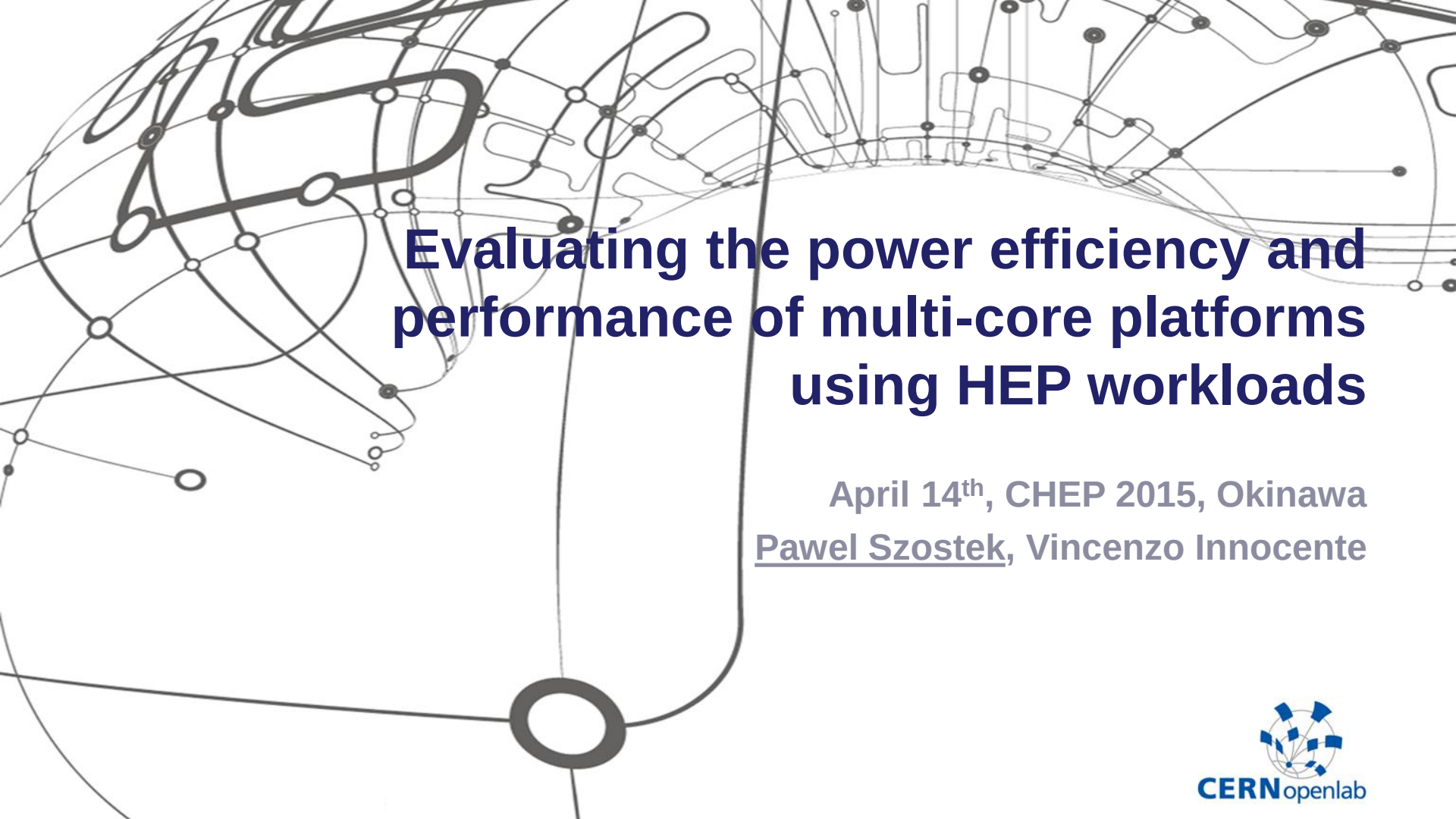


Evaluating the power efficiency and performance of multi-core platforms using HEP workloads

April 14th, CHEP 2015, Okinawa

Pawel Szostek, Vincenzo Innocente

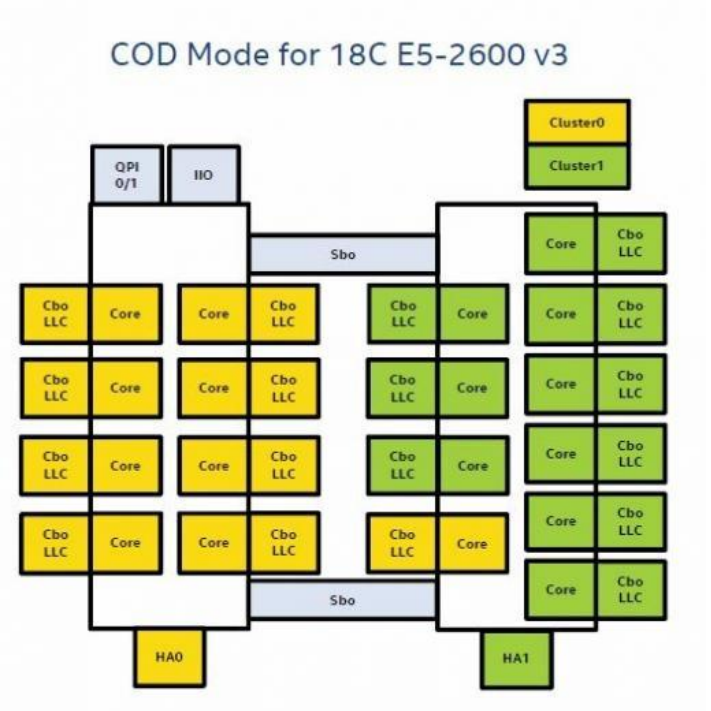
CERN openlab

Outline

- “Haswell-EP” microarchitecture refresher
- Description of the tested setup
- Experiments and results: scalability, vectorization, single core performance, power efficiency
- Conclusions

“Haswell-EP” microarchitecture refresher

- **dual-socket** platform, 22nm, AVX2 256b vectors, DDR4 support
- 10-core and higher core counts come with **two memory controllers** (Cluster-on-Die),
- beefed-up core (already in single socket HSW):
  - 8 instead of 6 execution units
  - bigger OoO execution buffer (192 vs. 168 entries)



“Haswell-EP” microarchitecture refresher (II)

