

# Benchmarking up-to-date x86 processors

Helge Meinhard / CERN-IT

HEPiX Spring 2009 Umea

26 May 2009

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

# Introduction



- Motivation
  - About 2.5 years ago, CERN started to tender for aggregate performance rather than boxes
  - About 1.5 years ago, CERN started to add an adjudication element based on power consumption
  - Massive tenders coming up (see CERN site report)
  - We need to know what to expect
  - Performance and power consumption now used for other types of machines as well
  - Also: curiosity…
- Scope: primarily worker nodes
  - Some applications for disk servers, midrange servers, …

# A word of warning…



- Results presented on these slides have been obtained with (next-to-production) beta samples
- Some results are inconsistent and need checking (in particular those obtained with gcc 4.3)
- Consider all numbers preliminary, and take them with a grain of salt…
  - And as the efficiency depends on the performance and the power consumption, take it with two (or square root thereof) grains of salt…

# Benchmarks used

- Performance
  - Previously: SPECint2000 with strict boundary conditions, phased out
  - Now: HEP-SPEC06 (based on all_cpp subset of SPECcpu 2006)
    - Collection of results in https://hepix.caspur.it/benchmarks
- Power consumption: VA in primary AC circuit
  - Idle: Just running Linux OS
  - Full load: half the cores run cpuburn, the other half runs linpack with optimal memory utilisation
  - For worker nodes, use 20% idle / 80% loaded weighted average

# Baseline



- Most recent acquisition of worker nodes
  - Supermicro twins with high-efficiency 780 W PSU
  - Per system:
    - Board with San Clemente chipset (X7DCT), BMC daughter card
    - Two low-voltage Harpertowns L5420 (2.5 GHz)
    - 4 modules of 4 GB DDR2
    - 2 hard drives SATA 7.2 krpm of 160 GB (3.5”)
  - Performance per system: 67.75 HEP-SPEC06
    - FZK measurement – ours were about 20% lower – to be investigated
  - Power consumption per twin: 410 VA
  - Power efficiency: 2 * 67.75 HEP-SPEC06 / 410 VA = 0.33 HEP-SPEC06 / VA

# Test systems



- Platform 1: Proprietary format for two Xeon 5500 (Nehalem EP) processors, equipped with 2 x E5530 (2.4 GHz), 12 DDR-3 modules of 2 GB each
  - Used for performance only
- Platform 2: Intel EPSD 1U system for two Xeon 5500 (Nehalem EP) processors, equipped with 2 x E5540 (2.53 GHz), 12 DDR-3 modules of 2 GB each, one SATA 3.5” 7.2 krpm drive
  - Used for performance and power measurements
- Additional Nehalem processors: L5520 (2.27 GHz), X5570 (2.93 GHz)
- Platform 3: Supermicro twin with two systems each with two AMD Opteron (Shanghai) HE processors, 8 x 2 GB DDR-2 modules, 2 SATA 3.5” 7.2 krpm drives
  - Used for performance and power measurements

# Standard running conditions



- System installed with SLC 5.3 (kernel 2.6.18-128.1.1.el5), x86_64 version
- Full memory configuration as delivered (e.g. 24 GB for Nehalem, 16 GB for Shanghai)
- For Nehalem: SMT off, Turbo on
- SPEC CPU 2006 compiled with system compiler gcc 4.1.2 as 32-bit application running with compatibility libraries
- 8 concurrent streams

# Standard running conditions

CERN IT Department

| Platform | Processor | HEP-SPEC06 |
|---|---|---|
| Baseline | 2 x L5420 2.5 GHz | 67.75 |
| 1 | 2 x E5530 2.4 GHz | 100.17 |
| 2 | 2 x E5540 2.53 GHz | 103.22 |
| 3 | 2 x 2376HE 2.3 GHz | 70.47 |

# Change of OS / compiler

CERN IT Department

| Platform | OS | Compiler | HEP-SPEC06 |
|---|---|---|---|
| 1 | SLC5 | gcc 4.1.2 | 100.17 |
| 1 | SLC4 (*) | gcc 3.4.6 | 84.88 |
| 2 | SLC5 | gcc 4.1.2 | 103.22 |
| 2 | SLC5 | gcc 4.3 | 106.45 |
| 3 | SLC5 | gcc 4.1.2 | 70.47 |
| 3 | SLC4 | gcc 3.4.6 | 67.33 |
| 3 | SLC5 | gcc 4.3 | 70.32 |

