

•业界动态

2016.05.17

韩岳岐

hanyueqi@cnic.cn

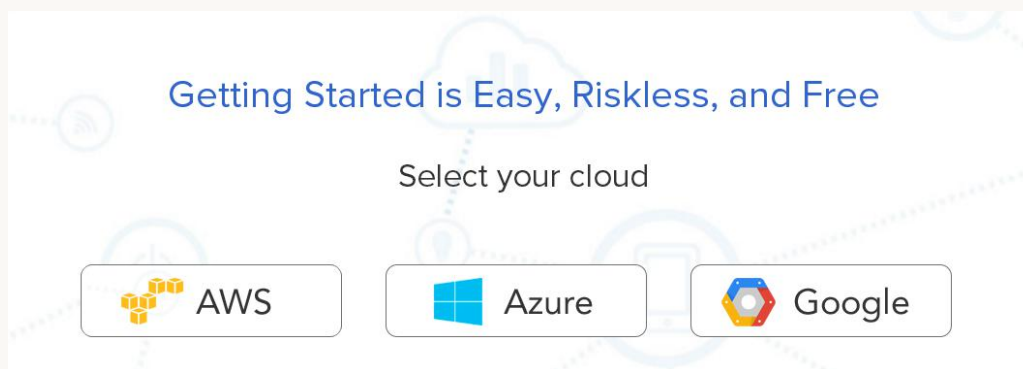


Enterprise Data World

- April 17-22, 2016
- <http://edw2016.dataversity.net/index.cfm>
 - 云中巨象——作为一种服务的Hadoop
 - 使用NoSQL八年经验总结

云中巨象——作为一种服务的Hadoop

- Ashish Thusoo, Qubole
- 数据性质
 - 海量交互数据，非传统结构化数据
- 分析内容，描述性->规范性
 - 描述性分析（发生了什么？）
 - 诊断性分析（为什么发生？）
 - 预测性分析（将会发生什么？）
 - 规范性分析（我们怎样才会让它发生？）
- 云端计算系统
 - 多租户数据访问
 - 引擎统一
 - 云计算服务
 - 可移植性服务
- 整合框架
 - 复杂批次处理SQL，Hive
 - 数据科学，Spark
 - 简单交互SQL查询，Presto
 - 批量数据清洗（ETL），Map Reduce



使用NoSQL八年经验总结

- Mike Bowers, 耶稣基督后期圣徒教会 (LDS)
- 应用场景
 - 189个应用, 几十亿事务
 - 1500万会员, 188中语言
 - 192个网站, 数十亿PV
- NoSQL实施策略
 - 经验1: 每个组织需要一个NoSQL champion
 - 经验2: 必须得到管理层的支持
 - 经验3: 必须得到开发者的支持
 - 经验4: 培训, 培训, 培训

