Benchmarking of (CPU) servers

Dr. Ulrich Schwickerath, CERN

See also:

https://twiki.cern.ch/twiki/bin/view/FIOgroup/ProcRefHyperthreading

Outline:

- Hardware used for HEPSPEC06 measurements
- Influence of SMT and Turbo mode on HEPSPEC06 results
- HEPSPEC06 Scaling behavior
- HEPSPEC06 Systematic errors
- First look at high statistics power measurements
- Summary

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Hardware used for HEPSPEC06 measurements

- **Chassis:**
  - SuperMicro CSE-827T-R1200B Twin2
  - equipped with 4 identical servers
- **Servers: 4**
  - Supermicro X8DTT-F board
  - Intel® 5520 “Tylersburg” chipset
  - 2 Intel® XEON L5520 CPUs, 2.26 GHz (Nehalem)
  - 16 GB RAM
  - 2X 500GB disks

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

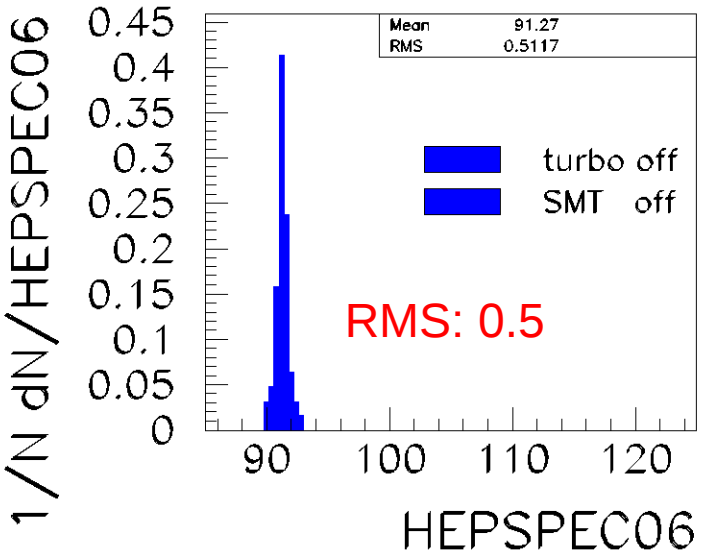
Can we gain from SMT or Turbo mode ?

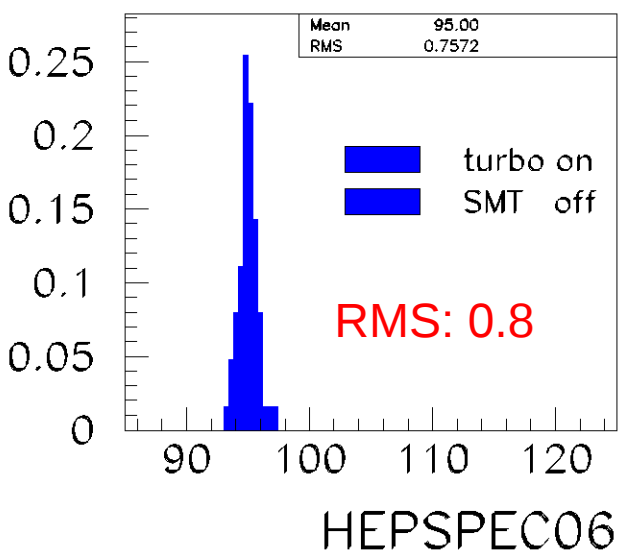
Idea of this test:

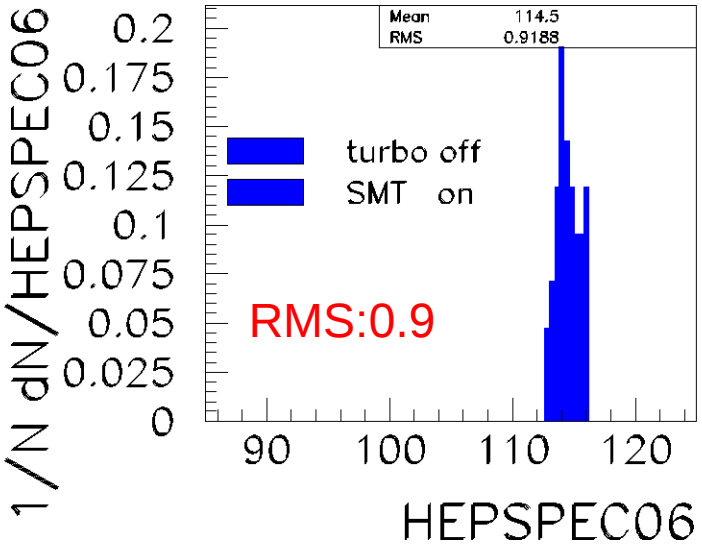
- Perform “large” number of HEPSPEC06 measurements on identical hardware but different BIOS settings
  - Turbo mode off, SMT off **(default)**
  - Turbo mode on, SMT off
  - Turbo mode off, SMT on
  - Turbo mode on, SMT on
- Histogram results by machine
- Measure relative improvement with respect to the default setting, and derive errors if possible

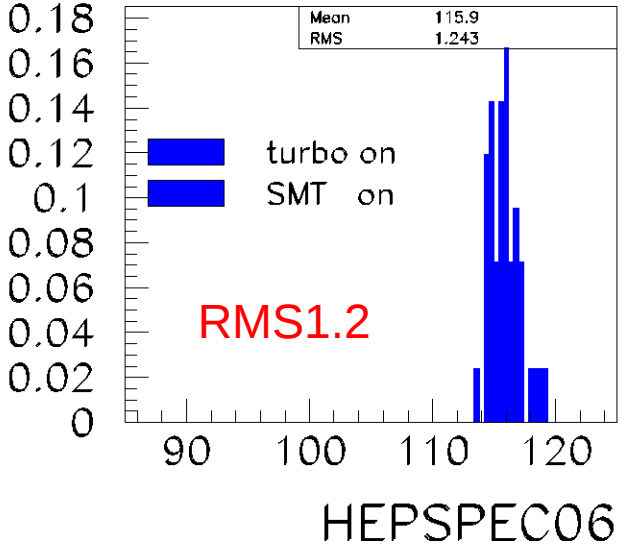
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

SMT and Turbo mode



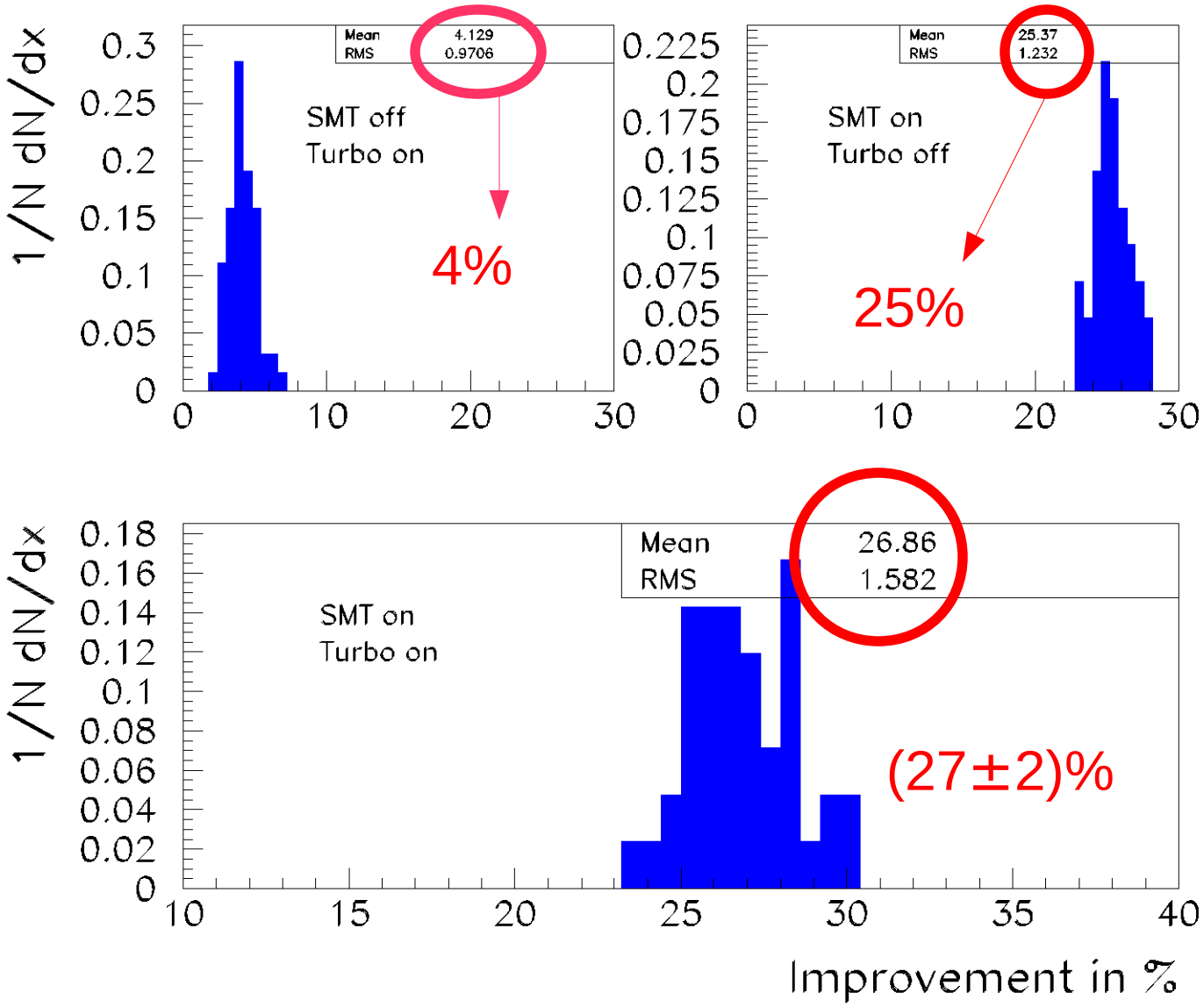






CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

SMT and Turbo mode: relative improvements



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

SMT and Turbo mode test: Observations

- The two options are correlated, the individual improvements don't sum up
- Turbo mode alone gives a (4.1±1.0)% improvement
- SMT alone results in (25.4±1.3)% performance improvement
- Both together result in (26.9±1.6)%
- Both options increase the uncertainty on of the measurements

