64-Bit Opteron systems in High Energy and Astroparticle Physics

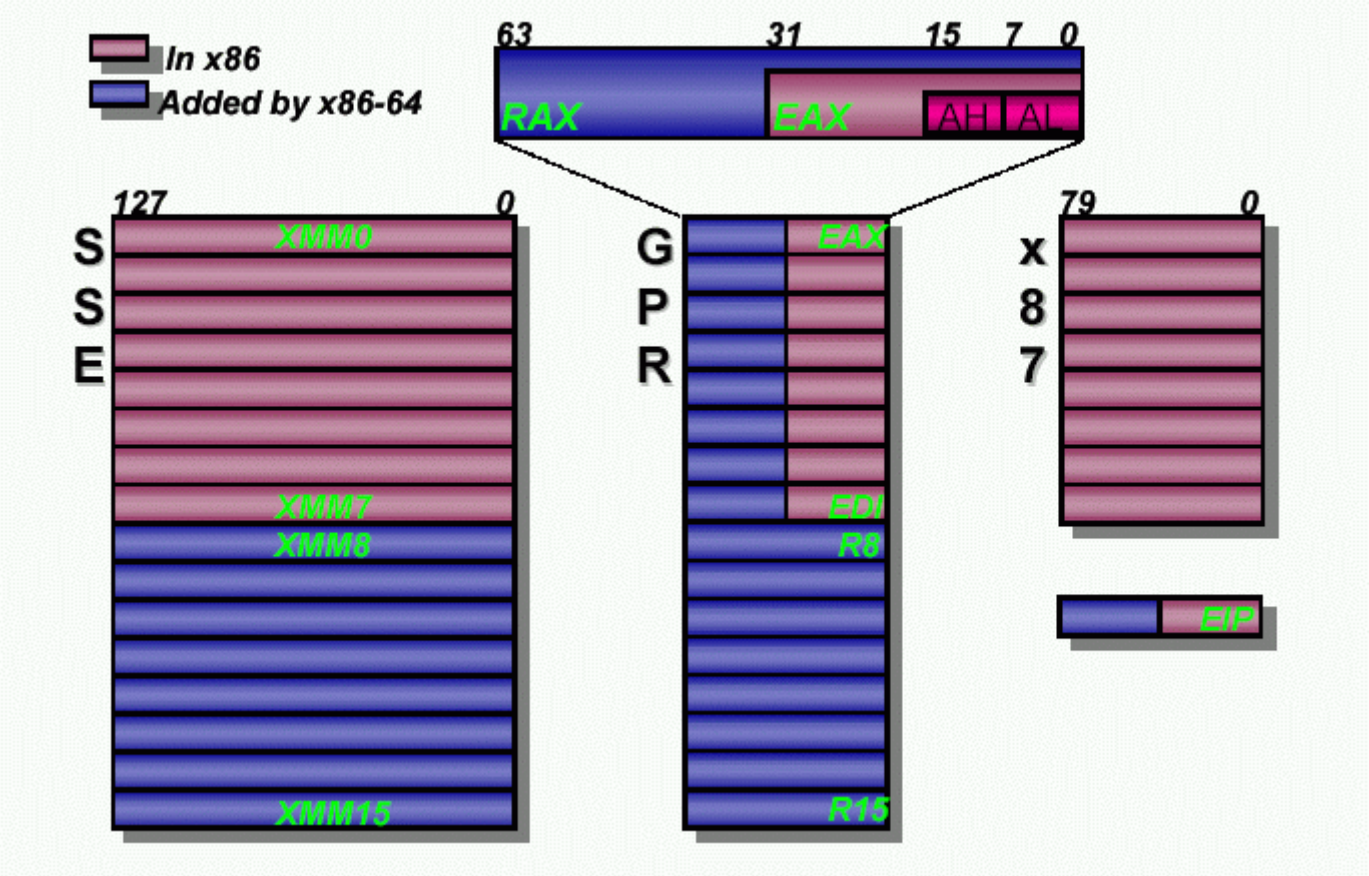
Peter Wegner, Stephan Wiesand

Outline

- brief introduction to AMD64
  - why these systems are fast
  - why they are faster in 64bit mode
  - where EM64T differs
- performance figures for physics applications
  - ROOT, Pythia, FORM, Sieglinde, LQCD
  - comparison with i386 and EM64T systems
- how to handle and use these systems
  - 64bit Linux distributions
  - common problems