(*) SMT was on, but the SLC4 kernel does not recognise virtual cores anyway

# 32-bit vs 64-bit applications

CERN IT Department

| Platform | OS | Compiler | Application | HEP-SPEC06 |
|---|---|---|---|---|
| 1 | SLC5 | gcc 4.1.2 | 32 bit | 100.17 |
| 1 | SLC5 | gcc 4.1.2 | 64 bit | 111.76 |
| 3 | SLC5 | gcc 4.1.2 | 32 bit | 70.47 |
| 3 | SLC5 | gcc 4.1.2 | 64 bit | 79.84 |
| 3 | SLC5 | gcc 4.3 | 32 bit | 70.32 |
| 3 | SLC5 | gcc 4.3 | 64 bit | 80.46 |
| 3 | SLC4 | gcc 3.4.6 | 32 bit | 67.33 |
| 3 | SLC4 | gcc 3.4.6 | 64 bit | 65.16 |

# SMT on vs off

CERN IT Department

| Platform | SMT | HEP-SPEC06 |
|---|---|---|
| 1 | off (8 streams) | 100.17 |
| 1 | on(16 streams) | 124.69 |
| 2 | off (8 streams) | 103.22 |
| 2 | on (16 streams) | 125.99 |

# Turbo on vs off



| Platform | Turbo | HEP-SPEC06 |
|---|---|---|
| 2 | on | 103.22 |
| 2 | off | 100.31 |

I have seen results suggesting that with different compilers (gcc 4.3, icc 11) or with 64-bit applications, the difference could be larger (order 10% or more)

# Different processor SKUs

CERN IT Department

| Platform | Processor | Speed / MHz | HEP-SPEC06 | HEP-SPEC06 / Speed |
|---|---|---|---|---|
| 2 | 2 x E5540 | 2530 | 103.22 | 0.041 |
| 2 | 2 x L5520 | 2270 | 94.25 | 0.042 |
| 2 | 2 x X5570 | 2930 | 118.21 | 0.040 |

# Different memory configurations

CERN IT Department

| Platform | Processor | Memory | HEP-SPEC06 |
|---|---|---|---|
| 1 | 2 x L5520 | 12 x 2 GB | 96.28 |
| 1 | 2 x L5520 | 8 x 2 GB | 94.32 |
| 1 | 2 x L5520 | 6 x 2 GB | 96.69 |

# Power consumption and efficiency

CERN IT Department

| Processor | Memory | Configuration | HEP-SPEC06 | Power | Efficiency |
|---|---|---|---|---|---|
| 2 x E5540 | 12 x 2 GB | Standard | 103.22 | 281.0 | 0.37 |
| 2 x E5540 | 8 x 2 GB | Standard | (103.22) | 262.8 | 0.39 |
| 2 x E5540 | 12 x 2 GB | SMT on | 125.99 | 294.3 | 0.43 |
| 2 x E5540 | 12 x 2 GB | Turbo off | 100.31 | 267.9 | 0.37 |
| 2 x L5520 | 12 x 2 GB | Standard | 94.25 | 245.1 | 0.38 |
| 2 x L5520 | 8 x 2 GB | Standard | (94.25) | 229.1 | 0.41 |
| 2 x X5570 | 12 x 2 GB | Standard | 118.21 | 320.9 | 0.37 |
| 2 x 2376HE | 8 x 2 GB | n/a | 70.47 | 252.6 | 0.28 |

Baseline (2 x L5420) had an efficiency of 0.33 or lower (to be confirmed)

# Summary



- For standard HEP applications, Nehalem offers a significant improvement in performance and performance/Watt
  - I expect Nehalem to play an important role in the replies to our current massive calls for tender
- Shanghai cannot compete (not even with Harpertown)
  - Curious to benchmark an Istanbul-based system
- With Nehalem-EP, phase space has become complex
  - Cache configurations, memory configurations, memory speed,…
- Problem persists that E processors, for which L versions exist, don't have a reproducible power consumption
- Measurements to be repeated and extended
  - Submit results to Caspur site https://hepix.caspur.it/benchmarks

# Acknowledgements


- Measurements, scripts, documentation:
  - Alex Iribarren
  - Manfred Alef
  - Ulrich Schwickerath
  - Julien Leduc
  - Axel Busch
- Hosting of Wiki:
  - Caspur team (Andrei Maslennikov, Monica Calori)
  - Alex Iribarren
- Test hardware:
  - Some interested vendors