Disclaimer

Intel provides these materials as-is, with no express or implied warranties.

All products, dates, and figures specified are preliminary, based on current expectations, and are subject to change without notice.

Intel, processors, chipsets, and desktop boards may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No product or component can be absolutely secure. Check with your system manufacturer or retailer or learn more at http://intel.com.

Some results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Intel and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.   For more complete information visit www.intel.com/benchmarks.

Performance results are based on testing as of April 2020 and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure.

*Other names and brands may be claimed as the property of others.

© Intel Corporation

Overview

Motivation: Compare performance impact of core scheduling across different workloads.

Experiments: A fixed configuration running a benchmark toggling the following: HT ON/HT OFF, default kernel/core-sched, .5 overcommit/1 overcommit/2 overcommit. Each VM has its own cgroup.

Overcommitting:

The ratio of total number of virtual CPUs in VM to CPU threads.

- .5 overcommit: number of vCPUs = half of the number of CPU threads
- 1 overcommit: number of vCPUs = number of CPU threads
- 2 overcommit: number of vCPUs = twice the number of CPU threads

For bare metal, it is the ratio of all software threads vs the number of CPU threads.
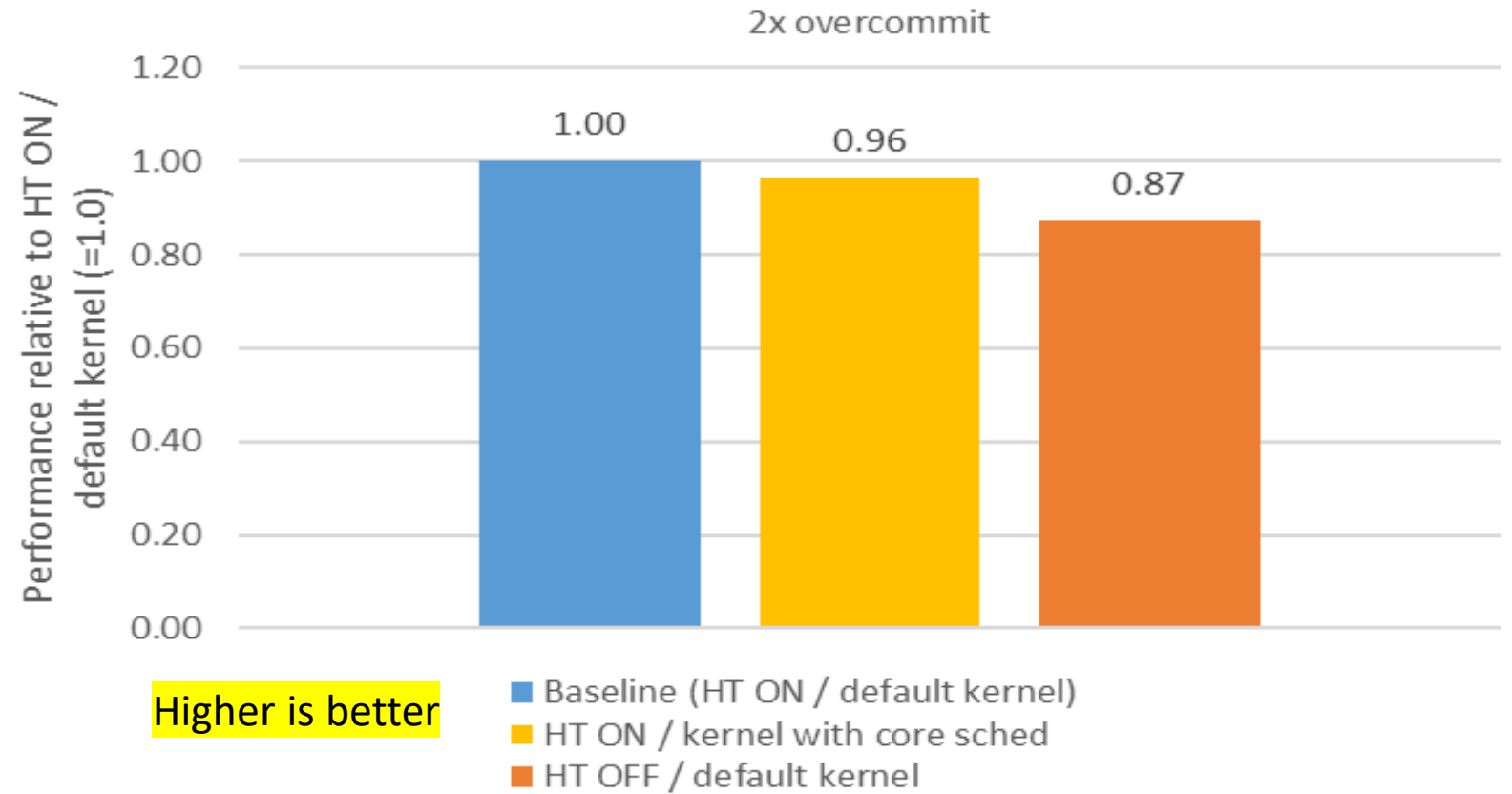
Kernel: https://github.com/digitalocean/linux-coresched/tree/coresched/v5-v5.5.y