Note: the benchmark was run with 1 process/logical CPU (8 or 16 processes)

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark scaling behavior

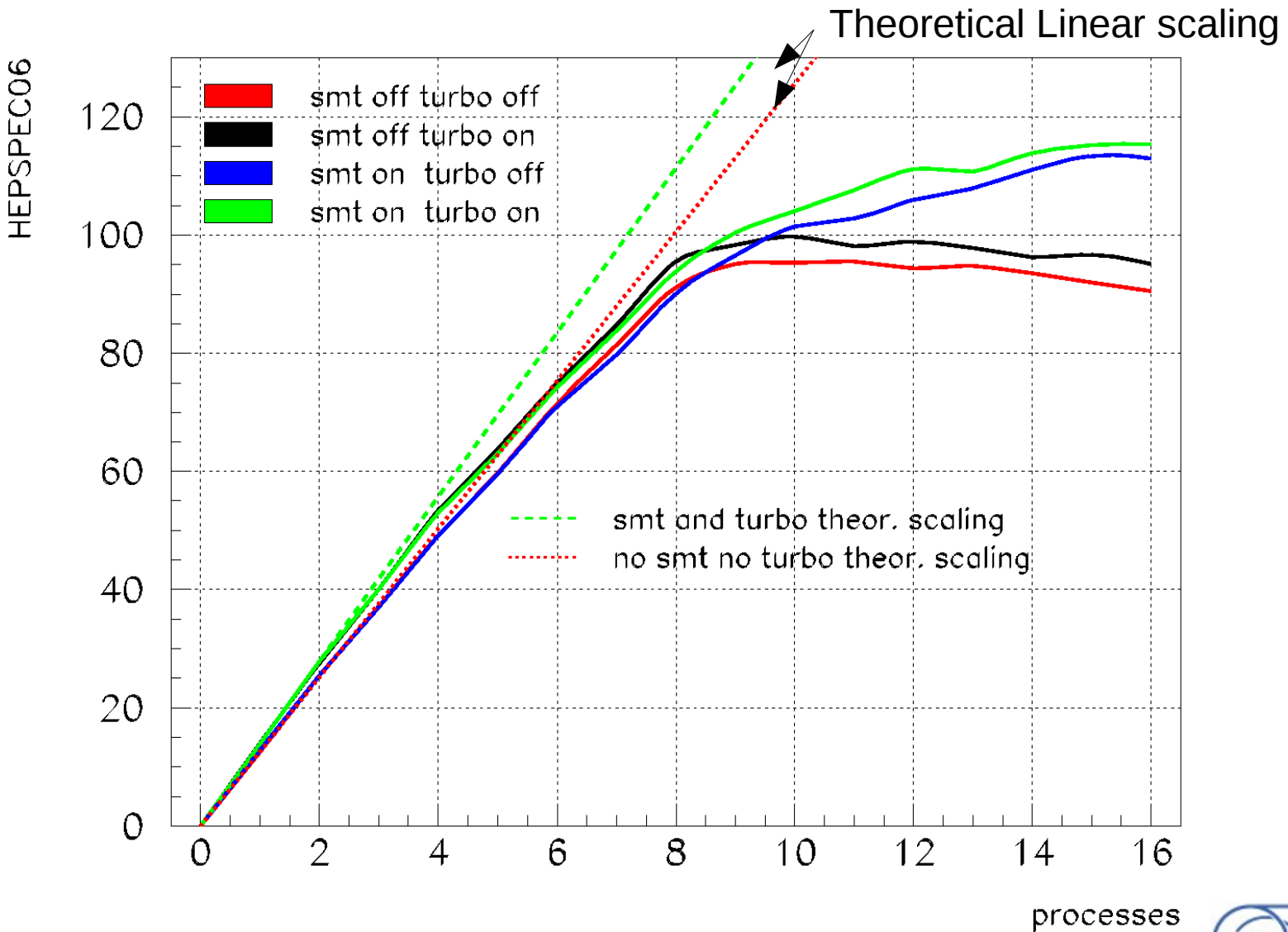
Idea of this test:

Test HEPSPEC06 performance as a function of the number of processes

- Use the same machines as in the previous test
- Scan up to 16 processes (=number of logical cores with SMT)
- Compare to theoretical linear scaling, based on the first two measurement points

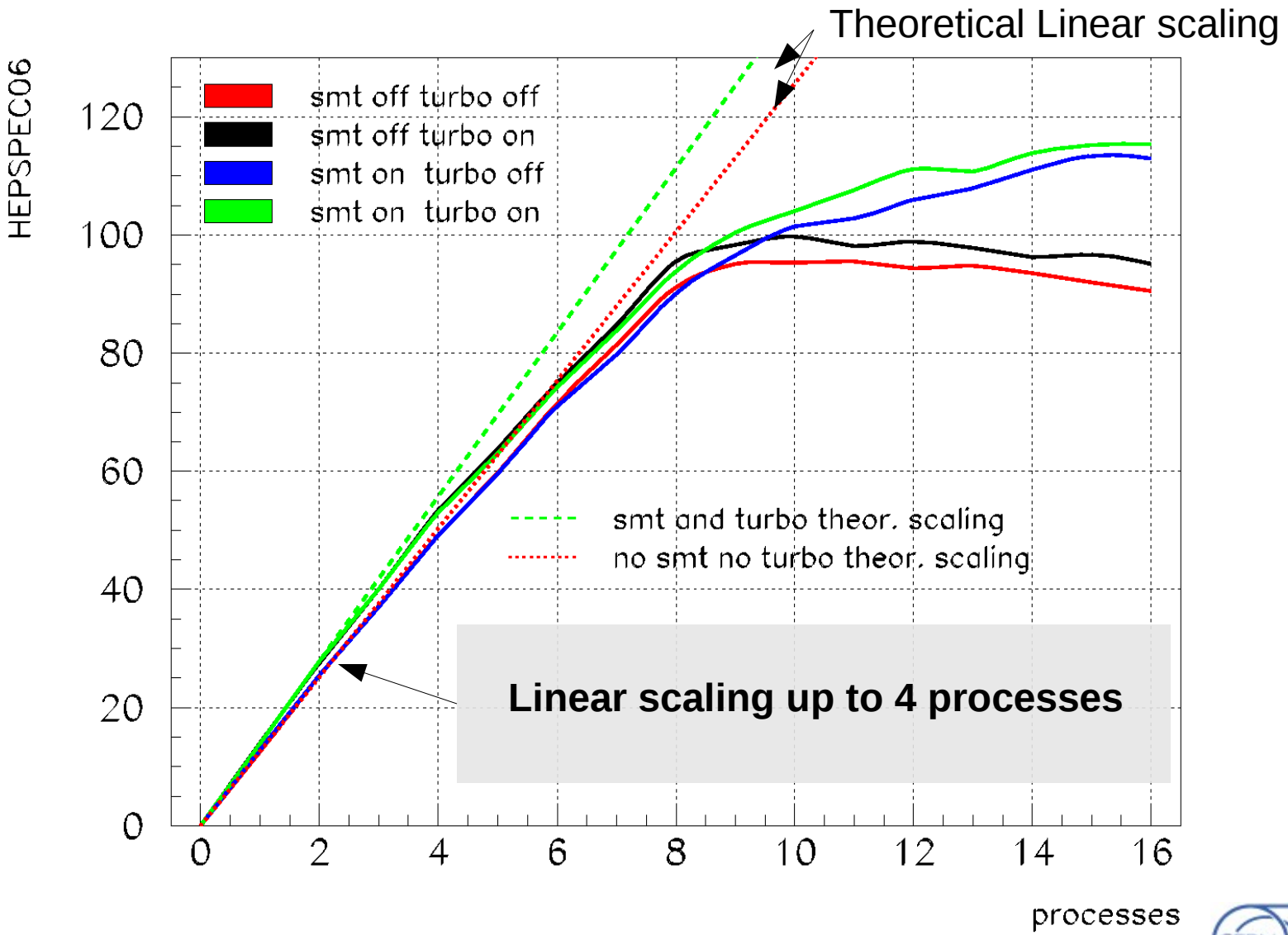
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark: scaling behavior



