INVESTIGATIONS OF **FUTURE COMPUTING PLATFORMS** FOR HIGH ENERGY PHYSICS

DAVID ABDURACHMANOV (UNL), FELICE PANTALEO (CERN), OMAR AWILE (CERN), MARCO GUERRI (CERN), ARITZ BROSA IARTZA (UNIVERSIDAD DE OVIEDO)

PRESENTED AT CHEP 2016

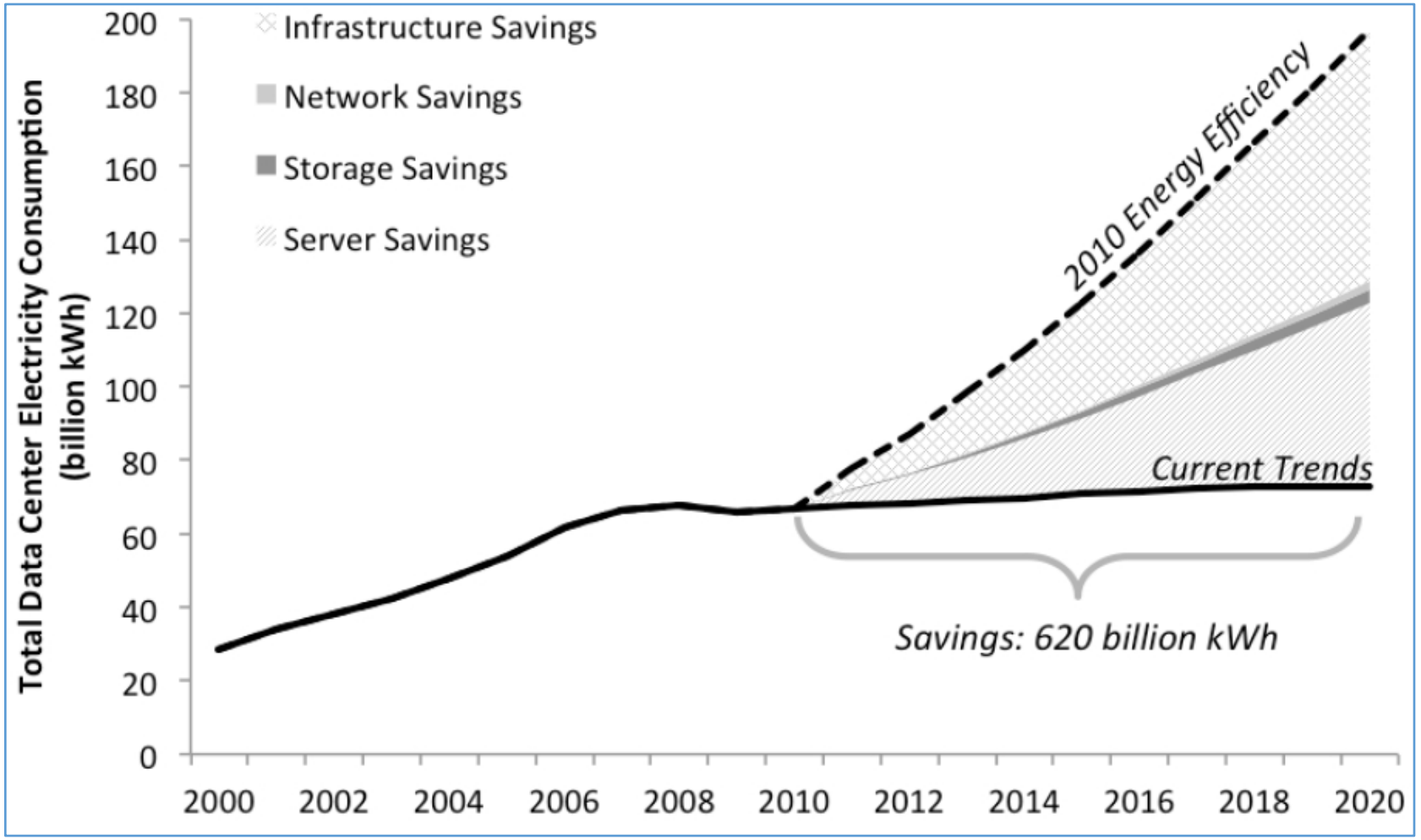
MOTIVATION

Gartner reported in 2010 that energy-related costs are the **fastest-rising cost in the data center**

New report:

Annual growth for server shipments have fallen significantly, same applies for power consumption, yet we continue to have "a drastic increase in demand for data center services"

United States Data Center Energy Usage Report (LBNL-1005775), June 2016



THE NEW ERA: THE CHANGE

The on-chip power density limitations are driving the computing market towards a greater variety of solutions

- ARM announced ARMv8.{0,1,2} which covers server market
  - Designs with up to 64 cores announced & to be used in supercomputers
- ARM recently announced Scalable Vector Extension (SVE) for HPC
  - Up to 2K vector sizes & binary compatibility
- IBM announced OpenPOWER Foundation which allows other companies to build PowerPC CPUs
- POWER9 (based on PowerISA 3.0) is the first one after OpenPOWER Foundation was launched
- IBM recently announced POWER8+ with NVLink 1.0 support
- Intel recently announced KNL (many-core and vector CPU) with legacy support for x86_64

THE NEW ERA: BEING BORING AND UNIQUE

Administrators and users shouldn't notice that they are running on different architectures – it's yet another PC

- All provide **LP64 data model** and **little-endian** support
- All can run RHEL/CentOS/Fedora distributions
- Minimal porting efforts (i.e. enablement)
- All have announced NVIDIA GPGPU (CUDA) support

But each vendor needs to differentiate

- POWER8+ has NVLink with up to 160GB/s bi-directional traffic
- KNL contains many-cores with 512-bit vectors
- ARMv8 SVE support from 128-bit to 2K-bit vectors in binary compatible way
- Bake custom silicon with general availability ISA

Geopolitics might play a role

SELECTED PROCESSORS

NOTE: All systems are production grade

	Cavium ThunderX	AMD A1100	APM X-Gene 1	Intel Xeon D-1540	Intel Xeon E5-2699 v3	Intel Xeon E5-2699 v4
Cores / Threads	48C/48T	8C/8T	8C/8T	8C/16C	18C/36T	22C/44T
Speed (GHz)	2.5	2.0	2.4	2.0 (2.6)	2.3 (3.6)	2.2 (3.6)
Cache (L3)	16MB	8MB	8MB	12MB	45MB	55MB
SMT	1	1	1	2	2	2
Year	Q4/15	Q1/2014	Q3/13	Q1/15	Q3/14	Q1/16

	Intel Xeon E5-2698 v3	Intel Xeon E5-2683 v4	Intel Xeon E5-2680 v4	Intel Xeon D-1581	IBM POWER8	IBM POWER8
Cores / Threads	16C/32T	16C/32T	14C/28T	16C/32T	8C/64T	10C/80T
Speed	2.3 (3.6)	2.1 (3.0)	2.4 (3.3)	1.8 (2.4)	3.857	3.491
Cache (L3)	40MB	40MB	35MB	24MB	64MB	80MB
SMT	2	2	2	2	8	8
Year	Q3/14	Q1/16	Q1/16	Q1/16	Q3/13	Q3/13

SELECTED PROCESSORS

	Cavium ThunderX	AMD A1100	APM X-Gene 1	Intel Xeon D-1540	Intel Xeon E5-2699 v3	Intel Xeon E5-2699 v4
Cores / Threads	48C/48T	8C/8T	8C/8T	8C/16C	18C/36T	22C/44T
Speed (GHz)	2.5	2.0	2.4	2.0 (2.6)	2.3 (3.6)	2.2 (3.6)
Cache (L3)	16MB	8MB	8MB	12MB	45MB	55MB
SMT	1	1	1	2	2	2
Year	Q4/15	Q1/2014	Q3/13	Q1/15	Q3/14	Q1/16

	Intel Xeon E5-2698 v3	Intel Xeon E5-2683 v4	Intel Xeon E5-2680 v4	Intel Xeon D-1581	IBM POWER8	IBM POWER8
Cores / Threads	16C/32T	16C/32T	14C/28T	16C/32T	8C/64T	10C/80T
Speed	2.3 (3.6)	2.1 (3.0)	2.4 (3.3)	1.8 (2.4)	3.857	3.491
Cache (L3)	40MB	40MB	35MB	24MB	64MB	80MB
SMT	2	2	2	2	8	8
Year	Q3/14	Q1/16	Q1/16	Q1/16	Q3/13	Q3/13

