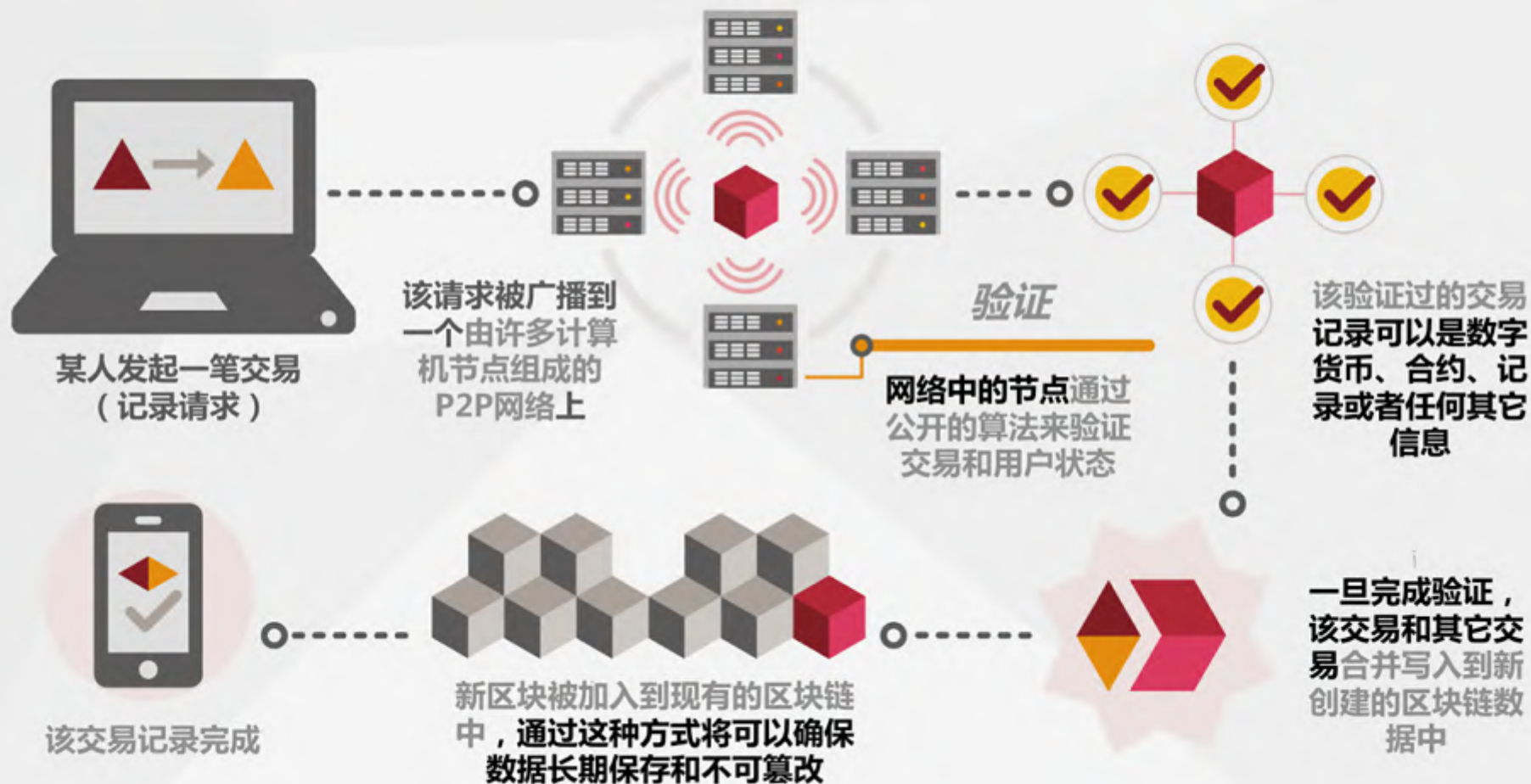


# 区块链与大数据

龚鸣



# 什么是区块链技术



# 区块链技术的优势

## 01 提升透明度

每个节点有相同的副本  
因此数据公开透明

## 02 精确追踪

区块链上的链式结构  
数据可完整回溯

## 03 长期保存

极高冗余使得系统  
具有极强的健壮性

01

06

02

05

03

04

## 不可篡改

除非同时控制大部分  
节点否则无法篡改数据

## 提升效率

无中介自动执行  
将极大提升效率

## 降低成本

无需中介可以极大  
下降管理成本



# 全新的数据库形式

## 关系型数据库

如Oracle, MySQL, SQL SERVER等



## 非关系型数据库

如HBase, Cassandra, MongoDB等



## 时间轴数据库 (区块链)

以比特币区块链为代表的区块链数据



## 已整理数据

可以进行数据管理和分析，  
