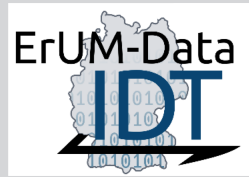
Benchmarking of compute resources for ATLAS

Albert-Ludwigs-Universität Freiburg

UNI FREIBURG

Benoît Roland, Anton Gamel, Markus Schumacher

ErUM Data IDT Collaboration Meeting
1st October 2019



Introduction

- Context and studies
- Compute resources under test and local environments
- Benchmarks used to estimate the CPU performance
- Benchmark results
- Validation of GRID benchmarking jobs
- Reliability of information in AGIS
- Conclusion and outlook

Context and studies

UNI FREIBURG

Context

- Use **fast benchmarks** as representative of CPU power for GRID activity
- Provide monitoring and prompt matching between workload and resources
- Provide prompt information about Virtual Machine performance
- **Optimize utilization of compute resources**

Studies

- Compare fast benchmarks to legacy HEP-SPEC06 (HS06) suite
- Study performance as a function of **core multiplicity**
- Study performance for **different machines and environments**
- Compare **local** benchmarking results to **GRID** benchmarking results

Compute resources under test and local environments

- **One CPU type** “Intel Xeon CPU E5-2630 v4 @ 2.20GHz” - 20 physical cores

  Tested in 3 different environments

- **ATLAS Tier-2/Tier-3 in Freiburg**

  Running bare metal
  Hyperthreading activated - Operating System SL6

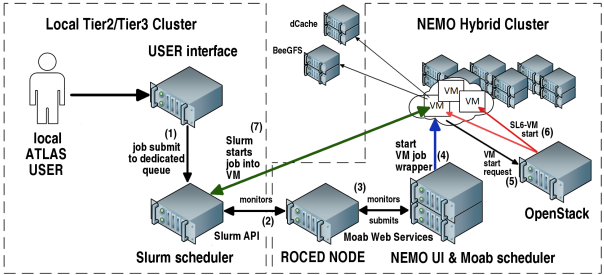
- **NEMO HPC Cluster - first configuration**

  Virtual Machine (VM)
  Hyperthreading activated - Operating System SL6

- **NEMO HPC Cluster - second configuration**

  Running bare metal
  No hyperthreading - Operating System CentOS7

NEMO HPC Cluster - 900 compute nodes - 20 cores per node
used as opportunistic resource by ATLAS Cluster



→ **Study impact of hyperthreading and virtualization on the benchmark behaviour**

CHEP 2018 proceedings

Benchmarks used to estimate the CPU performance

UNI FREIBURG

CERN Benchmarking Suite

- **HEP-SPEC06 (HS06):** Suite of benchmarks testing different CPU metrics (clock, bus,...)
  Weighted combination of the different benchmark results
  Legacy benchmark in HEP (Standard Performance Evalutation Corporation)
  Run time on the system under test / run time on a reference processor

- **Whetstone:** Floating-point arithmetic operations
  Millions of Whetstone Instructions Per Second - MWIPS
  Instructions in the intermediate language used by the compiler

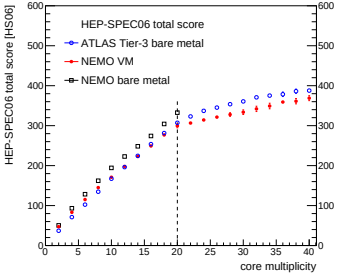
- **Dirac 2012 (DB12):** Operations on random numbers generated according to a Gaussian distribution
  HEP SPEC 2006 - HS06

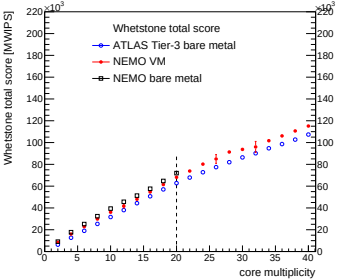
- **Kit Validation (KV):** GEANT4 simulation - interaction of single muon events in ATLAS detector
  Number of events produced per second
  Realistic HEP workload

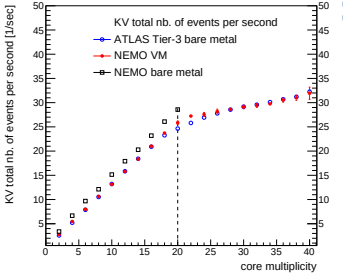
Benchmark results

- Scores of HEP-SPEC06, Whetstone and KV for the three configurations
- As a function of the number of cores used to run the benchmarks
- Benchmarks run 20 times for each core multiplicity in a varying random sequence
- Extract means and standard deviations of obtained distributions