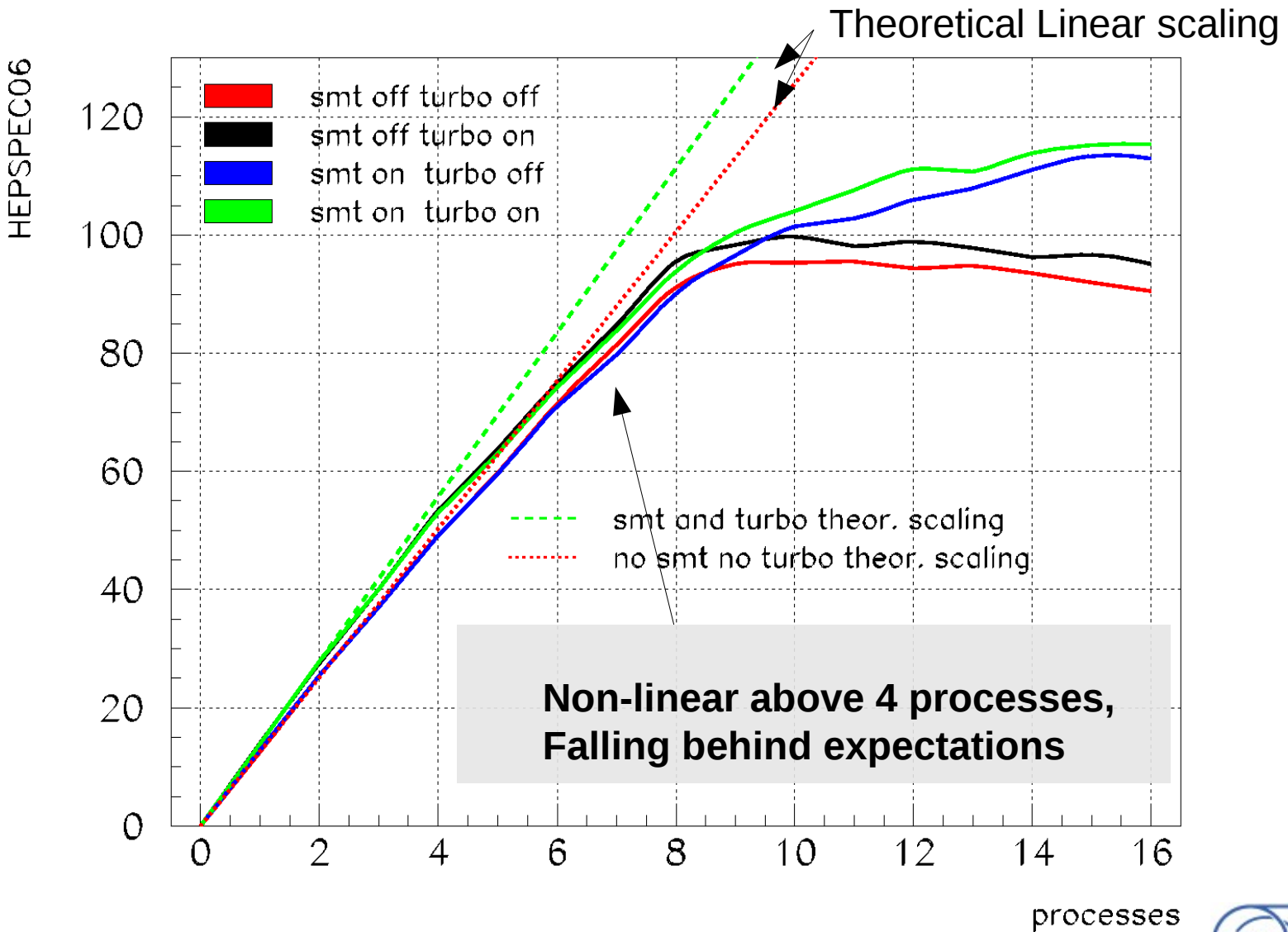
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark: scaling behavior



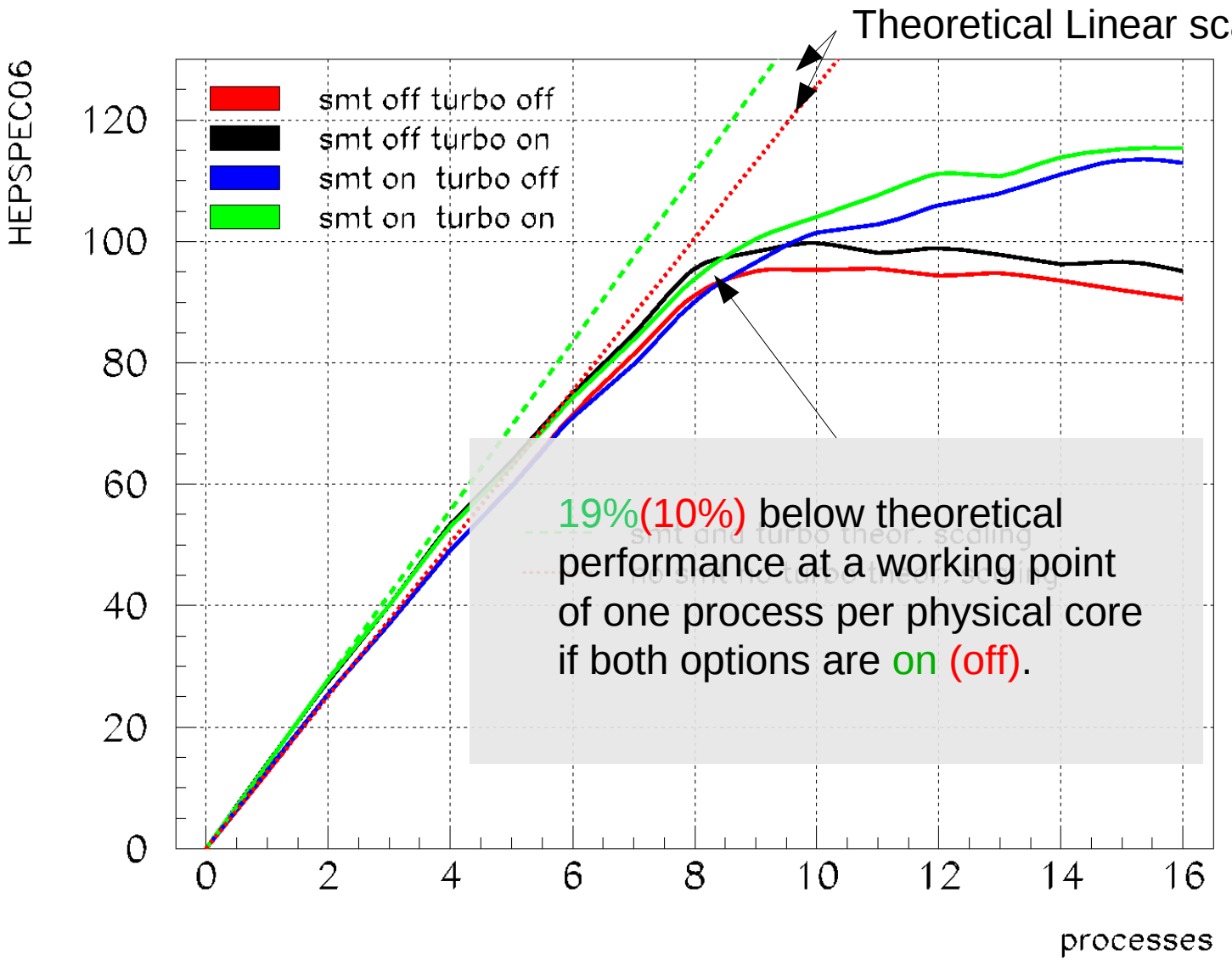
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark: scaling behavior



