Recommend
火龙果软件

从技术演变的角度看互联网后台架构

腾讯游戏
Tencent Games

About Myself

Recommend 火龙果软件

2008

University of Texas at San Antonio, PhD in Computer Science

2008 ~ 2011

Core Operating System, Microsoft, Redmond

2011 ~ 2012

Tagged/Funzio/Gree, San Francisco

2013 ~ 2015

回国创业，移动游戏

2015 ~ 2017

BCG, Uber
Data platform & Mobile

2017~现在

Tencent IEG
Turing Lab副总监

腾讯游戏 Tencent Games

Turing Lab

System Design？架构设计？Why？

when management/senior team leads/architects threw a copy of the book at junior developers, expecting that they could ***read the book and suddenly "level up"****, they were profoundly disappointed.*

-http://blogs.tedneward.com/post/reclaiming-design-patterns/

设计模式的意义

- 共通的语言和词汇，帮助沟通、编写文档和探索设计方案
- 从较高抽象层面思考，而非具体编程语言，从而让复杂系统变得更容易分析
- 提升和同事讨论分析时的切入层次
- 更容易理解现有系统
- 提供对设计本质的描述方式，而不仅仅是结果，从而引导更正确的决策
- 帮助理解设计思路，从而更有效地进行重构

-"Design Patterns"，GoF, 1994

System Design

Today

Tencent IEG Turing Lab

主题

- 后台架构的演化
- 中间件的发展

Tencent IEG Turing Lab
Tencent IEG Turing Lab
Tencent IEG Turing Lab

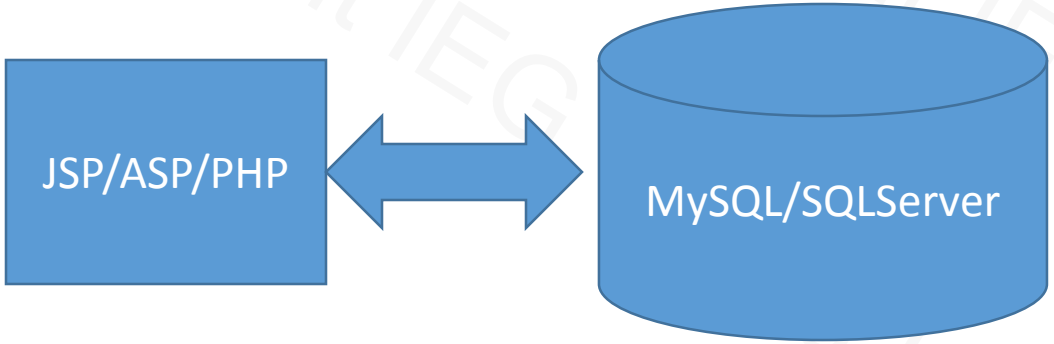
后台架构的演化：2000 ~ 2018

Tencent IEG Turing Lab

2000～2008:网站时代

PHP是最好的语言！Why？

JSP/ASP/PHP
- 业务逻辑和页面显示混写
- 两层结构（现在仍然流行于大量软件外包、业务兼职外包、淘宝网站代码中）
- 高手=懂得MVC，会写JSP/PHP+Servlet(Tomcat)+DB



