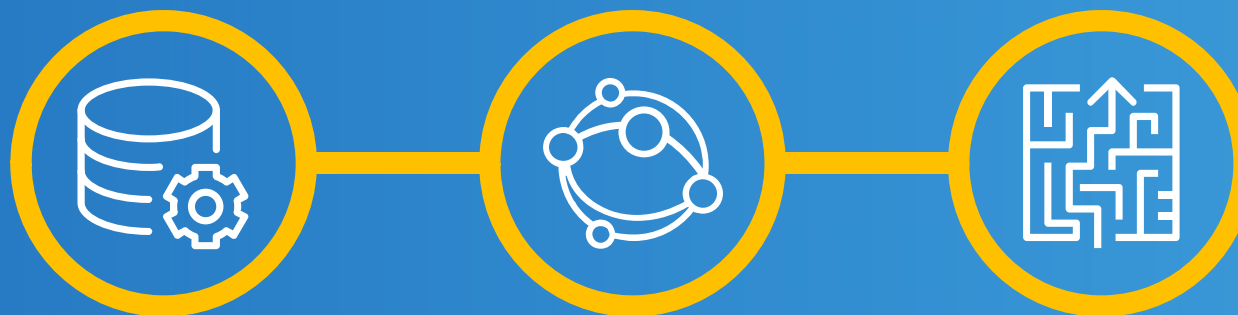


存储技术演进及融合人工智能的探索

先智数据公司

董唯元

2017年9月

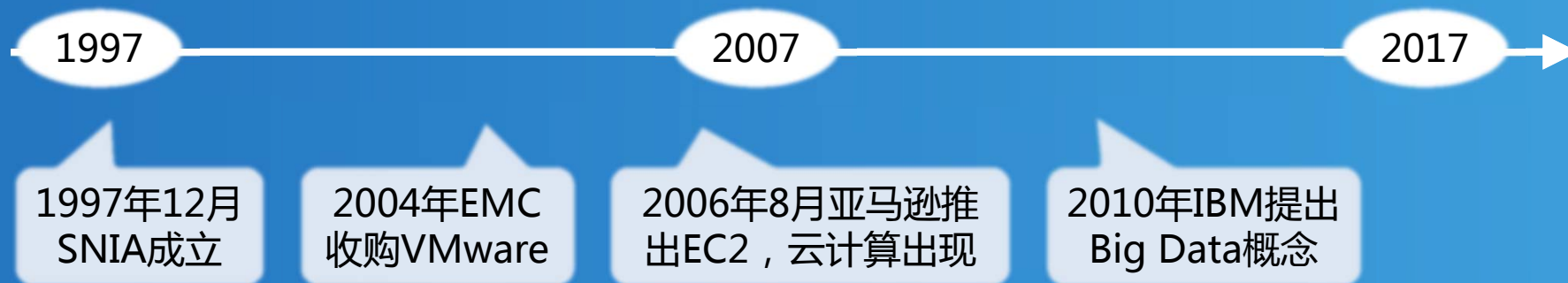
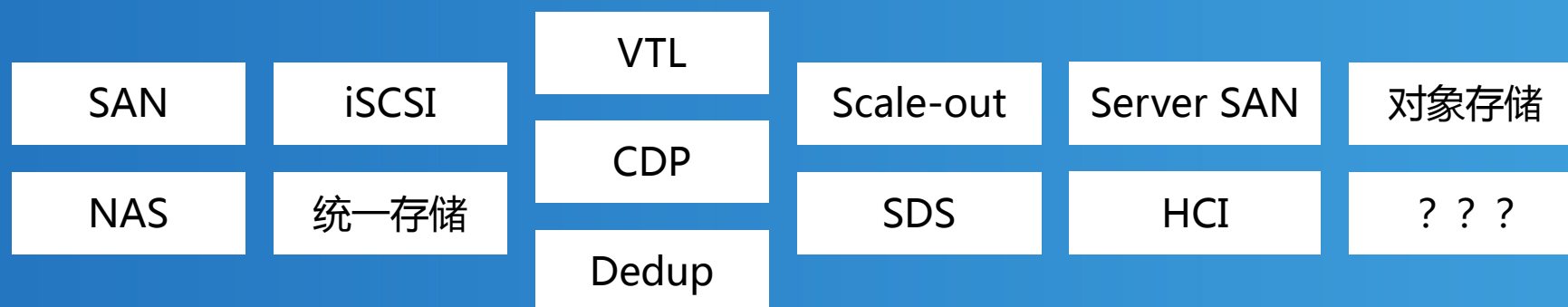