Average total benchmark scores versus core multiplicity





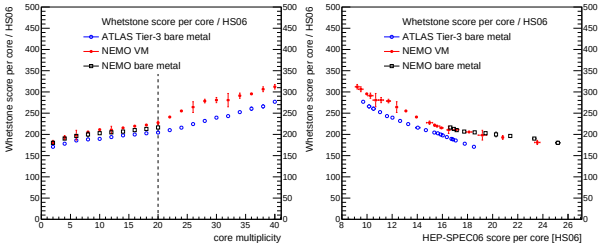


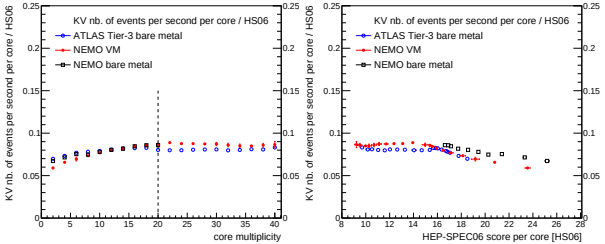
ATLAS Tier-3 bare metal, **NEMO VM**: both with hyperthreading and SL6 $\rightarrow$ **impact of virtualization**

- **HEP-SPEC06:** loss of performance due to virtualization does not exceed 5% at high core multiplicity
- **Kit Validation:** no impact from virtualization

Impact of hyperthreading: 25% increase in performance for HS06 and KV, 80% for Whetstone

Scaling of fast benchmarks with HEP-SPEC06





→ **Ratio between Whetstone and HS06 depends on core multiplicity and HS06**

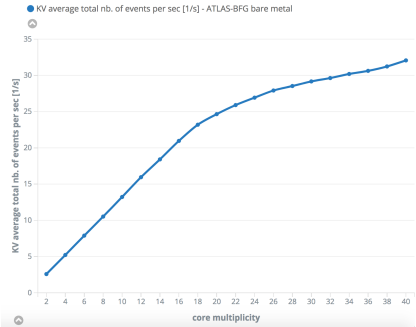
→ **Ratio between KV and HS06 flatter**

Benchmark results - Elasticsearch instance at CERN

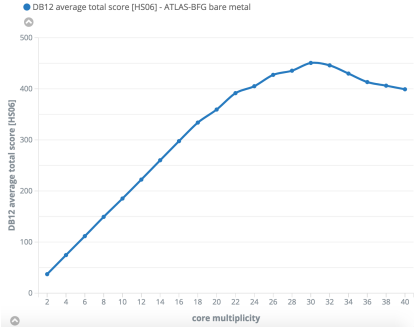
benchmark type, score per core, total score, benchmark duration, information about CPU, OS, grid sites,...

Benchmark results - Visualization with Kibana

KV average total score versus n cores



DB12 average total score versus n cores



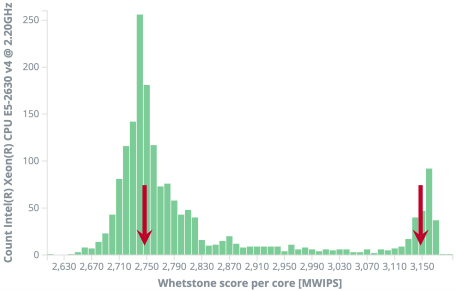
Visualizations and Dashboards available in Kibana

Average score per core, average total score, average benchmark duration, ...

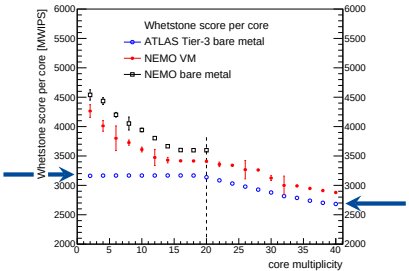
Validation of GRID benchmarking jobs

Whetstone benchmark run at Uni Freiburg - Intel Xeon CPU E5-2630 v4 @ 2.20GHz

GRID jobs - Whetstone score per core



Local jobs - Average Whetstone score per core as a function of core multiplicity



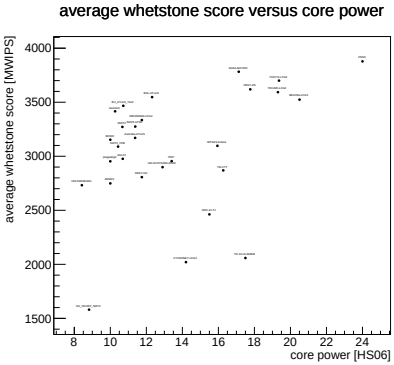
GRID score distribution shows 2 peaks - Small peak at 3150 MWIPS - Main peak at 2750 MWIPS