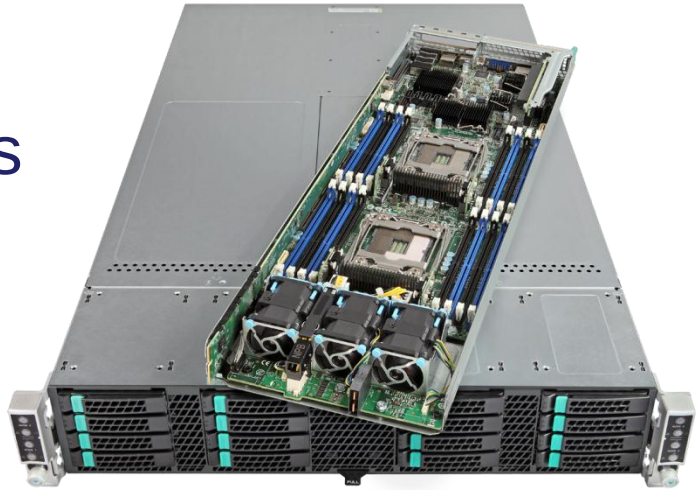
- comes with a couple of interesting power-saving features:
  - frequency/voltage scaled on a per-core basis
  - uncore frequency/voltage scaling independent from cores
  - configurable TDP
  - energy efficient turbo mode
- As in the single socket, AVX slows down the clock. Switching to AVX is penalized

Setup for the tests

- **Haswell-EP (dual socket):**
  - E5-2699v3 – 18 cores, 2.3Ghz, 145W TDP
  - E5-2698v3 – 16 cores , 2.3GHz, 135W
  - E5-2683v3 – 14 cores, 2.0GHz, 120W
- **Ivy Bridge-EP:**
  - E5-2695v2 – 12 cores, 2.4Ghz,115W
- **Sandy Bridge-EP:**
  - E5-2690 – 8 cores, 2.9GHz, 135W

Setup for the tests (II)

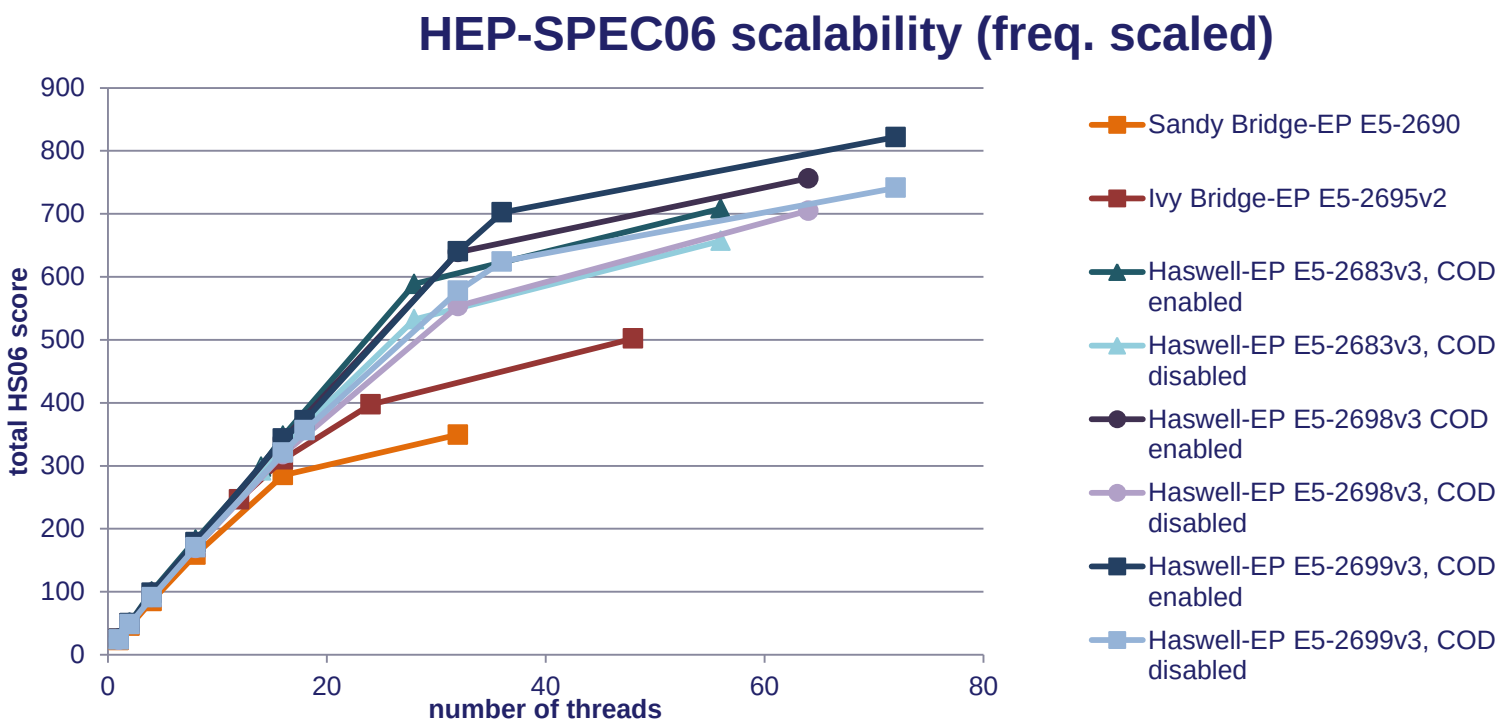
- 2U Intel Bobcat Pass chassis with 4 servers in an enclosure
- 8x8GB 2133MT/s DDR4 DIMMs
- 2xIntel SSD DC S3500 240GB, with LVM stripping
- Turbo Boost disabled, SMT enabled, P- and C- states disabled unless stated otherwise
- SLC 6.6 with the 2.6.32-504.12.2



