

# 数据中心技术

---

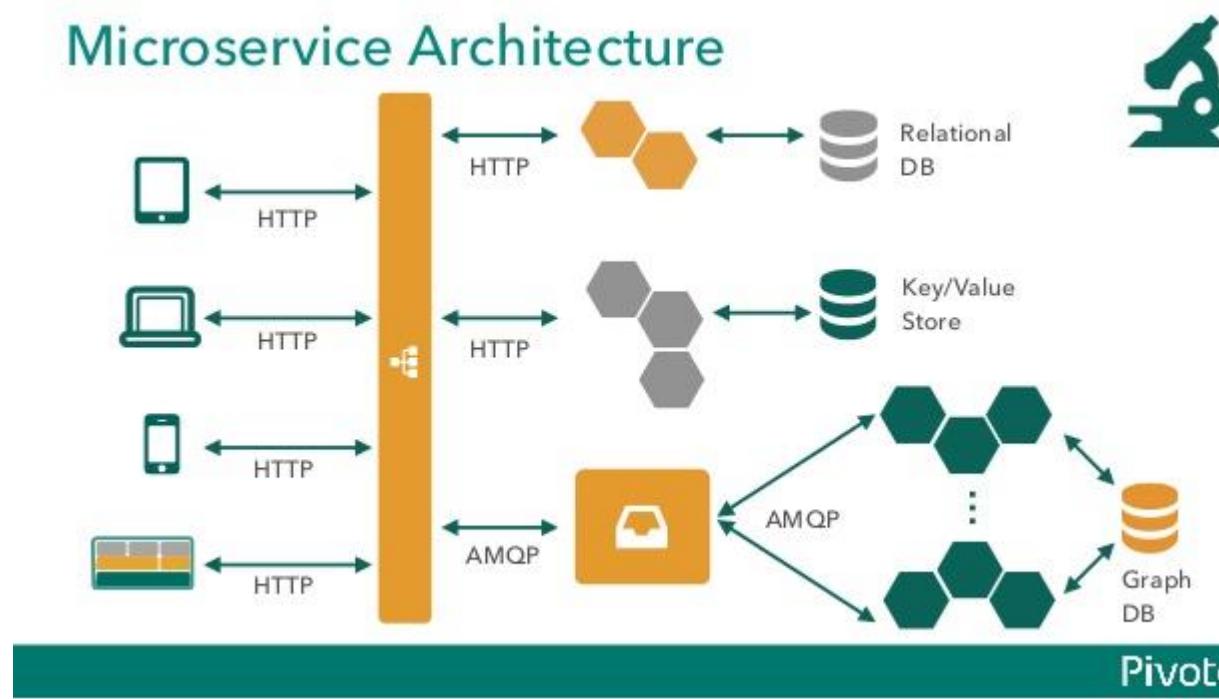
曾令仿、施展  
武汉光电国家实验室  
2019-09-11 至 2019-11-01



# 刚刚在后台发生了什么

- 有一整套服务架构在运行
  - 社交网络、消息队列、关系数据库、图数据库、键值对存储 .....

Microservice Architecture



# 刚刚在后台发生了什么

## ➤ 规模几何

看看腾讯的云



月活跃用户 >8亿  
同时在线用户 >2亿

月活跃用户 >6亿



QQ空间相册



图片 4000亿张 日上传



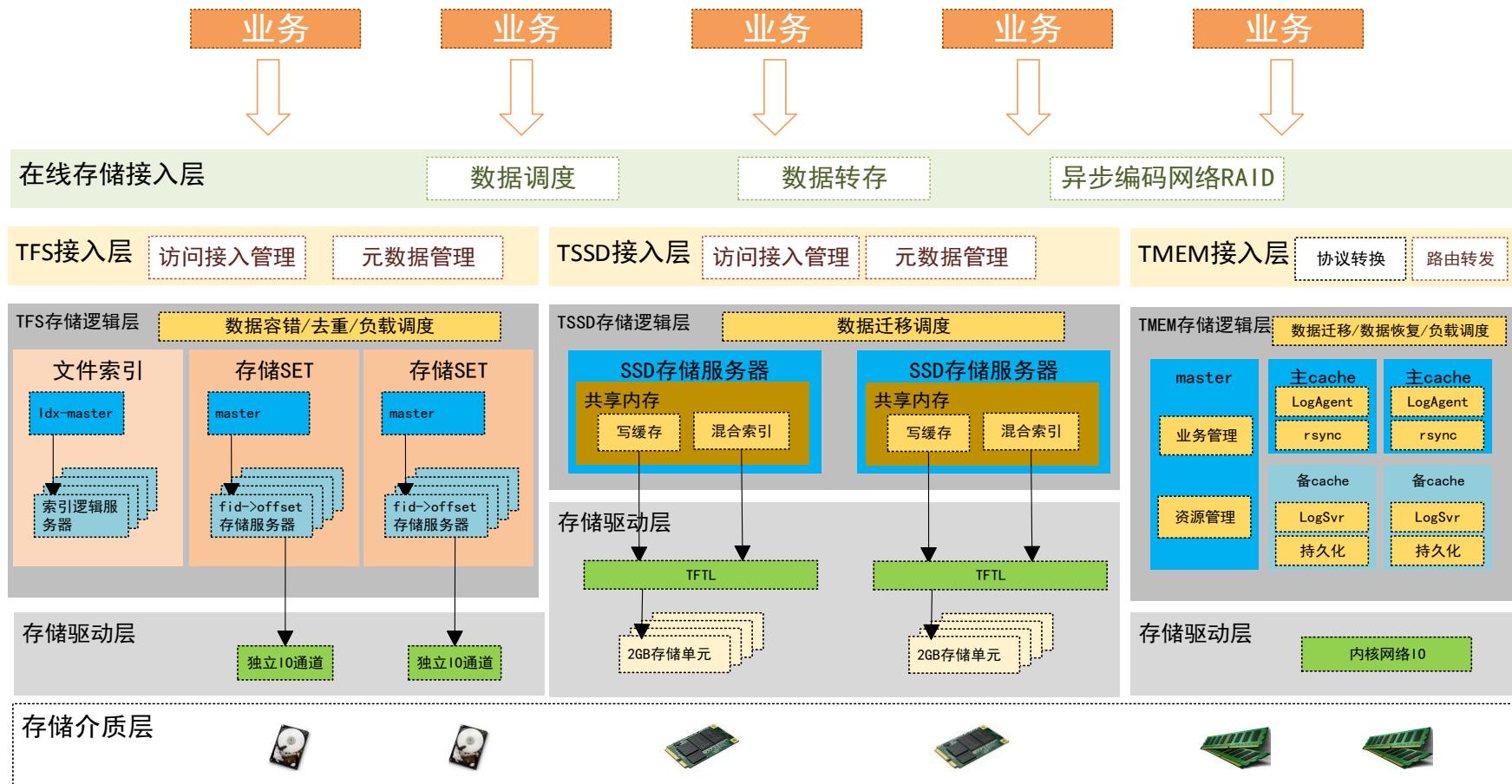
10亿张  
日下载 1200亿张

- 2015年，腾讯云存储已经达到万亿级文件数量，存储量达到500PB
- 2016，存储量达到1EB

- 社交、游戏，访问密度高达100万次/秒/100GB量级的数据读写；
- 在线业务，应保证良好的用户体验，不论数据访问密集程度如何，均要求延迟在100毫秒以内；
- 服务器数量10万级别。

