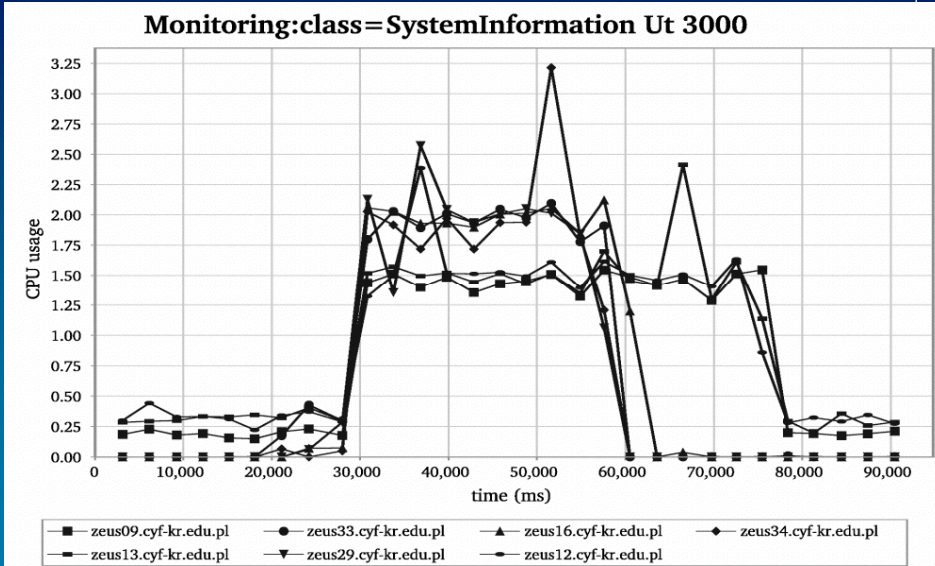


- GridBench retrieves monitoring information from **existing Grid monitoring services**, for a time window encapsulating a benchmarking experiment.
- The specification of the monitoring service and the data to be retrieved from it, is **included in the GBDL document** describing a specific experiment.

```
<component name="data-transfer" ID="xfer01">...</component>
<monitor type="RGMA" source="ccwp71.in2p3.fr:3306"
  query="select * from NetworkTCPThroughput
  where NMIdSource='adc0003.cern.ch'
  and NMIdDestination='ccwp7.in2p3.fr'
  <parameter name="begin">comp-begin="xfer01"</parameter>
  <parameter name="end">comp-end="xfer01"</parameter>
</monitor>
```
- Retrieval from monitoring services is conducted via **monitoring-client plug-ins** called by GridBench.

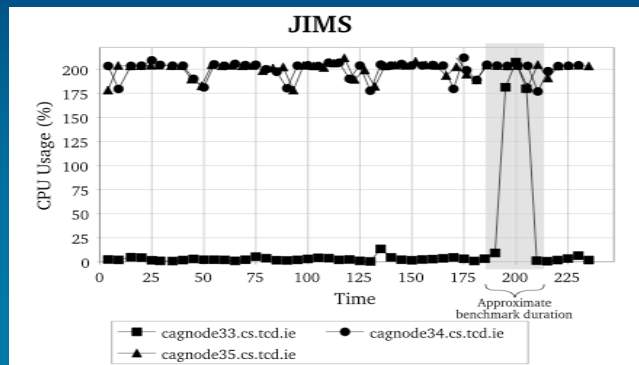
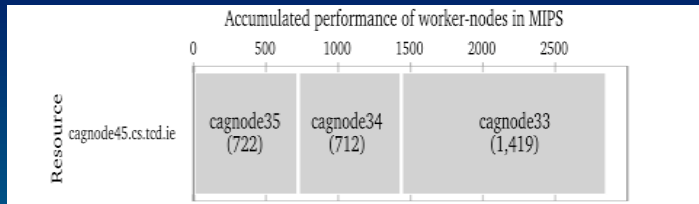
<http://grid.ucy.ac.cy>

Monitoring a benchmark



<http://grid.ucy.ac.cy>

Filtering polluted measurements



<http://grid.ucy.ac.cy>

Outline



- A simple reference architecture for the Grid.
- The GridBench suite of benchmarks.
- GridBench description language.
- Filtering polluted measurements.
- **The GridBench software.**
- Using GridBench to characterize Grid sites.
- Conclusions and Future Work.

<http://grid.ucy.ac.cy>

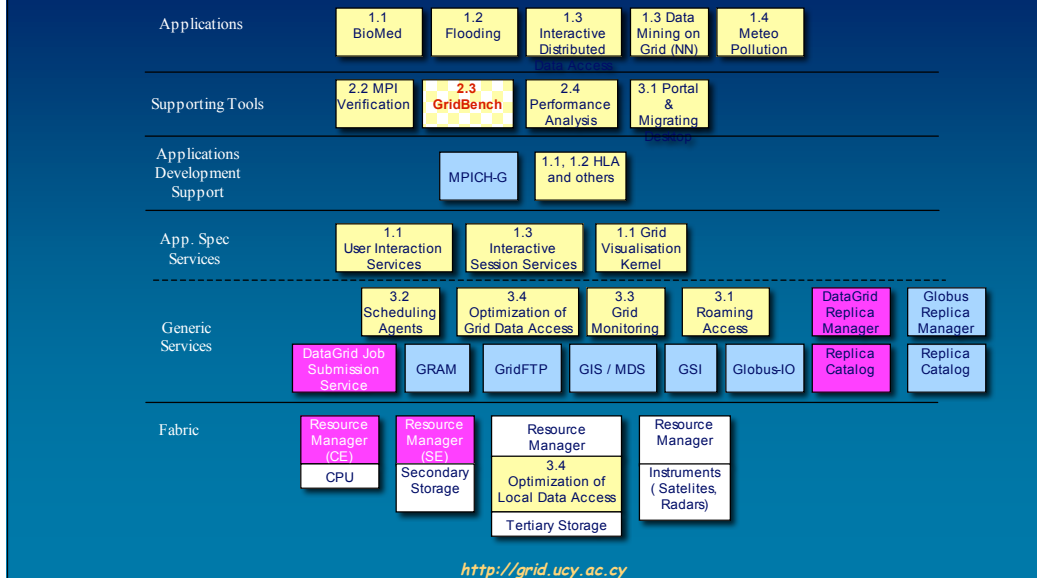
GridBench: requirements & functionality



- Supports:
 - **Specification** and **execution** benchmarks on a Grid.
 - **Collection** and **archival** of results.
- Archives together:
 - Benchmark specifications and measurements for publication and further analysis.
 - Monitoring information to help with result interpretation.
- Supports:
 - **Retrieval** and **graphical representation** of metrics.