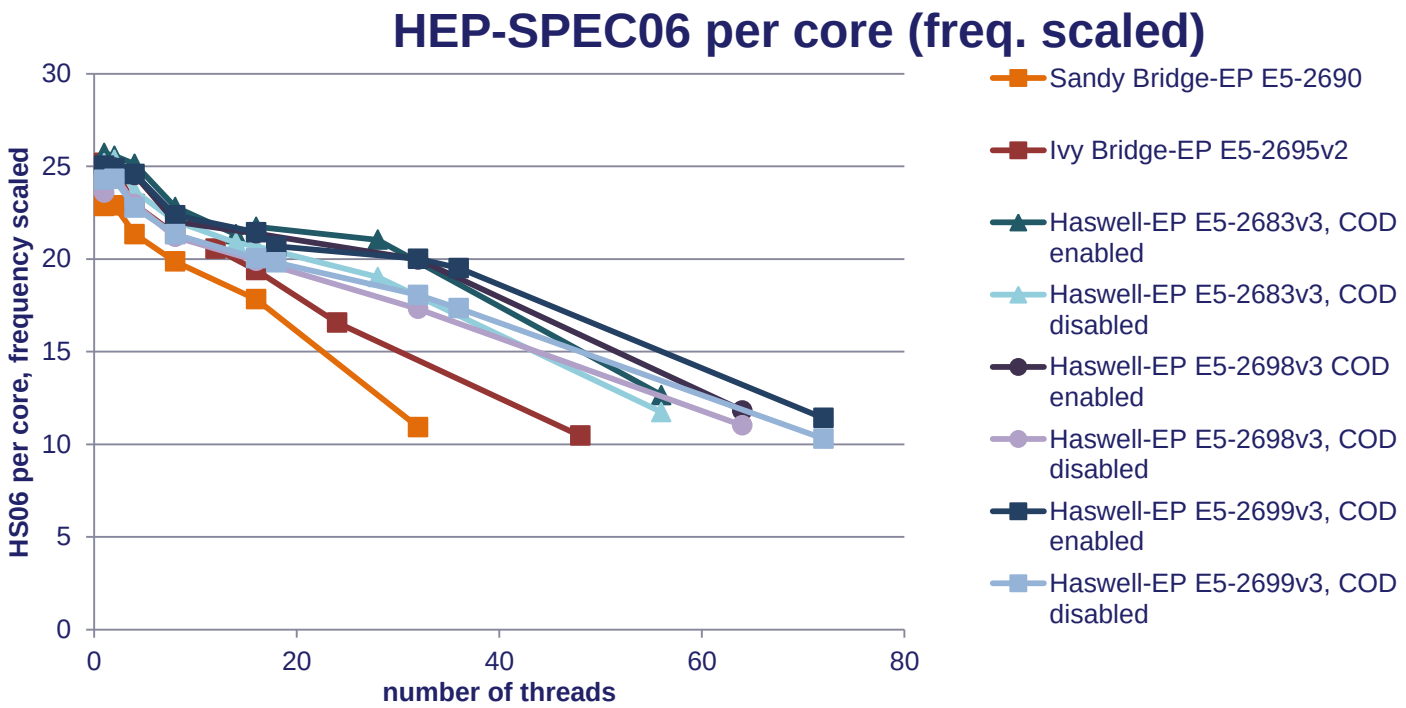
The benchmarks

- **Standard:** HEPSPEC06
  - not optimized for next-gen hardware
- **Analysis:**
  - MLFit
    - Threaded (pthreads, MPI, OpenMP, TBB)
    - Vectorized (Cilk+)
  - VIFit (Vincenzo Innocente Fit)
    - lightweight version of Mlfit
    - NUMA-aware memory management
- **Simulation:** up-to-date ParFullCMS with Geant 4 v.10.01 patch 01 [MT]
  - multi-threaded, not vectorized

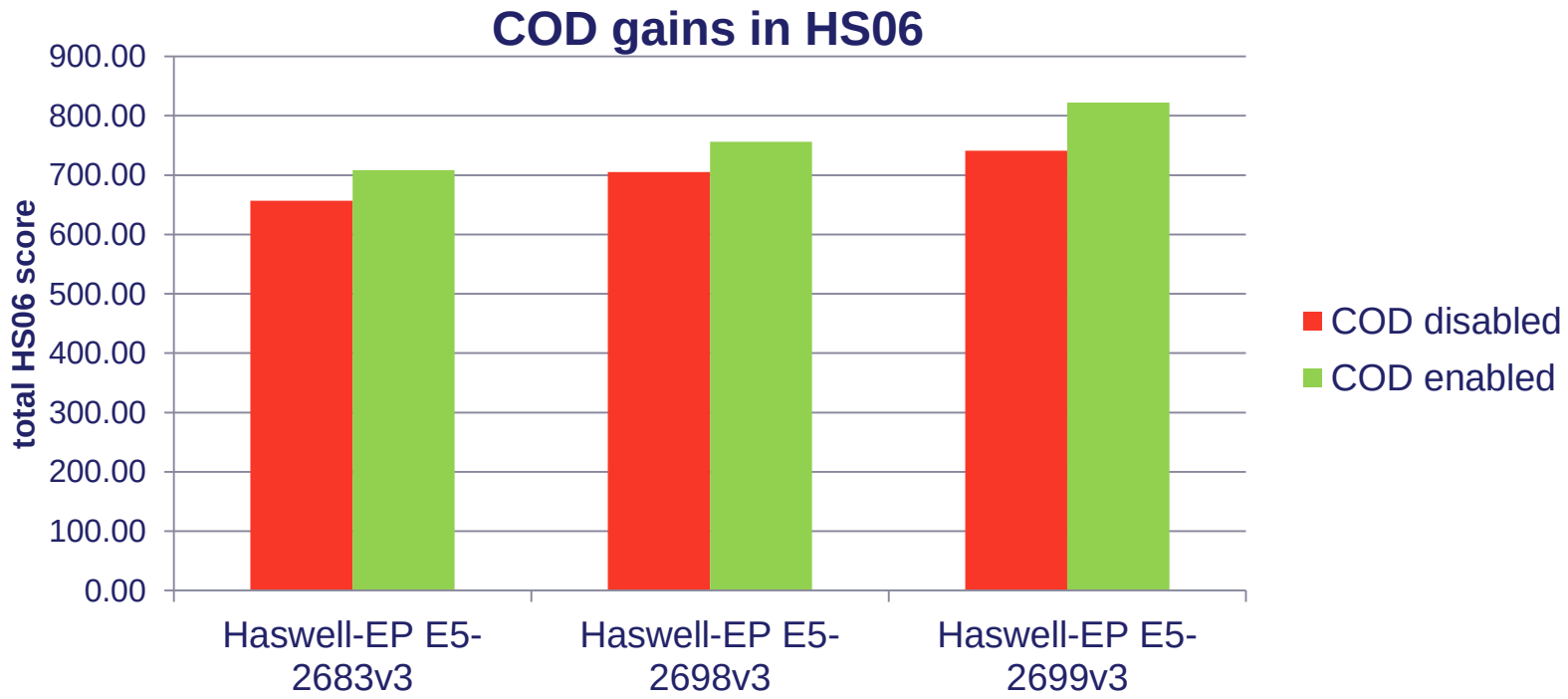
HEP-SPEC06 (I)


