



Co dělá Linux s mojí RAMkou?

Vlastimil Babka

OpenAlt 2025

The textbook command: free (1)

- `> free -m`

	total	used	free	shared	buff/cache	available
Mem:	31372	10740	8444	596	13234	20631
Swap:	2454	0	2454			

- `used + free != total` ? ($19184 < 31372$) – What's going on here ?
- Linux memory management principle: unused memory is wasted memory
 - some memory will be allocated directly by processes (`malloc()`, "new" → `mmap()`)
 - but the kernel will also cache all accessed file contents (and various metadata) in memory, as long as free memory exists
 - once free memory is (almost) depleted, the cache is relatively easy to discard
 - *used* tries to estimate how much is occupied that can't easily be discarded
 - otherwise the number wouldn't be really useful (always close to total)

The textbook command: free (1)

- `> free -m`

	total	used	free	shared	buff/cache	available
Mem:	31372	10740	8444	596	13234	20631
Swap:	2454	0	2454			

- `used + free + buff/cache != total ? (32418 > 31372)`
 - is perhaps part of *buff/cache* considered *used* ?
- `free + buff/cache != available ? (21678 > 20631)`
 - part of *free* or *buff/cache* is not considered *available* ?
- `used + available == total? (31371 ~= 31372)`
 - this looks promising
 - but what is *available* ? and what is *shared* ?

Kernel's accounting view: /proc/meminfo

MemTotal:	32125696 kB	AnonPages:	8922944 kB	AnonHugePages:	1363968 kB
MemFree:	8627220 kB	Mapped:	1683592 kB	ShmemHugePages:	0 kB
MemAvailable:	21107232 kB	Shmem:	611016 kB	ShmemPmdMapped:	0 kB
Buffers:	258820 kB	KReclaimable:	390056 kB	FileHugePages:	1062912 kB
Cached:	12902952 kB	Slab:	613748 kB	FilePmdMapped:	684032 kB
SwapCached:	0 kB	SReclaimable:	390056 kB	CmaTotal:	0 kB
Active:	5870672 kB	SUnreclaim:	223692 kB	CmaFree:	0 kB
Inactive:	16277940 kB	KernelStack:	35568 kB	Unaccepted:	0 kB
Active(anon):	8120 kB	PageTables:	82500 kB	Balloon:	0 kB
Inactive(anon):	9589756 kB	SecPageTables:	3248 kB	HugePages_Total:	0
Active(file):	5862552 kB	NFS_Unstable:	0 kB	HugePages_Free:	0
Inactive(file):	6688184 kB	Bounce:	0 kB	HugePages_Rsvd:	0
Unevictable:	80 kB	WritebackTmp:	0 kB	HugePages_Surp:	0
Mlocked:	80 kB	CommitLimit:	18575904 kB	Hugepagesize:	2048 kB
SwapTotal:	2513056 kB	Committed_AS:	21267524 kB	Hugetlb:	0 kB
SwapFree:	2513056 kB	VmallocTotal:	34359738367 kB	DirectMap4k:	894136 kB
Zswap:	0 kB	VmallocUsed:	84632 kB	DirectMap2M:	25239552 kB
Zswapped:	0 kB	VmallocChunk:	0 kB	DirectMap1G:	7340032 kB
Dirty:	844 kB	Percpu:	7680 kB		
Writeback:	196 kB	HardwareCorrupted:	0 kB		

MemTotal:	32125696 kB	AnonPages:	8922944 kB	AnonHugePages:	1363968 kB
MemFree:	8627220 kB	Mapped:	1683592 kB	ShmemHugePages:	0 kB
MemAvailable:	21107232 kB	Shmem:	611016 kB	ShmemPmdMapped:	0 kB
Buffers:	258820 kB	KReclaimable:	390056 kB	FileHugePages:	1062912 kB
Cached:	12902952 kB	Slab:	613748 kB	FilePmdMapped:	684032 kB
SwapCached:	0 kB	SReclaimable:	390056 kB	CmaTotal:	0 kB
Active:	5870672 kB	SUnreclaim:	223692 kB	CmaFree:	0 kB
Inactive:	16277940 kB	KernelStack:	35568 kB	Unaccepted:	0 kB
Active(anon):	8120 kB	PageTables:	82500 kB	Balloon:	0 kB
Inactive(anon):	9589756 kB	SecPageTables:	3248 kB	HugePages_Total:	0
Active(file):	5862552 kB	NFS_Unstable:	0 kB	HugePages_Free:	0
Inactive(file):	6688184 kB	Bounce:	0 kB	HugePages_Rsvd:	0
Unevictable:	80 kB	WritebackTmp:	0 kB	HugePages_Surp:	0
Mlocked:	80 kB	CommitLimit:	18575904 kB	Hugepagesize:	2048 kB
SwapTotal:	2513056 kB	Committed_AS:	21267524 kB	Hugetlb:	0 kB
SwapFree:	2513056 kB	VmallocTotal:	34359738367 kB	DirectMap4k:	894136 kB
Zswap:	0 kB	VmallocUsed:	84632 kB	DirectMap2M:	25239552 kB
Zswapped:	0 kB	VmallocChunk:	0 kB	DirectMap1G:	7340032 kB
Dirty:	844 kB	Percpu:	7680 kB		
Writeback:	196 kB	HardwareCorrupted:	0 kB		

MemTotal:	32125696 kB	AnonPages:	8922944 kB	AnonHugePages:	1363968 kB
MemFree:	8627220 kB	Mapped:	1683592 kB	ShmemHugePages:	0 kB
MemAvailable:	21107232 kB	Shmem:	611016 kB	ShmemPmdMapped:	0 kB
Buffers:	258820 kB	KReclaimable:	390056 kB	FileHugePages:	1062912 kB
Cached:	12902952 kB	Slab:	613748 kB	FilePmdMapped:	684032 kB
SwapCached:	0 kB	SReclaimable:	390056 kB	CmaTotal:	0 kB
Active:	5870672 kB	SUnreclaim:	223692 kB	CmaFree:	0 kB
Inactive:	16277940 kB	KernelStack:	35568 kB	Unaccepted:	0 kB
Active(anon):	8120 kB	PageTables:	82500 kB	Balloon:	0 kB
Inactive(anon):	9589752 kB			HugePages_Total:	0
Active(file):	5862552 kB			HugePages_Free:	0
Inactive(file):	6688184 kB			HugePages_Rsvd:	0
Unevictable:	8192 kB			HugePages_Surp:	0
Mlocked:	8192 kB			Hugepagesize:	2048 kB
SwapTotal:	2513056 kB	Committed_AS:	21107232 kB	Hugetlb:	0 kB
SwapFree:	2513056 kB	VmallocTotal:	3435973856 kB	DirectMap4k:	894136 kB
Zswap:	0 kB	VmallocUsed:	84632 kB	DirectMap2M:	25239552 kB
Zswapped:	0 kB	VmallocChunk:	0 kB	DirectMap1G:	7340032 kB
Dirty:	844 kB	Percpu:	7680 kB		
Writeback:	196 kB	HardwareCorrupted:	0 kB		

Property of kernel's direct mapping, not exactly occupied memory

MemTotal:	32125696 kB	AnonPages:	8922944 kB	AnonHugePages:	1363968 kB
MemFree:	8627220 kB	Mapped:	1683592 kB	ShmemHugePages:	0 kB
MemAvailable:	21107232 kB	Shmem:	611016 kB	ShmemPmdMapped:	0 kB
Buffers:	258820 kB			FileHugePages:	1062912 kB
Cached:	12902952 kB			FilePmdMapped:	684032 kB
SwapCached:				CmaTotal:	0 kB
Active:	5870672 kB			CmaFree:	0 kB
Inactive:	16277940 kB			Unaccepted:	0 kB
Active(anon):	8120 kB	PageTables:	8256 kB	Balloon:	0 kB
Inactive(anon):	9589756 kB	SecPageTables:	3248 kB	HugePages_Total:	0
Active(file):	5862552 kB	NFS_Unstable:	0 kB	HugePages_Free:	0
Inactive(file):	6688184 kB	Bounce:	0 kB	HugePages_Rsvd:	0
Unevictable:	80 kB	WritebackTmp:	0 kB	HugePages_Surp:	0
Mlocked:	80 kB	CommitLimit:	18575904 kB	Hugepagesize:	2048 kB
SwapTotal:	2513056 kB	Committed_AS:	21267524 kB	Hugetlb:	0 kB
SwapFree:	2513056 kB	VmallocTotal:	34359738367 kB	DirectMap4k:	894136 kB
Zswap:	0 kB	VmallocUsed:	84632 kB	DirectMap2M:	25239552 kB
Zswapped:	0 kB	VmallocChunk:	0 kB	DirectMap1G:	7340032 kB
Dirty:	844 kB	Percpu:	7680 kB		
Writeback:	196 kB	HardwareCorrupted:	0 kB		

Hugetlbfs details, beyond
the scope here

MemTotal:	32125696 kB	AnonPages:	8922944 kB	AnonHugePages:	1363968 kB
MemFree:				ShmemHugePages:	0 kB
MemAvailable:				ShmemPmdMapped:	0 kB
Buffers:				FileHugePages:	1062912 kB
Cached:				FilePmdMapped:	684032 kB
SwapCached:				CmaTotal:	0 kB
Active:				CmaFree:	0 kB
Inactive:	16277940 kB	KernelStack:	35568 kB	Unaccepted:	0 kB
Active(anon):	8120 kB	Pagetable:	82500 kB	Balloon:	0 kB
Inactive(anon):	9589756 kB	Section:	3248 kB	HugePages_Total:	0
Active(file):	5862552 kB	NonFileBacked:	0 kB	HugePages_Free:	0
Inactive(file):	6688184 kB	Backing:	0 kB	HugePages_Rsvd:	0
Unevictable:	80 kB	WritebackTmp:	0 kB	HugePages_Surp:	0
Mlocked:	80 kB	CommitLimit:	18575904 kB	Hugepagesize:	2048 kB
SwapTotal:	2513056 kB	Committed_AS:	21267524 kB	Hugetlb:	0 kB
SwapFree:	2513056 kB	VmallocTotal:	34359738367 kB	DirectMap4k:	894136 kB
Zswap:	0 kB	VmallocUsed:	84632 kB	DirectMap2M:	25239552 kB
Zswapped:	0 kB	VmallocChunk:	0 kB	DirectMap1G:	7340032 kB
Dirty:	844 kB	Percpu:	7680 kB		
Writeback:	196 kB	HardwareCorrupted:	0 kB		

Size of virtual region in the kernel
address space where `vmalloc()`
areas are placed
(rather useless number...)

MemTotal:	321256		AnonHugePages:	1363968	kB
MemFree:	8627		memHugePages:	0	kB
MemAvailable:	21107		memPmdMapped:	0	kB
Buffers:	258		leHugePages:	1062912	kB
Cached:	12902952	kB	FilePmdMapped:	684032	kB
SwapCached:	0	kB	CmaTotal:	0	kB
Active:	5870672	kB	CmaFree:	0	kB
Inactive:	16277940	kB	Unaccepted:	0	kB
Active(anon):	8120	kB	Balloon:	0	kB
Inactive(anon):	9589756	kB	HugePages_Total:	0	
Active(file):	5862552	kB	HugePages_Free:	0	
Inactive(file):	6688		HugePages_Rsvd:	0	
Unevictable:			HugePages_Surp:	0	
Mlocked:			Hugepagesize:	2048	kB
SwapTotal:	2513		Hugetlb:	0	kB
SwapFree:	2513056	kB	DirectMap4k:	894136	kB
Zswap:	0	kB	DirectMap2M:	25239552	kB
Zswapped:	0	kB	DirectMap1G:	7340032	kB
Dirty:	844	kB			
Writeback:	196	kB			

Confidential computing guests,
temporary after boot

RAM discarded as bad due
to HW reported errors (ECC)

HardwareCorrupted: 0 kB

MemTotal:	32125696	14 kB	AnonHugePages:	1363968	kB
MemFree:	862722	2 kB	ShmemHugePages:	0	kB
MemAvailable:	2110723	5 kB	ShmemPmdMapped:	0	kB
Buffers:	25882	5 kB	FileHugePages:	1062912	kB
Cached:	1290295	3 kB	FilePmdMapped:	684032	kB
SwapCached:	0	kB	CmaTotal:	0	kB
Active:	5870672	kB	CmaFree:	0	kB
Inactive:	16277940	kB	Unaccepted:	0	kB
Active(anon):	8120	kB	Balloon:	0	kB
Inactive(anon):	9589756	kB	HugePages_Total:	0	
Active(file):	5862552	kB	HugePages_Free:	0	
Inactive(file):	6688184	kB	HugePages_Rsvd:	0	
Unevictable:	80	kB	HugePages_Surp:	0	
Mlocked:	80	kB	Hugepagesize:	2048	kB
SwapTotal:	2513056	kB	Hugetlb:	0	kB
SwapFree:	2513056	kB	DirectMap4k:	894136	kB
Zswap:	0	kB	DirectMap2M:	25239552	kB
Zswapped:	0	kB	DirectMap1G:	7340032	kB
Dirty:	844	kB			
Writeback:	196	kB			

Used in the past, now
unused, hardcoded 0
for compatibility

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

(free -m)	total	used	free	shared	buff/cache	available
Mem:	31372	10740	8444	596	13234	20631
Swap:	2454	0	2454			

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

(free -m)	total	used	free	shared	buff/cache	available
Mem:	31372	10740	8444	596	13234	20631
Swap:	2454	0	2454			