SELECTED PROCESSORS

	Cavium ThunderX	AMD A1100	APM X-Gene 1	**Intel Xeon D-1540**	Intel Xeon E5-2699 v3	Intel Xeon E5-2699 v4
Cores / Threads	48C/48T	8C/8T	8C/8T	8C/16C	18C/36T	22C/44T
Speed (GHz)	2.5	2.0	2.4	2.0 (2.6)	2.3 (3.6)	2.2 (3.6)
Cache (L3)	16MB	8MB	8MB	12MB	45MB	55MB
SMT	1	1	1	2	2	2
Year	Q4/15	Q1/2014	Q3/13	Q1/15	Q3/14	Q1/16

	Intel Xeon E5-2698 v3	Intel Xeon E5-2683 v4	Intel Xeon E5-2680 v4	**Intel Xeon D-1581**	IBM POWER8	IBM POWER8
Cores / Threads	16C/32T	16C/32T	14C/28T	16C/32T	8C/64T	10C/80T
Speed	2.3 (3.6)	2.1 (3.0)	2.4 (3.3)	1.8 (2.4)	3.857	3.491
Cache (L3)	40MB	40MB	35MB	24MB	64MB	80MB
SMT	2	2	2	2	8	8
Year	Q3/14	Q1/16	Q1/16	Q1/16	Q3/13	Q3/13

SELECTED PROCESSORS

	Cavium ThunderX	AMD A1100	APM X-Gene 1	Intel Xeon D-1540	Intel Xeon E5-2699 v3	Intel Xeon E5-2699 v4
Cores / Threads	48C/48T	8C/8T	8C/8T	8C/16C	18C/36T	22C/44T
Speed (GHz)	2.5	2.0	2.4	2.0 (2.6)	2.3 (3.6)	2.2 (3.6)
Cache (L3)	16MB	8MB	8MB	12MB	45MB	55MB
SMT	1	1	1	2	2	2
Year	Q4/15	Q1/2014	Q3/13	Q1/15	Q3/14	Q1/16

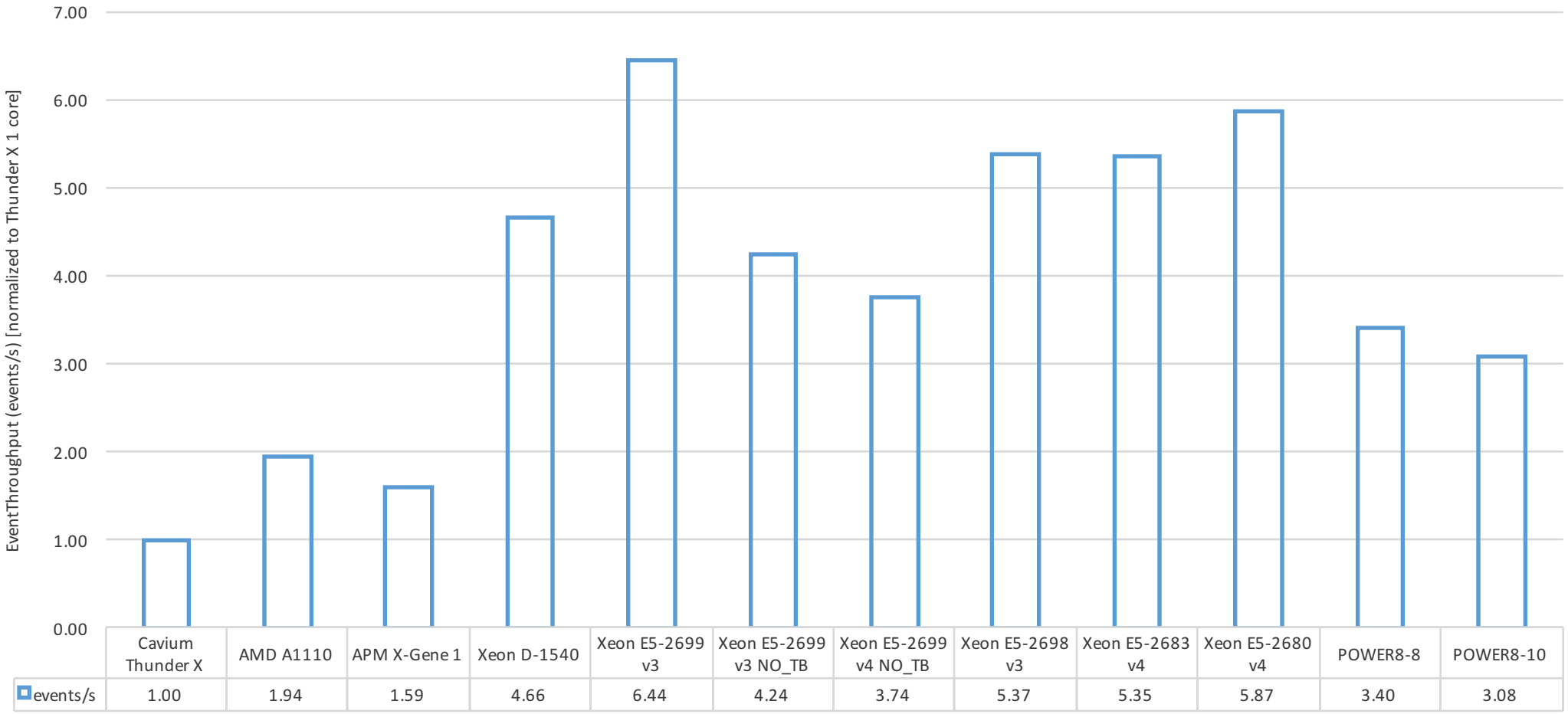
	Intel Xeon E5-2698 v3	Intel Xeon E5-2683 v4	Intel Xeon E5-2680 v4	Intel Xeon D-1581	IBM POWER8	IBM POWER8
Cores / Threads	16C/32T	16C/32T	14C/28T	16C/32T	8C/64T	10C/80T
Speed	2.3 (3.6)	2.1 (3.0)	2.4 (3.3)	1.8 (2.4)	3.857	3.491
Cache (L3)	40MB	40MB	35MB	24MB	64MB	80MB
SMT	2	2	2	2	8	8
Year	Q3/14	Q1/16	Q1/16	Q1/16	Q3/13	Q3/13

SELECTED PROCESSORS

	Cavium ThunderX	AMD A1100	APM X-Gene 1	Intel Xeon D-1540	Intel Xeon E5-2699 v3	Intel Xeon E5-2699 v4
Cores / Threads	48C/48T	8C/8T	8C/8T	8C/16C	18C/36T	22C/44T
Speed (GHz)	2.5	2.0	2.4	2.0 (2.6)	2.3 (3.6)	2.2 (3.6)
Cache (L3)	16MB	8MB	8MB	12MB	45MB	55MB
SMT	1	1	1	2	2	2
Year	Q4/15	Q1/2014	Q3/13	Q1/15	Q3/14	Q1/16

	Intel Xeon E5-2698 v3	Intel Xeon E5-2683 v4	Intel Xeon E5-2680 v4	Intel Xeon D-1581	IBM POWER8	IBM POWER8
Cores / Threads	16C/32T	16C/32T	14C/28T	16C/32T	8C/64T	10C/80T
Speed	2.3 (3.6)	2.1 (3.0)	2.4 (3.3)	1.8 (2.4)	3.857	3.491
Cache (L3)	40MB	40MB	35MB	24MB	64MB	80MB
SMT	2	2	2	2	8	8
Year	Q3/14	Q1/16	Q1/16	Q1/16	Q3/13	Q3/13

