

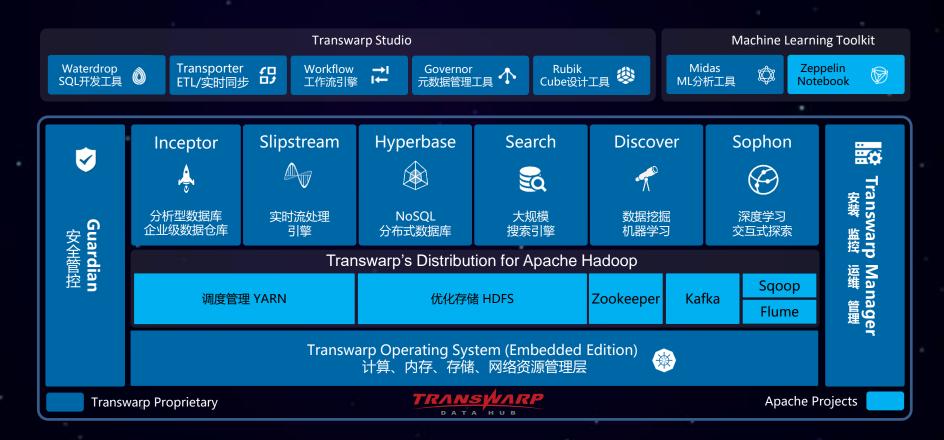
星环大数据产品介绍

111101011000000100g



星环科技 www.transwarp.io 2018年6月

Transwarp Data Hub 5.X Product Family



组件介绍

•基于Hadoop的数据仓库产品 •支持SQL 2003, 存储过程, Oracle/DB2/Teradata方言, 分布式事务 Inceptor •通过标准测试TPC-H/TPC-DS证明产品对性能的极限优化与提升 •混合负载管理与SLA管控 •融合事件驱动与微批处理的流式计算引擎 •完整支持SQL与存储过程 Slipstream •高可用支持与Exactly-Once语义支持 •针对物联网与低延时计算场景优化,如实时反欺诈 •完整支持使用SQL进行高并发业务 Hyperbase •支持建立全局/二级索引 •半/非结构化数据处理平台(JSON, Images etc.), 支持对象存储 •大规模数据全文检索引擎 Search •完整支持使用标准SQL扩展支持全文检索 • 支持混合存储模型,可以利用SSD存储加速 •机器学习与数据挖掘平台 •支持SQL/R/Python等开发接口 Discover •提供分布式机器学习算法 •提供丰富的行业模板 •集成TensorFlow与MxNet等深度学习框架 Sophon •多数据源支持,大数据与深度学习的完美结合 •帮助数据科学家方便地建立神经网络模型

分布式SQL引擎Inceptor





Apache Spark

基于内存的Map/Reduce计算引擎,即将成为新一代主流计算框架。 处理大数据像"光速"一样快,比Hadoop Map/Reduce快10x倍。

Holodesk

跨内存/闪存/磁盘等介质的分布式混合列式存储,常用于缓存数据供Spark高速访问。Holodesk内建内存索引,可提供比开源Spark更高的交互式统计性能;结合使用低成本的内存/SSD混合存储方案,可接近全内存存储的分析性能。

Most complete SQL support

兼容>95% ANSI SQL 2003, HiveQL和>90% PL/SQL语法,支持数据仓库、数据集市等分析系统中常用的复杂分析型语法,方便应用迁移。

SQL引擎

高度优化的高速SQL引擎,可运行在Spark或Map/Reduce上,可高速处理缓存在Holodesk上的列式数据。

丰富的工具支持

支持主流可视化和BI/挖掘工具,包括Tableau,IBM Cognos,SAP BO,Oracle BI,SAS等。支持Informatica,Pentaho/Kettle等ETL工具。

Data Federation

具备对多种关系数据库和Hadoop数据源进行交叉查询,聚合,以及关联操作等能力

2018/6/3 www.transwarp.io

实时流处理引擎Slipstream





Slipstream

- ➤ 99%的ANSI SQL 2003的支持率
- 强大的优化器提升性能
- ▶ 支持按时间切分滑动窗口和滑动步长

多数据源支持

- Socket
- > 文件
- Kafka
- ➣

多种输出方式支持

- Hyperbase
- ➤ Holodesk
- ➤ HDFS
- ×

企业数据总线支持

▶ 从一个流读入数据,再将其输入另一个 流,如常见的企业数据总线以Kafka为存 储中间件

在线数据挖掘支持

- ▶ 离线模型, 在线预测
- ▶ 时间窗口数据进行实时挖掘分析

NewSQL数据库Hyperbase



Real-time OLTP + OLAP + BATCH + Search + Graph Traversal Application
Mixed Workload 混合负载业务

