Introducing Last Level Cache Alignment in Kubernetes

New Static CPU Manager Policy for Performance Optimization

Motivation

Reduce Core to Core Latency (ns)

AMD EPYC™ 9755 Cores 0-15

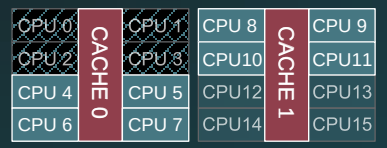
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CPU 0		30	28	32	29	31	26	26	121	122	121	122	117	120	115	122
CPU 1	30		30	30	29	31	29	29	122	122	121	122	115	118	115	117
CPU 2	28	30		30	27	31	23	29	120	122	119	122	117	116	115	115
CPU 3	32	30	30		29	29	25	29	122	122	121	122	121	122	116	117
CPU 4	29	29	27	29		25	26	27	119	121	119	121	117	121	115	115
CPU 5	31	31	31	29	25		31	29	119	121	119	120	115	121	115	115
CPU 6	26	29	23	25	26	31		27	118	121	116	116	115	119	115	115
CPU 7	26	29	29	29	27	29	27		121	121	117	122	115	117	116	117
CPU 8	121	122	120	122	119	119	118	121		30	28	31	29	31	26	25
CPU 9	122	122	122	122	121	121	121	121	30		30	30	29	31	30	29
CPU 10	121	121	119	121	119	119	116	117	28	30		30	27	31	23	29
CPU 11	122	122	122	122	121	120	116	122	31	30	30		29	29	25	29
CPU 12	117	115	117	121	117	115	115	115	29	29	27	29		25	26	27
CPU 13	120	118	116	122	121	121	119	117	31	31	31	29	25		31	29
CPU 14	115	115	115	116	115	115	115	116	26	30	23	25	26	31		27
CPU 15	122	117	115	117	115	115	115	117	25	29	29	29	27	29	27	

Default Static CPUManager

Cache Misaligned

8 CPUs / CPU 4-11

CPU 0-3 Reserved

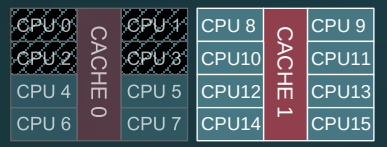


Align Uncore Cache Enabled

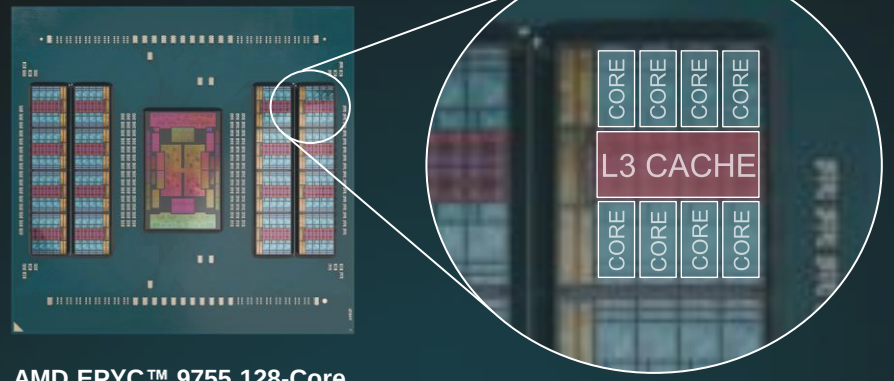
Cache Aligned

8 CPUs / CPU 8-15

CPU 0-3 Reserved



What is Uncore / Last Level Cache



AMD EPYC™ 9755 128-Core Processor

16x **CCDs** / Uncore Caches

"Core Complex Die" (CCD):

8 Cores per Uncore Cache

Modular Chiplet Architecture

Distributed Uncore Cache

x86 and ARM Processors

AMD® EYPC™

ARM® Neoverse™

2.8x Faster

	Average Core to Core Latency
Default Static CPUManager **Cache Misaligned** 8 CPUs / CPU 4-11	**80 ns**
Align Uncore Cache Enabled **Cache Aligned** 8 CPUs / CPU 8-15	**28 ns**

Solution

Enable Static Policy in K8s v1.32+

prefer-align-cpus-by-uncorecache

CPUManagerPolicyAlphaOptions

full-pcpus-only (if SMT-enabled)

Best Effort Alignment Policy

Minimizes CPU distribution across the fewest Uncore Caches for performance optimization

Vendor Agnostic

No impact when enabled on Monolithic Uncore Cache Processors

Performance Jitter Reduction

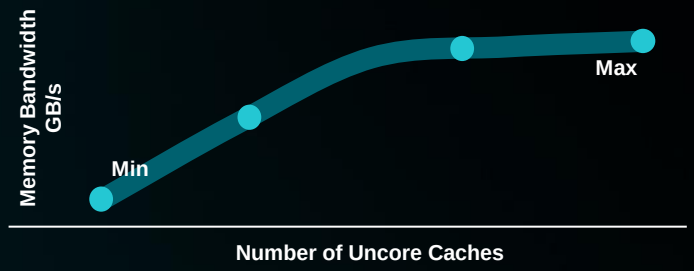
Cache Alignment can reduce performance variation

Use Cases

vRAN

5G Mobile Packet Core

Firewall



prefer-align-cpus-by-uncorecache

Enhancement Proposal

KubeCon CloudNativeCon Europe 2025

AMD together we advance_