Assignment-9

Linux System and its Applications

Systems and Storage Laboratory
Department of Computer Science and Engineering
Chung-Ang University

Assignment-9: Per-core vs Original Calclock

1.Use per-core calclock to profile pxt4_file_write_iter()

- Build pxt4 module and mount the testing device with pxt4
- Run FIO test, refering [Practical Class 7-a]
- Unmount the testing device and remove pxt4 module
- Run **dmesg** to check results

2. Use original calclock to profile pxt4_file_write_iter()

- Refer [Practical Class 7-b]
- Steps as the per-core calclock above

3. Compare pxt4 performance measured by FIO in two cas es

- Per-core calclock added to profile pxt4_file_write_iter()
- Original calclock added to profile pxt4_file_write_iter()

Assignment-9: Per-core vs Original Calclock

- **What to submit**

Take screenshots of

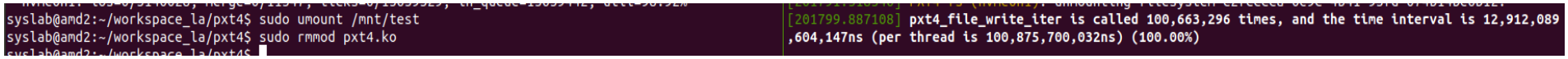
- **dmesg** result from **ktprint()** of per-core calclock
- **FIO** results after adding per-core calclock to profile pxt4_file_write_iter()
- A table of performance comparison

Throughput	Per-core Calclock	Original Calclock
IOPS		
MB/s		

- **Submit in pdf file including your name and student ID**
- "calclock.h" and "calclock.c" files are provided at https://github.com/loglamo/CAU-CSE-LinuxApplications-20232

Examples of screenshot 1

Dmesg result from ktprint()



