

# Technological Innovation at Alibaba

Alan Qi

Vice President of Ant Financial Service Group

# Outline



- 11-11 & Alibaba
- Technologies at Alibaba
- Conclusions



# What is Singles' day (11-11)?

# Black Friday vs. 11.11



## Black Friday (2014)

1 Country 1 Day

**2.3 Billion**

Dollar Online GMV

**9.1 Billion**

Dollar Offline GMV

**249M**

Online & Offline Customer

**27.9%**

On Mobile Client

## 11.11 Shopping Festival (2015)

1 Company 1 Day

**14.3 Billion**

Dollar Online GMV

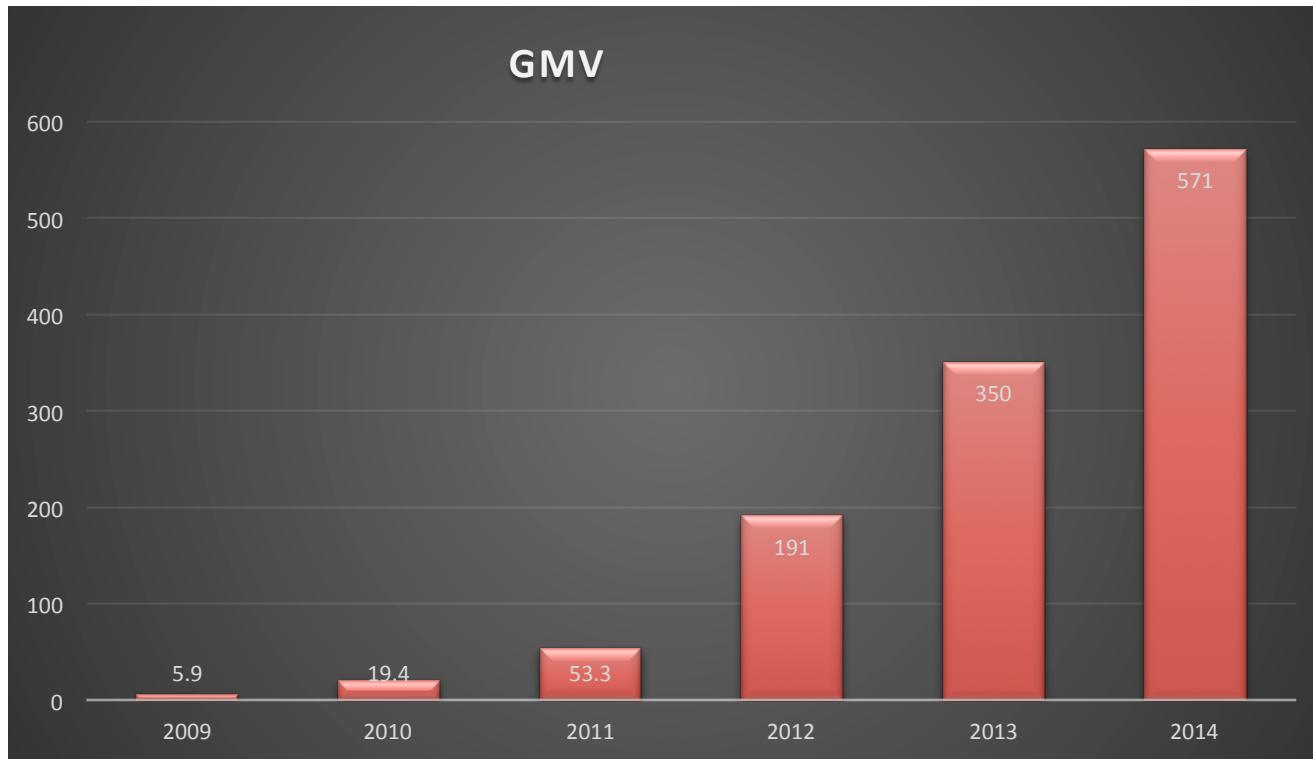
**>249M**

Online Customer

**68%**

On Mobile Client

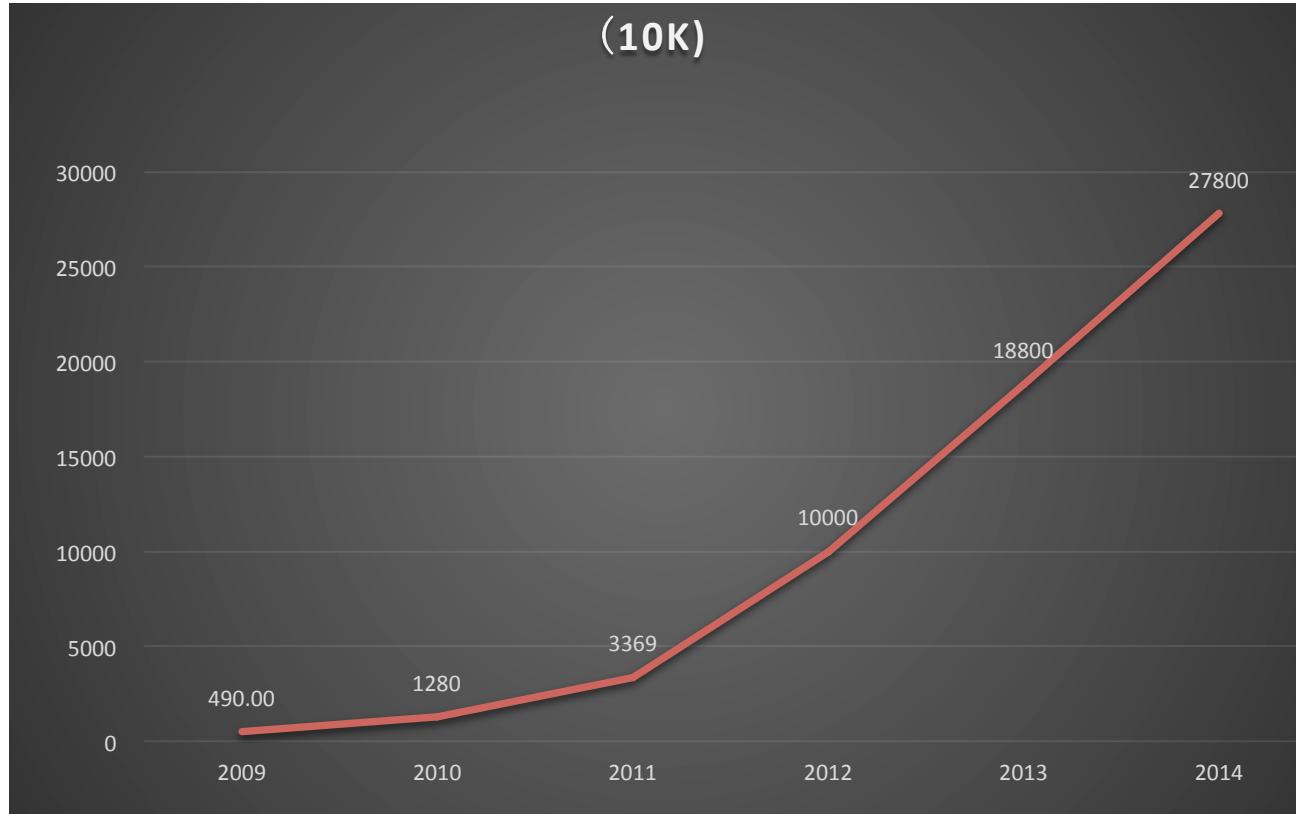
# "11-11" sales revenue explodes



In 7 years, sales revenue increases  
**154 times**

2015/11/11: 91.2 Billion Yuan (14.3 Billion USD)