HEP-SPEC06 scalability (freq. scaled)
total HS06 score
number of threads
0
100
200
300
400
500
600
700
800
900
0
20
40
60
80
Sandy Bridge-EP E5-2690
Ivy Bridge-EP E5-2695v2
Haswell-EP E5-2683v3, COD enabled
Haswell-EP E5-2683v3, COD disabled
Haswell-EP E5-2698v3 COD enabled
Haswell-EP E5-2698v3, COD disabled
Haswell-EP E5-2699v3, COD enabled
Haswell-EP E5-2699v3, COD disabled

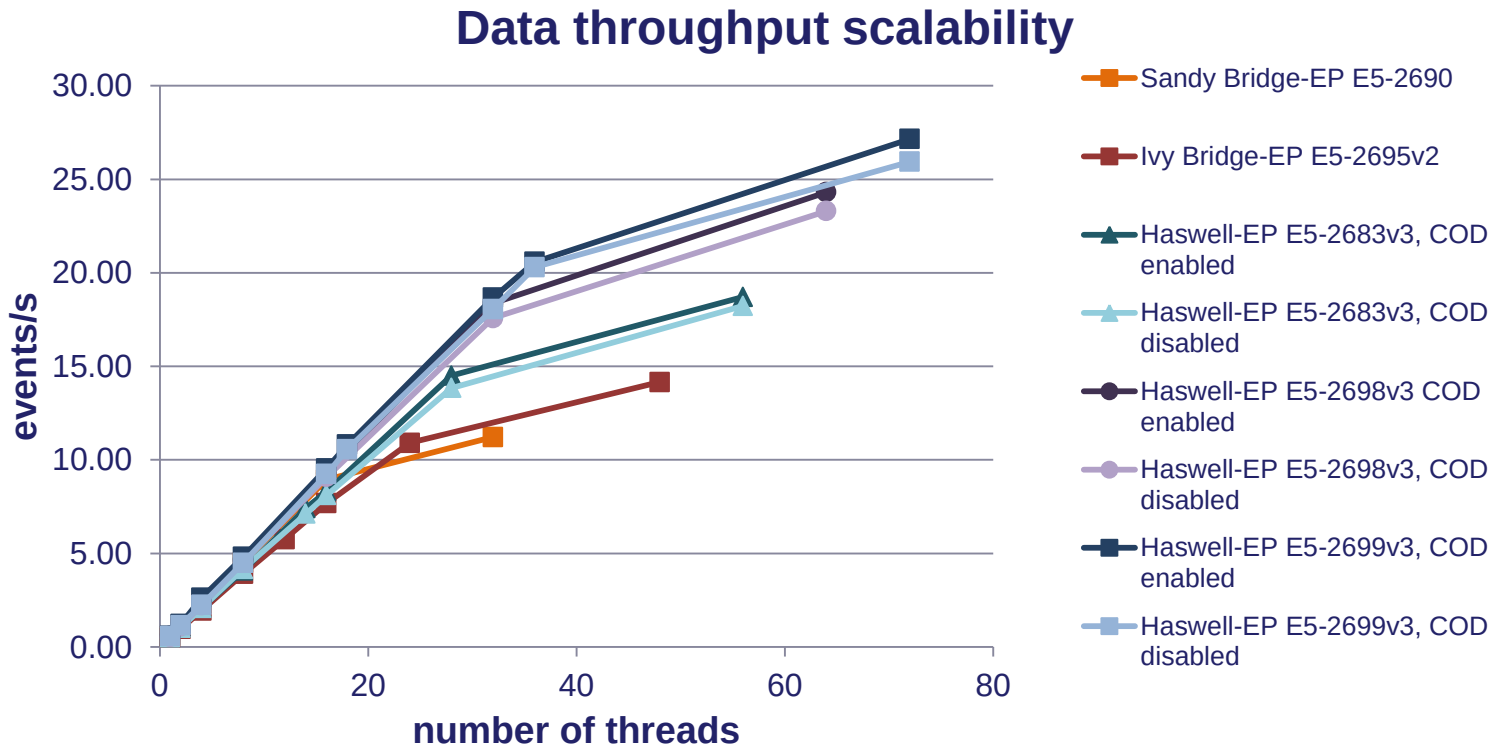
HEPSPEC06 (II)



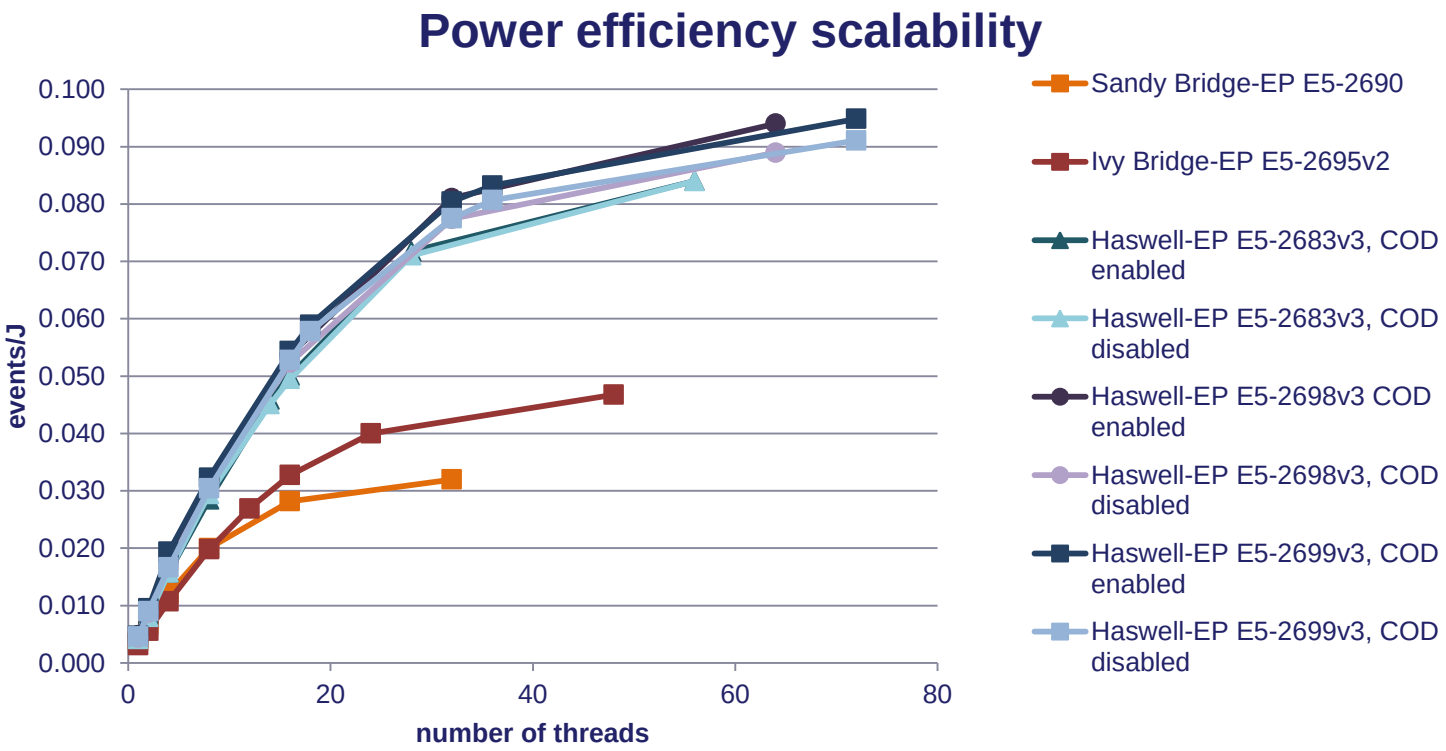
HEP-SPEC06 (III)