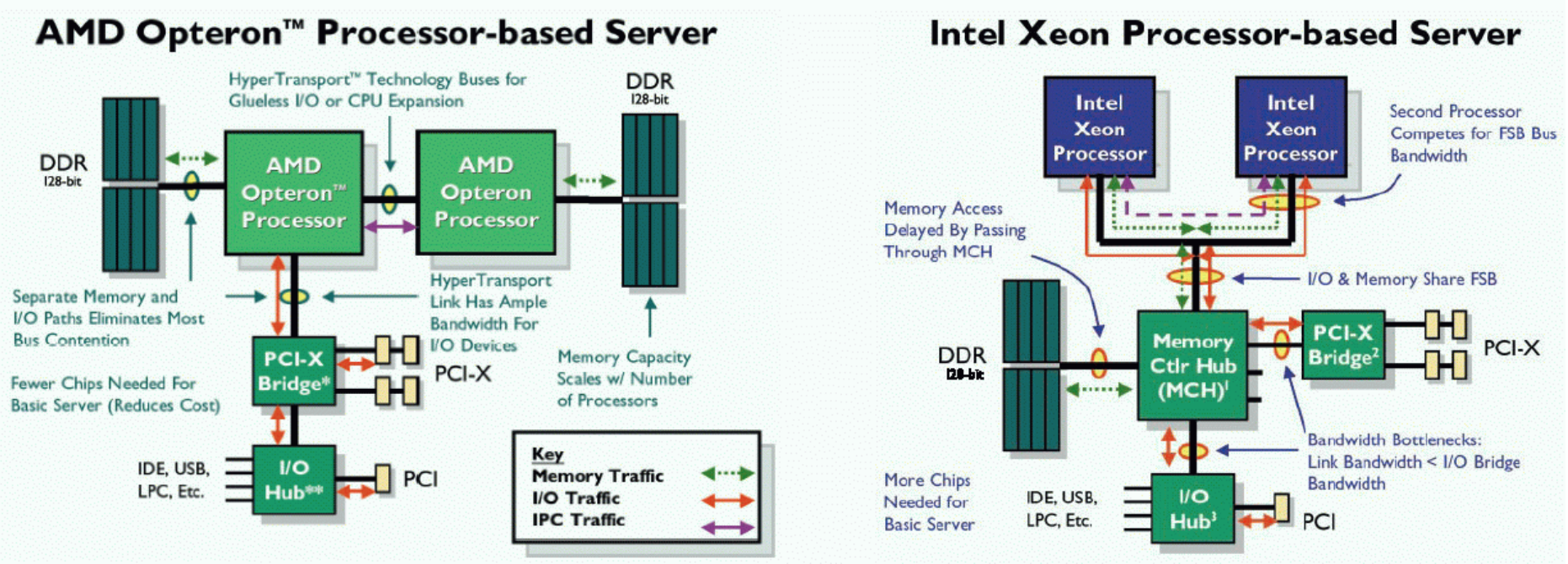
AMD64 Instruction Set Architecture

- extends the x86 CISC architecture
- executes i386 instructions in hardware
- legacy mode:
  - 32bit OS/apps
  - x86 register set
- long mode:
  - 64bit OS
  - 32/64bit apps
  - new registers
  - reduced PIC penalty



