Dziedzinowo zorientowane
usługi i zasoby infrastruktury
PL-Grid dla wspomagania
Polskiej Nauki w Europejskiej
Przestrzeni Badawczej

Benchmarking and Normalization of Computing Resources in PL-Grid Infrastructure

Wojciech Ziajka, Marcin Radecki,
Paweł Szepieniec and Tomasz Szepieniec

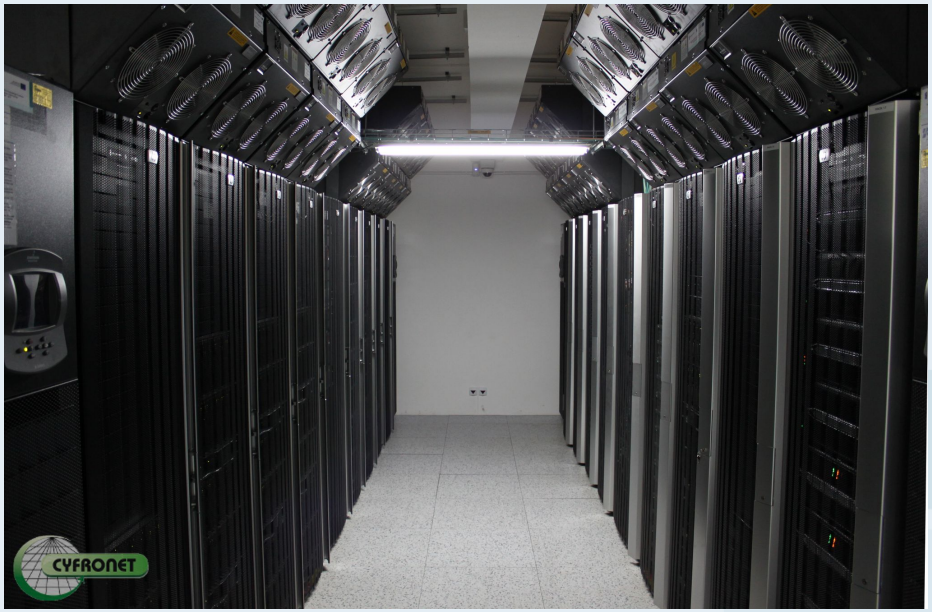
ACC Cyfronet AGH

CGW Krakow,
October 22-24, 2012

Differences in performance between clusters

CPU consumption on some less powerful system does not have the same value for the user as an equal consumption on an up-to-date resource.

„Odra”



„Zeus”

Motivation for normalization

- Fair accounting in heterogeneous infrastructure
  - **Different classes of nodes = different efficiency**
    - between computing centers
    - inside the clusters
  - The conversion of the use of resources, depending on nodes efficiency
- Support for users in applying for computational grants.
  - give information about the expected performance of the infrastructure
  - User should be able to estimate the size of the grant
    - Example: user experimets: sample test taks 6min on PC and user have to perform 100 000 simulations – **on PC it taks 10 000h, but how much time it will take on heterogeneous clusters ?**