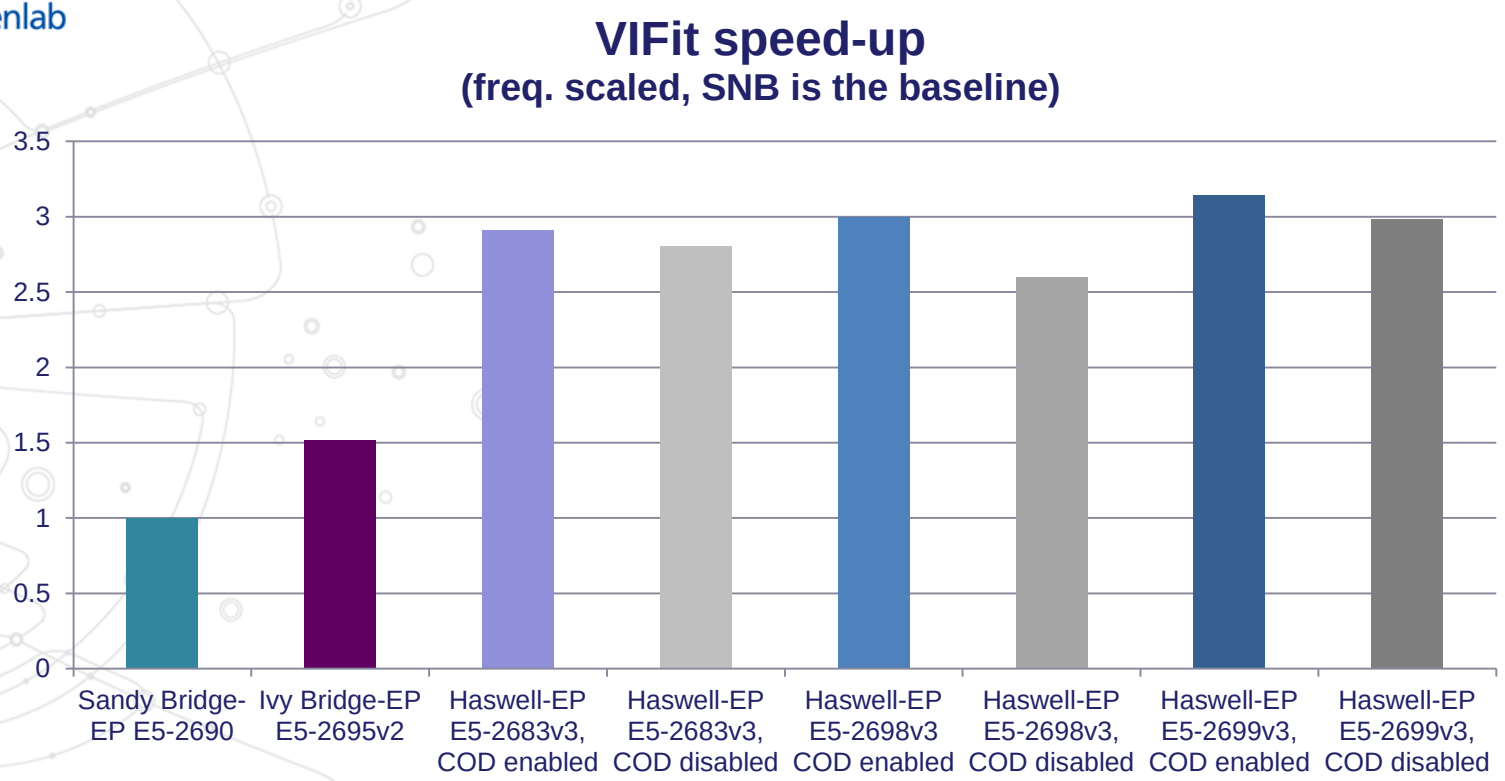
ParFullCMS (I)