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal: 32125696 kB

MemFree: 8627220 kB

MemAvailable: 21107232 kB

Let's look at these first

SwapCached: 0 kB

Active: 5870672 kB

Inactive: 16277940 kB

Active(anon): 8120 kB

Inactive(anon): 9589756 kB

Active(file): 5862552 kB

Inactive(file): 6688184 kB

Unevictable: 80 kB

Mlocked: 80 kB

SwapTotal: 2513056 kB

SwapFree: 2513056 kB

Zswap: 0 kB

Zswapped: 0 kB

Dirty: 844 kB

Writeback: 196 kB

AnonPages: 8922944 kB

Mapped: 1683592 kB

Shmem: 611016 kB

KReclaimable: 390056 kB

Slab: 613748 kB

SReclaimable: 390056 kB

SUnreclaim: 223692 kB

KernelStack: 35568 kB

PageTables: 82500 kB

SecPageTables: 3248 kB

CommitLimit: 18575904 kB

Committed_AS: 21267524 kB

VmallocUsed: 84632 kB

Percpu: 7680 kB

AnonHugePages: 1363968 kB

ShmemHugePages: 0 kB

ShmemPmdMapped: 0 kB

FileHugePages: 1062912 kB

FilePmdMapped: 684032 kB

CmaTotal: 0 kB

CmaFree: 0 kB

Balloon: 0 kB

Hugetlb: 0 kB

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal: 32125696 kB	SwapFree: 2513056 kB	CommitLimit: 18575904 kB
MemFree: 8627220 kB	Zswap: 0 kB	Committed_AS: 21267524 kB
MemAvailable: 21107232 kB	Zswapped: 0 kB	AllocUsed: 84632 kB
	Dirty: 7680 kB	
	Writeback: 0 kB	Pages: 1363968 kB
	AnonPages: 6522544 kB	ShmemHugePages: 0 kB
SwapCached: 0 kB	Mapped: 1683592 kB	ShmemPmdMapped: 0 kB
Active: 5870672 kB	Shmem: 611016 kB	FileHugePages: 1062912 kB
Inactive: 16277940 kB	KReclaimable: 390056 kB	FilePmdMapped: 684032 kB
Active(anon): 8120 kB	Slab: 613748 kB	CmaTotal: 0 kB
Inactive(anon): 9589756 kB	SReclaimable: 390056 kB	CmaFree: 0 kB
Active(file): 5862552 kB	SUnreclaim: 223692 kB	Balloon: 0 kB
Inactive(file): 6688184 kB	KernelStack: 35568 kB	Hugetlb: 0 kB
Unevictable: 80 kB	PageTables: 82500 kB	
Mlocked: 80 kB	SecPageTables: 3248 kB	
SwapTotal: 2513056 kB		

Let's look at these first

And these too

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Sub)*

MemTotal:	32125696 kB	SwapFree:	904 kB
MemFree:	8627220 kB	Zswap:	524 kB
MemAvailable:	21107232 kB	Dirty:	632 kB
Buffers:	258820 kB	Writeback:	680 kB
Cached:	12902952 kB	AnonPage:	968 kB
SwapCached:	0 kB	Mapped:	0 kB
Active:	5870672 kB	Shmem:	912 kB
Inactive:	16277940 kB	KReclaimable:	4032 kB
Active(anon):	8120 kB	Slab:	0 kB
Inactive(anon):	9589756 kB	SReclaimable:	390056 kB
Active(file):	5862552 kB	SUnreclaim:	223692 kB
Inactive(file):	6688184 kB	KernelStack:	35568 kB
Unevictable:	80 kB	PageTables:	82500 kB
Mlocked:	80 kB	SecPageTables:	3248 kB
SwapTotal:	2513056 kB		

Memory managed by kernel's main page allocator ("buddy allocator") only
Excludes early boot allocations
Changes with memory hotplug/hotremove
Changes with *in-kernel* guest memory ballooning drivers (except vmware?)

Legend: **Kernel** **Userspace** *Redundant (Sub)*

MemTotal:	32125696 kB	SwapFree:	904 kB
MemFree:	8627220 kB	Zswap:	524 kB
MemAvailable:	21107232 kB	Dirty:	632 kB
Buffers:	258820 kB	Writeback:	680 kB
Cached:	12902952 kB	AnonPage:	968 kB
SwapCached:	0 kB	Mapped:	0 kB
Active:	5870672 kB	Shmem:	912 kB
Inactive:	16277940 kB	KReclaimable:	4032 kB
Active(anon):	8120 kB	Slab:	0 kB
Inactive(anon):	9589756 kB	SReclaimable:	390056 kB
Active(file):	5862552 kB	SUnreclaim:	223692 kB
Inactive(file):	6688184 kB	Balloon:	0 kB
Unevictable:	80 kB	Hugetlb:	0 kB
Mlocked:	80 kB		
SwapTotal:	2513056 kB		

Memory managed by kernel's main page allocator ("buddy allocator") only
Excludes early boot allocations
Changes with memory hotplug/hotremove
Changes with *in-kernel* guest memory ballooning drivers (except vmware?)

Memory given back to host in balloon pages

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	2588 kB	Dirty:	844 kB	Percpu:	7680 kB
Occupied by the compressed userspace data		Back:	196 kB	AnonHugePages:	1363968 kB
		Pages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	2588 kB	Dirty:	8 kB	Percpu:	7680 kB
Active:	5870672 kB	Shmem:	8 kB	AnonHugePages:	1363968 kB
Inactive:	16277940 kB	ShmemHugePages:	0 kB	ShmemPmdMapped:	0 kB
Active(anon):	8120 kB	FileHugePages:	1062912 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	CmaTotal:	0 kB	CmaFree:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	Balloon:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Hugetlb:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB		
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Occupied by the compressed
userspace data

Size of the original
uncompressed
pages

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	Free:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB		
Unevictable:	80 kB	KernelStack:	35568		
Mlocked:	80 kB	PageTables:	82500		
SwapTotal:	2513056 kB	SecPageTables:	3248		

Slab allocator pages
Details in /proc/slabinfo
Also large kmalloc()

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	Free:	0 kB
		SUnreclaim:	223692 kB		
		KernelStack:	35568		
		PageTables:	82500		
		SecPageTables:	3248		

Slab caches with
shrinkers
that can discard
objects when asked

Slab allocator pages
Details in /proc/slabinfo
Also large `kmalloc()`

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal: 32125696 kB	SwapFree: 2513056 kB	CommitLimit: 18575904 kB
MemFree: 8627220 kB	Zswap: 0 kB	Committed_AS: 21267524 kB
MemAvailable: 21107232 kB	Zswapped: 0 kB	VmallocUsed: 84632 kB
Buffers: 258820 kB	Dirty: 844 kB	Percpu: 7680 kB
Cached: 12902952 kB	Writeback: 196 kB	AnonHugePages: 1363968 kB
SwapCached: 0 kB	AnonPages: 8922944 kB	ShmemHugePages: 0 kB
Active: 5870672 kB	Mapped: 1683592 kB	ShmemPmdMapped: 0 kB
Inactive: 16277940 kB	Shmem: 611016 kB	FileHugePages: 1062912 kB
Active(anon): 8120 kB	KReclaimable: 390056 kB	FilePmdMapped: 684032 kB
Inactive(anon): 0 kB	Slab: 613748 kB	CmaTotal: 0 kB
	SReclaimable: 390056 kB	CmaFree: 0 kB
	SUnreclaim: 223692 kB	Balloon: 0 kB
	KernelStack: 35568 kB	Hugetlb: 0 kB
	PageTables: 82500 kB	
	SecPageTables: 3248 kB	

Used to be
SReclaimable plus other
users – ION allocator,
now gone

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
...	...	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
...	...	...	1683592 kB	ShmemPmdMapped:	0 kB
...	...	...	611016 kB	FileHugePages:	1062912 kB
...	...	...	390056 kB	FilePmdMapped:	684032 kB
...	...	...	613748 kB	CmaTotal:	0 kB
...	...	...	390056 kB	CmaFree:	0 kB
Inactive(file):	668024 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Kernel stacks of
kernel threads (kthreads)
But all userspace
threads also have them

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal: 32125696 kB	SwapFree: 2513056 kB	CommitLimit: 18575904 kB
MemFree: 8627220 kB	Zswap: 0 kB	Committed_AS: 21267524 kB
MemAvailable: 21107232 kB	Zswapped: 0 kB	VmallocUsed: 84632 kB
Buffers: 258820 kB	Dirty: 844 kB	Percpu: 7680 kB
Cached: 12902952 kB	Writeback: 196 kB	AnonHugePages: 1363968 kB
SwapCached: 0 kB	AnonPages: 8922944 kB	ShmemHugePages: 0 kB
Active: 5870672 kB	Mapped: 1683592 kB	ShmemPmdMapped: 0 kB
Inactive: 16277940 kB	Shmem: 611016 kB	FileHugePages: 1062912 kB
	Reclaimable: 390056 kB	FilePmdMapped: 684032 kB
	Tab: 613748 kB	CmaTotal: 0 kB
	Reclaimable: 390056 kB	CmaFree: 0 kB
Inactive(file): 668	SUnreclaim: 223692 kB	Balloon: 0 kB
Unevictable: 80	KernelStack: 35568 kB	Hugetlb: 0 kB
Mlocked: 80 kB	PageTables: 82500 kB	
SwapTotal: 2513056 kB	SecPageTables: 3248 kB	

Page tables for userspace processes

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB		
		Reclaimable:	390056 kB		
		Tab:	613748 kB		
		Reclaimable:	390056 kB		
Inactive(file):	660000 kB	SUnreclaim:	223692 kB		
Unevictable:	80000 kB	KernelStack:	35568 kB		
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Page tables for userspace processes

Secondary page tables (IOMMU, KVM)

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

Amount of memory backing the actually used vmalloc() areas

MemTotal:	56 kB	CommitLimit:	18575904 kB
MemFree:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Percpu:	7680 kB
Cached:	12902952 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	Balloon:	0 kB
Unevictable:	80 kB	Hugetlb:	0 kB
Mlocked:	80 kB		
SwapTotal:	2513056 kB		
Dirty:	844 kB		
Writeback:	196 kB		
AnonPages:	8922944 kB		
Mapped:	1683592 kB		
Shmem:	611016 kB		
KReclaimable:	390056 kB		
Slab:	613748 kB		
SReclaimable:	390056 kB		
SUnreclaim:	223692 kB		
KernelStack:	35568 kB		
PageTables:	82500 kB		
SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

Amount of memory backing the actually used vmalloc() areas

Amount of memory backing percpu variables

MemTotal:	56 kB	CommitLimit:	18575904 kB
MemFree:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Percpu:	7680 kB
Dirty:	844 kB	AnonHugePages:	1363968 kB
Cached:	196 kB	ShmemHugePages:	0 kB
SwapCached:	8922944 kB	ShmemPmdMapped:	0 kB
Active(anon):	16277940 kB	FileHugePages:	1062912 kB
Inactive(anon):	8120 kB	FilePmdMapped:	684032 kB
Active(file):	9589756 kB	CmaTotal:	0 kB
Inactive(file):	5862552 kB	CmaFree:	0 kB
Unevictable:	6688184 kB	Balloon:	0 kB
Mlocked:	80 kB	Hugetlb:	0 kB
SwapTotal:	80 kB		
SwapFree:	2513056 kB		
Shmem:	611016 kB		
KReclaimable:	390056 kB		
Slab:	613748 kB		
SReclaimable:	390056 kB		
SUnreclaim:	223692 kB		
KernelStack:	35568 kB		
PageTables:	82500 kB		
SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB		8922944 kB	ShmemHugePages:	0 kB
			1683592 kB	ShmemPmdMapped:	0 kB
			611016 kB	FileHugePages:	1062912 kB
			390056 kB	FilePmdMapped:	684032 kB
			613748 kB	CmaTotal:	0 kB
			390056 kB	CmaFree:	0 kB
			223692 kB	Balloon:	0 kB
			35568 kB	Hugetlb:	0 kB
			82500 kB		
			3248 kB		

Sum of sizes of *accountable* `mmap()` areas of all processes (private, writable, not hugetlb)

- where writes by the processes can populate pages that are not from the page cache (AS = address spaces?)