使信息管理系统开始流行

## 海量数据管理

能够对海量级别数据进行管理和分析，使得  
谷歌、阿里巴巴这种世界级网站成为可能

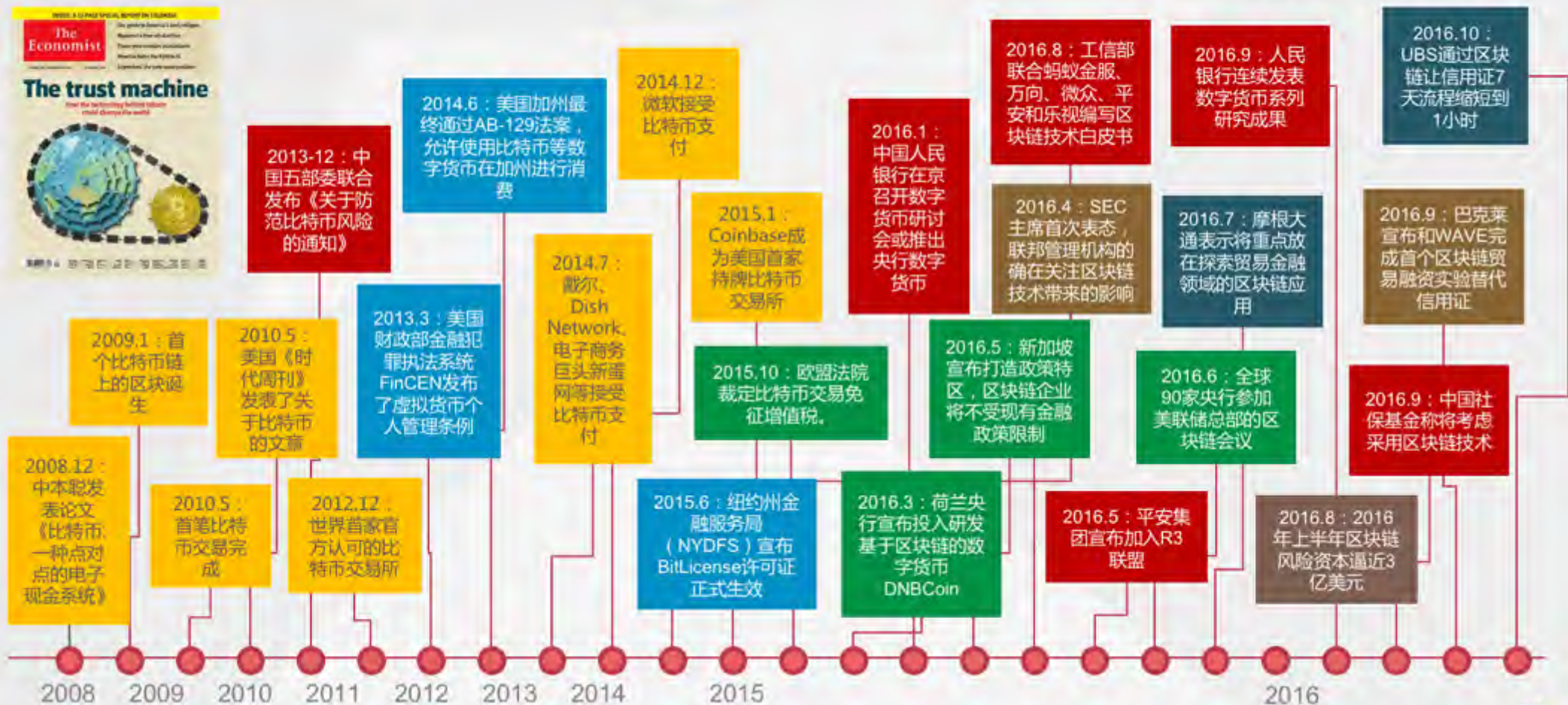
## 极高安全性数据

由于无法篡改和去信任机制，使价值转移  
成为可能，从互联网金融转向金融互联网



# 区块链技术在全球高速发展

\* 所有新闻均来自铅笔采编



# 区块链技术在各领域的应用

审计、基金、信托、投票、治理、资产证券化



## 支付结算

无中介的支付汇  
款与结算清算



## 物联网

自组网且自治  
的物联网方案



## 证券交易

无中心交易所的  
证券发行交易



## 供应链

供应链溯源和  
供应链金融



## 医疗健康

电子病历存证保  
管与药品溯源



## 电子商务

无中心的电子商  
务网站交易系统



## 去中心保险

自动执行的保  
险智能合约



## 分布式存储

无中心化的分  
布式存储系统

# 区块链

# BLOCKCHAIN

游戏、社交、存证、鉴定、身份识别、人工智能



## 区块链上的大数据

当比特币的区块链已经超过数百G的时候，对于很多人而言已经无法了解这里巨大数据的含义。当所有交易都保存在区块链时，这数据本身就是一个巨大的金矿，如果能够了解其中的含义，无论对于区块链的开发者还是投资者都有着巨大的意义。



随着各类区块链系统的越来越多，区块链上的数据已经出现了爆炸式的增长，我们认为对于数据的解读将会变得越来越重要，也会变得越来越困难。因为对于海量数据的存储和查询已经不是个人开发者和投资者能够完成的事情。但是却有大量的个人开发者和投资者对于数据结论有巨大的需求。

## OBSERVER观察者

OBSERVER是全球首个完全基于区块链技术的分布式大数据查询和商业系统。

每个人通过电脑或者手机，都可以很方便的  
对区块链上的海量数据进行查询，也可以直  
接查看别人建立完成的数据报告。并且对报  
告进行投资从而获得更多的回报。

