4.6 적재 파일럿 실행 4단계 – 적재 기능 테스트 플럼의 사용자 정의 Interceptor 추가

- FTP 클라이언트인 파일질라 실행해 Server02에 접속
- 업로드 경로: /opt/cloudera/parcels/CDH/lib/flume-ng/lib
- C://예제소스/bigdata2nd-master/CH04/bigdata.smartcar.flume-1.0.jar 파일을 /opt/cloudera/parcels/ CDH/lib/flume-ng/lib 경로에 업로드

4.6 적재 파일럿 실행 4단계 – 적재 기능 테스트 플럼의 Conf 파일 수정

- 01. CM 홈 → [Flume] → [구성]을 선택해서 예제 4.2의 내용을 플럼의 "구성파일"에서 직접 작성하거나 C://예제소스/ bigdata2nd-master/CH04/예제4.2/SmartCar_Agent.conf 파일을 열어 전체 내용을 복사한 후 붙여넣는다.
- 02. 플럼의 [변경 내용 저장] 버튼을 클릭한다.
- 03. CM 홈으로 다시 이동해서 플럼을 재시작한다.

☆ 4.6 적재 파일럿 실행 4단계 – 적재 기능 테스트

SmartCar 로그 시뮬레이터 작동

- **01.** 먼저 Server02에 SSH로 접속하고 bigdata.smartcar.loggen-1,0.jar가 위치한 경로인 /home/pilot-pjt/working으로 이동한다.
 - \$ cd /home/pilot-pjt/working
- 02. 스마트카 로그 시뮬레이터를 아래의 자바 명령으로 백그라운드 방식으로 실행한다. 2016년 1월 1일의 100대의 스마트카에 상태 정보 로그를 만들어 보자.
 - \$ java -cp bigdata.smartcar.loggen-1.0.jar com.wikibook.bigdata.smartcar.loggen.CarLogMain
 20160101 100 &
- 03. 정상적으로 작동됐는지 다음 내용으로 확인해 본다.
 - Inome/pilot-pjt/working/SmartCar 경로에 SmartCarStatusInfo_20160101.txt 파일이 생성됐는지 확인한다.
 2016년 1월 1일 날짜로 100대의 스마트카 상태 정보가 기록된 것을 확인할 수 있다. 최종 로그 파일의 크기는 100MB이고 생성되기까지 1∼2분 정도 걸린다.
 - \$ cd /home/pilot-pjt/working/SmartCar
 - \$ tail -f SmartCarStatusInfo_20160101.txt

4.6 적재 파일럿 실행 4단계 – 적재 기능 테스트 **플럼 이벤트 작동 (1/3)**

01. /home/pilot-pit/working/SmartCar 경로에 만들어진 SmartCarStatusInfo 20160101.txt 파일을 플럼의 SmartCarInfo의 SpoolDir 경로인 /home/pilot-pit/working/car-batch-log로 옮겨서 플럼의 File 이벤트가 작 동되게 해보자.

\$ mv /home/pilot-pjt/working/SmartCar/SmartCarStatusInfo_20160101.txt /home/pilot-pjt/ working/car-batch-log/

₩☆ 4.6 적재 파일럿 실행 4단계 - 적재 기능 테스트

- 02. 플럼의 실행 로그를 통해 SmartCarInfo_Agent가 정상적으로 작동하는지 확인한다. 아래의 디렉터리로 이동해서 tail 명령으로 플럼의 실행 로그를 확인해 보자.
 - \$ cd /var/log/flume-ng/
 - \$ tail -f /var/log/flume-ng/flume-cmf-flume-AGENT-server02.hadoop.com.log

그림 4.12와 같이 특별한 에러 없이 "Creating /pilot-pjt/collect···" 또는 "Updating checkpoint for file:···" 메시지가 나타나면 적상적으로 HDFS에 적재 중인 것이다. 완전히 적재되기까지 필자의 파일럿 PC를 기준으로 약5~10분 정도 소요됐다.

```
2016-03-13 14:15:09,474 INFO org.apache.flume.sink.hdfs.HDFSDataStream: Serializer = TFXT. HasBawl.ccalFileSystem = false
2016-03-13 14:15:11,394 INFO org.apache.flume.sink.hdfs.BucketWriter: Creating /pilot-pjt/collect/car-batch-log/20160313/car-batch-log.1457846109475.log.tmp
2016-03-13 14:15:17,140 WARN org.apache.flume.sink.hdfs.BucketWriter: Creating /pilot-pjt/collect/car-batch-log/20160313/car-batch-log.1457846109475.log.tmp
2016-03-13 14:15:17,392 INFO org.apache.flume.source.SpoolDirectorySource: The channel is full, and cannot write data now. The source will try again after 500 milliseco
nds
2016-03-13 14:15:20,433 WARN org.apache.flume.client.avro.ReliableSpoolingFileEventReader: Last read was never committed - resetting mark position.
2016-03-13 14:15:28,464 INFO org.apache.flume.channel.file.EventQueueBackingStoreFile: Updating checkpoint for /home/pilot-pjt/flume-ng/cl/checkpoint/checkpoint, elements
to sync = 10000
2016-03-13 14:15:28,575 INFO org.apache.flume.channel.file.EventQueueBackingStoreFile: Updating checkpoint metadata: logWriteOrderID: 1457846156084, queueSize: 9881, qu
eueHead: 6299
2016-03-13 14:15:28,575 INFO org.apache.flume.channel.file.Log: Updated checkpoint for file: /home/pilot-pjt/flume-ng/cl/data/log-14 position: 2577477 logWriteOrderID:
1457846156084
2016-03-13 14:15:15:
2016-03-13 14:15:58,617 INFO org.apache.flume.channel.file.EventQueueBackingStoreFile: Updating checkpoint metadata: logWriteOrderID: 1457846319142, queueSize: 8968, qu
2016-03-13 14:15:58,617 INFO org.apache.flume.channel.file.EventQueueBackingStoreFile: Updating checkpoint metadata: logWriteOrderID: 1457846319142, queueSize: 8968, qu
2016-03-13 14:15:58,617 INFO org.apache.flume.channel.file.EventQueueBackingStoreFile: Updating checkpoint metadata: logWriteOrderID: 1457846319142, queueSize: 8968, qu
```