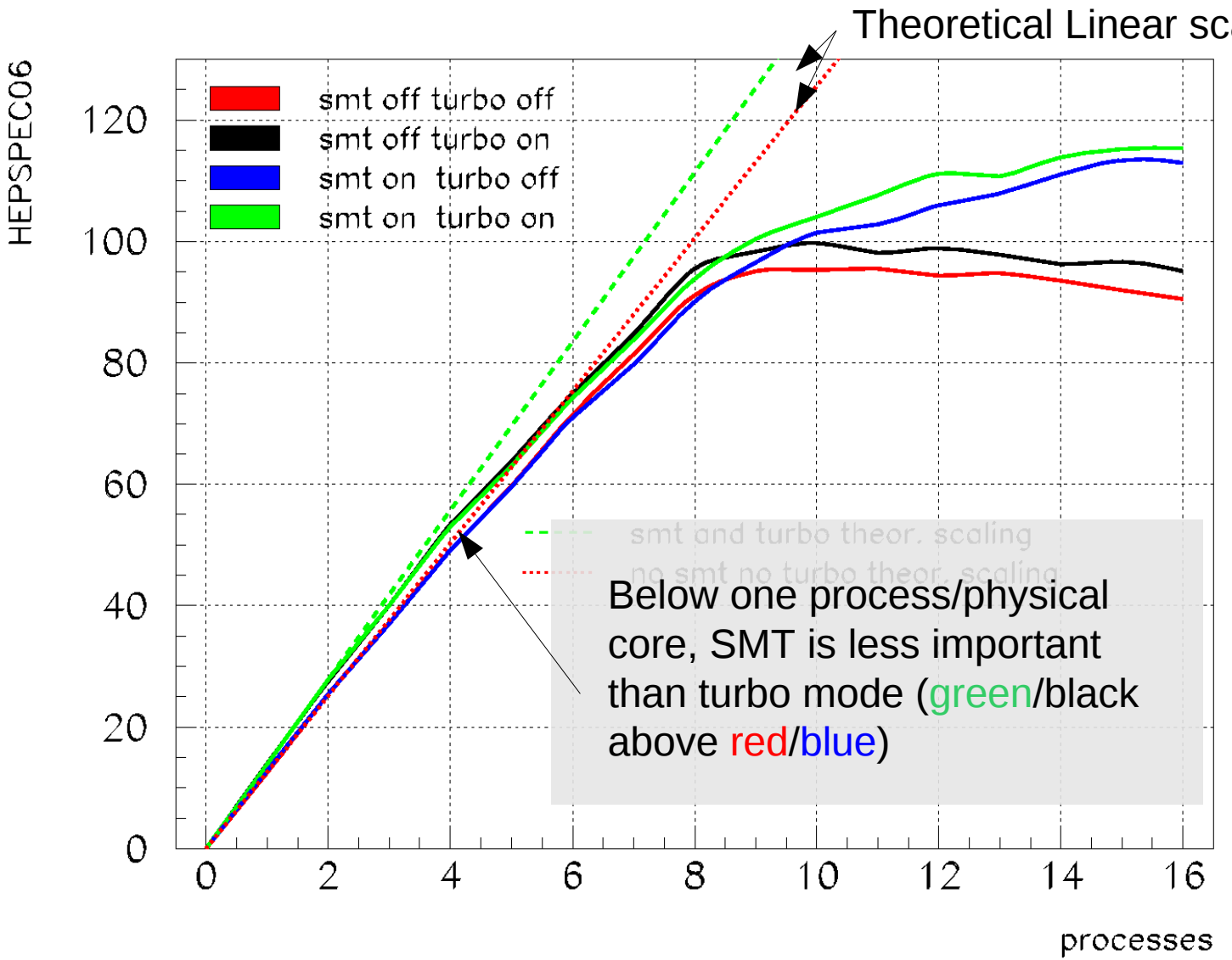
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark: scaling behavior



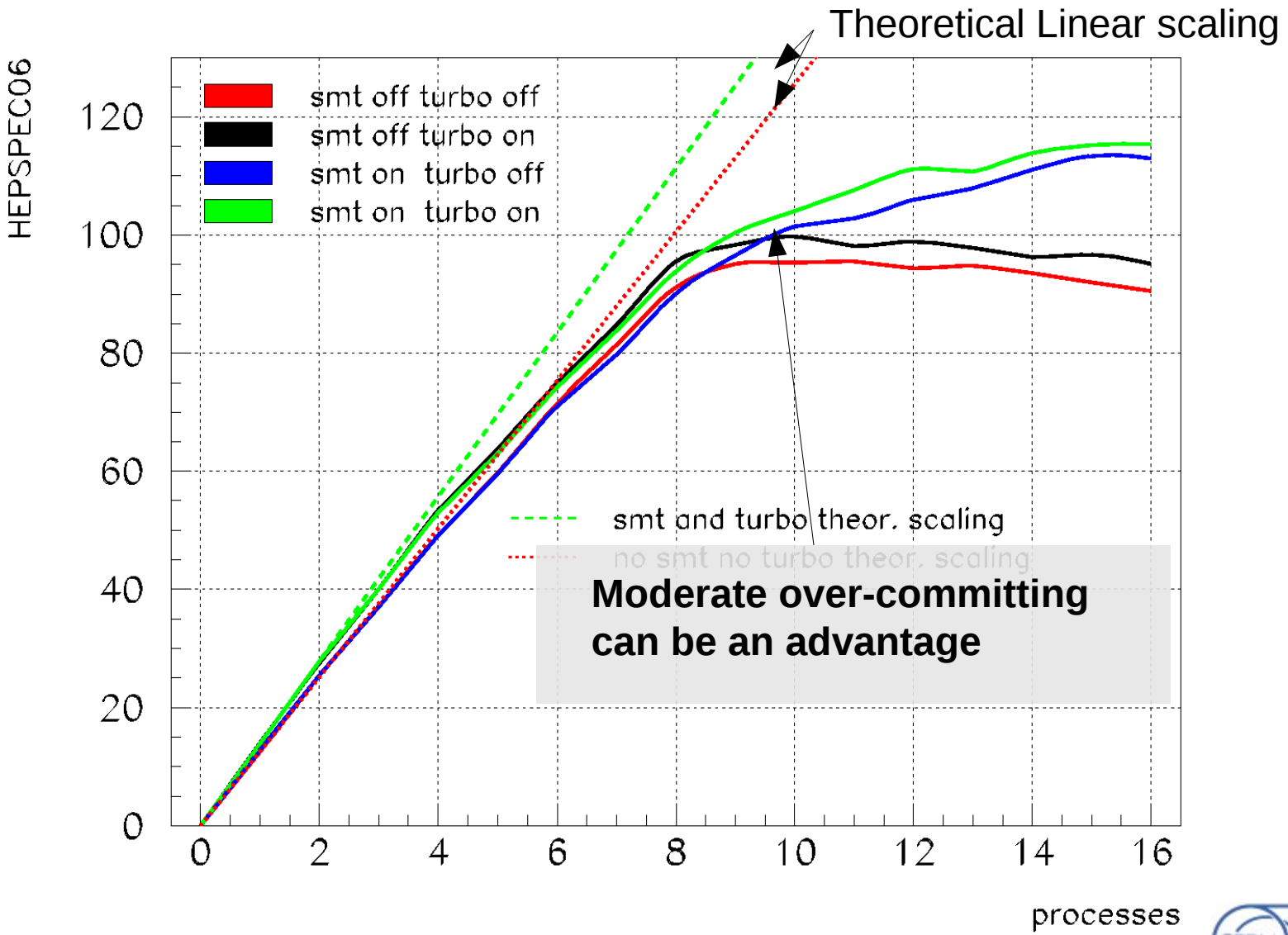
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark: scaling behavior



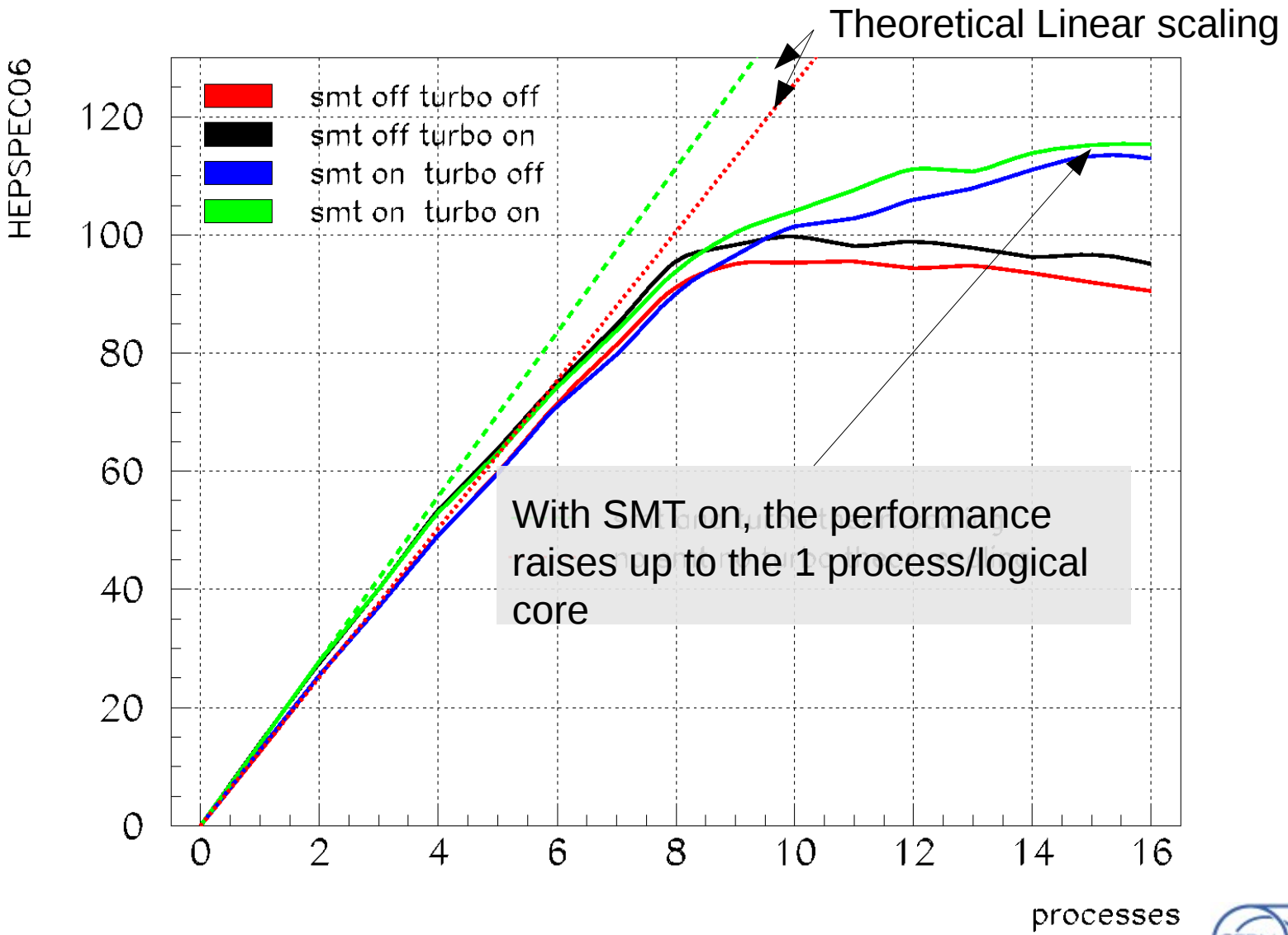
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark: scaling behavior