# 是什么在背后支撑

面向多存储介质和多访问密度的  
超大规模在线存储和处理



# 问题并不简单

- 12306
  - 高峰访问10亿PV(Page Visit)，集中在早8点到10点，每秒上千万。
  - 如此高的页面点击量对存储的读写（主要是读）要求非常高，因为高峰期IOPS会超过千万级。
  - 瓶颈在查询环节（读），而非想当然的购票环节（写）
    - 火车票查询业务占12306整个网站流量的90%以上，业务高峰期并发请求密集，性能要求是整个业务系统中最为重要的一环。

# 问题并不简单

## ➤ 阿里参与分析、解决

- 12306的访问峰谷的查询有天壤之别，几乎没有办法在成本和并发能力之间做一个好的平衡。
- 以往的一个做法是从几个关键入口流量控制，保障系统可用性，但是会影响用户体验。
- 通过云的弹性和按量付费的计量方式，来支持巨量的查询业务，把架构中比较‘重’(高消耗、低周转)的部分放在云上。这是一个充分利用云计算弹性的绝好实例。
- 2014年初双方团队就已开始讨论如何将余票查询系统放到云上
  - 十一黄金周期间进行了测试，效果显著。
  - 在春运售票高峰，12306最终将75%的余票查询业务切换到了阿里云上。

# 问题并不简单

- 据美国CNBC报道称，安全公司UpGuard在亚马逊云存储上发现了“被错误设为对公众可见”的机密文件，不过目前已经被移除，但是其中很多资料之前已经被不少网友看到。
- 2013年至2016年雅虎发生过三次泄露事故，10亿用户的个人信息受到影响，不过公司一直隐瞒，等了3年多时间才披露。后来Verizon因为数据泄露压低了收购价格。



# 何谓数据中心

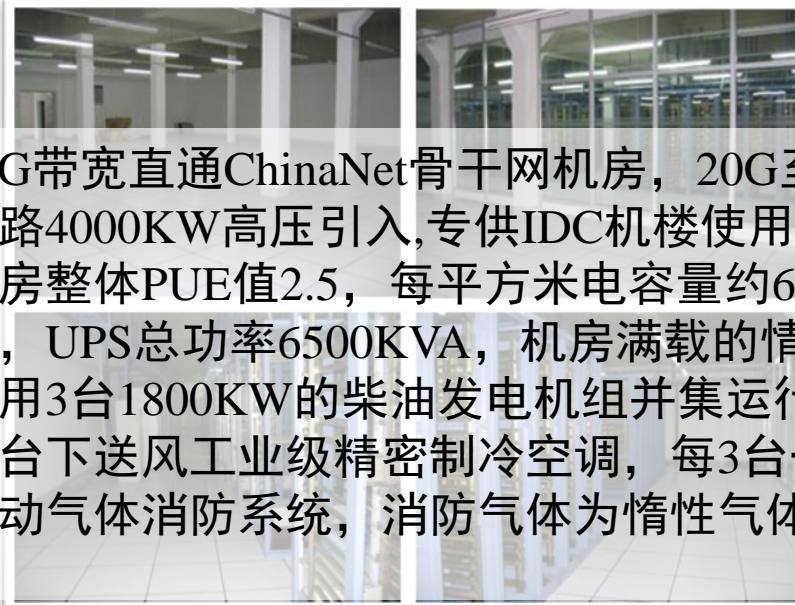
建行武汉数据中心数据机房



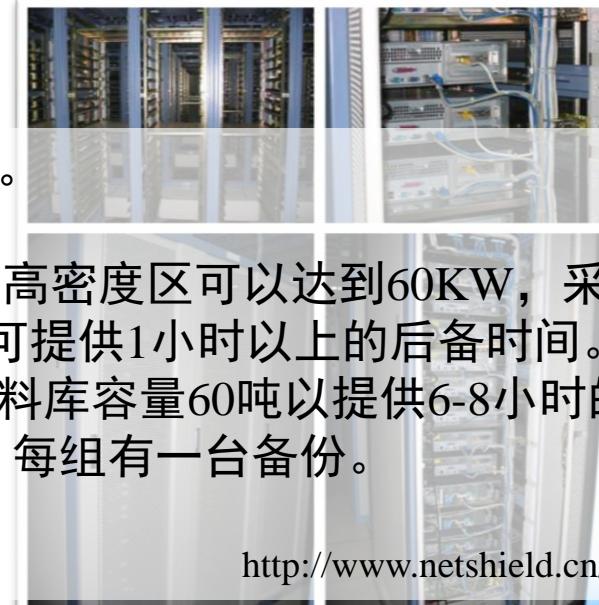
武汉热线IDC机房



# 武汉电信鲁巷IDC机房



- 40G带宽直通ChinaNet骨干网机房，20G至CN2。
- 双路4000KW高压引入,专供IDC机楼使用。
- 机房整体PUE值2.5，每平方米电容量约6KW，高密度区可以达到60KW，采用双路供电，UPS总功率6500KVA，机房满载的情况下可提供1小时以上的后备时间。
- 采用3台1800KW的柴油发电机组并集运行，油料库容量60吨以提供6-8小时的发电时间。
- 18台下送风工业级精密制冷空调，每3台一组，每组有一台备份。
- 自动气体消防系统，消防气体为惰性气体。



<http://www.netshield.cn/about/lx-idc.asp>

鲁巷IDC机房是武汉地区较早的数据中心。1997年6月开始投入使用并不断完善，交通便利，周边无任何产生腐蚀性气体、粉尘、噪音、强震动的厂矿企业,无超高压变电站、电气化铁道、大功率雷达站等强电磁干扰源。

鲁巷IDC机房面积约为900平米。从1997年一直托管和租用到现在的用户也有不少，现有客户，盛大、腾讯、新浪、TOM、千橡都曾落户在这里。

机房总面积900平方米，8级抗震建筑结构。

<http://www.idcquan.com/dianxin/525058.html>

# 何谓数据中心

- 数据中心 Data Center
- 维基百科
  - 数据中心是一整套复杂的设施，不仅仅包括计算机系统和其它与之配套的设备，包括通信和存储系统，还有冗余的数据通信连接、环境控制设备、监控设备以及各种安全装置。
- 谷歌《The Datacenter as a Computer》
  - 多功能的建筑物，能容纳多个服务器以及通信设备。
  - 这些设备被放置在一起是因为它们具有相同的对环境的要求以及物理安全上的需求，并且这样放置便于维护，而并不仅仅是些服务器的集合。