그림 4.12 플럼의 실행 로그 - SmartCarInfo_Agent 동작 확인

그림 4.13에서처럼 "···BucketWriter: Closing /pilot-pjt/···", "···BucketWriter: Renaming /pilot-pjt/···", 그리고 마지막으로 "···Writer callback called."라는 메시지가 보이면 모든 HDFS 적재가 성공적으로 끝난 것이다.

```
to sync = 10000
2016-03-13 15:08:11,655 INFO org.apache.flume.channel.file
Head: 392
2016-03-13 15:08:11,672 INFO org.apache.flume.channel.file
1457854783104
2016-03-13 15:08:20,628 INFO org.apache.flume.sink.hdfs
2016-03-13 15:08:20,668 INFO org.apache.flume.sink.hdfs
2016-03-13 15:08:20,677 INFO org.apache.flume.sink.hdfs
```

그림 4.13 플럼 실행 로그 - HDFS 적재 완료 확인

₩ 4.6 적재 파일럿 실행 4단계 - 적재 기능 테스트 ☑ HDFS 명령어 확인 (1/2)

\$ hdfs dfs -ls -R /pilot-pjt/collect/car-batch-log/

그림 4.14 HDFS 명령어로 적재 중인 스마트카의 로그 파일 확인



☑ HDFS 명령어 확인 (2/2)

- \$ hdfs dfs -cat /pilot-pjt/collect/car-batch-log/wrk_date=20160529/car-batch-log.1464491838559.log
- \$ hdfs dfs -tail /pilot-pjt/collect/car-batch-log/wrk_date=20160529/car-batch-log.1464491838559.log

```
- - X
root@server02:~
[root@server02 ~] # hdfs dfs -tail /pilot-pjt/collect/car-batch-log/wrk date=20160529/car-batch-log.1464491838559.log
20160101235852, V0099, 82, 86, 6, 83, 1, 1, 1, 1, A, A, 87, 20160529
20160101235856, V0099, 86, 91, 6, 90, 1, 1, 1, 1, A, A, 96, 20160529
20160101235900, V0099, 91, 83, 6, 89, 1, 1, 1, 1, B, A, 95, 20160529
20160101235904, V0099, 85, 93, 6, 84, 1, 1, 1, 1, A, A, 97, 20160529
20160101235908, V0099, 92, 89, 6, 100, 1, 1, 1, 1, A, A, 70, 20160529
20160101235912, V0099, 80, 82, 6, 82, 1, 1, 1, 1, A, A, 94, 20160529
20160101235916, V0099, 91, 82, 6, 88, 1, 1, 1, 1, A, A, 87, 20160529
20160101235920, V0099, 93, 93, 6, 80, 1, 1, 1, 1, A, A, 98
20160101235924, V0099, 84, 81, 6, 73, 1, 1, 1, 1, A, B, 85, 20160529
20160101235928, V0099, 89, 97, 6, 100, 1, 1, 1, 1, A, B, 84, 20160529
20160101235932, V0099, 78, 86, 6, 100, 1, 1, 1, 1, A, A, 97, 20160529
20160101235936, V0099, 79, 71, 6, 99, 1, 1, 1, 1, A, A, 94, 20160529
20160101235940, V0099, 86, 84, 6, 100, 1, 1, 1, 1, A, A, 9
20160101235944, V0099, 83, 95, 6, 86, 1, 1, 1, 1, A, A, 100, 20160529
20160101235948, V0099, 74, 85, 6, 75, 1, 1, 1, 1, A, A, 51, 20160529
20160101235952, V0099, 98, 86, 6, 89, 1, 1, 1, 1, A, A, 83, 20160529
20160101235956,V0099,87,90,6,93,1,1,1,1,A,B,89,20160529
20160101240000, V0099, 96, 75, 6, 81, 1, 1, 1, 1, A, A, 94, 20160529
[root@server02 ~]#
```

그림 4.15 HDFS 명령어로 적재 중인 스마트카의 로그 파일 내용 확인

```
$ ps -ef | grep smartcar.log
$ kill -9 [pid]
```

₩☆ 4.6 적재 파일럿 실행 4단계 - 적재 기능 테스트

실습