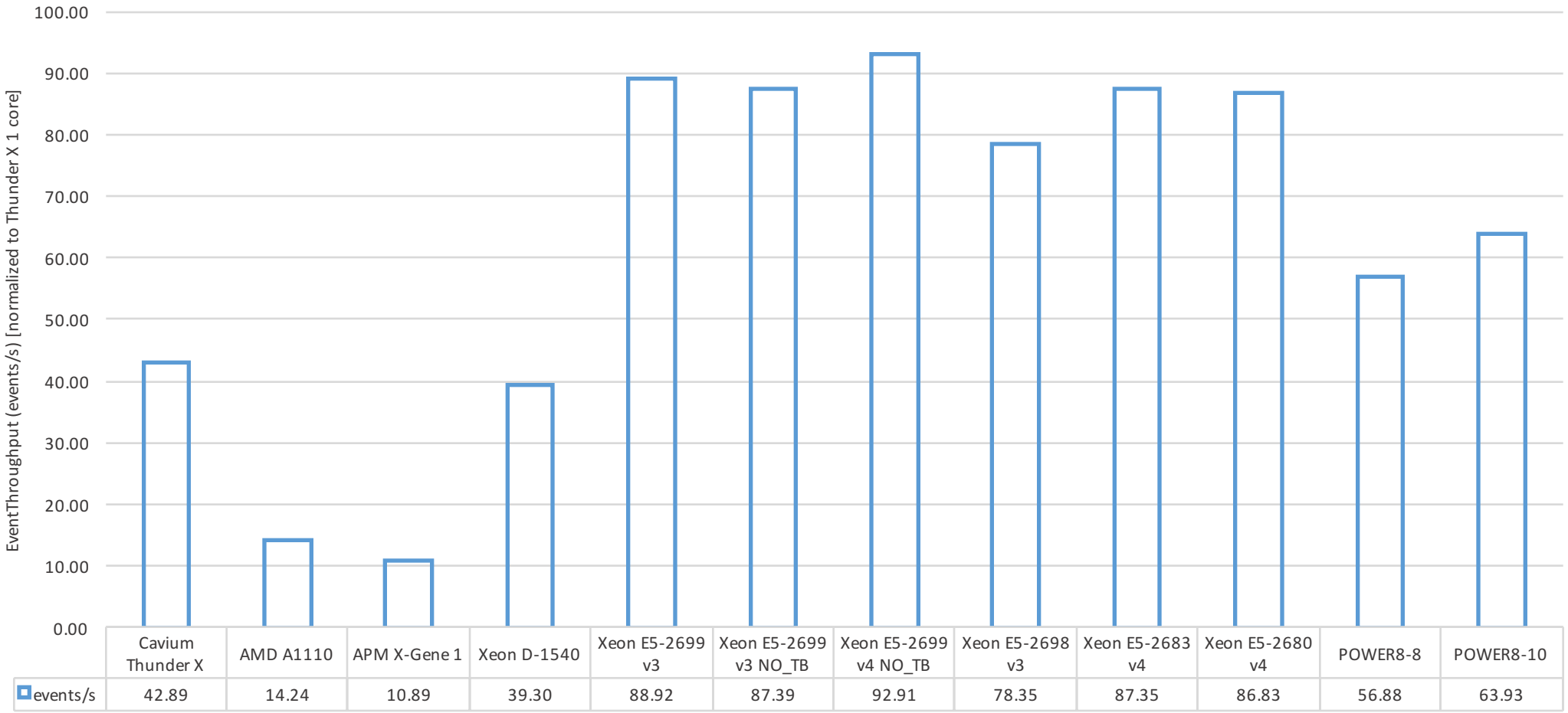
CMSSW RECO: RAW PERFORMANCE – SINGLE CORE



	Cavium Thunder X	AMD A1110	APM X-Gene 1	Xeon D-1540	Xeon E5-2699 v3	Xeon E5-2699 v3 NO_TB	Xeon E5-2699 v4 NO_TB	Xeon E5-2698 v3	Xeon E5-2683 v4	Xeon E5-2680 v4	POWER8-8	POWER8-10
events/s	1.00	1.94	1.59	4.66	6.44	4.24	3.74	5.37	5.35	5.87	3.40	3.08

NOTE: All numbers normalized to a single core Thunder X performance

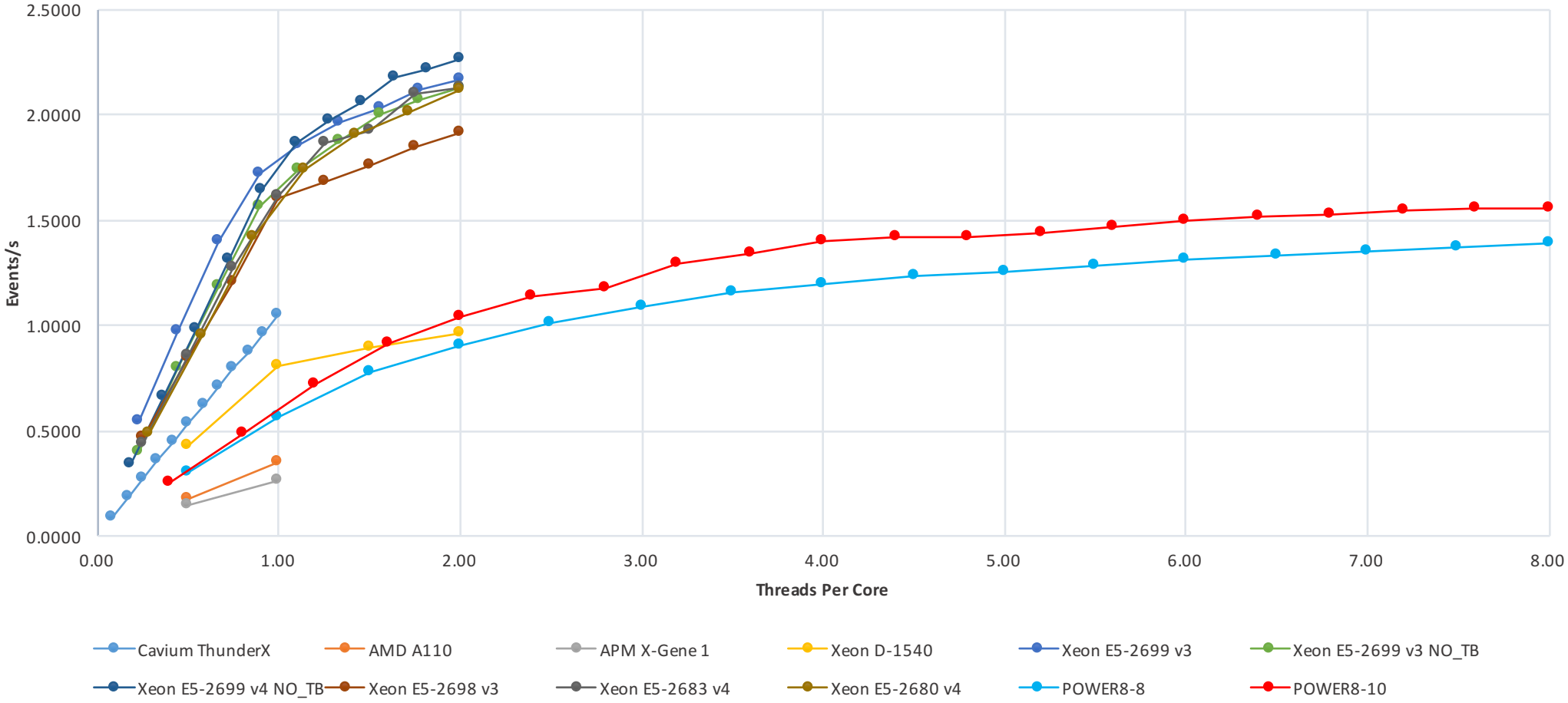
CMSSW RECO: RAW PERFORMANCE – SINGLE SOCKET



	Cavium Thunder X	AMD A1110	APM X-Gene 1	Xeon D-1540	Xeon E5-2699 v3	Xeon E5-2699 v3 NO_TB	Xeon E5-2699 v4 NO_TB	Xeon E5-2698 v3	Xeon E5-2683 v4	Xeon E5-2680 v4	POWER8-8	POWER8-10
events/s	42.89	14.24	10.89	39.30	88.92	87.39	92.91	78.35	87.35	86.83	56.88	63.93

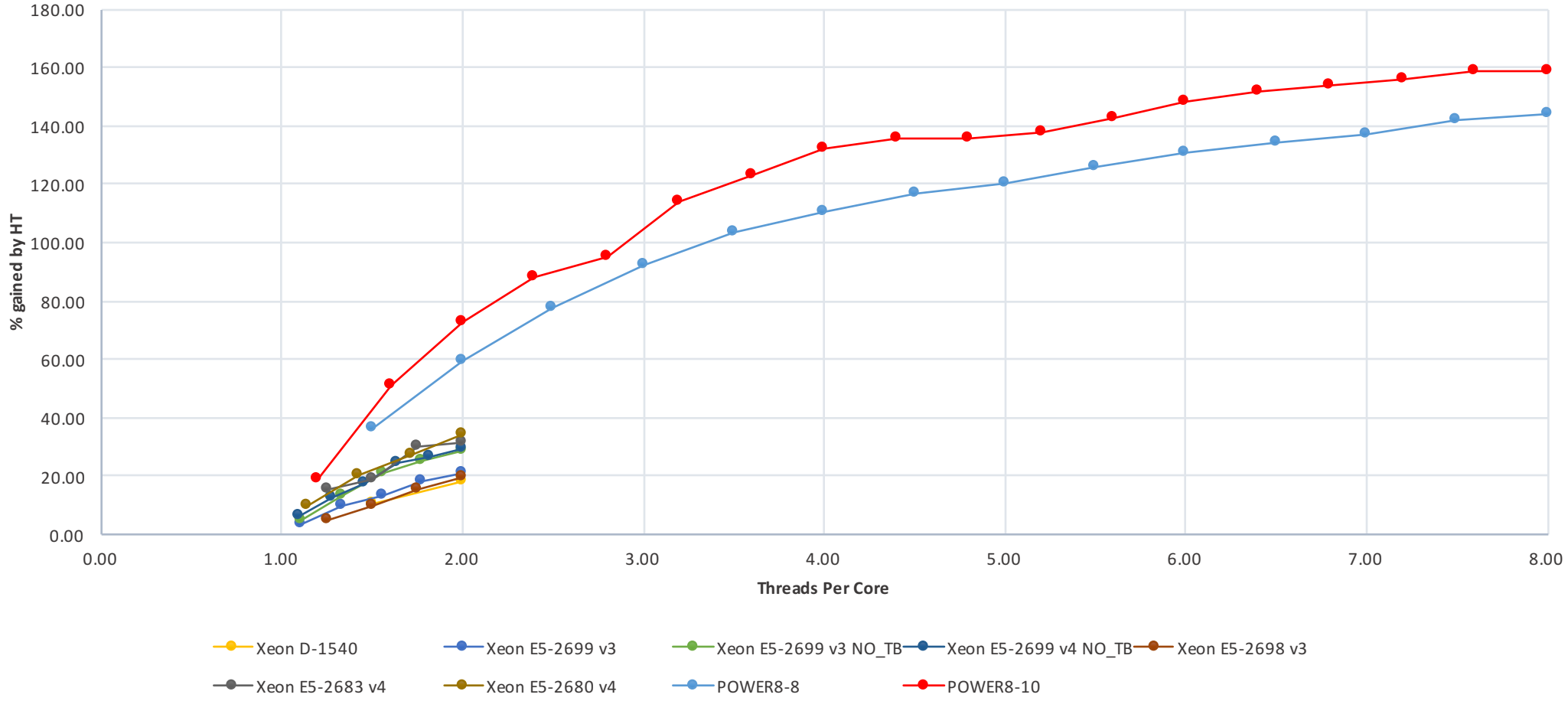
NOTE: All numbers normalized to a single core Thunder X performance.
Multi-threaded (**4 threads**) jobs were used to fill full socket due to memory constraints.

