第 0019 讲 6 内存管理 meminfo&zoneinfo 信息分析

内存管理 meminfo&zoneinfo 信息分析

一、详细分析/proc/meminfo 节点信息

```
vico@ubuntu: ~
File  Edit  View  Search  Terminal  Help
vico@ubuntu:~$ cat /proc/meminfo
MemTotal:       16362520 kB
MemFree:        14309592 kB
MemAvailable:   14846336 kB
Buffers:          111796 kB
Cached:           617944 kB
SwapCached:            0 kB
Active:          1152444 kB
Inactive:         389552 kB
Active(anon):     805940 kB
Inactive(anon):     6416 kB
Active(file):     346504 kB
Inactive(file):   383136 kB
Unevictable:          16 kB
Mlocked:              16 kB
SwapTotal:       2097148 kB
SwapFree:        2097148 kB
Dirty:               196 kB
Writeback:             0 kB
AnonPages:        812284 kB
Mapped:           236452 kB
Shmem:              7772 kB
KReclaimable:     121192 kB
Slab:             196964 kB
```

具体系统调用实现，在 Linux 内核源码函数设计如下：

fs > proc > C meminfo.c > meminfo_proc_show(seq_file *, void *)

```
34  static int meminfo_proc_show(struct seq_file *m, void *v)
35  {
36      struct sysinfo i;
37      unsigned long committed;
38      long cached;
39      long available;
40      unsigned long pages[NR_LRU_LISTS];
41      unsigned long sreclaimable, sunreclaim;
42      int lru;
43
44      si_meminfo(&i);
45      si_swapinfo(&i);
46      committed = percpu_counter_read_positive(&vm_committed_as);
47
48      cached = global_node_page_state(NR_FILE_PAGES) -
49              total_swapcache_pages() - i.bufferram;
50      if (cached < 0)
51          cached = 0;
52
53      for (lru = LRU_BASE; lru < NR_LRU_LISTS; lru++)
54          pages[lru] = global_node_page_state(NR_LRU_BASE + lru);
55
56      available = si_mem_available();
57      sreclaimable = global_node_page_state(NR_SLAB_RECLAIMABLE);
58      sunreclaim = global_node_page_state(NR_SLAB_UNRECLAIMABLE);
59
60      show_val_kb(m, "MemTotal:       ", i.totalram);
61      show_val_kb(m, "MemFree:        ", i.freeram);
```