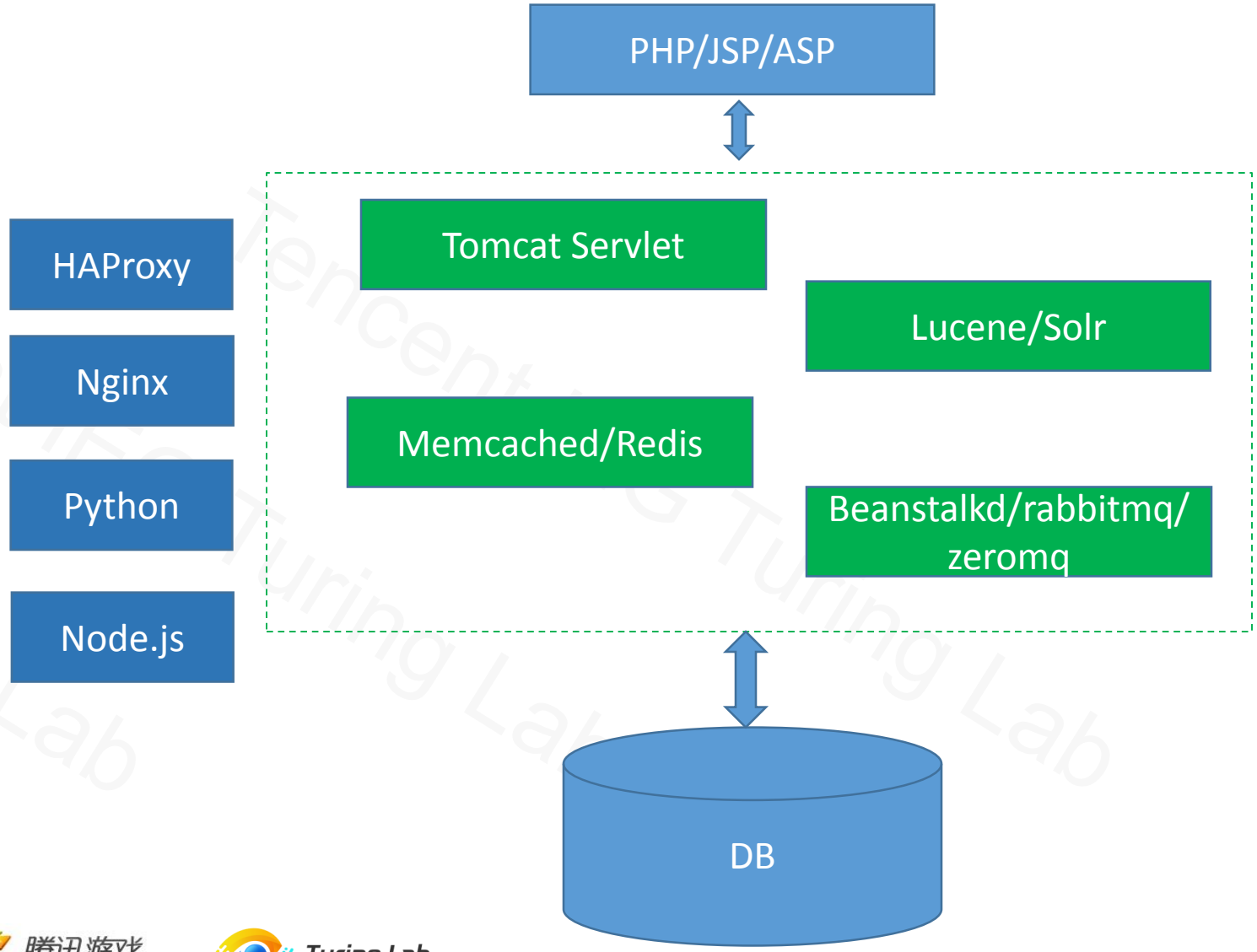

```php
<?php
$servername = "localhost";
$username = "username";
$password = "password";

$conn = new mysqli($servername, $username, $password);
?>
connected to MySQL!
Welcome <?php echo $username ?>
```

2009～2013:中间件时代

社交网站的崛起（Facebook/Twitter）

- 海量数据存储
- 海量数据搜索
- 初期的推荐系统（用户推荐）
- 并发能力、可用性要求大幅度提高
- 业务复杂度增加，开发人员数量大幅增加



Recommend
火龙果软件
Tencent IEG Turing Lab
腾讯游戏 Tencent Games
Turing Lab

2013 ~ 2017:大数据时代

Recommend 火龙果软件

移动互联网

- 大数据
- 数据挖掘
- 精准用户推荐
- 云计算到Docker化
- 微服务到Serverless
- 流计算(Stream Computing)
- DevOps
- 快速迭代/AB测试

客户端

ios | android | web | others

RESTFul Service | API Services

Routing/Dispatching

业务逻辑服务（Login） | 业务逻辑服务（Feeds） | 业务逻辑服务（Favorites）

中台

Transaction | User | recommendation | Map/location | Logging | Monitoring | Fraud

Infrastructure(基础架构)