System Requirements

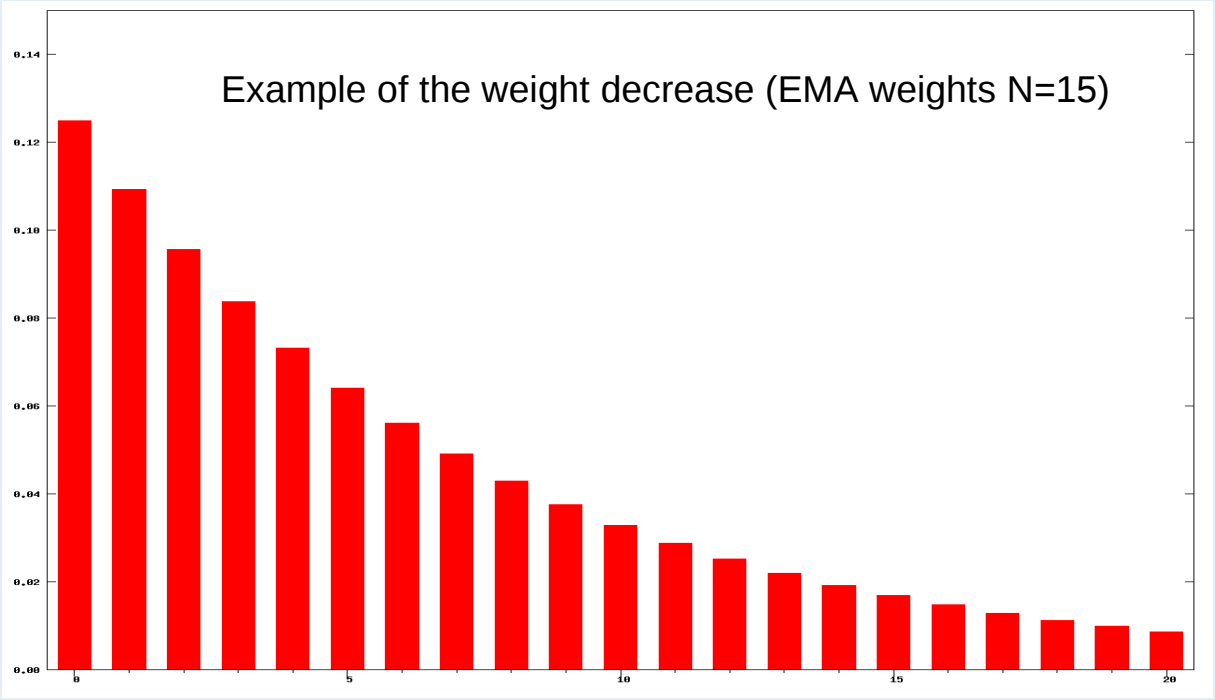
- Ability to **automatically adapt to dynamically changing** computing cluster infrastructure in following cases:
  - hardware modification,
  - addition or removal of nodes.
- System **automation**
  - Avoiding manual steps performed by administrators.
- **Real reflection** of user's feeling of the infrastructure
- Measuring of the infrastructure in production:
  - complements the current approach (tests on unloaded nodes),
  - involves performing tests on loaded nodes.