# “11-11” number of orders explodes



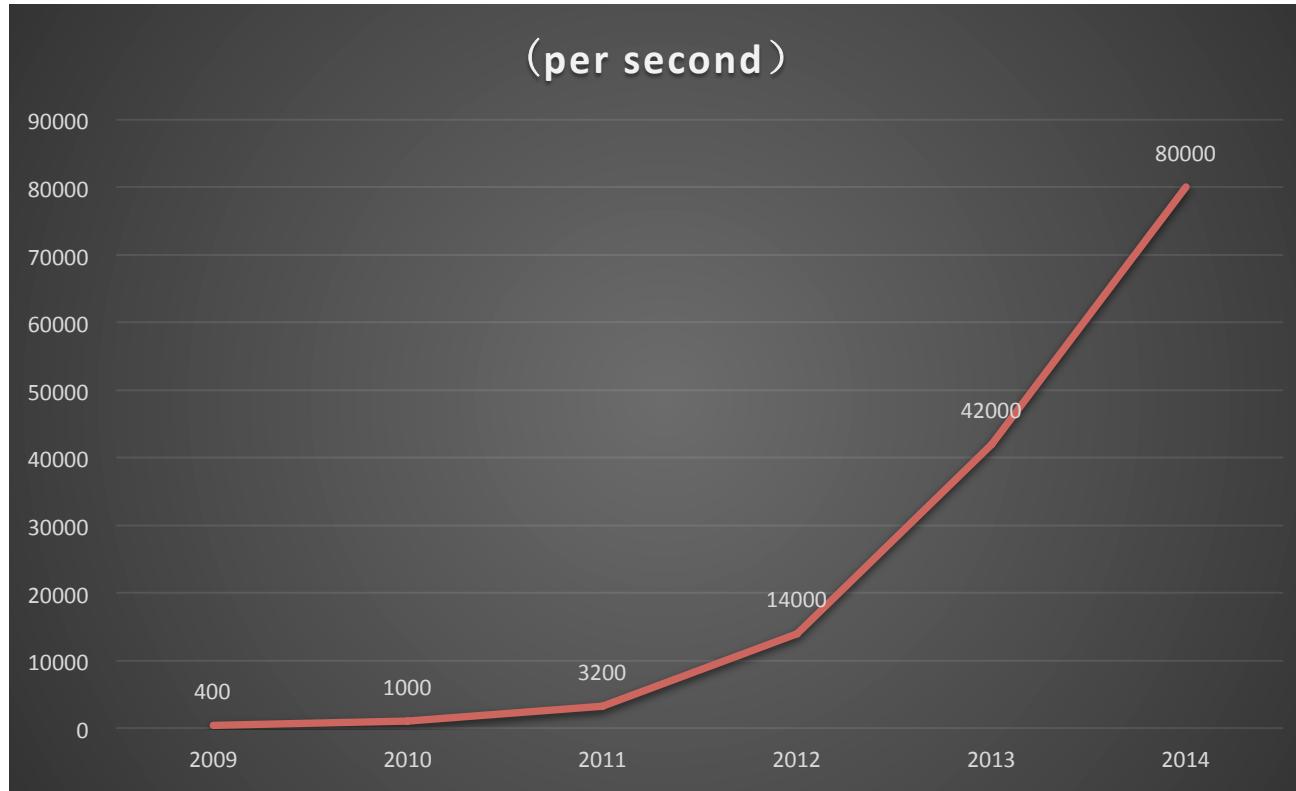
In 6 years, number of orders increases

**57 times**

2014/11/11  
Total orders: 278 Millions



# “11-11” order peak value explodes



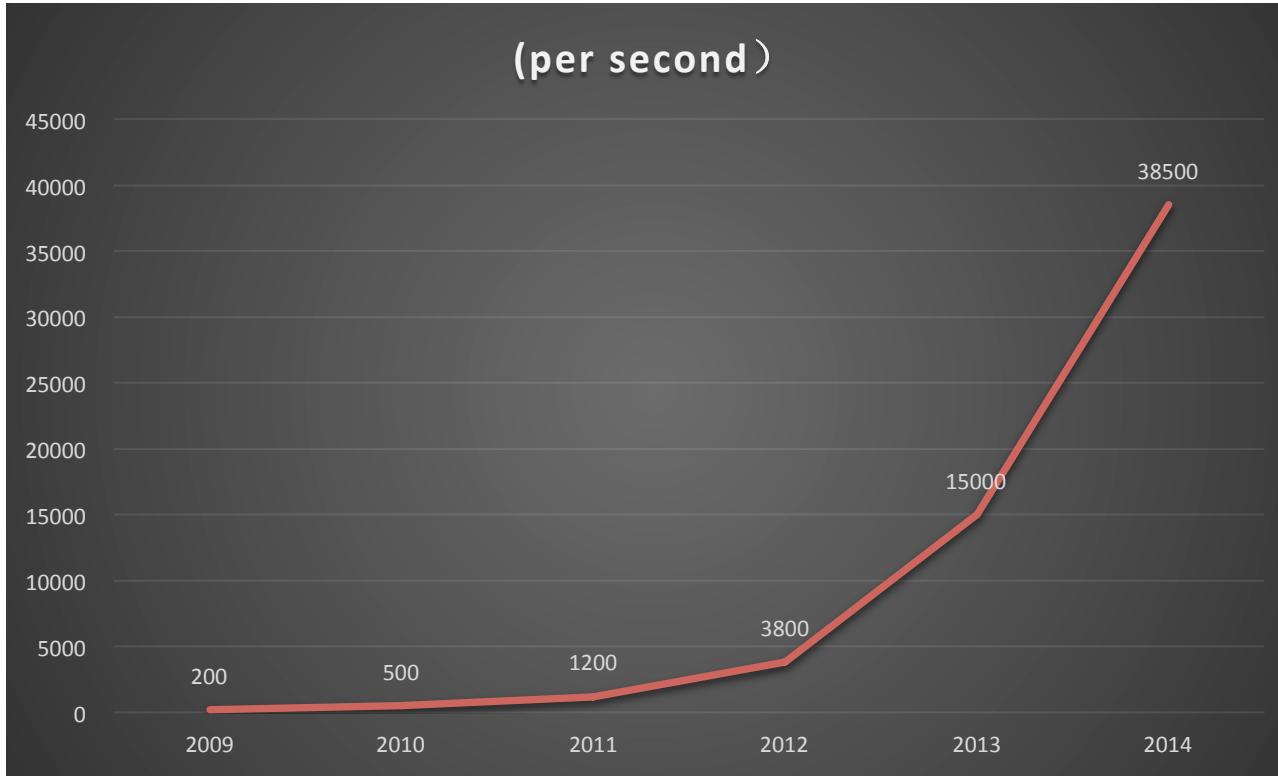
In 6 years, order peak values increases

**200 times**

2014/11/11  
Order peak value: 80K per sec.



# "11-11" transaction peak value explodes



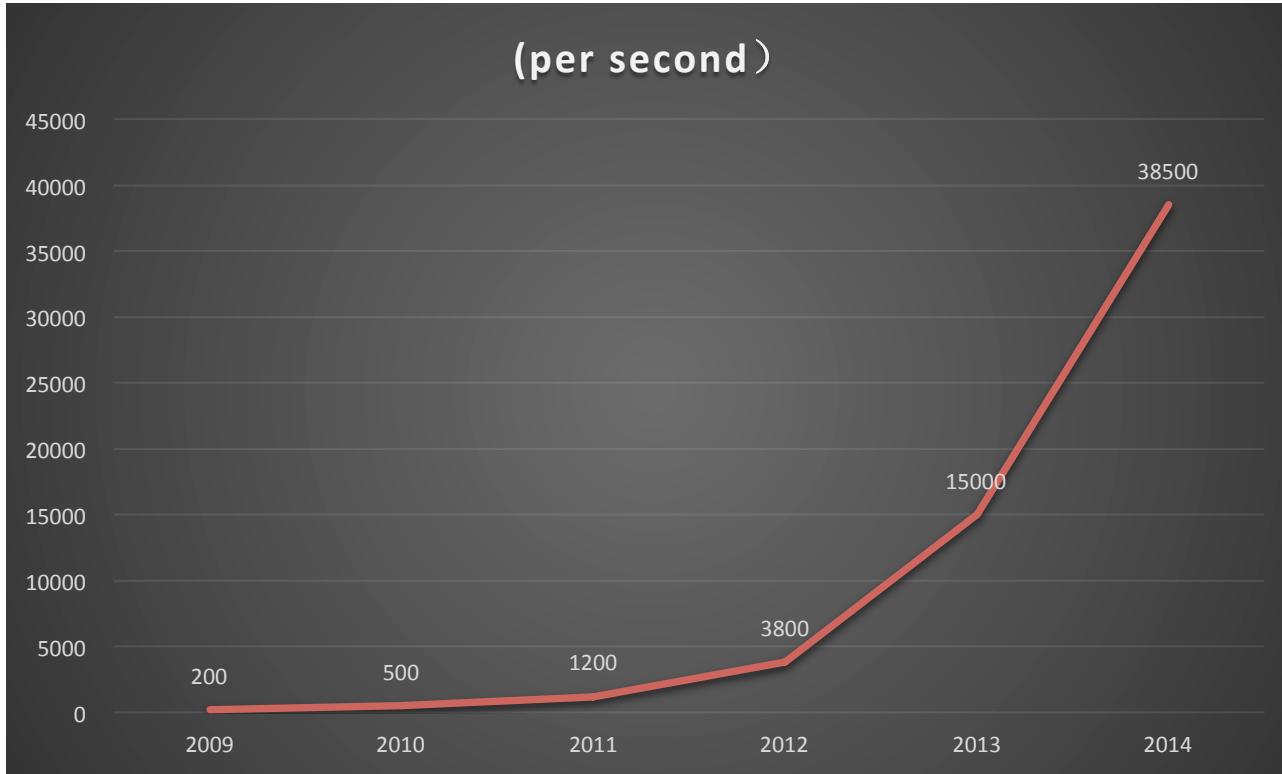
In 6 years, transaction peak values increases

**193 times**

2014/11/11  
Alipay: 38.5K per sec.



# "11-11" transaction peak value explodes



In 7 years, transaction peak values increases

**430 times**

2015/11/11  
Alipay: 85.9K per sec.