<http://grid.ucy.ac.cy>

GridBench Context: CrossGrid architecture



GridBench front-end

- **GBDL Translator**
 - XML benchmark description (GBDL) to "job description language"
 - Supports JDL (EDG, Condor) and RSL (Globus).
- **Benchmark Definition UI**
 - GUI for defining and executing benchmarks.
- **Benchmark Browser**
 - GUI for browsing and analyzing results.
- **Information Provider**
 - Publishes results to Metacomputing Directory Service (MDS).

<http://grid.ucy.ac.cy>

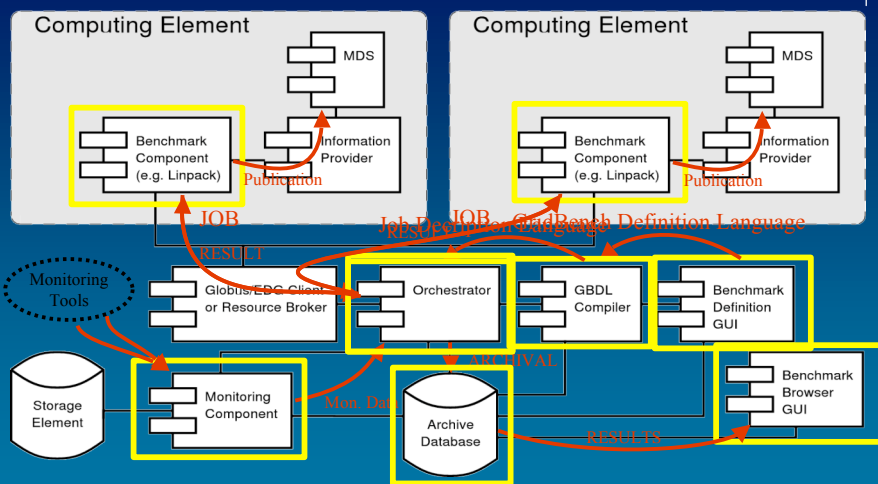
GridBench back-end



- **Orchestrator**
 - Manager of execution and result collection.
 - Web Service (it must submit the job and wait for the output).
 - Co-located with work-load management client of CrossGrid.
- **Archiver & Database**
 - Stores benchmark definition, results, and monitoring.
 - Web Service.
 - Requires network connectivity to the host running the apache database.
- **Benchmark Components**
 - Benchmark executable code.
- **Monitoring Component**
 - Collects information using e.g. R-GMA or OCM-G.

<http://grid.ucy.ac.cy>

GridBench at work



Building on the existing GBDL to JDL/RSL Translator
(Native Machine execution) collection of results

<http://grid.ucy.ac.cy>

Outline



- A simple reference architecture for the Grid.
- The GridBench suite of benchmarks.
- GridBench description language.
- Filtering polluted measurements.
- The GridBench software.
- **Using GridBench to characterize Grid sites.**
- Conclusions and Future Work.

<http://grid.ucy.ac.cy>

Characterizing Grid sites



- To characterize a site (cluster) and its resources, we are focusing on a small set of metrics, describing:
 - **CPUs performance**: OPS, FLOPS, INTOPS
 - **Cache performance**: MB/s
 - **Main Memory performance**: MB/s
 - **Local Interconnect**: latency, bandwidth
 - **I/O performance**: effective I/O bandwidth
 - **Maximum Available Memory**: available for dynamic allocation
- We are also looking at functionality aspects:
 - Are **local queuing systems** operating properly?
 - Is **MPI installed properly** and fully operable?
 - Is **SSH** working properly?

<http://grid.ucy.ac.cy>

Metrics and Benchmarks



Factor	Metric	Delivered By
CPU	Operations per second (mixture of floating point and integer arithmetic)	EPWhetstone
CPU	Floating-Point operations per second	EPFlops
CPU	Integer operations per second	EPDhrystone
memory	sustainable memory bandwidth in MB/s (copy,add,multiply,triad)	EPStream
memory	Available physical memory in MB	EPMemsize
cache	memory bandwidth using different memory sizes in MB/s	CacheBench
Interconnect	latency, bandwidth and bisection bandwidth	MPPTest
I/O	Effective I/O bandwidth	b_eff_io

- We are using widely known micro-benchmarks to derive the required metrics.
- We run those mbenchs in parallel on all measured resources.

<http://grid.ucy.ac.cy>

The screenshot displays the GridBench application interface, which is used for managing and monitoring benchmarks across a grid of resources. The interface is divided into several panes:

- Template List:** Shows a list of available benchmarks such as cachebench, epstream, epwhetstone, gbrtb, hpl, mppctest, vertig, epdhrystone, b_eff_io, epflops, bstream, and memsize.
- Resource List:** Lists available resources including cpge.fcra.org.es, ce0102zk.de, cluster.zsl.sav.sk, zeus-2.4.yf-hz.edupl, grid01.physics.aunz.nz, and asecegridsub.es.
- Result Matrix:** A table showing benchmark results for various resources. The columns include hpl, mppctest, vertig, epdhrystone, b_eff_io, epflops, bstream, and memsize. The rows correspond to the resources listed in the Resource List.
- EPStream 3:** A chart showing the accumulated Triad performance of worker-nodes. The x-axis represents the number of nodes (0 to 5,000), and the y-axis represents performance (0 to 5,000).
- MPPTest 4:** A chart showing the time taken to solve a problem versus the number of nodes. The x-axis represents the number of nodes (0 to 17,500), and the y-axis represents time (0 to 1,500).
- New epdhrystone Benchmark:** A dialog box for creating a new benchmark. It includes fields for definition, globus, and edg, and a text area for the benchmark definition.

