

分布式系统第二次作业

何泽 18340052

CRIU是一种在用户空间实现的进程或者容器 **checkpoint** 和 **restore** 的方法，从而实现进程或者容器的迁移。实际上，CRIU不仅可以实现冷迁移即离线迁移，还可以实现热迁移即在线迁移。请利用CRIU实现进程和容器的热迁移，并利用样例程序如web server等测试迁移过程中的性能损耗、观察发现，并撰写报告。

1. https://www.jianshu.com/p/2b288415896c?utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation;
2. <https://github.com/ZhuangweiKang/Docker-CRIU-Live-Migration>;
3. https://criu.org/Live_migration;

一、版本

- Ubuntu (18.04.1) :

```
heze@hezes-macbookpro:~$ cat /proc/version
Linux version 5.4.0-42-generic (bulld@lgw01-amd64-023) (gcc version 7.5.0 (Ubuntu 7.5.0-3ubuntu1~18.04)) #46~18.04.1-Ubuntu SMP Fri Jul 10 07:21:24 UTC 2020
heze@hezes-macbookpro:~$ uname -a
Linux hezes-macbookpro 5.4.0-42-generic #46~18.04.1-Ubuntu SMP Fri Jul 10 07:21:24 UTC 2020 x86_64 x86_64 x86_64 GNU/Linux
```

- Docker (已开启实验性功能)

```
heze@hezes-macbookpro:~$ docker version
Client:
 Version:      17.06.0-ce
 API version:  1.30
 Go version:   go1.8.3
 Git commit:   02c1d87
 Built:        Fri Jun 23 21:23:31 2017
 OS/Arch:      linux/amd64

Server:
 Version:      17.06.0-ce
 API version:  1.30 (minimum version 1.12)
 Go version:   go1.8.3
 Git commit:   02c1d87
 Built:        Fri Jun 23 21:19:04 2017
 OS/Arch:      linux/amd64
 Experimental: true
```

- CRIU (3.14) :

```
heze@hezes-macbookpro:~$ criu -V
Version: 3.14
```

二、进程迁移

- 首先写了如下脚本:

```
1 | i=0; while true; do echo $i; i=$((expr $i + 1)); sleep 1; done
```

每秒钟计数加一

- 查看当前进程PID

```
1 | pgrep -f test.sh
```

```
heze@hezes-macbookpro:~$ pgrep -f test.sh
1150
```

- 使用dump保存进程镜像

```
1 | sudo criu dump --t 1150 --images-dir ~/Desktop/test --shell-job
```

```
heze@hezes-macbookpro:~$ sudo criu dump -t 1150 --images-dir ~/Desktop/test --shell-job
[sudo] password for heze:
Warn (compel/arch/x86/src/lib/infect.c:281): Will restore 1231 with interrupted system call
heze@hezes-macbookpro:~$ |
```

- 运行状态

```
heze@hezes-macbookpro:~/Desktop$ bash ./test.sh
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Killed
heze@hezes-macbookpro:~/Desktop$ |
```

可以看到，运行至15时被停止、保存

- 可以看到此时生成了很多img文件：

```
heze@hezes-macbookpro:~/Desktop/test$ ls
core-1150.img  fdinfo-3.img  fs-1231.img  inventory.img  pagemap-1150.img  pages-2.img  stats-dump
core-1231.img  files.img     ids-1150.img  mm-1150.img   pagemap-1231.img  pstree.img   timens-0.img
fdinfo-2.img  fs-1150.img  ids-1231.img  mm-1231.img   pages-1.img       seccomp.img  tty-info.img
```

- 压缩：

```
1 | tar -cvf test.tar.gz ~/Desktop/test
```

```
heze@hezes-macbookpro:~/Desktop$ tar -cvf test.tar.gz ~/Desktop/test
tar: Removing leading `/' from member names
/home/heze/Desktop/test/
/home/heze/Desktop/test/fs-1150.img
/home/heze/Desktop/test/stats-dump
/home/heze/Desktop/test/seccomp.img
/home/heze/Desktop/test/files.img
/home/heze/Desktop/test/fdinfo-3.img
/home/heze/Desktop/test/pages-1.img
/home/heze/Desktop/test/inventory.img
/home/heze/Desktop/test/ids-1150.img
/home/heze/Desktop/test/fs-1231.img
/home/heze/Desktop/test/ids-1231.img
/home/heze/Desktop/test/pstree.img
/home/heze/Desktop/test/fdinfo-2.img
/home/heze/Desktop/test/mm-1150.img
/home/heze/Desktop/test/pagemap-1150.img
/home/heze/Desktop/test/pagemap-1231.img
/home/heze/Desktop/test/pages-2.img
/home/heze/Desktop/test/tty-info.img
/home/heze/Desktop/test/core-1150.img
/home/heze/Desktop/test/mm-1231.img
/home/heze/Desktop/test/core-1231.img
/home/heze/Desktop/test/timens-0.img
```

