About me

- Bùi Xuân Yên
- Lead Software Engineer and Head of Technology Research at Adbrix Korea
- Researcher at ANDA Lab (Advanced Network Design and Analysis) Soongsil University, Seoul, Korea
- Research interest: Low power wireless network protocol for control system, Network Design, Cloud-centric application structure, System architecture, Mobile marketing

Agenda

01.
What is Growth Marketer?

02.
Which Tool Do Growth Marketers Need?

03.
Mobile Marketing Big-data Architecture:
Top-down Design Approach Guideline

04.
Q&A

01.
What is Growth Marketer?

What Growth Marketers Do

User Acquisition Campaign → App Install → App Open → User Engagement → Revenue

Re-Engagement Campaign → App Open

Maximize ROI

How to make business decision?

02.

Which Tool Do Growth Marketers Need?
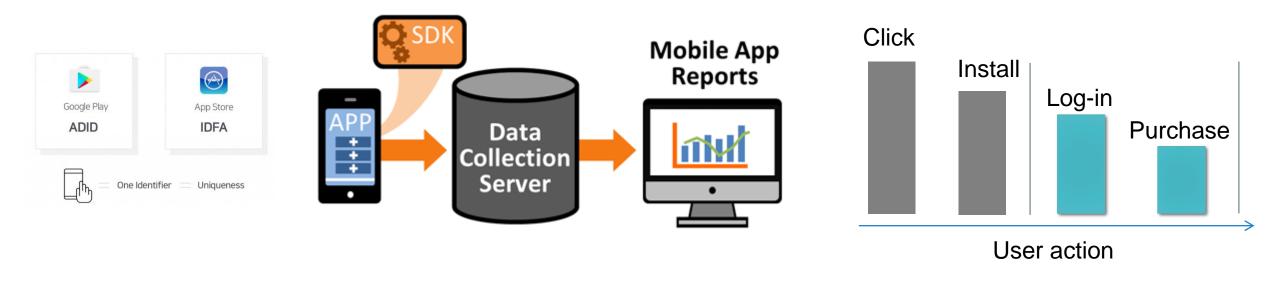
Mobile analytics: Data mining

	Reporting	Analytics
Purpose	Shows **what** is happening	Explains **why** it is happening
Tasks	Organizing Formatting Summarizing	Questioning Interpreting Exploring
Results	Results are **pushed** to users for review	Users **pull** results to answer questions
Value	Translate **data** into **information**	Offers **recommendations** to drive **action**

https://www.pardot.com/training

03.
Mobile Marketing Big-data Architecture
Top-down Design Approach Guideline

Mobile Marketing Big-data Architecture

Different problems needs different solutions and different tools!!! Welcome to BIG DATA technologies world!!!
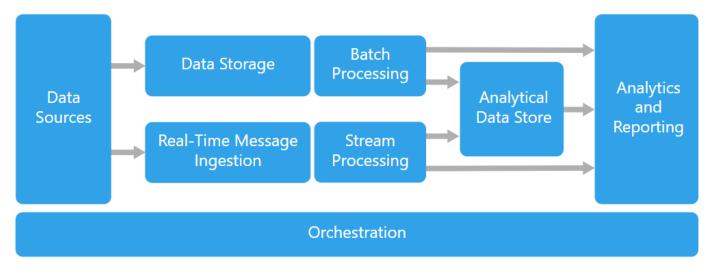
Mobile Marketing Big-data Architecture

Big data architecture style: A big data architecture is designed to handle the ingestion, processing, and analysis of data that is too large or complex for traditional database systems.

https://docs.microsoft.com/en-us/azure/architecture/
https://aws.amazon.com/architecture
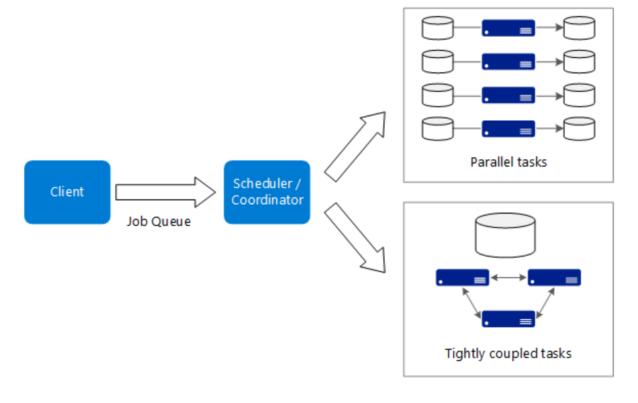
Mobile Marketing Big-data Architecture

Big compute architecture style: The term *big compute* describes large-scale workloads that require a large number of cores, often numbering in the hundreds or thousands.