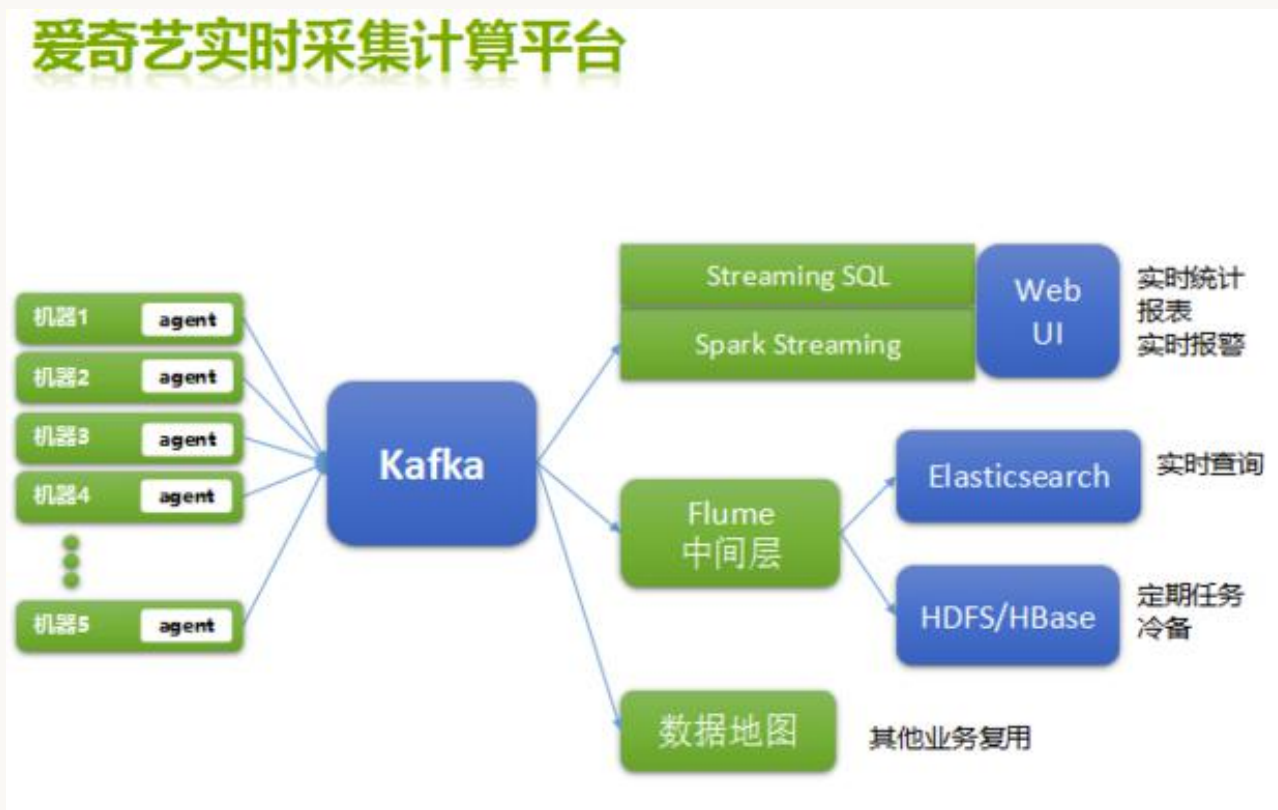


DTCC2016中国数据库技术大会

- May 12-14, 2016
- <http://dtcc.it168.com/>
 - 奇艺实时采集计算平台的功能架构
 - 大数据管理系统的发展与机遇

奇艺实时采集计算平台的功能架构

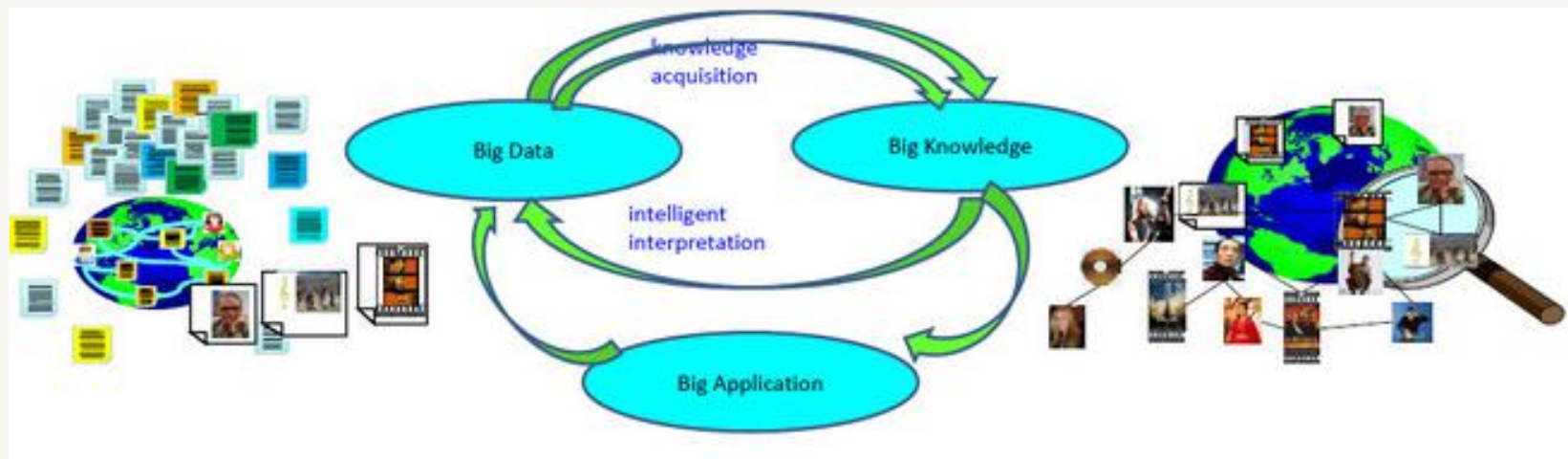
- 叶炜
- 互联网日志需求
 - 快速故障定位
 - 实时统计分析
 - 实时日志报警