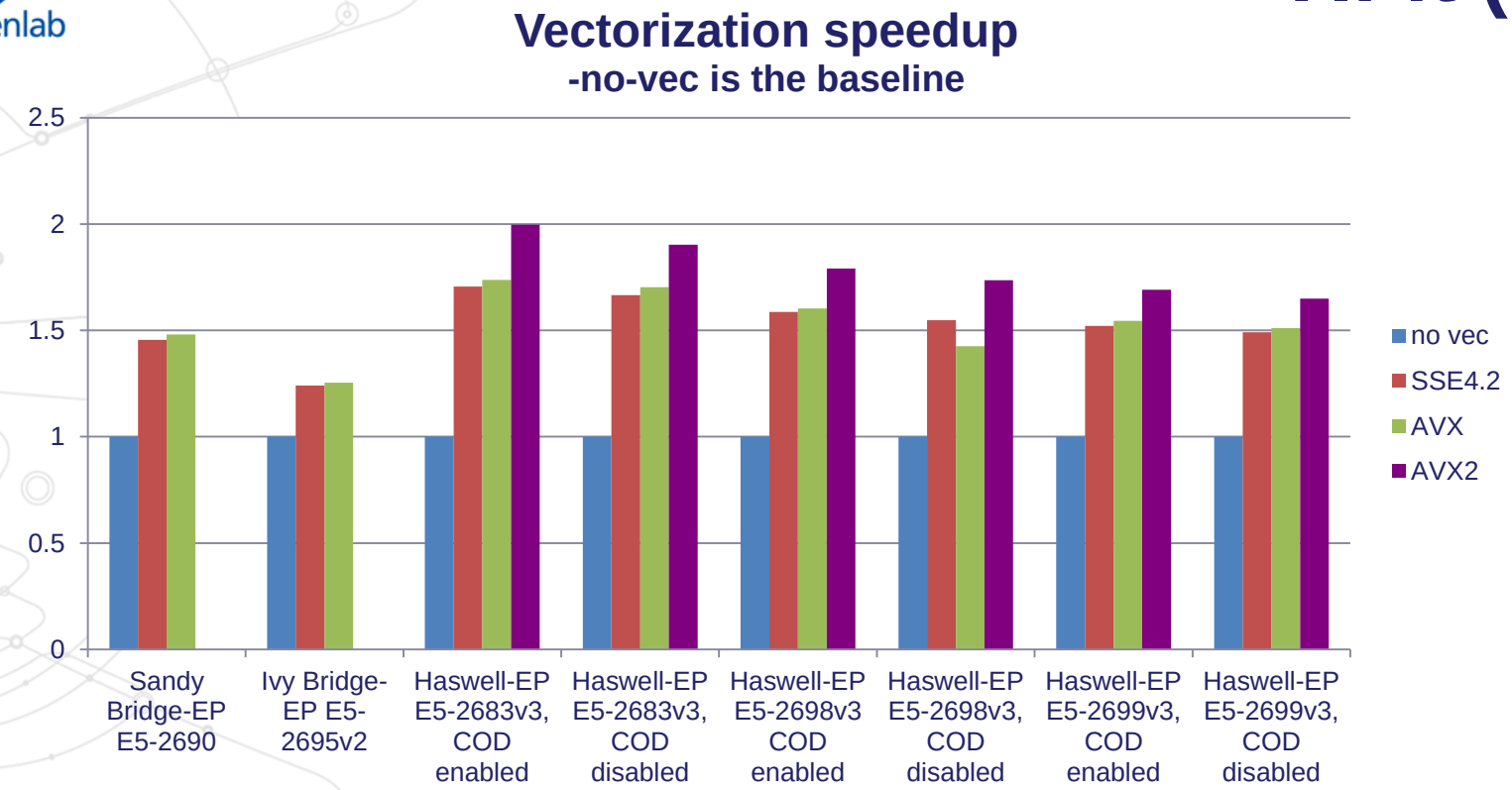
ParFullCMS (II)



VIFit (I)


VIFit speed-up
(freq. scaled, SNB is the baseline)
3.5
3
2.5
2
1.5
1
0.5
0
Sandy Bridge-EP E5-2690
Ivy Bridge-EP E5-2695v2
Haswell-EP E5-2683v3, COD enabled
Haswell-EP E5-2683v3, COD disabled
Haswell-EP E5-2698v3 COD enabled
Haswell-EP E5-2698v3, COD disabled
Haswell-EP E5-2699v3, COD enabled
Haswell-EP E5-2699v3, COD disabled

VIFit (II)

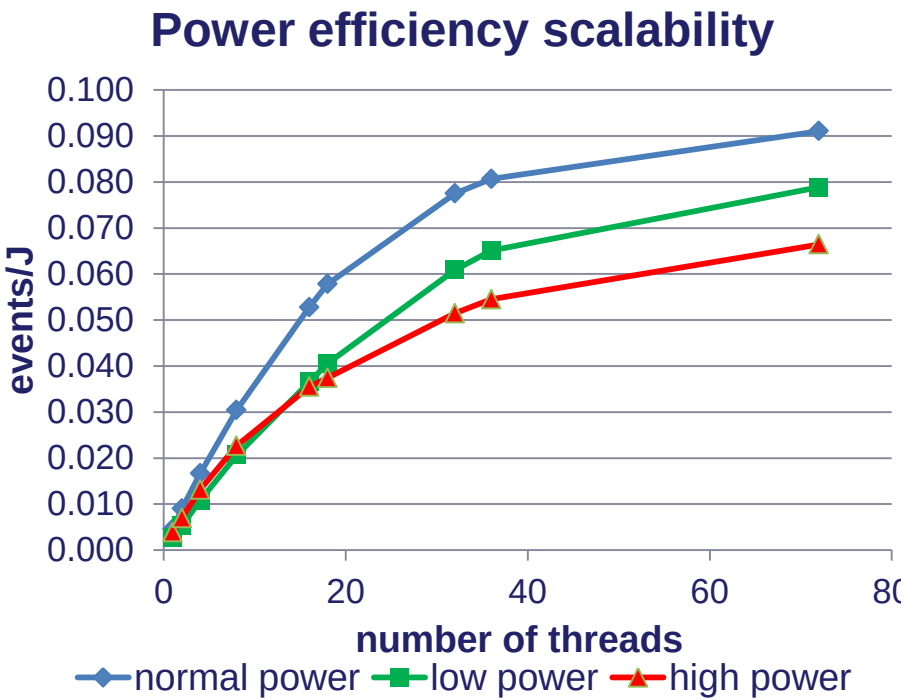

Vectorization speedup
-no-vec is the baseline
2.5
2
1.5
1
0.5
0
Sandy Bridge-EP E5-2690
Ivy Bridge-EP E5-2695v2
Haswell-EP E5-2683v3, COD enabled
Haswell-EP E5-2683v3, COD disabled
Haswell-EP E5-2698v3 COD enabled
Haswell-EP E5-2698v3, COD disabled
Haswell-EP E5-2699v3, COD enabled
Haswell-EP E5-2699v3, COD disabled
no vec
SSE4.2
AVX
AVX2