OBSERVER系统建立了一个完全自给自足的  
大数据生态系统，并且能够不断自我成长，  
最终将为全球大数据行业而提供全新的方向。

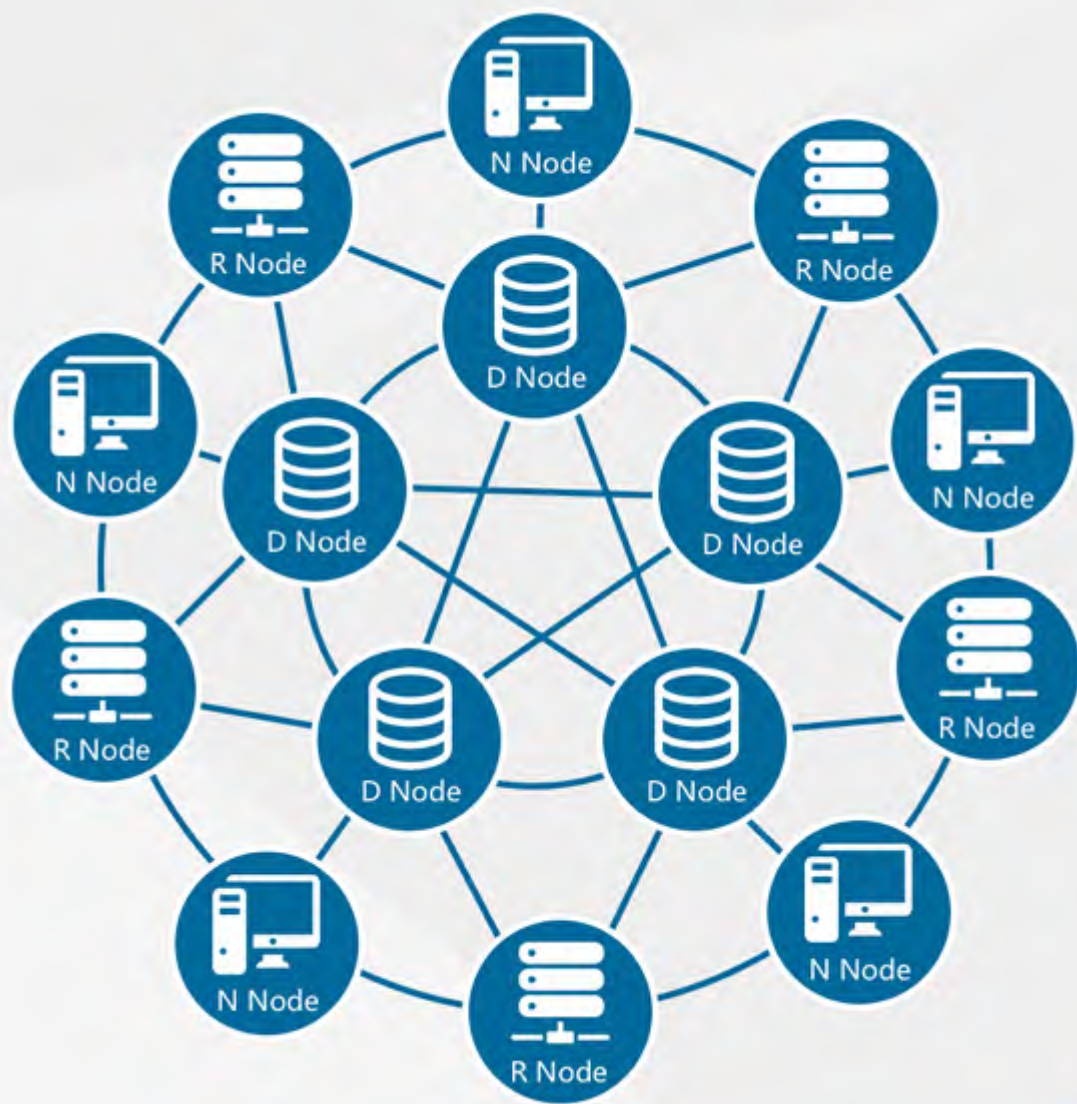


我们是大数据的观察者



## 系统架构

- 在整个系统中一共有三类数据节点，D类节点用来保证完整数据节点。R类节点保存查询数据，N类节点为普通节点。
- D类数据节点将会把类似于比特币的区块链数据提供存储和查询服务。
- R类数据节点负责提供存储查询结果服务。
- N类节点将能够对保存在D类节点的区块链数据进行查询，查询的结果能够保留在R类数据节点。



## 系统运作机制

- 用户支付代币OBS在D保存类节点进行查询，获得的查询结果可以保存在R类节点。
- 每份报告的存续需要消耗代币OBS，当本报告的代币被消耗完就自动被删除。
- 用户可以用代币OBS投资已经完成的查询报告，并且获取相应的代币收益。
- 用户可以直接用代币OBS购买在R类节点保存的各类报告，该报告的收益将会按比例分配给投资者。
- 用户可以在系统用代币OBS发布悬赏，其它用户可以生成报告并获得代币奖励。



# 全新的数据库形式

1



系统将会发放1亿OBScoin代币，这些代币将会在ICO过程中释放给投资者，投资者可以使用这些代币进行查询或者参与报告投资。由于系统整体代币有限，当需求增加，该代币将会获得升值。

2



D类节点将参与提供区块链数据的存储和查询服务。D类节点对硬件和网络有较高要求。运营团队将会维护5个D类节点，其他人也可以提供同样规格D类节点，确保数据的安全和公正。使用PoS机制，所有D类节点每60分钟将会获得系统万分之一的代币奖励