A limit for `Committed_AS`
 tunable by `sysctl`, defaults to half the
 RAM size plus all swap
 Only applies to `OVERCOMMIT_NEVER`
 (`vm.overcommit_memory = 2`)

SwapCached: 0 kB

Sum of sizes of *accountable*
`mmap()` areas of all processes
 (private, writable, not hugetlb)

- where writes by the processes
 can populate pages that are not
 from the page cache
 (AS = address spaces?)

other) Not consumed memory

0 kB

0 kB

844 kB

196 kB

8922944 kB

1683592 kB

611016 kB

390056 kB

613748 kB

390056 kB

223692 kB

35568 kB

82500 kB

3248 kB

CommitLimit: 18575904 kB

Committed_AS: 21267524 kB

VmallocUsed: 84632 kB

Percpu: 7680 kB

AnonHugePages: 1363968 kB

ShmemHugePages: 0 kB

ShmemPmdMapped: 0 kB

FileHugePages: 1062912 kB

FilePmdMapped: 684032 kB

CmaTotal: 0 kB

CmaFree: 0 kB

Balloon: 0 kB

Hugetlb: 0 kB

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Memory areas “reserved” for large contiguous allocations
 Recently also used for kdump kernel’s memory reservations
 When not needed for the intended purpose, can be used for userspace pages
 Migrated away when necessary (overwritten in case of kdump)

Not consumed memory	
kB	CommitLimit: 18575904 kB
kB	Committed_AS: 21267524 kB
kB	VmallocUsed: 84632 kB
kB	Percpu: 7680 kB
kB	AnonHugePages: 1363968 kB
kB	ShmemHugePages: 0 kB
kB	ShmemPmdMapped: 0 kB
kB	FileHugePages: 1062912 kB
kB	FilePmdMapped: 684032 kB
kB	CmaTotal: 0 kB
kB	CmaFree: 0 kB
kB	Balloon: 0 kB
kB	Hugetlb: 0 kB
Inactive(anon): 9589756 kB	Slab: 613748 kB
Active(file): 5862552 kB	SReclaimable: 390056 kB
Inactive(file): 6688184 kB	SUnreclaim: 223692 kB
Unevictable: 80 kB	KernelStack: 35568 kB
Mlocked: 80 kB	PageTables: 82500 kB
SwapTotal: 2513056 kB	SecPageTables: 3248 kB

Memory areas “reserved” for large contiguous allocations

Recently also used for kdump kernel’s memory reservations

When not needed for the intended purpose, can be used for userspace pages

Migrated away when necessary (overwritten in case of kdump)

How much of CmaTotal is currently not occupied
Either by contiguous allocation or userspace pages

Not consumed memory

kB	CommitLimit:	18575904 kB
kB	Committed_AS:	21267524 kB
kB	VmallocUsed:	84632 kB
kB	Percpu:	7680 kB
kB	AnonHugePages:	1363968 kB
kB	ShmemHugePages:	0 kB
kB	ShmemPmdMapped:	0 kB
kB	FileHugePages:	1062912 kB
kB	FilePmdMapped:	684032 kB
kB	CmaTotal:	0 kB
kB	CmaFree:	0 kB
kB	Balloon:	0 kB
kB	Hugetlb:	0 kB

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Legend: **Kernel** **Userspace** *Redundant (Subset of another)* **Not consumed memory**

MemTotal:	32125696 kB	SwapTotal:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB			Committed_AS:	21267524 kB
MemAvailable:	21107232 kB			VmallocUsed:	84632 kB
Buffers:	258820 kB	Dirty:	844 kB	Percpu:	7680 kB
Cached:	12902952 kB	Writeback:	196 kB	AnonHugePages:	1363968 kB
SwapCached:	0 kB	AnonPages:	8922944 kB	ShmemHugePages:	0 kB
Active:	5870672 kB	Mapped:	1683592 kB	ShmemPmdMapped:	0 kB
Inactive:	16277940 kB	Shmem:	611016 kB	FileHugePages:	1062912 kB
Active(anon):	8120 kB	KReclaimable:	390056 kB	FilePmdMapped:	684032 kB
Inactive(anon):	9589756 kB	Slab:	613748 kB	CmaTotal:	0 kB
Active(file):	5862552 kB	SReclaimable:	390056 kB	CmaFree:	0 kB
Inactive(file):	6688184 kB	SUnreclaim:	223692 kB	Balloon:	0 kB
Unevictable:	80 kB	KernelStack:	35568 kB	Hugetlb:	0 kB
Mlocked:	80 kB	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

Now let's look at these

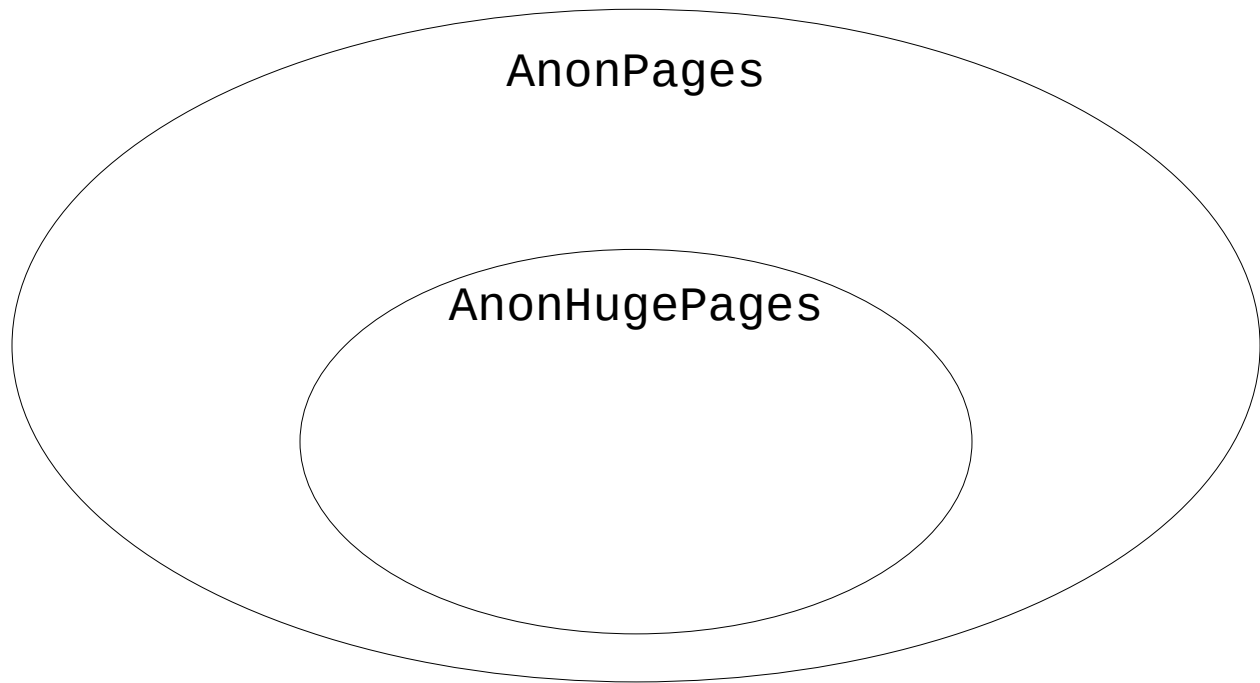
Anonymous memory vs page cache

- **Anonymous pages** – not backed by files
 - allocated mainly by faulting in `mmap(MAP_PRIVATE | MAP_ANONYMOUS)` areas
 - lifetime tied to the process(es) mapping them
 - can effectively become shared during `fork()`, due to copy-on-write (COW) semantics
 - cannot be simply written back to a file and discarded, only swapped out to swap
- **Page cache** – pages backed by files on a FS, or a block device
 - allocated by accessing files that are `mmap()`ed, or read/written to in other ways
 - lifetime independent of processes, can be discarded once not dirty, read back when needed again
- **Shared memory** (`shmem`, `tmpfs`) – a hybrid between anonymous and page cache
 - a RAM only filesystem (`/dev/shm`), also used for `shmget()` and `mmap(MAP_SHARED | MAP_ANONYMOUS)`
 - behaves a lot like page cache – lifetime independent of processes in most cases
 - but there are no persistent files backing it – must be swapped out like anonymous pages