- Opteron/Athlon64: SSE2, 3dNow!
- Xeon/Pentium 4: SSE2, SSE3

System Architectures



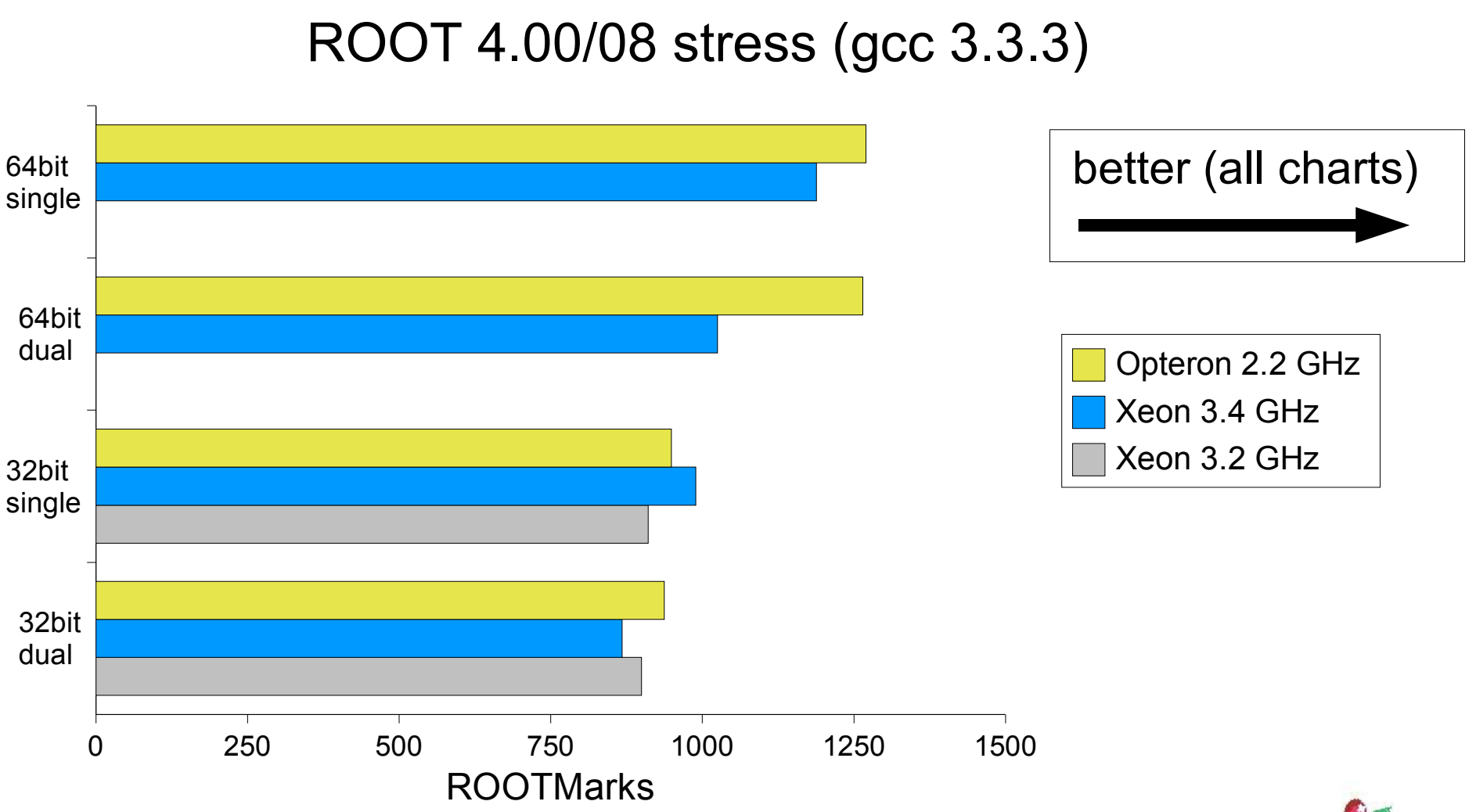
- Opterons have memory interfaces on chip
- memory bandwidth scales with number of CPUs
- but memory now more or less close to processes using it
- requires NUMA support from kernel for best efficiency

