

## 2.3.7 资源限制

讲师：汪洋





# 目录

1

内存资源限制

2

CPU 资源限制

3

限制实验



1

## 内存资源限制



- › CGroup 是 Control Groups 的缩写，是 Linux 内核提供了一种可以限制、记录、隔离进程组 (process groups) 所使用的物力资源 (如 cpu memory i/o 等等) 的机制。2007 年进入 Linux 2.6.24 内核，CGroups 不是全新创造的，它将进程管理从 cpuset 中剥离出来，作者是 Google 的 Paul Menage
- › 默认情况下，如果不对容器做任何限制，容器能够占用当前系统能给容器提供的所有资源
- › Docker 限制可以从 Memory、CPU、Block I/O 三个方面
- › OOME: Out Of Memory Exception
  - >> 一旦发生 OOME，任何进程都有可能被杀死，包括 docker daemon 在内
  - >> 为此，Docker 调整了 docker daemon 的 OOM 优先级，以免被内核关闭



- › 为应用做内存压力测试，理解正常业务需求下使用的内存情况，然后才能进入生产环境使用
- › 一定要限制容器的内存使用上限
- › 尽量保证主机的资源充足，一旦通过监控发现资源不足，就进行扩容或者对容器进行迁移
- › 如果可以（内存资源充足的情况），尽量不要使用 swap，swap 的使用会导致内存计算复杂，对调度器非常不友好



在 docker 启动参数中，和内存限制有关的包括（参数的值一般是内存大小，也就是一个正数，后面跟着内存单位 b、k、m、g，分别对应 bytes、KB、MB、和 GB）：

- `-m --memory`: 容器能使用的最大内存大小，最小值为 4m
- `--memory-swap`: 容器能够使用的 swap 大小
- `--memory-swappiness`: 默认情况下，主机可以把容器使用的匿名页（anonymous page）swap 出来，你可以设置一个 0-100 之间的值，代表允许 swap 出来的比例
- `--memory-reservation`: 设置一个内存使用的 soft limit, 设置值小于 `-m` 设置
- `--kernel-memory`: 容器能够使用的 kernel memory 大小，最小值为 4m。
- `--oom-kill-disable`: 是否运行 OOM 的时候杀死容器。只有设置了 `-m`，才可以把这个选项设置为 `false`，否则容器会耗尽主机内存，而且导致主机应用被杀死



--memory-swap	--memory	功能
正数S	正数M	容器可用总空间为S，其中ram为M，swap为(S-M)，若S=M，则无可用swap资源
0	正数M	相当于未设置swap (unset)
unset	正数M	若主机(Docker Host)启用了swap，则容器的可用swap为2*M
-1	正数M	若主机(Docker Host)启用了swap，则容器可使用最大至主机上的所有swap空间的swap资源

注意：在容器内使用free命令可以看到的swap空间并不具有其所展现出的空间指示意义。



2

## CPU 资源限制





Docker 提供的 CPU 资源限制选项可以在多核系统上限制容器能利用哪些 vCPU。而对容器最多能使用的 CPU 时间有两种限制方式：

- 一是有多个 CPU 密集型的容器竞争 CPU 时，设置各个容器能使用的 CPU 时间相对比例
- 二是以绝对的方式设置容器在每个调度周期内最多能使用的 CPU 时间



- `--cpuset-cpus=""` 允许使用的 CPU 集，值可以为 0-3, 0, 1
- `-c, --cpu-shares=0` CPU 共享权值（相对权重），默认值 1024
- `--cpuset-mems=""` 允许在上执行的内存节点（MEMs）
- `--cpu-period=0` 即可设置调度周期，CFS 周期的有效范围是 1ms~1s，对应的 `--cpu-period` 的数值范围是 1000~1000000
- `--cpu-quota=0` 设置在每个周期内容器能使用的 CPU 时间，容器的 CPU 配额必须不小于 1ms，即 `--cpu-quota` 的值必须  $\geq 1000$ ，单位微秒

## Example

```
docker run -it --cpu-period=50000 --cpu-quota=25000 ubuntu:16.04 /bin/bash
```

```
docker run -it --cpu-period=10000 --cpu-quota=20000 ubuntu:16.04 /bin/bash
```



--cpus 能够限制容器可以使用的主机 CPU 个数，并且还可以指定如 1.5 之类的小数



## NUMA

非统一内存访问（NUMA）是一种用于多处理器的电脑记忆体设计，内存访问时间取决于处理器的内存位置。在NUMA下，处理器访问它自己的本地存储器的速度比非本地存储器（存储器的地方到另一个处理器之间共享的处理器或存储器）快一些



# 3

## 限制实验



```
> docker run --name stress -it --rm -m 256m lorel/docker-stress-ng:latest stress -vm 2
```

```
> docker run --name stress -it --rm --cpus 2 lorel/docker-stress-ng:latest stress --cpu 8
```

```
> docker run --name stress -it --rm --cpuset-cpus 0 lorel/docker-stress-ng:latest stress --cpu 8
```



0 v e r