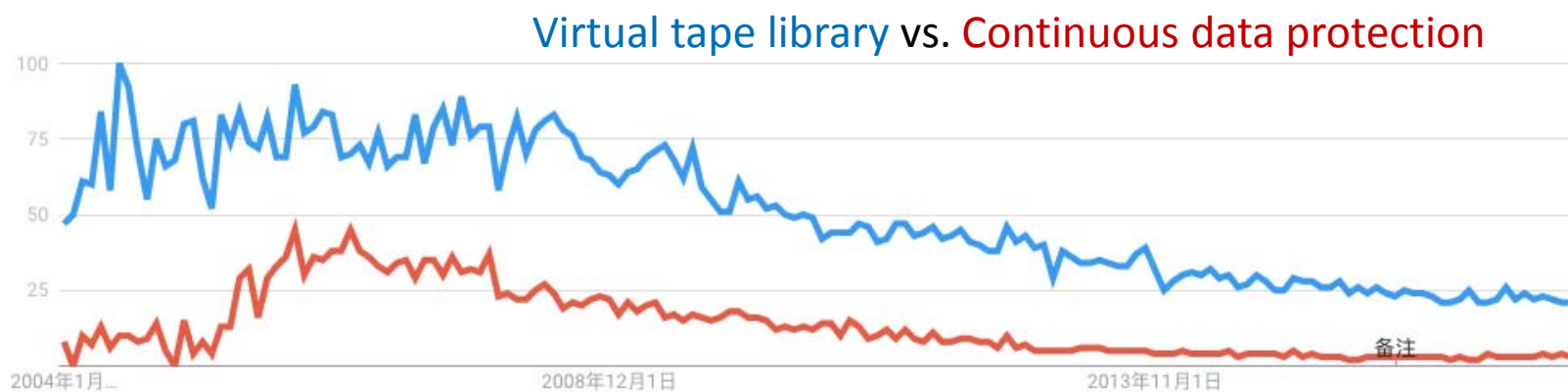
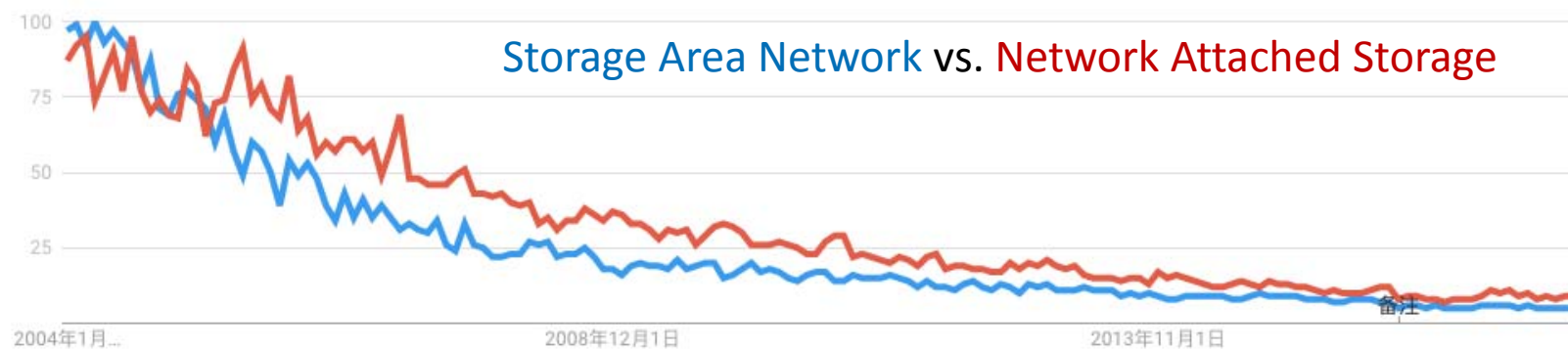
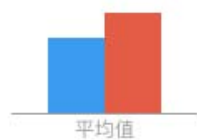
行业历史

