Snowflake 소개

Louis Lee (이장훈)
Korea Tech Lead | 이사
Sales Engineering, APJ
louis.lee@snowflake.com

© 2021 Snowflake Computing Inc. All Rights Reserved

3년간 스텔스 모드 후, 6년간 GA

120개 이상의 데이터 베이스
특허를 가진 업계의
베테랑들이 2012년 회사 설립

Series G Funding조달후
2020년9월 NYSE에 상장
Largest US software IPO in history

2014년 첫 고객 판매
2015년 GA

~2500명의 직원과
4500이상 고객

재미있는 팩트

Snowflake에서
하루에 처리되는 쿼리:
>9억 쿼리

Snowflake에서
관리하는 데이터 볼륨:
>1EB

하루에 스켄되는
데이터 볼륨:
>60PB

전세계 **4500이상** 고객
어카운드와 **80만**
사용자

© 2021 Snowflake Inc. All Rights Reserved

데이터 사일로가 가치 실현을 방해하고 있음

87%

의사 결정자는 외부 데이터 사용 능력을 확장하길 원합니다[1]

[1] "The Insights Professional's Guide To External Data Sourcing" Forrester, 2020

© 2021 Snowflake Inc. All Rights Reserved

데이터 클라우드는 글로벌 네트워크입니다

기업 및 데이터 공급자를 비즈니스에 가장 적합한 데이터에 연결하는
하나의 글로벌 통합 시스템

© 2021 Snowflake Inc. All Rights Reserved

데이터 클라우드의 이점

ACCESS

GOVERNANCE

ACTION

© 2021 Snowflake Inc. All Rights Reserved

PROVEN BY THOUSANDS OF CUSTOMERS

© 2021 Snowflake Inc. All Rights Reserved

THE DATA CLOUD TODAY

THOUSANDS OF ORGANIZATIONS ARE SHARING DATA WITH THEIR ECOSYSTEM

A software company shares terabytes of data with hundreds of customers

COVID-19 data is available live on Snowflake Data Marketplace from a US State, and other organizations

Today's financial data is accessible immediately without data pipelines

Thousands of companies share data with suppliers, partners, or other business units

*Visualization based on actual Data Cloud sharing activity as of April 2021

© 2021 Snowflake Inc. All Rights Reserved

데이터 클라우드의 요소

PLATFORM

+

CONTENT

© 2021 Snowflake Inc. All Rights Reserved

SNOWFLAKE 클라우드 데이터 플랫폼

DATA SOURCES

OLTP DATABASES

ENTERPRISE APPLICATIONS

THIRD-PARTY

WEB/LOG DATA

IoT

Google Cloud aws Azure

DATA CONSUMERS

DATA MONETIZATION

OPERATIONAL REPORTING

AD HOC ANALYSIS

REAL-TIME ANALYTICS

© 2021 Snowflake Inc. All Rights Reserved

GLOBAL SNOWFLAKE

© 2021 Snowflake Inc. All Rights Reserved

SNOWFLAKE ARCHITECTURE

스케일 아웃 서비스
멀티 클러스터 컴퓨팅 계층
중앙 저장소
클라우드 애그노스틱 계층

Google Cloud aws Azure

© 2021 Snowflake Inc. All Rights Reserved

Let’s dive in…

Informatica
talend
Stitch
Fivetran
MATILLION
ATTUNITY
striim
Segment
StreamSets
Amplitude
M
SIGMA
ThoughtSpot
SISENSE
tableau
Power BI
Qlik
MODE
looker
MicroStrategy
ETL/ELT
XS
Sales
M
정형 데이터
반정형 데이터
CSV
XML
JSON
데이터 레이크
Spark Streaming
스트리밍 데이터
kafka
Snowpipe
Tasks, Streams
aws
S
S
Data Science
alteryx
zepl
data iku
Amazon SageMaker
DataRobot
python
jupyter
Spark
L
Global Services
Metadata Repository
Logical Model
Security
Query Planning & Optimisation
Transactional Control

1. 워크로드 격리를 위한 Scale Across

2. 더 크고 복잡한 워크로드를 위한 Scale Up

SCALE UP - 10억건 데이터 로드 성능

서버 스펙을 두 배로 하면 실행 시간은 절반

서버 청구 단위는 **초단위**

그렇기 때문에 **같은 비용**으로 **8배 빠르게** 처리할 수 있습니다.

