

제플린과 오픈소스 생태계, 그리고 실리콘벨리

이문수

moon@nflabs.com / zeppelinx.io

ZeppelinX (aka NFLabs)



Apache Zeppelin,

어디쯤 와있을까?

2016.10.12



Apache Project Github 레파지토리 별 갯수

*Data 관련 프로젝트만

2016.10.12



1	Spark	10,310
2	PredictionIO	9,665
3	Kafka	3,898
4	Airflow	3,626
5	Storm	3,586
6	Cassandra	3,030
7	Thrift	3,026
8	CouchDB	2,718
9	Mesos	2,622
10	Hadoop	2,401
11	Zookeeper	2,118
12	Zeppelin	2,027

13	Flink	1,637
14	Hive	1,069
15	mahout	922
16	Hbase	896
17	Flume	745
18	Drill	697
19	Arrow	666
20	Geode	629
21	Kylin	620
22	Ignite	583
23	Lucene-Solr	470
24	Accumulo	252



Apache Project Github 레파지토리 컨트리뷰터

*이전 페이지에 있던 프로젝트들 중

2016.10.12



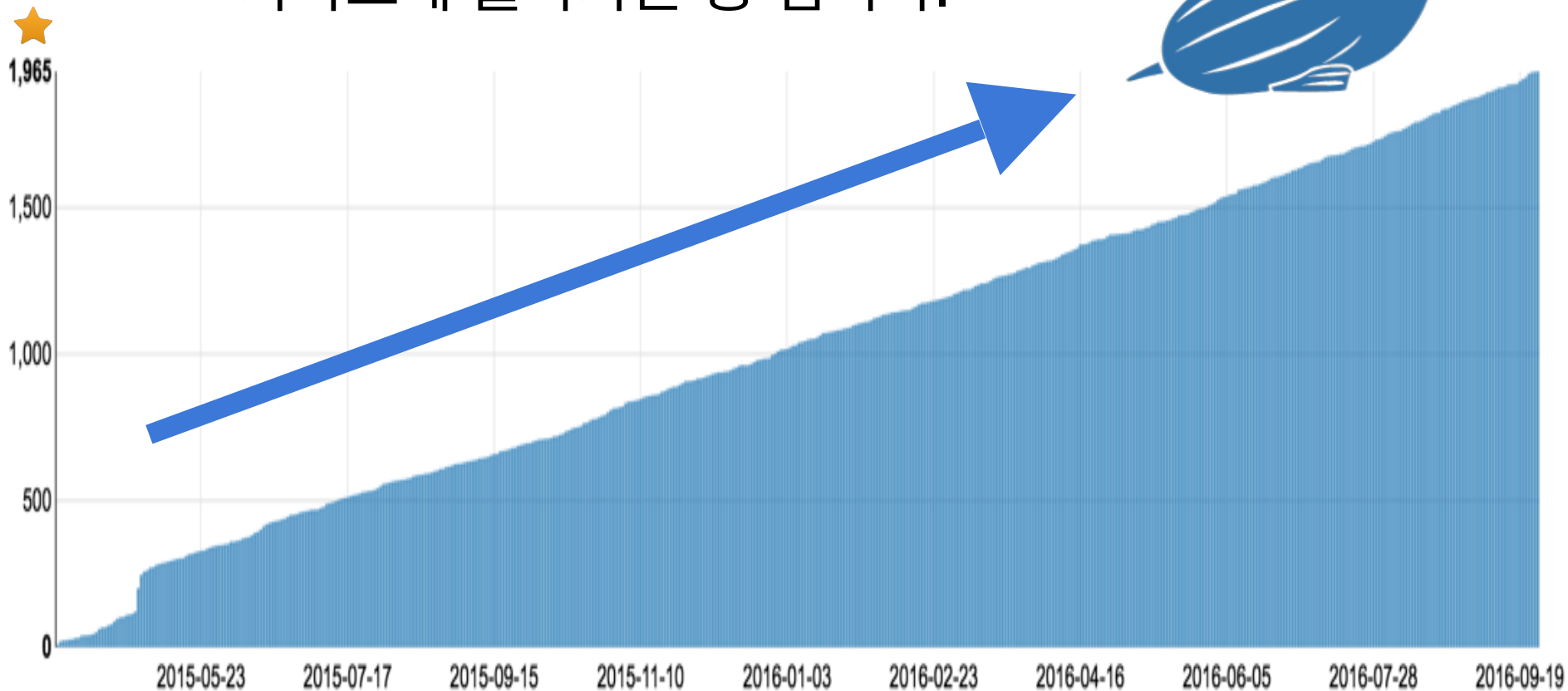
1	Spark	977
2	Kafka	232
3	Storm	230
4	Flink	230
5	Mesos	211
6	Airflow	192
7	Cassandra	187
8	Zeppelin	153
9	Thrift	126
10	CouchDB	106
11	PredictionIO	99
12	Hive	92