K8s/Mesos | Hive | Spark | Kafka | DB | Cache | Elasticsearch

腾讯游戏 Tencent Games

Turing Lab

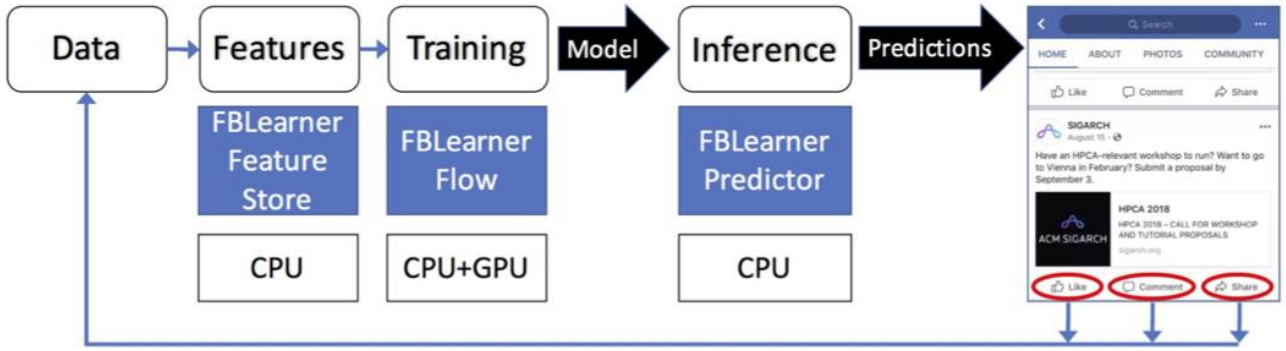
Tencent IEG Turing Lab

2018~??: AI 时代 - Machine Learning Driven

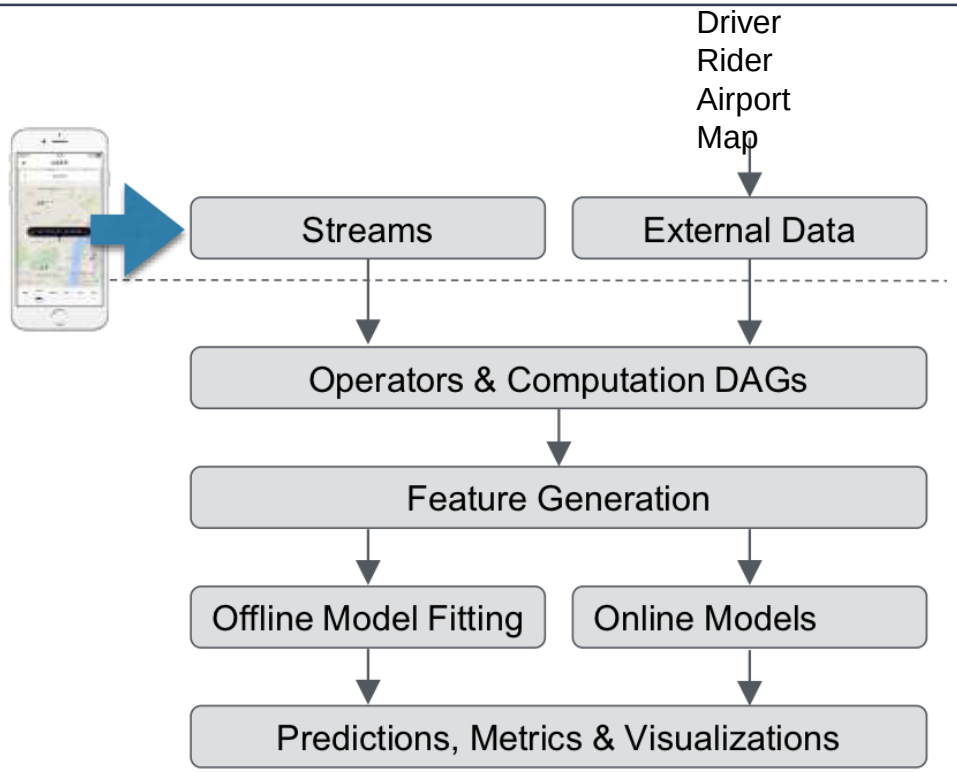
Recommend 火龙果软件

基于（转型）AI/机器学习的新形态公司

- Uber/滴滴
- 饿了吗/美团
- 头条/抖音



核心机器学习平台：FBFlow



核心机器学习平台：Michelangelo

新的挑战

Training System: Unified vs Independent
Deployment: Evaluation/AB Test/On Demand/ML Devops Pipeline/etc.
Serving: Tensorflow Serving and ??

“SysML: The New Frontier of Machine Learning System”

Stanford/CMU/Google/MSFT/Amazon/Intel/MIT/UCB/etc., 2019

腾讯游戏 Tencent Games　Turing Lab

Middleware

Tencent IEG Turing Lab

What is middleware?