Anonymous counters in detail

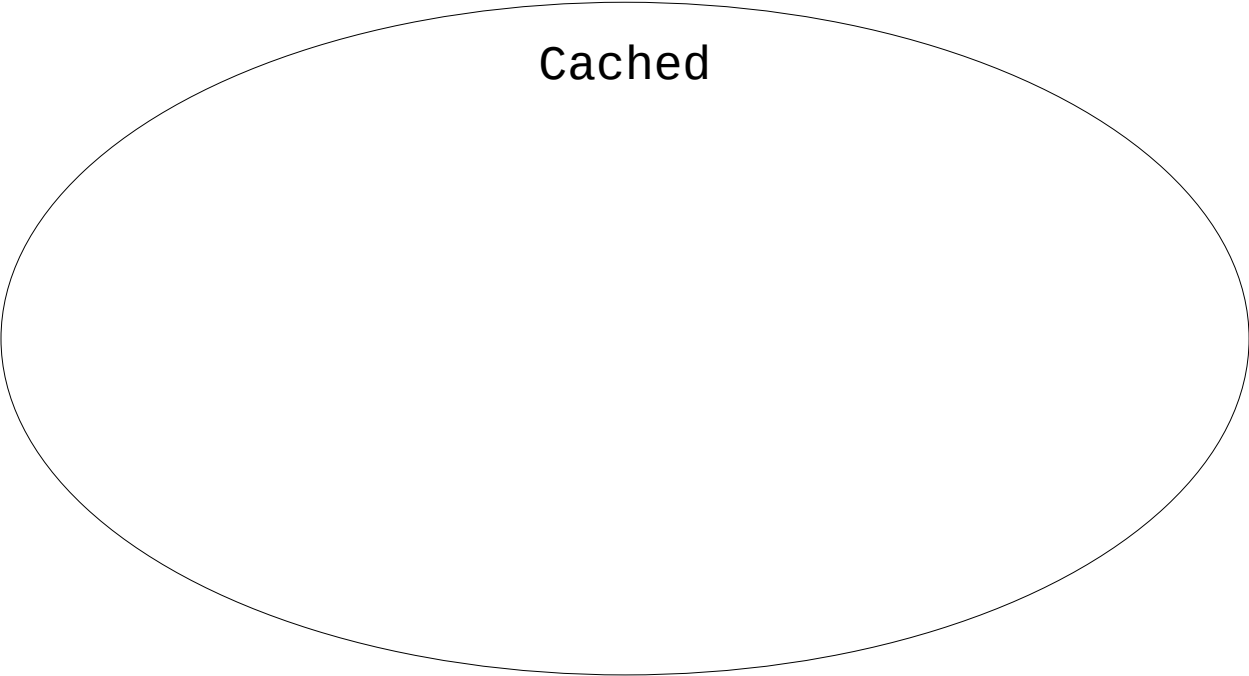
AnonPages

Anonymous counters in detail

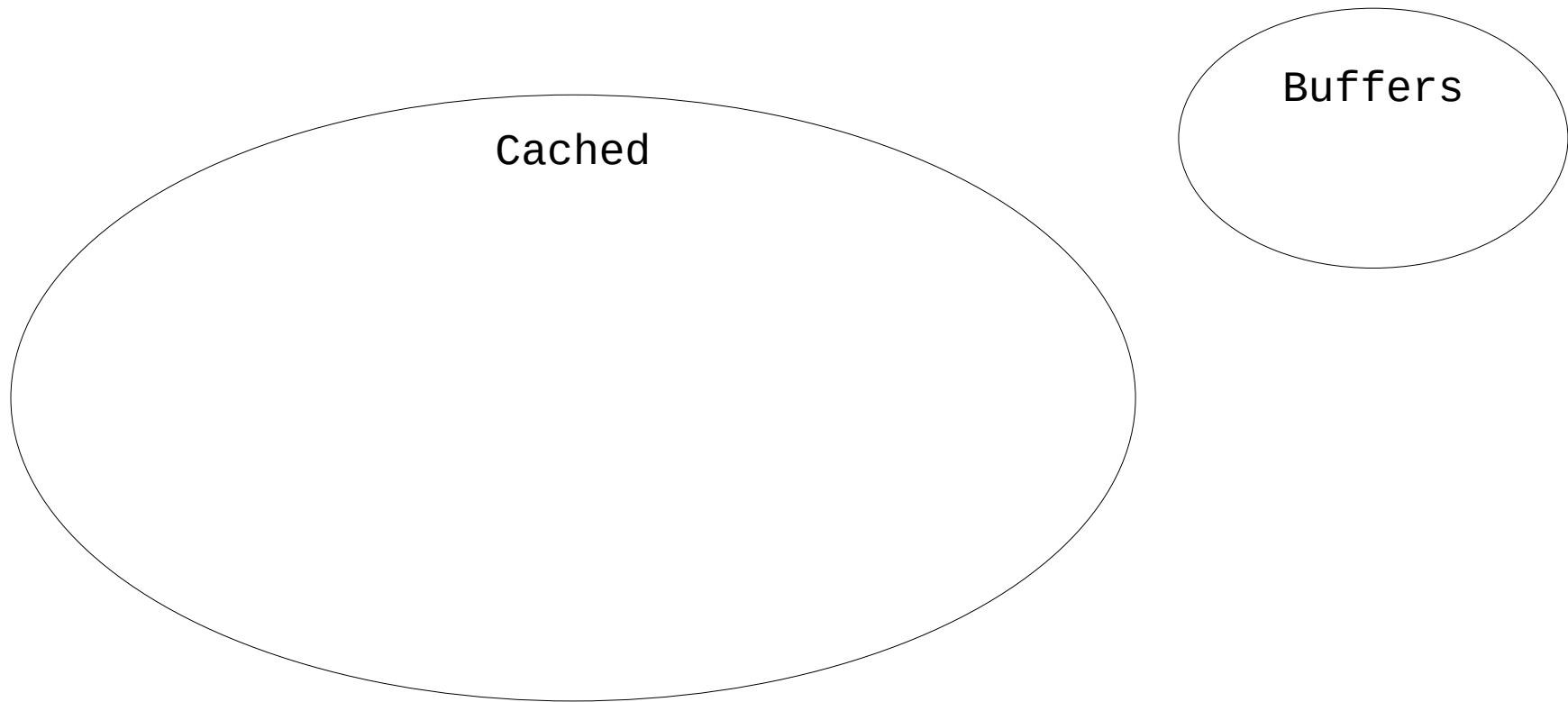


Page cache counters in detail

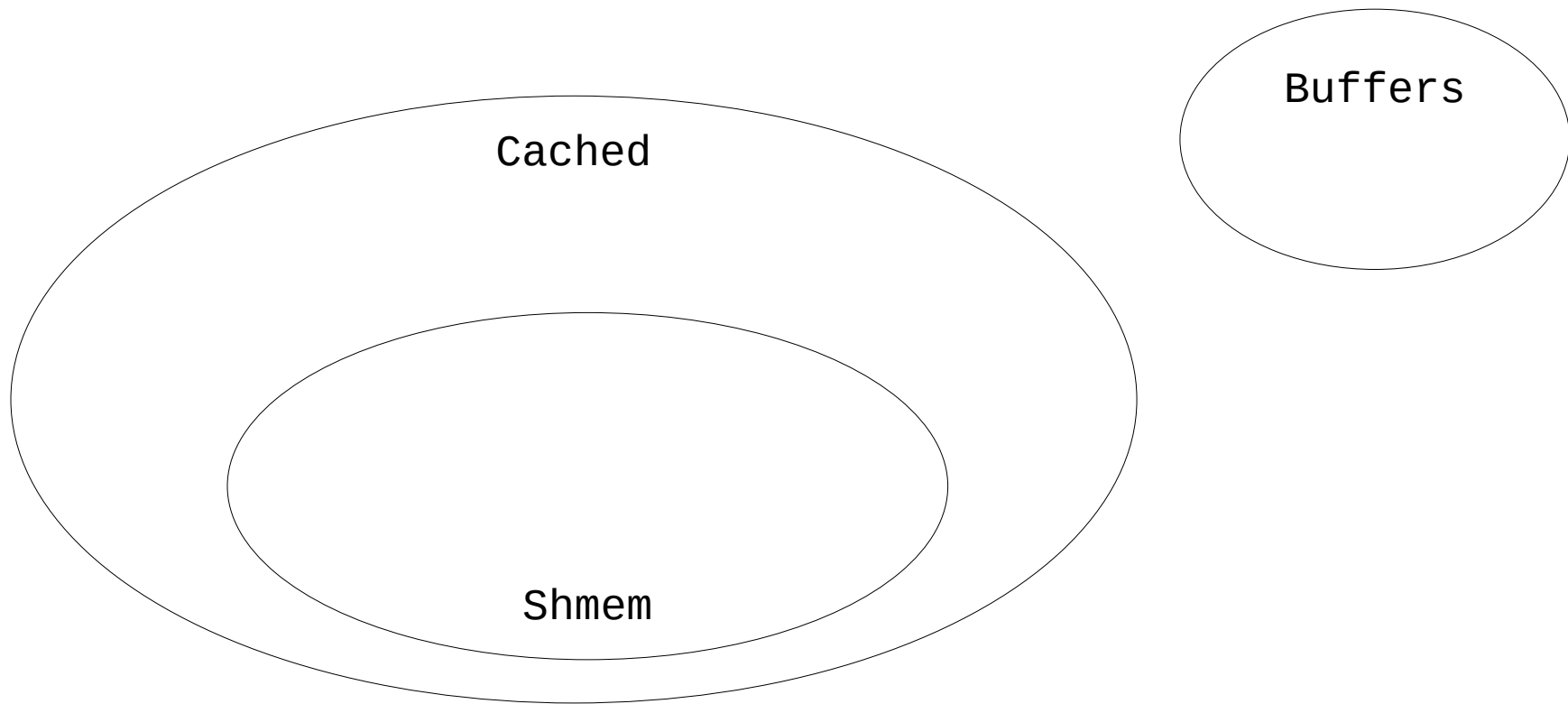
Cached



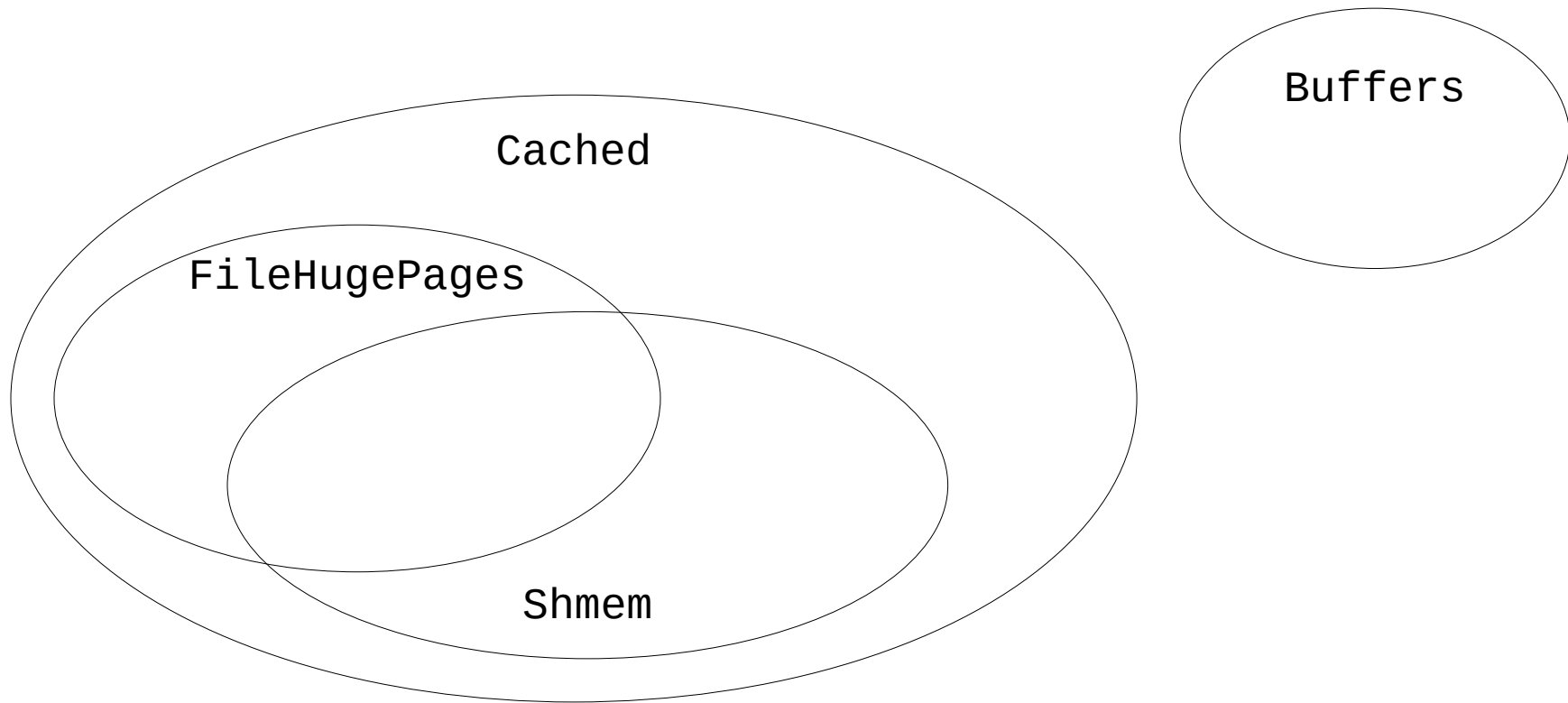
Page cache counters in detail



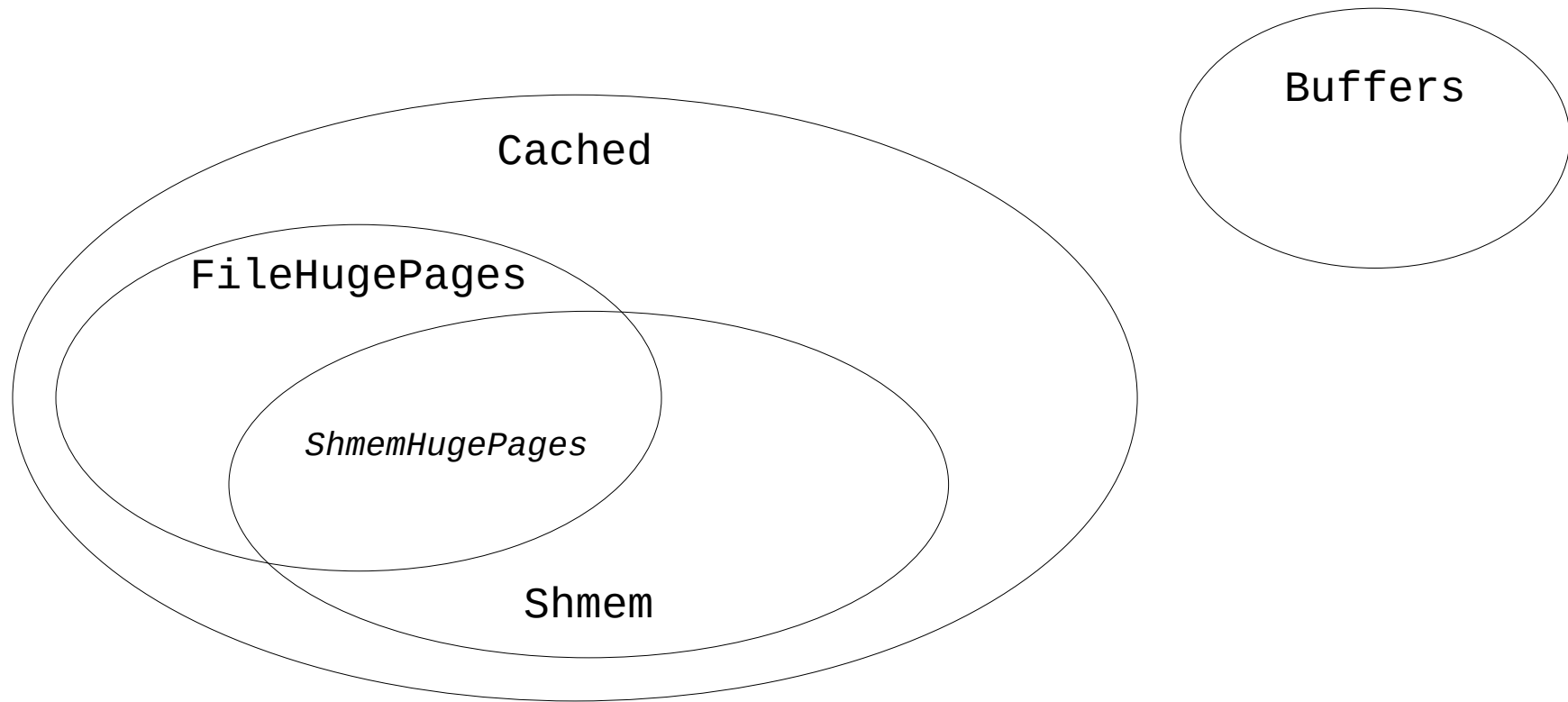
Page cache counters in detail



Page cache counters in detail

