练习 1: Flushes and Compactions

```
hbase(main):006:0> create 'mytable',NAME=>'cf',VERSIONS=>3
0 row(s) in 2.2810 seconds
hbase(main):008:0> flush 'mytable'
0 row(s) in 0.3760 seconds

hbase(main):009:0> put 'mytable','row1','cf:col','foo'
0 row(s) in 0.1340 seconds
hbase(main):011:0> flush 'mytable'
0 row(s) in 0.4400 seconds
```

#可以看到都一个文件
```
[ian@h1           ~]$         hdfs        dfs       -ls        -R
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf
-rw-r--r--        1   ian   supergroup            4906   2018-08-01   16:34
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/6b051aab448b4496af70
663ee4c92474
```

```
hbase(main):012:0> put 'mytable','row2','cf:col','too'
0 row(s) in 0.0090 seconds

hbase(main):013:0> flush 'mytable'
0 row(s) in 0.7490 seconds
```

#可以看到都两个文件
```
[ian@h1           ~]$         hdfs        dfs       -ls        -R
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf
-rw-r--r--        1   ian   supergroup            4906   2018-08-01   16:35
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/1aa2f94542a443f1ac2b8
92210c72884
-rw-r--r--        1   ian   supergroup            4906   2018-08-01   16:34
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/6b051aab448b4496af70
663ee4c92474
```

#查看文件内容
```
[ian@h1           ~]$         hbase        hfile       --printkv       --file
hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/6b051aab
448b4496af70663ee4c92474
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J:                    Found              binding                 in
[jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.c
lass]
SLF4J:                    Found              binding                 in
[jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf
4j/impl/StaticLoggerBinder.class]
```

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

2018-08-01 16:37:30,026 INFO [main] hfile.CacheConfig: Created cacheConfig: CacheConfig:disabled

K: row1/cf:col/1533112474031/Put/vlen=3/seqid=5 V: foo

Scanned kv count -> 1

---

hbase(main):014:0> major_compact 'mytable'

0 row(s) in 0.0860 seconds

#变回了一个文件

[ian@h1 ~]$ hdfs dfs -ls -R /hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf

-rw-r--r-- 1 ian supergroup 4941 2018-08-01 16:38 /hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/b12e07ef4aa4403d8db7cc9931495cae

#查看文件内容

[ian@h1 ~]$ hbase hfile --printkv --file hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/b12e07ef4aa4403d8db7cc9931495cae

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

2018-08-01 16:38:53,371 INFO [main] hfile.CacheConfig: Created cacheConfig: CacheConfig:disabled

K: row1/cf:col/1533112474031/Put/vlen=3/seqid=5 V: foo

K: row2/cf:col/1533112515743/Put/vlen=3/seqid=10 V: too

Scanned kv count -> 2

#查看 Write Ahead Log

[ian@h1 ~]$ hbase hlog hdfs://h1:9000/hbase/WALs/h2,16020,1532513475389/h2%2C16020%2C1532513475389..meta.1533111172710.meta