GridBench

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Template List

- bonnie
- cachebench
- epstream
- epwhetstone
- gdtb
- hpl
- mpptest
- verilog
- epwhetstone
- b_eff_io
- epfflops
- bstream
- memsize
- bstream1_1

Resource List

cgce@ca.orges	19/20 cpus 0/0 wns
ce010zkate	16/16 cpus 0/0 wns
cluster@sa.vsk	16/16 cpus 0/0 wns
xgrid@medupl	8/8 cpus 0/0 wns
gtbcg01@ca.unicar.es	8/8 cpus 0/0 wns
cms@uw.edu.pl	4/4 cpus 0/0 wns
ce02@ip.pt	3/3 cpus 0/0 wns
lok01@icarves	3/3 cpus 0/0 wns
ce.grid.es.gaes	2/4 cpus 0/0 wns
grid01.physics.auth.gr	2/2 cpus 0/0 wns
ce001.grid.ucy.ac.cy	1/1 cpus 0/0 wns
grid14.physics.auth.gr	1/1 cpus 0/0 wns
zeus-24cyf-kr.edu.pl	1/18 cpus 0/0 wns
cgnode00@luoa.gr	1/1 cpus 0/0 wns

Benchmark List

Resource List

<http://grid.ucy.ac.cy>

GridBench

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Resource List

cgce@ca.orges	19/20 cpus 0/0 wns
ce010zkate	16/16 cpus 0/0 wns
cluster@sa.vsk	16/16 cpus 0/0 wns
xgrid@medupl	8/8 cpus 0/0 wns
gtbcg01@ca.unicar.es	8/8 cpus 0/0 wns
cms@uw.edu.pl	4/4 cpus 0/0 wns
ce02@ip.pt	3/3 cpus 0/0 wns
lok01@icarves	3/3 cpus 0/0 wns
ce.grid.es.gaes	2/4 cpus 0/0 wns
grid01.physics.auth.gr	2/2 cpus 0/0 wns
ce001.grid.ucy.ac.cy	1/1 cpus 0/0 wns
grid14.physics.auth.gr	1/1 cpus 0/0 wns
zeus-24cyf-kr.edu.pl	1/18 cpus 0/0 wns
cgnode00@luoa.gr	1/1 cpus 0/0 wns

Source

- Update (PBS)
- Set Selected cpus...
- Last Output...
- Tests...
- JMS user

Retrieve more information from resource queues

<http://grid.ucy.ac.cy>

GridBench interface showing a list of resources. The 'Resource List' pane displays the following data:

Resource Name	CPU	WNS
cgce.jica.org.es	20/20 cpus	10/10 wns
ce01.02k.de	16/16 cpus	4/4 wns
cluster.aia.savk	16/16 cpus	16/16 wns
xgrid.cmedu.pl	8/8 cpus	4/4 wns
gibcg01.jica.mexico	8/8 cpus	4/4 wns
cms.fuw.edu.pl	4/4 cpus	0/0 wns
ce02.lip.pt	3/3 cpus	2/2 wns
lel01.jica.ru.es	3/3 cpus	3/3 wns
ce.grid.ces.gaes	4/4 cpus	2/2 wns
grid01.physics.auth.gr	2/2 cpus	1/1 wns
ce001.grid.ucy.ac.cy	1/1 cpus	1/1 wns
grid1-4.physics.auth.gr	1/1 cpus	1/1 wns
zeus-24cyf-hr.edu.pl	2/18 cpus	1/9 wns
cpnode06.liaoa.gr	1/1 cpus	1/1 wns
cedar.zros.grid.man.poznan.pl	0/2 cpus	0/1 wns

A red box highlights the 'WNS' column with the text: "Sites are reachable, queue status and WN numbers updated".

<http://grid.ucy.ac.cy>

GridBench interface showing a list of resources. A context menu is open over the 'grid1-4.physics.auth.gr' resource, with the following options:

- Source
 - Update (PBS)
 - Set Selected cpus...
 - Last Output...
- Tests...
 - MPI Test
 - MPI Test (G2)
 - SSH Test

A red box highlights the 'Tests...' menu item with the text: "Select some sites and perform an MPICH-G2 test".

<http://grid.ucy.ac.cy>

GridBench interface showing a Resource List and a pop-up window for MPL_G2 participating WNS.

Resource List:

Resource Name	CPU	WNS
ce010zkade	16/16	4/4
clusterjlaavsk	16/16	16/16
xgridicmedupl	8/8	4/4
gfbcg01jicaunicas	8/8	4/4
cms2uwedupl	0/0	0/0
ce02jipgt		
lok01jicarves		
ce.grid.ces.gaes		
grid01.physics.auth.gr		
ce001.grid.ucy.ac.cy		
grid1-4.physics.auth.gr		
zeus-24cyf-kz.edupl		
cgnode08.dlunaogr	1/1	1/1
cedar.res.grid.man.poznan.pl	0/2	0/1

MPL_G2 participating WNS:

- grid03.grid.ces.gaes
- grid03.physics.auth.gr
- wn001.grid.ucy.ac.cy

<http://grid.ucy.ac.cy>

GridBench interface showing a Resource List with a context menu open over 'ce010zkade'.

Resource List:

Resource Name	CPU	WNS
cgcejicaorges	20/20	10/10
ce010zkade	16/16	4/4
clusterjlaavsk		
xgridicmedupl		
gfbcg01jicaunicas		
cms2uwedupl	4/4	0/0
ce02jipgt	3/3	2/2
lok01jicarves	3/3	3/3
ce.grid.ces.gaes	4/4	2/2
grid01.physics.auth.gr	2/2	1/1
ce001.grid.ucy.ac.cy	1/1	1/1
grid1-4.physics.auth.gr	1/1	1/1
zeus-24cyf-kz.edupl	2/18	1/9
cgnode08.dlunaogr	1/1	1/1

Context Menu for ce010zkade:

- Source
 - plus 154 WNS
- Update IPBS
- Select cpus...
- Last Output...
- Tests...
 - cpus 4/4 WNS
 - JMS user

<http://grid.ucy.ac.cy>

GridBench interface showing the process of dragging a benchmark from the Template List to the Resource List. A red arrow points from the 'epstream' benchmark in the Template List to the 'cluster1.savsk' resource in the Resource List. A text box with a red border contains the instruction: "Drag a benchmark onto a resource".

