|  |
| --- |
| FCIS |
| Hadoop |
| Operation, Utilize |

|  |
| --- |
| SeokRae Kim  2017-8-22 |

내용

[I. 맵리듀스 튜닝 2](#_Toc491202995)

[1. 셔플 튜닝 2](#_Toc491202996)

[2. 콤바이너 클래스 적용 2](#_Toc491202997)

[3. 맵 출력 데이터 압축 2](#_Toc491202998)

[4. DFS 블록 크기 수정 2](#_Toc491202999)

[5. JVM 재사용 2](#_Toc491203000)

[6. 투기적인 잡 실행 2](#_Toc491203001)

[7. 압축 코덱 선택 2](#_Toc491203002)

[II. 하둡 운영 2](#_Toc491203003)

[1. 클러스터 계획 2](#_Toc491203004)

[2. 네임노드 HA 구성 2](#_Toc491203005)

[3. 파일 시스템 상태 확인 2](#_Toc491203006)

[4. 밸런서 2](#_Toc491203007)

[5. HDFS 어드민 명령어 사용 2](#_Toc491203008)

[6. 데이터 저장 공간 관리 3](#_Toc491203009)

[7. 데이터노드 제거 3](#_Toc491203010)

[8. 데이터노드 추가 3](#_Toc491203011)

[9. 네임노드 장애 복구 3](#_Toc491203012)

[10. 데이터노드 장애 복구 3](#_Toc491203013)

[11. 셸 스크립트를 이용한 배포 관리 3](#_Toc491203014)

[12. 리눅스 서버 설정 3](#_Toc491203015)

[III. 하둡 부가기능 활용 3](#_Toc491203016)

1. 맵리듀스 튜닝
   1. 셔플 튜닝
      1. 셔플이란
      2. 셔플 속성 수정
   2. 콤바이너 클래스 적용
   3. 맵 출력 데이터 압축
      1. Gzip 적용
      2. 스내피 설치
      3. 스내피 적용
   4. DFS 블록 크기 수정
   5. JVM 재사용
   6. 투기적인 잡 실행
   7. 압축 코덱 선택
2. 하둡 운영
   1. 클러스터 계획
      1. 하드웨어 선택
      2. 클러스터 규모 산정
   2. 네임노드 HA 구성
      1. 아바타노드의 등장 배경
      2. 아바타노드의 작동 방식
   3. 파일 시스템 상태 확인
   4. 밸런서
   5. HDFS 어드민 명령어 사용
      1. report
      2. safemode
      3. saveNamespace
      4. 파일 저장 개수 설정
      5. 파일 저장 용량 설정
   6. 데이터 저장 공간 관리
   7. 데이터노드 제거
   8. 데이터노드 추가
   9. 네임노드 장애 복구
      1. 네임노드와 보조네임노드 데이터구조
      2. 보조네임노드를 이용한 장애 복구
      3. NFS를 이용한 장애 복구
   10. 데이터노드 장애 복구
   11. 셸 스크립트를 이용한 배포 관리
       1. 작업 대상 파일 설정
       2. SSH 인증키 복사
       3. 슬레이브 서버로 파일 복사
       4. 슬레이브 서버에서 셸 명령어 실행
   12. 리눅스 서버 설정
       1. 파일 오픈 개수 설정
       2. 파일 시스템 종류 설정
       3. 메모리 설정
3. 하둡 부가기능 활용