# 何谓数据中心

- 通俗称谓
  - 服务器农场 Server Farm



# 数据中心起源

- 数据中心的概念起源于20世纪50年代末，当时美国航空公司与IBM合作，创建一个属于美国Sabre公司的乘客预定系统，使其主要商业领域的这一部分变得自动化。
- 1960年，数据处理系统的概念成为现实，它用于创建和管理飞机订座系统，让任何地方的任何代理点都可以及时获取电子数据，从此开启了企业级的数据中心大门。

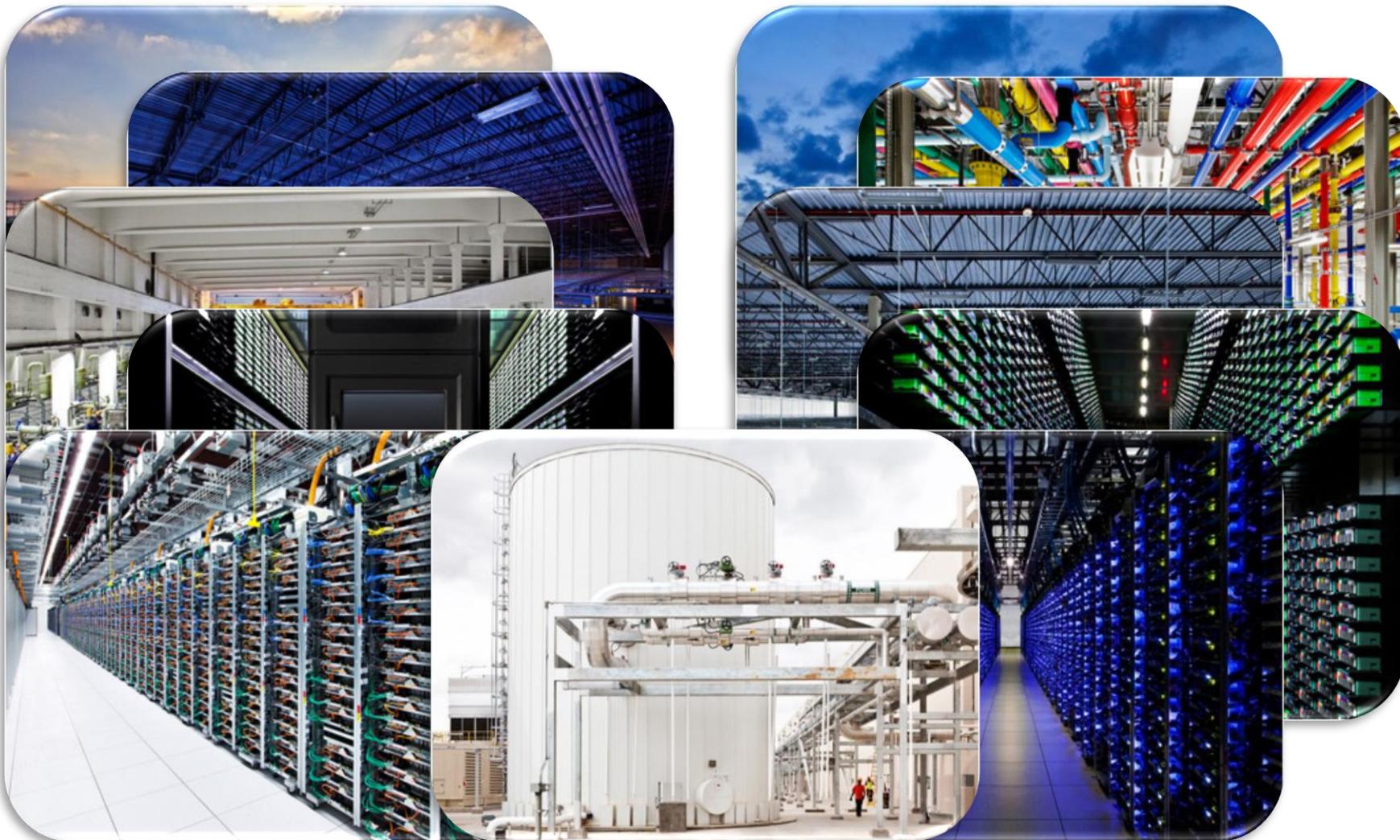
# 数据中心历史

- 1997 : 苹果公司创建了一款名为Virtual PC的程序并通过Connectix公司卖了出去。Virtual PC就像SoftPC一样允许用户在Mac电脑上运行窗口副本，以解决软件不兼容的问题。
- 1999 : VMware公司开始销售类似于Virtual PC的VMware Workstation。最初的版本只能在Windows系统上运行，后来也支持其他操作系统。Salesforce.com开创了通过一个简单网站交付企业应用的概念。
- 2001 : VMware ESX发行，这是一款裸机管理程序，可直接在服务器硬件上运行，无需额外的底层操作系统。
- 2002 : 亚马逊AWS开始发展一套以云技术为基础的服务，包括存储、计算和通过“Amazon Mechanical Turk”实现的人工智能。
- 2006 : 亚马逊AWS开始以网络服务的形式向企业提供IT基础设施服务，现在通常被称为“云计算”。

# 数据中心历史

- 2007 : Sun Microsystems公司采用了模块化数据中心，改变了企业计算的基础经济学。
- 2011 : Facebook发起了开放式计算项目，倡导全行业分享技术参数和实践经验，用以创建最节能、经济的数据中心。
- 2012 : 调查显示38%的企业已经使用云，28%的企业计划开始使用或扩建云。
- 2013 : Telcordia公司发布了有关电信数据中心设备和空间的一般要求。文件提出了对于数据中心设备和空间的最小空间和环境要求。
- 2017 : “数据孤岛” -> 数据中心整合
- 2019 : 边缘（数据中心）-云（数据中心）