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Template List	Resource List
bonnie	ce010fzk.de
cachebench	cluster1.savsk
epstream	cluster1.savsk
epwhetstone	cluster1.savsk
gbfb	cluster1.savsk
hpl	cluster1.savsk
mpptest	cluster1.savsk
verilog	cluster1.savsk
epdnytone	cluster1.savsk
b_eff_io	cluster1.savsk
epflops	cluster1.savsk
btstream	cluster1.savsk
memsize	cluster1.savsk
btstream1_1	cluster1.savsk

<http://grid.ucy.ac.cy>

GridBench interface showing the configuration panel for a new benchmark. A red arrow points to the 'New epstream Benchmark' dialog box. A text box with a red border contains the instruction: "Configuration panel opens: Set parameters.".

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Template List	Resource List
bonnie	ce010fzk.de
cachebench	cluster1.savsk
epstream	cluster1.savsk
epwhetstone	cluster1.savsk
gbfb	cluster1.savsk
hpl	cluster1.savsk
mpptest	cluster1.savsk
verilog	cluster1.savsk
epdnytone	cluster1.savsk
b_eff_io	cluster1.savsk
epflops	cluster1.savsk
btstream	cluster1.savsk
memsize	cluster1.savsk
btstream1_1	cluster1.savsk

New epstream Benchmark

Name: epstream Date: Sat May 15 02:39:17 EEST 2004 Type: mpi

Parameters	Resources	Monitoring
executable	epstream	value
execpath	/opt/ics/gridbench/bin	value
stage_executable	manual	value

new parameter

Description:

Submit using: globus 0% Advanced submit

<http://grid.ucy.ac.cy>

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Template List

- bonnie
- cachebench
- epstream
- epwhetstone
- gfbfb
- hpl
- mpptest
- verilog
- epdwhetstone
- b_eff_io
- epflops
- btstream
- memsize
- btstream_1

Resource List

Resource	# CPUs	# WNs	Job Manager
ce0102z.de	16/16	4/4	wms
cluster1.sav.sk	16/16	16/16	wms
xgrid1.med.upl	8/8	4/4	wms
gtbcg01.fca.mkanes	8/8	4/4	wms

New epstream Benchmark

Name: epstream Date: Sat May 15 02:39:17 EEST 2004 Type: mpi

Resource	# CPUs	# WNs	Job Manager
cluster1.sav.sk	16	16	jobmanager-pbs...

new resource

Description:

Submit using: globus 0% Advanced submit

Tune CPU /WN numbers

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Template List

- bonnie
- cachebench
- epstream
- epwhetstone
- gfbfb
- hpl
- mpptest
- verilog
- epdwhetstone
- b_eff_io
- epflops
- btstream
- memsize
- btstream_1

Resource List

Resource	# CPUs	# WNs	Job Manager
ce0102z.de	16/16	4/4	wms
cluster1.sav.sk	16/16	16/16	wms
xgrid1.med.upl	8/8	4/4	wms
gtbcg01.fca.mkanes	8/8	4/4	wms

New epstream Benchmark

Name: epstream Date: Sat May 15 02:39:17 EEST 2004 Type: mpi

Monitoring: jims | Monitoring.class=SystemInformation AverageIdle 2000

SVG

Description:

Submit using: globus 0% Advanced submit

Setup Monitoring (JIMS)

The screenshot shows the GridBench application window. On the left is the 'Template List' with various benchmarks like bonnie, cachebench, epstream, etc. The main area is divided into 'Resource List' and 'New epstream Benchmark'. The 'Resource List' shows four resources: ce0102z.de, cluster.la.svk, xgridic.medupl, and gtcg01.fcaimkanes. The 'New epstream Benchmark' window is open, showing the GBDL definition. A red box labeled 'Preview GBDL' points to the GBDL tab, and another red box labeled 'Advanced' points to the 'Advanced' button in the submission bar.

Submit using: globus 0% **Advanced** submit

<http://grid.ucy.ac.cy>

This screenshot is similar to the one above but shows the 'New epstream Benchmark' window with the 'RSL' tab selected. The RSL definition is visible, showing resource manager contact information and job parameters. A red box labeled 'Preview RSL (GLOBUS)' points to the RSL tab. The 'Advanced' button in the submission bar is also visible.

Submit using: globus 0% **Advanced** submit

<http://grid.ucy.ac.cy>

GridBench interface showing the 'New epstream Benchmark' dialog. The dialog displays the JDL code for the benchmark:

```

definition | GBDL | globus | edg
#Automatically Generated by GridBench
#####
StdOutput = "std out";
StdError = "std err";
InputSandbox = ("jopt/cg/gridbench/bin/epstream");
Requirements = other.GlueCEUniqueID == "cluster.ui.sav.sk:2119/jobmanager-pbs-workq";
Executable = "epstream";
JobType = "mpioch";
NodeNumber = 16;
OutputSandbox = ("std out","std err");
  
```

A red box highlights the text "Preview JDL (CG,EDG,LCG)" with an arrow pointing to the JDL code.

<http://grid.ucy.ac.cy>

GridBench interface showing the 'New epstream Benchmark' dialog. The 'Parameters' tab is selected, showing the following table:

name	value	type	def
executable	epstream	value	def
execspath	/opt/cg/gridbench/bin	value	def
stage_executable	manual	value	def

The progress bar shows 21% completion. A red box highlights the text "Benchmark running" with an arrow pointing to the progress bar. Another red box highlights the "SUBMIT" button.

<http://grid.ucy.ac.cy>



The screenshot shows the GridBench application window. The 'Resource List' pane on the right contains the following entries:

Resource Name	Cpus	Wns
ce0102k.de	16/16	4/4
cluster.alavak	16/16	16/16
xgridic.med.kupl	8/8	4/4
gtbcg01.fca.mkanes	8/8	4/4

The 'New epstream Benchmark' dialog is open, showing the 'Results' tab with the following data:

name	value	type
completion	2.215129	value
memory_bandwidth	<vector>	value

Buttons for 'RESULTS' and 'Done.' are highlighted with red boxes and arrows.

<http://grid.ucy.ac.cy>



The screenshot shows the GridBench application window with the 'New epstream Benchmark' dialog open. The 'Parameters' tab is active, showing the following table:

hostname	copy	scale	add	triad
n1	1600.000000	533.333333	1600.000000	685.714286
n3	1600.000000	533.333333	1600.000000	685.714286
n2	1600.000000	533.333333	1600.000000	685.714286
n4	1600.000000	533.333333	1600.000000	685.714286
n12	1600.000000	533.333333	1600.000000	685.714286
n6	1600.000000	533.333333	1600.000000	685.714286

The 'Add to chart' button is highlighted with a red box and arrow, with a callout box containing the text: 'Add metric to new chart (chart type determined based on metric)'. The 'Done.' button is also highlighted with a red box and arrow.