Hardware for Performance Comparisons

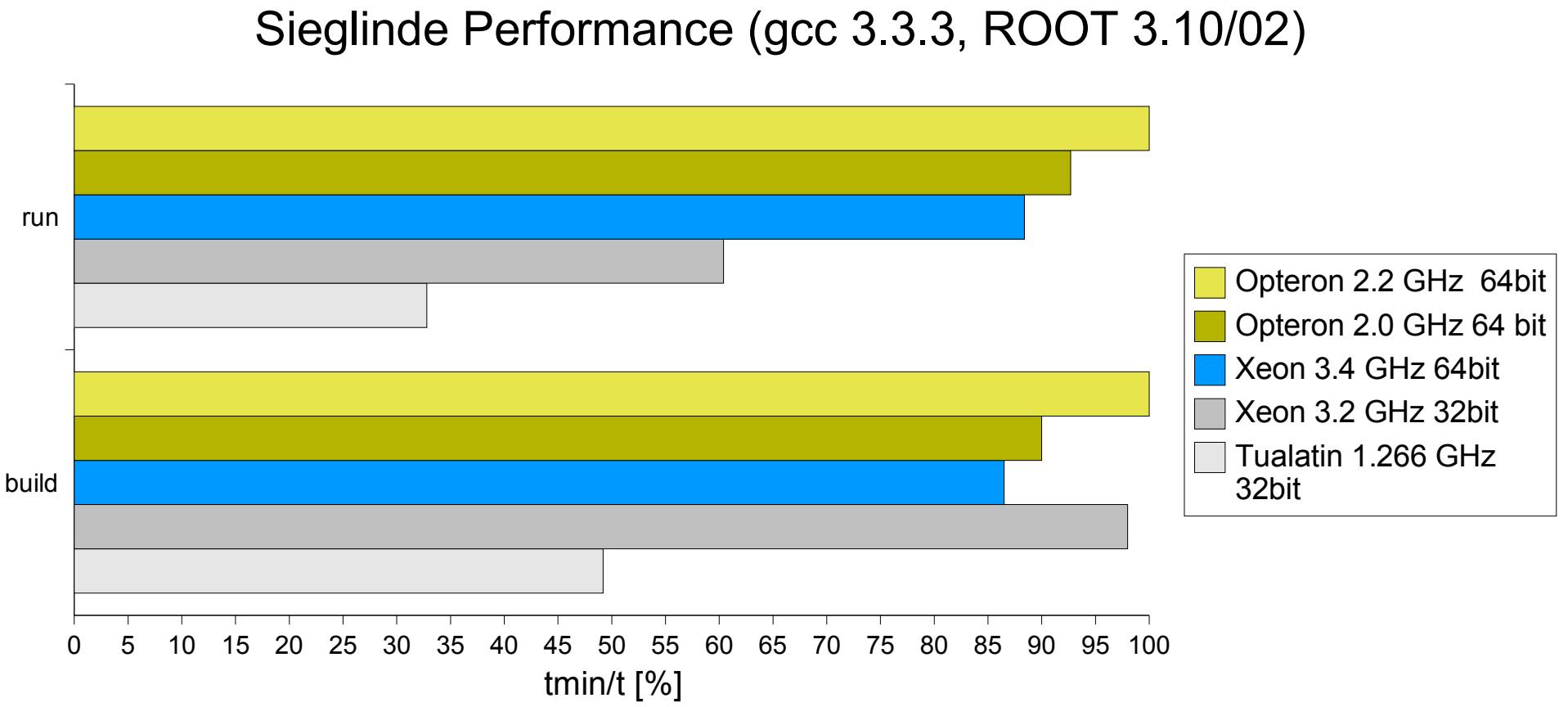
- All equipped with 2 CPUs and SCSI disk:
  - Opteron 2.0 GHz: IBM eServer 325, 4 GB
    - SuSE 9.0 professional, kernel 2.4.21-215-smp
  - Opteron 2.2 GHz: Sun Fire V20z, 4GB
    - SuSE 9.0 professional, kernel 2.4.21-231-smp
  - Xeon 3.4 GHz: Supermicro 7044H-X8R, 4GB
    - SuSE 9.1 professional, kernel 2.6.4-52-smp
  - Xeon 3.2 GHz: Sun Fire V65x, 2 GB
    - SuSE 8.2 professional, kernel 2.4.26
  - Tualatin 1.266 GHz: Supermicro 6013H, 1GB
    - SuSE 8.2 professional, kernel 2.4.25