Benchmarking procedure

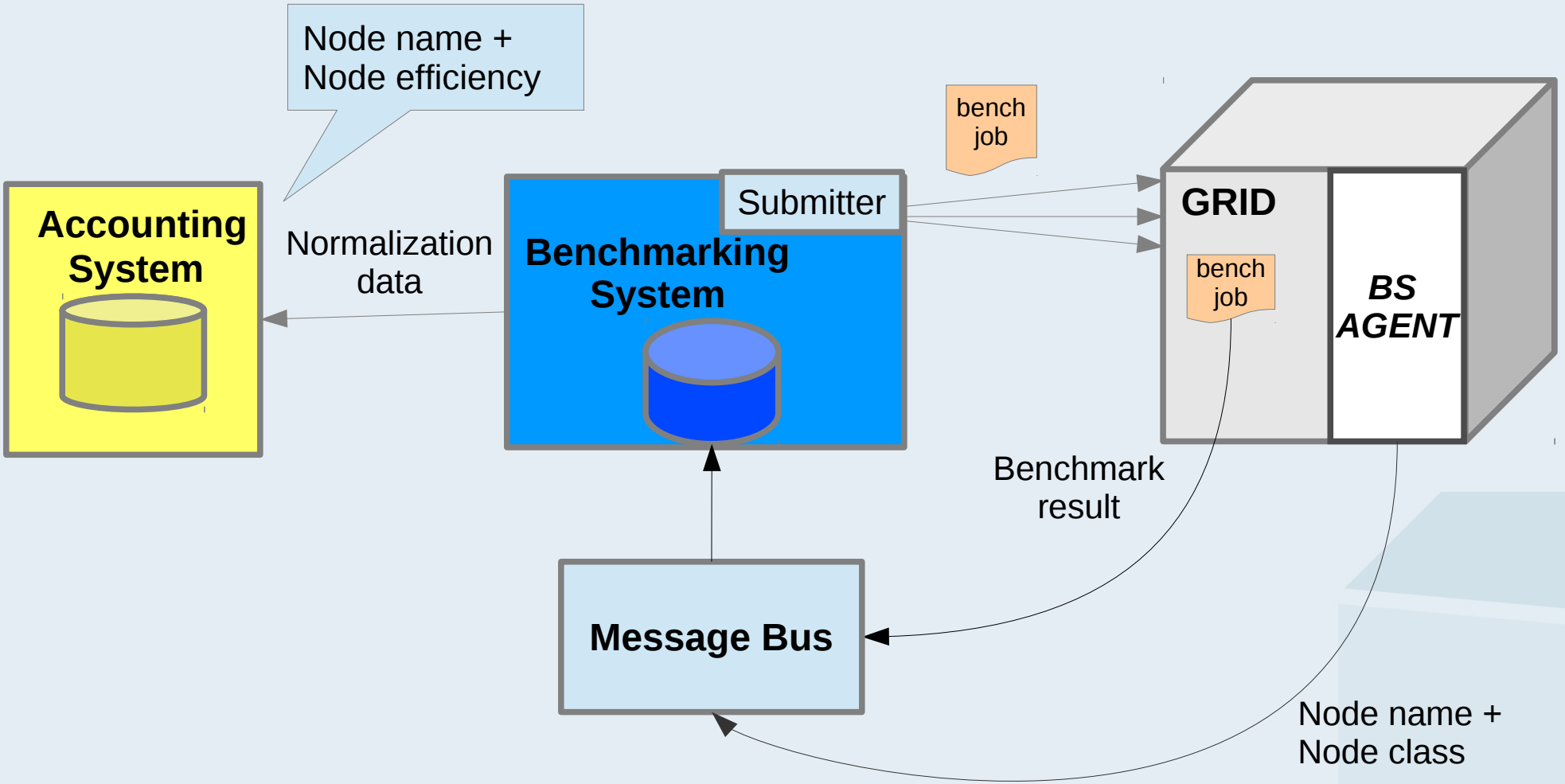
- Testing during the **production use** of the cluster.
- **HPL benchmark** is used - implementation of Linpack, which is the basis of the TOP500 ranking
  - This benchmark solves a dense system of linear equations for floating point double precision. By which mainly tests the floating-point arithmetic and memory access.
- Benchmark will run in **sequential** manner
  - Single bechmark test **taks 3-5 minutes** (using 256MB memory)
  - Perform tests **every 3 hours**
- Cluster administrators will be able to offer a set of options – it will allow for the most effective benchmark running.

Averaging benchmark results

- **Exponential moving average (EMA)**
  - Weights of older benchmarks results decrease exponentially