Local average score ranges from 3150 MWIPS (below 20 cores) to 2700 MWIPS (at 40 cores)

→ **GRID and local scores compatible - jobs distribution well optimized!**

Reliability of information in AGIS

- Core power values manually entered in AGIS by site responsibles
- Information used to quantify site resources for GRID job distribution



- Correlation between AGIS core power and average Whetstone results
- **Values in AGIS not representative of the actual core performance**

Conclusion and outlook

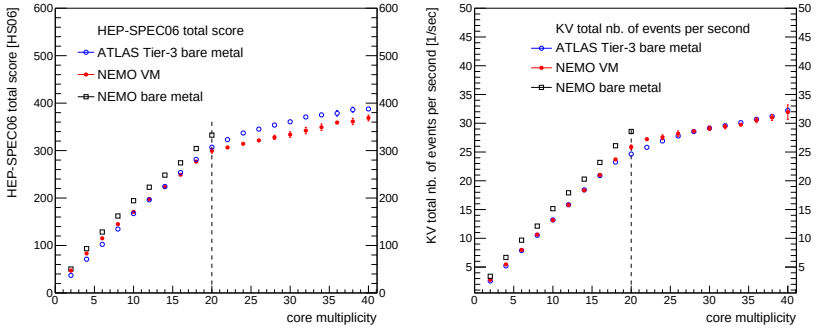
UNI FREIBURG

- Fast benchmarks are **reliable** and follow a behavior **similar** to HEP-SPEC06
- KV benchmark - **realistic** HEP workload - behaviour **closest** to HEP-SPEC06
- Virtualization has **little to no impact** with respect to running bare metal
- Scores retrieved from **GRID** agree with results obtained **locally** on the Freiburg cluster
- **Core powers manually entered in AGIS could be favourably replaced by the outputs of fast benchmarks run periodically in the Hammercloud framework** code on gitlab

UNI FREIBURG

Backup

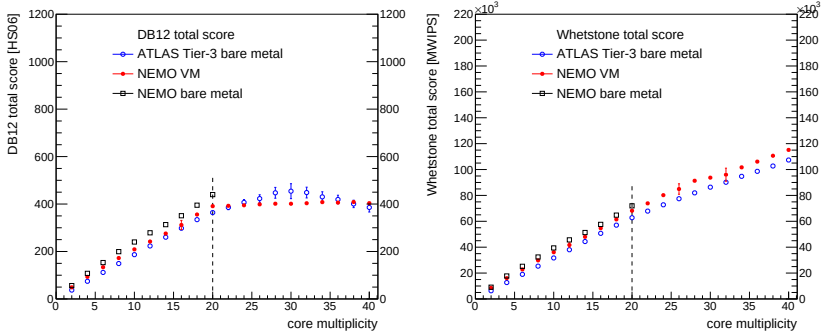
Average total benchmark scores versus core multiplicity - HEP-SPEC06 - KV



ATLAS Tier-3 bare metal, **NEMO VM**: both with hyperthreading and SL6 $\rightarrow$ **impact of virtualization**

- **HEP-SPEC06:** loss of performance due to virtualization does not exceed 5% at high core multiplicity
- **Kit Validation:** no impact from virtualization

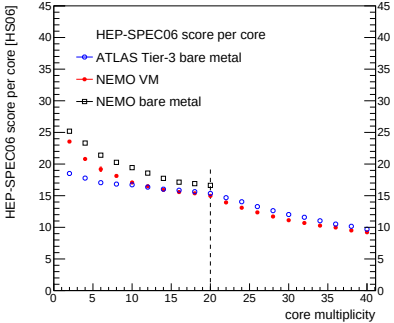
Average total benchmark scores versus core multiplicity - DB12 - Whetstone

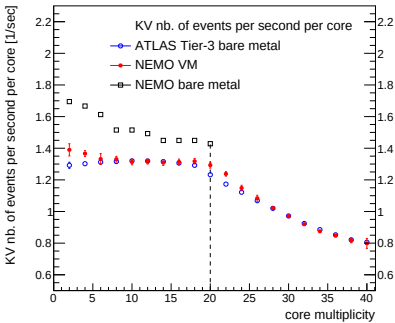


ATLAS Tier-3 bare metal, **NEMO VM**: both with hyperthreading and SL6 → **impact of virtualization**