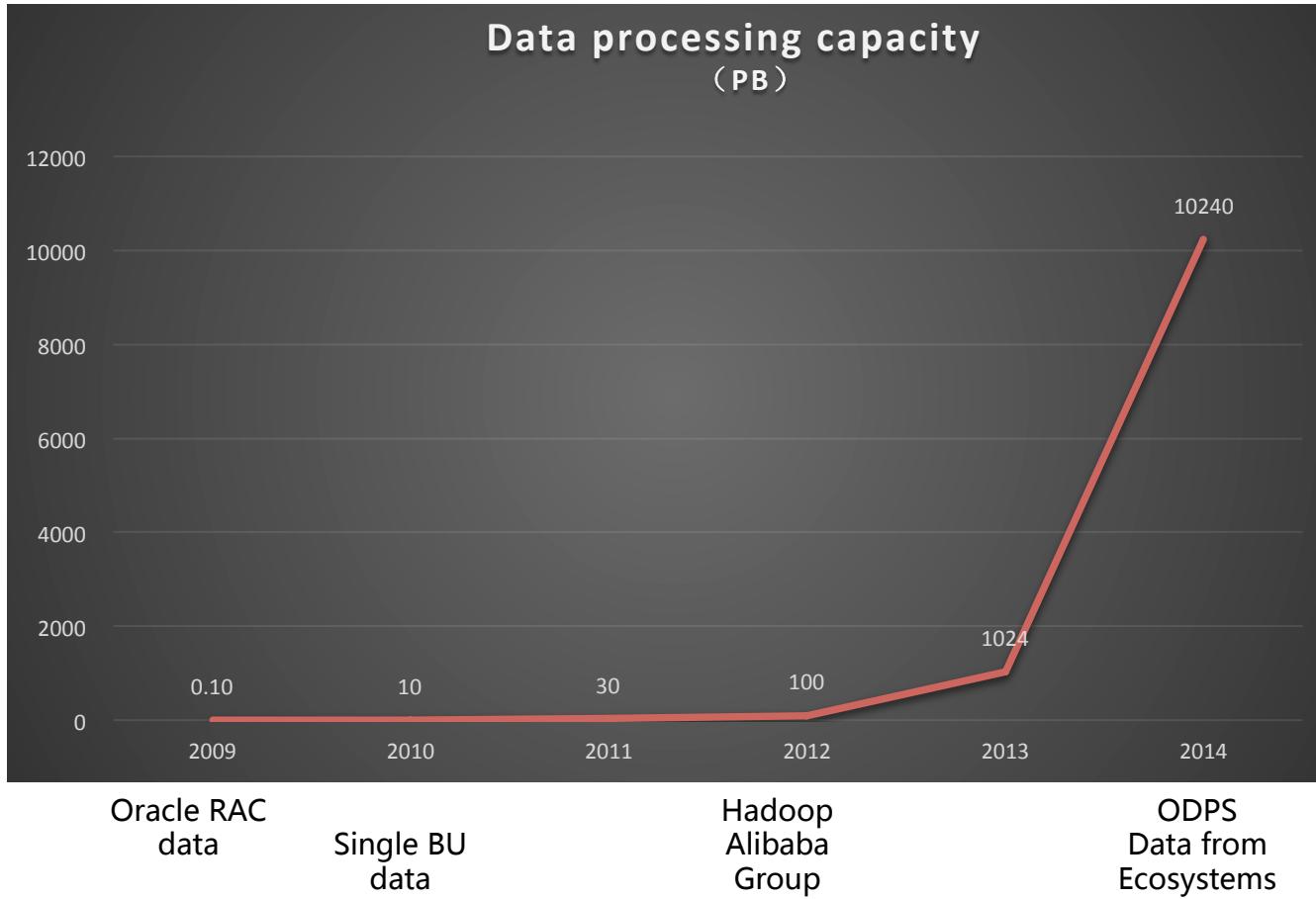
Visa: 56K per sec. in lab  
14K per sec in real

Master: 40K per sec. in lab

**Alipay: Strongest processing power among all payment systems**



# Capacity of processing financial big data explodes



In 6 years, data processing capacity increases

**100K times**

2014/11/11 :  
Total processed data :  
100PB

# “11-11” needs information security





# The company behind 11-11?

# Alibaba History



Alibaba founded in Jack Ma's apartment in Hangzhou

1999



RMB1M profit daily; Alipay is launched

Company breaks even

2002



Alibaba.com lists in HK

2004



Alibaba Cloud Computing is founded

2007

Taobao is re-organized into Taobao Marketplace, Tmall.com and eTao

2011



CAINIAO 菜鸟  
Alibaba Group is reorganized into 25 business units; announces a plan to set up Small and Micro Financial Services Group; leads the formation of Cainiao

2013

Alibaba announced "Hangzhou +Beijing" dual-center strategy

2015



2001

Alibaba.com surpasses 1M registered users

2003

RMB1M revenue daily; Taobao is founded



2005

RMB1M tax paid daily; Acquired China Yahoo!



2008

Tmall.com is introduced  
**Tmall TMALL.COM**

2010

AliExpress is launched; Juhuasuan and eTao are introduced  
**AliExpress**



2012

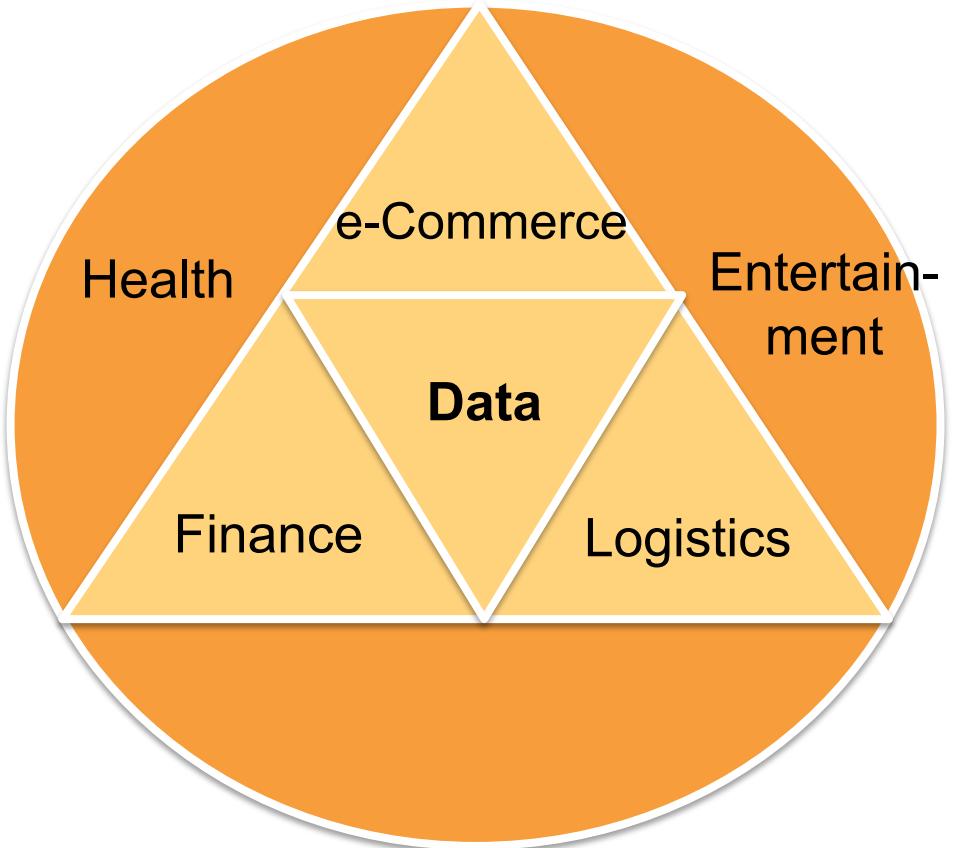
Alibaba Group establishes the Alibaba Foundation with a sizeable fund dedicated to social causes