CMSSW RECO: RAW PERFORAMNCE – SCALABILITY



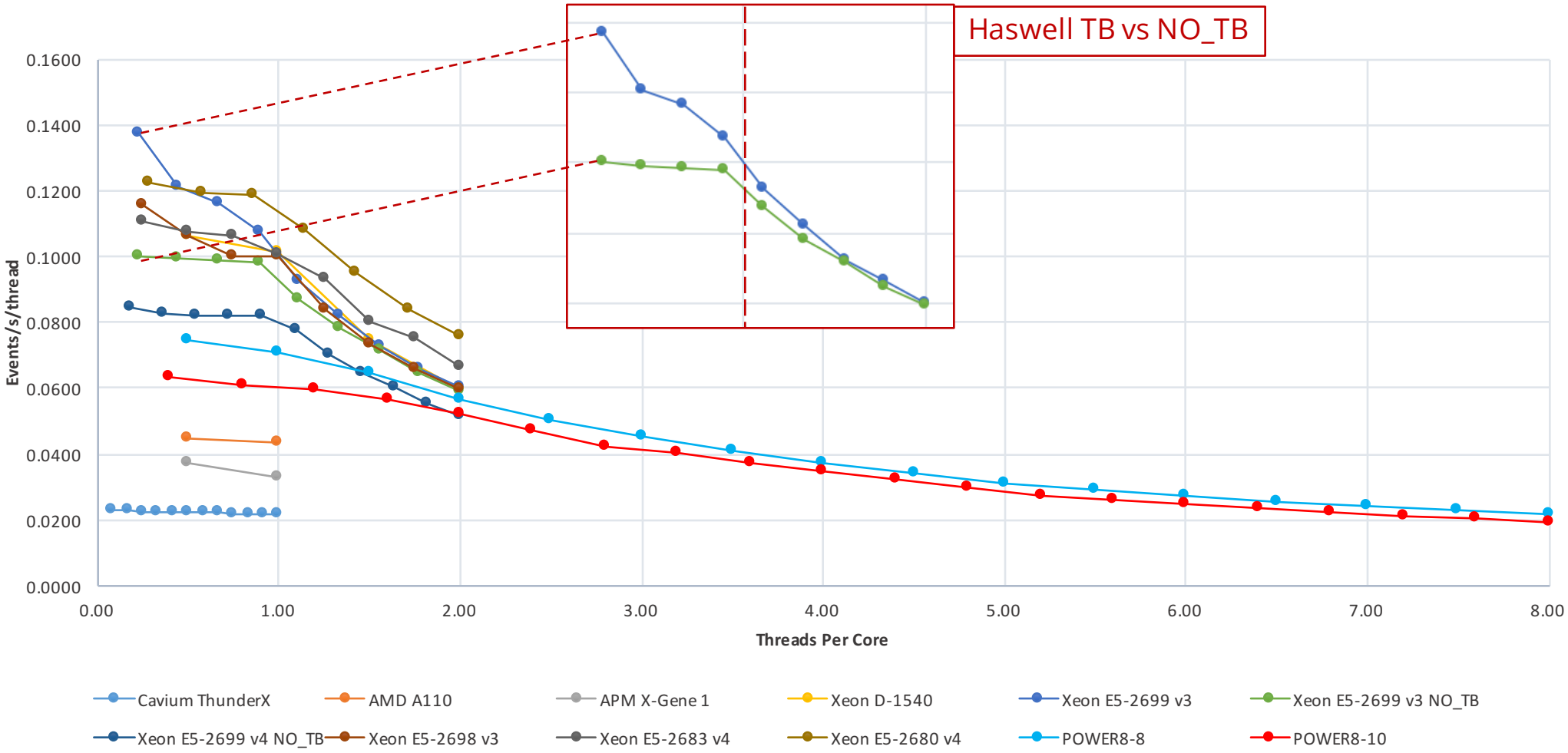
NOTE: Multi-threaded (**4 threads**) jobs were used to fill full socket due to memory constraints.

CMSSW RECO: RAW PERFORAMNCE – HYPERTHREADING



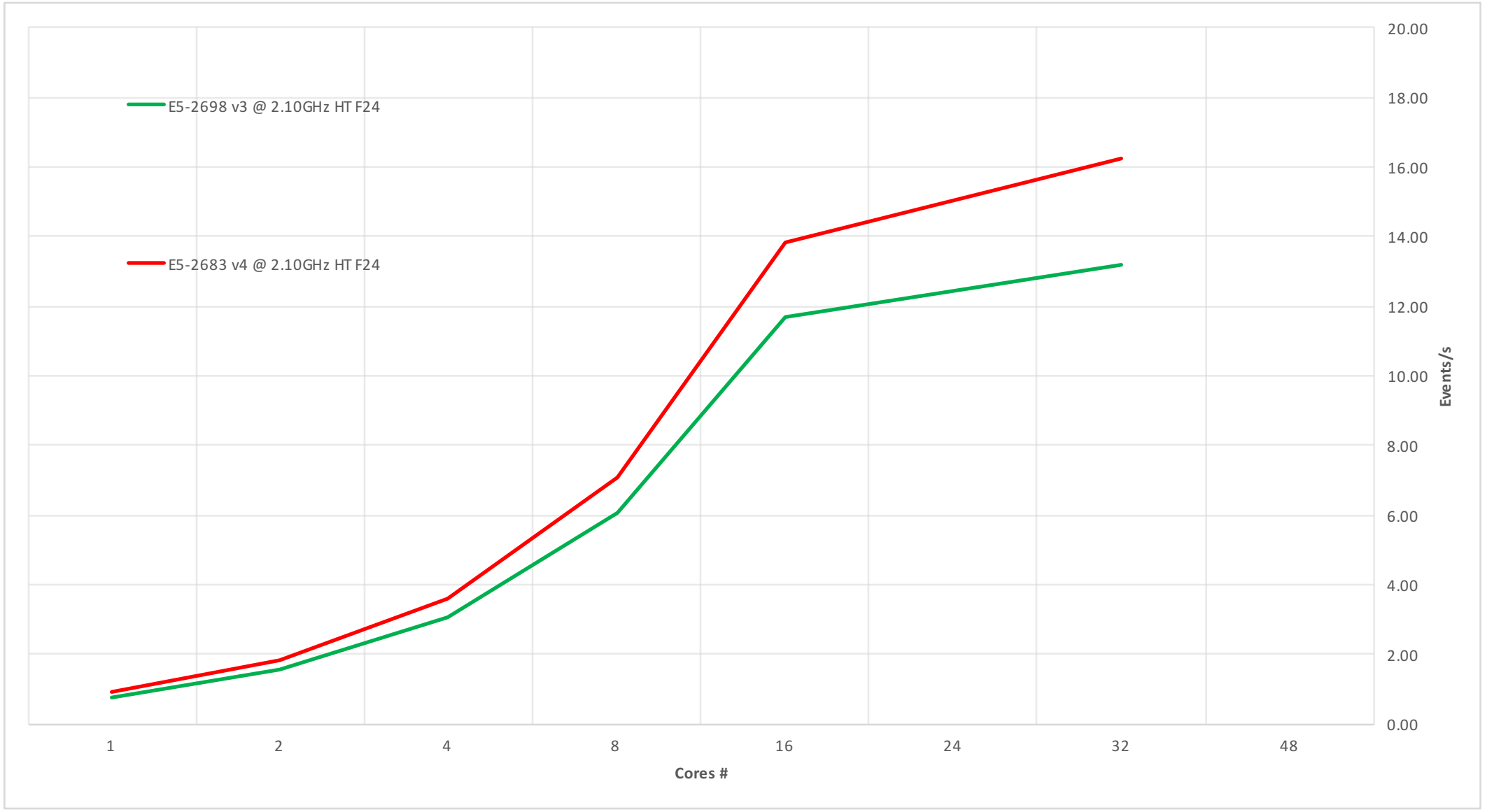
NOTE: Multi-threaded (**4 threads**) jobs were used to fill full socket due to memory constraints.