<http://grid.ucy.ac.cy>

New epstream Benchmark

Name: epstream Date: Wed Oct 13 19:01:01 EEST 2004 Type: mpi

Parameters	Resources	Monitoring	Results
name	value	value	type
completion	2 2 15 12 9		value
memory_bandwidth			value

hostname	copy	scale	add	triad
n1	1600.000000	533.333333	1600.000000	685.714286
n3	1600.000000	533.333333	1600.000000	685.714286
n2	1600.000000	533.333333	1600.000000	685.714286
n4	1600.000000	533.333333	1600.000000	685.714286
n12	1600.000000	533.333333	1600.000000	685.714286
n6	1600.000000	533.333333	1600.000000	685.714286

Chart with selected metric

EPStream 1

Accumulated Triad performance of worker-nodes in MB/s

Resource: cluster.ui.sav.sk

<http://grid.ucy.ac.cy>

1. Browse database

cachebench	epstream	epstream	gputo	mpi	mpirun	verilog	epdhystone	b_eff_io
cluster.ui.sav.sk	0	6	1	3	0	8	0	1
ce010.tzk.de	0	0	1	8	0	0	0	0
xgrid.icm.edu.pl	0	4	2	2	0	4	0	1
cgce.ftca.org.es	0	8	1	2	0	0	3	0
grid01.physics.auth.gr	0	0	0	1	0	0	0	0
gtbcg01.ftca.unican.es	0	0	0	0	0	0	0	0
gtbcg01.ftca.unican.es	0	0	0	0	0	0	0	0
ce02.lip.pt	0	0	0	0	0	0	0	0
ce001.grid.ucy.ac.cy	0	0	0	0	0	0	0	0
ce101.grid.ucy.ac.cy	0	0	0	0	0	0	0	0
cgnode45.cs.tcd.ie	0	0	0	2	0	0	0	0

2. Select resource and benchmark

3. Add metric to existing chart (below)

EPStream 1

Accumulated Triad performance of worker-nodes in MB/s

Resource: cluster.ui.sav.sk

<http://grid.ucy.ac.cy>



GridBench

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Templates Resources Results

Result Matrix

	cachebench	epstream	epwhetstone	gbftb	hpl	mpprest	vertiq	epdhystone	b_eff_io
cluster.ui.sav.sk	0 6	1	3	0	8	7	0	1	2
ce010.fzk.de	0 0	1	8	0	0	0	0	0	0
xgrid.icm.edu.pl	0 4	2	2	0	0	4	0	1	1
caice.ifra.org.es	0 8	1	2	0	0	3	0	1	2
grid01.physics.auth.gr	0 0	0	1	0	0	0	0	0	0
gtbcg01.ifca.uvic.ca	0 0	0	0	0	0	0	0	0	0
ce02.lip.pt	0 0	0	0	0	0	2	0	0	0
ce001.grid.ucy.ac.cy	0 0	0	0	0	0	0	0	0	0
ce101.grid.ucy.ac.cy	0 0	0	0	0	0	2	0	0	0
cagnode45.cs.tcd.ie	0 0	0	2	0	0	0	0	0	0

ID Date CPU# execpath executable stage_exe...

79	9/21/04 ...	4	/opt/cg/p...	epstream	manual
78	9/21/04 ...	8	/opt/cg/p...	epstream	manual

Update

ce001.grid.ucy.ac.cy 1/1 cpus 1/1 wns

grid1-4.physics.auth.gr

zeus 2

cagnode

cedar 2

EPStream 1

Accumulated "Trips" performance of worker-nodes in MB/s

Metric added

New epstream...

<http://grid.ucy.ac.cy>



GridBench

Hint: Drag a benchmark 'Template' onto a 'Resource'!