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

```
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Writer Classes: ProtobufLogWriter
Cell Codec Class: org.apache.hadoop.hbase.regionserver.wal.WALCellCodec
Sequence=72 , region=1588230740 at write timestamp=Wed Aug 01 16:28:34 CST 2018
row=mytable,,1533112113539.1a70be711154321068b158fdef4454e3., column=info:regioninfo
Sequence=73 , region=1588230740 at write timestamp=Wed Aug 01 16:28:34 CST 2018
row=mytable,,1533112113539.1a70be711154321068b158fdef4454e3., column=info:regioninfo
row=mytable,,1533112113539.1a70be711154321068b158fdef4454e3., column=info:server
row=mytable,,1533112113539.1a70be711154321068b158fdef4454e3.,
column=info:serverstartcode
row=mytable,,1533112113539.1a70be711154321068b158fdef4454e3.,
column=info:seqnumDuringOpen
Sequence=74 , region=1588230740 at write timestamp=Wed Aug 01 16:29:57 CST 2018
row=mytable,,1533112113539.1a70be711154321068b158fdef4454e3., column=info:
Sequence=75 , region=1588230740 at write timestamp=Wed Aug 01 16:30:04 CST 2018
row=mytable,,1533112203632.1ae1b85d43f326f0aae64a3cbc84b2ca., column=info:regioninfo
Sequence=76 , region=1588230740 at write timestamp=Wed Aug 01 16:30:04 CST 2018
row=mytable,,1533112203632.1ae1b85d43f326f0aae64a3cbc84b2ca., column=info:regioninfo
row=mytable,,1533112203632.1ae1b85d43f326f0aae64a3cbc84b2ca., column=info:server
row=mytable,,1533112203632.1ae1b85d43f326f0aae64a3cbc84b2ca.,
column=info:serverstartcode
row=mytable,,1533112203632.1ae1b85d43f326f0aae64a3cbc84b2ca.,
column=info:seqnumDuringOpen
Sequence=79 , region=1588230740 at write timestamp=Wed Aug 01 16:34:40 CST 2018
row=\x00, column=METAFAMILY:HBASE::FLUSH
Sequence=80 , region=1588230740 at write timestamp=Wed Aug 01 16:34:40 CST 2018
row=\x00, column=METAFAMILY:HBASE::FLUSH
```

练习 2: Compactions and Data

```
hbase(main):015:0> put 'mytable','row3','cf:col','row3value'
0 row(s) in 0.0220 seconds

hbase(main):016:0> flush 'mytable'
0 row(s) in 0.7820 seconds

hbase(main):017:0> major_compact 'mytable'
0 row(s) in 0.0630 seconds
```

#使用 hfile － printkv 确认 row3 数据是否存在
```
[ian@h1          ~]$          hdfs          dfs          -ls          -R
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf
-rw-r--r--        1    ian    supergroup              4984    2018-08-01    17:10
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/01d98929abcf4a8d988f3
```

```
30660640e1d
[ian@h1          ~]$          hbase          hfile          --printkv          --file
hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/01d98929
abcf4a8d988f330660640e1d
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J:                       Found                       binding                       in
[jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.c
lass]
SLF4J:                       Found                       binding                       in
[jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf
4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2018-08-01     17:11:30,823     INFO     [main]     hfile.CacheConfig:     Created     cacheConfig:
CacheConfig:disabled
K: row1/cf:col/1533112474031/Put/vlen=3/seqid=5 V: foo
K: row2/cf:col/1533112515743/Put/vlen=3/seqid=10 V: too
K: row3/cf:col/1533113195669/Put/vlen=9/seqid=16 V: row3value
Scanned kv count -> 3
```

```
hbase(main):018:0> deleteall 'mytable','row3'
0 row(s) in 0.0500 seconds
```
hbase(main):020:0> flush 'mytable'

0 row(s) in 0.7420 seconds

#使用 hfile －printkv 确认 row3 数据存在
```
[ian@h1          ~]$          hbase          hfile          --printkv          --file
hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/01d98929
abcf4a8d988f330660640e1d
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J:                       Found                       binding                       in
[jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.c
lass]
SLF4J:                       Found                       binding                       in
[jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf
4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2018-08-01     17:16:33,198     INFO     [main]     hfile.CacheConfig:     Created     cacheConfig:
CacheConfig:disabled
K: row1/cf:col/1533112474031/Put/vlen=3/seqid=5 V: foo
K: row2/cf:col/1533112515743/Put/vlen=3/seqid=10 V: too
```
K: row3/cf:col/1533113195669/Put/vlen=9/seqid=16 V: row3value
```
Scanned kv count -> 3
[ian@h1          ~]$          hdfs          dfs          -ls          -R
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf
```

-rw-r--r-- 1 ian supergroup 4984 2018-08-01 17:10 /hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/01d98929abcf4a8d988f330660640e1d