root@ubuntu:/home/vico# cat /proc/meminfo

MemTotal: 16362520 kB // 表示当前系统可用物理内存总数

MemFree: 14455324 kB // 表示当前系统剩余空闲物理内存

// 表示系统中可使用页面的数量（空闲页面、文件映射页面、可回收的页面）

公式：

可使用页面数量=空闲页面+文件映射页面+可回收页面-系统保留的页面

MemAvailable: 14873756 kB

Buffers: 54688 kB // 用于块层的缓存

Cached: 585220 kB // 用于页面高速缓存的页面

SwapCached: 0 kB // 统计交换缓存的数量

// 活跃的匿名页面（LRU_ACTIVE_ANON）和活跃的文件映射页面（LRU_ACTIVE_FILE）

Active: 1091584 kB

// 不活跃的匿名页面（LRU_INACTIVE_ANON）和不活跃的文件映射页面（LRU_INACTIVE_FILE）

Inactive: 372708 kB

Active(anon): 819572 kB // 活跃的匿名页面

Inactive(anon): 6376 kB // 不活跃的匿名碳

Active(file): 272012 kB // 活跃的文件映射页面

Inactive(file): 366332 kB // 不活跃的文件映射碳

Unevictable: 16 kB // 不能回收的碳

Mlocked: 16 kB // 不会被交换到交换分区的碳

SwapTotal: 2097148 kB // 交换分析的大小

SwapFree: 2097148 kB // 交换分区的空闲空间大小

Dirty: 776 kB // 脏页的数量

Writeback: 0 kB // 正在回写的页面数量

AnonPages: 824428 kB // 统计有反向映射的页面

Mapped: 238420 kB // 统计所有映射到用户地址空间的内存缓存页面

Shmem: 7724 kB // 共享内存页面的数量

KReclaimable: 67156 kB // Linux 内核可回收的内存，包含回收的 slab 页面

Slab: 143680 kB // 所有 slab 页面

SReclaimable: 67156 kB // 可回收的 slab 页面

SUnreclaim: 76524 kB // 不可回收的 slab 页面

KernelStack: 12480 kB // 所有进程内核栈大小

PageTables: 42316 kB // 所有用于页表的页面数量

NFS_Unstable: 0 kB // 在 NFS 中，发送到服务器端但是还没有写入磁盘的页面

Bounce: 0 kB // 退回

WritebackTmp: 0 kB // 回写过程中的使用临时缓存

CommitLimit: 10278408 kB // 虚拟内存限值

Committed_AS: 4527116 kB // 表示系统已经分配的内存情况

VmallocTotal: 34359738367 kB // vmalloc 区域的总的大小

VmallocUsed: 27596 kB // 已使用的 vmalloc 区域总的大小

VmallocChunk: 0 kB // 最大的连续未被使用 vmalloc 区域

Percpu: 46080 kB // percpu 机制使用的页面，由 pcpu_nr_pages()函数来统计

HardwareCorrupted: 0 kB // 当系统检测到内存的硬件故障时，会把有问题的页面删除

AnonHugePages: 0 kB // 统计透明巨型页的数量

ShmemHugePages: 0 kB

ShmemPmdMapped: 0 kB // 使用透明巨型页并且映射到用户空间

FileHugePages: 0 kB

FilePmdMapped: 0 kB

CmaTotal: 0 kB

CmaFree: 0 kB

HugePages_Total: 0

HugePages_Free: 0

HugePages_Rsvd: 0

HugePages_Surp: 0

Hugepagesize: 2048 kB

Hugetlb: 0 kB

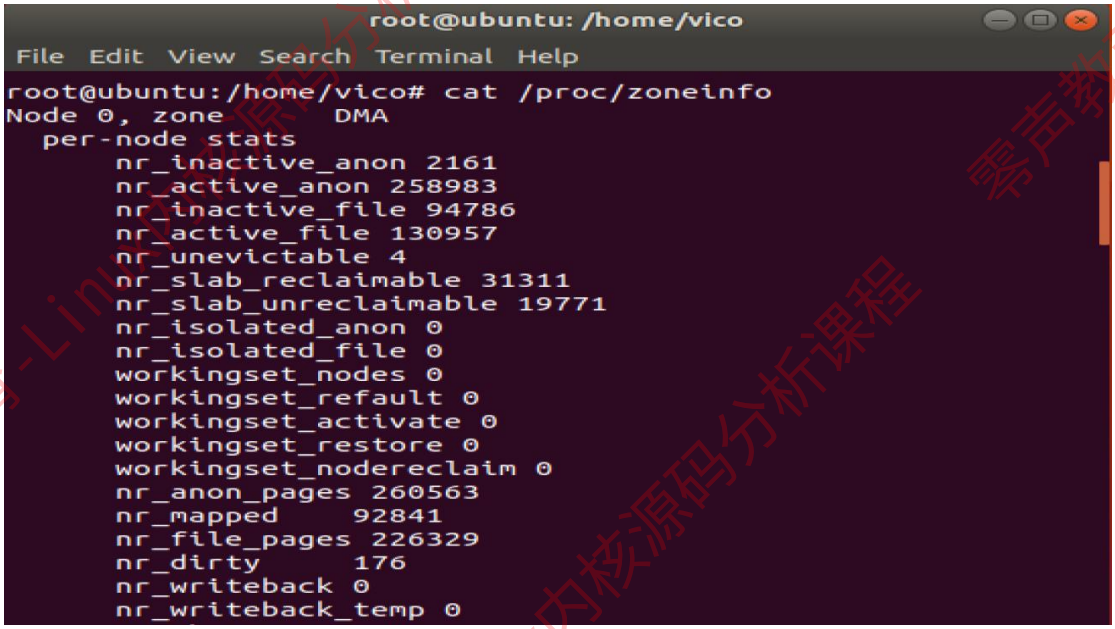
DirectMap4k: 161600 kB

DirectMap2M: 5081088 kB

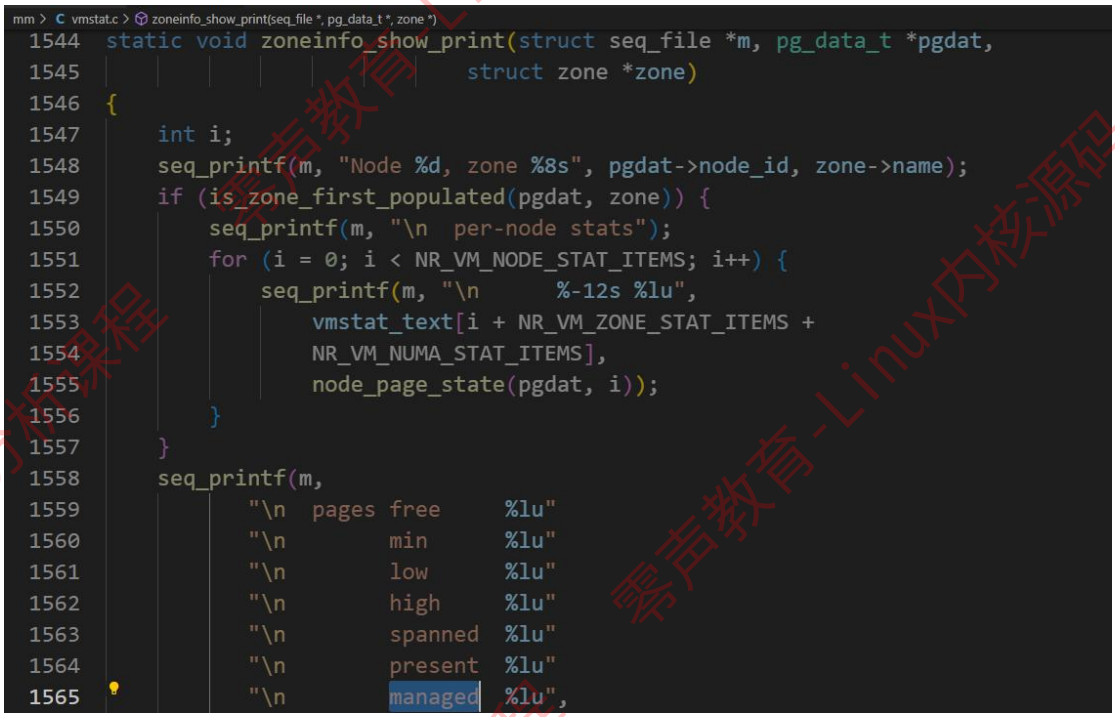
DirectMap1G: 13631488 kB

root@ubuntu:/home/vico#

二、详细分析/proc/zoneinfo 节点信息



具体系统调用实现，在 Linux 内核源码函数设计如下：



```
root@ubuntu:/home/vico# cat /proc/zoneinfo
Node 0, zone            DMA
  per-node stats
      nr_inactive_anon 1592