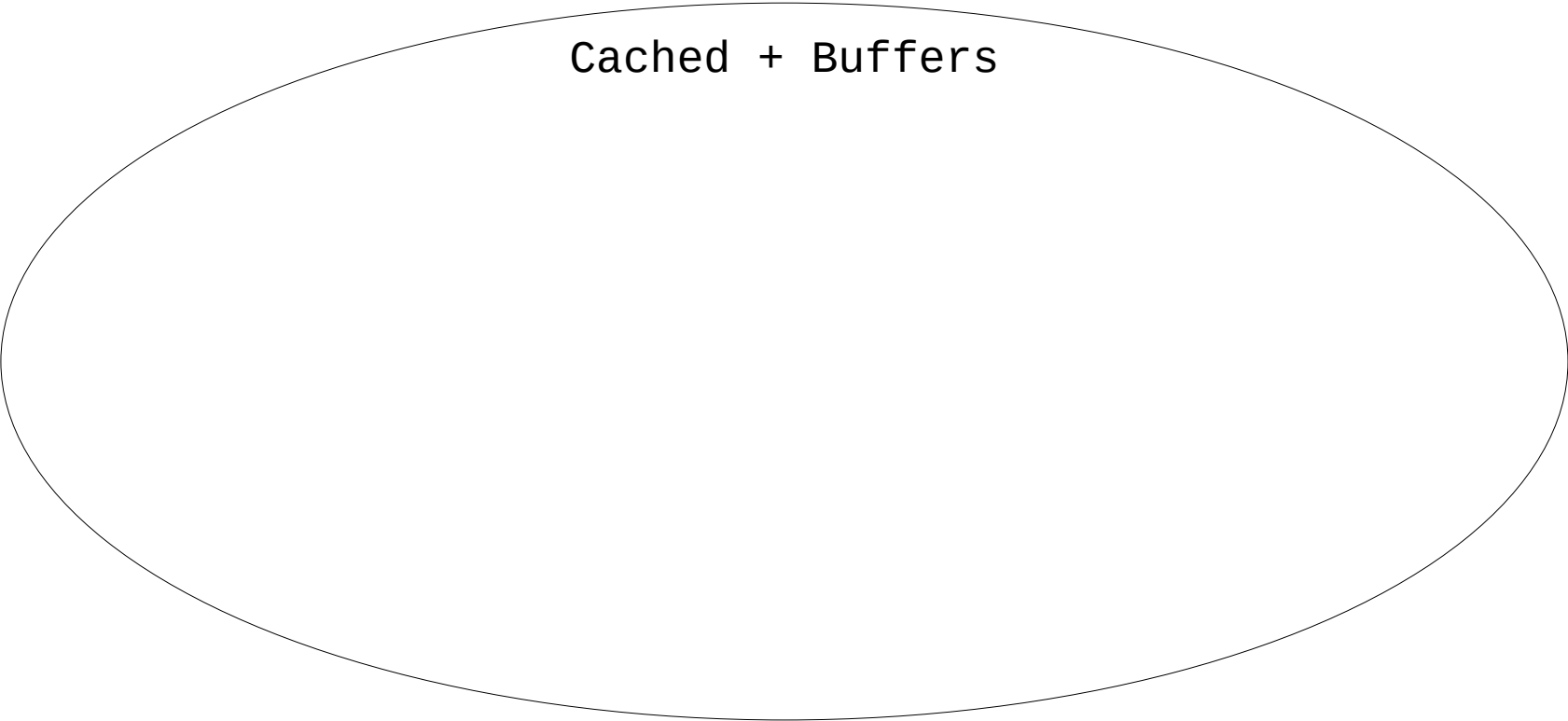


Page cache counters in detail

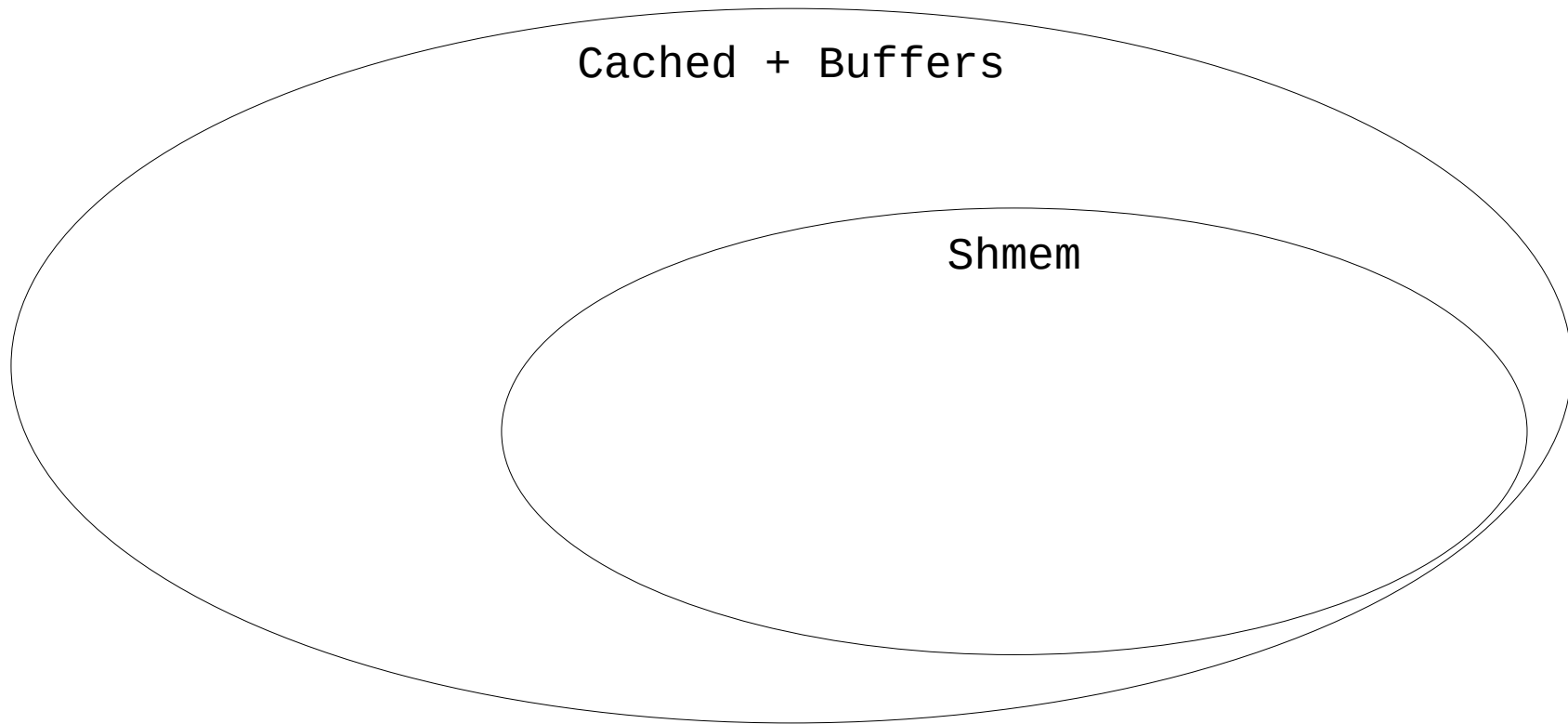


Page cache counters in detail – (Pmd)Mapped

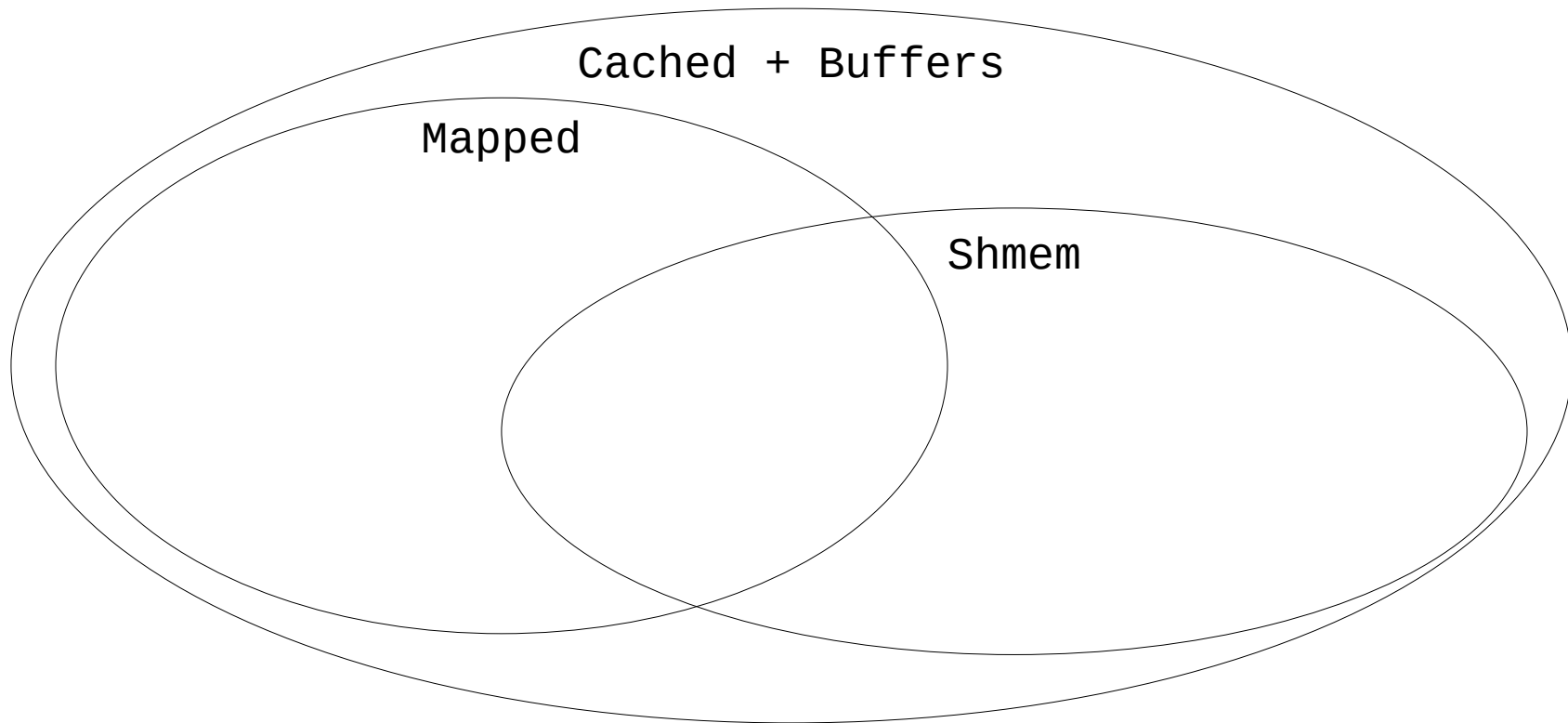
Cached + Buffers



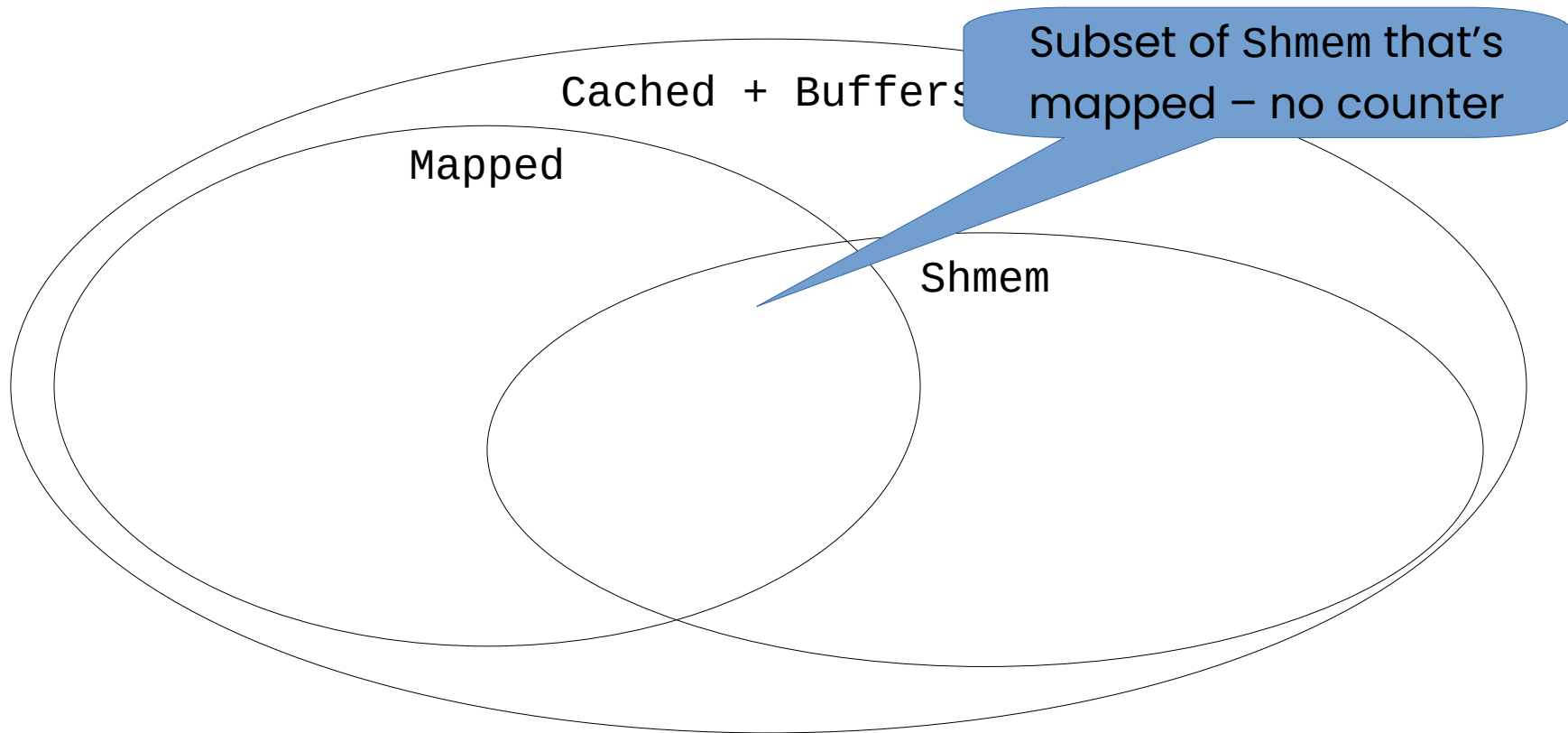
Page cache counters in detail – (Pmd)Mapped



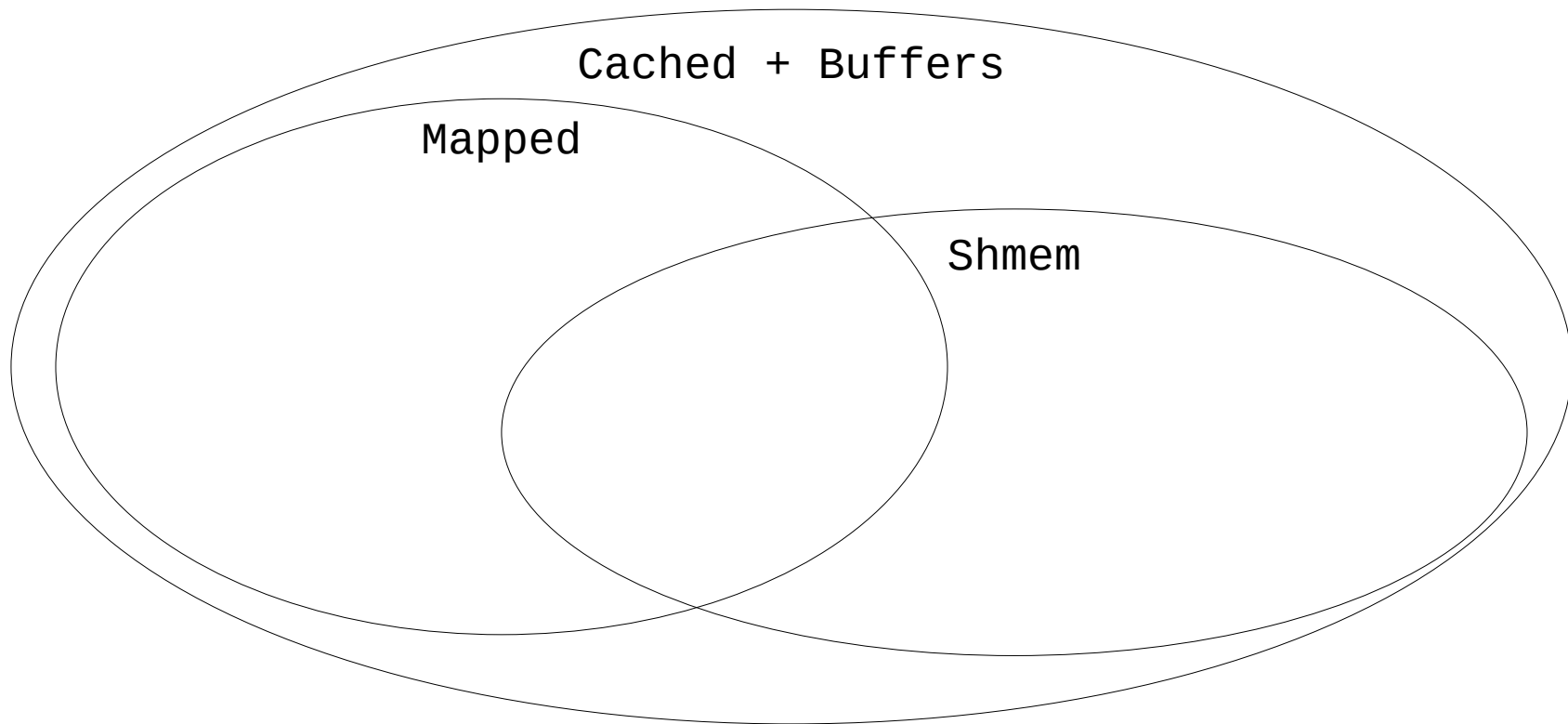
Page cache counters in detail – (Pmd)Mapped