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEPSPEC06 benchmark: scaling behavior



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Scaling behavior: observations

- Linear scaling up to 4 processes
- Non-linear scaling behavior above 4 processes, below theoretical curves
- 19%(10%) below theoretical performance at a working point of one process per physical core if both options are on (off).
- Below 1 process/physical core, SMT is less important than turbo mode
- Even with SMT off, moderate **over-committing can be an advantage**
- With SMT on, the performance raises up to the 1 process/logical core

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Alternative application:CDB benchmarking

CDB benchmark: central piece of Quattor tool suite

- Metric: real time needed, in seconds
- Compilation of large number of small files which include each other
- Resolving dependencies
- Used at CERN for node management

	turbo off	turbo on
SMT off	443s	445s
SMT on	534s	519s

- Turbo mode does not help
- SMT is harmful !

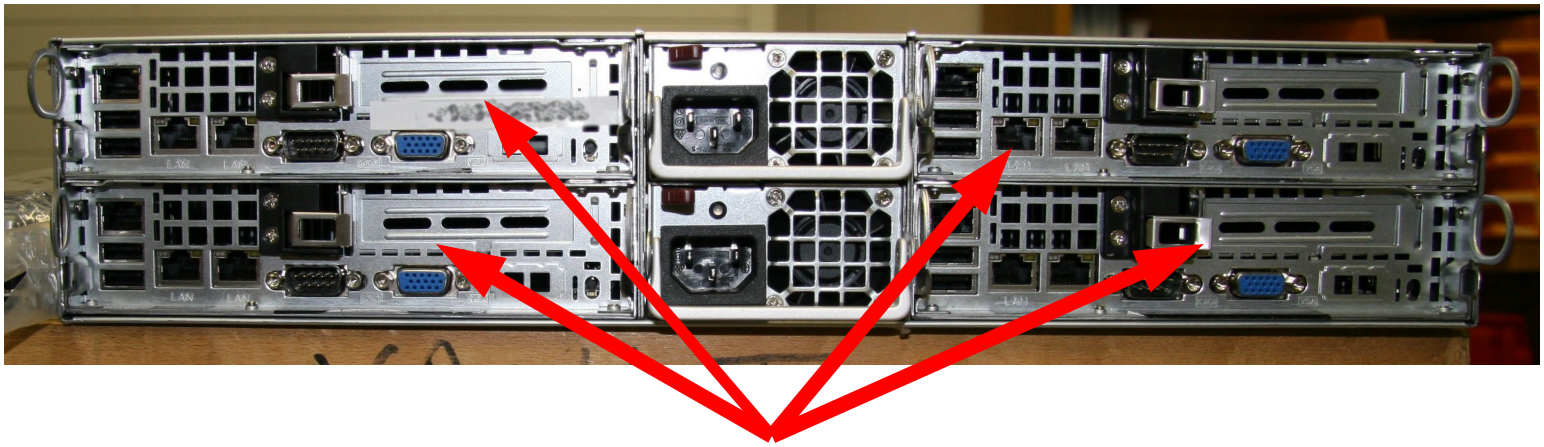
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Work in progress: systematic errors

Idea of this test: test for systematic variations

- Run the benchmark on **all 4 machines** with the same settings
- Any significant differences should be treated as systematic errors

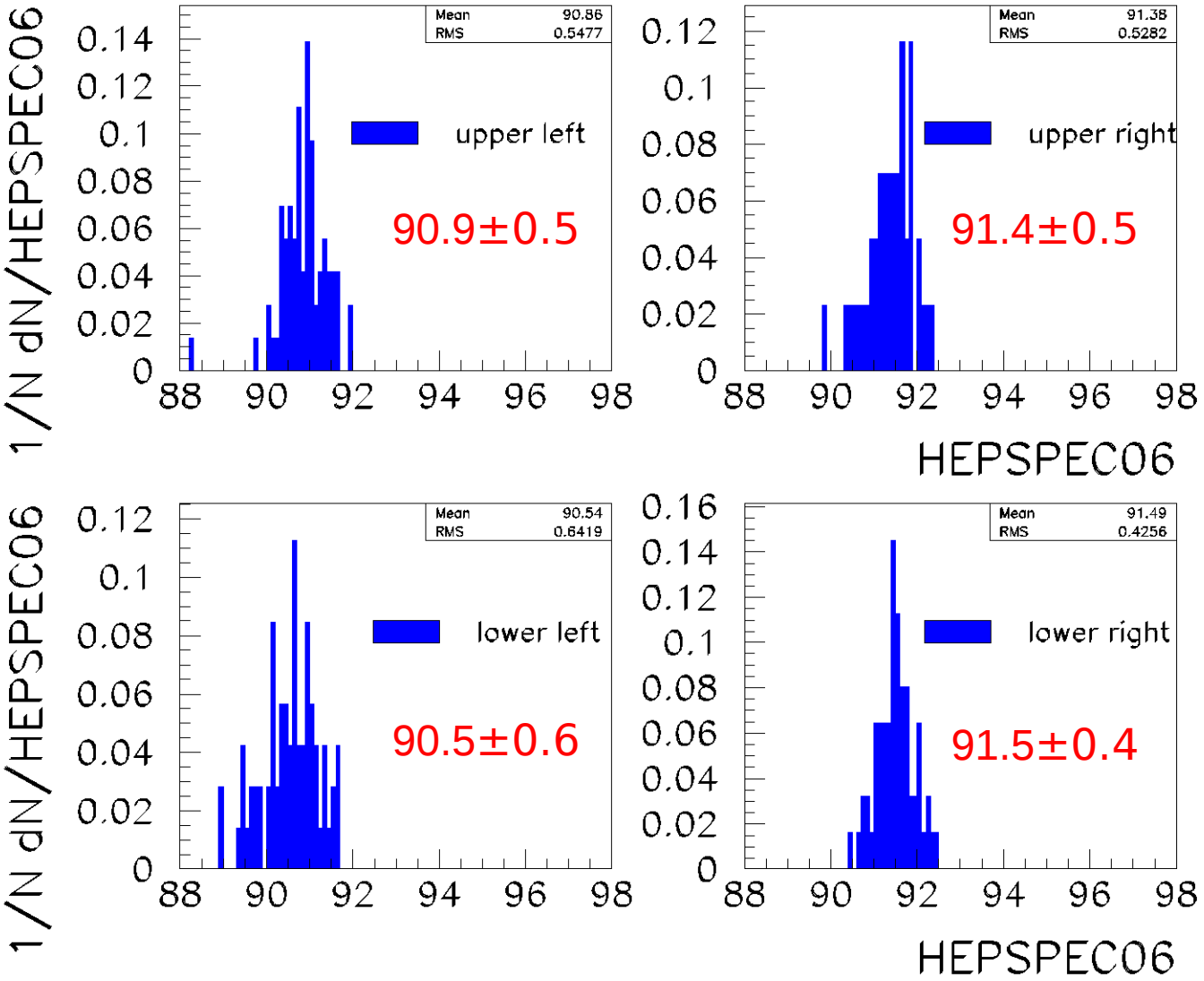
Test system enclosure rear view



Individual test systems

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Systematic errors: cont.



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Systematic errors: cont.

- Left-right asymmetry seen for this hardware type
  - Left hand side machines: **90.7±0.6**
  - Right hand side machines **91.5±0.5**

- Cross checked with a different twin2 system from a different vendor
  - Distributions are smaller
  - No left-right asymmetry seen !

NOT really significant, still … possible reasons:

- Statistics ?
- Real general hardware issue ?
- Bad cabling of the test system ?
- Just a bad sample box … ?