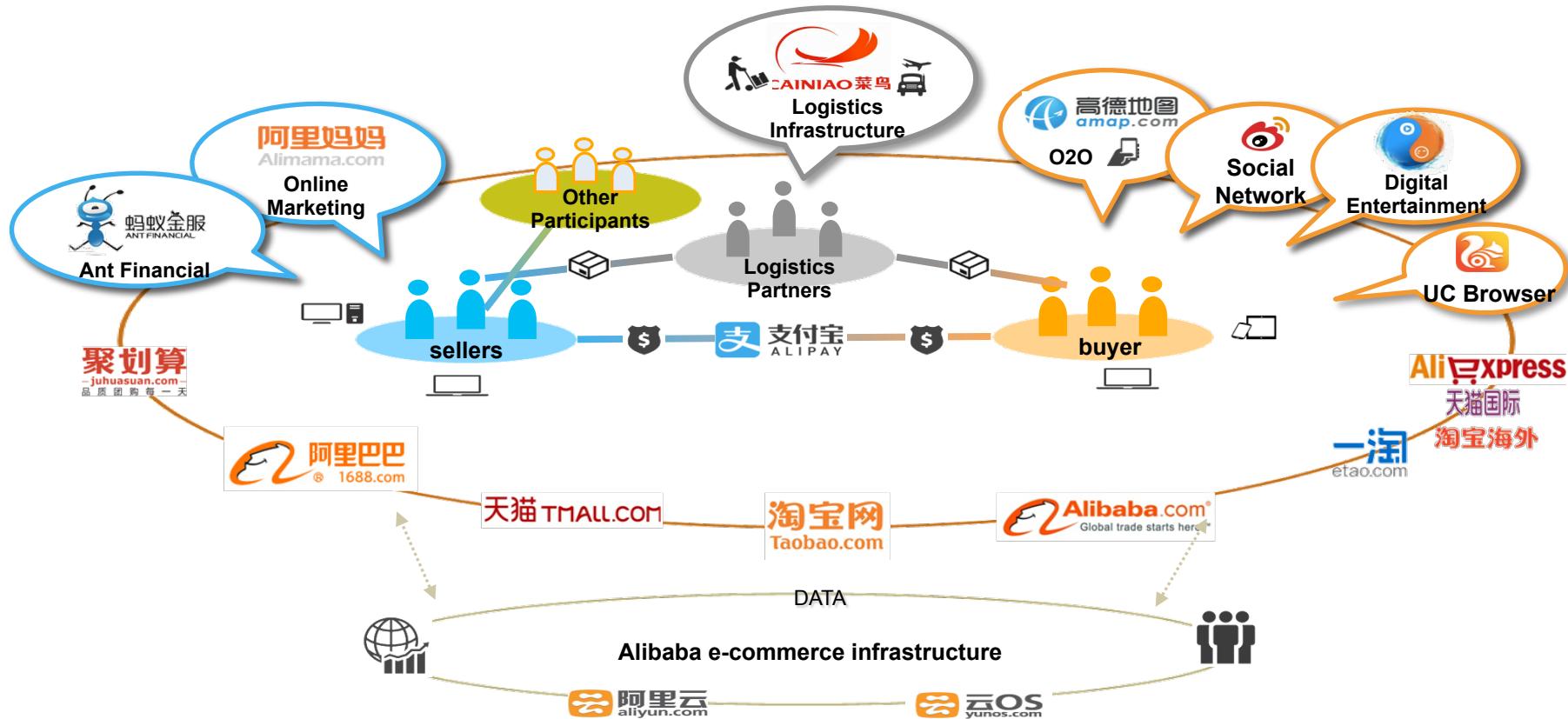
2014

September 19<sup>th</sup> : Alibaba Group's IPO (Nasdaq: BABA)

# Alibaba Strategy



# Alibaba Ecosystem



# *Technologies at Alibaba*

# Key Technologies

Deep user profiling & interest mining to better connect advertisers to their target audience

Large-scale real-time “search by image” technology that powers camera-based product search

9 large data centers.  
Active-active data centers across multiple cities technology

Break 4 world records on Sort Benchmark 2015.  
3 time faster than Apache Spark

Advertising

Infrastructure

Search

Cloud Computing

Security

Mobile

Machine Learning

Middleware

Dynamic restoration & deployment ability of mobile clients

Super-large scale learning from data with billions of attributes and billions of samples

A secure e-commerce ecosystem for online transactions, data exchanges, cloud computing and mobile interactions.

Ten billion level RPC and message processing capability



Alibaba Group

存

## High hardware cost

IOE (IBM servers, Oracle databases and EMC storage): expensive for scalability. **In a few years, the hardware cost could have bankrupted the whole company!**

通

## Data islands

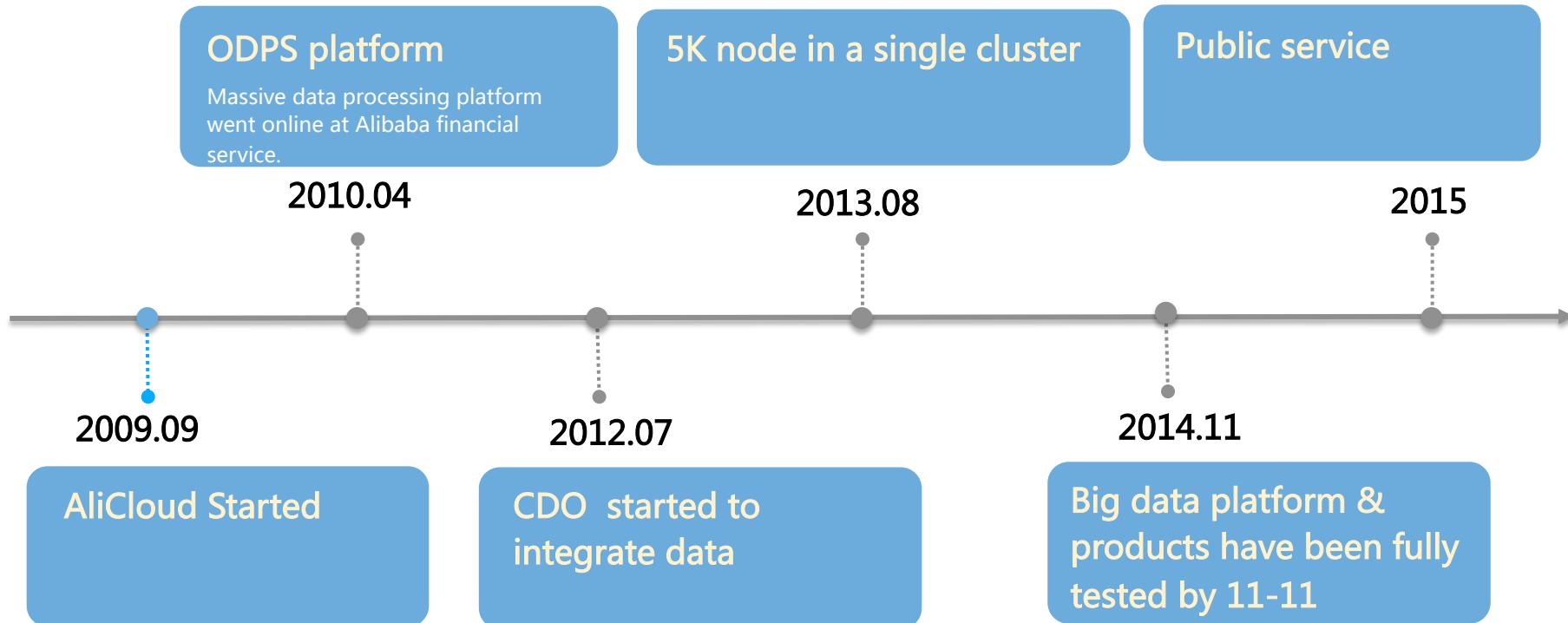
Data from different business units are located in different clusters. Extremely hard to share data with each other!

用

## No common data standard for usage

Due to isolated storage, data are moved between clusters with repeated storage and computations.

# Milestones



# Data-driven business at Alibaba



Dozens of business units  
Hundreds of PB data  
Thousands of engineers

# *Cloud Computing & Infrastructure*

**08/27/2010**

Apsaras became the platform of four applications: Search, Mail, Image Storage, AliFinance

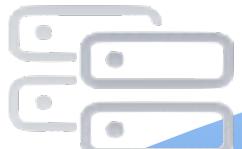


**09/10/2009**

Aliyun.com Inc established

**08/15/2013**

5000-node Apsaras cluster went into production



**07/28/2011**

Aliyun.com went online, releasing 1<sup>st</sup> cloud service: ECS

**5K**



**07/2014**

Launched ODPS, which is capable of processing 100 petabytes of data in six hours.



**10/15/2015**

Achieved ITU Telecom World Entrepreneurship Award for excellence in providing and promoting innovative ICT solutions with social impact

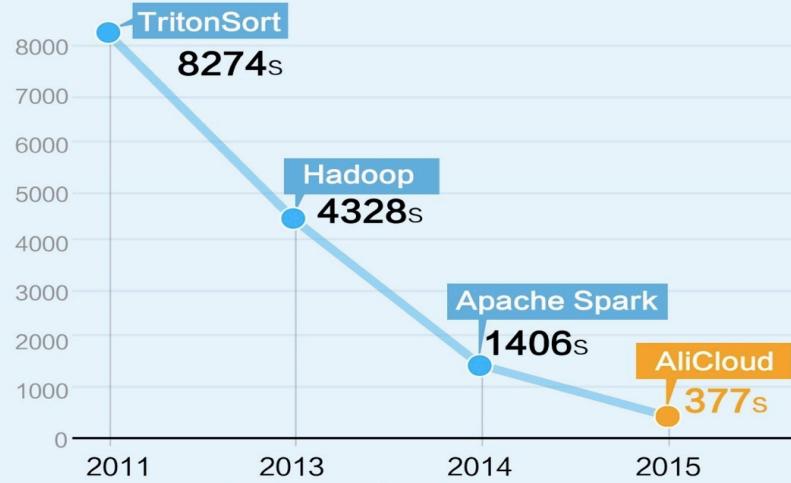
# World Records



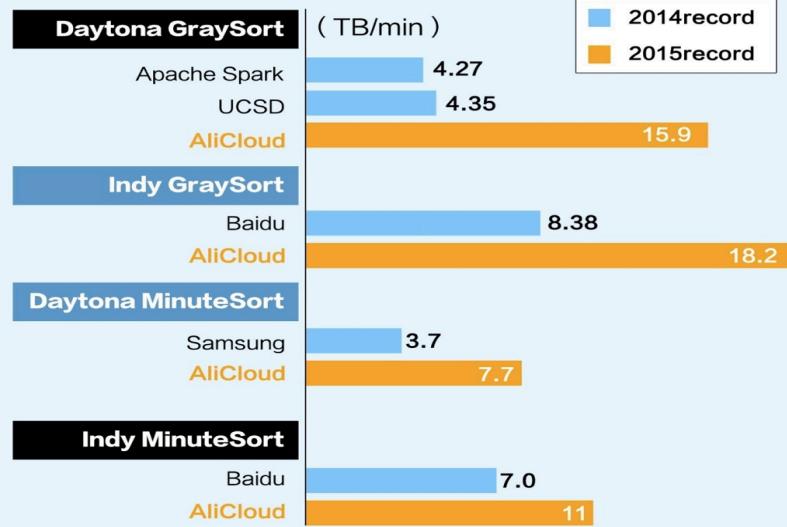
- Break **4 world records** on Sort Benchmark 2015
- Sort 100 TB of data in 377 seconds, **3 times faster** than the record set by Apache Spark in 2014

