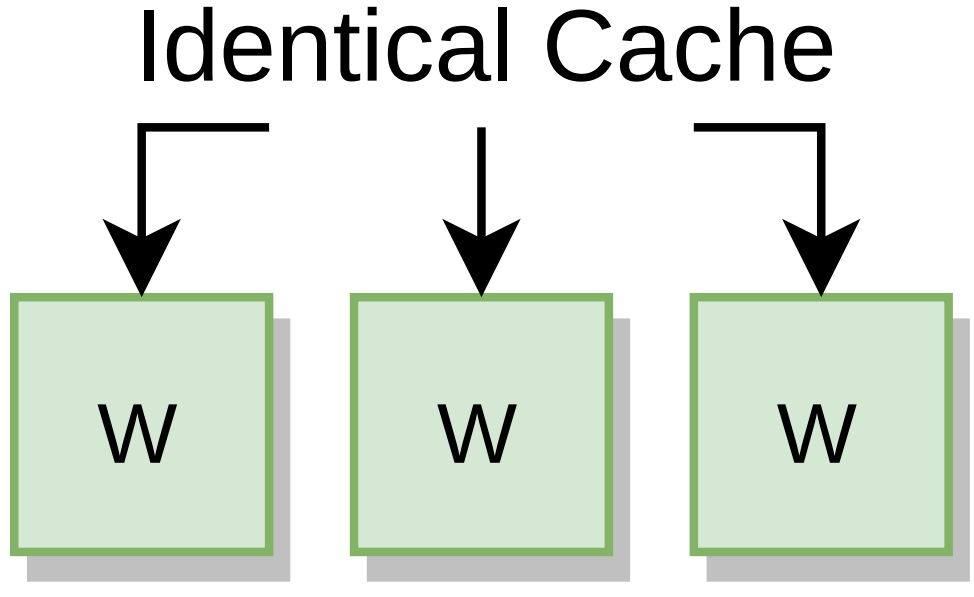
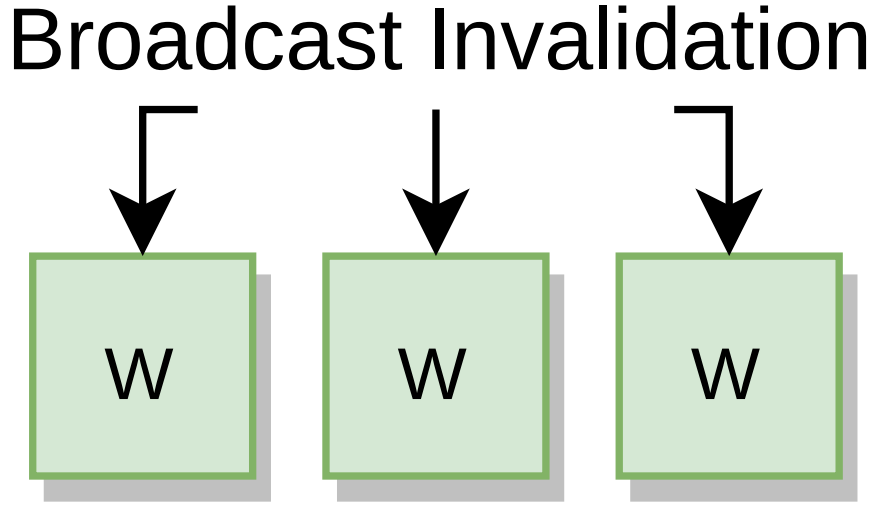
CACHING BEYOND RAM