Examples of screenshot 2

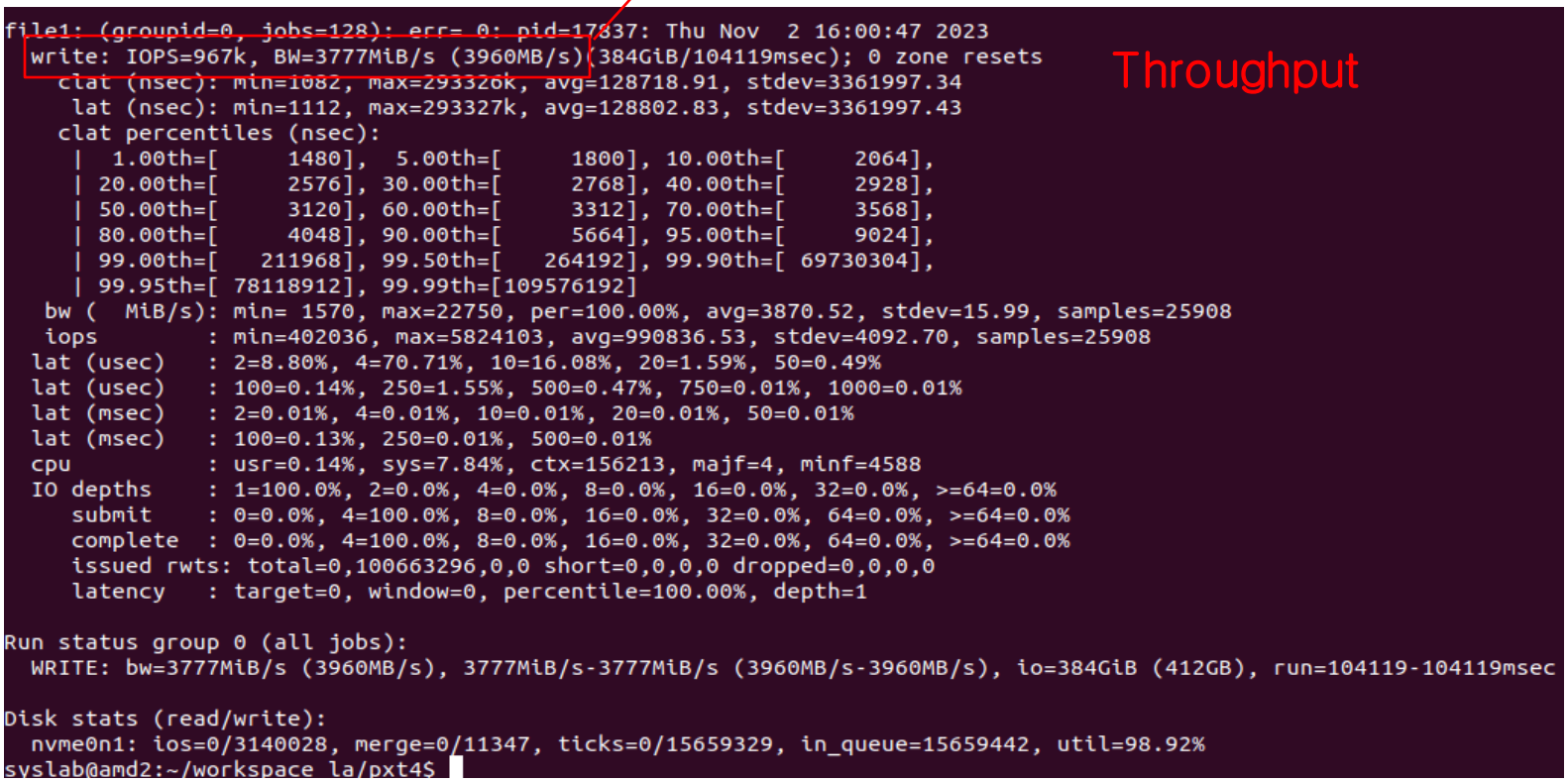
FIO result with random writes after performing sequential writes while adding per-core calclock to profile pxt4_file_write_iter()

```
file1: (groupid=0, jobs=128): err= 0: pid=17837: Thu Nov  2 16:00:47 2023
  write: IOPS=967k, BW=3777MiB/s (3960MB/s)(384GiB/104119msec); 0 zone resets
    clat (nsec): min=1082, max=293326k, avg=128718.91, stdev=3361997.34
     lat (nsec): min=1112, max=293327k, avg=128802.83, stdev=3361997.43
    clat percentiles (nsec):
     |  1.00th=[     1480],  5.00th=[     1800], 10.00th=[     2064],
     | 20.00th=[     2576], 30.00th=[     2768], 40.00th=[     2928],
     | 50.00th=[     3120], 60.00th=[     3312], 70.00th=[     3568],
     | 80.00th=[     4048], 90.00th=[     5664], 95.00th=[     9024],
     | 99.00th=[   211968], 99.50th=[   264192], 99.90th=[ 69730304],
     | 99.95th=[ 78118912], 99.99th=[109576192]
   bw (  MiB/s): min= 1570, max=22750, per=100.00%, avg=3870.52, stdev=15.99, samples=25908
   iops        : min=402036, max=5824103, avg=990836.53, stdev=4092.70, samples=25908
  lat (usec)   : 2=8.80%, 4=70.71%, 10=16.08%, 20=1.59%, 50=0.49%
  lat (usec)   : 100=0.14%, 250=1.55%, 500=0.47%, 750=0.01%, 1000=0.01%
  lat (msec)   : 2=0.01%, 4=0.01%, 10=0.01%, 20=0.01%, 50=0.01%
  lat (msec)   : 100=0.13%, 250=0.01%, 500=0.01%
  cpu          : usr=0.14%, sys=7.84%, ctx=156213, majf=4, minf=4588
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=0,100663296,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=1

Run status group 0 (all jobs):
  WRITE: bw=3777MiB/s (3960MB/s), 3777MiB/s-3777MiB/s (3960MB/s-3960MB/s), io=384GiB (412GB), run=104119-104119msec

Disk stats (read/write):
  nvme0n1: ios=0/3140028, merge=0/11347, ticks=0/15659329, in_queue=15659442, util=98.92%
syslab@amd2:~/workspace_la/pxt4$
```