大数据管理系统的发展与机遇

- 孟晓峰
- 大数据管理系统
 - 运营式系统--数据源是被动产生：数据规范，有秩序，强调数据的一致；
 - 互联网系统--数据源是主动产生：数据结构复杂，无秩序，不强调数据的一致性或只强调弱一致性；
 - 感知式系统--数据源是自动产生：数据呈现多源异构、分布广泛、动态演化等。
- 大数据管理（数据管理技术+数据库技术）
- 大数据应用三要素
- 大数据管理系统分类

大数据管理系统的发展与机遇



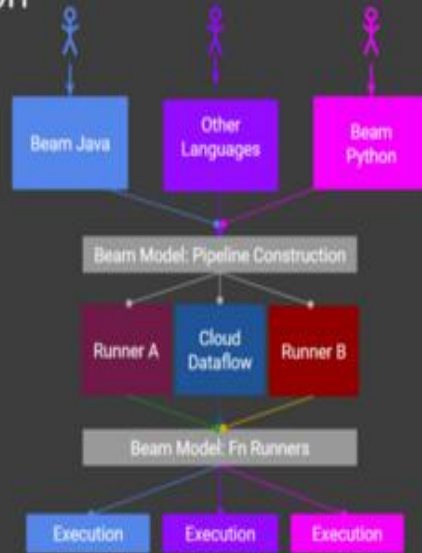
图数据管理系统	Prege、Giraph、PowerGraph、GraphChi、XStream、Giraph+	数据的“广度”和“深度”
流数据管理系统	S4、Storm、Puma、Samza	数据的“密度”
众包数据管理系统	Mturk、CrowdFlower、samasource、CloudCrowd	数据的“广度”
在线分析管理系统	Dremel、Drill、Impala、BlinkDB	数据的“广度”
商业数据管理系统	Oracle、MySQL、SQL Server、MongoDB、HBase、Cassandra、Redis、Spanner、OceanBase、RDS、Cloud SQL、Azure、RDS、BigTable、DynamoDB、SimpleDB	数据的“广度”