Recommend
火龙果软件

Cache

- 快速读取（大量读取，较少写入）
- Redis, Memcached

Message Queue

- 非实时：异步处理信息（数据分析、机器学习）
- 实时：消息分发（聊天、推送）
- Kafka, RabbitMQ, Zeromq

Service Registry

- 服务注册、发现、管理
- Zookeeper

Security&Authentication

- 通常由应用框架自行实现

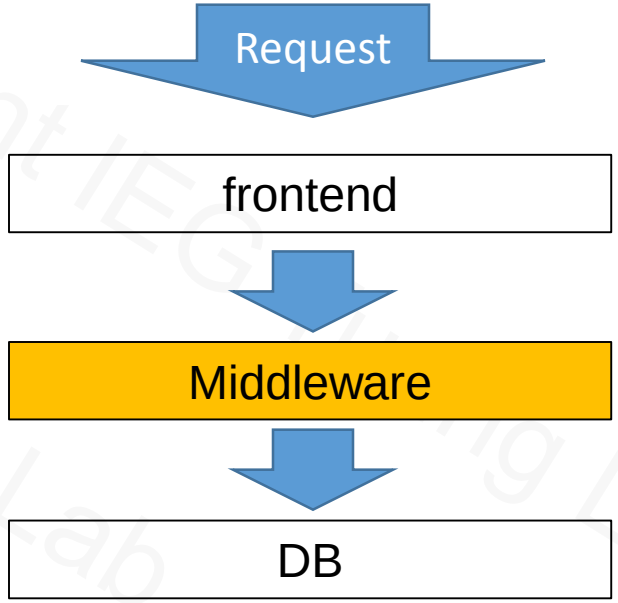
机器学习框架

- 负责机器学习实现
- Tensorflow

RPC Protocol

- 服务接口定义和统一（开发阶段）
- Protocol Buffer, Thrift

从架构设计出发的Middleware



腾讯游戏 Tencent Games
Turing Lab

Middleware v.s. 中台

	Middleware	中台
针对对象	技术/开发	业务/组织
和业务逻辑绑定	No	Yes
使用者	开发人员	产品/管理/PM/市场/非技术人员
能独立实现业务需求	No	Yes
示例	消息队列 Tensorflow Springcloud(整合方案)	用户管理 交易后台管理 推荐平台

Recommend 火龙果软件

腾讯游戏 Tencent Games Turing Lab

Why middleware? ----好友推荐系统的演化

早期代码

Requests

```
rs = conn.query('SELECT * FROM
users
WHERE sex='male'
AND age>20 AND age<40
AND city='Beijing'
ORDER BY clicks DESC,
comments DESC
LIMIT 100')

result = shuffle(rs)
```

Mysql

Sorting: O(nlgn)

O(n)

*仍然存在于大量开源论坛、CMS、新闻聚合及淘宝出售的代码系统中

Tencent IEG Turing Lab

Why middleware? ----好友推荐系统的演化

初步优化代码 – 引入缓存

复杂，过于主观

Requests

```
query_key= generate_query_key(user_info)

arr = redis.get(query_key)

If arr == NULL {
 rs = conn.query('select … limit 10000')
 arr = process_rs(rs)
 redis.put(query_key, arr)
}

result = shuffle(arr).slice(0,10)
```

Mysql

Redis

offline key-value calculation

缺乏灵活性

Tencent IEG Turing Lab

Why middleware? ----好友推荐系统的演化

引入离线排序/搜索引擎

新用户相关信息

Mysql

Batch Process

Requests

```
query_key= generate_query_key(user_info)

arr = redis.get(query_key)

If arr == NULL {
 rs = query_search_engine(user_info, type = 'suggestions')
 arr = process_rs(rs)
 redis.put(query_key, arr)
}

result = shuffle(arr).slice(0,10)
```

Solr/Lucene
（inverted index)