3. 더 높은 동시성을 위한 Scale Out

Informatica
talend
Stitch
Fivetran
MATILLION
ATTUNITY
striim
Segment
StreamSets
Amplitude
SIGMA
ThoughtSpot
SISENSE
tableau
Power BI
Qlik
MODE
looker
MicroStrategy
ETL/ELT
XS
정형 데이터
반정형 데이터
데이터 레이크
스트리밍 데이터
kafka
Snowpipe
Tasks, Streams
Sales
M
M
Multi-cluster
aws
Clone
XSL
DevOps
Data Protection,
Replication & Time Travel
XL
Data
Science
L
Finance/DBAs
Global Services
Metadata Repository
Logical Model
Security
Query Planning & Optimisation
Transactional Control
alteryx
zepl
data
iku
Amazon
SageMaker
DataRobot
python
jupyter
Spark

Informatica
talend
Stitch
Fivetran
MATILLION
ATTUNITY
striim
Segment
StreamSets
Amplitude
외부
M
SIGMA
ThoughtSpot
SISENSE
tableau
Power BI
Qlik
MODE
looker
MicroStrategy
ETL/ELT
XS
정형 데이터
반정형 데이터
데이터 레이크
스트리밍 데이터
kafka
Snowpipe
Tasks, Streams
aws
데이터 공유,
데이터 Exchange
Sales
M
M
Multi-cluster
Clone
XL
DevOps
Data Protection,
Replication & Time Travel
L
Finance/DBAs
XL
Data
Science
alteryx
zepl
data
iku
Amazon
SageMaker
DataRobot
python
jupyter
Spark
Global Services
Metadata Repository
Logical Model
Security
Query Planning & Optimisation
Transactional Control

SNOWFLAKE DATA MARKETPLACE

LL au_demo12
Louis Lee

- Worksheets
- Dashboards
- Discover Data
- Admin

Search Worksheets

Classic Console

PREVIEW APP

< Discover

Snowflake Data Marketplace

Listings | Requests

CATEGORIES

- All
- Business
- Demographics
- Financial
- Government
- Health
- Marketing
- Security
- Transportation
- Weather

Drop Technologies Inc.

Verified Transaction Data

Access Millennial Centric 1st Party Transaction Data. Listing Powered by Highland Math

Daily | Personalized

Highland Math

Covid19 Impact and Discourse

Analyze Covid19 metrics and resulting societal effects

Daily

Starschema

COVID-19 Epidemiological Data

Daily unpivoted, analytics-ready data on COVID-19.

Daily

Weather Source

Premier Weather Data for Analytics

A curated continuum of hyper-local past, present, forecast, and climatology weather data.

Daily | Personalized

Auryc

Web and App User Journey Analytics Data

A continuously refreshed customer experience analytics data.

Near Real-time | Personalized

ThinkData Works

Covid-19 Canadian Outbreak

A daily refreshed feed of Covid-19 case counts by Province and Regional Health Boundaries

Daily | Personalized

PRIVATE DATA EXCHANGE

비즈니스 이코 시스템

보호된 DATA EXCHANGE

외부 데이터 제공자

© 2021 Snowflake Inc. All Rights Reserved

DATA CLOUD GROWTH

April 2020

April 2021

© 2021 Snowflake Inc. All Rights Reserved

*Visualizations based on actual Data Cloud sharing activity as of April 30, 2020 and April 30, 2021

Demo time...

Citi Bike Station map
Citi Bike System Data | Citi Bik
S3 Management Console
Worksheet - Reset v2 (37/43)
Worksheet - Data Sharing
member.citibikenyc.com/map/
Paused
Apps
SFC
Pers
SFC
Trello
Mind
Intro
Demo
Docs
Snowflake
Services
Personal
Internal
Useful
SE Onboarding_wi...
citi bike
How it Works
Pricing
Station Map
Explore NYC
Help
Log In
Join
Find a station
Enter a station name, street name or address.
Lyndhurst
Bloomfield
Belleville
Secaucus
West New York
Union City
Kearny
East Orange
Harrison
Jersey City
Hillside
Elizabeth
Bayonne
EWR
LGA
Rikers Island
Flushing Meadows-Corona Park
mapbox
© Mapbox © OpenStreetMap Improve this map

Citi Bike Station map | Citi Bike System Data | Citi Bike NYC | S3 Management Console | Worksheet - Reset v2 (37/43) | Worksheet - Data Sharing

citibikenyc.com/system-data

citi bike

How it Works | Pricing | Station Map | Explore NYC | Help | Log In | Join

System Data

Where do Citi Bikers ride? When do they ride? How far do they go? Which stations are most popular? What days of the week are most rides taken on? We've heard all of these questions and more from you, and we're happy to provide the data to help you discover the answers to these questions and more. We invite developers, engineers, statisticians, artists, academics and other interested members of the public to use the data we provide for analysis, development, visualization and whatever else moves you.

This data is provided according to the NYCBS Data Use Policy.

Citi Bike Trip Histories

We publish downloadable files of Citi Bike trip data. The data includes:

- Trip Duration (seconds)
- Start Time and Date
- Stop Time and Date
- Start Station Name
- End Station Name
- Station ID
- Station Lat/Long

S3 Management Console

s3.console.aws.amazo...

