



arm

NUMA topology limitations

Valentin Schneider <valentin.schneider@arm.com>

25/08/2020

Outline

- LPC19 topology talk recap
- LPC20: "unusual" NUMA topologies
 - sched_group construction issues
 - How much do we care?

Previously

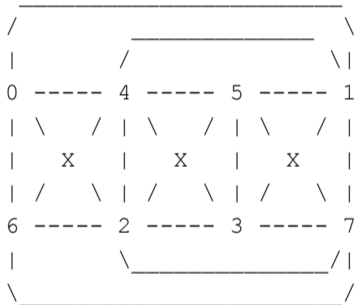
- Cavium ThunderX cache topology quirks^a
- Only relevant arm64 platform; shelved for now
- Recent thing about PowerPC^b looking somewhat similar, might resurface...

^a: <https://linuxplumbersconf.org/event/4/contributions/484/>

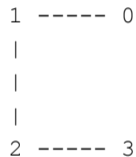
^b: <https://lore.kernel.org/lkml/20200729061355.GA14603@linux.vnet.ibm.com/>

NUMA Diameter

- Longest shortest path (in # hops) between any two nodes



Diameter == 2 [1]



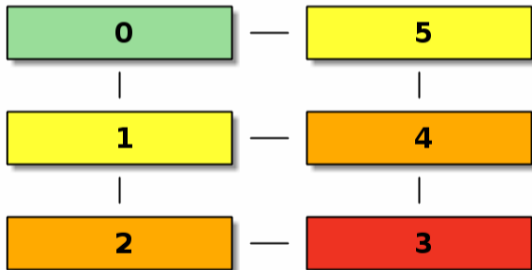
Diameter == 3 [2]

[1] A decade of wasted cores: <http://www.ece.ubc.ca/~sasha/papers/eurosys16-final29.pdf>

[2] [RFC] sched/topology: NUMA topology limitations: <https://lkml.kernel.org/r/jhjtux5edo2.mognet@arm.com>

Example topology - sched_domains

- 1 CPU / node
- NUMA domains: all CPUs within ∞ distance (at most ∞ hops away)



- CPU0 NUMA domains:
 - 0 hops: span=0
 - 1 hops: span=0, 1, 5
 - 2 hops: span=0, 1, 5, 2, 4
 - 3 hops: span=0, 1, 5, 2, 4, 3

Example topology - sched_groups

- sched_domain_span(<2 hops>) == 0-2, 4-5

0, 1, 5, 2, 4



~~0, 1, 5,~~ 2, 4



~~0, 1, 5,~~ 2, 4



- No system crash, but poor resource usage
- Load balance at 2 hops domain:
 - Won't pull tasks from 3 hops away
 - Load of CPUs 3 hops away impacts balancing 2 hops away

In conclusion

- Problem only arises when diameter > 2
 - Hard to find any such system running Linux ATM
- Not trivially fixable
 - New unique groups, which CPU should own them (`sched_group_capacity`)?
 - Won't be the local group of any CPU, never visited by `update_group_capacity()`
- How likely are we to see more diameter > 2 systems?
 - Consider documenting + warning at boot?
 - Feel free to pitch in at <https://lkml.kernel.org/r/jhjtux5edo2.mognet@arm.com>

arm

Thank You

Danke

Merci

谢谢

ありがとう

Gracias

Kiitos

감사합니다

धन्यवाद

شكراً

ধন্যবাদ

תודה

The logo for Arm Limited, consisting of the word "arm" in a lowercase, white, sans-serif font.

The Arm trademarks featured in this presentation are registered trademarks or trademarks of Arm Limited (or its subsidiaries) in the US and/or elsewhere. All rights reserved. All other marks featured may be trademarks of their respective owners.

www.arm.com/company/policies/trademarks