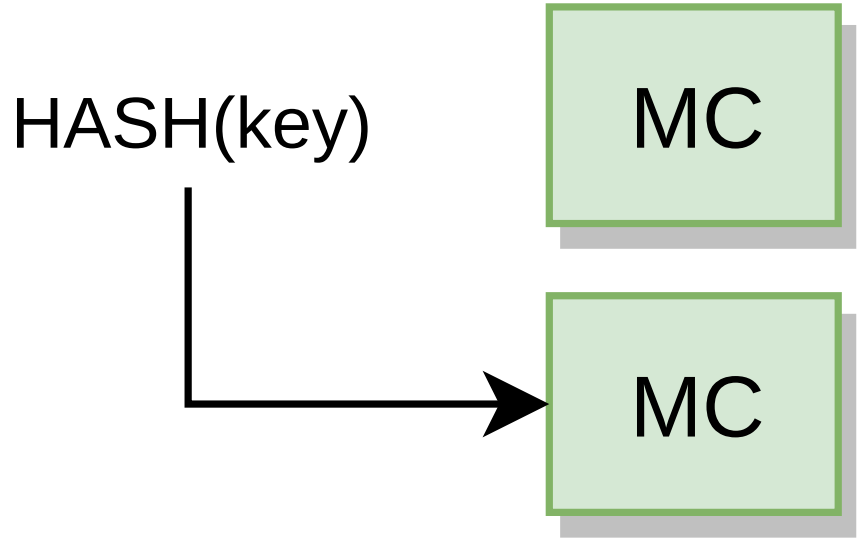
memcached.org/blog

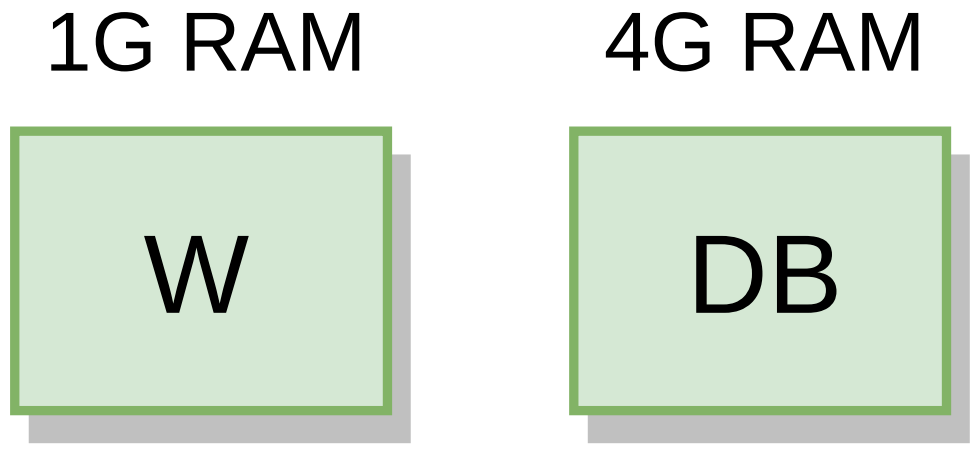
@dormando

WHY RAM?

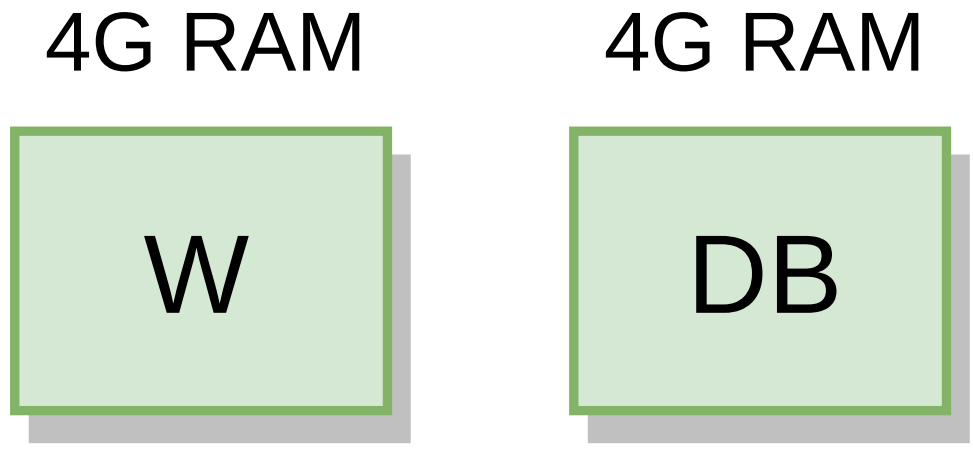

Identical Cache
w
w
w


Broadcast Invalidation
W
W
W

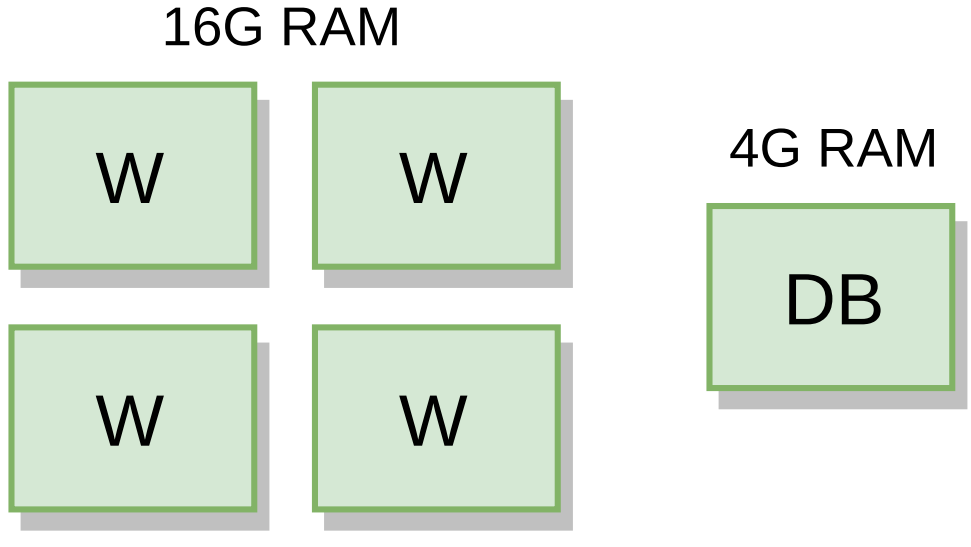

HASH(key)
MC
MC



32bit OS!