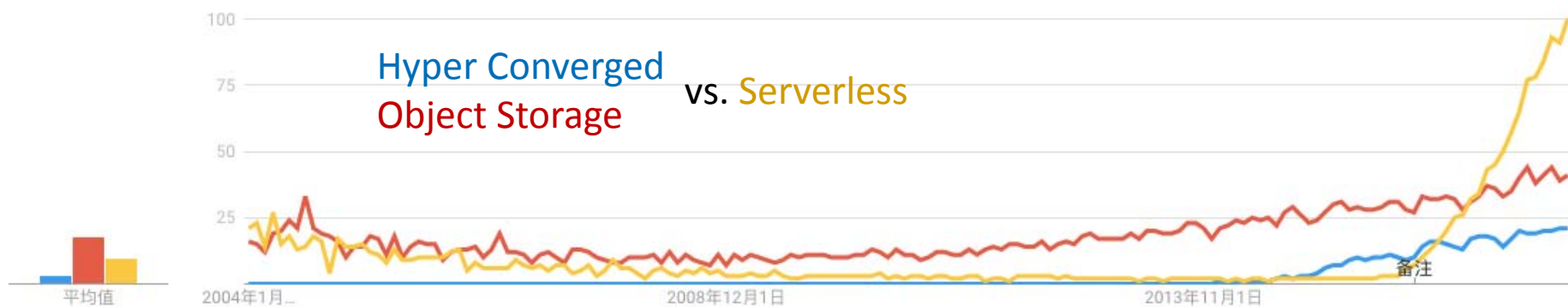
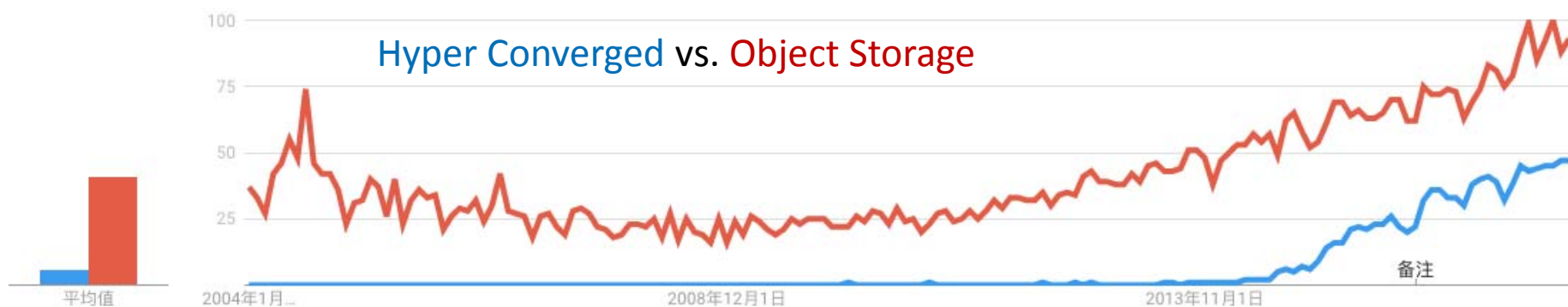
时代冲击