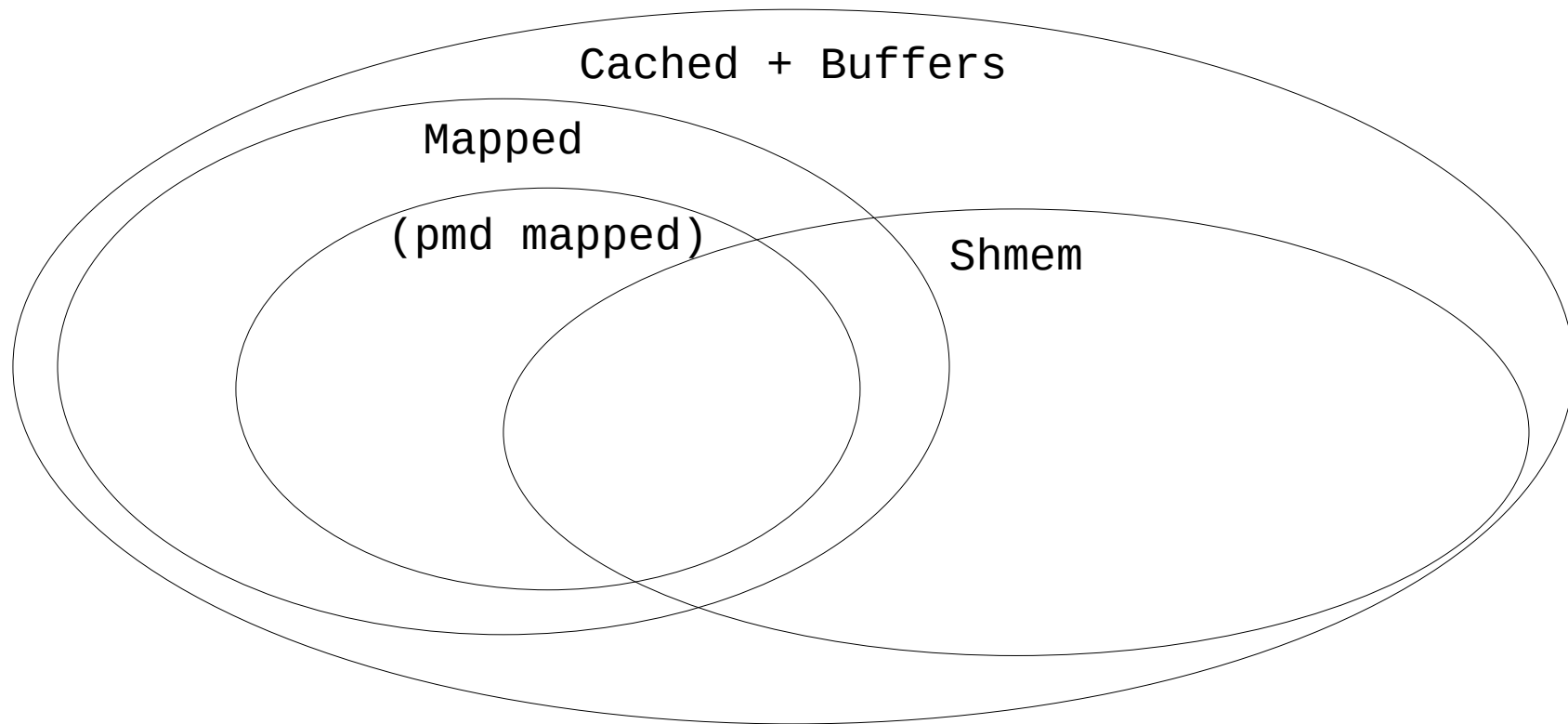
Page cache counters in detail – (Pmd)Mapped



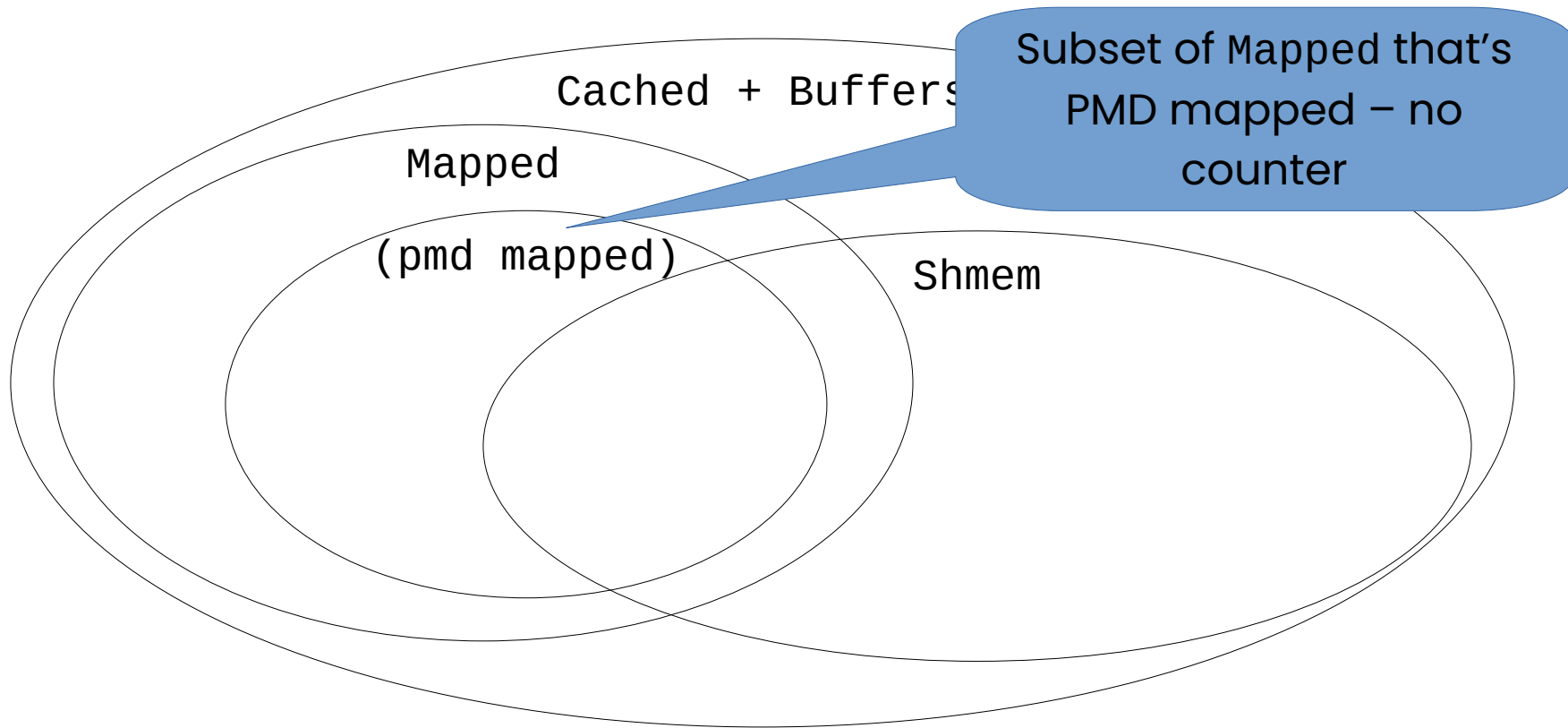
Page cache counters in detail – (Pmd)Mapped



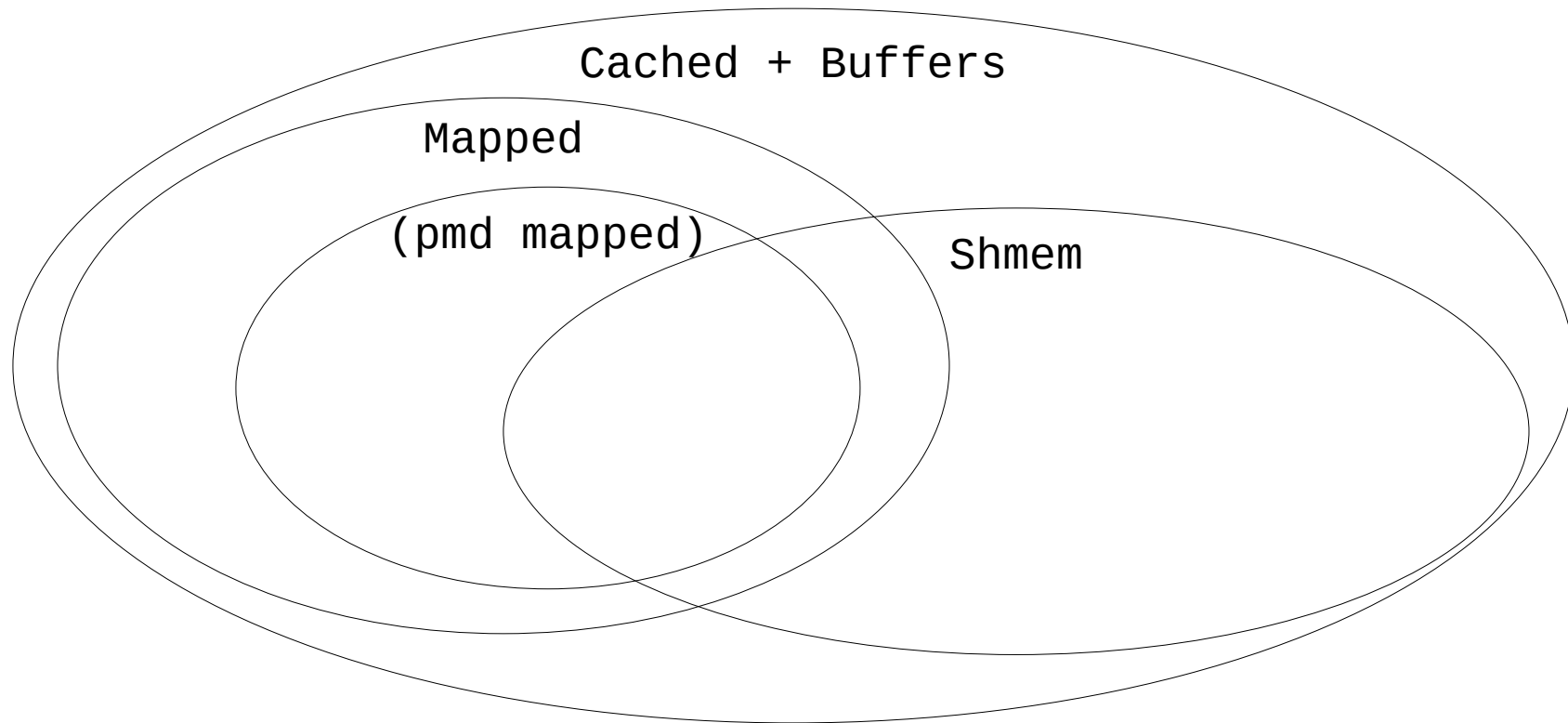
Page cache counters in detail – (Pmd)Mapped



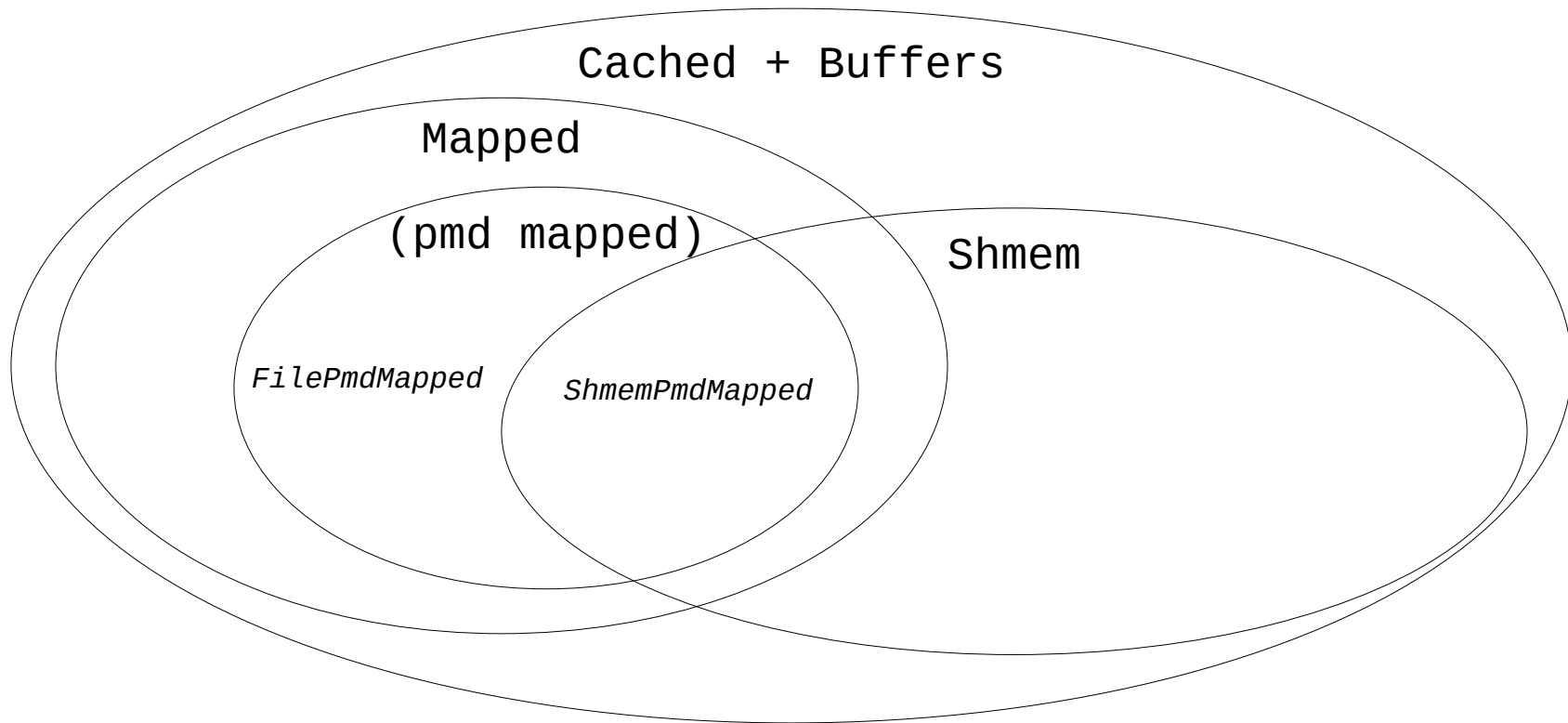
Page cache counters in detail – (Pmd)Mapped