Redis

太慢
不满足实时场景(滴滴/摩拜/饿了么）

Tencent IEG Turing Lab

Why middleware? ----好友推荐系统的演化

Recommend 火龙果软件

大数据时代

系统复杂度大幅增加，架构开始messy

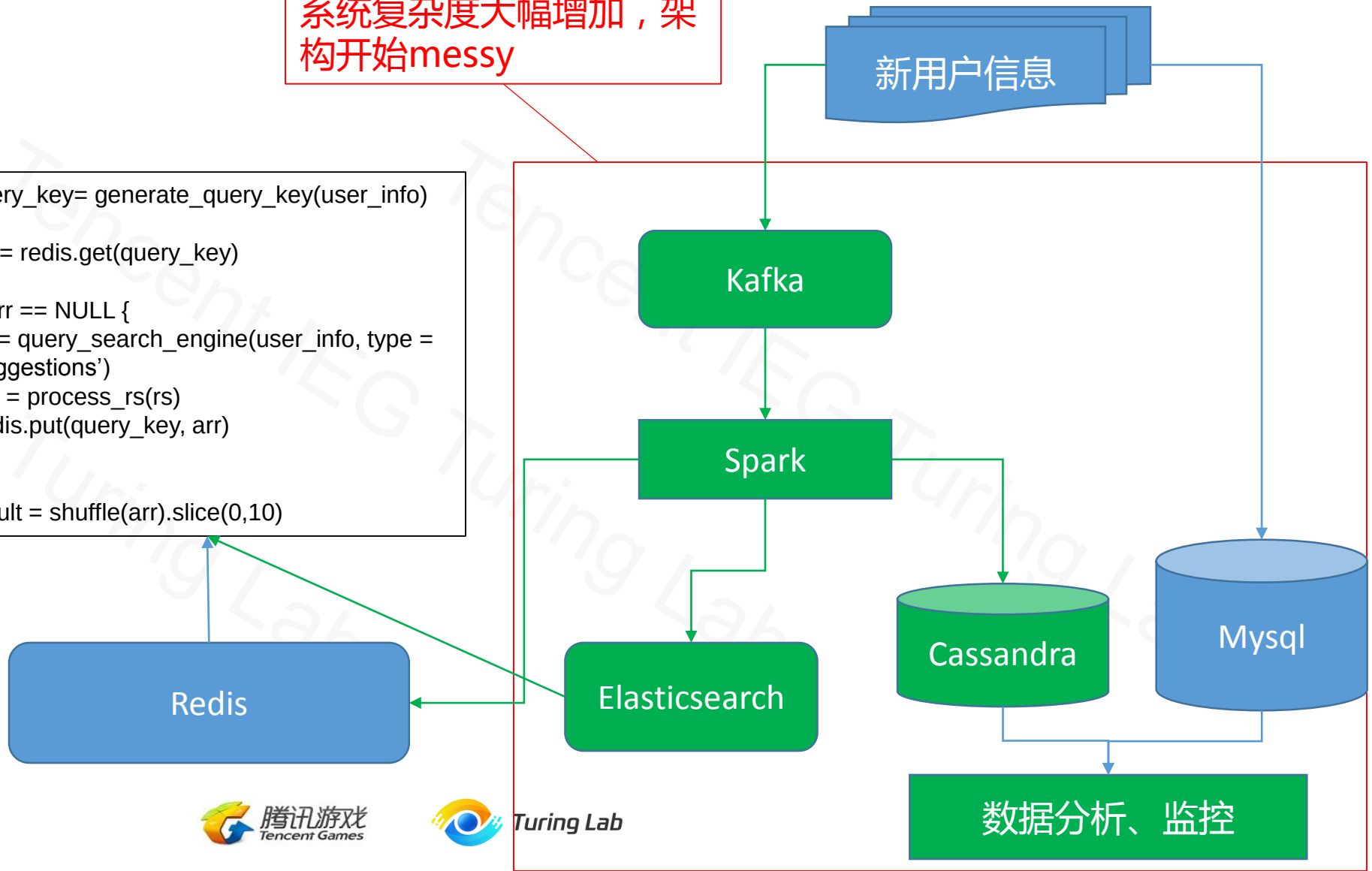
Requests

```
query_key= generate_query_key(user_info)

arr = redis.get(query_key)

If arr == NULL {
 rs = query_search_engine(user_info, type =
'suggestions')
 arr = process_rs(rs)
 redis.put(query_key, arr)
}

result = shuffle(arr).slice(0,10)
```