Data Architecture

Understand your data usage: How data is processed and stored

Traditional RDMS workloads: These workloads include online transaction processing (OLTP) and online analytical processing (OLAP). Data in OLTP systems is typically relational data with a predefined schema and a set of constraints to maintain referential integrity.
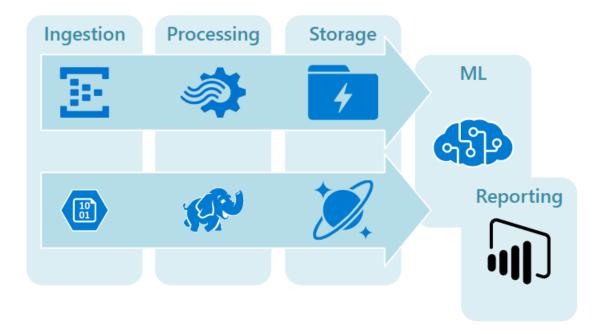
| Property	Transaction processing systems (OLTP)	Analytic systems (OLAP)
Main read pattern	Small number of records per query, fetched by key	Aggregate over large number of records
Main write pattern	Random-access, low-latency writes from user input	Bulk import (ETL) or event stream
Primarily used by	End user/customer, via web application	Internal analyst, for decision support
What data represents	Latest state of data (current point in time)	History of events that happened over time
Dataset size	Gigabytes to terabytes	Terabytes to petabytes

Data Architecture

Big data stack: A big data architecture is designed to handle the ingestion, processing, and analysis of data that is too large or complex for traditional database systems. The data may be processed in batch or in real time. Big data solutions typically involve a large amount of non-relational data, such as key-value data, JSON documents, or time series data.

Data Architecture
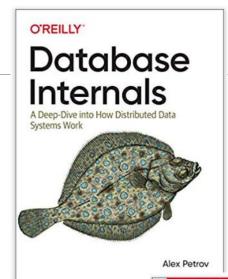
Deep dive into big data principle.

"*Data is at the center of many challenges in system design today. Difficult issues need to be figured out, such as scalability, consistency, reliability, efficiency, and maintainability. In addition, we have an overwhelming variety of tools, including relational databases, NoSQL datastores, stream or batch processors, and message brokers. What are the right choices for your application? How do you make sense of all these buzzwords?*"

Book reference:

Database Internals: A deep-dive into how distributed data systems work

Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems

Google Stack

Data processing

Data Storage (Columnar database, distributed file system, …)

Monitoring

Cluster Management & Coordination

data processing

FlumeJava [CRP+10]	**Tenzing** [CLL+11]	**MillWheel** [ABB+13]	**Pregel** [MAB+10]
parallel programming	*SQL-on-MapReduce*	*stream processing*	*graph processing*

MapReduce [DG08]
parallel batch processing

Percolator [PD10]
incremental processing

PowerDrill [HBB+12]
query UI & columnar store

data storage

MegaStore [BBC+11]
cross-DC ACID database

Spanner [CDE+13]
cross-DC multi-version DB

Dremel [MGL+10]
columnar database

BigTable [CDG+06]
row-consistent multi-dimensional sparse map

monitoring tools

Dapper [SBB+10]
pervasive tracing

GFS/Colossus [GGL03]
distributed block store and file system

$\textbf{CPI}^2$ [ZTH+13]
interference mitigation

coordination & cluster management

Chubby [Bur06]
locking and coordination

Borg [VPK+15] and **Omega** [SKA+13]
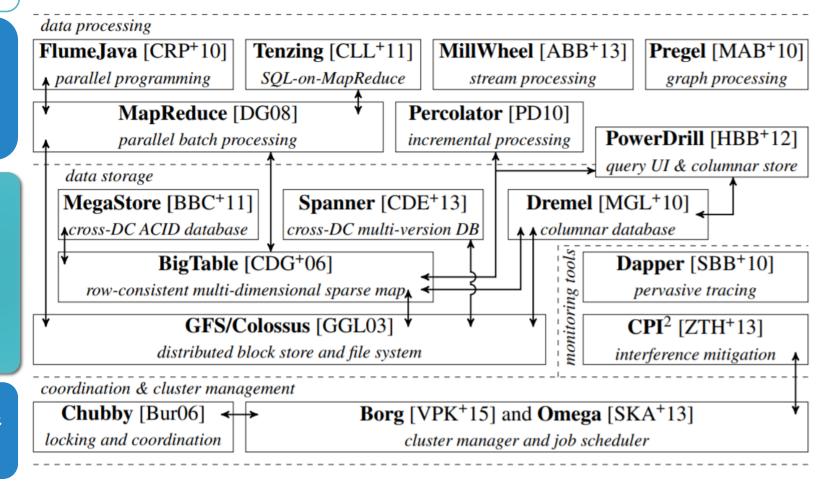*cluster manager and job scheduler*