CHEP'04
INTERLAKEN

ROOT Performance

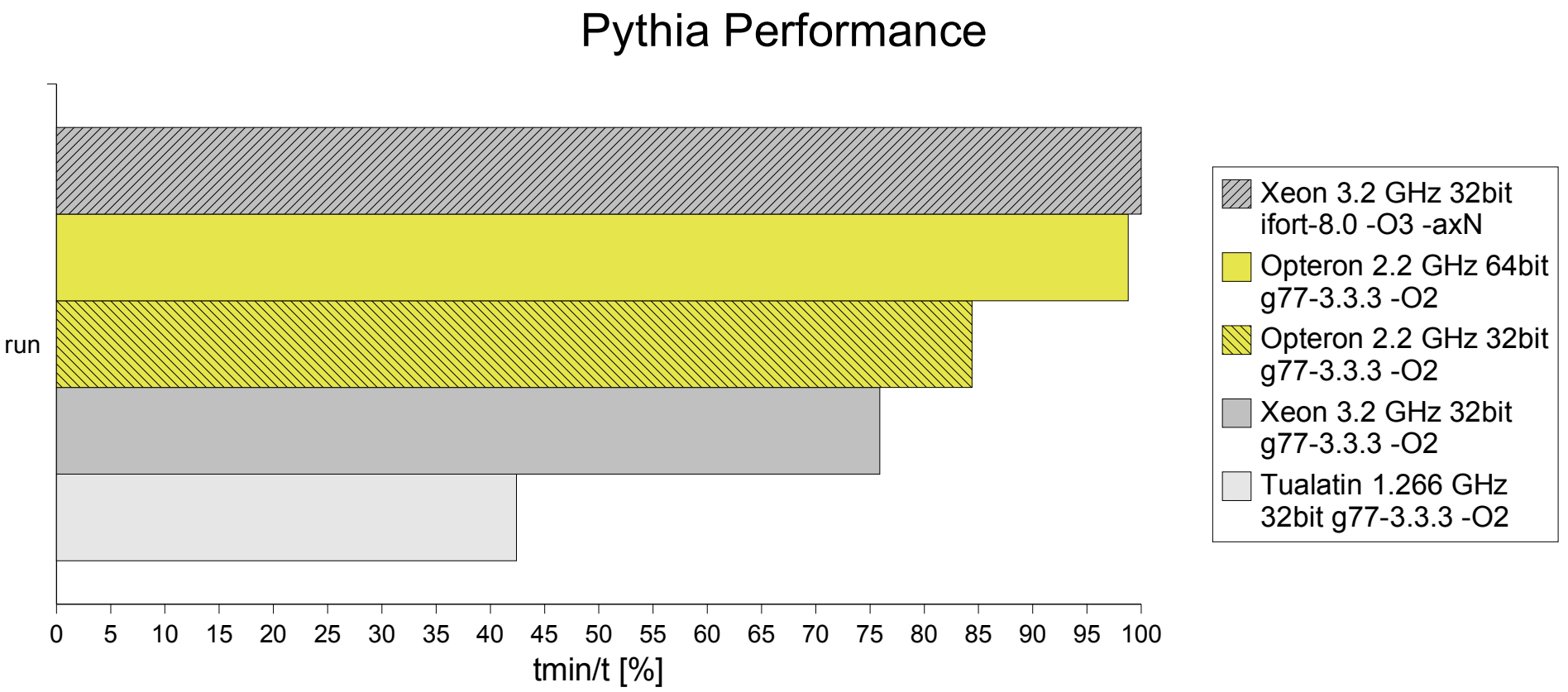


Sieglinde Benchmark



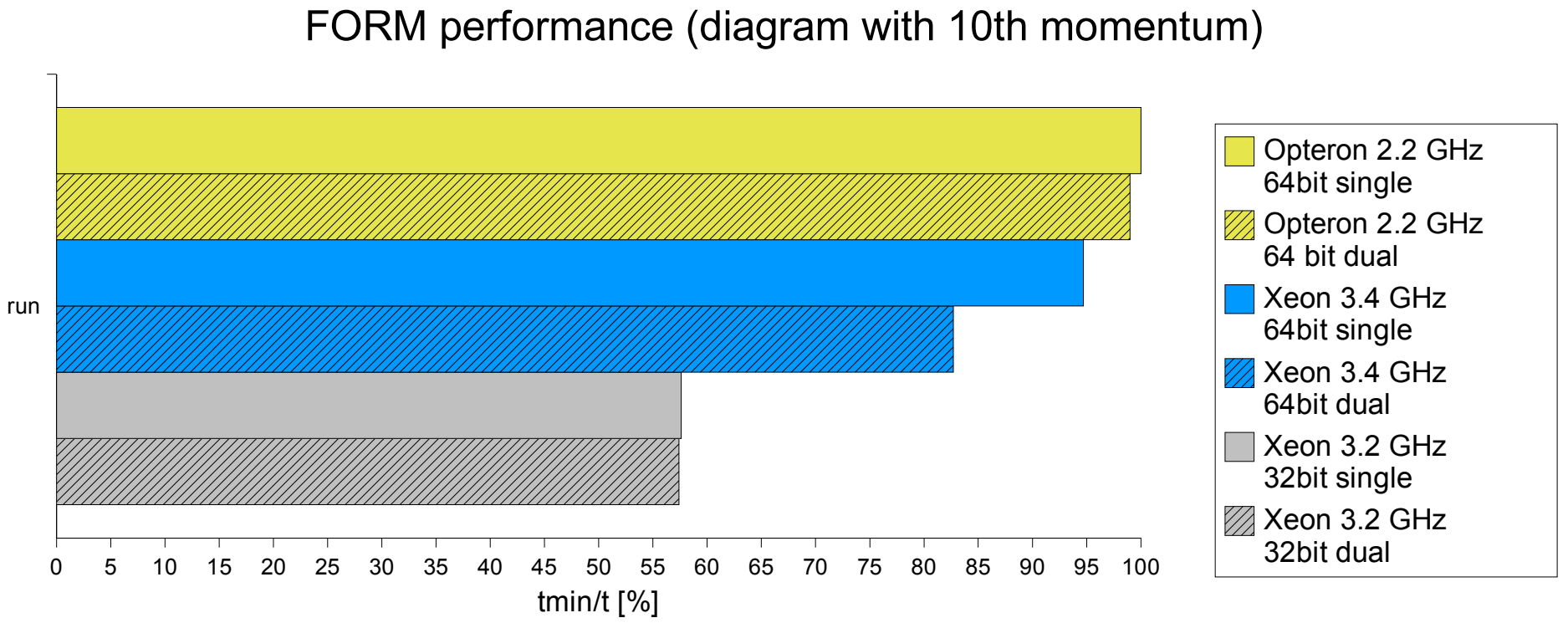
- Amanda experiment's neutrino reconstruction / filtering software
- single process, but uses a MySQL server on same host