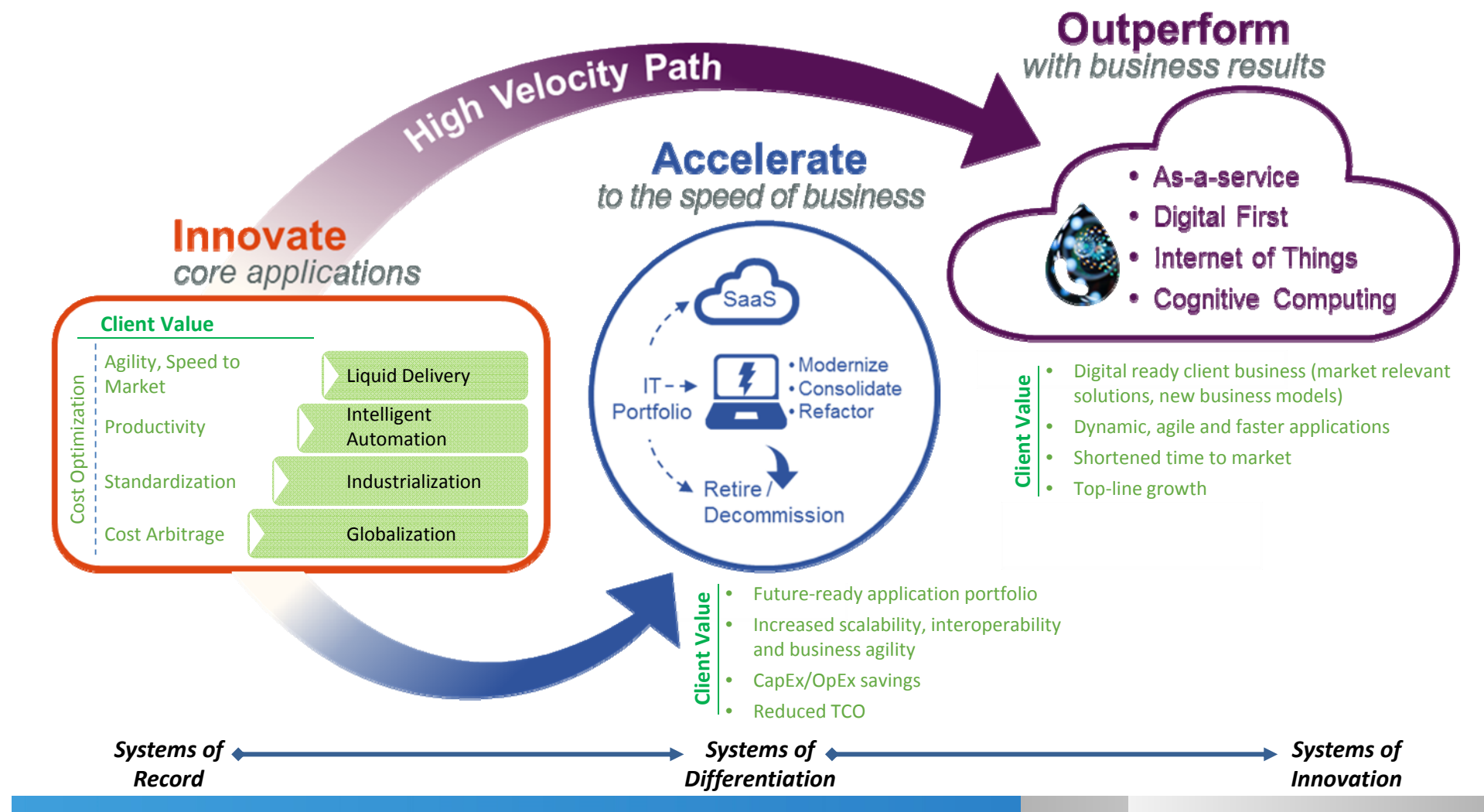
未来展望

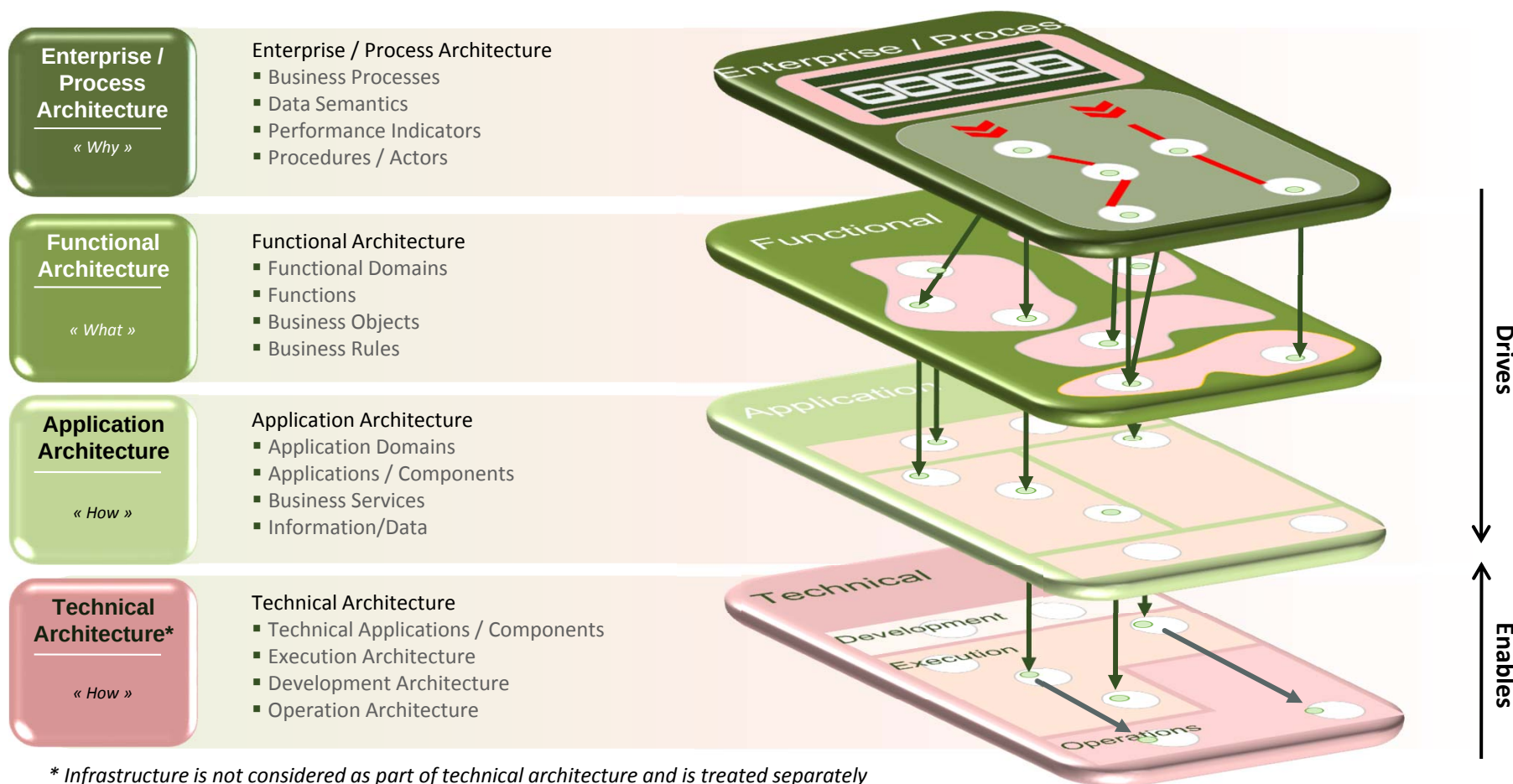
存储行业历史