Examples of screenshot 3

Table of performance comparison with random writes

Throughput	Per-core Calclock	Original Calclock
IOPS	967k	*[From previous class]*
MB/s	3960MB/s	*[From previous class]*




```
file1: (groupid=0, jobs=128): err= 0: pid=17837: Thu Nov  2 16:00:47 2023
  write: IOPS=967k, BW=3777MiB/s (3960MB/s)(384GiB/104119msec); 0 zone resets
    clat (nsec): min=1082, max=293326k, avg=128718.91, stdev=3361997.34
     lat (nsec): min=1112, max=293327k, avg=128802.83, stdev=3361997.43
    clat percentiles (nsec):
     |  1.00th=[     1480],  5.00th=[     1800], 10.00th=[     2064],
     | 20.00th=[     2576], 30.00th=[     2768], 40.00th=[     2928],
     | 50.00th=[     3120], 60.00th=[     3312], 70.00th=[     3568],
     | 80.00th=[     4048], 90.00th=[     5664], 95.00th=[     9024],
     | 99.00th=[   211968], 99.50th=[   264192], 99.90th=[ 69730304],
     | 99.95th=[ 78118912], 99.99th=[109576192]
   bw (  MiB/s): min= 1570, max=22750, per=100.00%, avg=3870.52, stdev=15.99, samples=25908
   iops        : min=402036, max=5824103, avg=990836.53, stdev=4092.70, samples=25908
  lat (usec)   : 2=8.80%, 4=70.71%, 10=16.08%, 20=1.59%, 50=0.49%
  lat (usec)   : 100=0.14%, 250=1.55%, 500=0.47%, 750=0.01%, 1000=0.01%
  lat (msec)   : 2=0.01%, 4=0.01%, 10=0.01%, 20=0.01%, 50=0.01%
  lat (msec)   : 100=0.13%, 250=0.01%, 500=0.01%
  cpu          : usr=0.14%, sys=7.84%, ctx=156213, majf=4, minf=4588
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=0,100663296,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=1

Run status group 0 (all jobs):
  WRITE: bw=3777MiB/s (3960MB/s), 3777MiB/s-3777MiB/s (3960MB/s-3960MB/s), io=384GiB (412GB), run=104119-104119msec

Disk stats (read/write):
  nvme0n1: ios=0/3140028, merge=0/11347, ticks=0/15659329, in_queue=15659442, util=98.92%
syslab@amd2:~/workspace_la/pxt4$
```

Notes

- **calclock.h** files used in per-core and original calclocks may have the same name, but different contents. When profiling functions with the per-core calclock, please move **calclock.h** file of original calclock outside the pxt4 folder.
- Include **calclock.h** in ./file.c and ./super.c
- Add **calclock.o** to "pxt4-y" in ./Makefile

Notes

- **Create pxt4_file_write_iter_internal()** according to the original pxt4_file_write_iter() when adding per-core calclock to ./file.c

Original pxt4_file_write_iter() in ./file.c

```
#include "calclock.h"
static ssize_t pxt4_file_write_iter(struct kiocb *iocb, struct iov_iter *from) {
      struct inode *inode = file_inode;
      …
}
```

Add pxt4_file_write_iter_internal() in ./file.c

```
#include "calclock.h"
static ssize_t pxt4_file_write_iter_internal(struct kiocb *iocb, struct iov_iter
*from) {
      struct inode *inode = file_inode;
      … }
KTDEF(pxt4_file_write_iter);
static ssize_t pxt4_file_write_iter(struct kiocb *iocb, struct iov_iter *from) {
       ssize_t ret;
       ktime_t localclock[2];
       ktget(&localclock[0]);
       ret = pxt4_file_write_iter_internal(iocb,from);
       ktget(&localclock[1]);
       ktput(localclock, pxt4_file_write_iter);
       return ret;}
```

Notes

- **Do not forget** main steps after successfully making pxt4 module:
  - $insmod [module.ko]
  - $mkfs –t ext4 [device]

(E.g., ***mkfs -t ext4 /dev/nvme0n1***)

  - $mount –t pxt4 [device] [mount point]

(E.g., ***mount –t pxt4 /dev/nvme0n1 /mnt/test***)

  - Run FIO
  - $umount [mount point]

(E.g., ***umount /mnt/test***)

  - $rmmod [module.ko]
  - Check **dmesg**