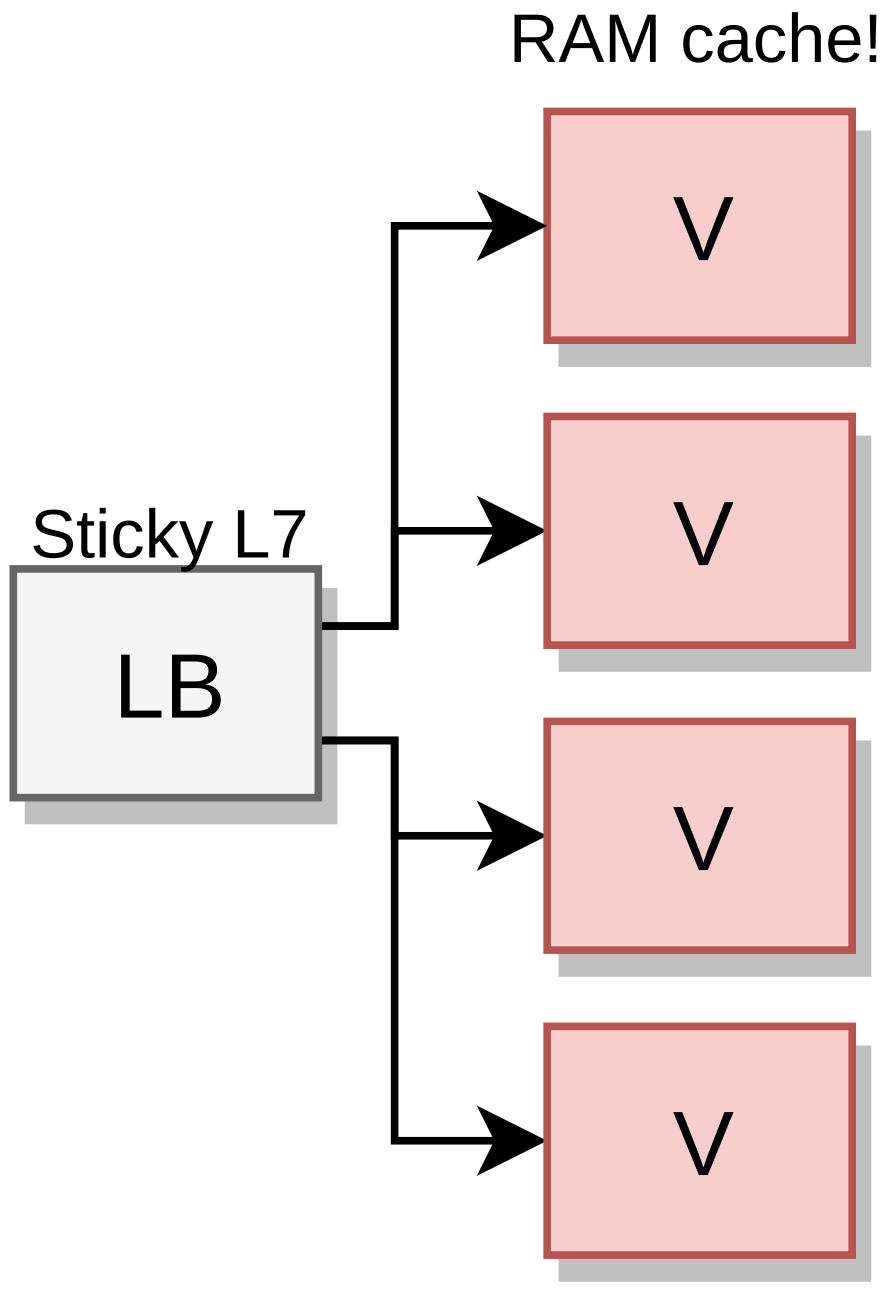
Filled Empty
RAM Slots

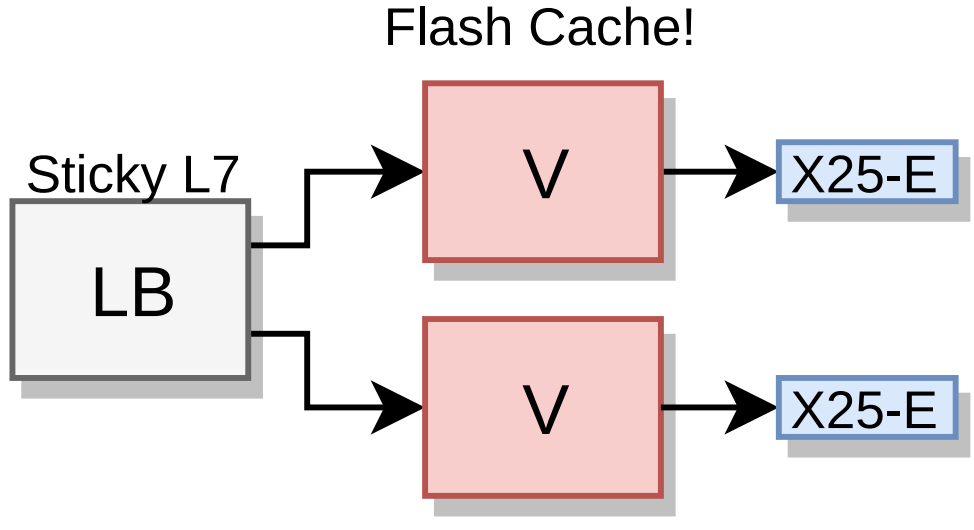

16G RAM
W
W
W
W
4G RAM
DB

ENTER FLASH

FLASH! NOT JUST FOR HEAVY MACHINERY

- X25-M 80G IOPS: 35k read, 300 write
- X25-E 64G IOPS: 35k read, 3300 write