13	Hbase	91
14	Hadoop	82
15	Ignite	76
16	Dril	67
17	Accumulo	65
18	Kylin	51
19	Lucene-Solr	50
20	Geode	46
21	Arrow	22
22	Flume	26
23	mahout	23
24	Zookeeper	9



2015, 2016 The best opensource bigdata software!

Apache Zeppelin, 가파르게 올라가는 중 입니다.





Apache Zeppelin,
누가 기여하고 있나?



aka. NFLabs



150 Contributors world wide



Apache Zeppelin,
누가 쓰고 있을까?

★ Between

세계 최초 Zeppelin 유저



전체 약 4000 명 직원 중 1000명이 Zeppelin
사용. Tableau 대체중

The Netflix logo, the word "NETFLIX" in white capital letters on a red rectangular background, is positioned to the left of the text.

NETFLIX

내부 개발자 portal 에 Zeppelin 제공
Recommendation 팀에서 사용



1.5PB / sec 처리하는 인프라 관리 하기 위해
사용



서울시립대학교
UNIVERSITY OF SEOUL

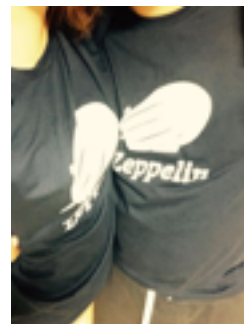
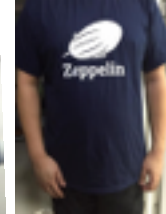
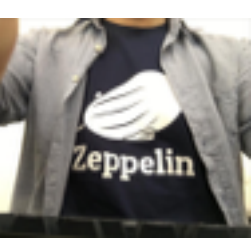


AT&T



JPL

National Aeronautics and Space
Administration
Jet Propulsion Laboratory
California Institute of Technology



*사진
제플린과 친구들과
빅데이터 비행단



Apache Zeppelin,
누가 돈벌고 있을까?



EMR 에서 Zeppelin 서비스



Azure HD Insight 에서 Zeppelin 서비스



Google cloud data proc 에서 Zeppelin 스크립트 제공



Zeppelin 서비스로 제공



HDP 패키지에 Zeppelin 탑재



Juju 이용해 Zeppelin 을 on-prem/cloud 에 배포가능



분석 인터페이스로 Zeppelin 제공



Datascience workbench 에서 분석 인터페이스로 Zeppelin 제공



Apache Zeppelin, 생태계

Contributors



Users



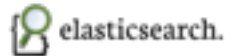
서울시립대학교
UNIVERSITY OF SEOUL



150 Contributors world wide



Service provider



Technology Integration



Zeppelin 이 가져다준 것

회사이름

NFlabs



zeppelin X

위치

한국



실리콘 벨리 & 한국

투자

Seed (한국)



Series A (미국)

컨퍼런스/
밋업

서울



서울, 샌프란시스코, 산호세, 오스틴, 시애틀, 밴쿠버, 시드니, 벨번, 부다페스트, 베를린, 암스테르담, 더블린, 상하이, 싱가포르, 런던, 도쿄, 세비아....



무엇이 지금까지 여정을 가능하게 했을까요?