-rw-r--r-- 1 ian supergroup 5079 2018-08-01 17:16 /hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/def6a5cadf6e425793b2227f4c9e30f2

[ian@h1 ~]$ hbase hfile --printkv --file hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/def6a5cadf6e425793b2227f4c9e30f2

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

2018-08-01 17:17:08,592 INFO [main] hfile.CacheConfig: Created cacheConfig: CacheConfig:disabled

K: row3/cf:/1533114974811/DeleteFamily/vlen=0/seqid=22 V:

Scanned kv count -> 1

#使用 hfile －printkv 确认 row3 数据存在

[ian@h1 ~]$ hdfs dfs -ls -R /hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf

-rw-r--r-- 1 ian supergroup 4941 2018-08-01 17:20 /hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/45472beb375e407b991b322006d96d35

[ian@h1 ~]$ hbase hfile --printkv --file hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/45472beb375e407b991b322006d96d35

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

2018-08-01 17:20:52,840 INFO [main] hfile.CacheConfig: Created cacheConfig: CacheConfig:disabled

K: row1/cf:col/1533112474031/Put/vlen=3/seqid=5 V: foo

K: row2/cf:col/1533112515743/Put/vlen=3/seqid=10 V: too

Scanned kv count -> 2

hbase(main):022:0> put 'mytable','row4','cf:col','first data'

0 row(s) in 0.0180 seconds

hbase(main):023:0> put 'mytable','row4','cf:col','2nd data'

0 row(s) in 0.0090 seconds

hbase(main):024:0> put 'mytable','row4','cf:col','third data'

0 row(s) in 0.0070 seconds

hbase(main):025:0> put 'mytable','row4','cf:col','4th data'

0 row(s) in 0.0060 seconds

hbase(main):027:0> flush 'mytable'

0 row(s) in 0.7250 seconds

#注意：在 major_compact 之前已经只剩下 3 条数据！！！

[ian@h1   ~]$ hbase   hfile   --printkv   --file
hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/a766a995
d70e407ba534918ee8948507

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found   binding   in
[jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.c
lass]

SLF4J: Found   binding   in
[jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf
4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

2018-08-01   17:24:22,771   INFO   [main]   hfile.CacheConfig: Created   cacheConfig:
CacheConfig:disabled

K: row4/cf:col/1533115424842/Put/vlen=8/seqid=31 V: 4th data

K: row4/cf:col/1533115418800/Put/vlen=10/seqid=30 V: third data

K: row4/cf:col/1533115411683/Put/vlen=8/seqid=29 V: 2nd data

Scanned kv count -> 3

---

hbase(main):028:0> major_compact 'mytable'

0 row(s) in 0.0690 seconds

[ian@h1   ~]$ hdfs   dfs   -ls   -R
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf

-rw-r--r--   1   ian   supergroup   5061   2018-08-01   17:27
/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/6fd107862c7240a3a973e
313d18507b0

[ian@h1   ~]$ hbase   hfile   --printkv   --file
hdfs://h1:9000/hbase/data/default/mytable/1ae1b85d43f326f0aae64a3cbc84b2ca/cf/6fd10786
2c7240a3a973e313d18507b0

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/ian/installed/hbase/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/ian/installed/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

2018-08-01 17:28:02,073 INFO [main] hfile.CacheConfig: Created cacheConfig: CacheConfig:disabled

K: row1/cf:col/1533112474031/Put/vlen=3/seqid=5 V: foo

K: row2/cf:col/1533112515743/Put/vlen=3/seqid=10 V: too

K: row4/cf:col/1533115424842/Put/vlen=8/seqid=31 V: 4th data

K: row4/cf:col/1533115418800/Put/vlen=10/seqid=30 V: third data

K: row4/cf:col/1533115411683/Put/vlen=8/seqid=29 V: 2nd data

Scanned kv count -> 5