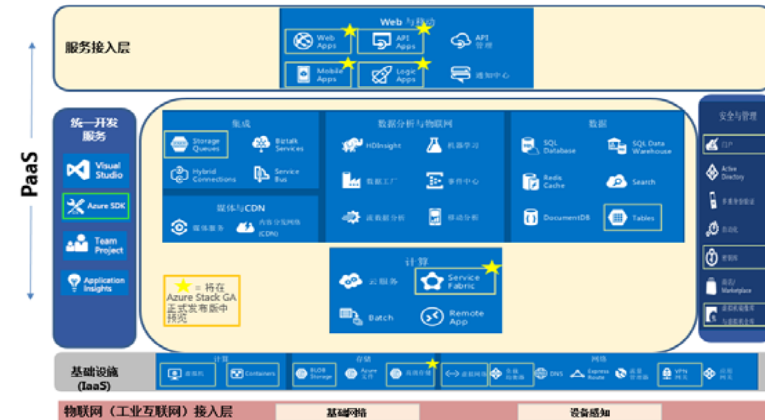
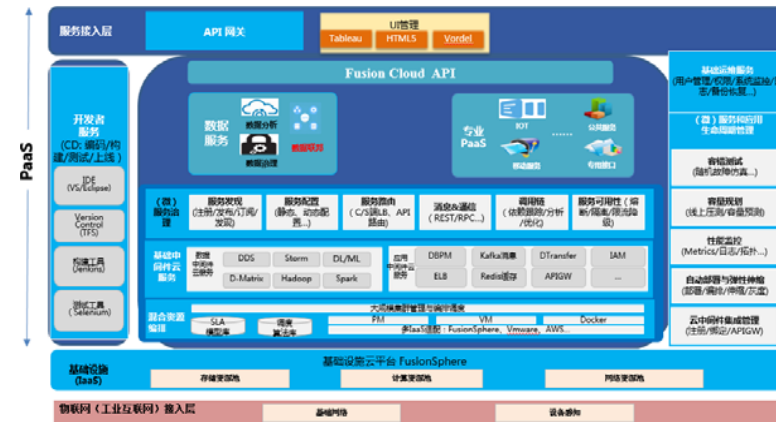
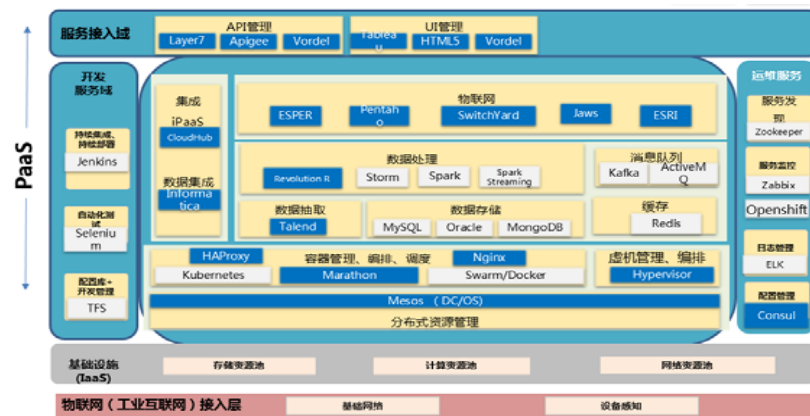