Page cache counters in detail – (Pmd)Mapped



Page cache counters in detail – (Pmd)Mapped



Active and Inactive

- Userspace pages are sorted by their “hotness” on least recently used (LRU) lists
 - divided to anon and file due to their different characteristics
 - divided to active and inactive parts for reclaim optimization purposes, not workload reality
 - pages near the tail of the inactive list are the first to be evicted by kswapd or direct reclaim
- The counters are an orthogonal way to see how much memory is used by the page cache and anonymous pages
 - With a twist: *Shmem* is in page cache (*Cached*) but on *anon* LRU (because it’s swap backed)
 - Unevictable/Mlocked may account both page cache and anonymous pages
- $\text{AnonPages} + \text{Shmem} = 8922944 + 611016 = 9533960 \text{ kB}$
 $\text{Active(anon)} + \text{Inactive(anon)} = 9589756 + 8120 = 9597876 \text{ kB}$
- $\text{Cached} + \text{Buffers} - \text{Shmem} = 12902952 + 258820 - 611016 = 12550756 \text{ kB}$
 $\text{Active(file)} + \text{Inactive(file)} = 5862552 + 6688184 = 12550736 \text{ kB}$
- Significant differences (e.g. larger active+inactive sum) may be due to a memory leak

Legend: **Kernel** **Userspace** *Redundant* *Not consumed memory* **Sums to MemTotal**

MemTotal:	32125696 kB	SwapFree:	2513056 kB	CommitLimit:	18575904 kB
MemFree:	8627220 kB	Zswap:	0 kB	Committed_AS:	21267524 kB
MemAvailable:	21107232 kB	Zswapped:	0 kB	VmallocUsed:	84632 kB
Buffers:	258820 kB	<i>Dirty:</i>	<i>844 kB</i>	Percpu:	7680 kB
Cached:	12902952 kB	<i>Writeback:</i>	<i>196 kB</i>	<i>AnonHugePages:</i>	<i>1363968 kB</i>
<i>SwapCached:</i>	<i>0 kB</i>	AnonPages:	8922944 kB	<i>ShmemHugePages:</i>	<i>0 kB</i>
<i>Active:</i>	<i>5870672 kB</i>	<i>Mapped:</i>	<i>1683592 kB</i>	<i>ShmemPmdMapped:</i>	<i>0 kB</i>
<i>Inactive:</i>	<i>16277940 kB</i>	<i>Shmem:</i>	<i>611016 kB</i>	<i>FileHugePages:</i>	<i>1062912 kB</i>
<i>Active(anon):</i>	<i>8120 kB</i>	<i>KReclaimable:</i>	<i>390056 kB</i>	<i>FilePmdMapped:</i>	<i>684032 kB</i>
<i>Inactive(anon):</i>	<i>9589756 kB</i>	Slab:	613748 kB	CmaTotal:	0 kB
<i>Active(file):</i>	<i>5862552 kB</i>	<i>SReclaimable:</i>	<i>390056 kB</i>	CmaFree:	0 kB
<i>Inactive(file):</i>	<i>6688184 kB</i>	<i>SUnreclaim:</i>	<i>223692 kB</i>	Balloon:	0 kB
<i>Unevictable:</i>	<i>80 kB</i>	KernelStack:	35568 kB	Hugetlb:	0 kB
<i>Mlocked:</i>	<i>80 kB</i>	PageTables:	82500 kB		
SwapTotal:	2513056 kB	SecPageTables:	3248 kB		

So which fields sum to MemTotal?

- $\text{MemFree} + \text{Buffers} + \text{Cached} + \text{Zswap} + \text{AnonPages} + \text{Slab} + \text{KernelStack} + \text{PageTables} + \text{SecPageTables} + \text{VmallocUsed} + \text{Percpu} + \text{Hugetlb}$
= 31539312 kB
- $\text{MemTotal} = 32125696 \text{ kB}$
 - 572 MB not accounted for by any /proc/meminfo counter
 - Might be allocations by drivers, filesystems, networking... some expose the info elsewhere
 - Unaccounted field proposed as RFC in 2016 – [mixed feedback](#)
- So what exactly is MemAvailable then?
 - An estimate of how much is free or reclaimable (except anonymous)
 - $\text{MemFree} \text{ (minus some reserve)}$
+ $\text{Active(file)} + \text{Inactive(file)} \text{ (minus some reserve)}$
+ $\text{KReclaimable} \text{ (minus some reserve)}$
= 21568012 kB (reserves not subtracted)
 - $\text{MemAvailable: } 21107232 \text{ kB}$

Determining process memory usage – ps?

- ps, top, htop... all have columns for process memory usage
 - VSZ or VIRT (virtual memory size)
 - sizes of `mmap()` areas (VMA's) only
 - may count towards `Committed_AS`, but that does not say how much memory is actually occupied
 - RSS or RES (resident set size)
 - sum of the populated parts (actually backed by memory) of VSZ
 - better, but also not perfect
 - SHR, CODE, DATA... and other subsets of RSS
 - More specific use for some of the resident memory

How can RSS be misleading ?

- Demo: an [simple program](#) mapfork
 - Create a 1GB area with `mmap(MAP_ANONYMOUS | MAP_PRIVATE)`
 - Fault in all the pages in the area by writing a byte in a loop with 4kB step
 - `fork()` 10 children processes of the main process
 - Rewrite the area in each child
 - Observe system memory usage and RSS of the test processes in each step
 - `free -m ; ps -C mapfork u`

How can RSS be misleading ?

- Create a 1GB area with `mmap(MAP_ANONYMOUS | MAP_PRIVATE)`

	total	used	free	shared	buff/cache	available
Mem:	31372	12229	10060	647	10179	19143

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	0.0	1051220	1504

- VSZ is 1GB but RSS just 1.5MB

How can RSS be misleading ?

- Create a 1GB area with `mmap(MAP_ANONYMOUS | MAP_PRIVATE)`

	total	used	free	shared	buff/cache	available
Mem:	31372	12229	10060	647	10179	19143

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	0.0	1051220	1504

- VSZ is 1GB but RSS just 1.5MB
- Fault in all the pages in the area by writing a byte in a loop with 4kB step

How can RSS be misleading ?

- Create a 1GB area with `mmap(MAP_ANONYMOUS | MAP_PRIVATE)`

	total	used	free	shared	buff/cache	available
Mem:	31372	12229	10060	647	10179	19143

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	0.0	1051220	1504

- VSZ is 1GB but RSS just 1.5MB

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

	total	used	free	shared	buff/cache	available
Mem:	31372	13360	8926	647	10183	18012

PID	%CPU	%MEM	VSZ	RSS
78893	0.1	3.2	1051220	1050200

- RSS is now also 1GB, used memory increased by 1GB

How can RSS be misleading ?

- Create a 1GB area with `mmap(MAP_ANONYMOUS | MAP_PRIVATE)`

	total	used	free	shared	buff/cache	available
Mem:	31372	12229	10060	647	10179	19143

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	0.0	1051220	1504

Increased by ~1GB

- VSZ is 1GB but RSS just 1.5MB

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

	total	used	free	shared	buff/cache	available
Mem:	31372	13360	8926	647	10183	18012

PID	%CPU	%MEM	VSZ	RSS
78893	0.1	3.2	1051220	1050200

- RSS is now also 1GB, used memory increased by 1GB

How can RSS be misleading ?

- Create a 1GB area with `mmap(MAP_ANONYMOUS | MAP_PRIVATE)`

	total	used	free	shared	buff/cache	available
Mem:	31372	12229	10060	647	10179	19143

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	0.0	1051220	1504

Increased by ~1GB

- VSZ is 1GB but RSS just 1.5MB

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

	total	used	free	shared	buff/cache	available
Mem:	31372	13360	8926	647	10183	18012

PID	%CPU	%MEM	VSZ	RSS
78893	0.1	3.2	1051220	1050200

Increased by ~1GB

- RSS is now also 1GB, used memory increased by 1GB

How can RSS be misleading ?

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

	total	used	free	shared	buff/cache	available
Mem:	31372	13360	8926	647	10183	18012

PID	%CPU	%MEM	VSZ	RSS
78893	0.1	3.2	1051220	1050200

How can RSS be misleading ?

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

	total	used	free	shared	buff/cache	available
Mem:	31372	13360	8926	647	10183	18012

PID	%CPU	%MEM	VSZ	RSS
78893	0.1	3.2	1051220	1050200

- fork() 10 children processes of the main process

How can RSS be misleading ?

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

```
-      total      used      free      shared  buff/cache   available
Mem: 31372    13360      8926        647      10183      18012
```

```
      PID      %CPU      %MEM     VSZ     RSS
78893      0.1       3.2   1051220  1050200
```

- fork() 10 children processes of the main process

```
-      total      used      free      shared  buff/cache   available
Mem: 31372    13459      8823        653      10192      17912
```

```
      PID      %CPU      %MEM     VSZ     RSS
78893      0.0       3.2   1051220  1050328
79302      0.0       3.2   1051220  1049144
...
79311      0.0       3.2   1051220  1049108
```

- Sum of RSS is now 11GB but the used memory almost didn't change

How can RSS be misleading ?

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

```
-      total      used      free      shared  buff/cache   available
Mem: 31372    13360      8926        647      10183      18012
```

```
      PID      %CPU      %MEM     VSZ     RSS
78893        0.1       3.2   1051220  1050200
```

Almost unchanged

- fork() 10 children processes of the main process

```
-      total      used      free      shared  buff/cache   available
Mem: 31372    13459      8823        653      10192      17912
```

```
      PID      %CPU      %MEM     VSZ     RSS
78893        0.0       3.2   1051220  1050328
79302        0.0       3.2   1051220  1049144
...
79311        0.0       3.2   1051220  1049108
```

- Sum of RSS is now 11GB but the used memory almost didn't change

How can RSS be misleading ?

- Fault in all the pages in the area by writing a byte in a loop with 4kB step

-

	total	used	free	shared	buff/cache	available
Mem:	31372	13360	8926	647	10183	18012

PID	%CPU	%MEM	VSZ	RSS
78893	0.1	3.2	1051220	1050200

Almost unchanged

- fork() 10 children processes of the main process

-

	total	used	free	shared	buff/cache	available
Mem:	31372	13459	8823	653	10192	17912

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	3.2	1051220	1050328
79302	0.0	3.2	1051220	1049144
...				
79311	0.0	3.2	1051220	1049108

Sum of these is ~11GB

- Sum of RSS is now 11GB but the used memory almost didn't change

How can RSS be misleading ?

- fork() 10 children processes of the main process

-	total	used	free	shared	buff/cache	available
	Mem: 31372	13459	8823	653	10192	17912

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	3.2	1051220	1050328
79302	0.0	3.2	1051220	1049144
...				
79311	0.0	3.2	1051220	1049108

How can RSS be misleading ?

- fork() 10 children processes of the main process

-	total	used	free	shared	buff/cache	available
	Mem: 31372	13459	8823	653	10192	17912

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	3.2	1051220	1050328
79302	0.0	3.2	1051220	1049144
...				
79311	0.0	3.2	1051220	1049108

- Rewrite the area in each child

How can RSS be misleading ?

- fork() 10 children processes of the main process

-	total	used	free	shared	buff/cache	available
Mem:	31372	13459	8823	653	10192	17912
PID	%CPU	%MEM	VSZ	RSS		
78893	0.0	3.2	1051220	1050328		
79302	0.0	3.2	1051220	1049144		
...						
79311	0.0	3.2	1051220	1049108		

- Rewrite the area in each child

-	total	used	free	shared	buff/cache	available
Mem:	31372	23741	226	647	8501	7630
PID	%CPU	%MEM	VSZ	RSS		
78893	0.1	3.2	1051220	1050328		
79302	0.2	3.2	1051220	1049144		
...						
79311	0.0	3.2	1051220	1049532		

- Used memory increased by 10GB, RSS unchanged

How can RSS be misleading ?

