

课程总结

范颖捷 | 2018年7月



目录 >

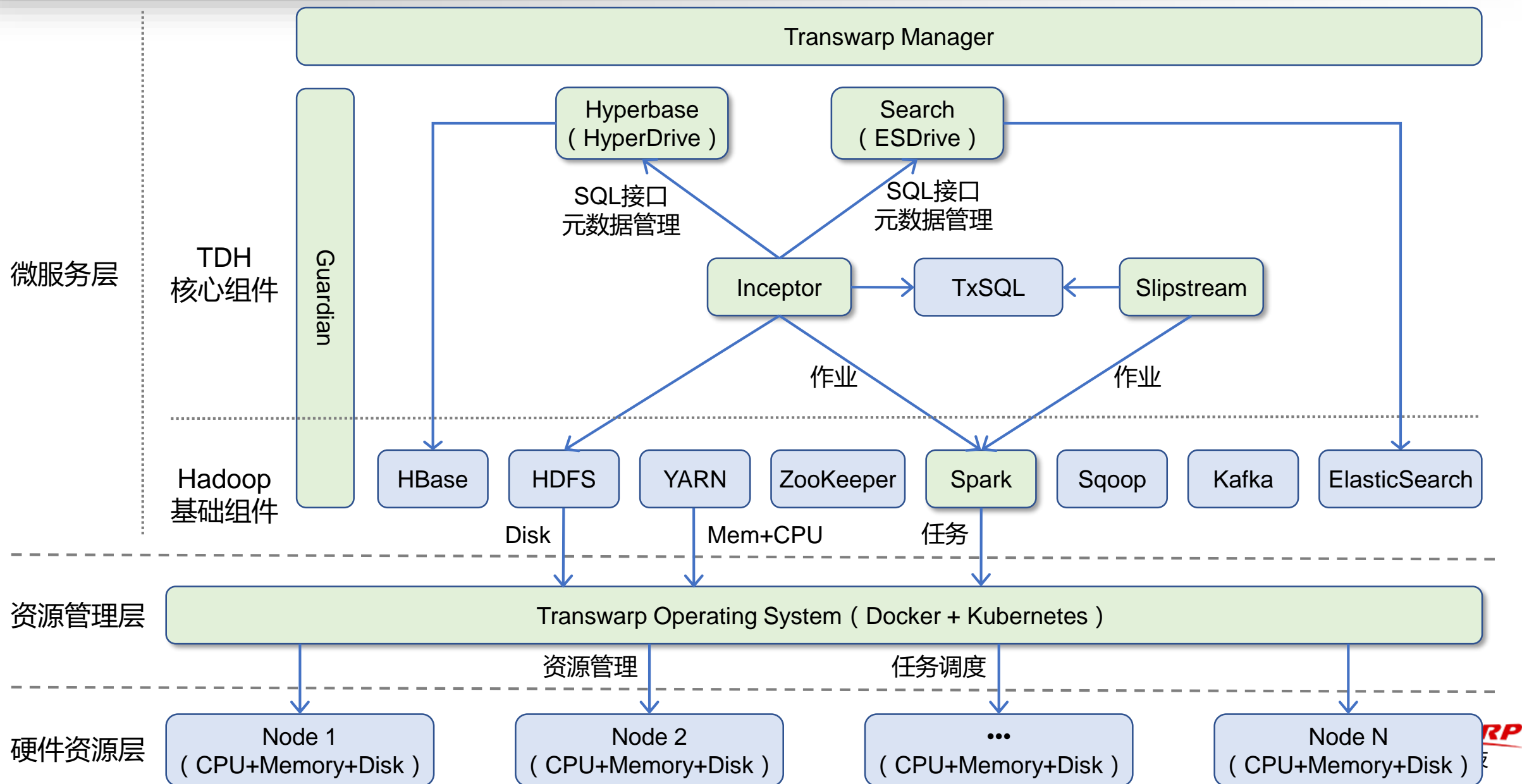
CONTENTS


- 1 TDH平台的技术架构
- 2 大数据技术的典型应用场景
- 3 总结与展望



1 chapter

TDH平台的技术架构

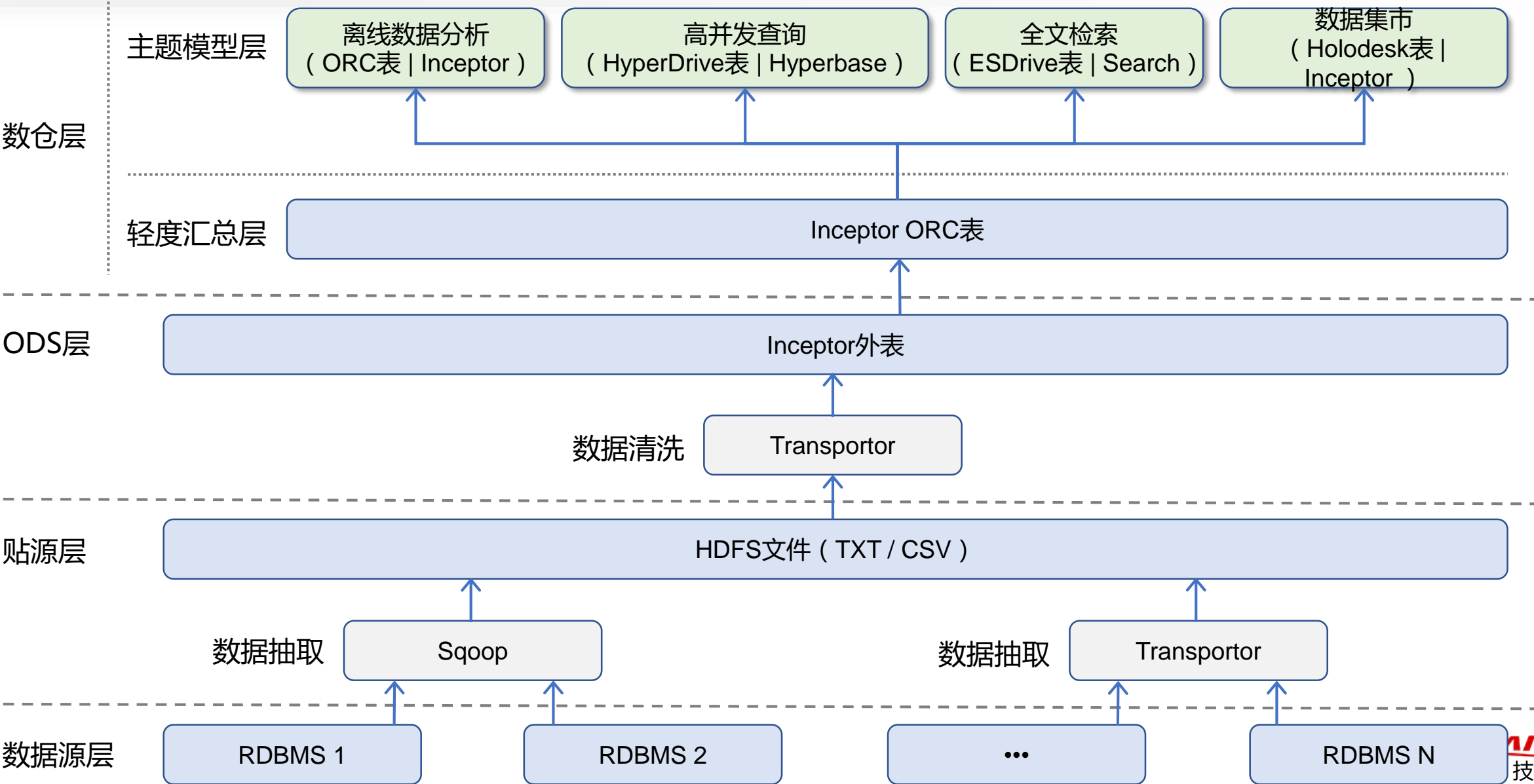




2 chapter

大数据技术的典型应用场景

- ✓ 数据仓库
- ✓ 实时流处理



数据再分析

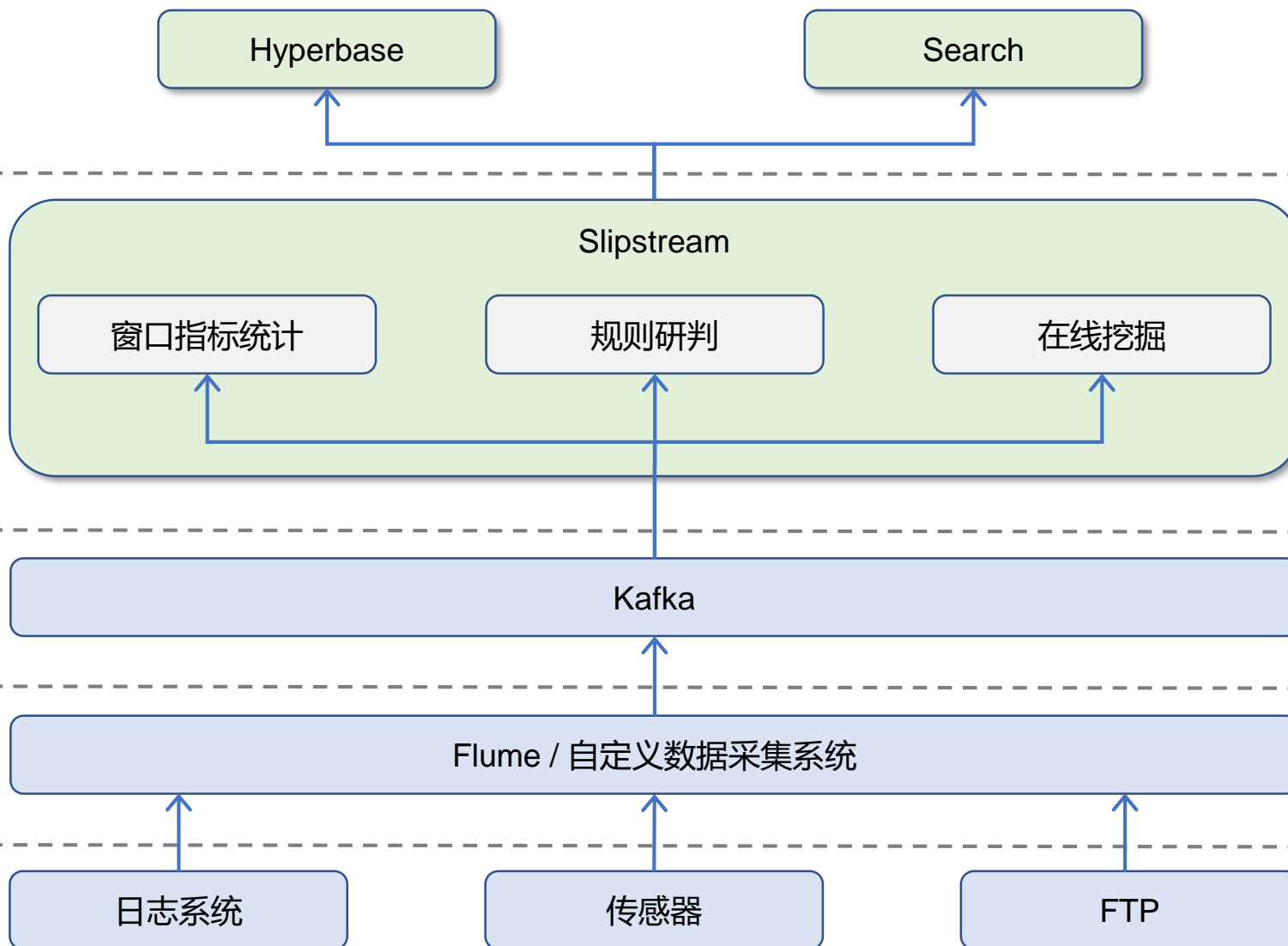
数据存档


流处理引擎

消息中间件

数据采集

数据源





3 chapter

总结与展望

- ✓ 回顾与总结