CMSSW RECO: RAW PERFORAMNCE – TURBO BOOST

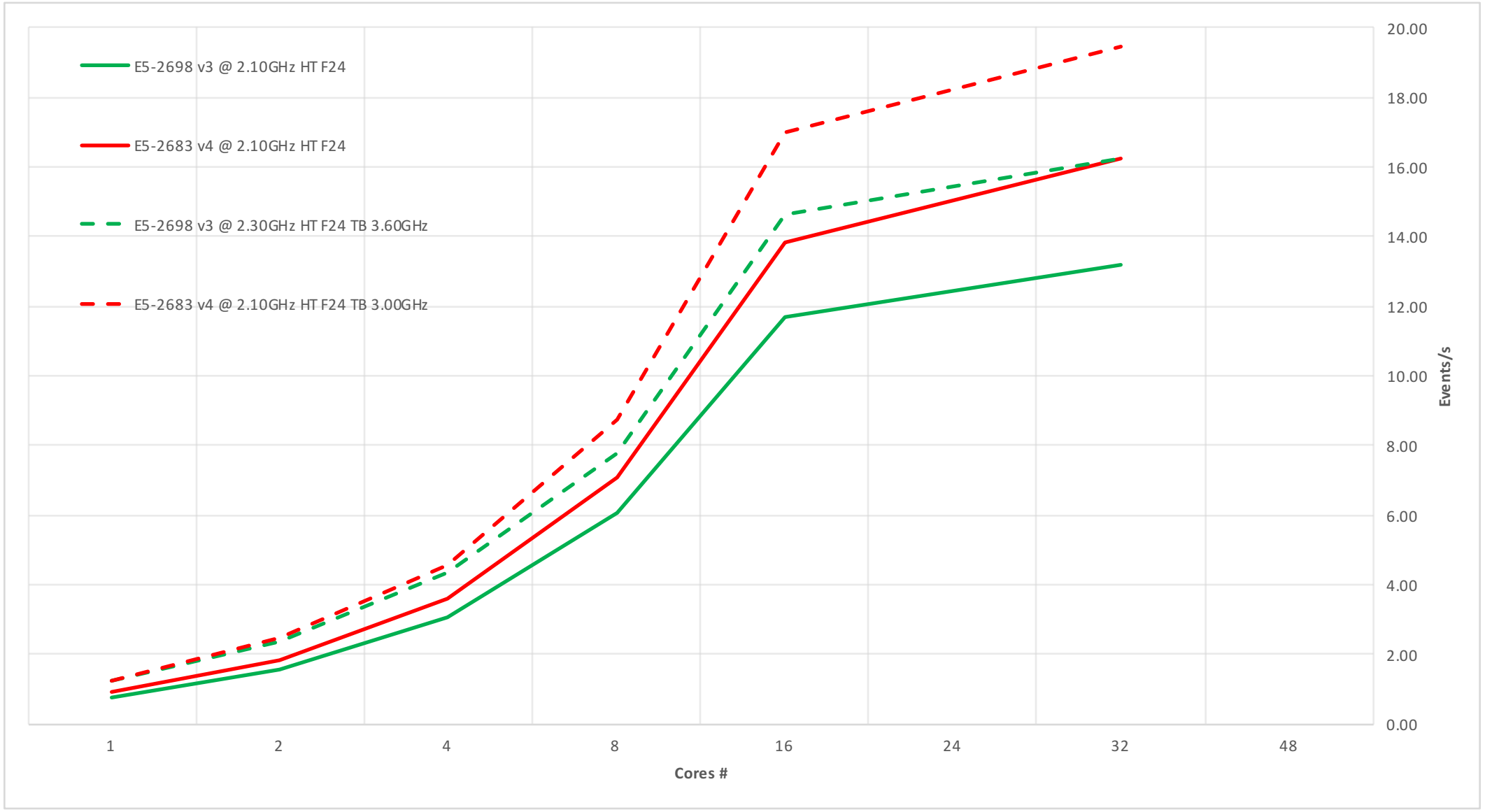


NOTE: Multi-threaded (**4 threads**) jobs were used to fill full socket due to memory constraints.

