

Policy zones: memory partitioning for fun and profit

Yu Zhao <yuzhao@google.com>



Status quo

There are three types of zones:

1. First four zones partition the physical address space of CPU memory.
2. Device zone provides interoperability between CPU and device memory.
3. Movable zone commonly represents a memory allocation policy.

Other specialized pools somewhat similar to zones:

1. CMA
2. HugeTLB



Design goals

The short-term goals are:

1. Support allocation time policies, e.g., TAO.
2. Support runtime policies, e.g., Tetris.

The long-term goals are:

1. Eliminate the need of specialized pools.
2. Eliminate DMA/DMA32 zones.

Fun fact: there is a new MC this year– Zoned Storage MC!



LINUX
PLUMBERS

CONFERENCE Vienna, Austria / Sept. 18-20, 2024

Use cases

Community:

- Android (OPPO).

At Google:

- Leveraging HVO to reduce THP struct page overhead in data centers.

Future opportunities:

- 1GB THPs.
- Larger device HMM page sizes, e.g., 4KB CPU + 64KB GPU base page sizes.



Zones vs pageblocks

Why not new pageblock types?

	Purpose	Lifetime	Granularity	Controllability	Observability
Pageblocks	Grouping by mobility	Allocation dependent	Small, immutable	Within buddy allocator	Limited
Zones	Partitioning memory	System dependent	Large, mutable	Within core MM	Acceptable