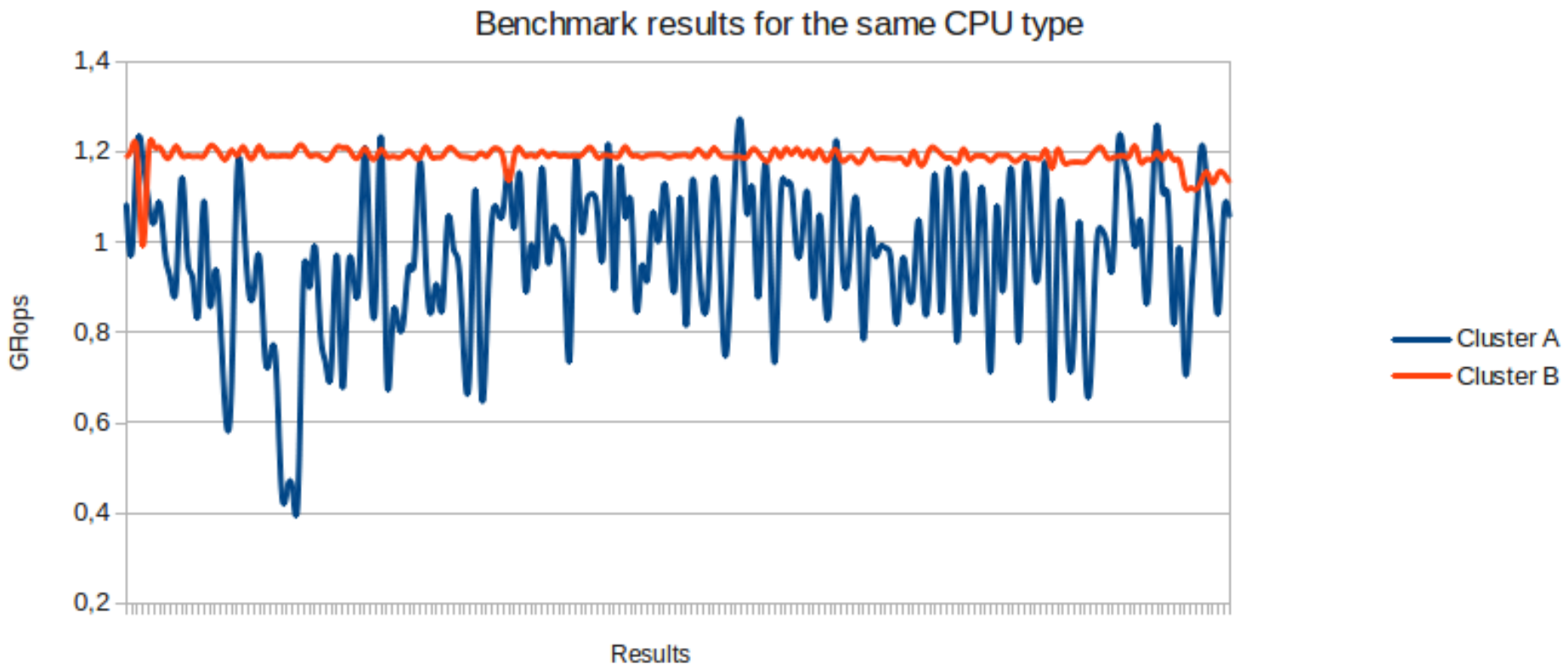
System Architecture



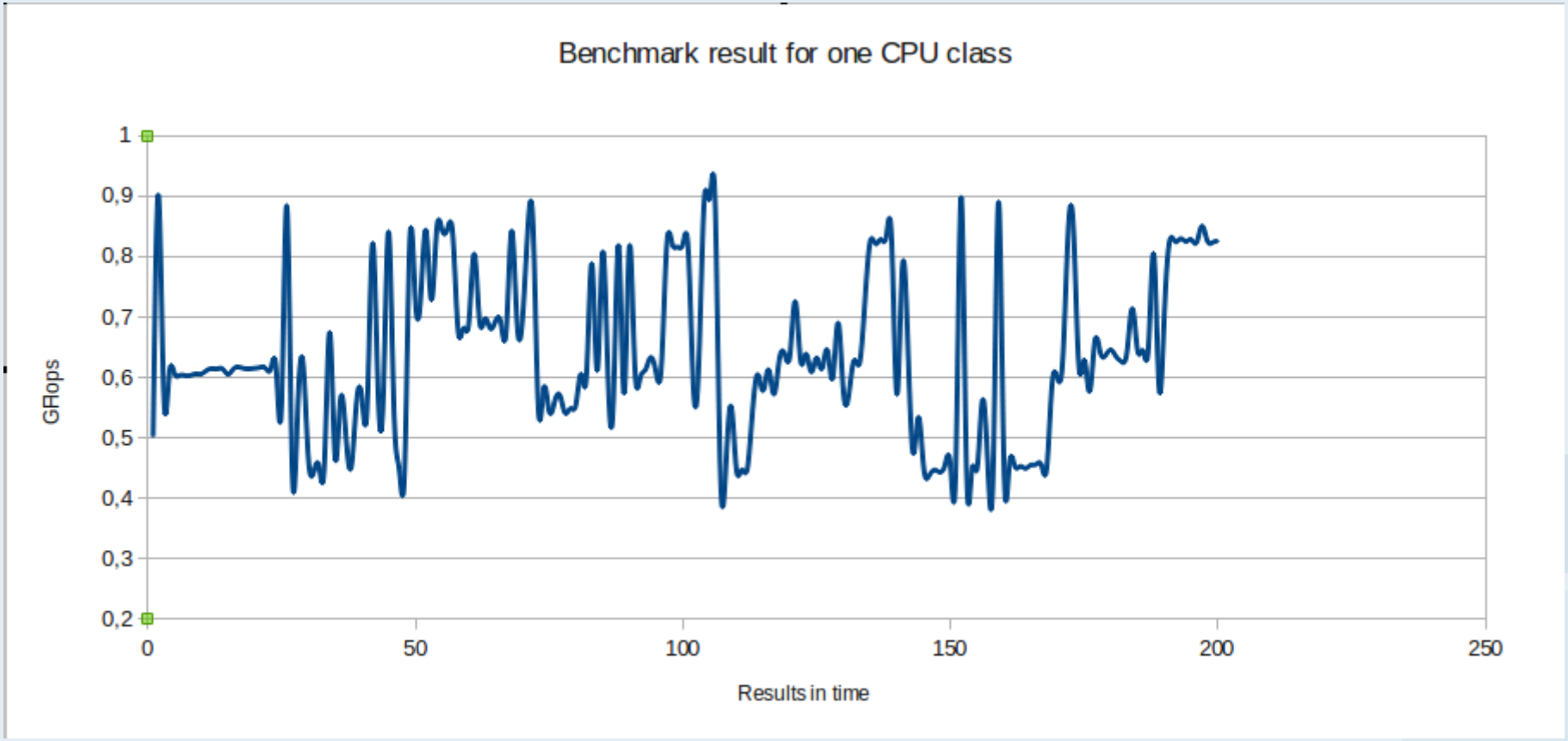
Benchmark results for two clusters

CPU model name	CPU Count	Current EMA	AVG	DEV	MIN	MAX	No. tests
Intel L5420 @ 2,50GHz	8	**0,5**	0,51	0,12	0,1	0,64	397
Intel L5640 @ 2,27GHz	12	**0,99**	0,96	0,16	0,45	1,26	200
Intel X5650 @ 2,67GHz	12	**1,09**	1,09	0,17	0,54	1,6	442
Intel E5645 @ 2,40GHz	12	**1,01**	0,98	0,21	0,38	1,44	287
Intel L5640 @ 2,27GHz	24	**1,02**	1,15	0,13	0,5	1,2	68
Intel E5345 @ 2,33GHz	8	**0,49**	0,49	0,05	0,2	0,51	275
Intel L5640 @ 2,27GHz	12	**1,05**	1,06	0,22	0,35	1,21	929
Intel E5530 @ 2,40GHz	8	**1,04**	1,04	0	1,04	1,04	1