## Top Results

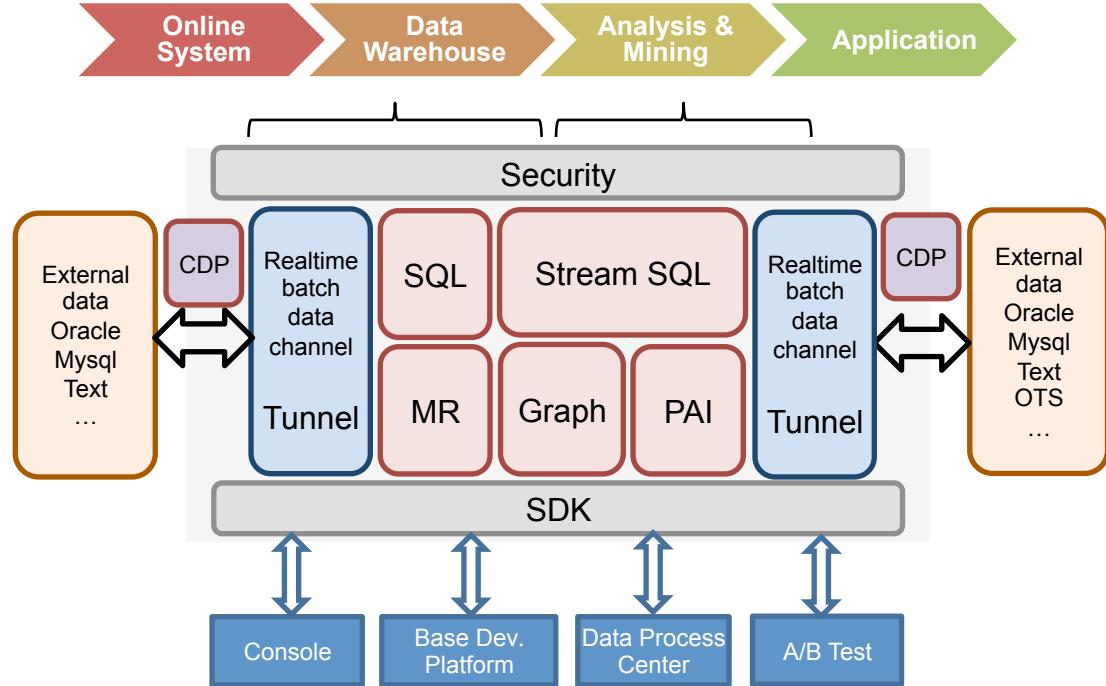
(time taken to process 100 terrabytes of data)



## Sort Benchmark's world records



Capable of processing **100 petabytes** of data in **6 hours**.



### Distributed



### Safety



### Easy to Use



### Authorization and Access





## 中国铁路客户服务中心

www.12306.cn是中国铁路客户服务中心唯一网站，截至目前，没有授权其他网站。

2013年3月17日 星期日

首 页 | 客运服务 | 货运服务 | 行包服务 | 车站引导 | 铁路常识 | 站车风采 | 客户信箱

### 站车风采



高寒通勤  
动车组席位  
更多>>>

### 最新动态

为保障您顺利购票，请下载安装相证书。

- 关于银行系统升级暂停发售电子支付业务的通知 (2013-03-15)
- 关于调整部分车站互联网、电话订票起售时间的公告 (2013-01-24)
- 铁路旅客服务质量网上调查问卷 (2013-01-16)
- 铁路旅客服务质量网上调查问卷第三次抽奖结果 (2013-01-11)
- 青藏铁路公司关于加开3月21、23日西宁西—拉萨K9015次旅客列车的通知 (2013-03-14)
- 青藏铁路公司关于加开西宁西—拉萨旅客列车的通知 (2013-03-13)
- 成都铁路局关于部分列车停运的公告 (2013-03-11)

更多>>>

### 中国铁路货运电子商务平台

快运货物班列  
快运货物订票服务  
货运站综合信息  
货运调度查询  
货运业务咨询  
相关链接

网上购票用户注册  
购票 / 预约  
退 票  
全 票 查 询  
旅客列车时刻表查询  
旅客乘车证查验

全文搜索: 请输入搜索条件   搜索

铁路客运 法律法规及规范性文件  
铁路货运 法律法规及规范性文件

网上购票常见问题 铁路常识 货运网上受理常见问题

- 在互联网购买了儿童票，如何在售票窗口换取纸质车票？
- 注册用户时，系统提示身份证信息重复怎么办？
- 购买实名制车票后丢失了怎么办？

CNTV: online resource integration for CCTV (biggest TV station in China) and online service. Cost saving: >50%



The screenshot shows the homepage of the CNTV online service. The top banner features the CNTV logo and the text "玩游戏 看民俗 赢大奖" (Play Games, Watch Folk Customs, Win Prizes) with fireworks and traditional Chinese figures. Below the banner, the "活动流程" (Activity Flow) section explains the process: New users register, log in, answer questions, invite friends, and participate in draws. It also highlights that CNTV integrates and converts content from various channels through the internet to provide online services. The "登录" (Login) section includes fields for username and password, and links for registration and password recovery. A large red button labeled "登录" (Login) is prominent. The main content area features a "1+1双人大礼包!" (1+1 Double Big Gift Pack) offer, showing a stack of four tablets and a magazine cover for "悦己SELF". It also displays a tablet displaying a news feed. A "一等奖: 1名" (First Prize: 1 person) for an Apple iPad MB29 2CH/A 9.7英寸平板电脑 is listed. The "合作伙伴" (Partners) section lists several companies: 悅己SELF, 酷狗音乐 (Kugou Music), PPTV, 酷六网 (Ku6.com), 木蚂蚁应用市场 (Mumayi.com), and 暴风影音 (BaoFeng YingYin).

CNTV 中国网络电视台

玩游戏 看民俗 赢大奖

活动时间: 2011年1月19日-2月17日

活动流程

新用户请先注册, 登录后方可参与游戏、答题、邀请好友等环节。

CNTV肩负对央视众多频道内容资源进行整合和转换通过互联网提供在线服务

登录

密码

注册 忘记密码

登录

登录 > 答题 > 邀请 > 抽奖

1+1双人大礼包!

如果您邀请的站外好友中奖, 您将得到与他一样的奖品1份!