→ needs verification with final machines and more statistics

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Implications for procurement procedures

- Both SMT and Turbo mode increase the spread of the HEPSPEC06 measurements
- Not all CPUs support SMT and turbo mode

Recommendations:

- If the proposed CPUs support SMT and/or Turbo mode, these have to be switched off for the HEPSPEC performance evaluation for the bid
- For **production** purpose, and depending on the application to be run, **turbo mode should be switched on**
- Switching on SMT mode on batch worker nodes can help for multi-threaded applications but ...
  - No additional İncrease of job slot number (limited by disk/memory requirements)
  - Possible implications on licensing for commercial batch system solutions
  - Implications for CPU count publishing on the GRID
- Switching on SMT also for dedicated machines needs case by case study as it can be harmful

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Power measurements (preliminary)

Hardware: "small disk server" test machine with
- CPU and memory layout similar to the CPU servers
- Two (redundant) power supply units
- Hardware raid controller

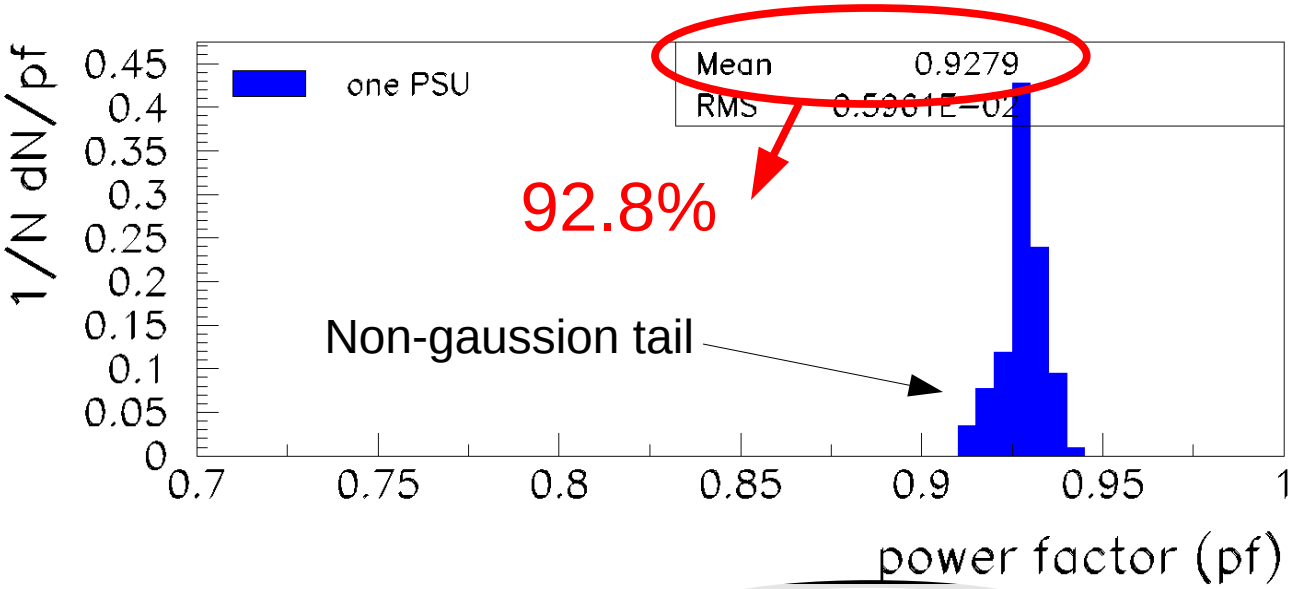
Power measurement unit:
- ZES/Zimmer Electronics LMG500, 8 channels
- 2 channels used (one per power supply)

- Measurement conditions
  - Machine idle
  - Each measurement is an average over 20min time

Main interest of this study:
shape of distributions and **variations of mean values,** rather than the absolute mean values

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Power factors: using one PSU only or both

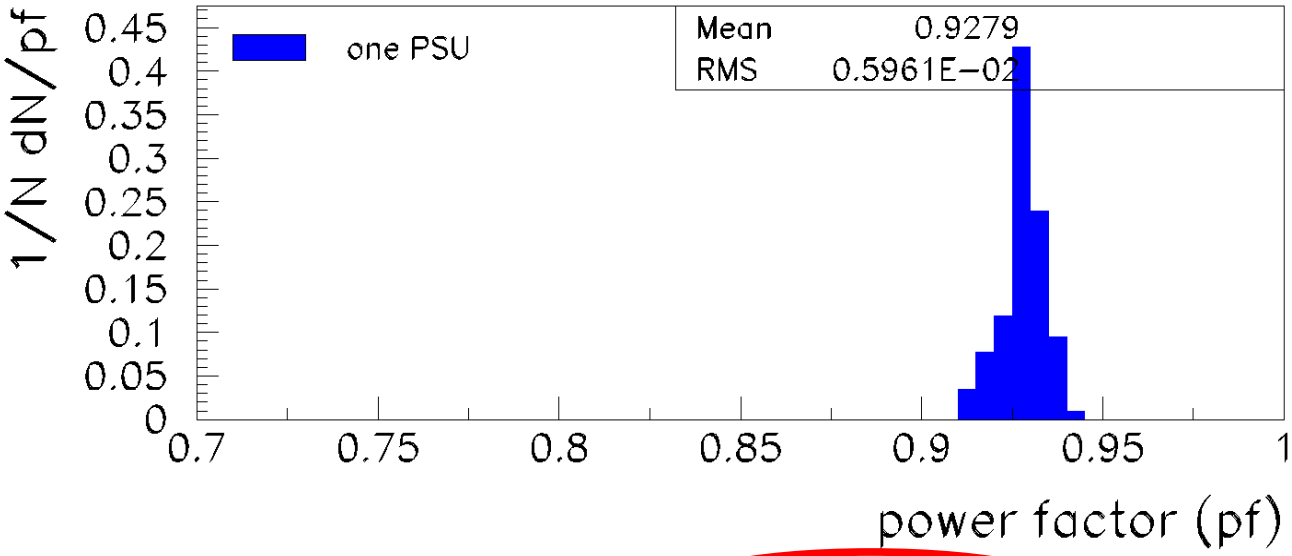


Ratio of effective and apparent power

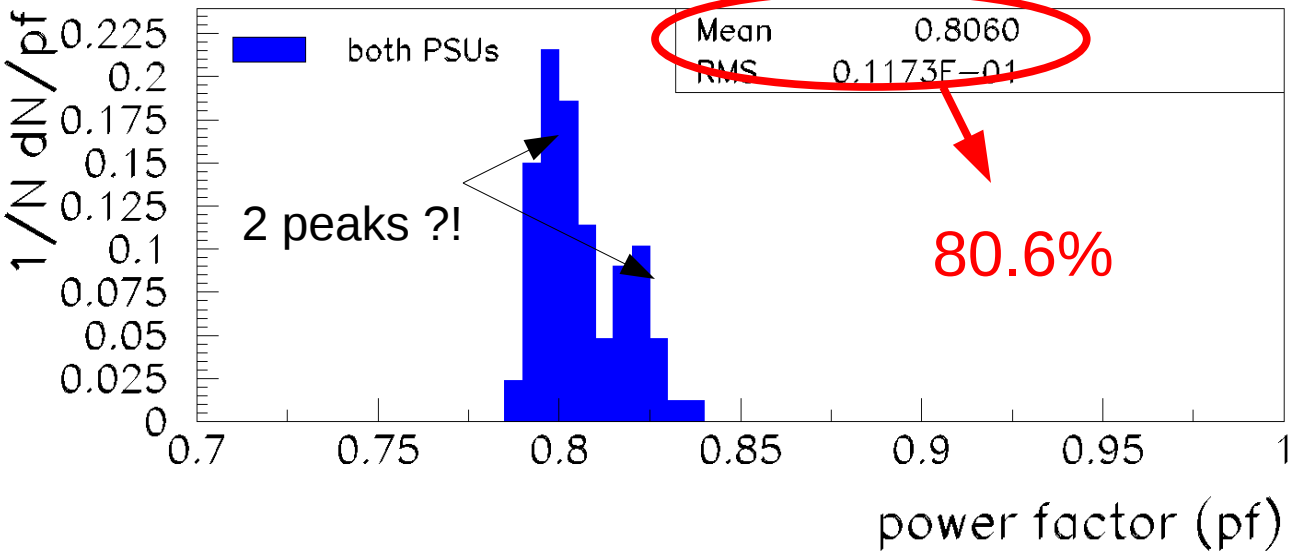
Ratio of the sum of effective and apparent power