# 前沿



# 美国

地图 卫星图像



# 德国

地图 卫星图像





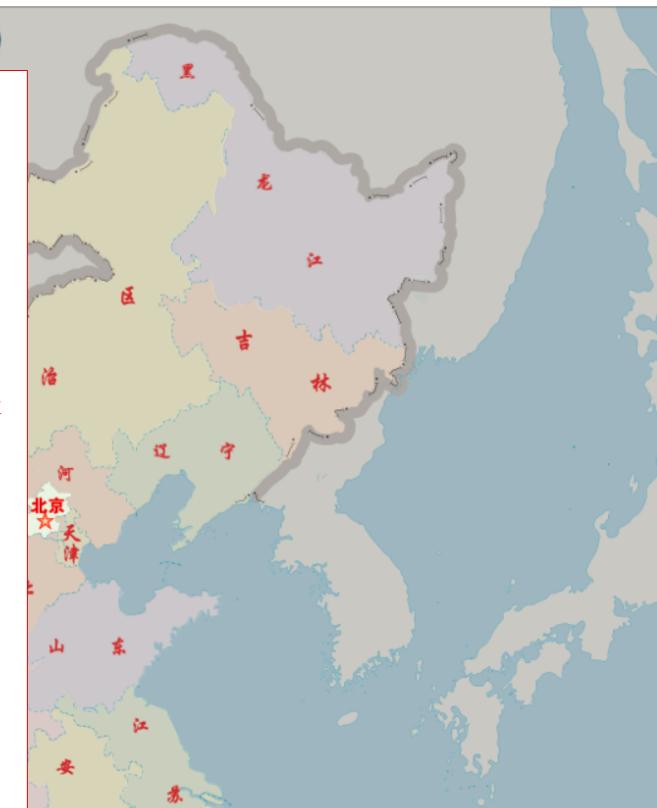
# 中国



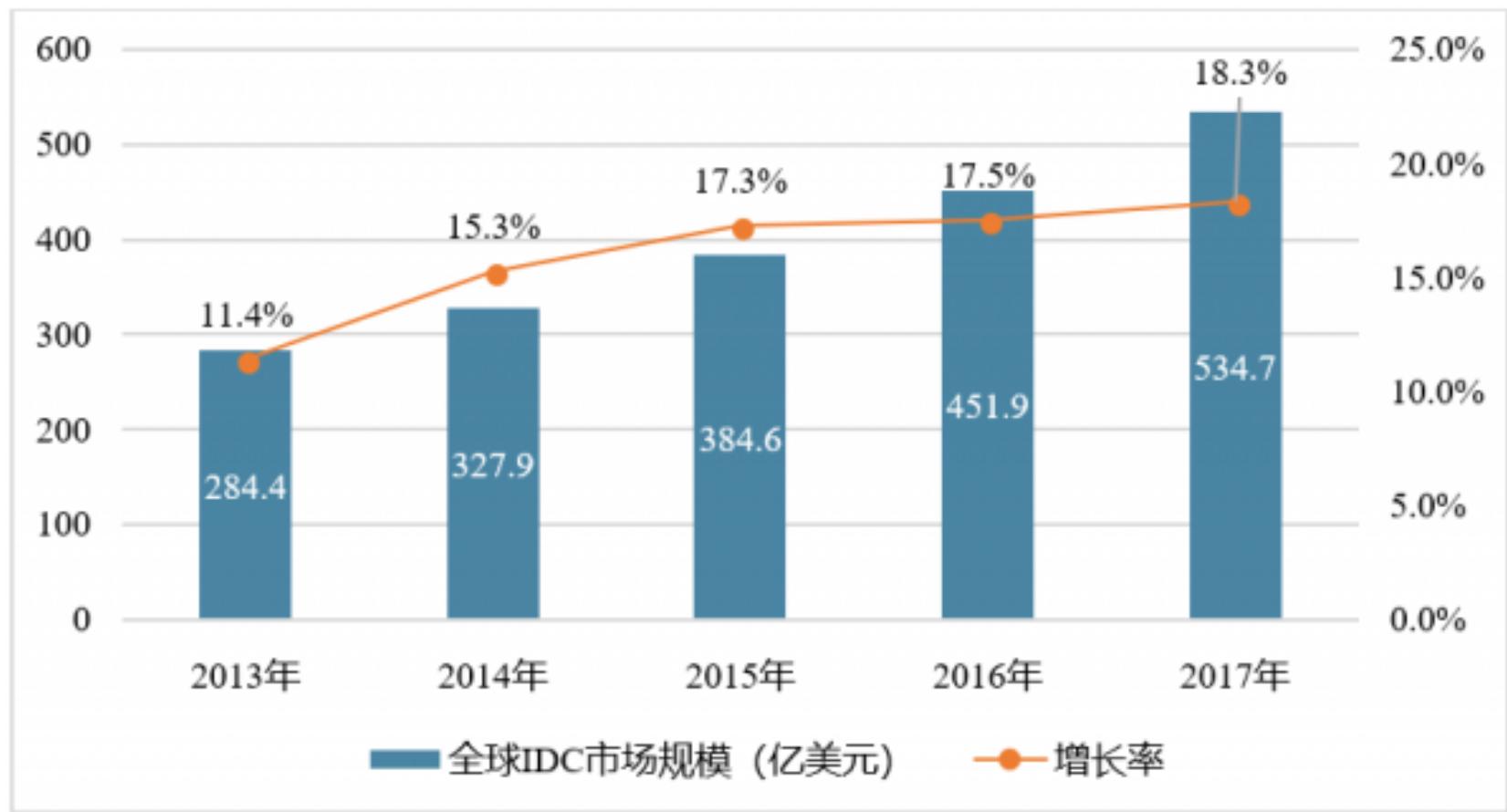
## 全国数据中心分布图

北京 天津 上海 提交信息

<b>中国联通IDC机房</b>	<b>硅谷亮城IDC机房</b>
北京酒仙桥IDC (四星级)	北京市联通鲁谷数据中心 (五星级)
北京联通回龙观电话局数据中心 (三星级)	北京联通中关村1+1大厦数据中心 (三星级)
数字北京大厦IDC (五星级)	北京顺义区林河联通数据中心 (三星级)
北京亦庄电话局IDC (四星级)	北京西三旗电话局IDC (四星级)
北京西红门电话局IDC (四星级)	北京石景山电话局IDC (四星级)
北京上地电话局IDC (四星级)	北京南苑电话局IDC (四星级)
北京广内电话局IDC (四星级)	北京东四电话局IDC (四星级)
北京联通京门大厦多线机房 (四星级)	北京五棵松联通机房
北京北苑电话局数据中心 (三星级以下)	北京联通长途电话局机房
北京龟君庙IDC (五星级)	北京土城IDC (五星级)
北京联通三元桥数据中心 (五星级)	北京联通电银大楼数据中心 (五星级)
北京分公司亦庄国际IDC (五星级)	北京联通方庄机房
<b>Newt</b> 北京定福庄电话局IDC (四星级)	<b>Newt</b> 蓝汛MIDS北京通州东古城数据中心 (五星级)
<b>中国电信IDC机房</b>	北京电信上地机房
中国电信酒仙桥双线机房	中国电信静安里机房
中国电信西二旗机房	263数据港机房
<b>第三方IDC机房</b>	企商在线华威桥数据中心
企商在线广渠门数据中心	北京电信通雅和宫机房
北京光环新网酒仙桥机房(T4标准)	北京鹏博士数据酒仙桥 (NGDC) 机房
北京鹏博士数据中关村机房	北京鹏博士数据苏州桥机房
<b>数字化机房</b>	中关村软件园数据中心 (五星级)
中关村硅谷机房 (三星级)	网联无限IDC数据中心
北京鹏博士数据普惠大厦IDC机房	北京电信通三元大厦IDC机房
北京总部基地IDC机房	北京通管网联马连道IDC机房
北京中央电视塔双线机房	北京国研双线机房
北京鼓楼双线接入机房	北京石景山科技馆双线机房
北京铜牛四线机房	光环新网东直门IDC数据中心
歌华多媒体数据中心 (MDC)	中电华通工体机房