System setup

HammerDB
core-sched-v5_v5.5.y

Platform	S2600WFD
Number of Sockets	1
Processors	Intel(R) Xeon(R) Gold 6258R CPU
Number of Cores per Socket	28
Last Level Cache	28.5 MB
Processor Base Frequency	2.70 GHz
Memory	384GB (12 x 32 GB DDR4 @2666)
Number of VMs	2VMs, 56 vcpu/VM, 1vspu/vsocket
OS Distribution	RHEL 8.0 (0otpa)
Host kernel	**core-sched-v5_v5.5.y**
Guest kernel	4.18.0-80.el8.x86_64
BIOS version	SE5C620.86B.0X.02.0141.121620190242
Microcode version	0x500002c
Storage	1x3.2T P4610 for each VM, 2VMs total
Networking	1Gbe NIC
Benchmark version	HammerDB v3.2
Data base	PostgreSQL version 11.5 NOPM
Number of warehouses	1000
Number of virtual users	30 per VM
Testing date	April 2020

Measurements provided for research purposes (on Intel internal reference platforms)

HammerDB
core_sched_v5_v5.5.y
2x overcommit

2x overcommit

Performance relative to HT ON / default kernel (=1.0)

- 1.00 — Baseline (HT ON / default kernel)
- 0.96 — HT ON / kernel with core sched
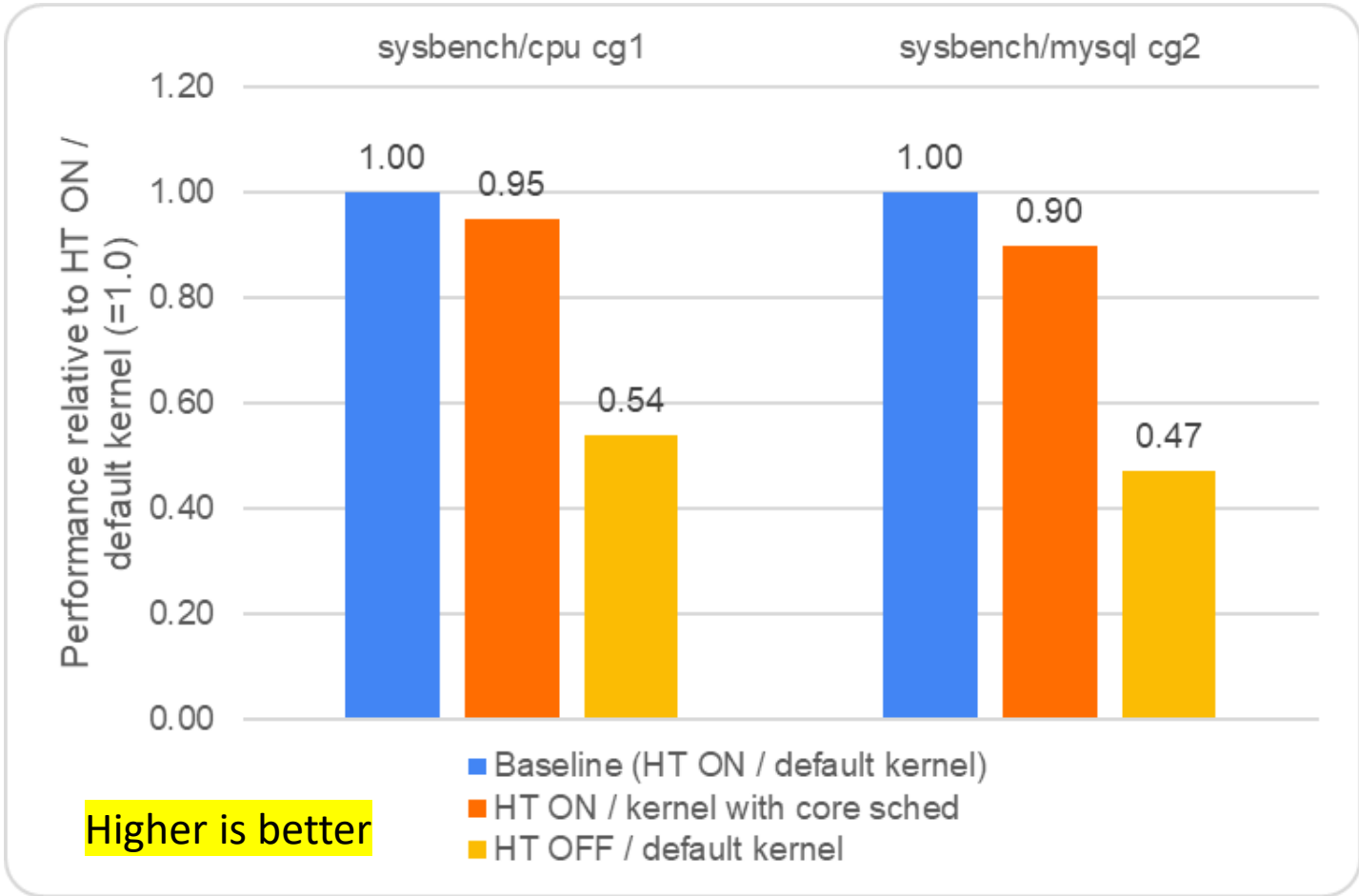- 0.87 — HT OFF / default kernel