CERN IT Departmer
CH-1211 Genève 2
Switzerland
www.cern.ch/it

Power factors: using one PSU only or both



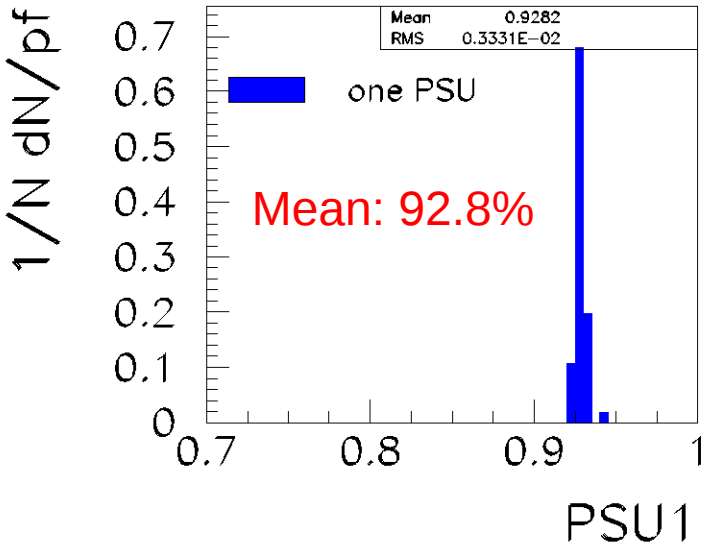
Ratio of effective and apparent power

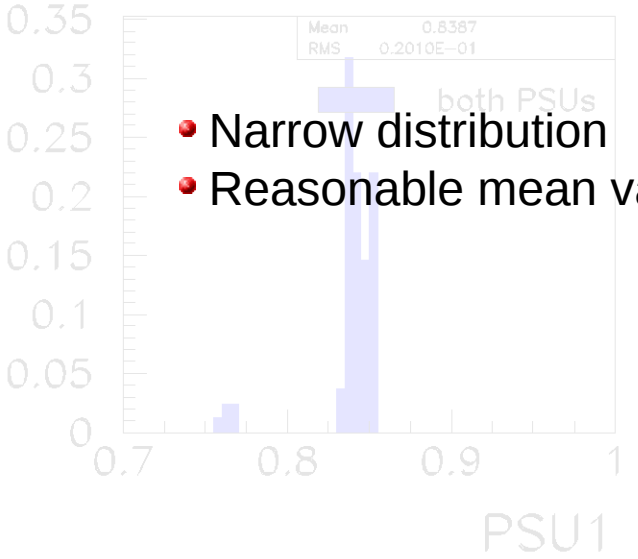


Ratio of the sum of effective and apparent power

CERN IT Departmer
CH-1211 Genève 2
Switzerland
www.cern.ch/it

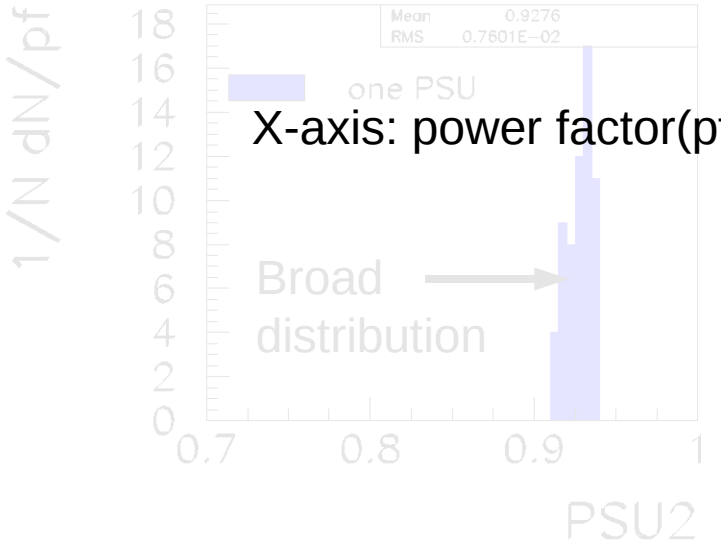
Power factors of the two power supplies





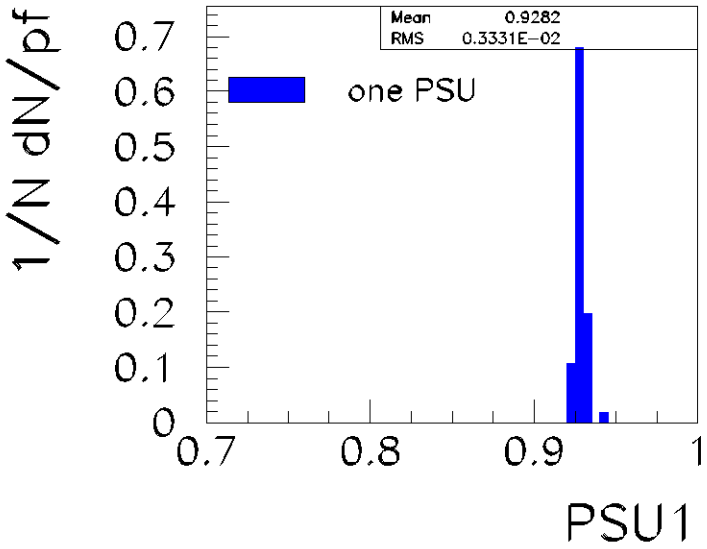
- Narrow distribution
- Reasonable mean value

X-axis: power factor(pf)

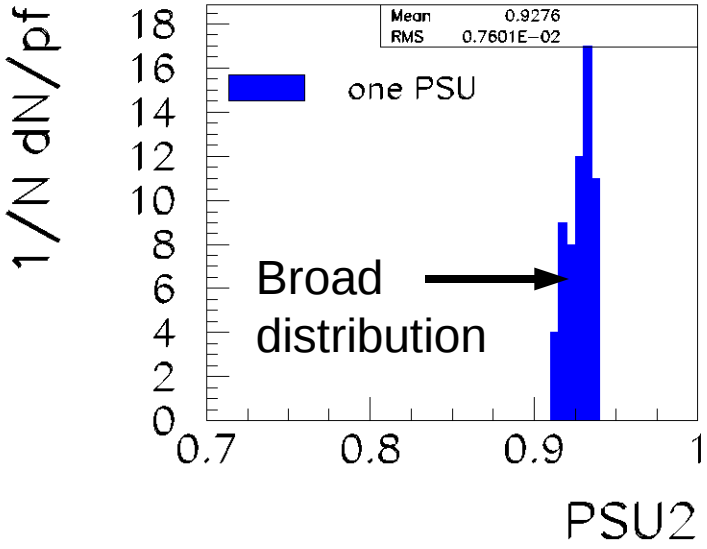


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Power factors of the two power supplies



X-axis: power factor(pf)



Using the other power supply

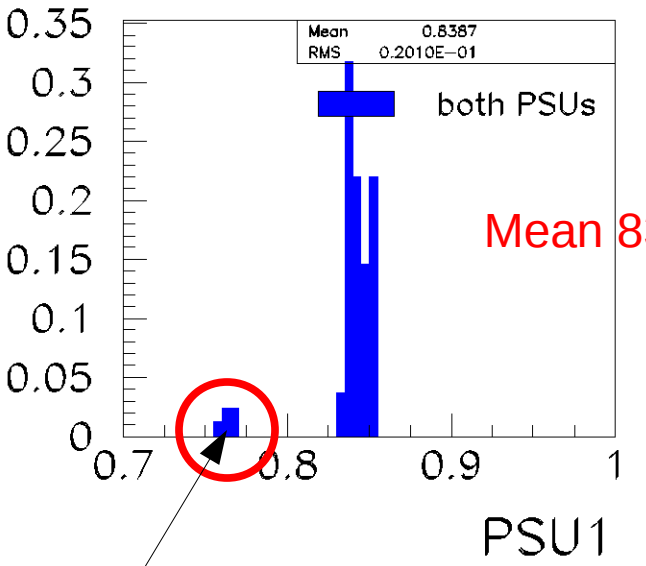
- Same mean value
- But twice broader distribution

Bad power supply ???

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

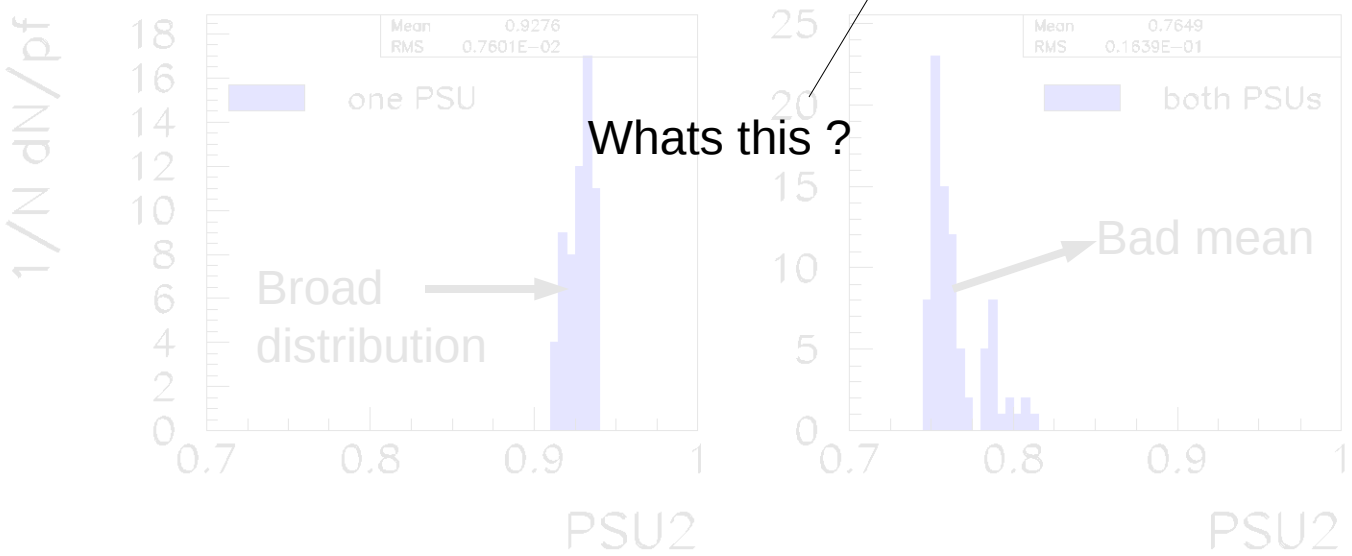
Power factors of the two power supplies

- Both units connected,
- Power factor PSU2
- Fairly narrow for PSU1
- Not really Gaussian shape



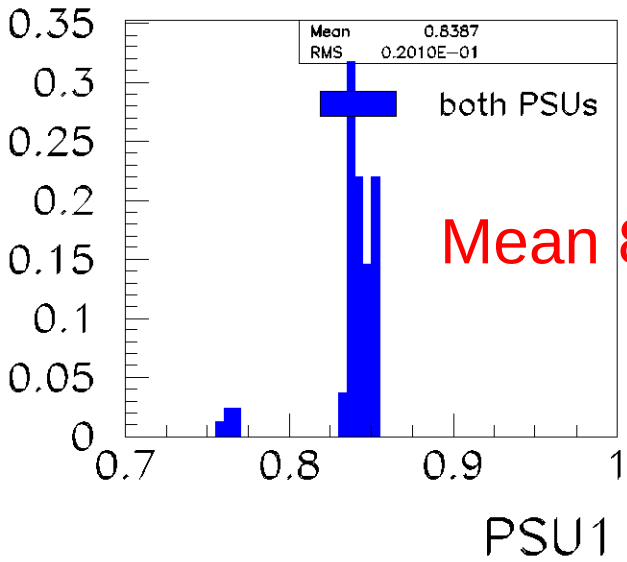
Mean 83.7%

Whats this ?