Template List Resources Results

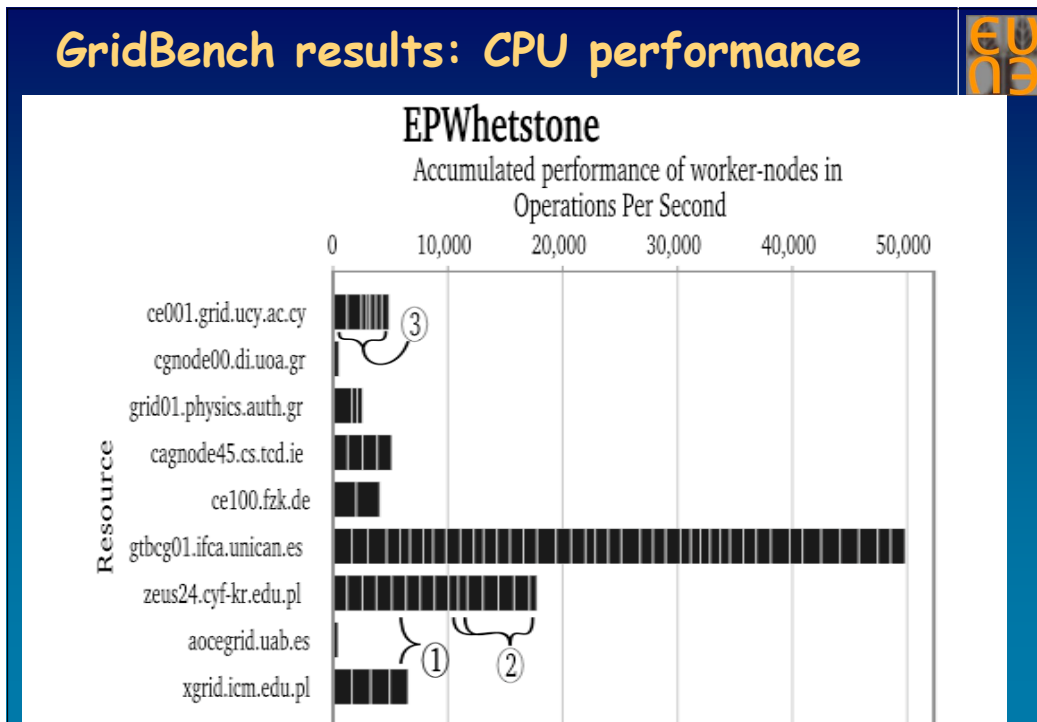
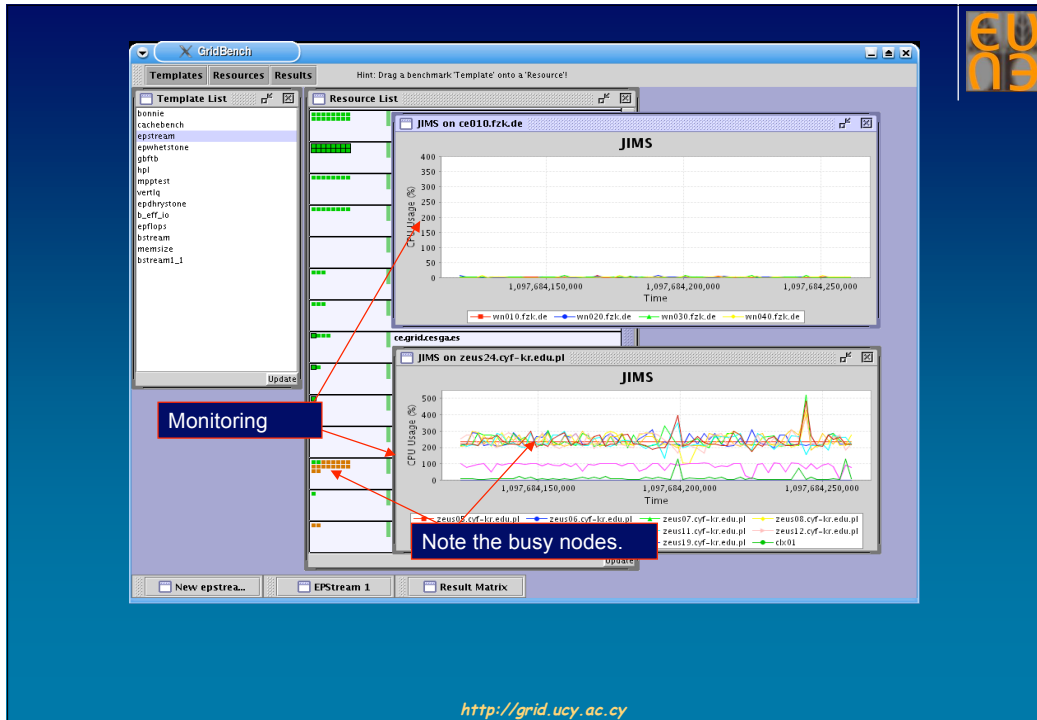
Resource List

ce010.fzk.de	16/16 cpus 4/4 wns
cluster.ui.sav.sk	16/16 cpus 16/16 wns
xgrid.icm.edu.pl	8/8 cpus 4/4 wns
gtbcg01.ifca.uvic.ca	8/8 cpus 4/4 wns
cms2u.wedupl	0/0 cpus 0/0 wns
ce02.lip.pt	3/3 cpus 2/2 wns
lkh01.ifcar.vuz	3/3 cpus 3/3 wns
ce.grid.es.gaes	4/4 cpus 2/2 wns
grid01.physics.auth.gr	2/2 cpus 1/1 wns
ce001.grid.ucy.ac.cy	1/1 cpus 1/1 wns
grid1-4.physics.auth.gr	1/1 cpus 1/1 wns
zeus 24.cyl-kr.edupl	1/1 cpus 1/1 wns
cagnode06.liaoa.gr	
cedar.zros.grid.man.poznan.pl	