Transwarp Inceptor

Transwarp Hyperdrive

Transwarp ESDrive

Transaction	Document	Object	Index	Search
SQL & API	SQL & API	SQL & API	SQL & API	SQL & API
Transaction Execution Engine 分布式事务 处理引擎	Document Store json/bson 文档存储	Object Store image/files, etc 对象存储	Global/Local /Full-text Index 全局/局部/全文 索引	Distributed Full-text Search 全文搜索

Transwarp HBase – Hadoop Database 分布式 Big Table

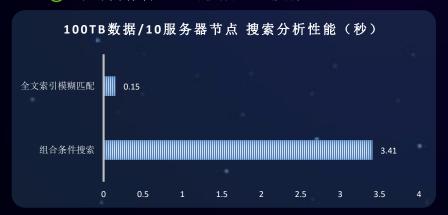
Elastic Search 分布式全文索引

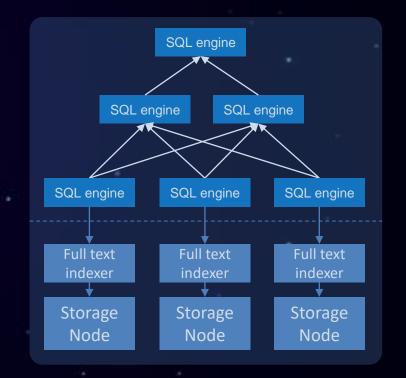
Hyperbase

- ➤ 完整支持使用SQL进行高并发业 务
- ▶ 支持建立全局/二级索引
- ➤ 半/非结构化数据处理平台 (JSON, Images etc.), 支持对象存储

大规模统计和搜索引擎Search

- ✓ SQL引擎和搜索引擎融合
- ◇ 计算紧贴数据和索引,亚秒级在线分析
- ◇ 冷热数据分级存储,毫秒级实时关键字检索
- ✓ 企业级数据访问控制,稳定可靠
- ✓ 堆外内存管理,可扩展至PB级别



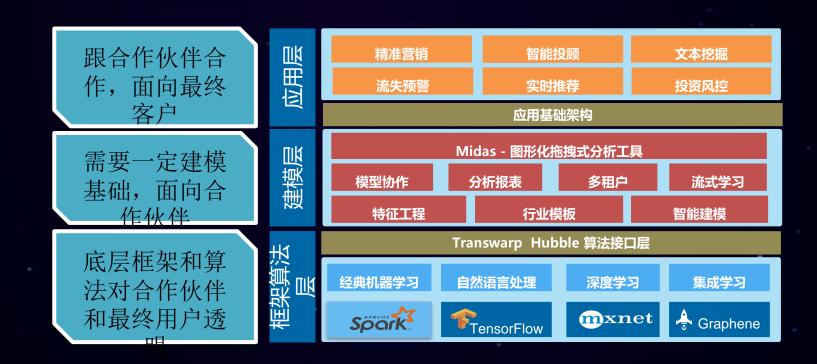


数据挖掘平台Discover



Transwarp Discover Platform		担保链分析	推	荐系统	X	.险分析	Б	豆欺诈	Ż	文本分类
	使用层	web JS Interface web展示接口		Language Interfaces 多语言接口			Midas <mark>b</mark> midas 图形界面		R Studio R Studio IDE开发环境	
		Transwarp Hubble – Programming Interfaces to connect algorithms								
	算法模型层	Feature Engineering 特征工程	Statistics 统计算法	Classification 分类算法	Clustering 聚类算法	Regression 回归算法	Genetic 遗传算法	Graph 图计 算	Sampling 采样算法	Q-Learning 增强学习
		Industry Templates 行业模板	Sequential Analysis 时序分析	Dimension Reduction 主成分分析	Association Mining 关联/推荐	Deep Learning 深度机器学习	Discriminate Analysis 判别分析	Factor Analysis 因子分析	Belief Network 信念网络	Decision Methods 决策方法
		Customized Plugins 自定义插件	Distributed Execution Engine 分布式执行引擎		Deep Learning Frameworks 深度学习框架		R Runtime Library R语言动态运行库			
	数据层	Transwarp Connector – SQL Interfaces to connect data sources File Interface								
		Inceptor			yperbase Stream			HDFS		

数据分析平台Sophon



星环科技实用链接



- 星环科技技术论坛:
 http://support.transwarp.cn/
- 星环科技产品手册:
 https://docs.transwarp.cn/

大数据技术公众号