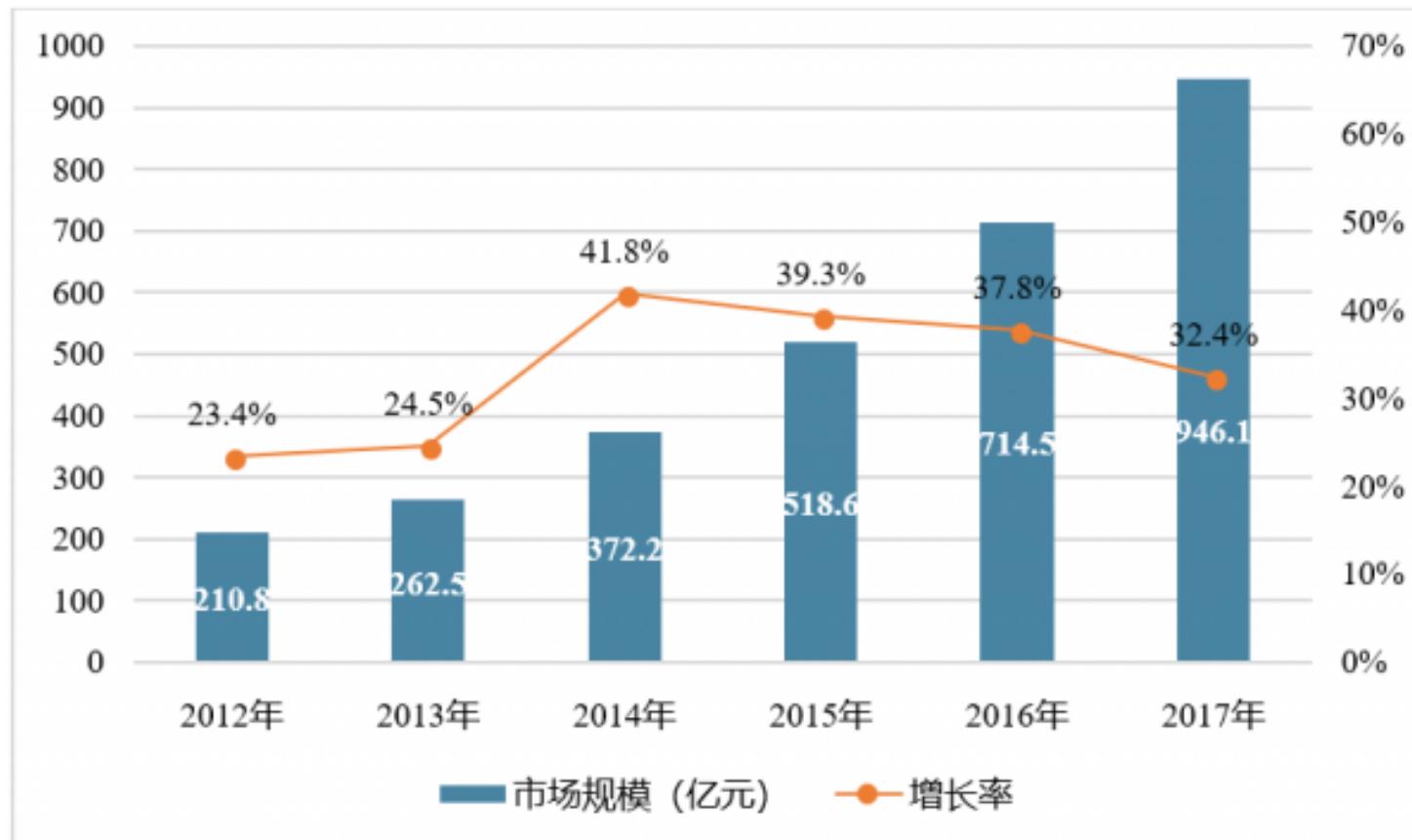


# 发展



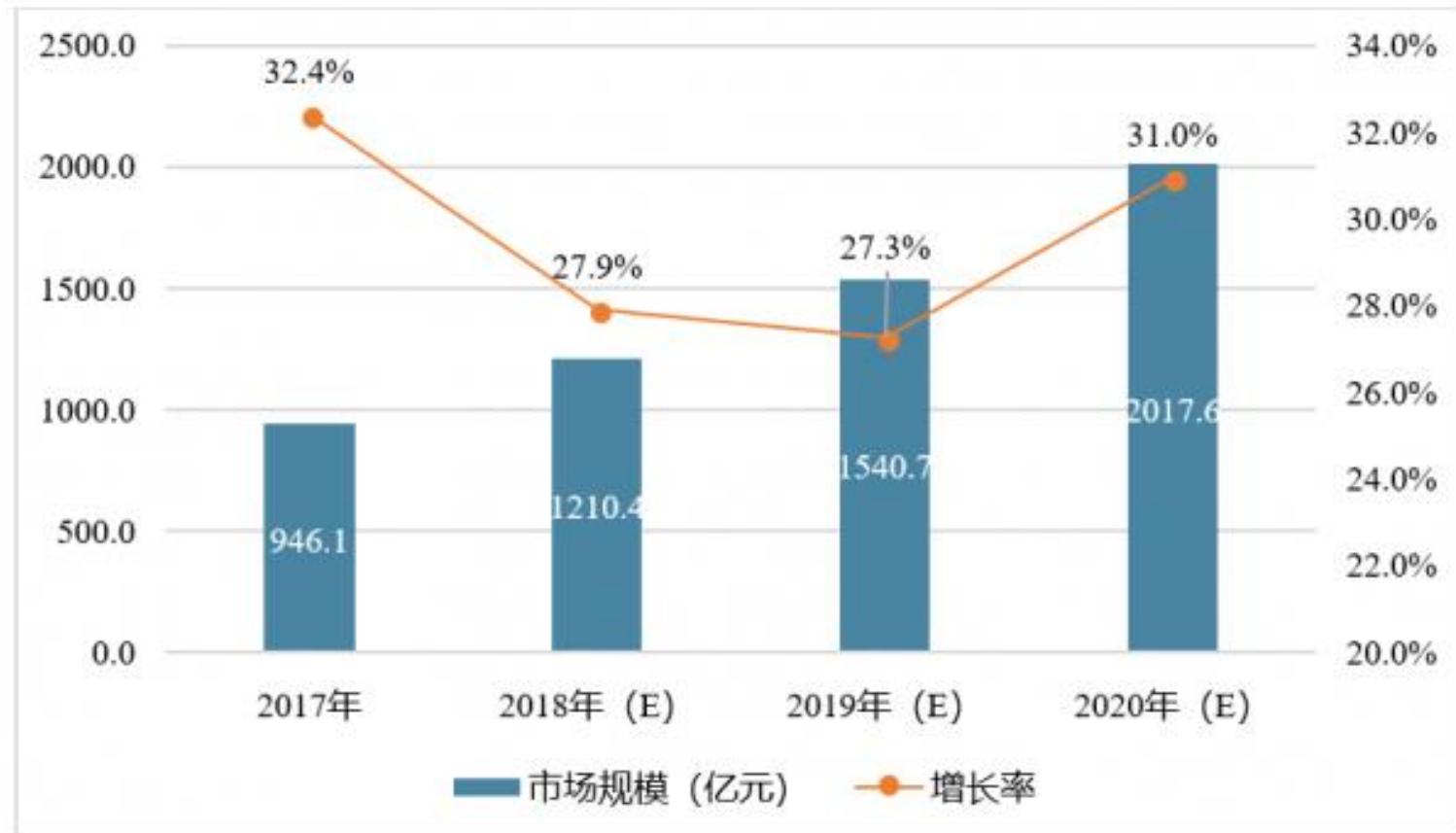
数据来源：科智咨询（中国IDC圈），2018.02

# 发展



数据来源：科智咨询（中国IDC圈），2018.03。

# 发展



数据来源：科智咨询（中国IDC圈），2018.03.