Apache Beam

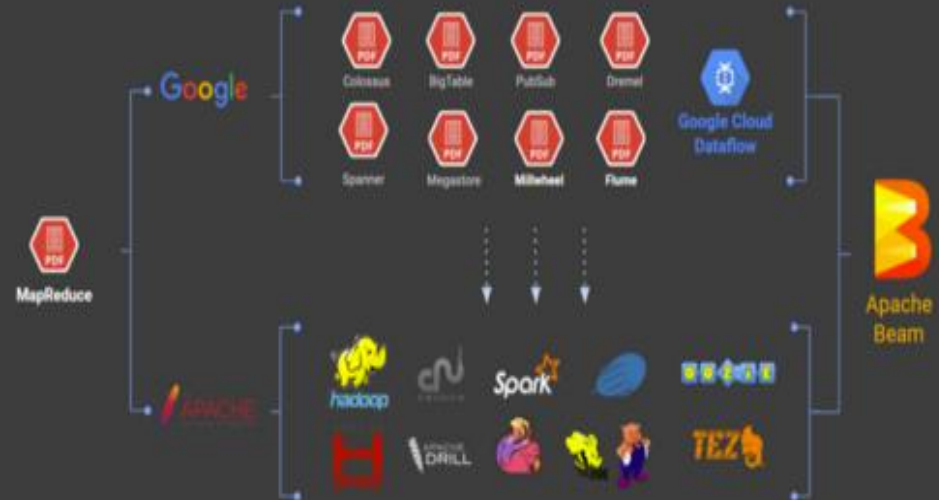
- Google-DataFlow
- Incubator

Apache Beam Technical Vision

1. **End users:** who want to write pipelines in a language that's familiar.
2. **SDK writers:** who want to make Beam concepts available in new languages.
3. **Runner writers:** who have a distributed processing environment and want to support Beam pipelines



The Evolution of Apache Beam



海普洛斯完成5000万元A轮融资 发起“万人癌症基因测序计划”

- 国家超级计算中心深圳云计算中心
- 癌症基因检测和健康人群基因检测
 - 通过对癌症患者的血液或者癌症组织的样本进行基因检测，并为其提供靶向治疗、预后监测
 - 通过分析受检者的遗传基因，给出其可能患有遗传性疾病，特别是遗传性肿瘤的风险评估
- 超微量肿瘤液体活检
- 截至目前，已有30多家全国大型医院或科室加盟“万人癌症基因测序计划”，并且已经完成了将近5000例肿瘤患者或高危人群的基因检测。

网络内容推荐引擎Outbrain获4500万美元融资

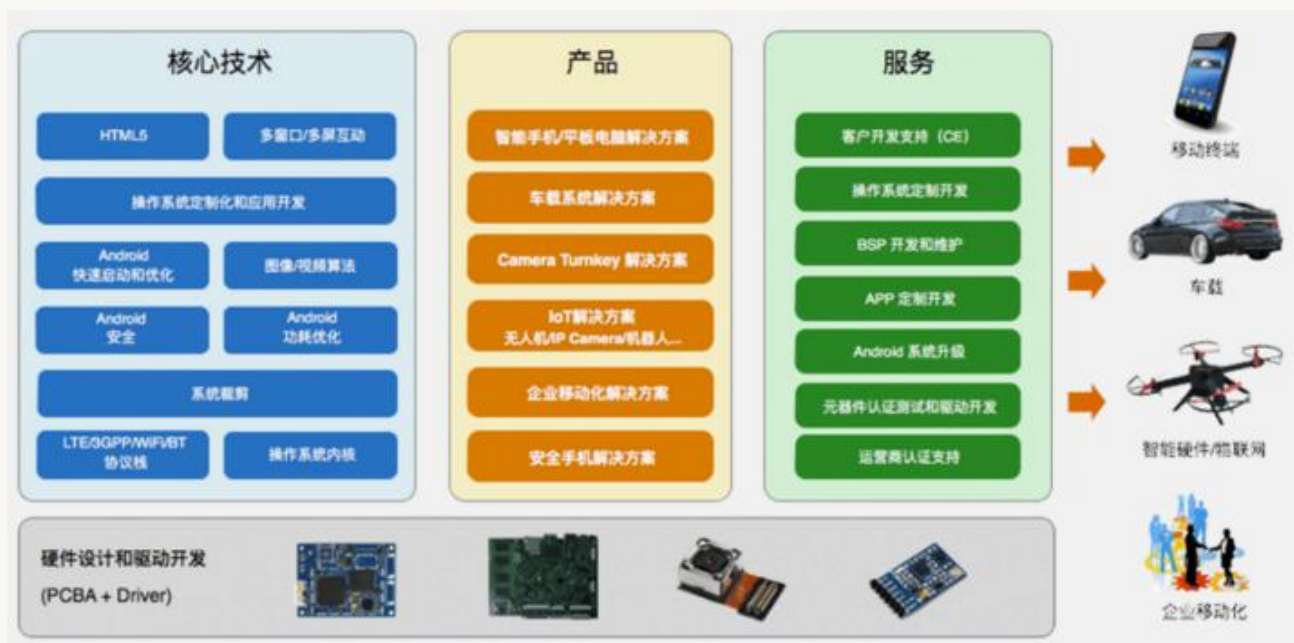
- 个性化的网络内容推荐引擎
- Outbrain 会依据读者的兴趣，在文章末尾呈现一系列的推荐阅读内容
- CNN、ESPN、Le Monde、福克斯新闻、卫报、Slate、The Telegraph 和纽约邮报
- Outbrain 服务的网站超过 10 万家，付费用户近千家。Outbrain 每个月的内容推荐条数超过 2000 亿条，触达全球 5.61 亿的独立访客