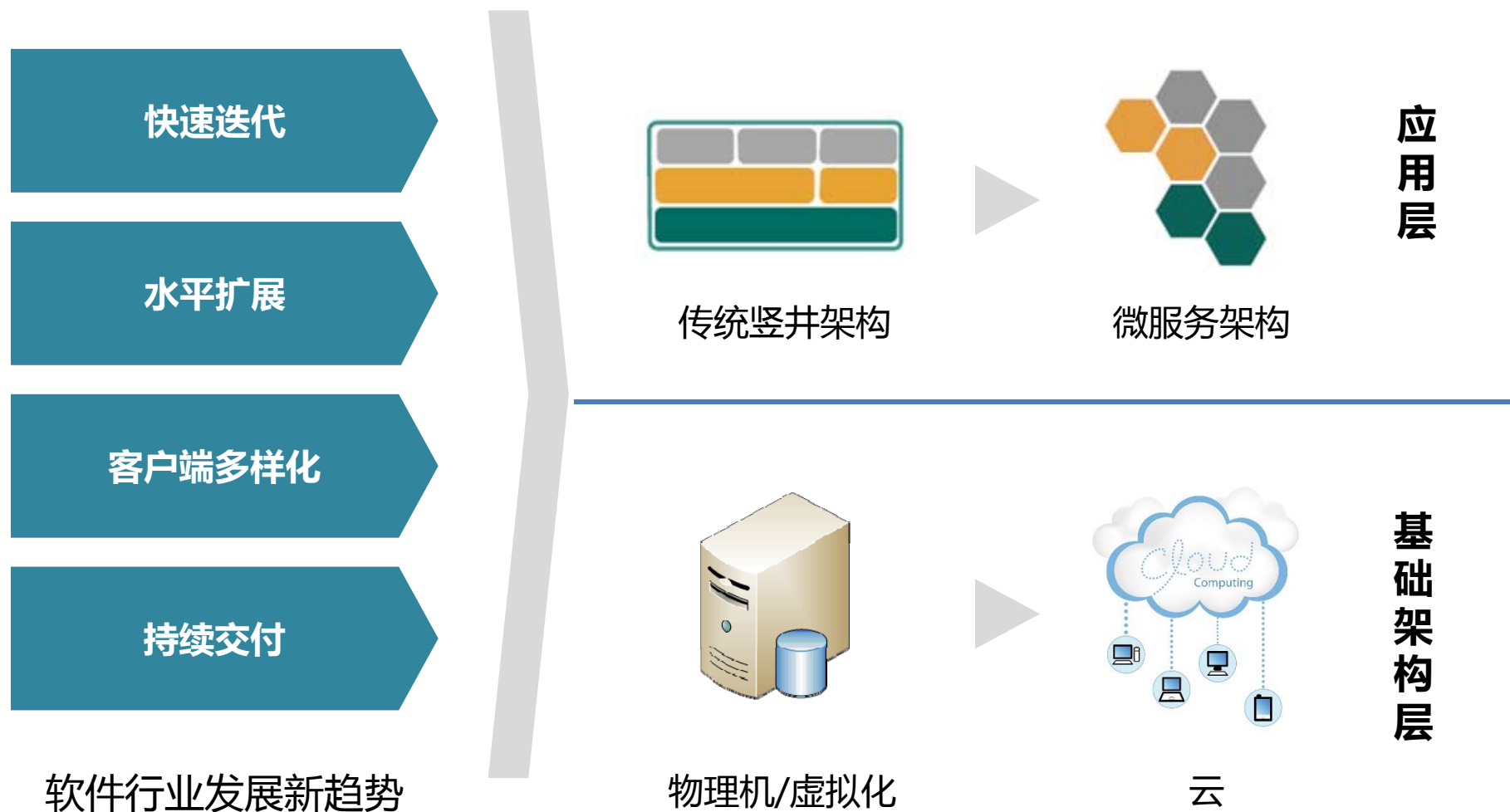






* Infrastructure is not considered as part of technical architecture and is treated separately





Server SAN、HCI、分布式块存储——过渡期的暂时性需求



- 传统应用互联网化的起步引擎
 - 兼容传统应用
 - 实现分布式资源池化
- 互联网原生应用并不存在此类需求

业务无关性的思考

- 通用性是优势吗？
- 存储技术的发展，是否距离业务诉求越来越远？
- Alluxio (Tachyon) 的例子

人工智能可以帮助存储系统做什么？

 PROPHETSTOR

传统存储设备
强专业性运维



智能化调度
简化运维流程
弱专业性依赖



智能化存储
突破机制限制
消除互斥问题

问题示例——普通软件定义存储的尴尬

PROPHETSTOR



软件定义存储

多副本

纠删码

快速修复

快照

复制

.....

.....以稳定性和可靠性考量，软件定义存储只适合替代中低端磁盘阵列.....

——某最终用户

开放性与可靠性互斥难题！！！！

空间利用率与可靠性互斥难题！！！！

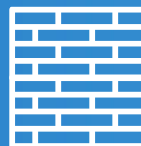
可管理性与性能互斥难题！！！！

可靠性与性能互斥难题！！！！

问题示例——磁盘阵列维系可靠性的方式不再适用

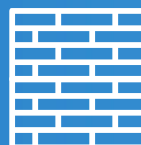
PROPHETSTOR

限制、限制、再限制



牺牲开放性换取可靠性
已经有悖技术趋势

冗余、冗余、再冗余



多盘故障概率随系统
规模增加指数级上升

每当磁盘故障的时候.....

- 降级状态将持续多久？
- 降级状态下哪些操作需要注意？
- 性能下降多少？
- 修复动作占用多少资源？

 PROPHETSTOR

糟糕却流行的方式

Rebuilding Priority :