# 案例 - Google Data Center



# 案例 - Google Data Center



Singapore



Lenoir, North Carolina



Hong Kong



Hamina, Finland

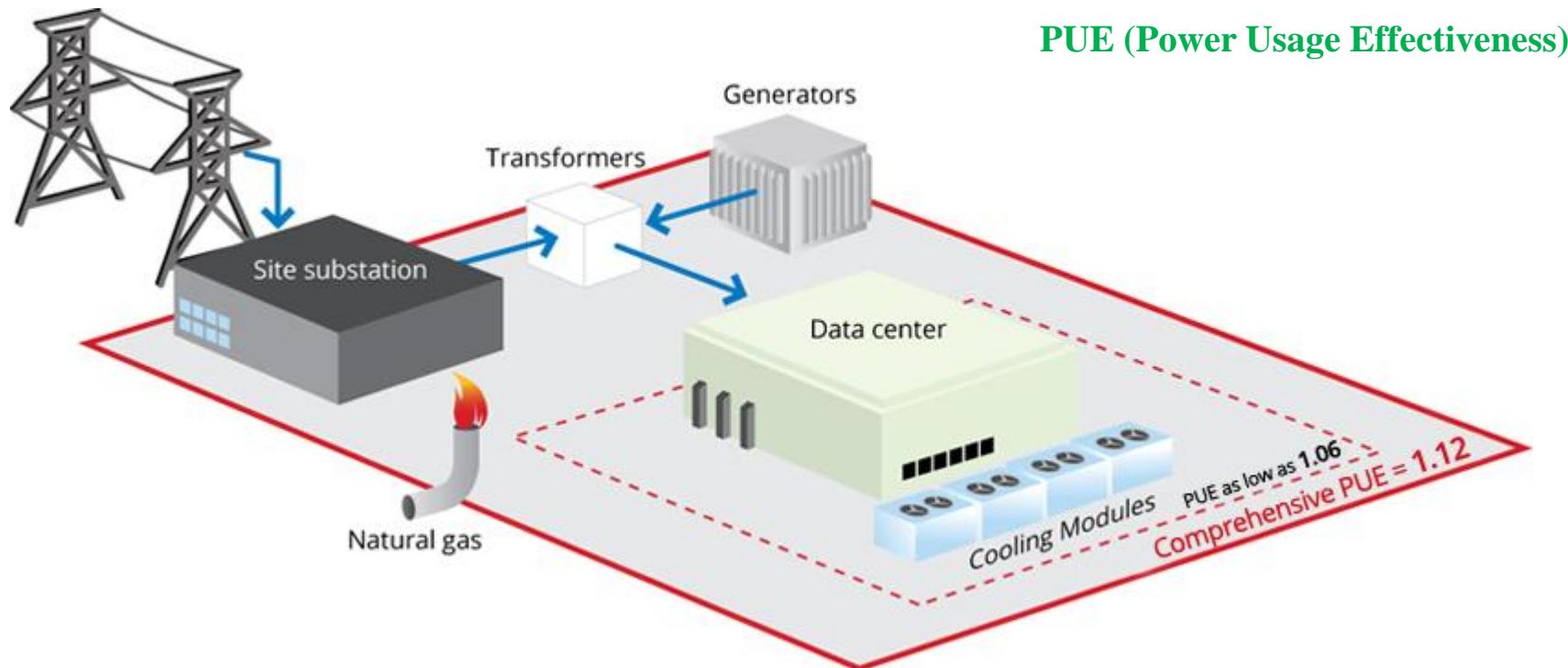


The Dalles, Oregon



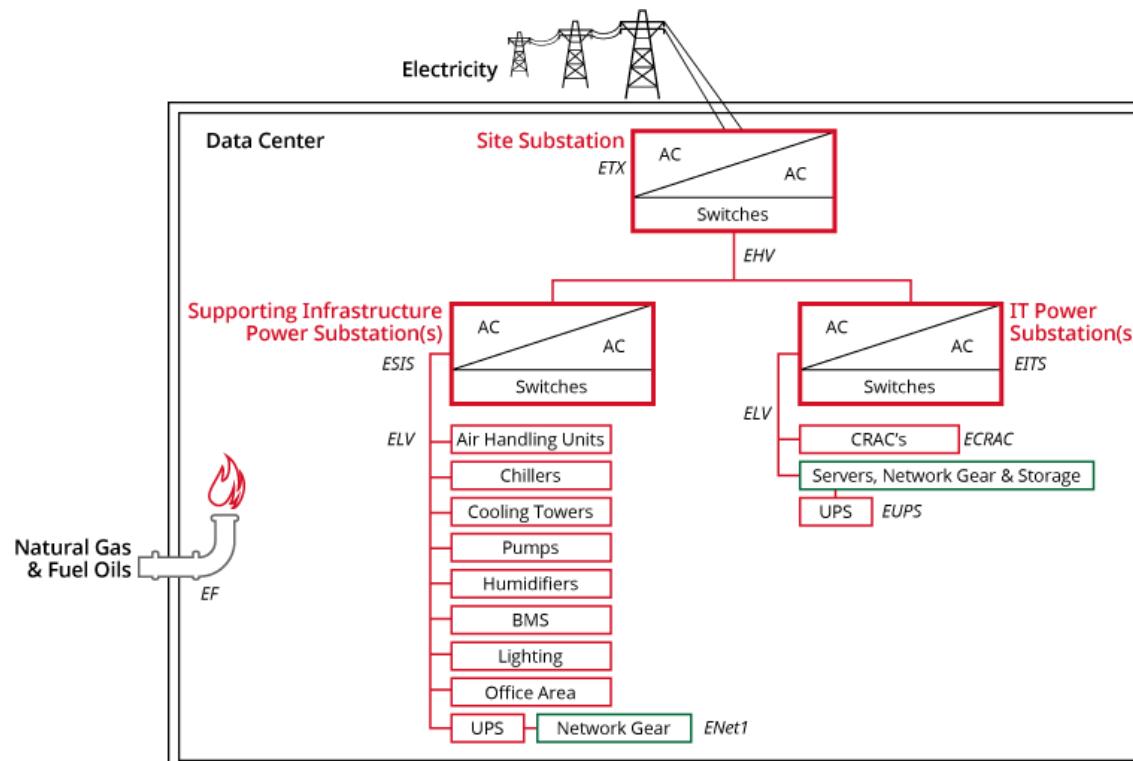
Council Bluffs, Iowa

# 案例 - Google Data Center

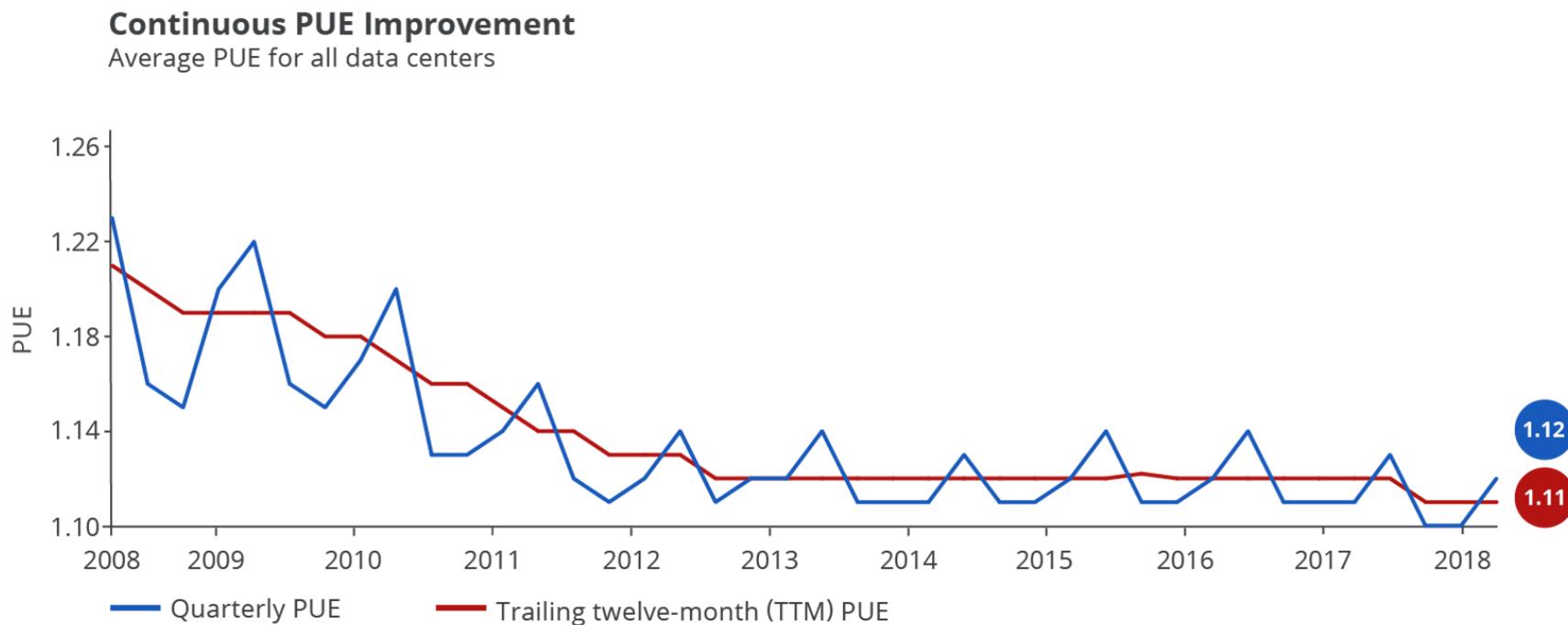