Services | Resource Groups | SE-Contributor/alan.eldridge

Amazon S3 > snowflake-corp-citibike-demo-master > V2 > tables > trips

Overview

Type a prefix and press Enter to search. Press ESC to clear.

Upload | + Create folder | Download | Actions

US West (Oregon)

Viewing 1 to 18

Name	Last modified	Size	Storage class
data_0_2_0.csv.gz	Aug 9, 2019 11:43:04 AM GMT+1000	2.5 MB	Standard
data_0_3_0.csv.gz	Aug 9, 2019 11:43:08 AM GMT+1000	7.6 MB	Standard
data_0_4_0.csv.gz	Aug 9, 2019 11:42:55 AM GMT+1000	2.8 MB	Standard
data_0_7_0.csv.gz	Aug 9, 2019 11:42:53 AM GMT+1000	10.9 MB	Standard
	Aug 9, 2019		

Feedback | English (US) | Privacy Policy | Terms of Use

S3 Management Console

s3.console.aws.amazo...

Services | Resource Groups | SE-Contributor/alan.eldridge

Amazon S3 > snowflake-corp-citibike-demo-master > V2 > tables > weather 2019

Overview

Type a prefix and press Enter to search. Press ESC to clear.

Upload | + Create folder | Download | Actions

US West (Oregon)

Viewing 1 to 104

Name	Last modified	Size	Storage class
data_0_0_0.json.gz	Aug 9, 2019 11:44:46 AM GMT+1000	16.2 MB	Standard
data_0_0_1.json.gz	Aug 9, 2019 11:44:49 AM GMT+1000	8.0 MB	Standard
data_0_1_0.json.gz	Aug 9, 2019 11:44:46 AM GMT+1000	16.1 MB	Standard
data_0_1_1.json.gz	Aug 9, 2019 11:44:45 AM GMT+1000	3.7 MB	Standard
	Aug 9, 2019		

Feedback | English (US) | Privacy Policy | Terms of Use

EVER EXPANDING ECOSYSTEM

© 2021 Snowflake Inc. All Rights Reserved

Adobe

Situation / Pain

- Ingest 100+ TB every Friday and perform complex segmentation analysis but not meeting SLA
- Failure led to missed SLAs, Audience Manager customers did not get their reports on time
- Poor customer experience

Value

- Every Friday, spin up 12 x 6XL warehouses (~6000 servers) for ingestion and processing
- Job now completes in less than 2 hours and spin warehouses down when process is done
- Achieving SLAs and happy customers

snowflake

AS OF 4/30/2021 - Q1 FY22

PRODUCT REVENUE [1]

$213.8M

+ 110% YoY Growth

NET REVENUE RETENTION RATE [2]

168%

TOTAL CUSTOMERS [2]

4,532

+ 67% YoY Growth

$1M CUSTOMERS [2]

104

+ 117% YoY Growth
Customers with trailing 12-month product revenue greater than $1M

FORTUNE 500 CUSTOMERS [3]

187

+ 30% YoY Growth

DATA MARKETPLACE LISTINGS [4]

498

Total Listings
31% QoQ Growth

CUSTOMER SATISFACTION

FORRESTER CONSULTING TOTAL ECONOMIC IMPACT™ STUDY [5]

612%

Customer ROI Delivered Over 3 years

NET PROMOTER SCORE (NPS) [6]

71

Most Customers Would Recommend Snowflake to a Friend or Colleague

DRESNER CUSTOMER SATISFACTION SCORE [7]

100%

Of Customers Recommend Snowflake for Fourth Consecutive Year

RISE OF THE DATA CLOUD

1. For the three months ended April 30, 2021. 2. As of April 30, 2021. Please see our Q1 FY22 earnings press release for definitions of net revenue retention rate, total customers, and customers with trailing 12-month product revenue greater than $1 million. 3. As of April 30, 2021. Based on the 2020 Fortune 500 list. Our Fortune 500 customer count is subject to adjustments for annual updates to the Fortune 500 list by Fortune, as well as acquisitions, consolidations, spin-offs, and other market activity with respect to such customers. 4. As of April 30, 2021. Each live dataset, package of datasets, or data service published by a data provider as a single product offering on the Snowflake Data Marketplace is counted as a unique listing. A listing may be available in one or more regions where the Snowflake Data Marketplace is available 5. Forrester Consulting Total Economic Impact™ Study Commissioned By Snowflake, August 2020. 6. As of May 2020. 7. Dresner Advisory Services: 2021 Analytical Data Infrastructure Market Study, January 2021.

snowflake®

TRANSFORMING YOUR DATA STRATEGY WITH SNOWFLAKE