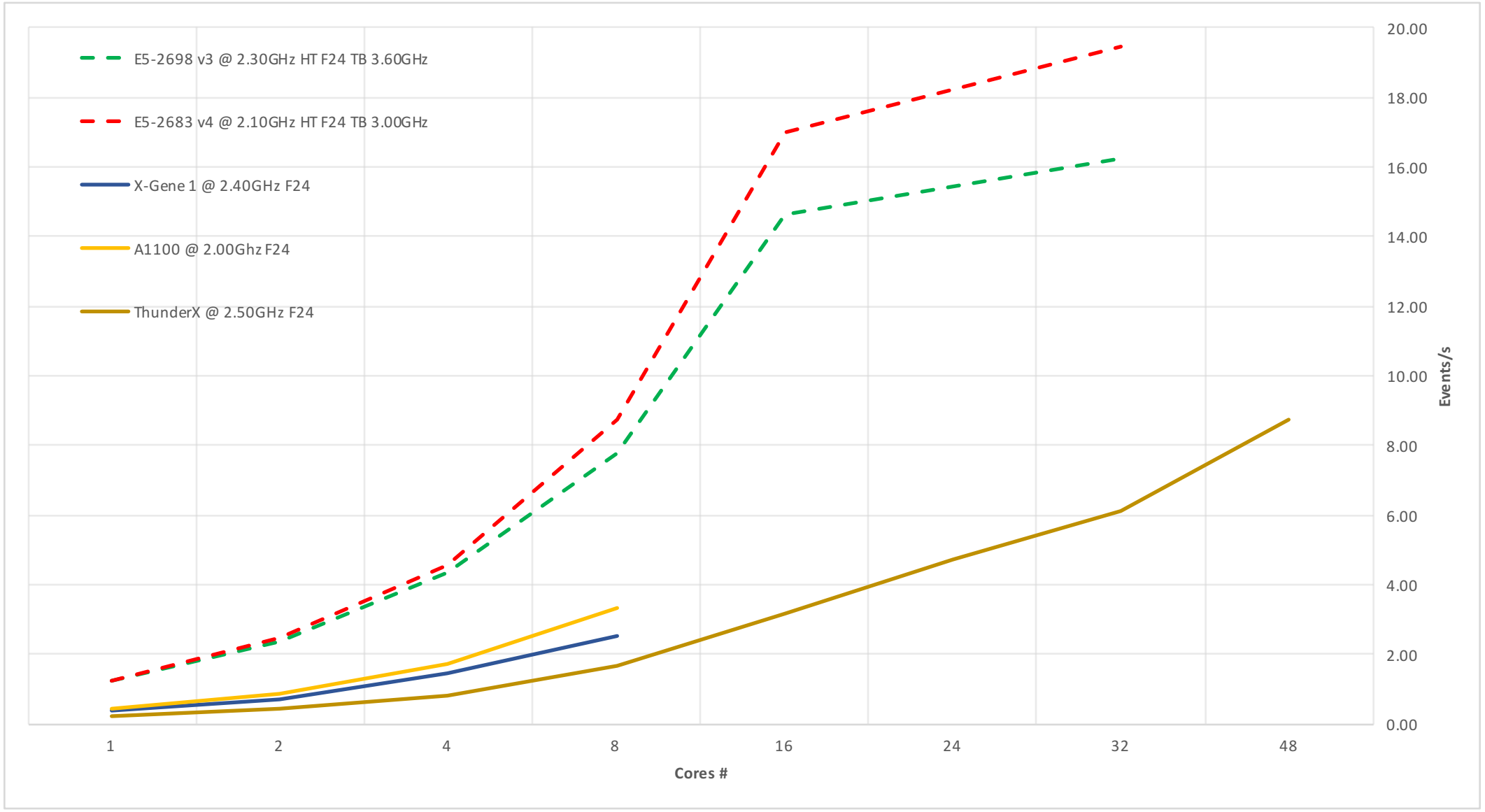
PARFULLCMS: SCALABILITY – XEON V3 VS V4, FREQ LOCKED



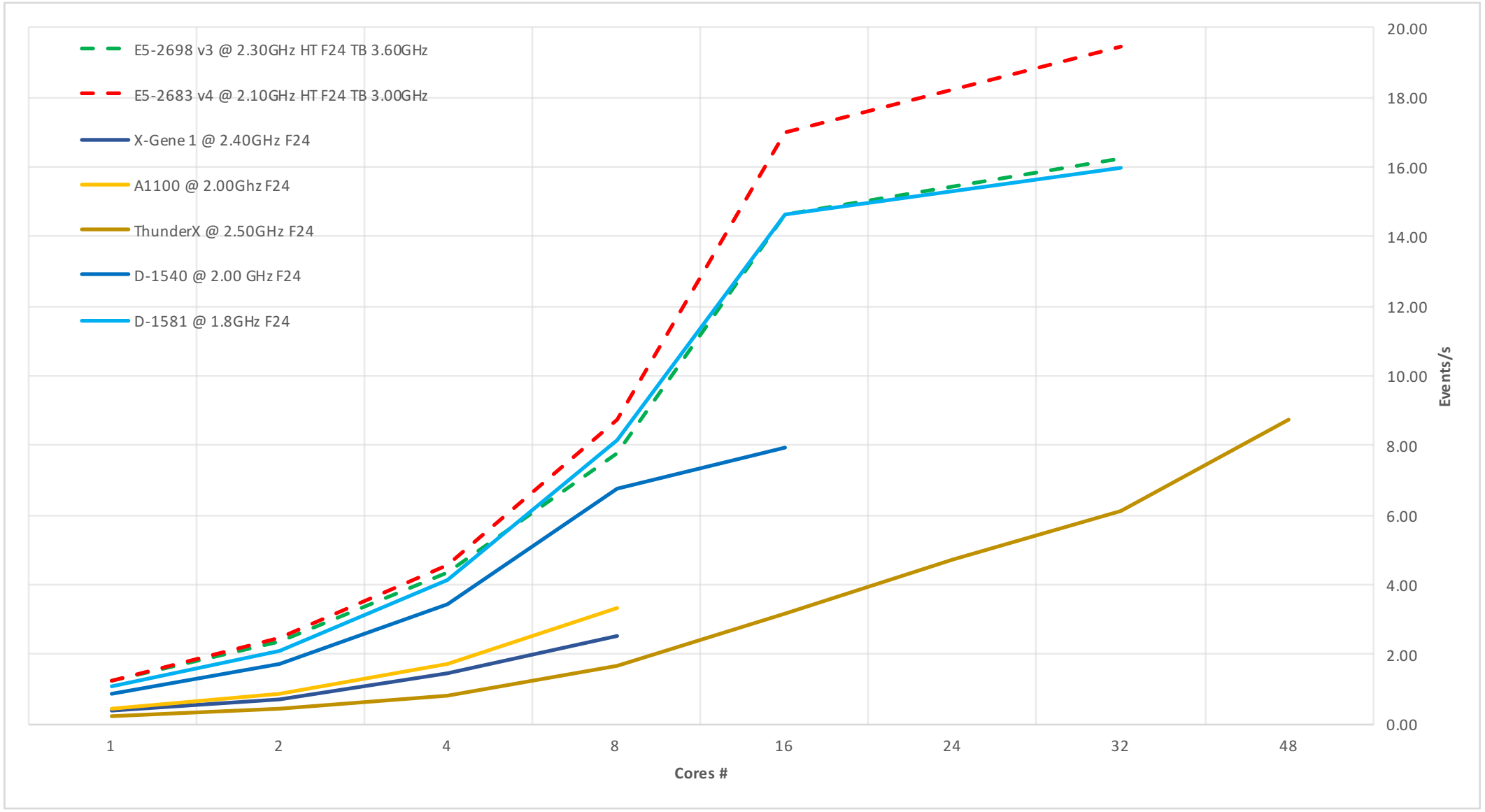
PARFULLCMS: SCALABILITY – XEON V3 VS V4, TURBOBOOST



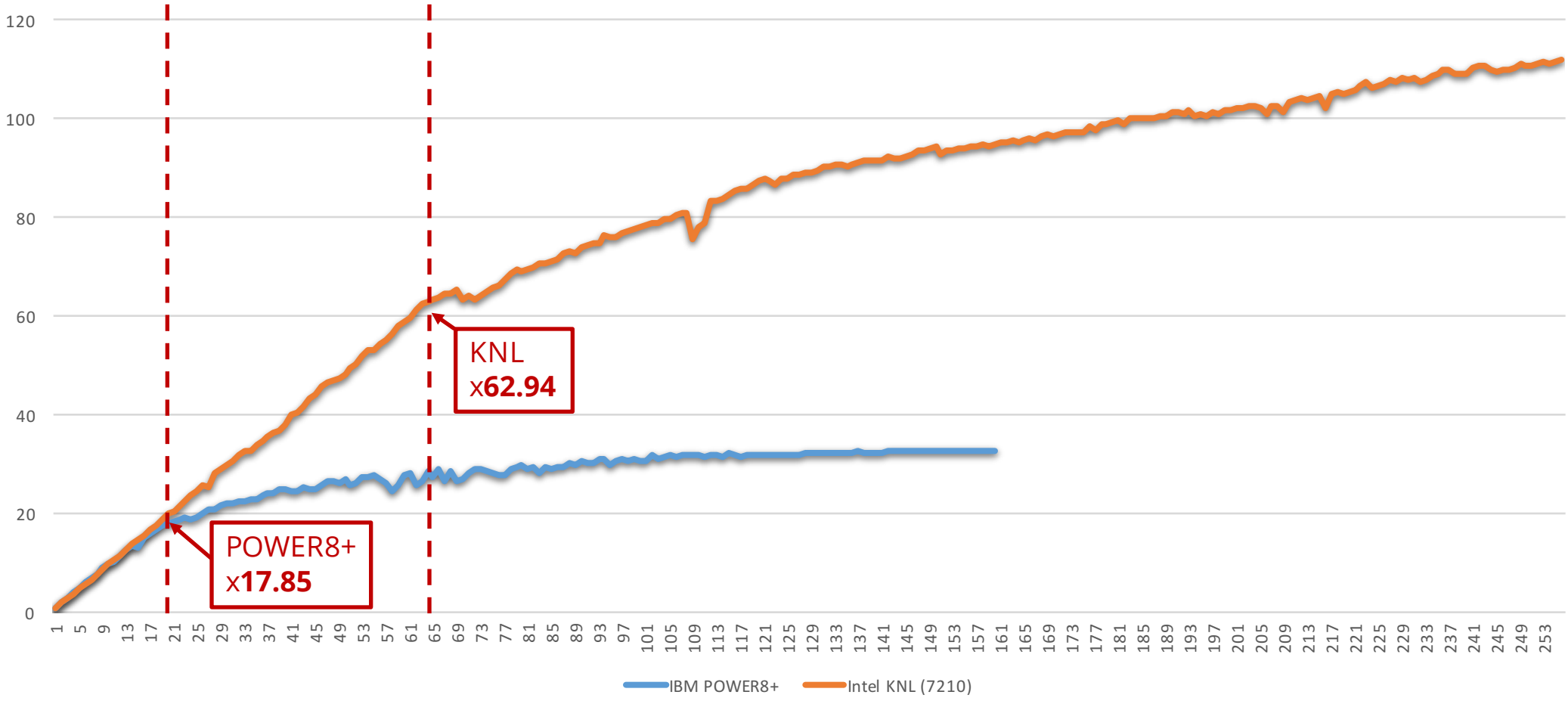
PARFULLCMS: SCALABILITY – ARM64/AARCH64



PARFULLCMS: SCALABILITY – XEON-D (BROADWELL)