- software made available by Peter Nießen, Univ. of Delaware

Pythia 6.2



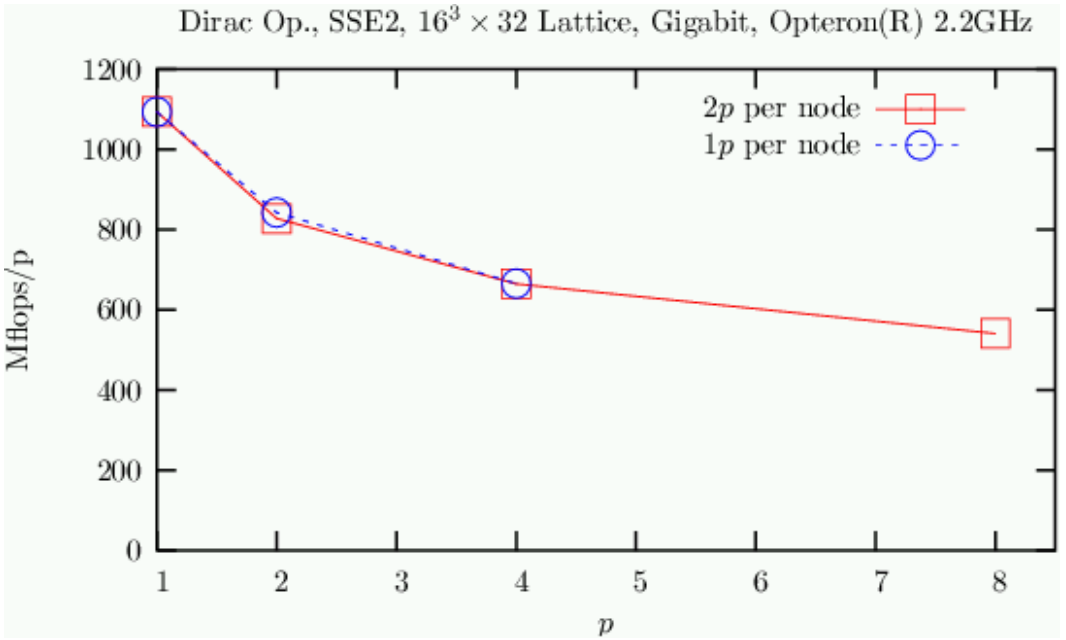
- Pythia 6.2 example 4
  "study of W mass shift by colour rearrangement at LEP 2"
- ifort executable refused to run on Opterons