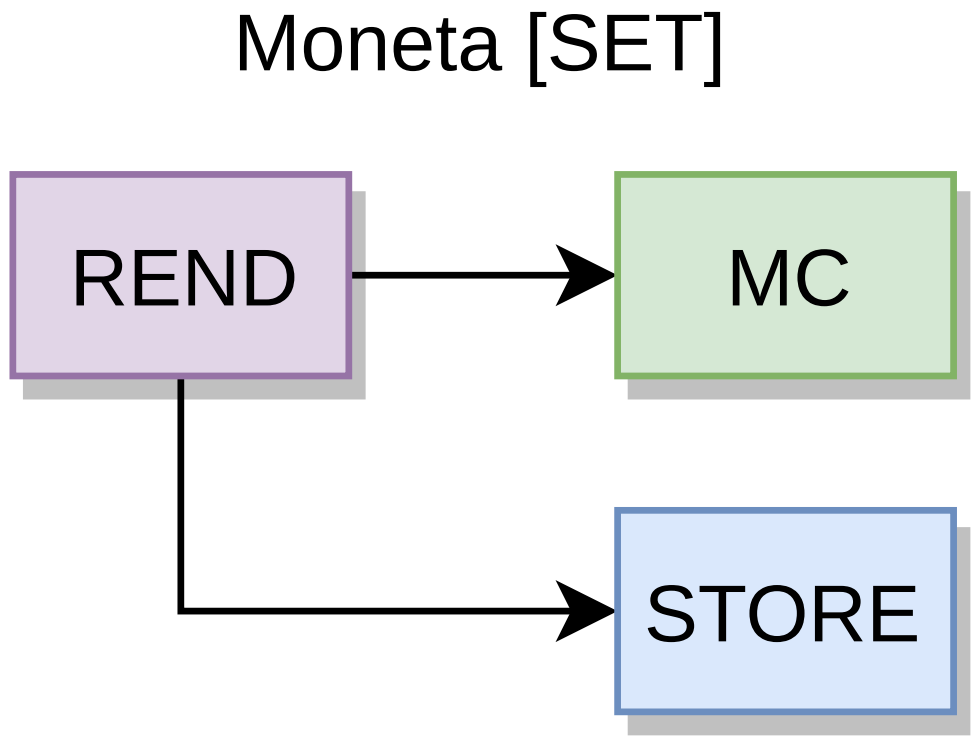

RAM cache!
V
V
V
V
Sticky L7
LB

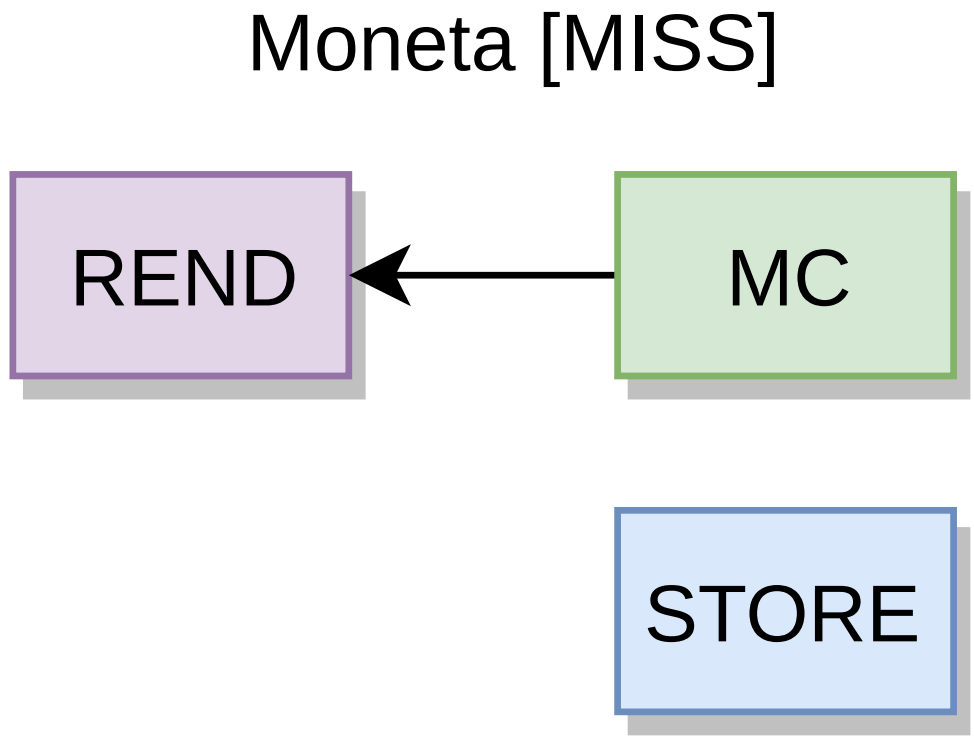

Flash Cache!
Sticky L7
LB
V
V
X25-E
X25-E

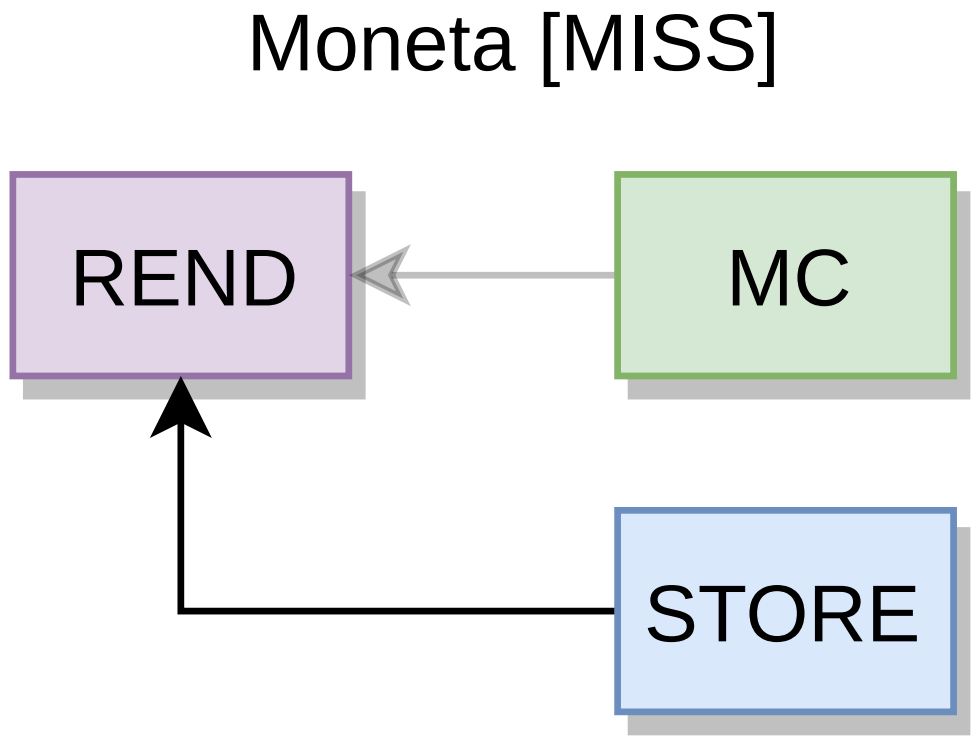
BAM! CDN

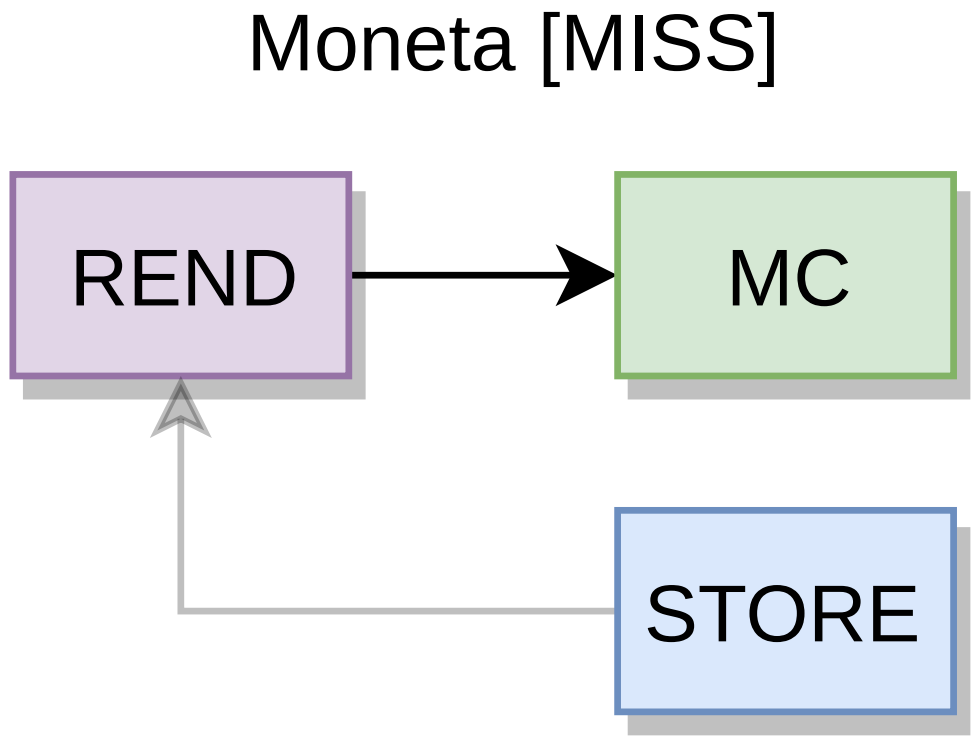