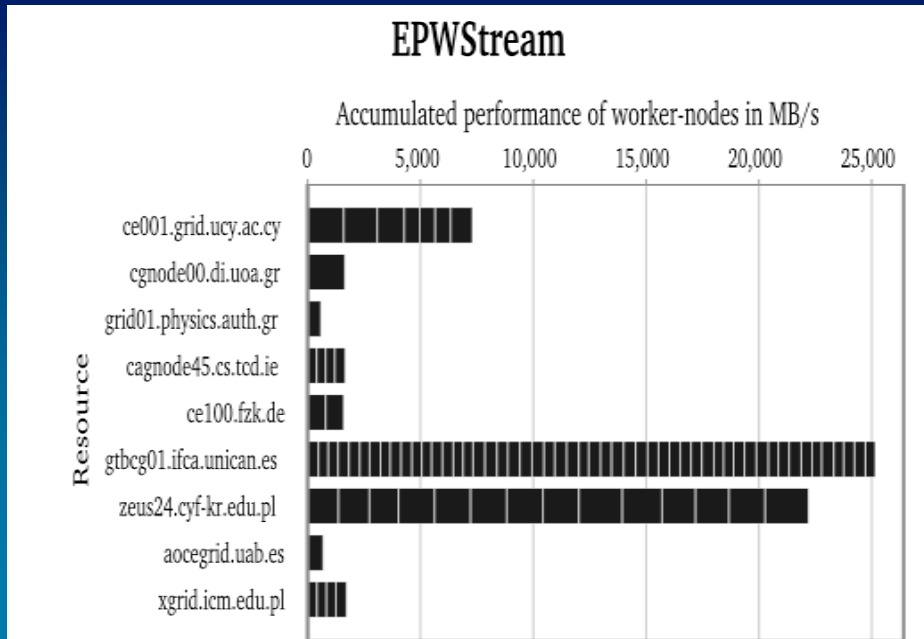
Select some resources for monitoring

New epstream... EPStream 1 Result Matrix

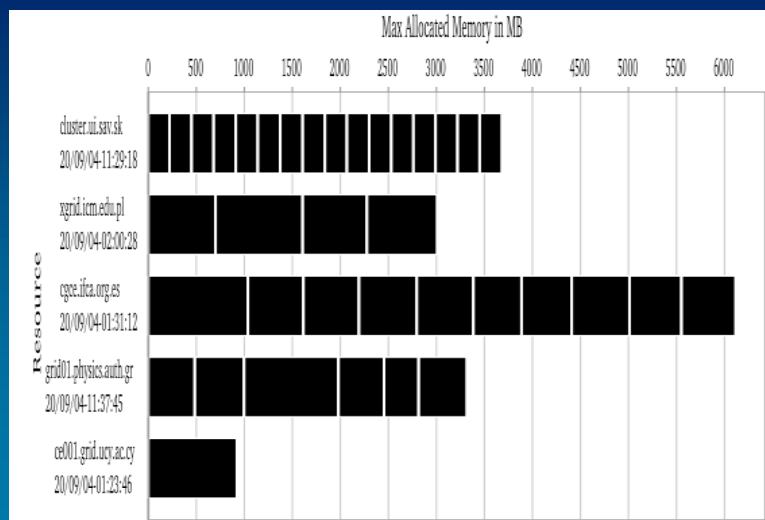
<http://grid.ucy.ac.cy>



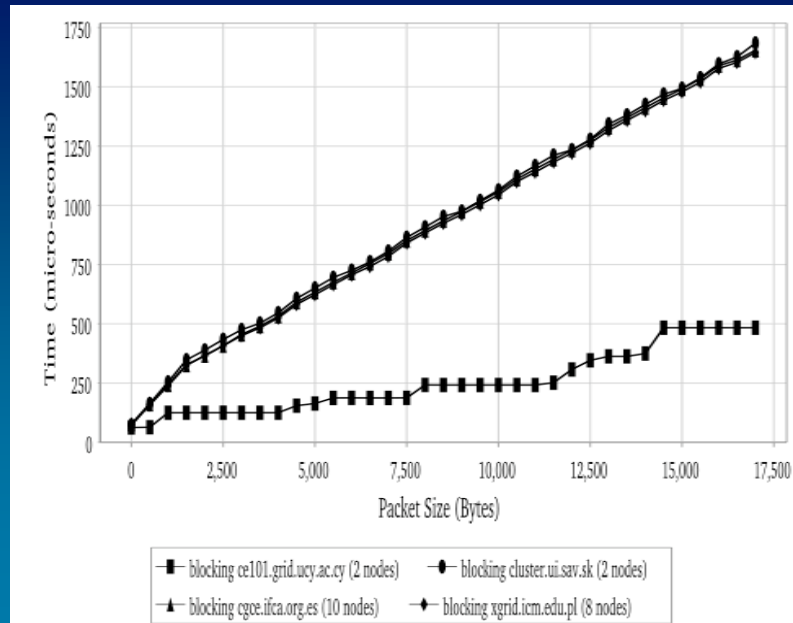
GridBench results: Memory performance



GridBench results: Available Memory

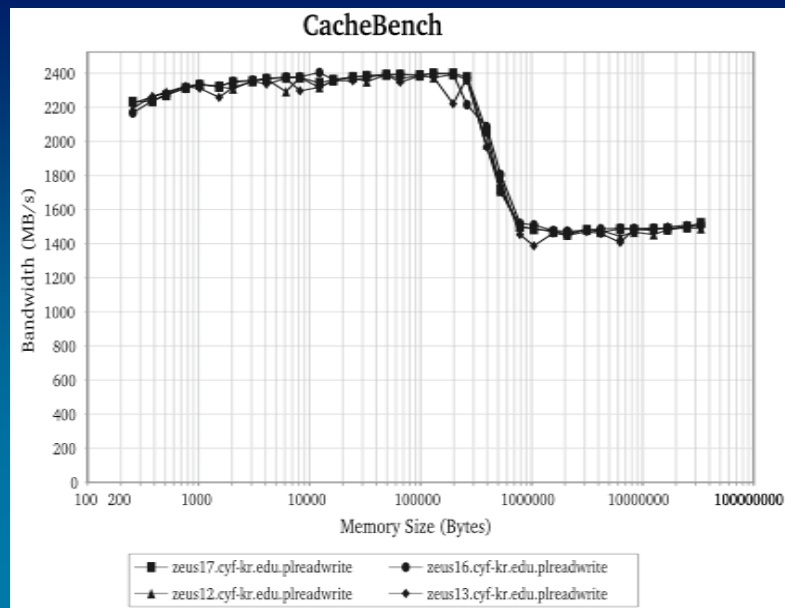


GridBench results: Communication



<http://grid.ucy.ac.cy>

GridBench results: cache

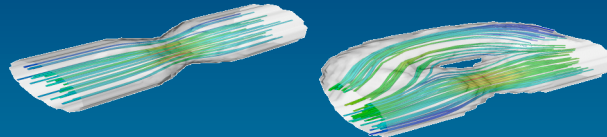


<http://grid.ucy.ac.cy>

Resource Selection with GridBench



- Where should I run my fluid-flow code for Surgical Planning?



Simulated flows

<http://grid.ucy.ac.cy>

The screenshot shows the GridBench application window with three tabs: Templates, Resources, and Results. The Resources tab is active, displaying a list of resources. A context menu is open over the 'grbcg011' resource, showing options: 'Update IPBS...', 'Set Selected cpus...', 'Last Output...', 'Test...', and 'JMS user'. A blue callout box points to the 'Set Selected cpus...' option with the text: 'Select some CPUs, or leave the default (default is "all free CPUs")'. The resource list includes columns for resource name, status, and available resources (cpus, wms).

Resource Name	Status	CPUs	WMS
cgce2fcaorges	Available	19/20	0/0
ce010zkade	Available	6/6	0/0
cluster2savsk	Available	6/6	0/0
xgridcmadupl	Available	6/6	0/0
grbcg011	Available	6/6	0/0
cagnode	Available	6/6	0/0
cms4uw...	Available	4/4	0/0
ce02lippt	Available	3/3	0/0
lok01fikarves	Available	3/3	0/0
ce.grid.ces.gaes	Available	2/4	0/0
grid01.physics.auth.gr	Available	2/2	0/0
ce001.grid.ucy.ac.cy	Available	1/1	0/0
grid14.physics.auth.gr	Available	1/1	0/0
cgnode06.lucea.gr	Available	1/1	0/0



<http://grid.ucy.ac.cy>

GridBench interface showing benchmark configuration. A red arrow points from a benchmark in the Template List to a resource in the Resource List. A callout box says "Drag a benchmark onto a resource". Another callout box points to the Monitoring tab in the "New bstream1_1 Benchmark" dialog, saying "Optional: Set parameters or leave defaults".