- fork() 10 children processes of the main process

-	Mem:	total 31372	used 13459	free 8823	shared 653	buff/cache 10192	available 17912
	PID	%CPU	%MEM	VSZ	RSS		
	78893	0.0	3.2	1051220	1050328		
	79302	0.0	3.2	1051220	1049144		
	...						
	79311	0.0	3.2	1051220	1049108		

- Rewrite the area in each child

-	Mem:	total 31372	used 23741	free 226	shared 647	buff/cache 8501	available 7630
	PID	%CPU	%MEM	VSZ	RSS		
	78893	0.1	3.2	1051220	1050328		
	79302	0.2	3.2	1051220	1049144		
	...						
	79311	0.0	3.2	1051220	1049532		

Increased by ~10GB

- Used memory increased by 10GB, RSS unchanged

How can RSS be misleading ?

- fork() 10 children processes of the main process

- Mem: total 31372 used 13459 free 8823 shared 653 buff/cache 10192 available 17912

PID	%CPU	%MEM	VSZ	RSS
78893	0.0	3.2	1051220	1050328
79302	0.0	3.2	1051220	1049144
...				
79311	0.0	3.2	1051220	1049108

- Rewrite the area in each child

- Mem: total 31372 used 23741 free 226 shared 647 buff/cache 8501 available 7630

PID	%CPU	%MEM	VSZ	RSS
78893	0.1	3.2	1051220	1050328
79302	0.2	3.2	1051220	1049144
...				
79311	0.0	3.2	1051220	1049532

Increased by ~10GB

No change here

- Used memory increased by 10GB, RSS unchanged

What was happening in the kernel ?

- Kernel has per-process (mm) RSS counters of several (mutually exclusive) types
 - resident file pages, anonymous pages, shmem pages, and also swap entries
 - increased and decreased as respective pages (swap entries) get mapped or unmapped in the page tables
 - exported by e.g. `/proc/$PID/status` along with others
- `mmap( )` without `MAP_POPULATE` creates a VMA but doesn't populate the page tables – no RSS increase
- Writing with 4kB step populates page tables with anonymous pages – RSS increased
- `fork( )` copies both the VMA's and page tables of the parent to the child (with exceptions)
 - the RSS counters will have almost the same values
 - but the pages with data are not duplicated, instead they become transparently shared
- Once child process writes, new pages are allocated by the copy on write (CoW) mechanism
 - No change to the RSS, but memory usage increased

Why is there no accurate counter ?

- Could we maintain a per-process counter where memory usage of shared pages would be weighted by how many processes share them?
 - Defined as **PSS (proportional set size)**
 - Sum over all processes would match the total used memory
 - It would be expensive – i.e. a CoW in one process would have to increment counters in other processes
 - RSS is far from perfect, but simple and efficient to maintain
- BTW, how does OOM killer determine which process to kill ?
 - Also RSS (all types of entries, including swap), plus memory occupied by task's page tables
 - With almost all pages shared, killing one of our test processes would actually free nothing
 - We would have to take kill all of them to free any memory
 - Better metric – USS (unique/unshared set size) – count pages not shared with anyone else
 - Not maintained as that would be similarly expensive to maintaining PSS

Pay the price for accuracy on-demand?

- PSS and USS can be provided by the kernel, but it's a snapshot calculated on the fly
 - `/proc/$PID/smmaps(_rollup)`
 - Has to walk the page tables, and inspect individual mapped pages' mapcount and other attributes
 - Holds `mmap_lock` for read, which can block the inspected process taking `mmap_lock` for write, and result in priority inversion
 - Low priority monitoring task might be blocking high priority workload
 - There is an ongoing work to fix this

Pay the price for accuracy on-demand?

- Example (some fields omitted) for our demo program after the fork():

```
- > cat /proc/69788/smmaps_rollup  
00400000-7ffdb60d9000 ---p 00000000 00:00 0
```

[rollup]

```
Rss: 1050416 kB  
Pss: 95386 kB  
Pss_Dirty: 95374 kB  
Pss_Anon: 95365 kB  
Pss_File: 12 kB  
Pss_Shmem: 8 kB  
Shared_Clean: 1700 kB  
Shared_Dirty: 1048676 kB  
Private_Clean: 0 kB  
Private_Dirty: 40 kB  
Referenced: 1050416 kB  
Anonymous: 1048704 kB  
AnonHugePages: 1048576 kB  
ShmemPmdMapped: 0 kB  
FilePmdMapped: 0 kB  
Swap: 0 kB  
SwapPss: 0 kB  
Locked: 0 kB
```

Pay the price for accuracy on-demand?

- Example (some fields omitted) for our demo program after the children writes

```
- > cat /proc/69788/smmaps_rollup  
00400000-7ffdb60d9000 ---p 00000000 00:00 0
```

[rollup]

```
Rss: 1050204 kB  
Pss: 1048629 kB  
Pss_Dirty: 1048617 kB  
Pss_Anon: 1048616 kB  
Pss_File: 11 kB  
Pss_Shmem: 1 kB  
Shared_Clean: 1488 kB  
Shared_Dirty: 108 kB  
Private_Clean: 0 kB  
Private_Dirty: 1048608 kB  
Referenced: 1050204 kB  
Anonymous: 1048704 kB  
AnonHugePages: 16384 kB  
ShmemPmdMapped: 0 kB  
FilePmdMapped: 0 kB  
Swap: 0 kB  
SwapPss: 0 kB  
Locked: 0 kB
```

Pay the price for accuracy on-demand?

- Example (some fields omitted) for our demo program after the children writes

```
- > cat /proc/69788/smmaps_rollup  
00400000-7ffdb60d9000 ---p 00000000 00:00 0
```

[rollup]

```
Rss: 1050204 kB  
Pss: 1048629 kB  
Pss_Dirty: 1048617 kB  
Pss_Anon: 1048616 kB  
Pss_File: 11 kB  
Pss_Shmem: 1 kB  
Shared_Clean: 1488 kB  
Shared_Dirty: 108 kB  
Private_Clean: 0 kB  
Private_Dirty: 1048608 kB  
Referenced: 1050204 kB  
Anonymous: 1048704 kB  
AnonHugePages: 16384 kB  
ShmemPmdMapped: 0 kB  
FilePmdMapped: 0 kB  
Swap: 0 kB  
SwapPss: 0 kB  
Locked: 0 kB
```

(dirty) pages that only this
process is mapping
~ USS

More user friendly usage – the smem tool

- The demo program after populating the area:

	PID	Command	Swap	USS	PSS	RSS
-	72922	./mapfork	0	1048728	1048756	1050288

- After fork()s

-	PID	Command	Swap	USS	PSS	RSS
	73362	./mapfork	0	28	95369	1049404
...	72922	./mapfork	0	36	95388	1050416

- After children writes

-	PID	Command	Swap	USS	PSS	RSS
	73362	./mapfork	0	1048604	1048618	1049412
...	73371	./mapfork	0	1048620	1048636	1049732

How to debug *missing* kernel memory?

- Unaccounted amount too large or growing over time?
 - Something is allocating memory directly and not freeing it back. But who?
- CONFIG_PAGE_OWNER=y and boot with page_owner=on
 - Some CPU and memory overhead, but none unless enabled on boot
 - /sys/kernel/debug/page_owner – per-page allocation stacks and other info, requires postprocessing to aggregate
 - /sys/kernel/debug/page_owner_stacks/show_stacks – new aggregate output
 - Not sorted, but echo X > .../count_threshold can limit the output
 - freeing stacks also collected, can be dumped in case some page corruption is detected (i.e. debug_pagealloc)

How to debug *missing* kernel memory?

- Example `/sys/kernel/debug/page_owner_stacks/show_stacks`

- `get_page_from_freelist+0x1004/0x1210`
`__alloc_pages_noprof+0x18f/0x10b0`
`new_slab+0xc7/0x3c0`
`__slab_alloc+0x75f/0xbe0`
`kmem_cache_alloc_lru_noprof+0x1e8/0x210`
`fuse_alloc_inode+0x26/0x110`
`alloc_inode+0x1b/0x60`
`iget5_locked+0x97/0xe0`
`fuse_iget+0x63/0x230`
`fuse_lookup_name+0x16d/0x250`
`fuse_lookup+0x7f/0x1a0`
`__lookup_slow+0x89/0x130`
`lookup_one_unlocked+0x93/0xb0`
`ovl_lookup_single+0x7b/0x370`
`ovl_lookup_layer+0x12d/0x150`
`ovl_lookup+0x72e/0xd80`
`nr_base_pages: 18880`

How to debug *missing* kernel memory?

- What if we know it's a particular slab cache? (from /proc/slabinfo)
 - For example, `kmalloc-*` caches have many different users
 - Boot with a `slab_debug` option, e.g. `slab_debug=U, kmalloc-*`
 - `cat /sys/kernel/debug/slab/kmalloc-64/alloc_traces`
 - No minimum threshold that could be set, but it is sorted from the largest user
 - ```
182 scsi_dev_info_list_add_keyed+0x99/0x1a0 waste=1456/8
age=98885/98885/98885 pid=1 cpus=7
 scsi_dev_info_list_add_keyed+0x99/0x1a0
 scsi_init_devinfo+0x5b/0xa0
 init_scsi+0x1a/0x90
 do_one_initcall+0x5b/0x230
 kernel_init_freeable+0x1d3/0x220
 kernel_init+0x1a/0x130
 ret_from_fork+0x34/0x50
 ret_from_fork_asm+0x1a/0x30
```

# How to debug *missing* kernel memory?

- A new possible alternative to both of page\_owner and slab\_debug mechanisms:
- CONFIG\_MEM\_ALLOC\_PROFILING, sysctl.vmemprofiling=0|1|never
  - Can be enabled/disabled after boot, but the memory overhead paid on boot
  - Not capturing full stacks, but callsites (file, line, function)
    - If some helper is used by different users, it needs to be wrapped in a macro to distinguish them
- # sort -g /proc/allocinfo | tail -n 11 | numfmt --to=iec
  - |      |       |                                   |                              |
|------|-------|-----------------------------------|------------------------------|
| 1.9M | 9923  | fs/dcache.c:1636                  | func:__d_alloc               |
| 2.1M | 513   | mm/execmem.c:31                   | func:__execmem_alloc         |
| 2.1M | 655   | kernel/fork.c:313                 | func:alloc_thread_stack_node |
| 2.1M | 528   | drivers/virtio/virtio_ring.c:319  | func:vring_alloc_queue       |
| 2.5M | 616   | mm/mm_init.c:2428                 | func:alloc_large_system_hash |
| 3.7M | 925   | mm/shmem.c:1722                   | func:shmem_alloc_folio       |
| 4.5M | 1134  | mm/memory.c:1062                  | func:folio_prealloc          |
| 8.0M | 2048  | drivers/char/virtio_console.c:442 | func:alloc_buf               |
| 9.4M | 2384  | mm/memory.c:1060                  | func:folio_prealloc          |
| 28M  | 2557  | mm/slub.c:2325                    | func:alloc_slab_page         |
| 45M  | 11299 | mm/readahead.c:248                | func:page_cache_ra_unbounded |

# How to debug *missing* kernel memory?

- What if I don't want to reboot or even rebuild the kernel?
- Tracing! Capture memory allocation and freeing events.
  - Store stacktrace of the allocation events and allocated page address
  - Drop the stored event when freeing the same address
    - Postprocessing of the trace, or in the kernel itself by eBPF
  - At any point, report outstanding allocations (not yet freed)
    - If the memory leak is still ongoing, it will eventually dominate the report
- Example: [memleak.py](#) in bcc-tools

# Conclusion: what to make of memory usage

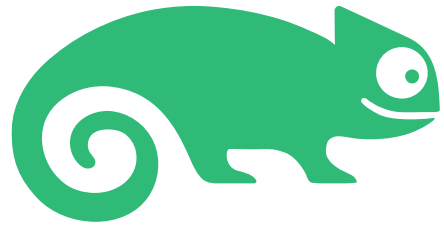
- “MemFree” isn’t a useful value, will converge to near-zero eventually on most systems
- “MemAvailable” and “used” make much more sense to report the system state at any given moment
- During the workload’s steady operation, the exact values are rarely important
- Values may change on distro upgrades due to kernel/glibc implementation details
  - kernel/user memory usage – changes e.g. due to scalability, sizing of structures
  - active/inactive LRU sizes – reclaim algorithm may change, doesn’t say much about the workload
- Limits for memcgs / alerts determined on an older version may need updating
  - memcg kmem consumption – new kinds of kernel objects may get accounted

# Conclusion: what to make of memory usage

- What matters more than exact values – trends over time or dynamic changes
  - a particular counter keeps growing over time – possible memory leak?
  - “available” memory exhausted, only a little page cache remaining, thrashing or OOM kills?
  - is this due to kernel bug, userspace bug or my workload is simply too big?
- Looking at `/proc/vmstat` event counters can tell much more information and often explain the changes in memory counters
  - See what actions the kernel is doing to keep up with the memory demand
  - But what exactly are the events that we can observe? Topic for the next year ;)

# Final note: non-zero swap usage is not a bug

- Consider a system that has enough memory to fit its workload and some page cache
- An overnight activity (backup, virus scanning) results in accessing many more files
- No swap usage with older kernels, all swap is free
- Suddenly, some swap becomes used after a kernel/distro upgrade – why?
  - Same workload, same amount of memory?
- The answer: kernel got better at tracking unused file vs anon memory around v5.9
  - Workingset change tracking expanded to anon memory
  - Previously reclaim very biased towards page cache, reluctant to swap out anon
  - Now able to detect anonymous page not touched *at all* by their processes
    - Vs file pages touched at least once a day due to the overnight activity
    - Swapping out is the correct thing to do – more memory for actually used page cache
- TID: <https://www.suse.com/support/kb/doc/?id=000021196>



**SUSE**

**Thank you**

Questions?