GOOD (ENOUGH) SSD'S ARE EVERYWHERE

TRADEOFFS


Moneta [SET]
REND
MC
STORE


Moneta [MISS]
REND
MC
STORE


Moneta [MISS]
REND
MC
STORE

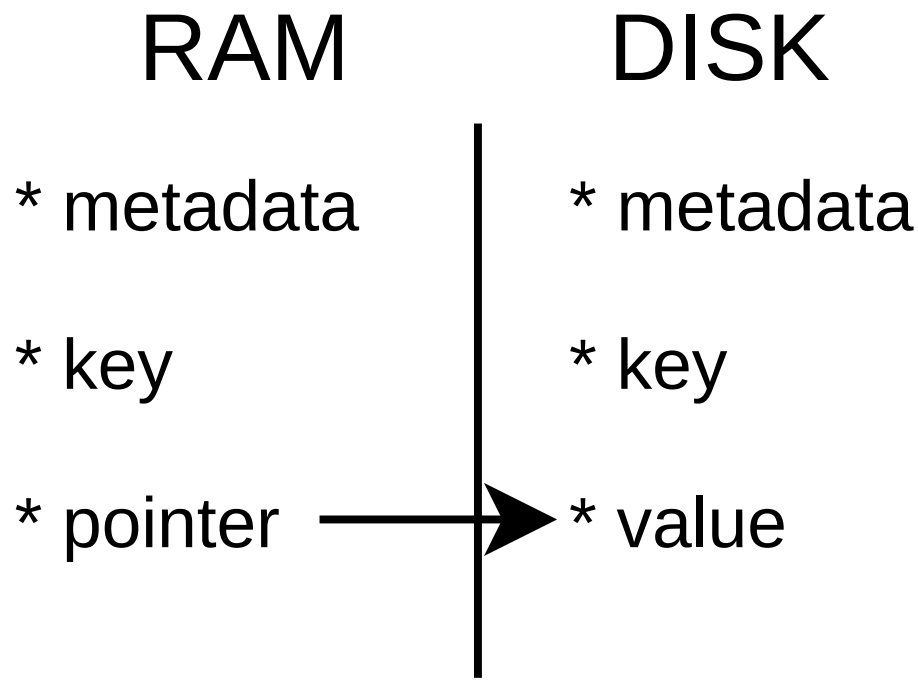

Moneta [MISS]
REND
MC
STORE

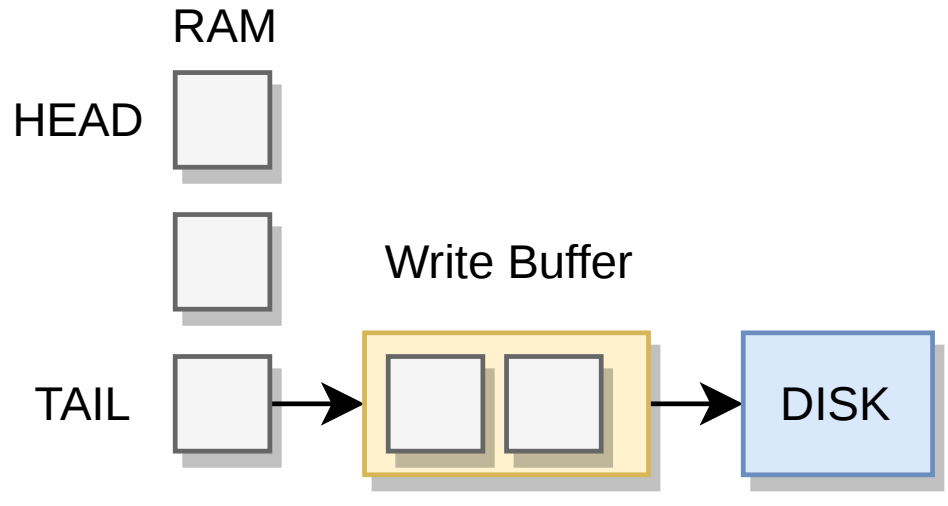
CACHE -> DB -> DB

ARE SMALL ITEMS VALUABLE ON DISK CACHE?

NO.

MEMCACHED EXTSTORE


RAM
DISK
* metadata
* key
* pointer
* metadata
* key
* value


RAM
HEAD
Write Buffer
TAIL
DISK

Bad: Still limited by RAM.
Good: Much less writing, consistent reads.