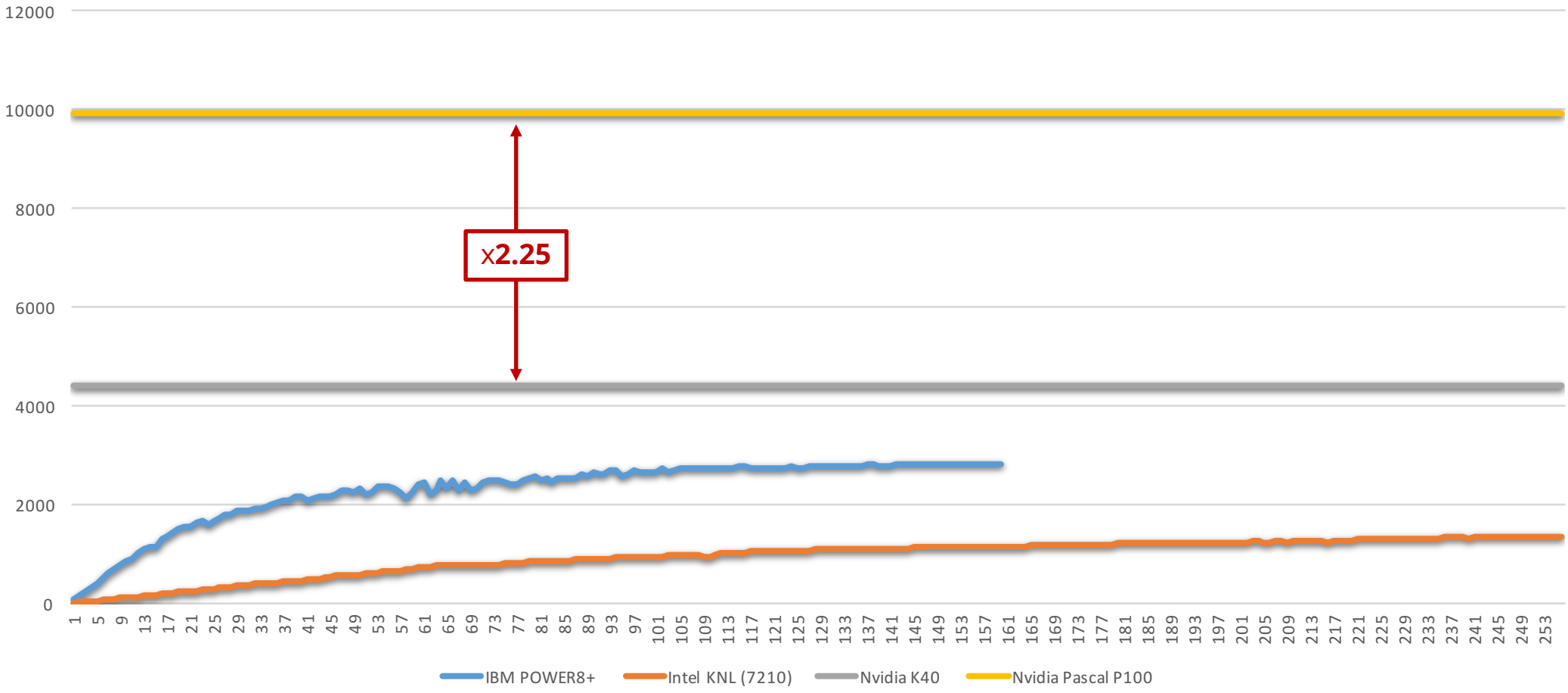


FKDTREE: SCALABILITY - SPEEDUP



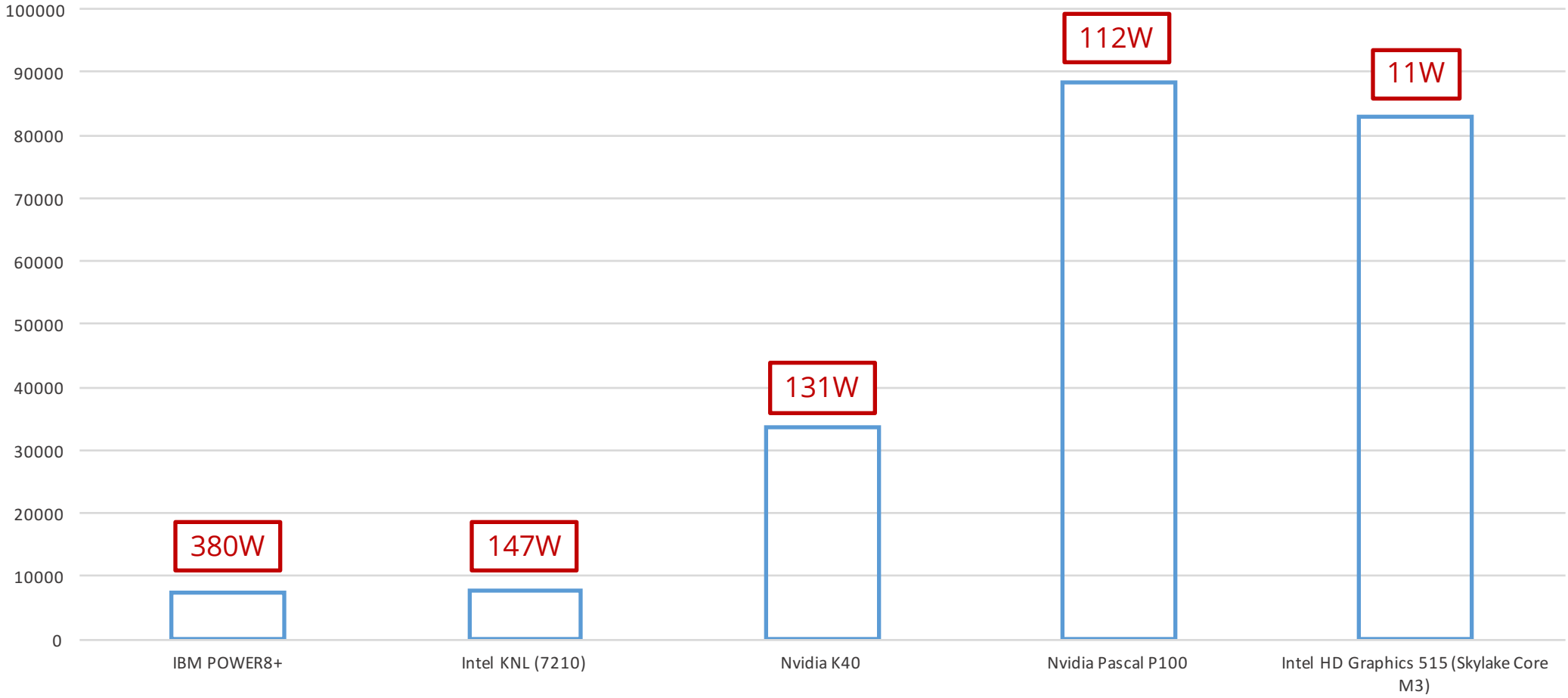
NOTE: FKDtree is a heapified KDTree with parallel search implemented in OpenMP, TBB, OpenCL, and CUDA. It is used for nearest neighbor finding in tracking and clustering.
This is **box to box comparison**, **POWER8+ is 2S** Minsky system.

FKDTREE: THROUGHPUT - SEARCHES/MS



NOTE: P100-SXM2 is connected with NVLink 1.0, but we only transfer data once.
Implementation is using SP.

FKDTREE: POWER EFFICIENCY - SEARCHES/J



NOTE: Power measurements are for CPUs/GPGPUs, not for the full box.

RACE IS HEATING UP

- Intel continues to improve their process node (**industry leader**) and update microarchitecture **at steady pace**
- Intel continues to push boundaries with many-core & vector CPUs (KNL)
- 2017-2018 (next two years) are important for ARMv8 64-bit and PowerPC (POWER9)
- Nvidia P100 provides a major improvement from K40 in terms of performance and power efficiency