Higher is better

- **Database benchmark. 2VMs in 2 cgroups. ~75%-80% cpu utilization in host.**

System setup

Sysbench
core-sched-v5_v5.5.y

Platform	Intel Xeon Server platform
CPU	CLX-AP
Number of Sockets	2
Processors	CLX-AP B0
Number of Cores per Socket	48
CPU max/min MHz	3800/1000 MHz
Host kernel	**core_sched_v5-v5.5.y**
Benchmark version	sysbench 1.1.0-bd4b418
Testing date	April 2020

Sysbench
2x overcommit

sysbench/cpu cg1   sysbench/mysql cg2

Higher is better

- Baseline (HT ON / default kernel)
- HT ON / kernel with core sched
- HT OFF / default kernel

- **2x overcommit, 192 cpu threads, 192 sysbench-cpu in cgroup 1 and 192 sysbench-mysql software threads in cgroup2.**

Sysbench
4x overcommit

Performance relative to HT ON / default kernel (=1.0)

	sysbench/cpu cg1	sysbench/cpu cg2	sysbench/mysql cg3	sysbench/mysql cg4
Baseline (HT ON / default kernel)	1	1	1	1
HT ON / kernel with core sched	0.9	1.32	0.42	0.42
HT OFF / default kernel	0.47	0.66	0.89	0.89