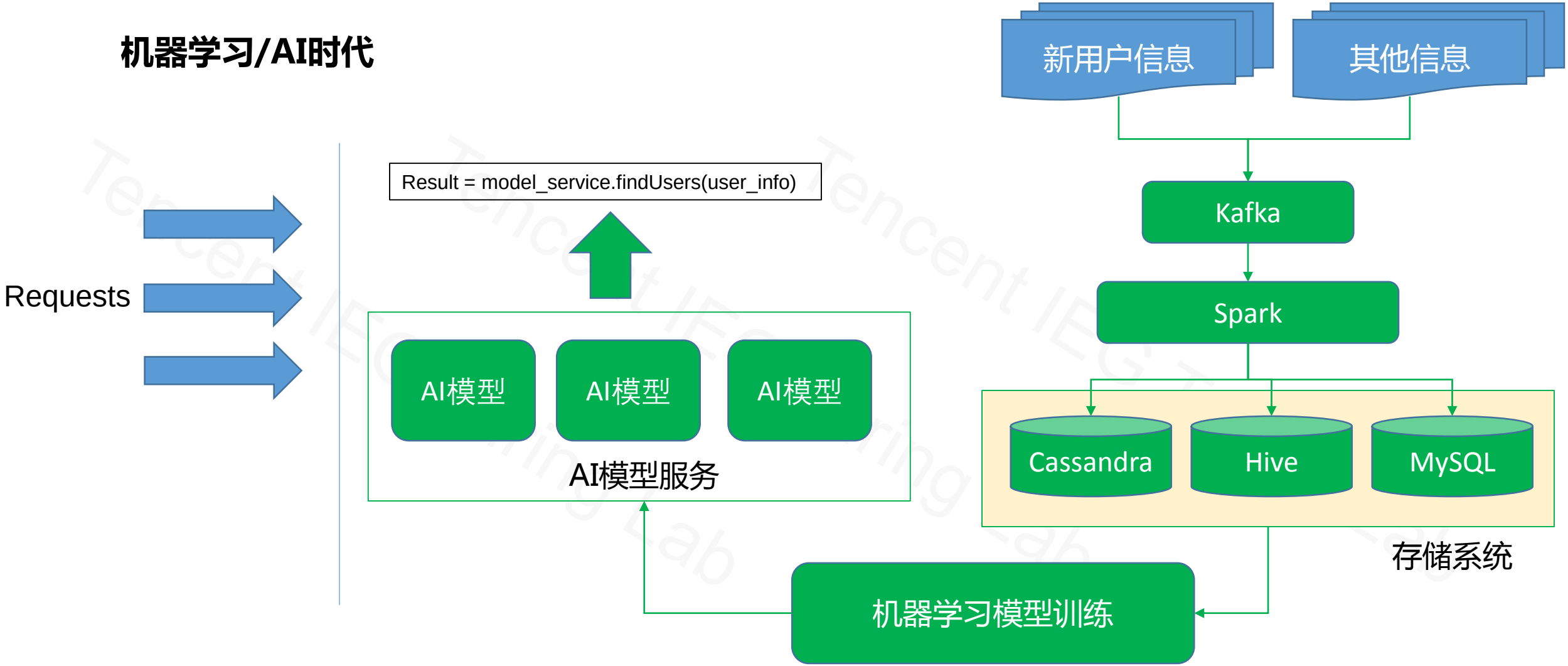


Tencent IEG Turing Lab

腾讯游戏 Tencent Games Turing Lab

Why middleware? ----好友推荐系统的演化

机器学习/AI时代

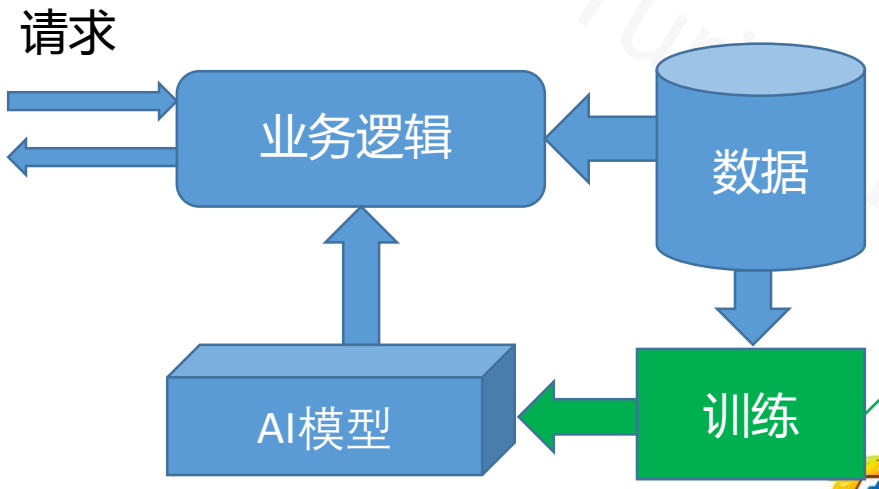


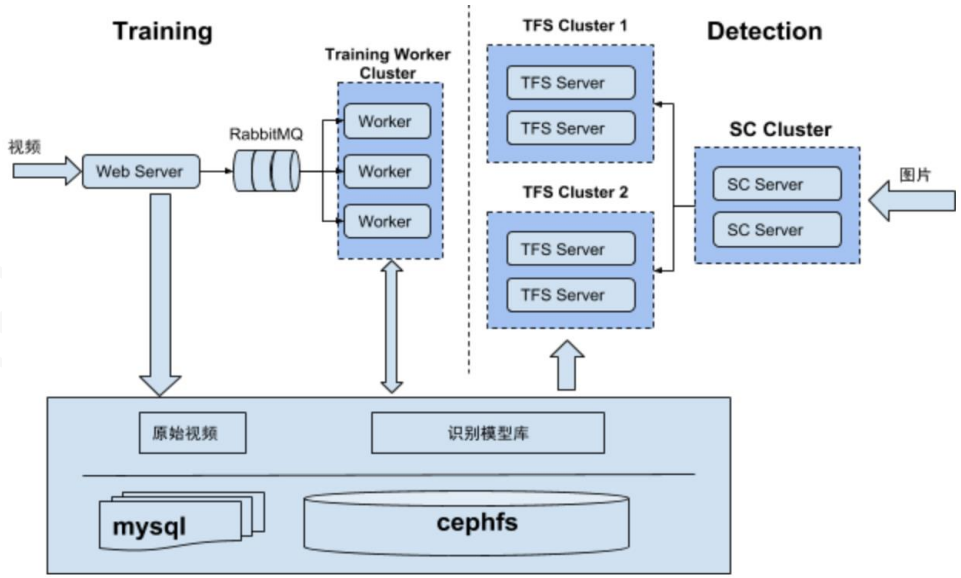
Machine Learning and Backend Design

Recommend 火龙果软件

传统后端系统

基于机器学习





一个图片AI检测系统框架

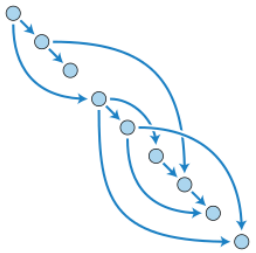
一般web请求：<1秒
大数据读取(Hive): 分钟
文件上传/下载：几秒至几小时
视频处理：几分钟
AI模型训练：**几小时或几天**

- 异步结果(Asynchronous Result)
- 并行任务(Parallel Tasks)

Tencent IEG Turing Lab