High/Low



副本分布边界问题

PROPHETSTOR

磁盘MTBF
决定单盘健康概率

两副本可保护上限 **96** 颗磁盘！
三副本可保护上限 **512** 颗磁盘！

系统可靠性
要求达到
99.999%

n 磁盘总数

k 副本数

p 单盘健康概率

P 系统健康概率

$$P = \sum_{i=0}^{k-1} \binom{i}{n} p^{n-i} (1-p)^i$$

基于智能预测的主动式故障处理

PROPHETSTOR

主动式
故障处理

预测故障时间点
估算数据修复窗口
根据负载制定修复计划
提前进行修复操作

平稳应对故障事件
应用负载不受影响

被动式
故障处理

机械性故障检测

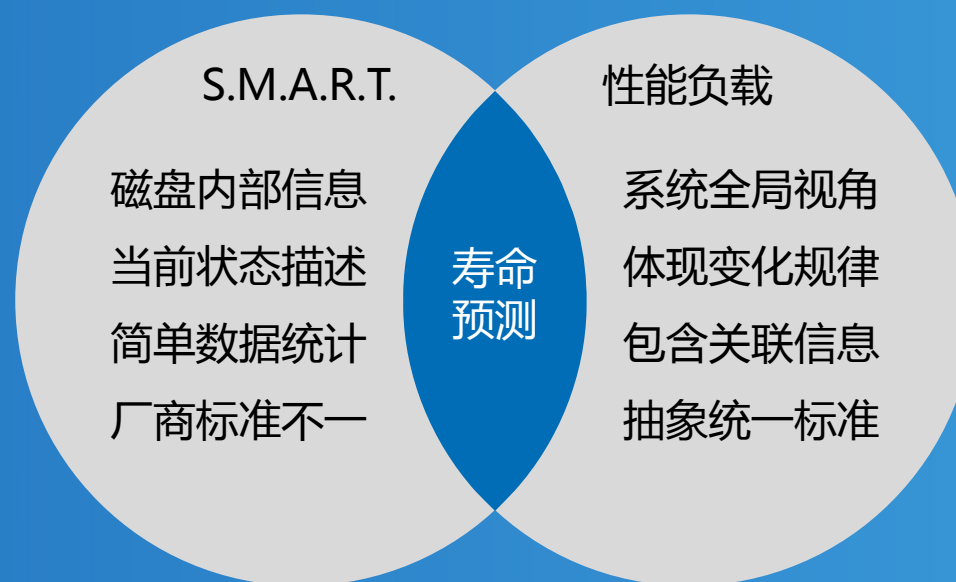
经受考
验时刻

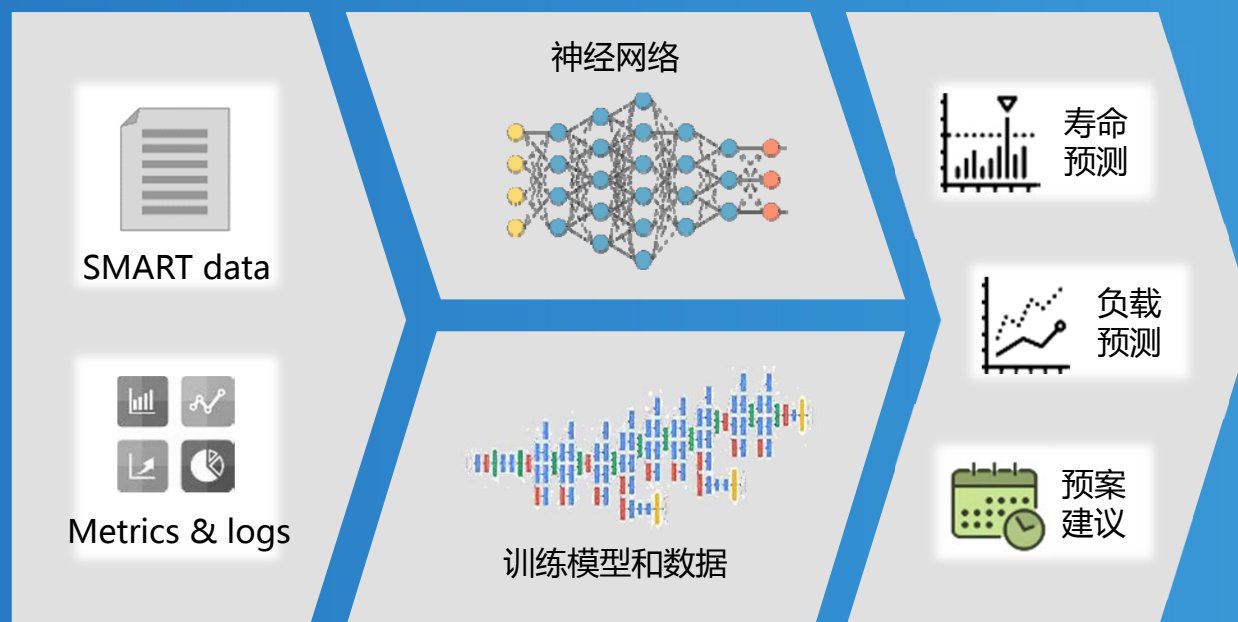
系统进入降级工作状态
被动触发数据修复操作
修复过程抗故障能力下降
按固定设置占用计算资源