      nr_active_anon 206450
      nr_inactive_file 96845
      nr_active_file 84139
      nr_unevictable 4
      nr_slab_reclaimable 30210
      nr_slab_unreclaimable 19212
      nr_isolated_anon 0
      nr_isolated_file 0
      workingset_nodes 0
      workingset_refault 0
      workingset_activate 0
      workingset_restore 0
      workingset_nodereclaim 0
      nr_anon_pages 207679
      nr_mapped       59666
      nr_file_pages 181354
      nr_dirty        115
      nr_writeback 0
      nr_writeback_temp 0
      nr_shmem        1930
      nr_shmem_hugepages 0
      nr_shmem_pmdmapped 0
      nr_file_hugepages 0
      nr_file_pmdmapped 0
      nr_anon_transparent_hugepages 0
      nr_unstable  0
      nr_vmscan_write 0
      nr_vmscan_immediate_reclaim 0
      nr_dirtied      27385
      nr_written      8105
      nr_kernel_misc_reclaimable 0
    pages free          3976
          min           16
          low           20
          high          24
          spanned    4095
          present    3997
          managed    3976
          protection: (0, 2950, 15914, 15914, 15914)
```

```
        nr_free_pages 3976
        nr_zone_inactive_anon 0
        nr_zone_active_anon 0
        nr_zone_inactive_file 0
        nr_zone_active_file 0
        nr_zone_unevictable 0
        nr_zone_write_pending 0
        nr_mlock     0
        nr_page_table_pages 0
        nr_kernel_stack 0
        nr_bounce    0
        nr_zspages   0
        nr_free_cma  0
        numa_hit     0
        numa_miss    0
        numa_foreign 0
        numa_interleave 0
        numa_local   0
        numa_other   0
    pagesets
      cpu: 0
                count: 0
                high:  0
                batch: 1
    vm stats threshold: 2
    node_unreclaimable:  0
    start_pfn:           1
Node 0, zone      DMA32
    pages free     765443
          min      3128
          low      3910
          high     4692
        spanned  1044480
        present  782288
        managed  765904
        protection: (0, 0, 12964, 12964, 12964)
        nr_free_pages 765443
        nr_zone_inactive_anon 0
        nr_zone_active_anon 0
        nr_zone_inactive_file 0
        nr_zone_active_file 0
        nr_zone_unevictable 0
        nr_zone_write_pending 0