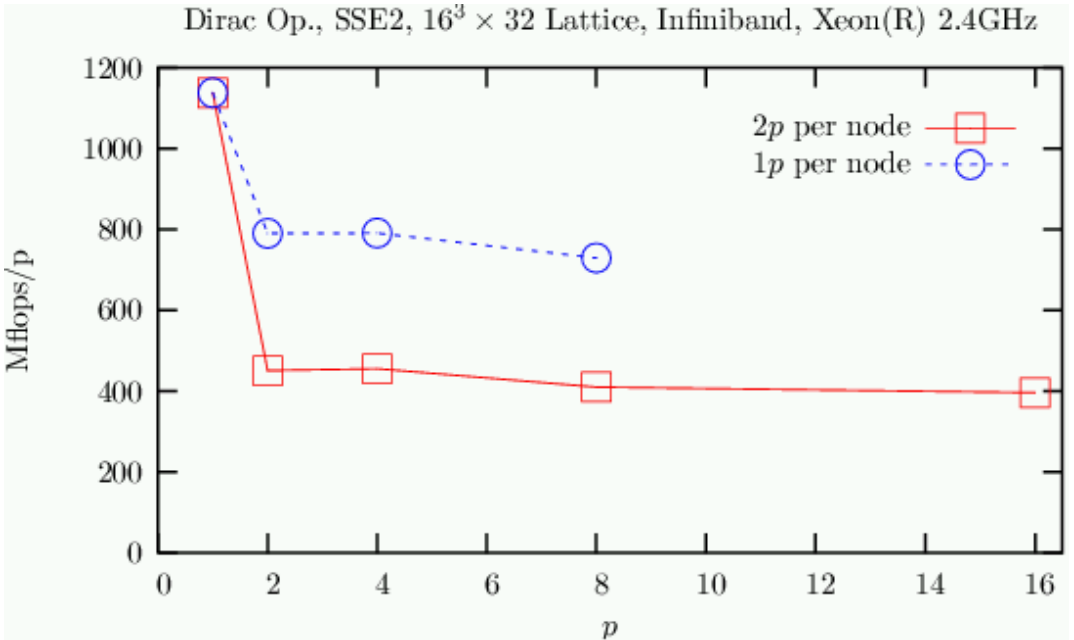
FORM 3.1



- symbolic formula manipulation, C, huge data sets
- implements own “paging” of data to disk
- 64bit executable built by author J. Vermaseren on DESY test system
- 32bit executable built with icc (www.nikhef.nl/~form)

Dual/Single CPU Performance in Clusters





- measurements by C. Urbach, FU Berlin
- 32bit Lattice QCD, MPI
- performed on clusters with Gigabit Ethernet and Infiniband interconnects
  - not our test systems
- p = number of processes

Handling Linux on AMD64 (EM64T)

- system looks, feels and behaves like “a linux PC”
  - BIOS, boot loaders, installation procedures are all unchanged
- most common porting problem: longs and pointers are 64 bit
  - 64bit executables can't use 32bit libraries and shared objects
    - cernlib is not available for this platform (yet?)
- decent distributions are available:

	SuSE			Red Hat		SL
	9.0	9.1	SLES8	SLES9	FC2	EL3	3.0.2
free download	yes	yes	no	no	yes	source	yes
kernel	2.4	2.6	2.4	2.6	2.6	2.4	2.4
stable AFS client	yes	no	yes	no	no	yes	yes
NUMA kernel	yes	yes	yes	yes	yes	no	no
full 64bit port	yes	yes	yes	yes	yes	yes	yes
32bit compat packages	yes	yes	yes	yes	(1)	(1)	(1)
works well	yes	yes	?	?	?	?	(2)

(1) Red Hat installs the packages from the i386 distribution in addition
(2) having trouble with AFS/krb5 integration

Conclusions

- 64bit commodity hardware is a reality
- performs very well in our HEP/Astroparticle physics applications
  - even in 32bit legacy mode
  - even better in 64bit long mode
  - CPU utilization on SMP systems is very good for Opterons
- Linux handling is identical to 32bit systems
  - no extra skills required
  - providing 32bit compatibility is some overhead
  - ecosystem is in place (distributions, compilers, libraries)
    - few problems remain to be solved