- ✓ 大数据技术的发展趋势

深度思考

- 哈希取模在哪些技术和产品中使用过，分别发挥什么作用？
- ZooKeeper在哪些技术和产品中使用过？ 分别发挥什么作用？
- 计算框架与资源管理框架是如何协同工作的？
- 分治思想在大数据技术中主要体现在哪些方面，分别有哪些具体应用？

➤ 融合化

- Hadoop

- 产品：Cloudera CDH、Hortonworks HDP、MapR、星环TDH、华为FI
- 适用场景：大集群（1000+）、高并发（HBase）、PB级以上的结构化、半结构化、非结构化数据

- MPP（Massively Parallel Processor，大规模并行计算）

- 产品：EMC GreenPlum、HP Vertica、Teradata Aster Data、Sybase IQ、南大通用GBase
- 适合场景：小集群（100左右）、低并发（50左右）、结构化数据

➤ 微服务化

- 容器技术：Docker、CoreOS

- 数据中心操作系统：Kubernetes、TOS

➤ 云服务化

- PaaS：为用户提供面向主题的大数据平台组件包，如：数仓组件包
- SaaS：为用户直接提供大数据应用产品，如：报表

➤ 智能化

- 大数据处理技术与人工智能算法的深度整合
- 大数据基础平台与人工智能平台的深度整合

➤ 大数据一体机

- 一种专为大数据分析处理而设计的软、硬件结合的产品
- 由一组集成的服务器、存储设备、操作系统、数据库管理系统，以及为数据查询、处理、分析而特别预装和优化的软件组成
- 高性能大数据一体机具有良好的稳定性和纵向扩展性



Q&A

TRANSWARP
星环科技