Zeppelin 을 Spark 이 한창 떠오르는 시점에, 사람들이 필요로 하는 기능을 넣어서 공개할 수 있었던 이유

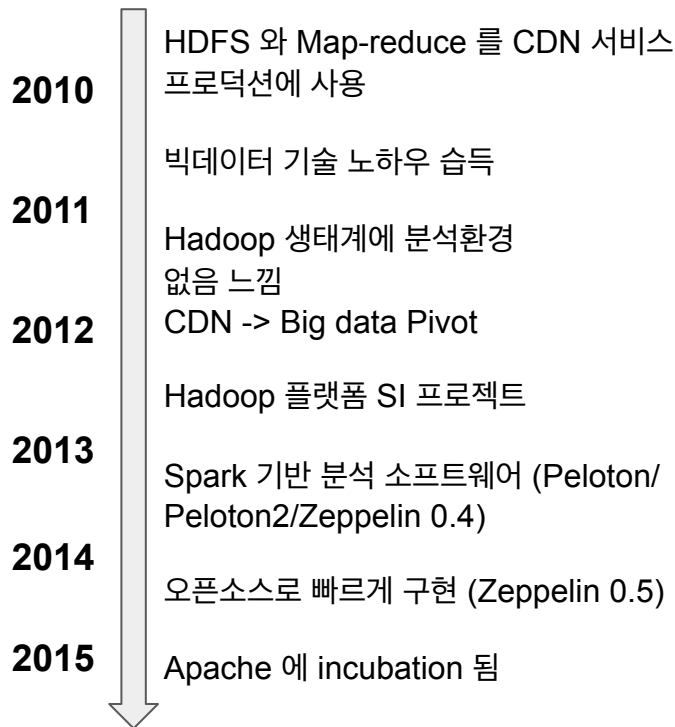


잡질질해서, 어슬렁거려서,
책보고 공부해서, 똑똑해서

실제 사용해보고

여러번 시행착오를 겪었기 때문

(Peloton, Peloton2, Zeppelin-0.4)



Great team
Great culture



첫 번째 오픈소스프로젝트 임에도 성공적인 오픈소스로 키울 수 있었던 이유 이미 오픈소스 프로젝트와 같은 방식으로 일하고 있었기 때문



원래부터 가지고 있던 팀 / 문화

오픈소스 프로젝트

다양성

- 4개국에서 온 개발자들
- 서로다른 경험과 문화



- 전세계의 서로다른 개발 경험, 서로다른 문화를 가진 개발자

문화

- 원격 근무, Flexible time
- 모든 커뮤니케이션은 영어로



- 전세계에 분산된 개발자들과 원격으로 협업

구글, 아마존, 마이크로소프트등 전세계 기업이 기여하고 서비스에 이용하는 이유

비즈니스를 독점하려고 하지 않았기 때문

Support
Consulting
Training
Cloud service
Enterprise Packaging
OEM
System integration
Opensource development