3



R类节点将参与提供对于查询报告的加密存储和查询服务。由于查询报告的存续是需要消耗代币，而R类节点将会通过提供服务而获得代币奖励。

通过OBScoin的代币机制，将让系统中的每个角色都能够获得激励，从而能让系统完全自主运行。



# Kylin引擎

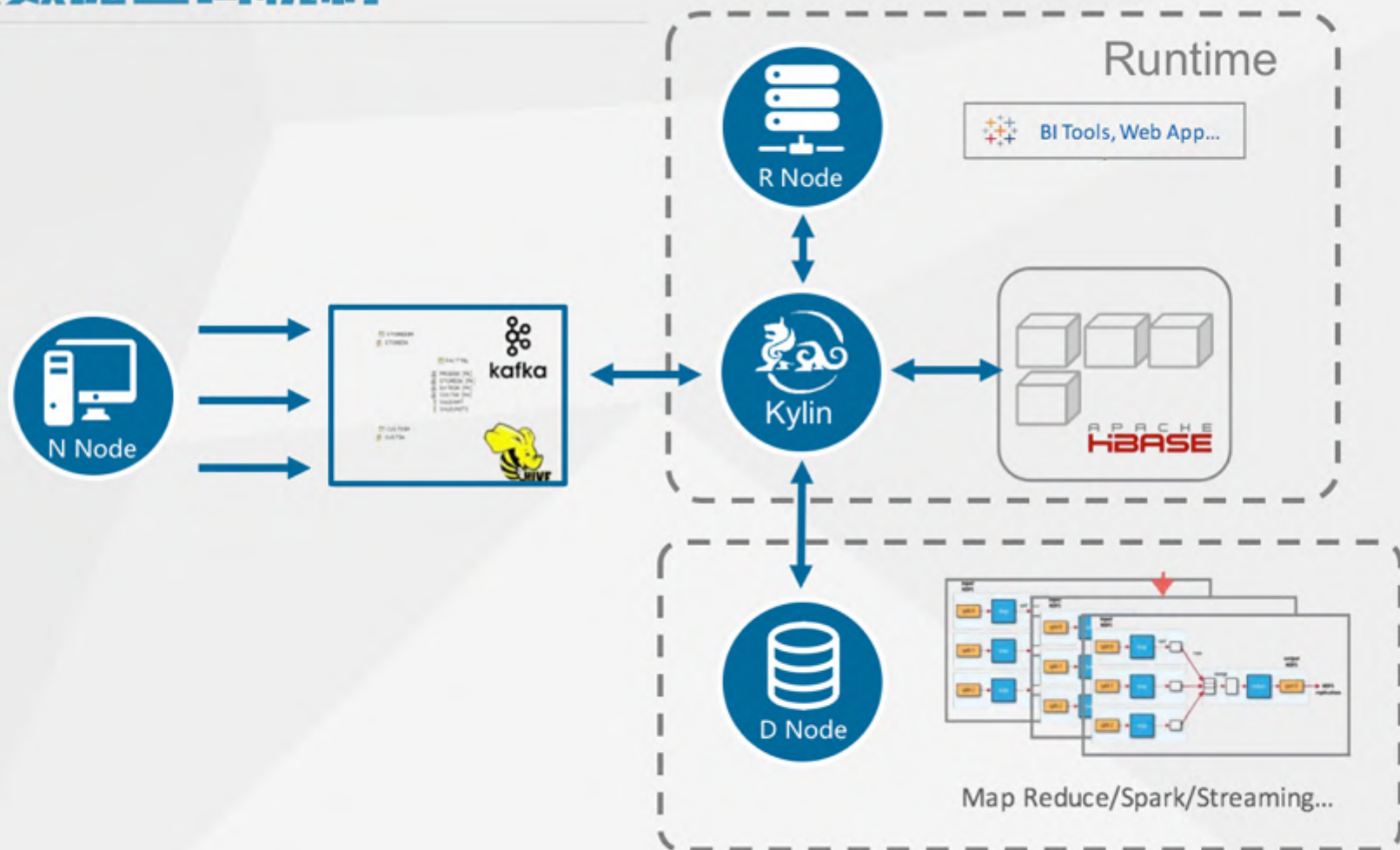


随着Hadoop为代表的大数据技术的普及，越来越多的数据被收集、存储起来，并进一步进行各种处理以满足不同的业务分析需求。业界的实践证明，分布式大数据平台可以有效的进行各种批量处理、数据加工甚至挖掘等，为“机器”使用和处理大量数据带来了前所未有的便利和能力。

但大数据的飞速发展并没有为传统的数据分析师带来更多的好处。其原因在于Hadoop等大数据平台能很好的满足批量数据处理需求但很难让“人”以“交互式”的方式在超大规模数据集上进行各种不同维度的快速分析，特别是Hive等最终将SQL查询翻译成MapReduce的方式无法让用户在秒级时间内获得他们所需要的结果，而且很多分析师不得不使用Shell终端等方式访问和运行相关脚本，远远超越了对一个分析师、BI人员的要求。同时，由于处理时间和方式的差异，很难在短时间内获得分析结果，从而加大了企业内大数据分析平台及应用在推行上的难度。