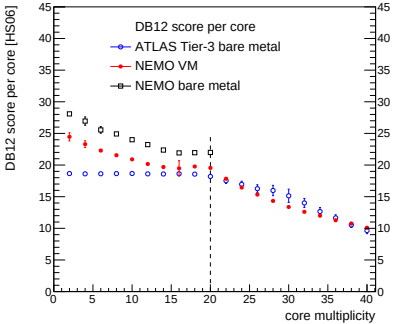
- **DB12:** small impact from virtualization
- **Whetstone:** VM has better performance than running bare metal

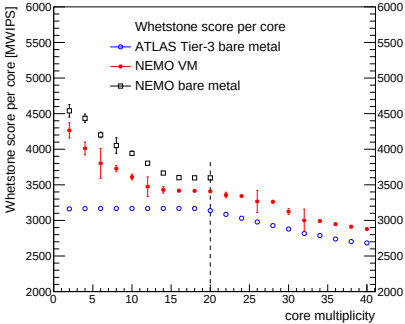
Average benchmark scores per core versus core multiplicity - HEP-SPEC06 - KV





Average benchmark scores per core versus core multiplicity - DB12 - Whetstone





Validation of GRID benchmarking jobs

UNI FREIBURG

Benchmark run at Uni Freiburg through GRID benchmarking jobs

- Intel Xeon CPU E5-2630 v4 @ 2.20GHz - Score distribution
- Three queues - Distribution over the last 6 months

  MC reconstruction - 8 cores jobs → **benchmarks run on 8 cores**

  MC generation - 1 core jobs → **benchmarks run on 1 core**

  User analysis - 1 core jobs → **benchmarks run on 1 core**
- **not possible to know exact CPU workload** because presence of extra local jobs

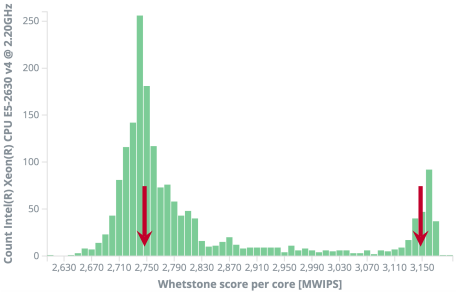
Benchmark run locally at Uni Freiburg

- Intel Xeon CPU E5-2630 v4 @ 2.20GHz - Average score distribution versus core multiplicity

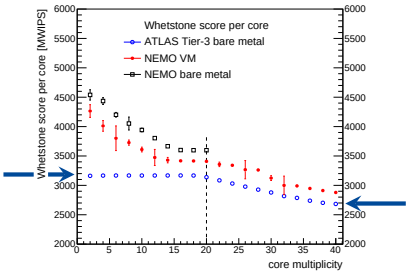
Validation of GRID benchmarking jobs - Whetstone

Whetstone benchmark run at Uni Freiburg - Intel Xeon CPU E5-2630 v4 @ 2.20GHz

GRID jobs - Whetstone score per core



Local jobs - Average Whetstone score per core as a function of core multiplicity



GRID score distribution shows 2 peaks - Small peak at 3150 MWIPS - Main peak at 2750 MWIPS

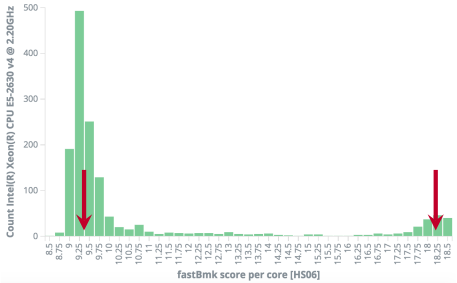
Local average score ranges from 3150 MWIPS (below 20 cores) to 2700 MWIPS (at 40 cores)

→ **GRID and local scores compatible - jobs distribution well optimized!**

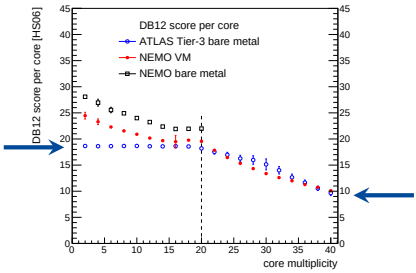
Validation of GRID benchmarking jobs - fastBmk - DB12

fastBmk - DB12 benchmarks run at Uni Freiburg - Intel Xeon CPU E5-2630 v4 @ 2.20GHz

GRID jobs - fastBmk score per core



Local jobs - Average DB12 score per core as a function of core multiplicity



GRID score distribution shows 2 peaks - Small peak around 18.25 HS06 - Main peak around 9.25 HS06

Local average score ranges from 18.5 HS06 (below 20 cores) to 10 HS06 (at 40 cores)

→ **GRID and local scores compatible - jobs distribution well optimized!**