邀请好友

本次活动奖品由悦己SELF网提供

一等奖: 1名  
苹果 (Apple) iPad MB29 2CH/A 9.7英寸平板电脑

合作伙伴

悦己SELF

酷狗音乐

PPTV

酷六网

木蚂蚁应用市场

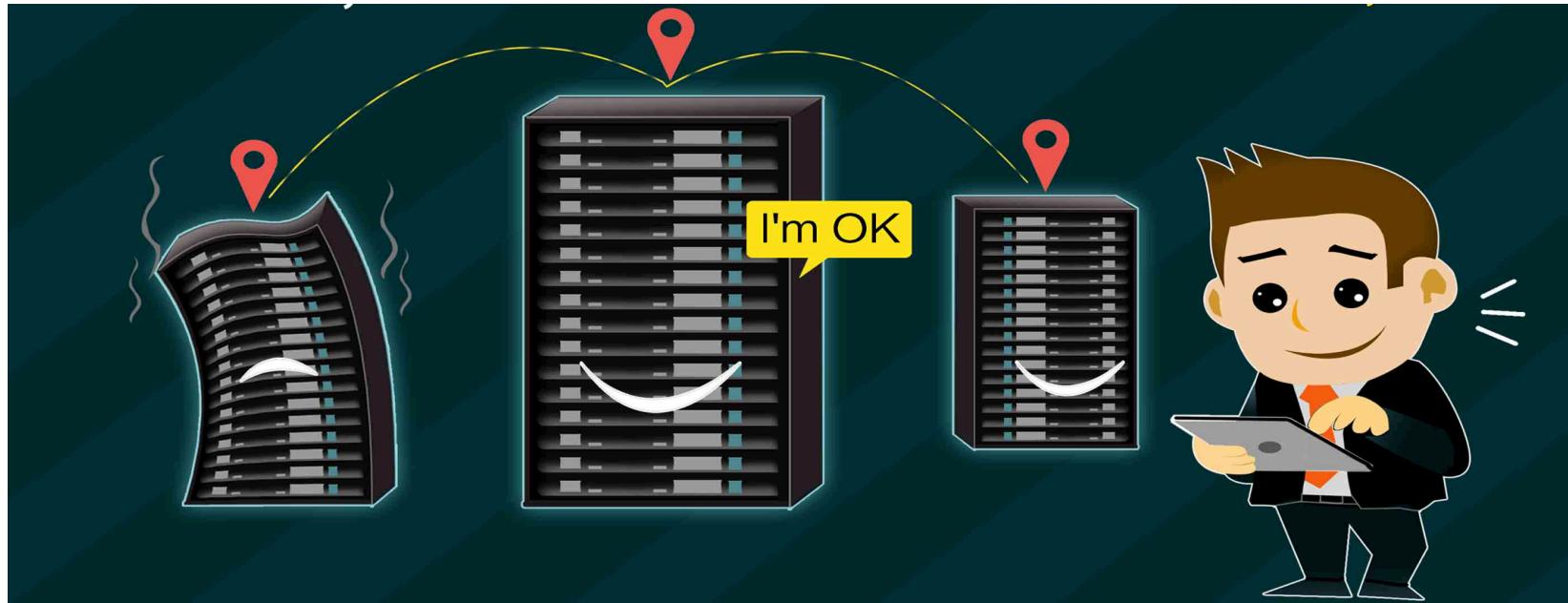
暴风影音

# Reliable service via data sharing



## Active-active data centers across multiple cities

We can distribute network traffic between our Shanghai, Hangzhou and Shenzhen data center without any interruption.



# World-wide Data Centers



9 large data centers in China Mainland, Hong Kong, Singapore and America.

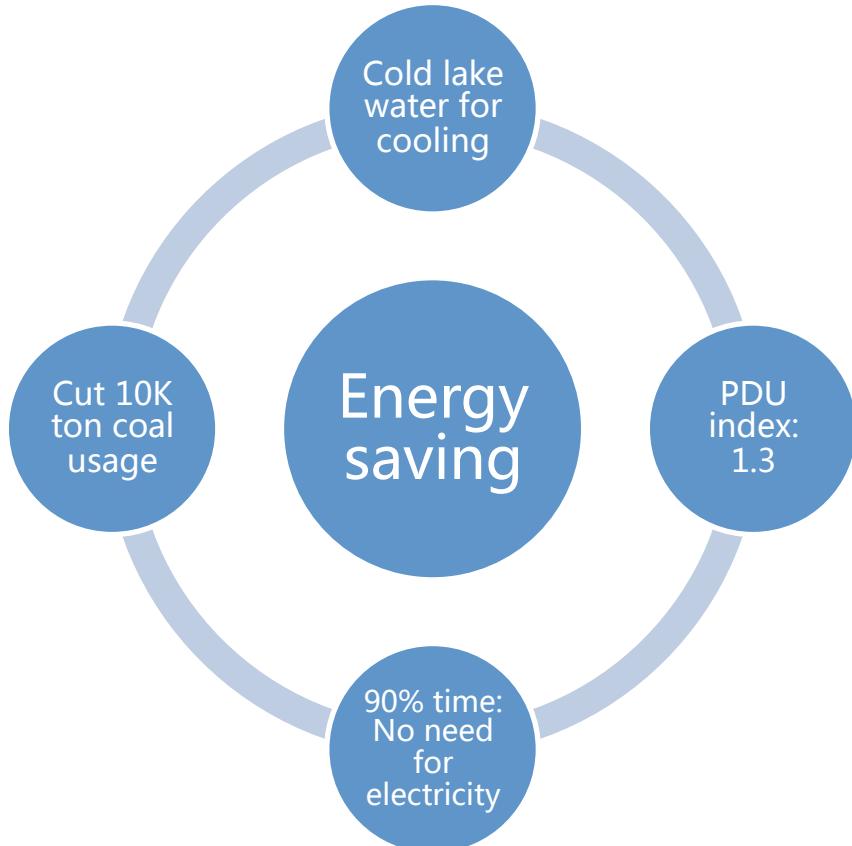


# Green Data Center: Zhang Bei



- **Zhang Bei: Strong wind, averaged temperature < 37 degree F.**
- **Deploy solar energy, wind power and cold water cooling solutions: 100% green**
- **Estimated at least 45% energy saving for cooling.**

# Green Data Center: Thousand Island Lake



# Information Security

- Intercept 3 million deceptive transaction requests every day (user identification, etc.).
- Totally successful defense of DDoS attack : **546,859GB**
- Totally successful defense of attack since 2014: **9,383,772,384** times
- Dec. 20<sup>th</sup>, 2014, Defeated a DDoS attack , which lasted **14 hours** and peak traffic reached **453.8 Gb per second**



# *Big Data & Machine Learning*



切换到商户版 >

首页

常见问题

自助服务

联系客服



## 在线客服

推荐指数: ★★★★★

适用人群: 喜欢使用在线交流方式的用户

立即使用 >



## 电话客服

推荐指数: ★★

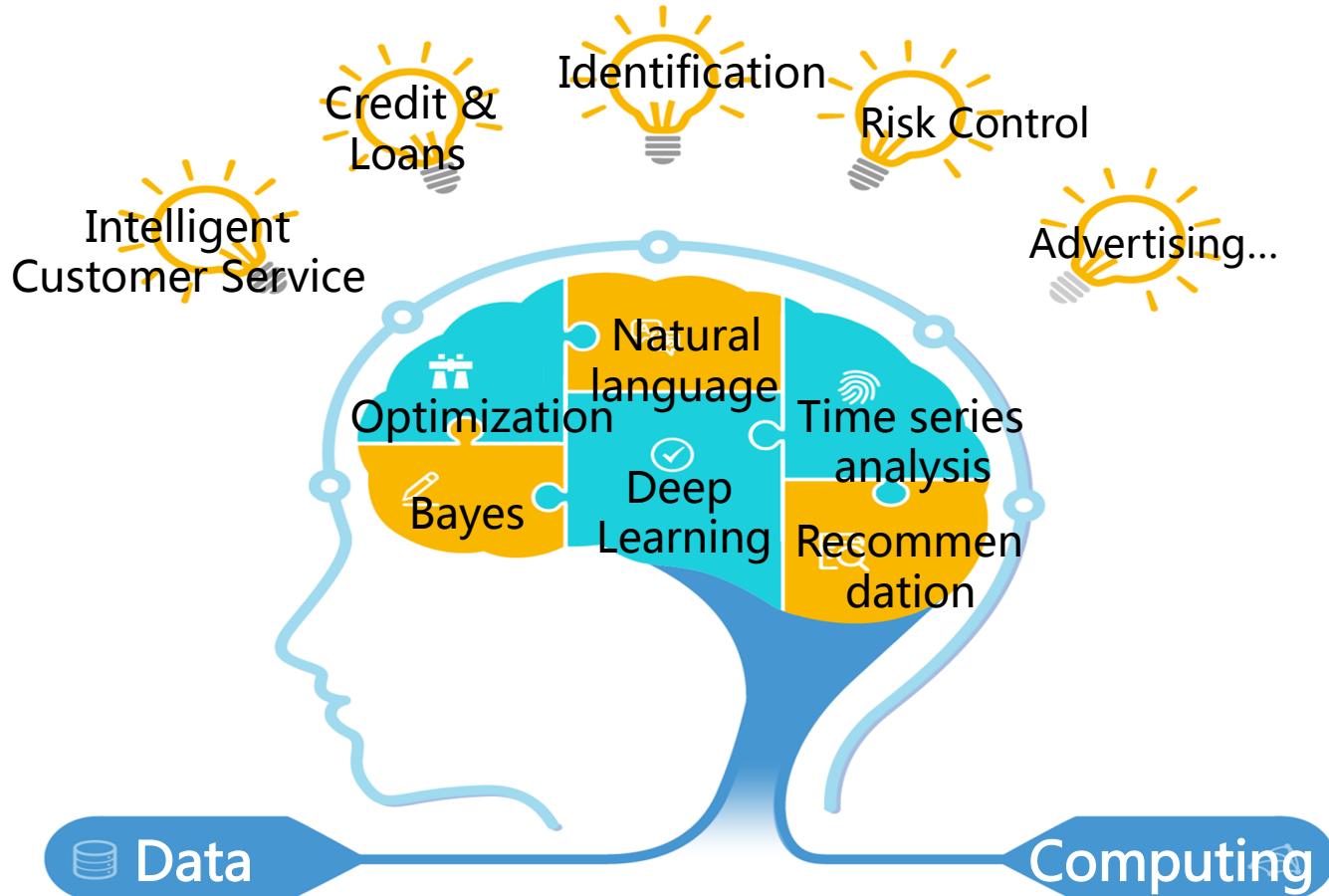
适用人群: 喜欢使用电话方式解决问题的用户

📍 数字代表按键, 如3代表3号键。客服主要服务时间: 9:00—24:00

## On 11/11/2015

- **Total service volume: > 10 Millions**
- **> 95%: supported by intelligent robots**
- **< 5% : supported by human**
- **>98% phone call: went through**