Higher is better
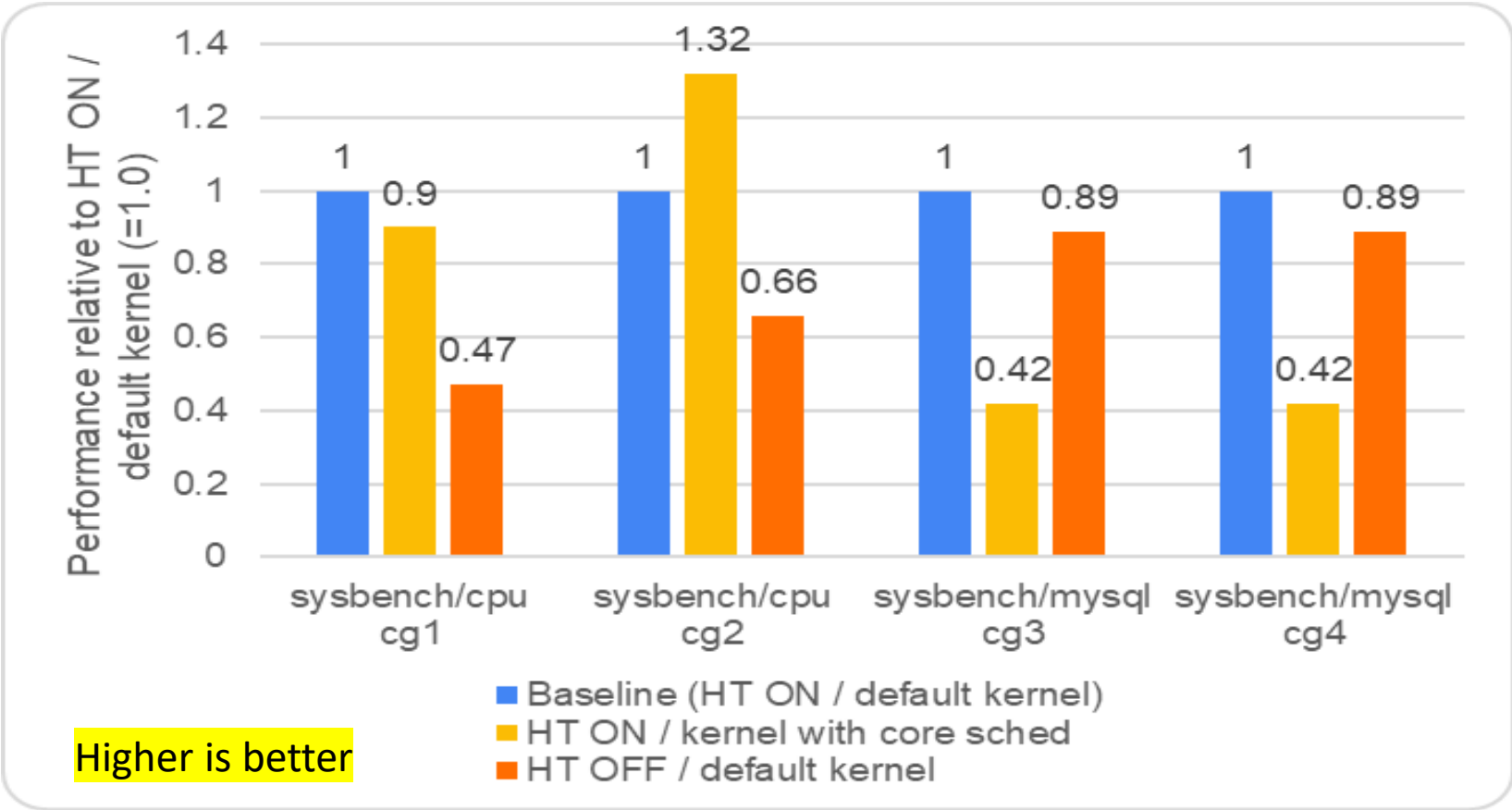
- **4X overcommit. 2 sysbench-cpu cgroups and 2 sysbench-mysql cgroups.**
- Uneven load distribution between the two workload types. Bias towards the cpu intensive workload. Load balancing needs more work.

Measurements provided for research purposes (on Intel internal reference platforms)
For more complete information about performance and benchmark results, visit www.intel.com/benchmarks

8

System setup

Build kernel
core-sched-v5_v5.5.y

Platform	CoffeeLake S 82 UDIMM RVP
Number of Sockets	1
Processors	Intel(R) Core(TM) i9-9900K CPU
Number of Cores per Socket	8
Last Level Cache	16 MB
Processor Base Frequency	3.6 GHz
Memory	32 GB (4 x 8 GB DDR4 @2667)
Memory speed achieved	2666 MHz
OS Distribution	Ubuntu 16.04.6 LTS
Host kernel	**core_sched-v5_v5.5.y**
BIOS version	CNLSFWR1.R00.X214.B00.1910160758
Microcode version	0xca
Storage	1 x 8 GB SSD 1 x 460 GB SSD 1 x 4 GB AEP
Networking	Up to 40Gb NIC
Testing date	April 2020

Kernel build
core_sched_v5_v5.5.y