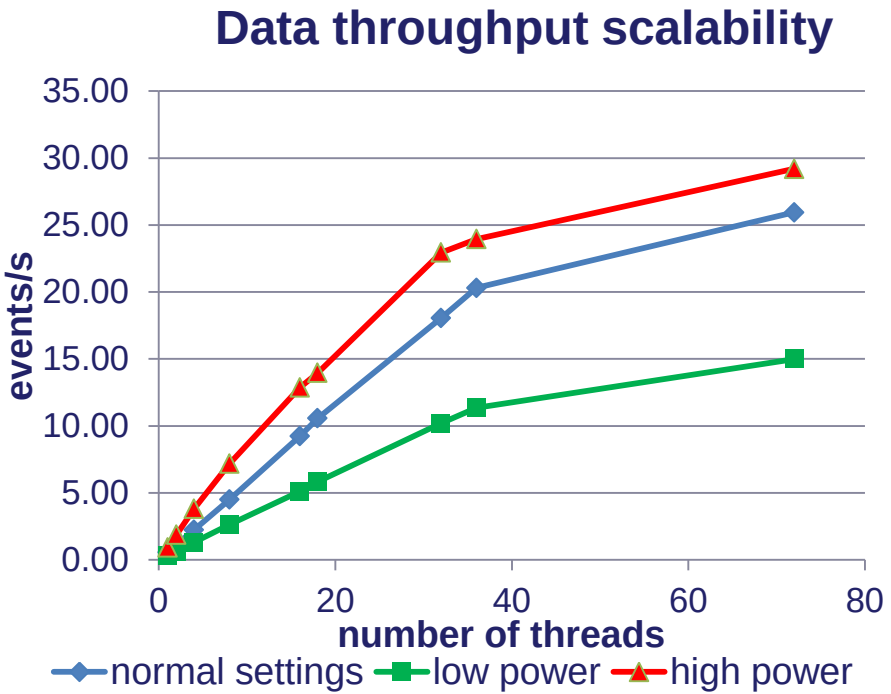
ParFullCMS – energy efficiency

- Three power settings (on Haswell-EP E5-2699v3 – 18 cores/socket)

	Turbo	P- and C-states	Uncore scaling	Energy Efficient Turbo
“standard”	NO	NO	YES	NO
Low power	YES	YES	YES	YES
High power	YES	NO	NO	NO

ParFullCMS – energy efficiency (II)





Conclusions

- Haswell-EP offers major improvements in the **uncore** part, which can be mainly exploited by multi-threaded (NUMA-aware) applications
- Otherwise, we “just” get a massive **core-count increase**
- Haswell-EP provides **higher vectorization gains** than previous dual socket platforms
- New power saving features allow lowering down the **absolute power consumption**, but for a cost of lower power efficiency

Questions?

質問？

Other questions?

pawel.szostek@cern.ch

BACK-UP SLIDES

Detailed COD diagram

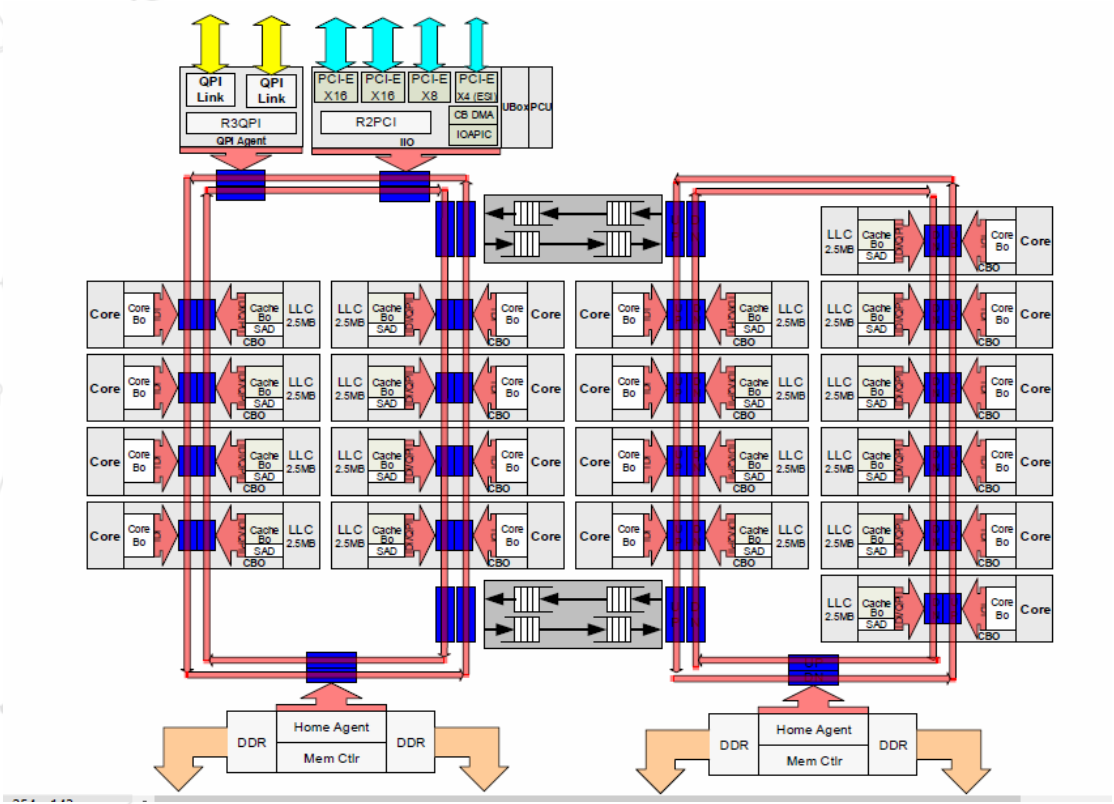
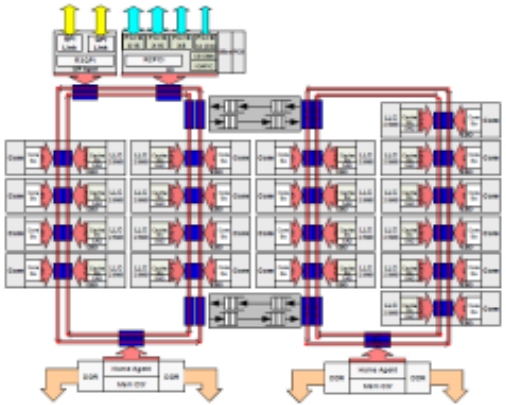


Image: Intel

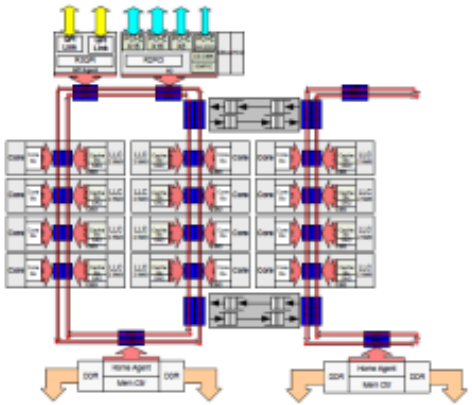
Detailed COD diagram (II)