WORKLOADS

- ML facts / computed data / templated data
- Sessions :P
- Saving the 50% of RAM used in 8k+ items

FUTURE

JBOD [JUST A BUNCH OF DEVICES]

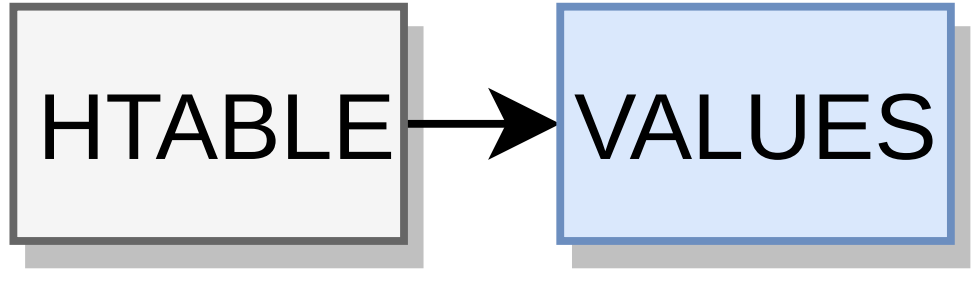
```
memcached -o
ext_path=/m/f:64g,ext_path=/m2/f:64g
```

TIERED STORAGE

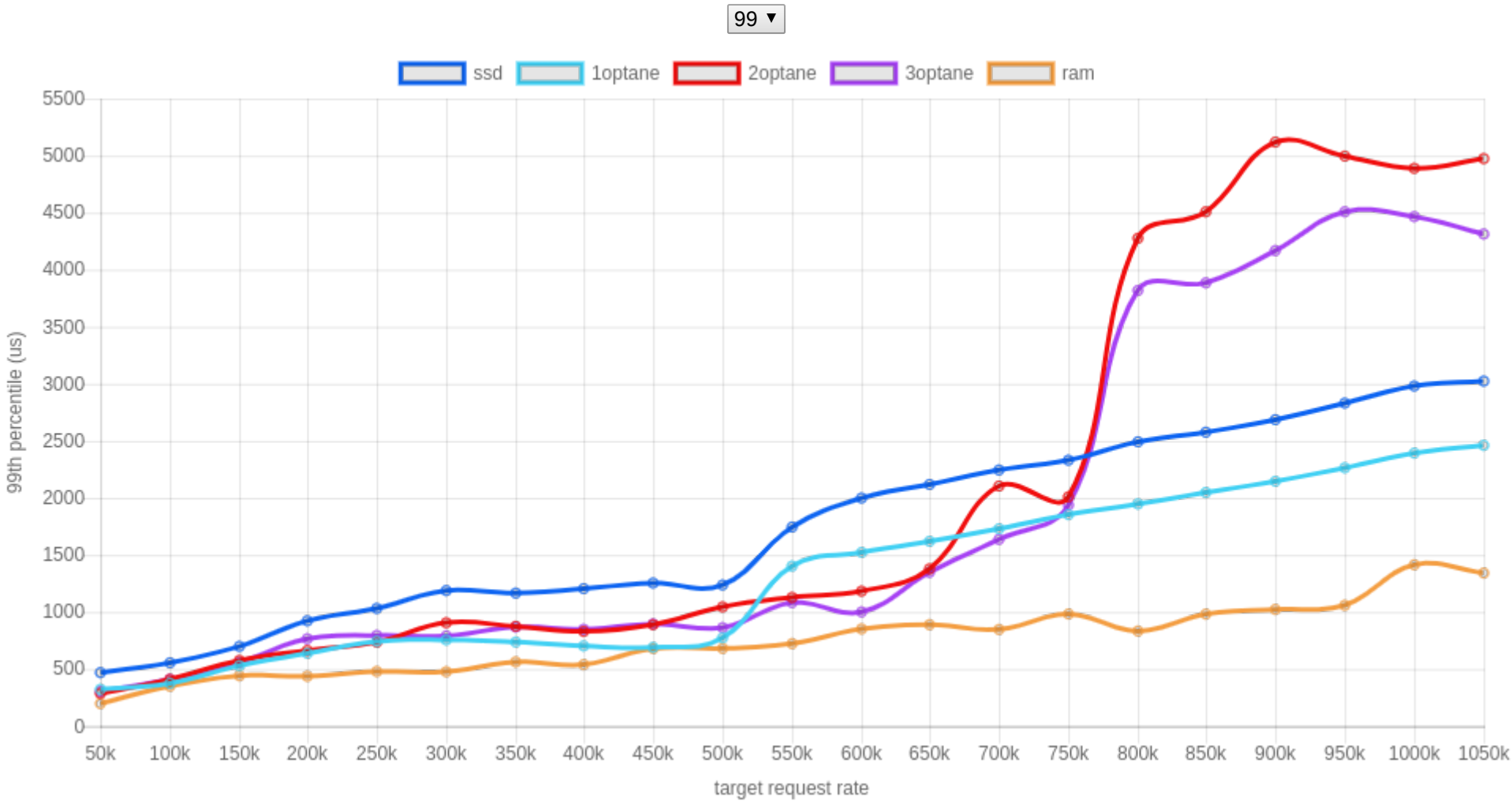
github.com/memcached/memcached/pull/432

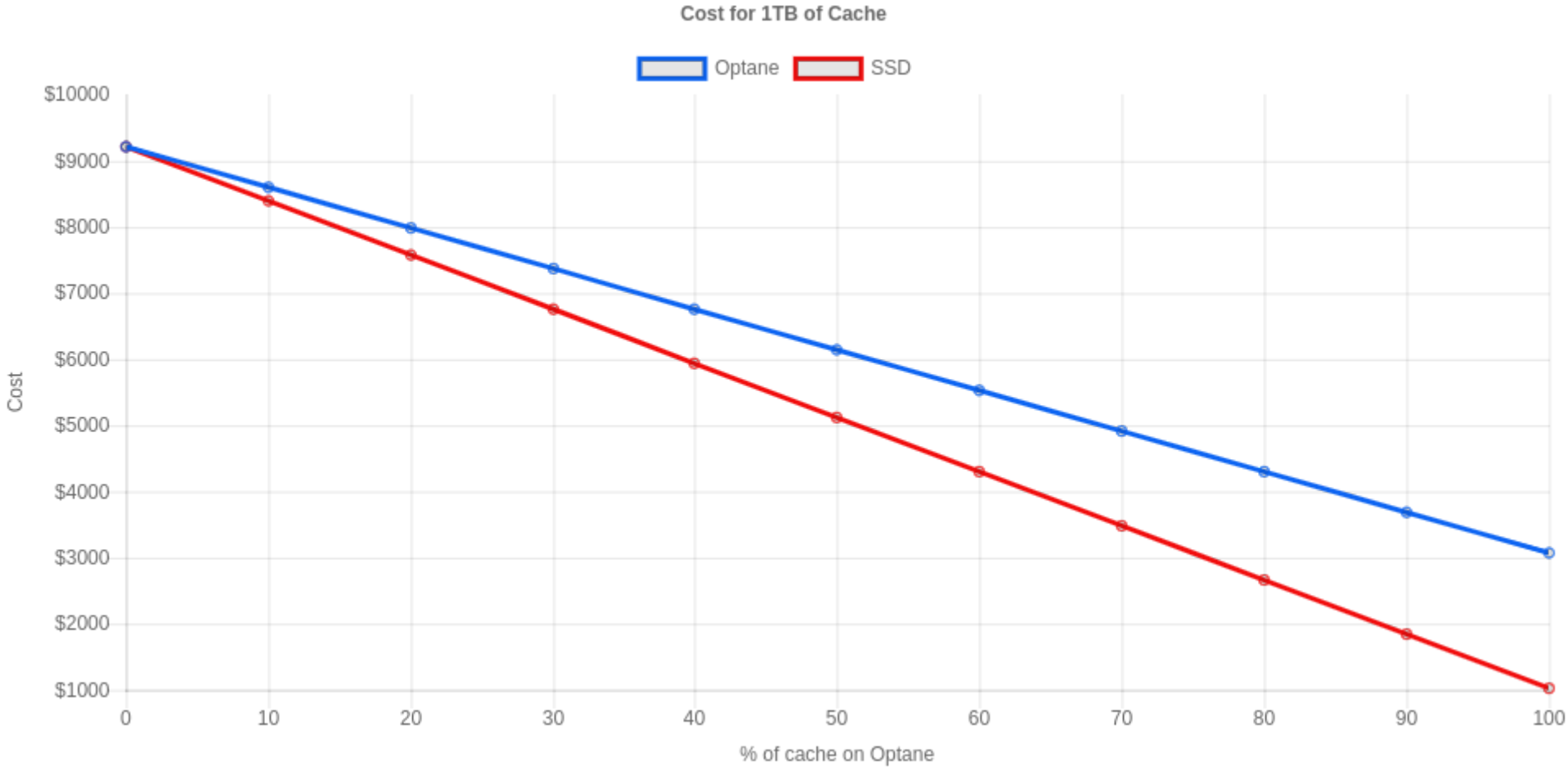
```
ext_path=/m/f:64g:compact
ext_path=/m/f:64g:lowttl
```

PERSISTENT MEMORY



WRAP UP: PERFORMANCE


99
ssd
1optane
2optane
3optane
ram
99th percentile (us)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
50k
100k
150k
200k
250k
300k
350k
400k
450k
500k
550k
600k
650k
700k
750k
800k
850k
900k
950k
1000k
1050k
target request rate


Cost for 1TB of Cache
Optane
SSD
$10000
$9000
$8000
$7000
$6000
$5000
$4000
$3000
$2000
$1000
Cost
0
10
20
30
40
50
60
70
80
90
100
% of cache on Optane

THANKS!

memcached.org/blog

@dormando

github.com/memcached/memcached/wiki/Extstore