

아파치 스파크
데이터 레이크 풍경

데이터 레이크의 개념
구글 파일 시스템
하둡 분산 파일 시스템

데이터 레이크의 개념
맵 리듀스
HDFS
하둡

데이터 레이크의 개념
1. 테라데이터
2. 익사데이터

데이터 레이크의 개념
1. 수직 확장성
2. 높은 자본 비용
1. 수평 확장성
2. 낮은 자본 비용

데이터 레이크?

데이터 레이크의 개념
데이터 레이크

데이터 레이크의 개념
데이터 과학자
나머지 인터페이스

m

t

s

u

e

파일 다운로드

YARN | 쿠버네티스 | 메소스

g s

n

n

o

l

C JDBC/ODBC

저장소

검색

HDFS 아마존 아저씨 구글

S3 Blob 클라우드

보안 | 워크플로우 | 메타데이터 | 데이터 수명주기 | 모니터링 | 운영