# AI at Alibaba

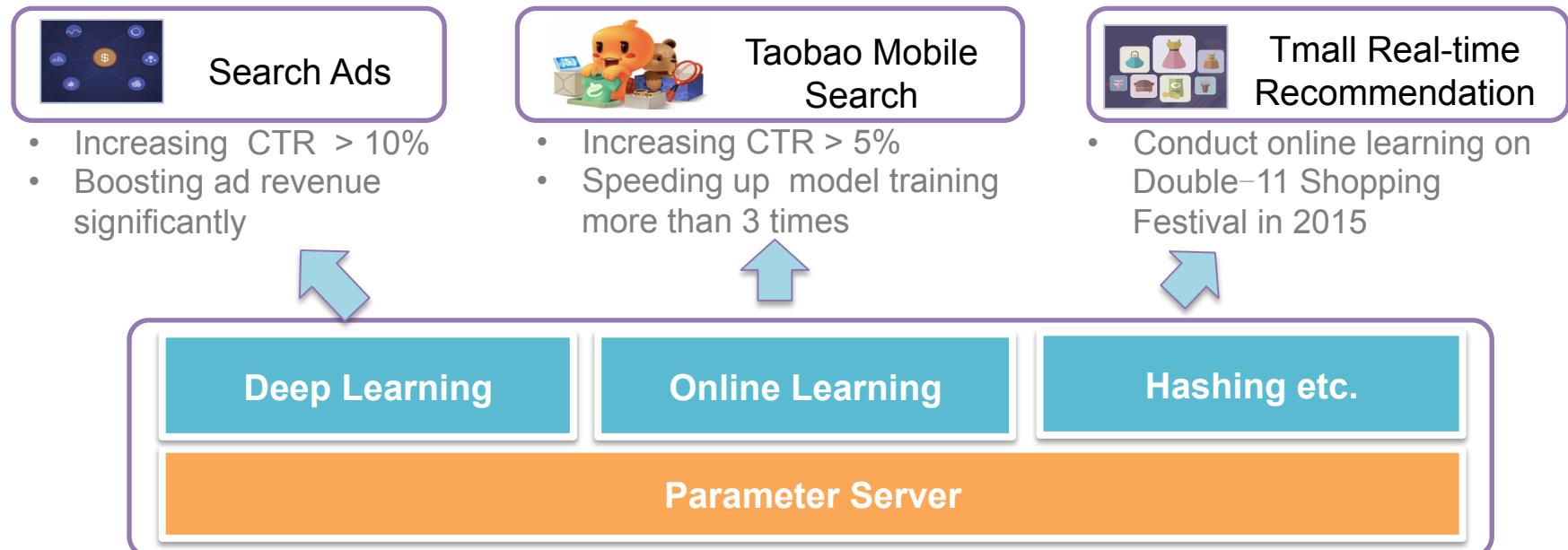


# Intelligent customer service

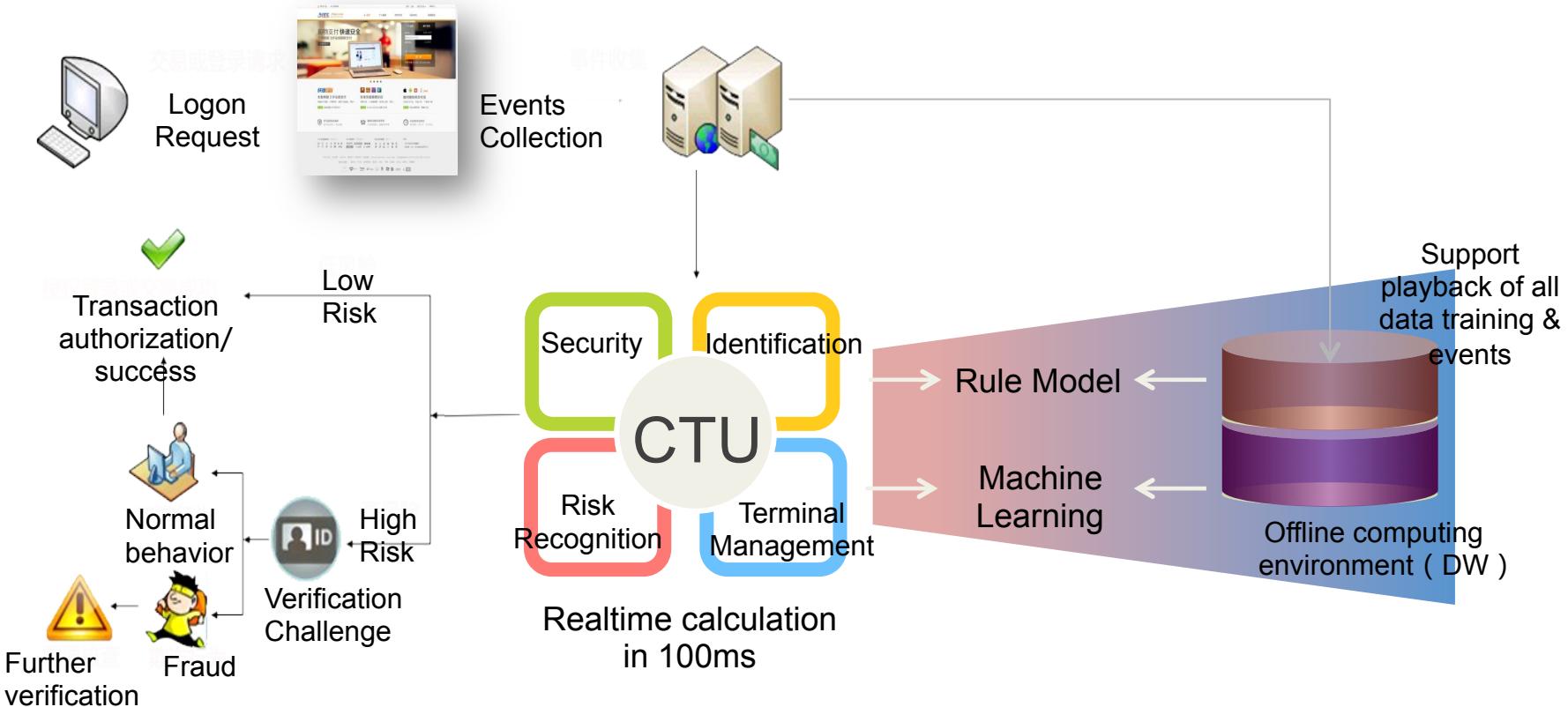


- Speech recognition
- Deep Learning for problem identification & QA
- Natural language processing
- Knowledge base construction

# Large-scale Machine Learning



# Risk Control (CTU)



- Using big data technologies to discover risks and identify abnormal behaviors.



- Address challenges for small and medium enterprises to meet their financial need:
  - Small loans
  - Tedious loan approval procedure
  - Low approval rate
  - Long cycle

Farmers need loans for tractors

# Data-driven micro loans for small enterprises



- 10 millions plus loans
- 90 billion+ RMB portfolio
- Over 300 thousand enterprises

- No human intervention based on machine learning methods
- Default rate way less than traditional banks