准故障
出现

故障
发生

时间



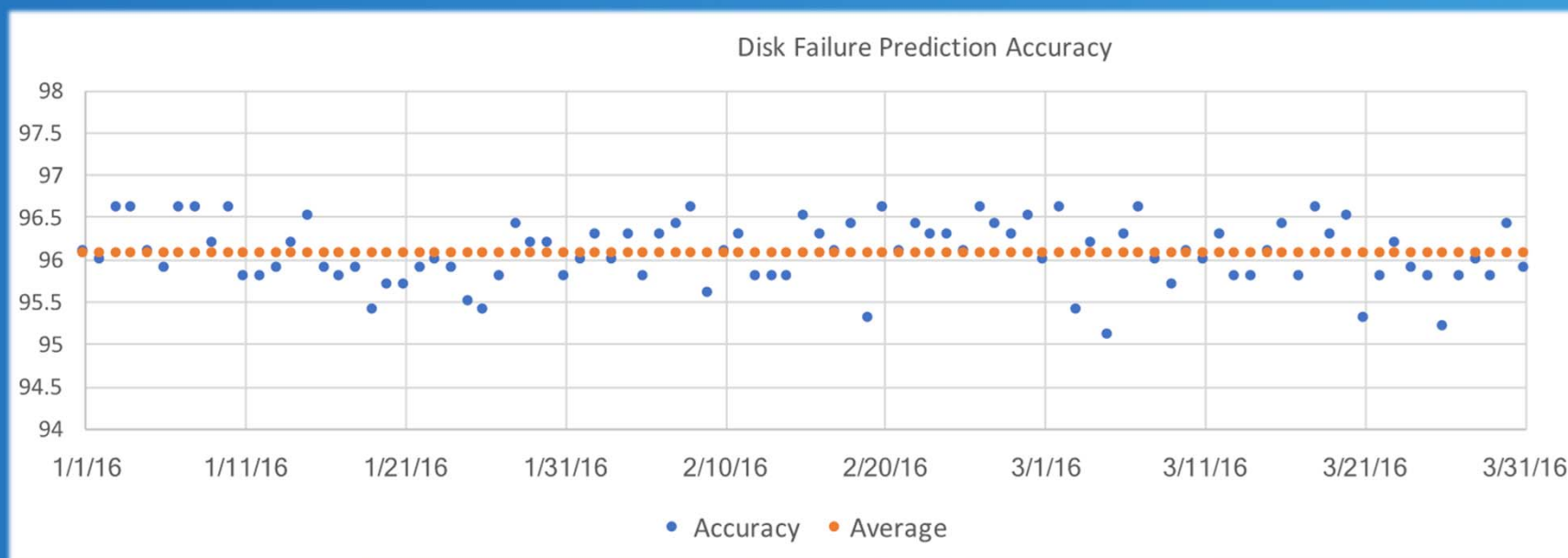


DiskProphet 预测准确率 (Cisco测试报告)

PROPHETSTOR

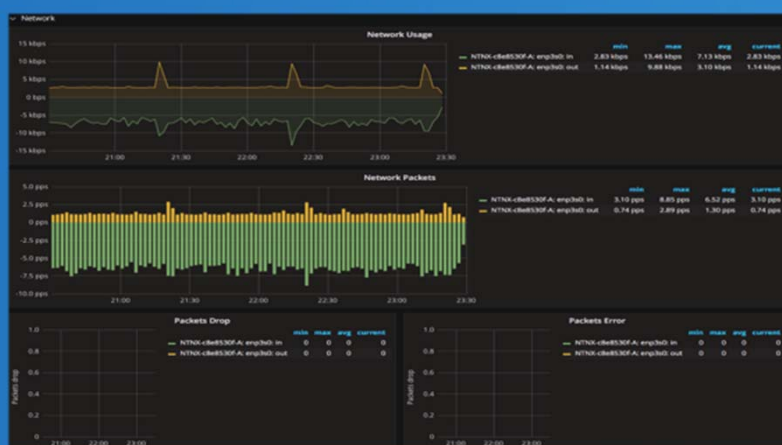
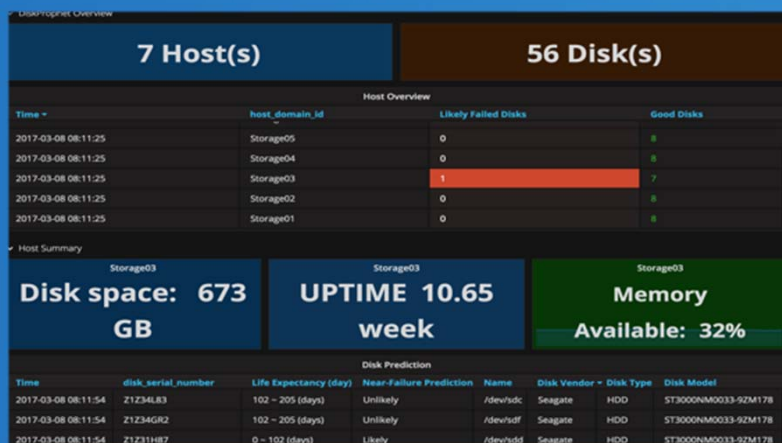
测试周期：2016年1月1日 - 2016年3月31日

平均预测准确率 **96.1%**，最低准确率在95%以上



DiskProphet GUI

PROPHETSTOR



Thank you !

先智数据公司

董唯元