        nr_mlock     0
```

```
        nr_page_table_pages 0
        nr_kernel_stack 0
        nr_bounce       0
        nr_zspages      0
        nr_free_cma     0
        numa_hit        16
        numa_miss       0
        numa_foreign 0
        numa_interleave 1
        numa_local      16
        numa_other      0
    Pagesets //   // 表示每个 CPU 内存分配器中每个 CPU 缓存的页面数据信息
      cpu: 0
                count: 400
                high:    378
                batch: 63
    vm stats threshold: 12
    node_unreclaimable:    0 // 表示页面回收失败的次数
    start_pfn:             4096 // 表示内存管理区的起始页帧号
Node 0, zone      Normal
    pages free      2806651
          min       13750
          low       17187
          high      20624
          spanned   3407872
          present   3407872
          managed   3320750
          protection: (0, 0, 0, 0, 0)
        nr_free_pages 2806651
        nr_zone_inactive_anon 1592
        nr_zone_active_anon 206450
        nr_zone_inactive_file 96845
        nr_zone_active_file 84139
        nr_zone_unevictable 4
        nr_zone_write_pending 115
        nr_mlock        4
        nr_page_table_pages 10576
        nr_kernel_stack 11904
        nr_bounce       0
        nr_zspages      0
        nr_free_cma     0
        numa_hit        877306
        numa_miss       0
        numa_foreign 0
```

```
        numa_interleave 32733
        numa_local      877306
        numa_other      0
    pagesets
      cpu: 0
                  count: 191
                  high:    378
                  batch: 63
    vm stats threshold: 16
    node_unreclaimable:   0
    start_pfn:            1048576
Node 0, zone    Movable
    pages free      0
          min       0
          low       0
          high      0
          spanned   0
          present   0
          managed   0
          protection: (0, 0, 0, 0, 0)
Node 0, zone      Device
    pages free      0 // 表示此内存管理区中空闲页面的数量
          min       0 // 表示此内存管理区处于最低水位的页面数量
          low       0 // 表示此内存管理区处于低水位的页面数量
          high      0 // 表示此内存管理区处于高水位的页面数量
          spanned   0 // 表示此内存管理区包含的页面数量
          present   0 // 表示此内存管理区里实际管理的页面数量
          managed   0 // 表示此内存管理区中被伙伴系统管理的页面数量
          protection: (0, 0, 0, 0, 0) // 表示此内存管理区预留的内存
root@ubuntu:/home/vico#
```