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

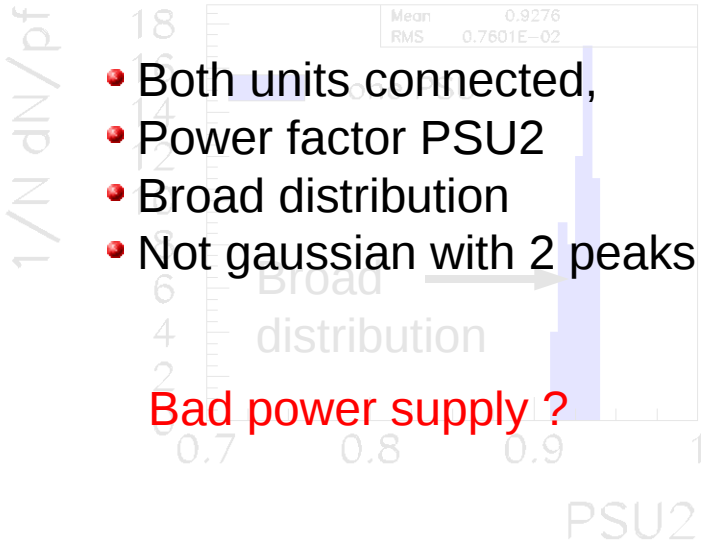
Power factors of the two power supplies

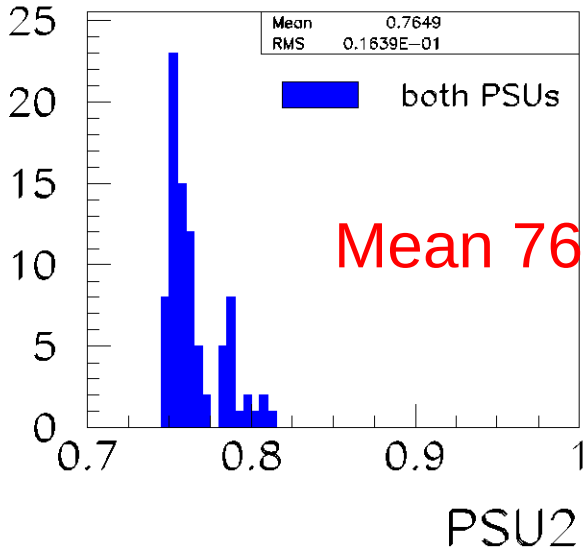


Mean 83.7%

- Both units connected,
- Power factor PSU2
- Broad distribution
- Not gaussian with 2 peaks !

Bad power supply ?

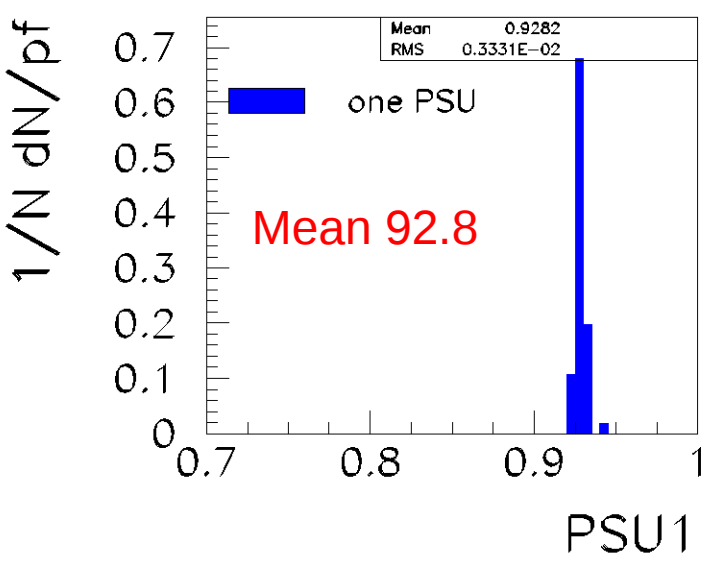


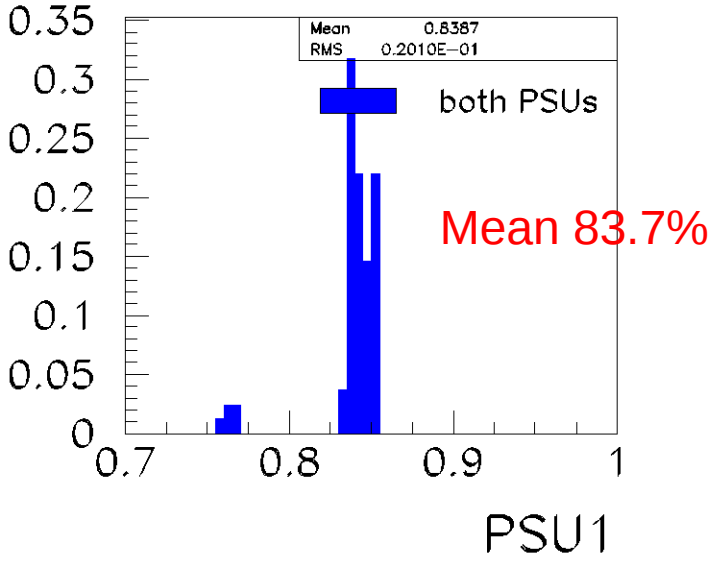


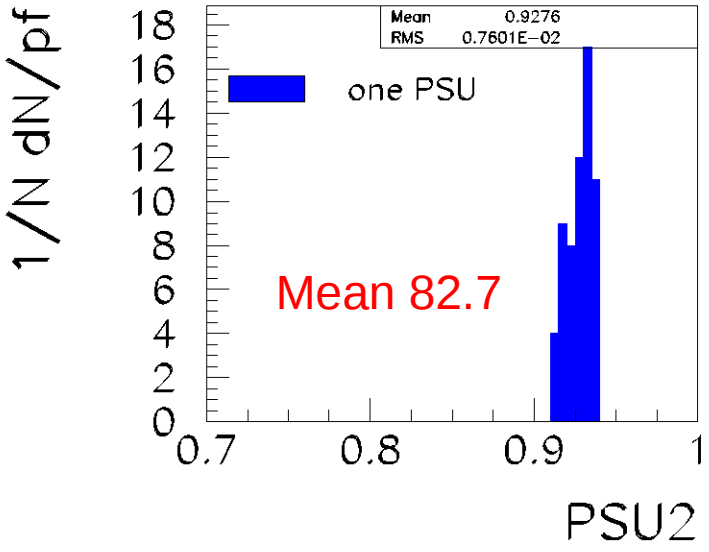
Mean 76.5%

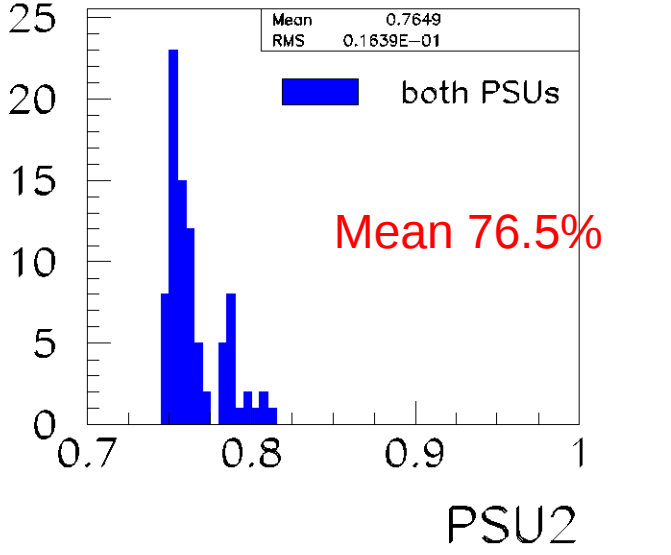
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Power factors of the two power supplies









CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Power measurements summary

Disclaimer:
The results are not very conclusive. Further verification is needed !!!!!

- Using only one power supply instead of both seems to be an advantage for power saving
- The two power supplies, although the same model, showed a very different performance
- The width of the distributions is non-negligible
- The shape of the distributions is non-gaussian

Lessons learned:

- Seen variations of up to 4% in apparent and effective power measurements
- Error estimation difficult due to non-gaussion distributions
- Significant power factor variations between PSU of the same model
- Power factor variations of several per-cent points are not unlikely

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Summary

- Up to **(27±2) %** performance gain by **SMT** and **turbo mode**, in terms of HEPSPEC06
- **Scaling** in terms of HEPSPEC06 remains **below expectations**
- Hints of systematic performance variations, depending on the hardware type
- Enhanced statistics power measurements are not very conclusive but interesting nevertheless

All presented results may have implications for procurement procedures.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it