Haswell EP Die Configurations

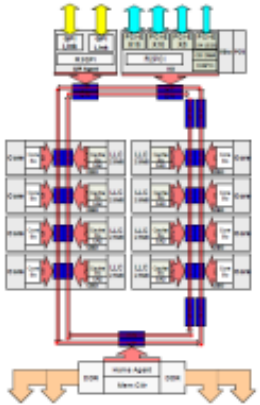
14-18 Core (HCC)



10-12 Core (MCC)



4-8 Core (LCC)



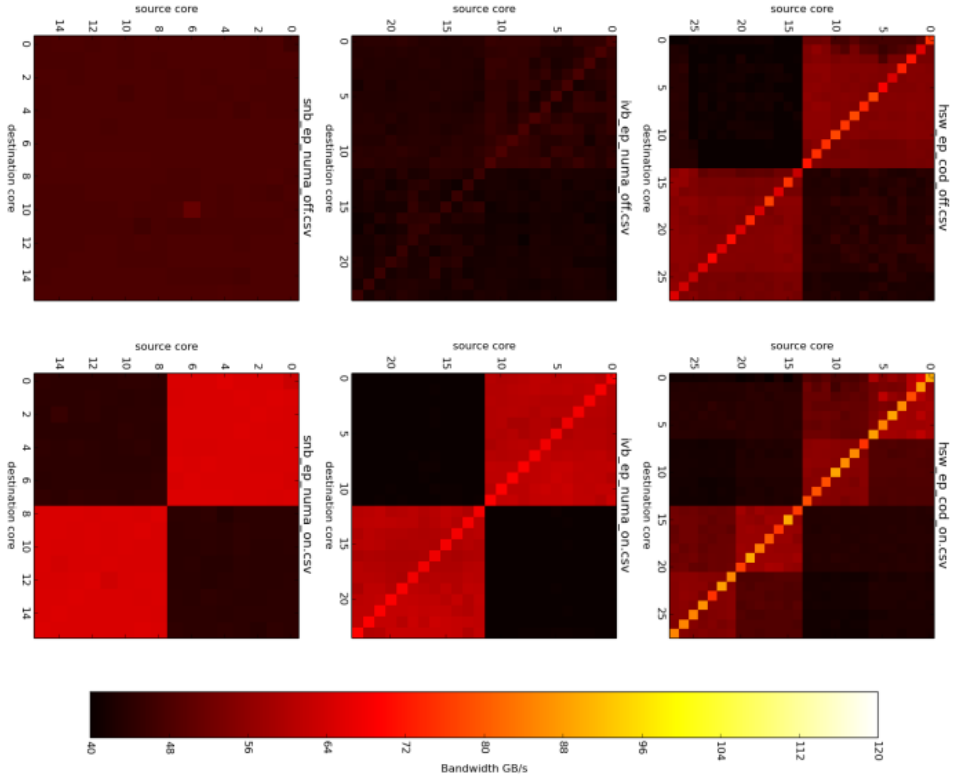
Not representative of actual die-sizes, orientation and layouts – for informational use only.

Chop	Columns	Home Agents	Cores	Power (W)	Transitors (B)	Die Area (mm²)
HCC	4	2	14-18	110-145	5.69	662
MCC	3	2	6-12	65-160	3.84	492
LCC	2	1	4-8	55-140	2.60	354

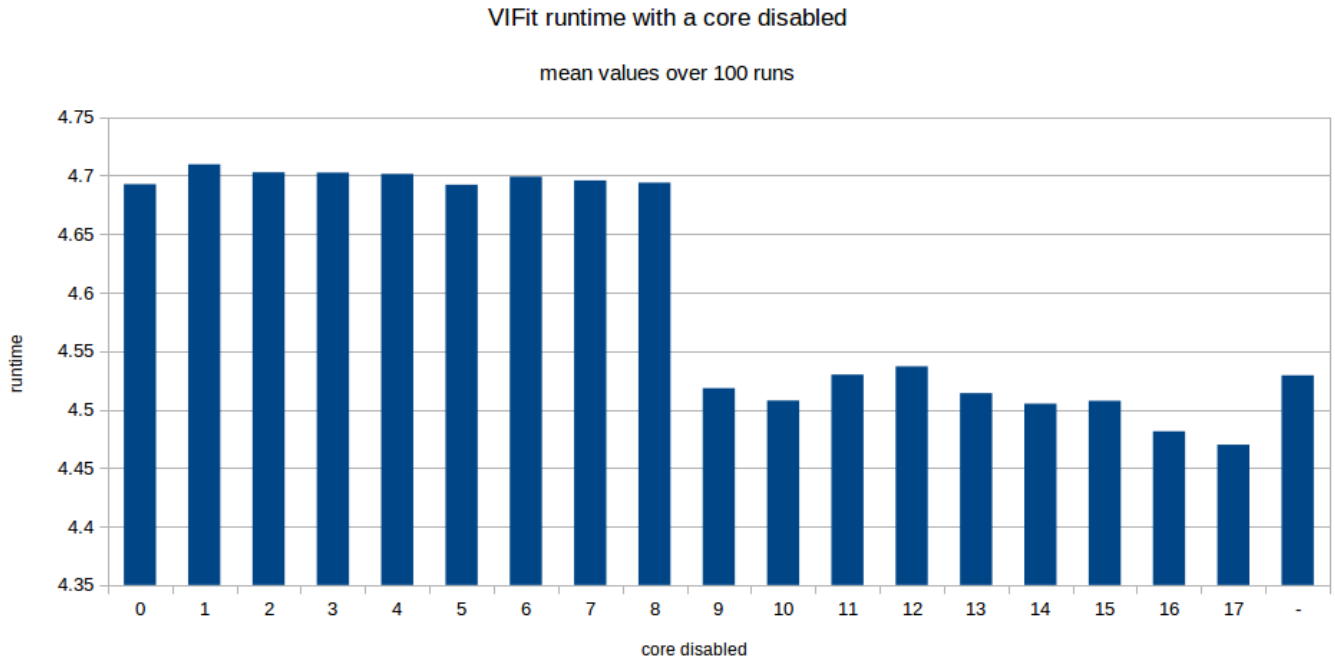
intel | 12

Image: Intel

Core-to-core bandwidth



Backup – VIFit with a core disabled



Backup – power consumption