- 我使用U盘将压缩文件拷贝至另一台机器，并解压缩：

```
1 | tar -xvf ~/Desktop/test.tar.gz
```

```
heze@hezes-dell:~/Desktop$ tar -xvf ~/Downloads/test.tar.gz
home/heze/Desktop/test/
home/heze/Desktop/test/fs-1150.img
home/heze/Desktop/test/stats-dump
home/heze/Desktop/test/seccomp.img
home/heze/Desktop/test/files.img
home/heze/Desktop/test/fdinfo-3.img
home/heze/Desktop/test/pages-1.img
home/heze/Desktop/test/inventory.img
home/heze/Desktop/test/ids-1150.img
home/heze/Desktop/test/fs-1231.img
home/heze/Desktop/test/ids-1231.img
home/heze/Desktop/test/pstree.img
home/heze/Desktop/test/fdinfo-2.img
home/heze/Desktop/test/mm-1150.img
home/heze/Desktop/test/pagemap-1150.img
home/heze/Desktop/test/pagemap-1231.img
home/heze/Desktop/test/pages-2.img
home/heze/Desktop/test/tty-info.img
home/heze/Desktop/test/core-1150.img
home/heze/Desktop/test/mm-1231.img
home/heze/Desktop/test/core-1231.img
home/heze/Desktop/test/timens-0.img
```

- 恢复:

```
1 | sudo criu restore -t 1200 --images-dir ~/Desktop/test --  
    shell-job
```

```
heze@hezes-dell:~/Desktop$ sudo criu restore -t 1200 --images-dir ~/Desktop/test  
--shell-job  
[sudo] password for heze:  
Warn (criu/crtools.c:217): Using -t with criu restore is obsoleted  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32
```

可以看到从16开始计数，恢复成功

三、Docker迁移

- 首先运行一个容器，每5秒加一

```
1 | docker run -d --name looper2 --security-opt  
    seccomp:unconfined busybox \  
2      /bin/sh -c 'i=0; while true; do echo $i; i=$((expr $i  
    + 1)); sleep 5; done'
```

```
heze@hezes-macbookpro:~$ docker run -d --name looper2 --security-opt seccomp:unconfined busybox  
/bin/sh -c 'i=0; while true; do echo $i; i=$((expr $i + 1)); sleep 5; done'  
1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f
```

- 使用 `checkpoint` 保存状态，保存到 `/home` 目录下

```
1 | docker checkpoint create --checkpoint-dir=/home looper2  
    checkpoint2
```

```

heze@hezes-macbookpro:~$ docker checkpoint create --checkpoint-dir=/home looper2 checkpoint2
checkpoint2
heze@hezes-macbookpro:~$ docker logs looper2
0
1
2
3
4
5
6
7

```

- 此时 `/home` 目录下会生成ID名的文件夹，进去 `/checkpoints/checkpoint2` 就可以看到镜像文件

```

heze@hezes-macbookpro:/home$ ls
1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f  heze  ubuntu
heze@hezes-macbookpro:/home$ cd 1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f/
heze@hezes-macbookpro:/home/1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f$ ls
checkpoints
heze@hezes-macbookpro:/home/1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f$ cd checkpoints/
heze@hezes-macbookpro:/home/1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f/checkpoints$ ls
checkpoint2
heze@hezes-macbookpro:/home/1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f/checkpoints$ cd checkpoint2/
heze@hezes-macbookpro:/home/1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f/checkpoints/checkpoint2$ ls
cgroup.img      files.img        mm-22.img        seccomp.img
config.json     fs-1.img         mountpoints-13.img  timens-0.img
core-1.img      fs-22.img        pagemap-1.img     tmpfs-dev-59.tar.gz.img
core-22.img     ids-1.img        pagemap-22.img    tmpfs-dev-62.tar.gz.img
criu.work       ids-22.img       pages-1.img       tmpfs-dev-63.tar.gz.img
descriptors.json  inventory.img    pages-2.img       utsns-12.img
heze@hezes-macbookpro:/home/1aa6d67b323a1c07ed75cbb145d1d56df11685413e11fd2e79255afd31fea13f/checkpoints/checkpoint2$ |

```

- 和上面一样，压缩，用U盘传到另一台电脑，再解压缩
- 首先只创建但不运行一个容器，命名为 `looper3`

```

1 | docker create --name looper3 --security-opt
   | seccomp:unconfined busybox \
2   | /bin/sh -c 'i=0; while true; do echo $i; i=$(expr $i
   | + 1); sleep 1; done'

```

```

heze@hezes-dell:~$ docker create --name looper3 --security-opt seccomp:unconfine
d busybox /bin/sh -c 'i=0; while true; do echo $i; i=$(expr $i + 1);
sleep 5; done'
0797371c2a1b5407526313128b077f743eba3a140e76332b4071741d893e2c34

```

- 恢复：

```

1 | docker start --checkpoint-dir=/home --checkpoint=checkpoint2
   | looper3

```

```
heze@hezes-dell:~$ docker start --checkpoint-dir=/home/1aa6d67b323a1c07ed75cbb14
5d1d56df11685413e11fd2e79255afd31fea13f/checkpoints/ --checkpoint=checkpoint2 lo
oper3
heze@hezes-dell:~$ docker logs looper3
8
9
heze@hezes-dell:~$ docker logs looper3
8
9
10
```

可以看到从前面的断点7继续，从8开始，迁移成功

至此，进程和容器的热迁移全部成功。