$$PUE = \frac{ESIS + EITS + ETX + EHV + ELV + EF}{EITS - ECRAC - EUPS - ELV + ENet1}$$

# 案例 - Google Data Center



# 案例 - Google Data Center



# 案例 - Microsoft Data Center



# 案例 - Facebook Data Center

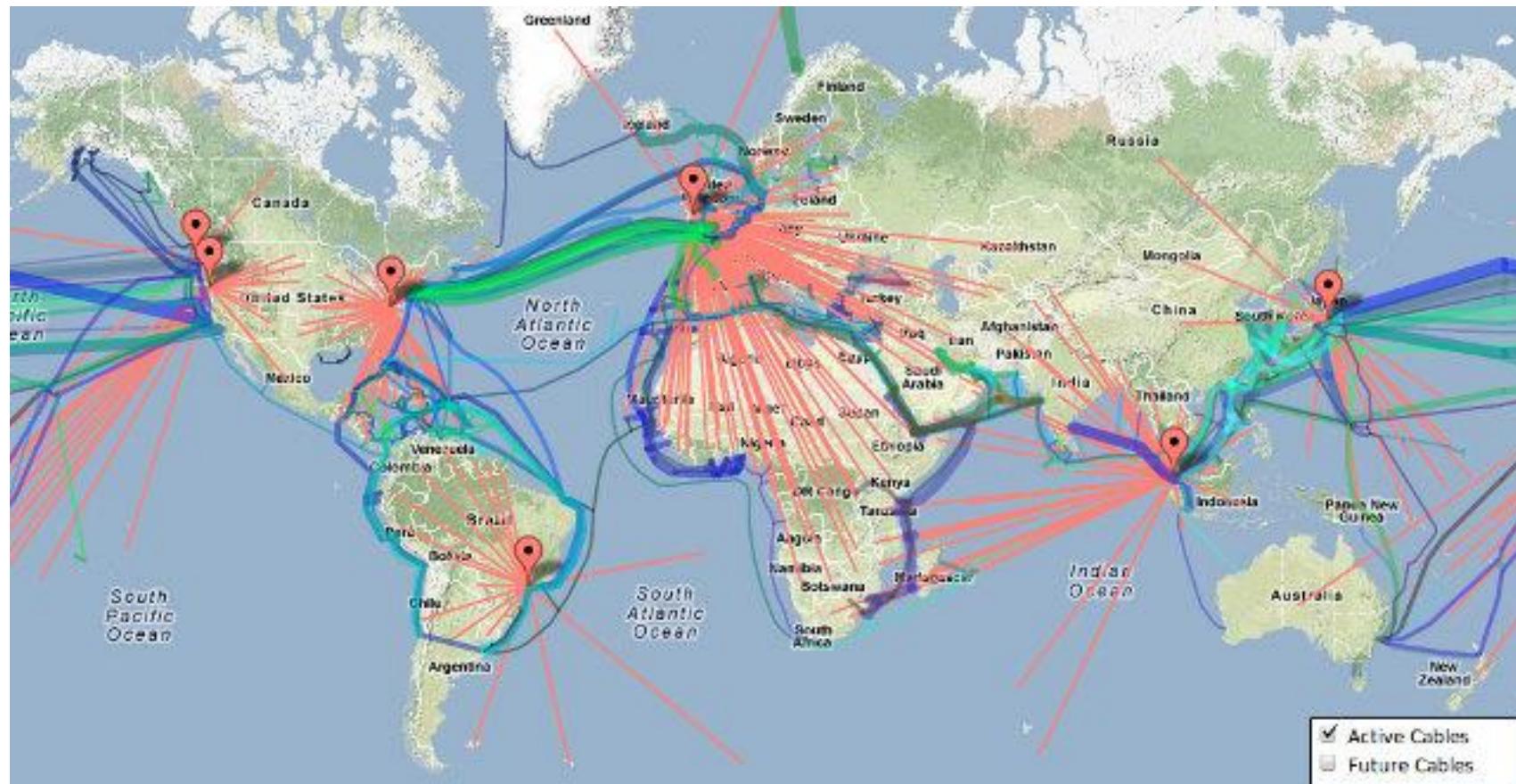
Altoona, Iowa; Prineville, Oregon; Forest City, North Carolina; and Lulea, Sweden.