Figure 1: The Google infrastructure stack. I omit the F1 database [SOE+12] (the back-end of which was superseeded by Spanner), and unknown front-end serving systems. Arrows indicate data exchange and dependencies between systems; simple layering does *not* imply a dependency or relation.

21

Google Big-data Stack vs Open Source Alternatives vs AWS

	Google	Open source	AWS
Locking and Coordination	Chubby	Zookeeper	-
Quản lý cluster và job scheduler	Borg / Omega	Docker + Kubernetes	ECS or EKS
Hệ thống quản lý File phân tán (Distributed File-System)	GFS/Colossus	HDFS	S3
Quản lý dữ liệu (Data Storage)	BigTable	HBase	DynamoDB
Quản lý dữ liệu (Data Storage)	Spanner	CockroachDB	DynamoDB
Columnar Store	Dremel	haeinsa	DynamoDB
Map/Reduce Data processing	MapReduce	Hadoop / YARN	EMR
Graph processing	Pregel	Hama	-
Stream processing	MillWheel	Spark	EMR
Parallel Pipelines	FlumeJava	Apache Crunch	Data pipeline And Glue
SQL-on-MapReduce	Tenzing	Impala, Prestodb	Athena And Redshift

Domain Driven Design: Mobile Marketing Context

Session:

Domain Expert: "We need to track when users use app".

Developer: *"Every time users launch app, there is start session event, when app state is from foreground to background for 1 minute, there is a stop session event"*

In-app event:

Domain Expert: *"Users do action in games, buy items in app, we need to track it."*

Developer: *"Ok, it is also an event. Event has attribute to store meta data"*

Domain Expert   Developer

```
[
  {
    "event_id": "unique_id_1",
    "event_type": "session_start",
    "identity": {
      "user_id": "UUID",
      "token_id": "apns_token",
      "email": "speaker@vmd2019.vn",
      "device_info": {
        "os": "ios",
        "country": "VN",
        "time_zone": "utc+7"
      }
    },
    "event_attribute": {
      "timestamp": "1558593947"
    }
  },
  {
    "event_id": "unique_id2",
    "event_type": "game_action",
    "event_attribute": {
      "timestamp": "1558594947",
      "category": "level_up",
      "level": 100
    }
  },
  {
    "event_id": "unique_id3",
    "event_type": "session_end",
    "event_attribute": {
      "timestamp": "1558595947"
    }
  }
]
```

Simplified Log Processing System

Event ingestion

- Install event
- In-app event
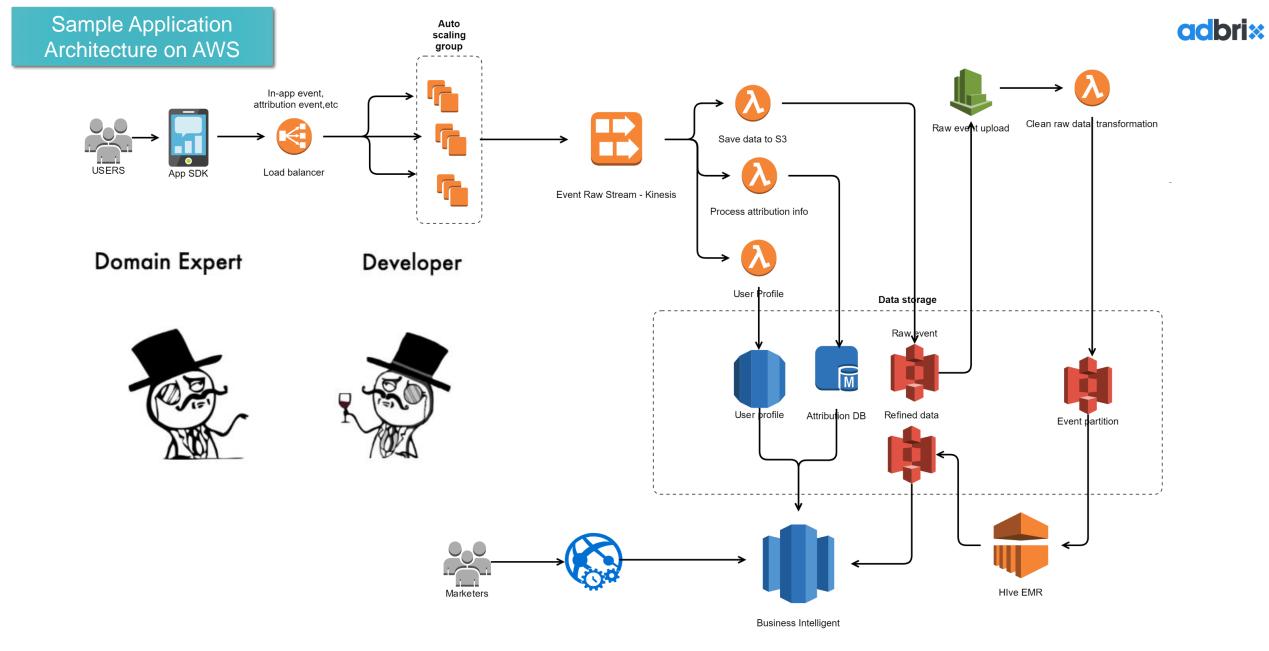- Purchase event
- Custom event