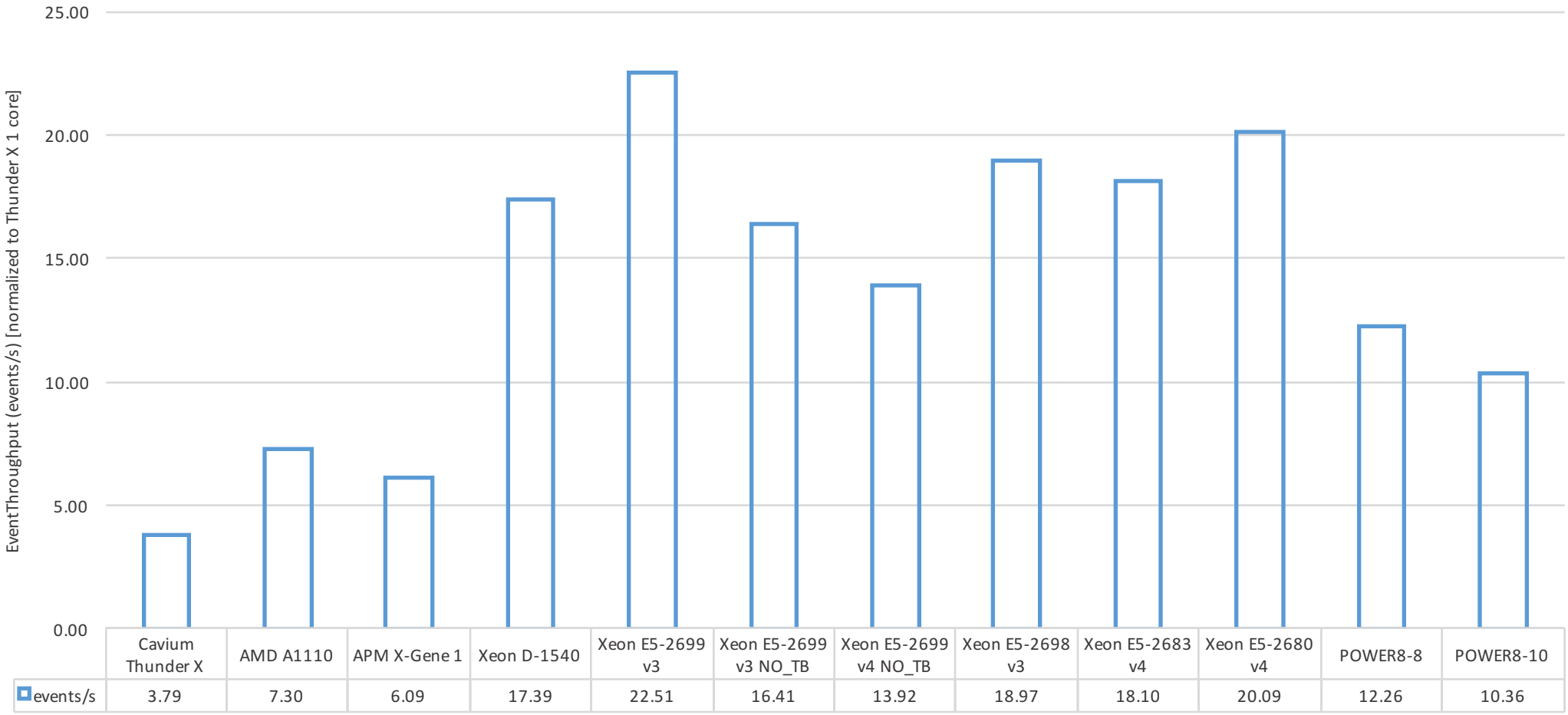
THANK YOU!

For providing hardware and / or helping setting it up we would like to thank:

- Intel
- CERN Techlab and Openlab
- Supermicro
- IBM
- Rackspace

BACKUP

CMSSW RECO: RAW PERFORMANCE – SINGLE JOB

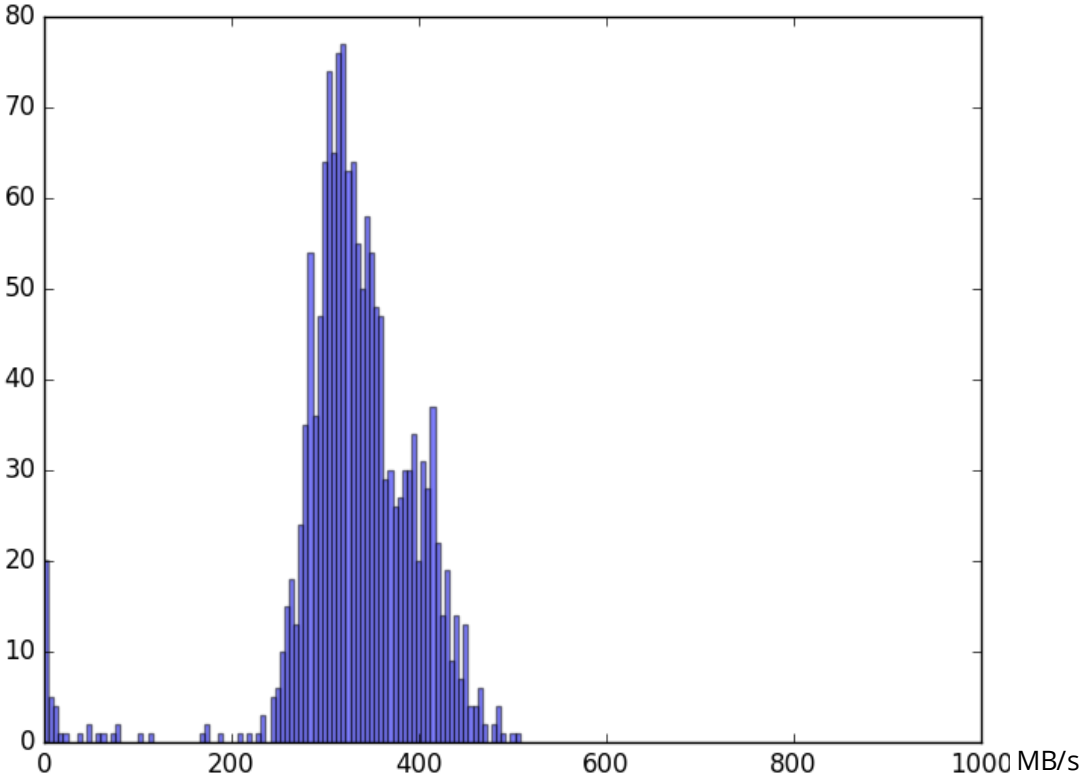


	Cavium Thunder X	AMD A1110	APM X-Gene 1	Xeon D-1540	Xeon E5-2699 v3	Xeon E5-2699 v3 NO_TB	Xeon E5-2699 v4 NO_TB	Xeon E5-2698 v3	Xeon E5-2683 v4	Xeon E5-2680 v4	POWER8-8	POWER8-10
events/s	3.79	7.30	6.09	17.39	22.51	16.41	13.92	18.97	18.10	20.09	12.26	10.36

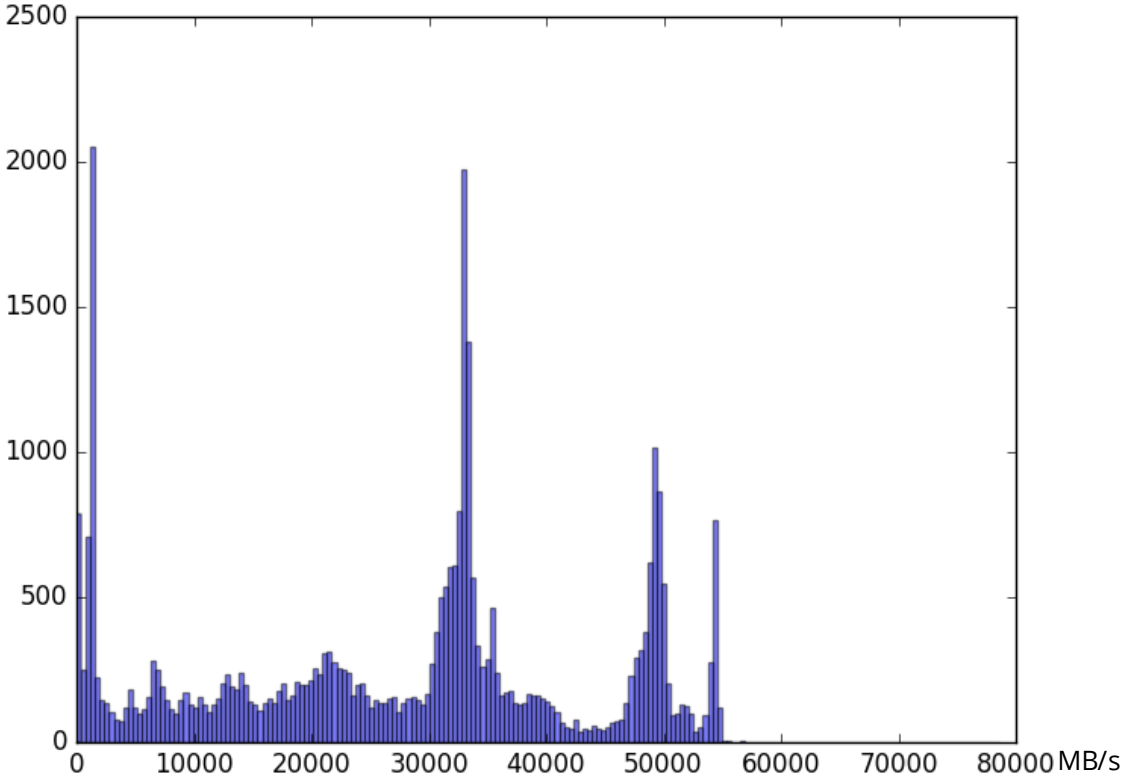
NOTE: All numbers normalized to a single core Thunder X performance.
The job is configured with **4 threads** to represent 2016 executions.

MAIN MEMORY BANDWIDTH (INTEL RDT) – 1S, E5-2683 V4

ParFullCMS



HEP SPEC 2006



STREAM Triad best rate: 40329.8 MB/s (75%) or 53773 MB/s (100%) [Similar number reported by RDT]
All are running with 32 threads (16C/32T, full 1S), 40MB LLC