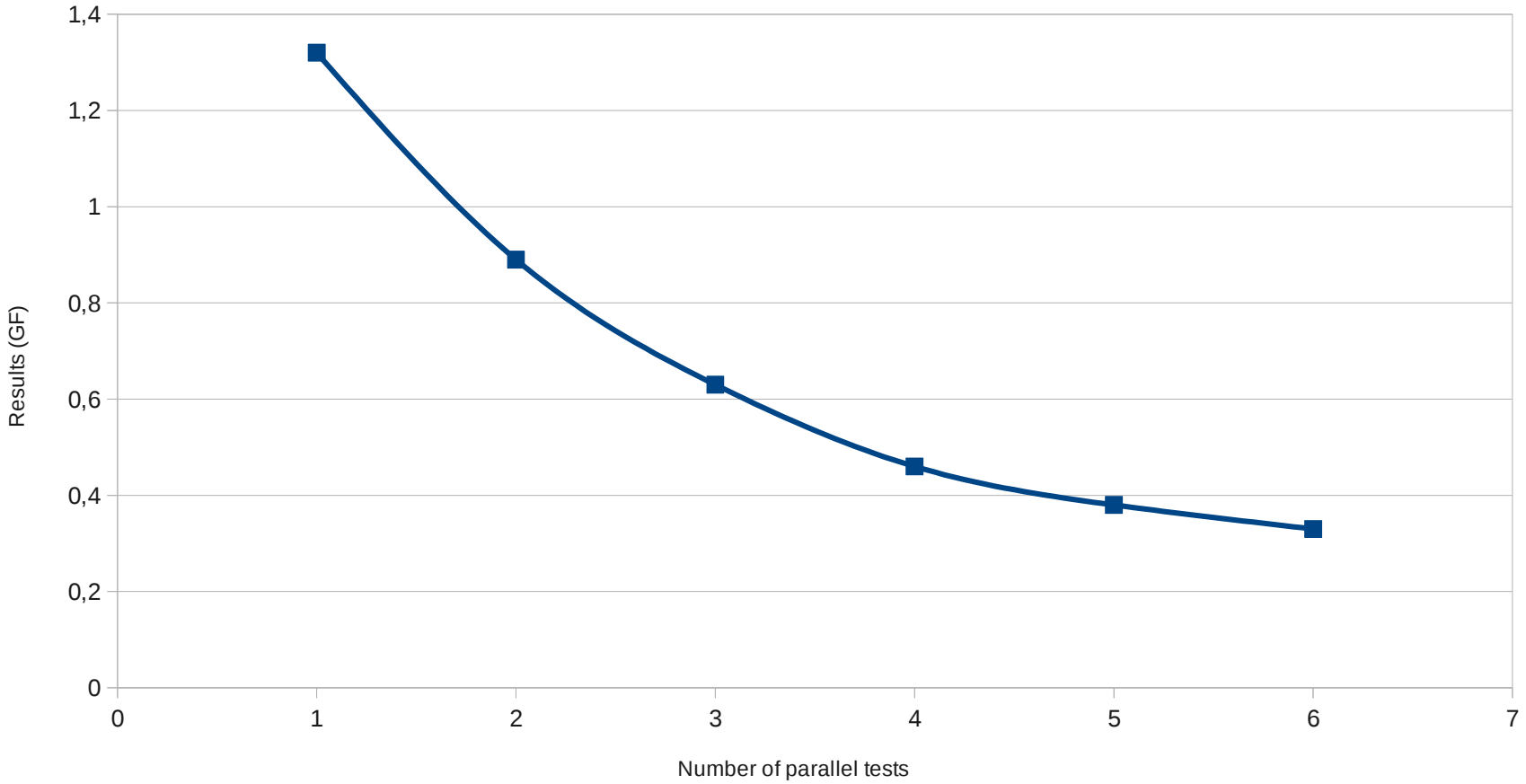
Results for two clusters with the same CPU class