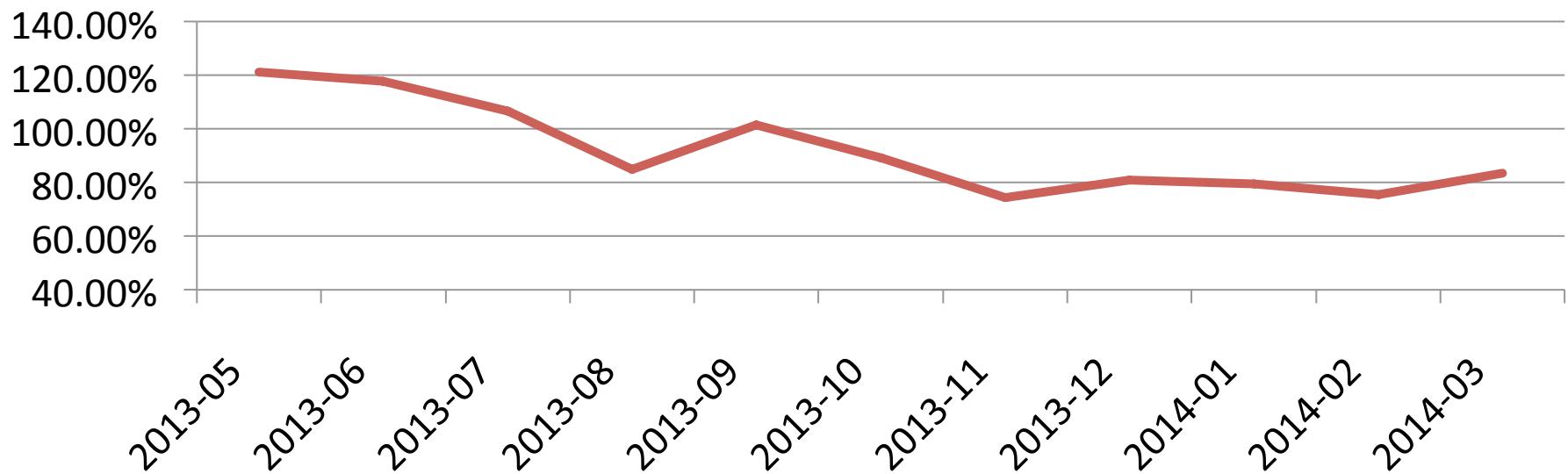
# Microloans for rural areas in China



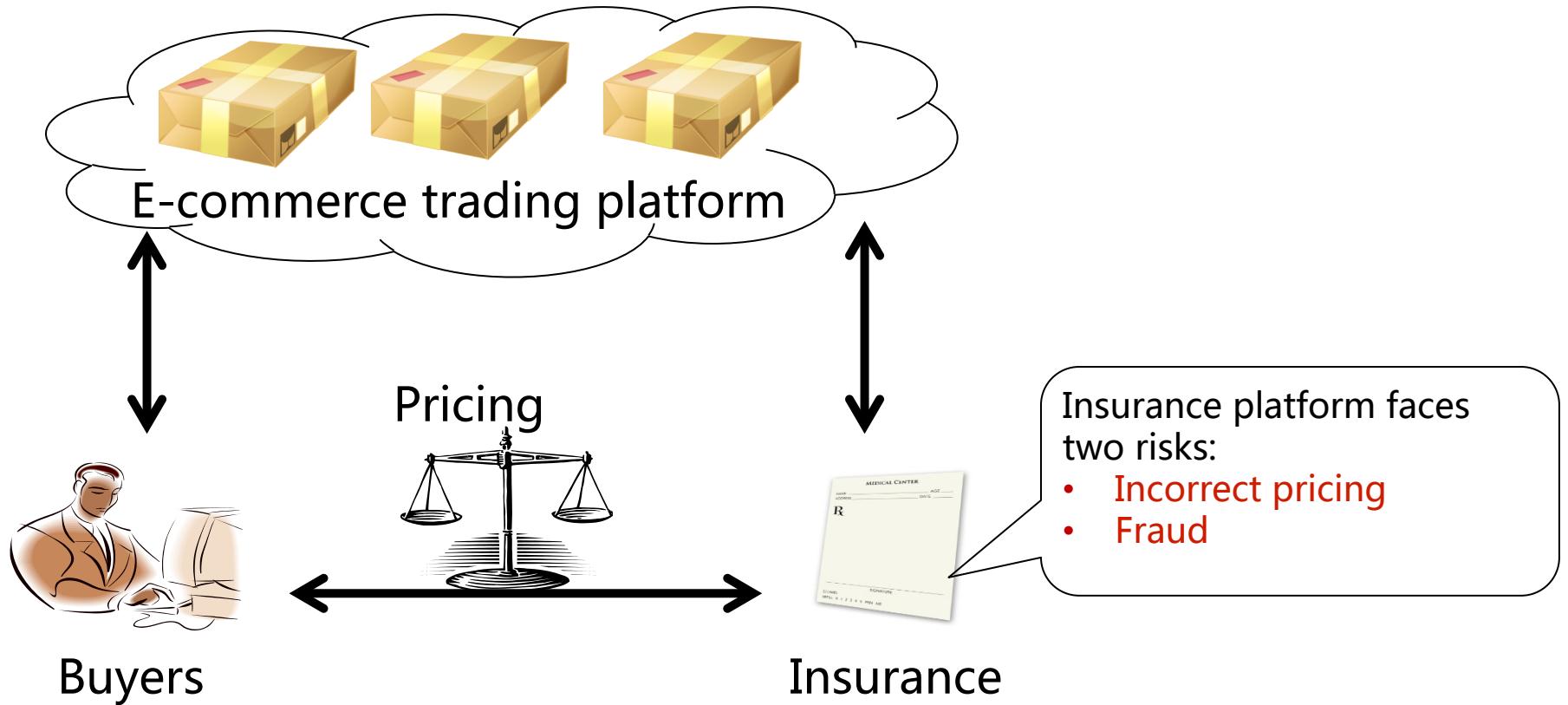
180 Thousand loans for small and medium enterprises in rural areas, in total 17.8 Billion Yuan (~2.8 Billion USD)

- Insurance contracts has year-on-year growth rate of 100%.
- Over 1 billion contracts in 2013
- Record-breaking over 100 million contracts one day on November 11, 2013

## Overall rate of compensation



# Shipping Insurance



# Shipping Insurance for Returned Products

Fixed rate

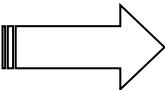
- ☐ Uniform 5% fixed rate

Simple

Actuarial approach

- ☐ Solely based on historical data and demographics

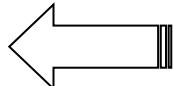
Easy to explain



Dynamic pricing based on big data

- ☐ Millions of features, real time pricing
- ☐ Machine learned models

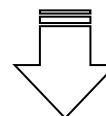
High accuracy



Model driven pricing

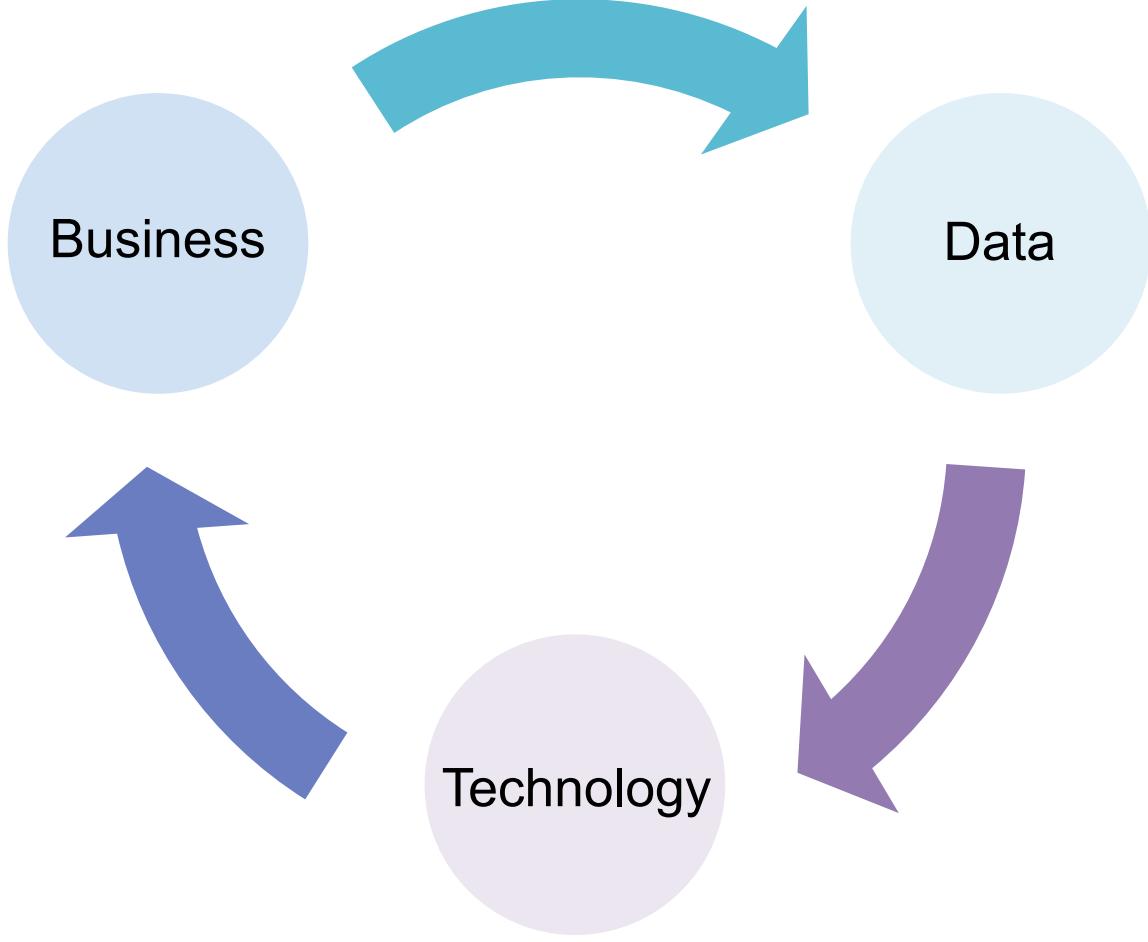
- ☐ Pricing model based on a few parameters

Improved accuracy



# *Conclusions*

# Feedback loop



# Technology Expands Business Boundary



- Enable devices to think & learn like human-being.



Artificial  
Intelligence

- Secure communication & high-performance computation.



Software  
Defined  
Network

- New network framework providing an innovation solution for network virtualization.

Quantum  
Comm. &  
Comp.

Smart City

- Future cities fusion of cloud comp., big data analysis, mobile Internet and social networks.



**THANK YOU**