CSC494 Commands

Fengjia Zhang

Yuhan Shao

1. Environment variable set up for spark

In ${SPARK\_HOME}/conf/spark-env.sh:

SPARK\_EXECUTOR\_MEMORY=126G

SPARK\_DRIVER\_MEMORY=126G

export SPARK\_LOCAL\_DIRS=spark.local.dir /home/zhangf68/spark\_folder/sparktmp

// /home/zhangf68/spark\_folder is set to be the home directory for spark, or otherwise spark may run on root

1. Setup
   1. Blocktrace

mkdir blocktrace

sudo mkfs.ext4 /dev/nvme2n1

sudo mount /dev/nvme2n1 blocktrace

sudo mount // to check whether the file system is formatted correctly

// sudo umount spark

* 1. Spark on different file system

mkdir spark\_folder

sudo mkfs.ext4 /dev/nvme1n1 //ext4

sudo mkfs.btrfs -f /dev/nvme1n1 // Btrfs

sudo mkfs.f2fs /dev/nvme1n1 // F2FS

sudo mount /dev/nvme1n1 ./spark\_folder

sudo mount // to check whether the file system is formatted correctly

// sudo umount spark

// copy the spark to spark\_folder

scp -r zhangf68@swift-014.syslab.sandbox:/home/zhangf68/spark ./spark\_folder

1. Running
   1. Blocktrace

sudo blktrace -d /dev/nvme1n1 -w 3600 -o - | blkparse -a fs -i - > blocktrace/filename

* 1. Spark on different file system

In SPARK\_HOME directory(zhangf68@mel-13:~/spark\_folder/spark/spark-2.4.0-bin-hadoop2.7$):

time ./bin/run-example graphx.Analytics pagerank ../ sx-stackoverflow.edges --numEPart=-1 --vertexStorageLevel=MEMORY\_AND\_DISK --edgeStorageLevel=MEMORY\_AND\_DISK