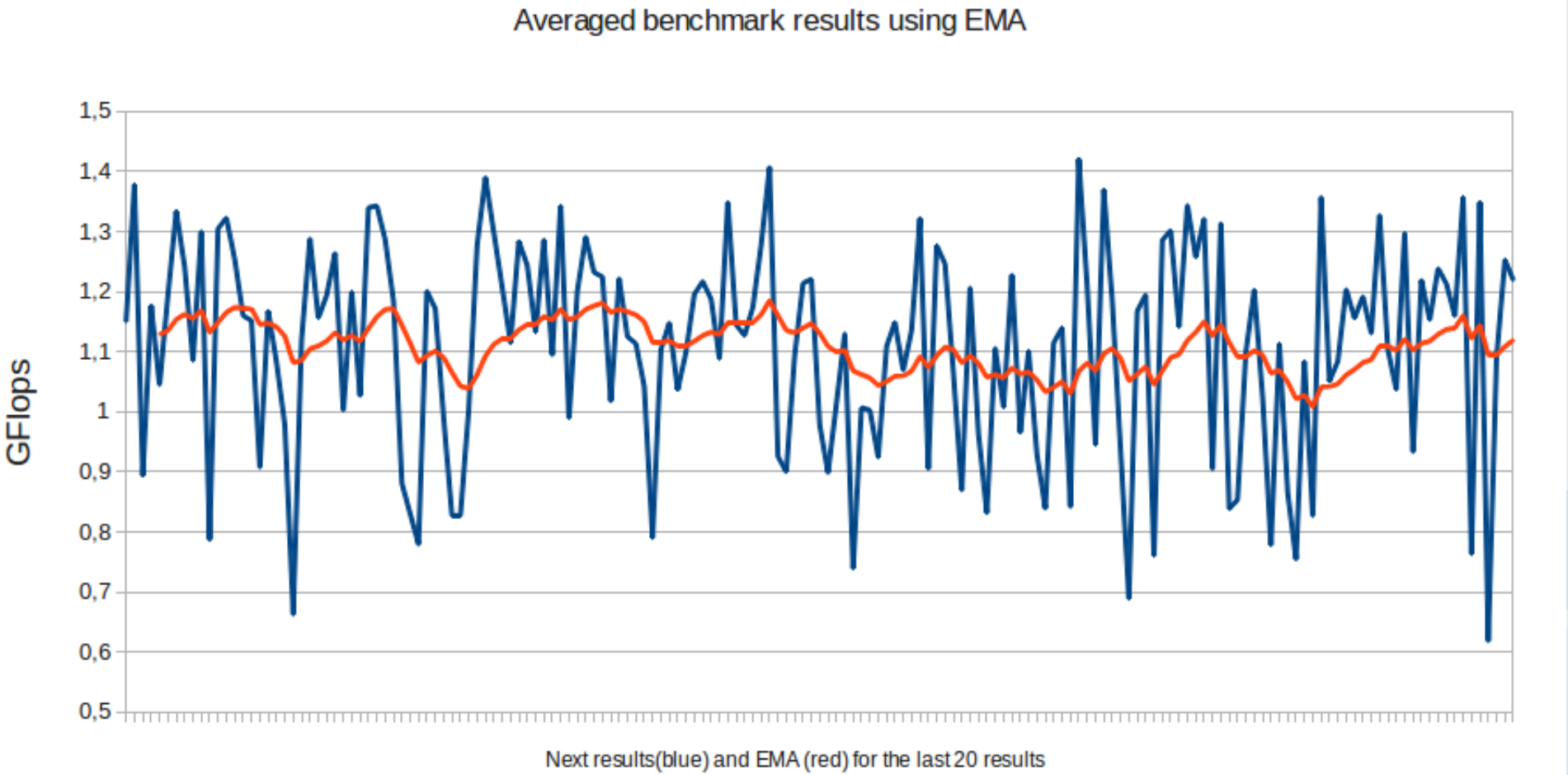
Sample results for one node



Running tests parallelly on a single 6-core machine



Averaging results of benchmark in time



Normalization of sample job

- Class A – nodes performance (EMA) = 0.49 GFlops
  - Job properties:
    - Execution time = 72h
    - Processors count = 64
    - Normalized Accounted walltime = 72h*64*0.49 = **2257.92 PLGh**
- Class B - nodes performance (EMA) = 1.09 GFlops
  - Job properties:
    - Execution time = 72h
    - Processors count = 64
    - Normalized Accounted walltime = 72h*64*1.09 = **5022.72 PLGh**

Future work

- Extend the benchamark to be more comprehensive
  - Include RAM and I/O characteristics
- Tuning benchmark using compilation options