# 案例 - Apple Data Center



# 案例 - AWS Data Center



# 案例 – 阿里巴巴



千岛湖数据中心全景

阿里云新一代绿色数据中心，设计年均 PUE 低于 1.3

千岛湖数据中心因地制宜采用了湖水制冷。深层湖水通过密闭管道流经数据中心，帮助服务器降温，再流经2.5公里的青溪新城中轴溪，作为城市景观呈现，自然冷却后最终回到千岛湖。得益于千岛湖地区年平均气温17度，其常年恒定的深层湖水水温，阿里方面称可以让数据中心90%的时间都不依赖湖水之外的制冷能源，制冷能耗节省超过8成。

# 等一下

- 前面的案例聚焦于哪几个方面？
  - 规模、成本
    - 域内扩展、跨域
  - 效率
  - 环境
- 然后呢？
  - 供电！



# 再谈谈PUE

## ➤ 2006年以谷歌为首

➤ 评估数据中心总用电量的衡量指标主要是：

- 电力使用效率 ( PUE )
- 主要为考察资源高效而建立的指标

➤ 但是

- PUE 只考虑数据中心的内部操作，未揭露电力来源与实际用电量
- Google声称其数据中心只占全球用电量的0.01%，但并无公布实际千瓦时。

# 算一算

- 若以2012年全球用电量为197.1亿千瓦时计算，**Google**数据中心用电量相当**土耳其**每年用电量。

Google



- 2013年可持续发展报告中，**Facebook**数据中心使用9.86亿千瓦时的电力，相当于**布吉纳法索**用电量。

facebook



# 算一算

## ➤ 绿色和平组织几年前揭露

- Facebook 数据中心用电来源超过一半是燃煤发电
- 当时 Facebook 强调他们的电力也是来自电网，社会大众应关心数据中心用电效率。

PUE



## ➤ 现在Facebook数据中心完全使用再生能源发电



苹果的北卡罗来纳州数据中心的电能来自于两台20MW的太阳能电池阵列，一台18MW的太阳能电池阵列和10MW的沼气燃料电池。



苹果公司与当地杜克能源公司合作建设五个太阳能项目，总峰值容量为20MW。在晚上和在一年中较冷的月份，该数据中心直接使用外部空气进行冷却，此举可以保证在大部分时间关闭冷水机组。



加州数据中心使用的能源，主要来自加利福尼亚州的风力发电场，通过与加州政府制定接入计划将电力直接接入数据中心。该数据中心自2013年以来一直以100%的可再生能源进行供电。

位于加州蒙特利市的130MW California Flats太阳能项目在2017正式交付使用，苹果公司直接把太阳能发电厂的电力接入纽瓦克数据中心。

苹果公司与俄勒冈州一个新的风力发电场Montague签署了200MW的购电协议。该发电厂每年产生的电能超过5.6亿千瓦时。Montague风电场预计将在今年年底前正式交付使用。

苹果还与距离数据中心数千米的俄勒冈二号光伏阵列签署了一项电力采购协议。该项目在2016年底已交付使用，每年可生产1.4亿千瓦时的电能。



# 阿里巴巴千岛湖数据中心

- 数据中心外部有两台湖水处理器
  - 通过密封管道从湖中取水
  - 取水口符合环保标准，选择中间层，既不会太深避免泥沙等问题，也不会太浅有较多浮游生物。
  - 取水层基本无杂质，水温基本在13度左右。
  - 经过缓冲池进入湖水处理器，基本不需要进行水处理。
- 高压柴油发电机作为应急电力设备
  - 发生突发情况，常规油管储备可以保证8-10小时供电。
  - 阿里会人为断电+监控的方式，实现负载状态的实时柴油发电机的电力切换测试。

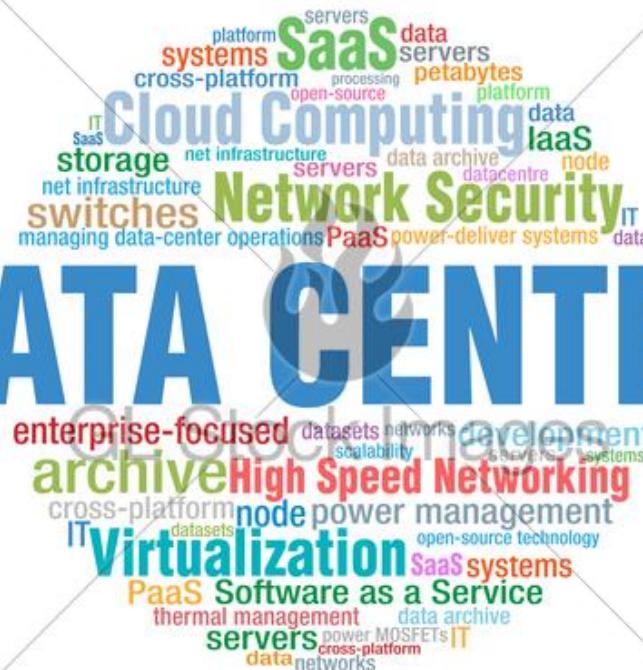


# 阿里巴巴千岛湖数据中心



慢着

- ## ➤ 好像有不少术语



# 还有

- 不少相关项目



# 背景调查

- 装机
  - 裸机装系统
  - 自己用零件装配台式机
- Linux
  - 只用桌面
    - GNOME, KDE
  - 会用命令行模式
    - bash, zsh, fish ...
  - 用过5个以上GNU工具
    - grep, sed, awk, sort, uniq, ps, top, watch, nc, jobs ...
  - 远程使用主机
    - ssh, mosh
  - 拥有VPS、云主机

# 背景调查

- C/C++
  - 作业程序
  - 多模块工程
  - 知名程序库
    - STL, Boost, OpenGL, CUDA, MPI ...
- Java
  - 作业程序
  - 知名工程
    - Hadoop, Spark, Giraph, Storm ...
  - 构建工具
    - Ant, Maven, SBT ...
- Python
  - 一般脚本
  - 管理过系统、运行过网站
  - 跑过数据分析、机器学习任务
    - Tensorflow, torch, caffe ...

# 平台技术：容器

- 专业语言
  - R, Matlab
- 流行语言
  - JS, Ruby
- 另类语言
  - LISP, Closure, Scala, Haskell, Erlang

# 背景调查

## ➤ 专业网站

- Stackflow
- Github, Bitbucket
- Linkedin
- LeetCode
- Release Mirrors, Vagrant Cloud, Docker Hub

# 后续内容

- 泛读论文 USENIX ATC2019, HPCA2019, SC2019, ASPLOS2019
- 了解环境 Linux, Git, SSH, Python, OpenStack, K8S, Docker ...
- 作业分组 根据班级选课人数，安排8组以内
- 预选方案 2周后公开，亦可结合兴趣提出本组实验设计
- 讲解 “大规模高性能分布式块存储系统数据中心部署实例”