Kylin是一个开源的分布式分析引擎，提供Hadoop之上的SQL查询接口及多维分析（OLAP）能力以支持超大规模数据。Apache Kylin是首个完全由中国团队设计开发，并贡献到Apache软件基金会(ASF)的顶级项目，开源一年左右的时间，已经在国内国际多个公司被采用作为大数据分析平台的关键组成部分，包括eBay、Expedia、Exponential、百度、京东、美团、明略数据、网易、中国移动、唯品会等。

# 大数据查询机制



## 技术特点



### 亚秒级查询

在百亿及以上规模数据集上为业务用户及分析师提供亚秒级的查询速度，并同时支持高并发，使得在大数据平台上对超大规模数据进行交互式分析成为可能；支持ANSI SQL查询标准，使得业务用户及分析师无需重新学习新的技术即可掌握在海量超大规模数据集上快速分析的能力。



### 无缝集成

支持与企业级商业智能(BI)及可视化工具无缝集成，提供标准的ODBC、JDBC驱动及REST API接口等以连接流行的数据分析、展示工具，如Tableau、Microsoft PowerBI、Microsoft Excel、Apache Zeppelin、Saiku等



### 自助服务

通过电脑或者移动设备，使用简单易用的界面，使得分析师及用户能以简洁而快速的方式分析海量数据。易于使用的Web界面允许用户自己构建数据集而无需知晓底层技术。



### 可扩展架构

全新设计的可扩展架构从根本上解耦了对特定技术的依赖，将计算框架，数据源以及底层存储等扩展到更多的技术领域，为不同的技术栈提供可配置的优化解决方案。



# 核心优势

## 数据公开透明

完全基于区块链技术，确保数据的真实性，没人可以操纵、伪造或者篡改数据或者查询结果。

用户仅需浏览器，以熟悉的可拖拽方式交互地探复杂数据源，支持数十种可视化报表技术，简易地多格式数据分享。

## 海量数据简易查询

### 核心优势

## 代币驱动

通过代币进行激励，在无需中介介入的情况下能够确保其中的所有人都可以从中获得激励，来确保系统稳定运行。

该模式在未来兼容更多的大数据形态，来重构分布式大数据的商业模式。

## 重构商业模式

# 报告的生命周期

## 定价

报告的定价权属于报告的生成者和报告的投资者共同拥有。生成者也可以将报告的以价格设置为零，那么报告生成者将无法从该报告获得任何收益。

## 生成

用户可以选择根据悬赏或者自己通过支付代币获得查询结果，并生成大数据查询报告。该数据报告类型如果为可分享，则可以设定报告价格。

## 悬赏

用户能够发布悬赏 (Reward) 来征求数据结果。

03

02

01



04

## 保存

报告将会加密保存在多个R类节点，并且每天消耗一定数量代币。由于是多节点保存，可以确保所有信息不可篡改，而节点本身无法读取保存的数据报告。

05

## 投资

用户能够选择有投资潜力的报告来进行投资，一部分投资代币将用于维持报告的存续，投资金额越高则获得回报的比例也越高。

06

## 购买

可以选择购买任意在分享中的报告，根据每个报告的价格支付系统代币，并且可以对报告进行评价和评级。该评价和评级可以在今后报告生成者的记录中可以查询

07

## 评价

每个购买者将可以对报告进行全方位的评价，对数据分析的专业性、深度、广度、准确度和客观性进行评价，并且该评价将可以被所有人可见，而且成该报告生成者的历史记录。

08

## 删除

当报告最终维持代币被彻底消耗完，则表明其生命期自动终结，该报告将会被系统自动删除。

# 目标用户



01

## 数字货币投资者

查看各类数字货币的真实交易情况，并且及时追踪数字货币在交易所的进出情况

02

## 相关商户

能够确认进出的数字货币是否来自于非法活动，并且了解数字货币整体使用清苦康

03

## 监管部门

能够通过大数据追踪非法的数字货币交易以及洗钱行为，对交易进行画像

04

## 开发者

了解数字货币的使用情况以及存在问题，并能够进行针对性开发

05

## 相关企业和机构

能够通过系统开发接口进行区块链相关开发并建立相关应用。



# OBSERVER的价值



# 区块链铅笔

希望能够和您共同书写区块链的历史