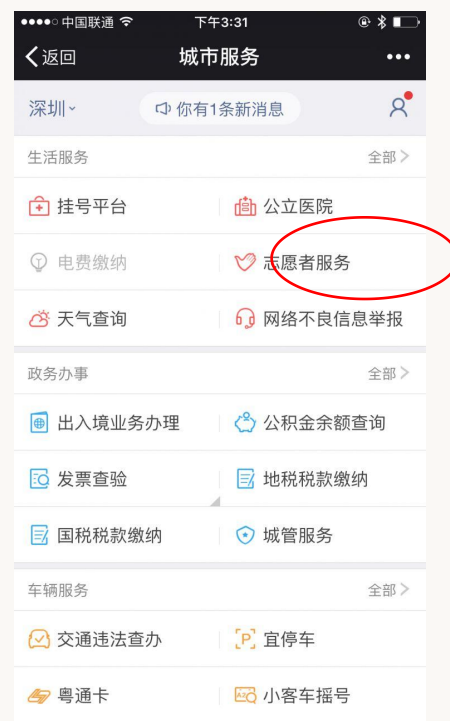
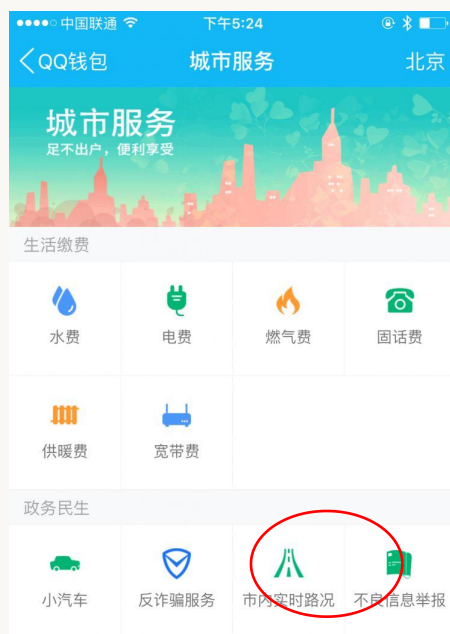
中科创达发布骁龙820VR一体机参考设计

产业链上游发力拉低VR一体机门槛

- 基于骁龙 820 VR 一体机参考设计,包括屏幕、光学、到 PCBA、散热设计等多方面,同时搭载了自主研发的 VR OS,针对底层驱动及能效比优化、延时畸变动作追踪算法、Android M 裁剪、优化及上层应用和 launcher 等的全套 turnkey 解决方案。



腾讯携手12321试水互联网+城市服务新模式



互联网产品、大数据、云计算这几年真的是这么过来的

- 产品/服务层
 - 只注重产品
 - 只注重数据
 - 注重产品和数据的均衡
- 大数据层
 - Player类型
 - 平台类型
 - 专用工具型
 - 咨询服务类型
- 云计算层
- 价值应用层