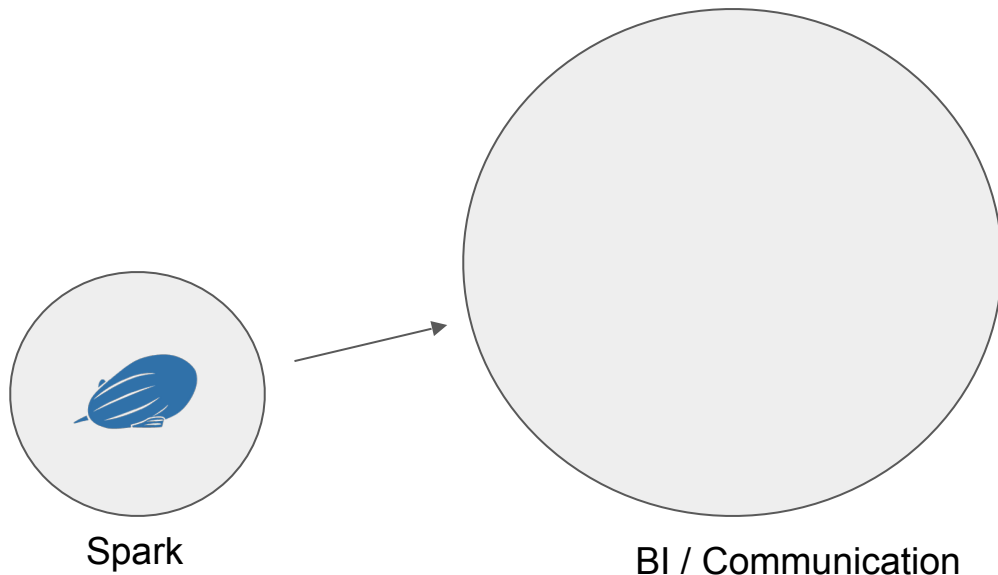


Sharing and Collaboration
service on top of Zeppelin



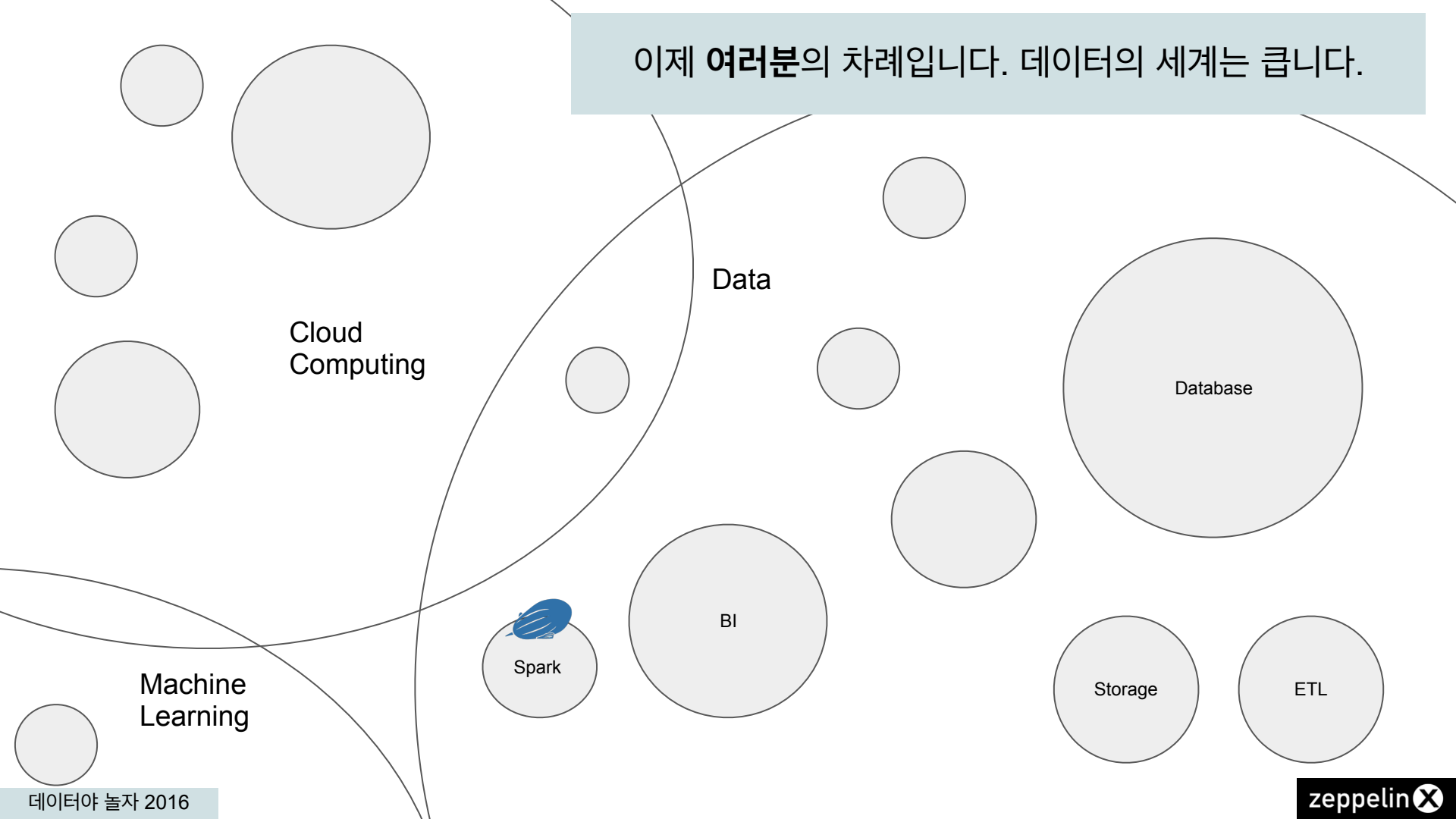


제플린은 어디로 갈까요?



더 큰 세상으로 갑니다.

이제 여러분의 차례입니다. 데이터의 세계는 큼니다.



Cloud Computing

Data

Database

BI

Spark

Storage

ETL

Machine Learning