Data processing

- Real-time processing attribution data, user profile
- Run batch data ETL job to save data to S3
- Send post-back event

Business Intelligent

- Ad-hoc query: Active user, DAU, Retention rate, user group data
- Data Visualization and save report to CSV

Sample Application Architecture on AWS

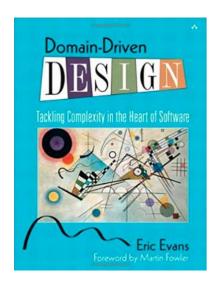# Learn More about Design Pattern
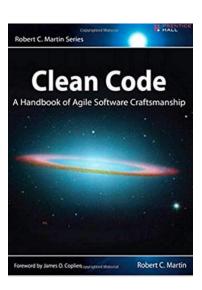
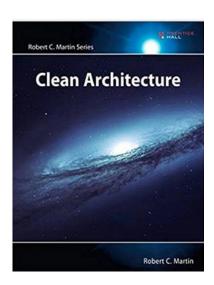*"Good architect is required for good service".*

Book:

Domain-Driven Design: Tackling Complexity in the Heart of Software

Clean Code: A Handbook of Agile Software Craftsmanship

Clean Architecture: A Craftsman's Guide to Software Structure and Design
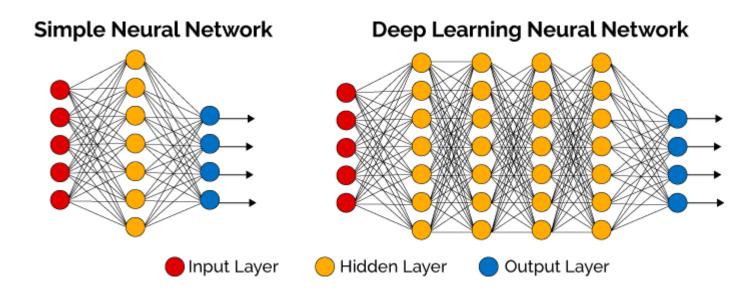
4. Q&A

Future of mobile analytics

Technology changes, more requirements from marketers → We need to support more solutions to make marketer's life easy

Machine learning is booming. By using machine learning (deep learning) , we can create new value from existing data.

Future of Mobile Marketing

Attribution

Click/Impression tracking

Install attribution

Event postback

Improved Fraud detection

Đo lường hiệu năng chiến
dịch quảng cáo
Phản ánh rõ tham số ROI

Analytics

Active user report

Retention rate report

Custom reports

Đánh giá hành vi người
dùng và quản lý chất lượng
ứng dụng

Action

Define targeting audience

Automation, intelligent,
personalized engagement
marketing campaign

Thực hiện chiến lược chăm
sóc khách hang, chuyển hóa
khách hàng tiềm năng

Machine
Learning

What's next?

Design and deploy big data marketing platform is not easy task (more requirements, marketing expert's knowledge requirement, system scalability, reliability, stability) and it is costly

For marketer: Adbrix provides solid tool for all your need. Using it and make your KPI high. For more detail information, please come to visit our camp to talk with our marketing experts.

For developer: Customer expects real value from your product. Don't need to re-invent the wheel. I hope that information in this talk helps you do better product which your customer is waiting.

http://adbrix.io/

www.adbrix.io

console.adbrix.io

abx@igaworks.com