Parameters	Resources	Monitoring	Type
name	bstream1_1	value	
executable	bstream1_1	value	
execpath	/opt/cg/gridbench/bin	value	
datapath	/opt/cg/gridbench/d_	value	
stage_executable	manual	value	
iterations	2	value	
Reynolds	2.0	value	del
data_id	tube38x40x40	value	del
nx	36	value	del
ny	38	value	del
nz	38	value	del
den	1.0	value	del

Description:
The B_stream 1.0 - Light is the dressed down version of B_Stream 1.0, a full blown parallel 3D implementation of the D3Q19 Lattice Boltzmann method. B_Stream 1.0 - Light allows simulation of the steady flow in a small 3D tube at a fixed Re-number.

Submit using: globus 0% Advanced submit

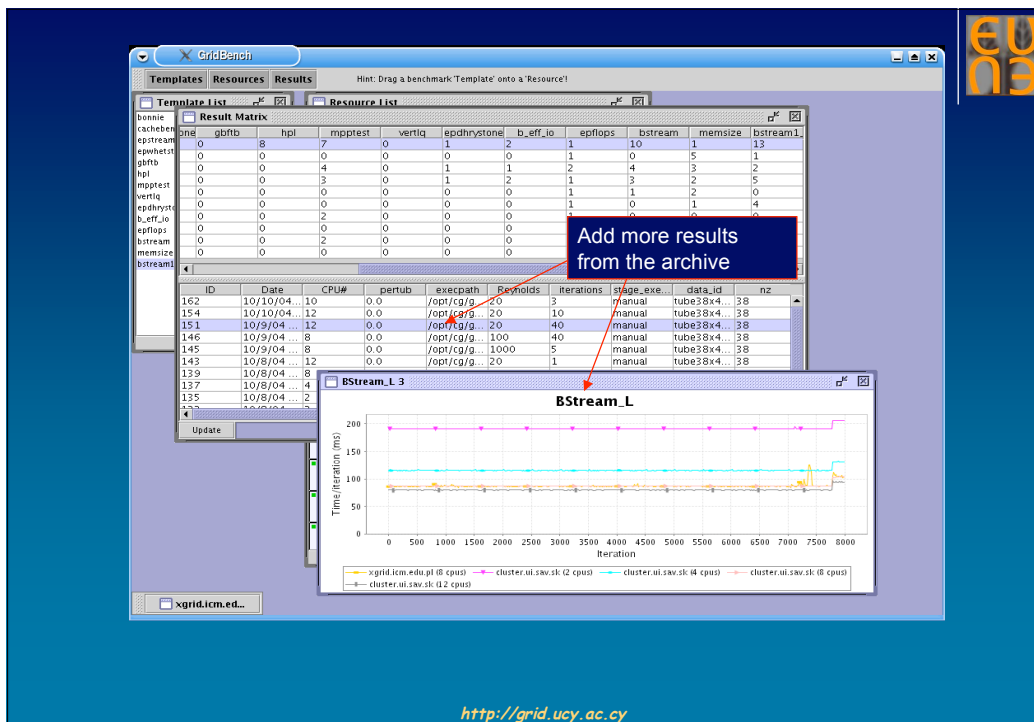
<http://grid.ucy.ac.cy>

GridBench interface showing benchmark results. A callout box points to the Results tab in the "xgrid.icm.edu.pl bstream1_1" dialog, saying "Plot result when benchmark finishes". A plot titled "BStream_L 3" shows "Time (iteration (ms))" vs "Iteration".

Parameters	Resources	Monitoring	Results
name	bstream1_1	value	
iteration_times	<vector>		
completion_time	698.786075		

Description:
This benchmark implements the steady flow in a small 3D tube at a fixed Re-number.

<http://grid.ucy.ac.cy>



Conclusions

- **Virtualization and resource heterogeneity** turn Grid Benchmarking into a:
 - Challenging and expensive process.
 - Necessary undertaking for performance-based decisions.
- **Isolated metrics** are of little use. We need instead, **structured sets of metrics**:
 - Describing collectively the performance capacity of an abstract representation of a Grid infrastructure.
 - Amenable to statistical distillation, to derive higher-level, qualitative metrics.
 - Whose storage, organization, visualization, and interpretation raises several difficulties.
- **Virtualization and the lack of central control**, put the accuracy of benchmarking measurements to question. The combination of metrics with monitoring information is required to filter-out invalid measurements.

Conclusions



- The **size**, the **openness**, and the **complexity** of the Grid, make it susceptible to a variety of **frequent, partial failures**. Thus:
 - Functionality benchmarking is equally important from an end-user's perspective.
 - Benchmarking a mechanism for driving **automatic remote healing**.
- The dynamic nature of Grid infrastructures necessitate a **periodic "refreshment"** of performance metrics. Thus:
 - Grid benchmarking can be established as an **automated central Grid service** running with **special privileges**.
 - Benchmarks can be used as a quick, "end-to-end" test of a Grid's **"health."**
 - Benchmarks can be used for the automatic **auditing** of resource providers by a VO administration: compliance to policies, reliability of information services, etc.

<http://grid.ucy.ac.cy>

Current and Future Work



- Extending **GBDL** and the **GBDL translator** to support:
 - The description of configuration-parameter selection constraints and guidelines.
 - The automatic selection of configuration parameters.
 - Interoperability with UNICORE middleware.
- Expanding the **GridBench suite**.
- Automating the process of **metrics filtering** and **decision support**.
- Using GridBench-metrics for **resource selection and brokerage**.
- Deriving **higher-level metrics** to express "quality features" of Grid infrastructures: Homogeneity, Trustworthiness of GIS, Health of the infrastructure, Reliability & Robustness.

<http://grid.ucy.ac.cy>

ΤΕΛΟΣ

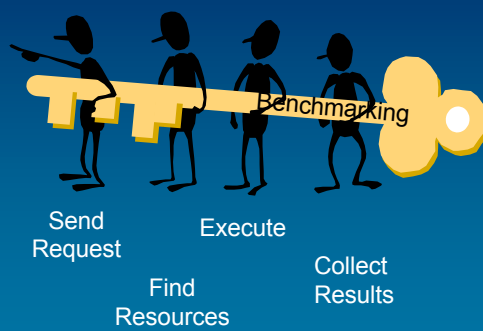
<http://grid.ucy.ac.cy>

Questions?

GridBench



Grid System



GridBench



<http://grid.ucy.ac.cy>