腾讯游戏 Tencent Games　Turing Lab

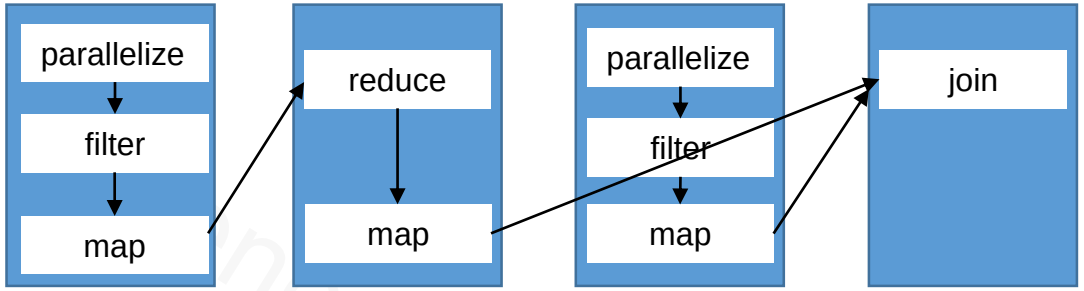
DAG（有向无环图）in Backend Architecture



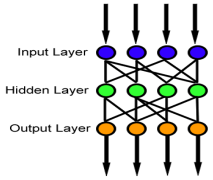
DAG
(from Wikipedia)

DAG(Directed Acyclic Graph)
- Immutable Data
- No Cycle
- Can only go from earlier to later (topological ordering)

Spark



Neural Network



Machine Learning Platform

TensorFlow

Tesla.oa.com

Programming Practice

ReactJS

腾讯游戏 Tencent Games

Turing Lab

Reactive Programming

RxJava | RxSwift | RxPy

Tencent IEG Turing Lab

THANKS