



# 华南某市供电局全景可视化大数据平台

## 项目背景：

随着电力企业信息化建设和实用化的逐步推进，信息化数据已经日趋完善，先后完成了生产调度、电力营销和客户服务、经营管理、资产管理等各种应用系统，积累了大量数据。随着数据量持续增加、数据类型日益丰富、以及部分业务产生了对实时数据的应用需求，传统的关系型数据平台已难以继续有效支撑数据管理和应用的需求，主要体现在：

1) 缺乏对非结构化数据的支持,大量数据资产未被合理有效利用。当前数据存储主要针对结构化数据的采集、存储、计算和应用提供支持，而无法对图档、文档、日志等非结构化数据提供高效支持。华南某市供电局的输变电、配电、营销系统具备对接线图、设备图、用户档案等非结构化数据进行存储和检索的需求，而这类数据具有类型多、体量大、价值密度低的特点。使用传统集中式的方式即便实现对此类数据的存储，也存在检索效率低下的问题，同时，由于集中式存储的成本高昂、扩展空间有限，也必须按照一定周期用磁带进行归档处理，对于归档后的数据一般难以再次有效应用，使大量非结构化数据资源被闲置。

2) 缺少流式处理的手段，无法支撑各专业对实时数据的应用需求。按输变电、营配信息集成的管理及应用要求，包括SCADA在内的多种实时数据对于华南某市供电局相关业务的实时监控、评估与分析具有重要意义。而目前现有数据处理架构无法支持对时序数据的实时处理。

通过引入大数据相关技术组件，开展输变电设备展示的大数据平台搭建部署，在现有数据存储基础之上，对数据采集、存



储、计算和应用层，分别构建对结构化、非结构化以及流式和批量数据的处理能力；基于X86 服务器，搭建功能全面、性能高、扩展性好、安全稳定、成本低廉的分布式处理集群；建成分布式框架的大数据管理及服务平台。为输变电设备信息全景展现及高端应用提供一个支撑平台。通过平台建设，可以快速实现各类信息的自助查询、高端分析，为公司的分析和决策提供最方便、直观、可操作的可视化窗口，进而提升华南某市供电局资产绩效管理的优化。

## 项目目标：

本次华南某市供电局全景可视化大数据平台的建设，主要有两个目标。

首先是引入大数据相关技术组件，开展输变电设备展示的大数据平台搭建部署，在现有数据存储基础之上，对数据采集、存储、计算和应用层，分别构建对结构化、非结构化以及流式和批量数据的处理能力。

另一个目标是基于X86 服务器，搭建功能全面、性能高、扩展性好、安全稳定、成本低廉的分布式处理集群；建成分布式框架的大数据管理及服务平台。为输变电设备信息全景展现及高端应用提供一个支撑平台。

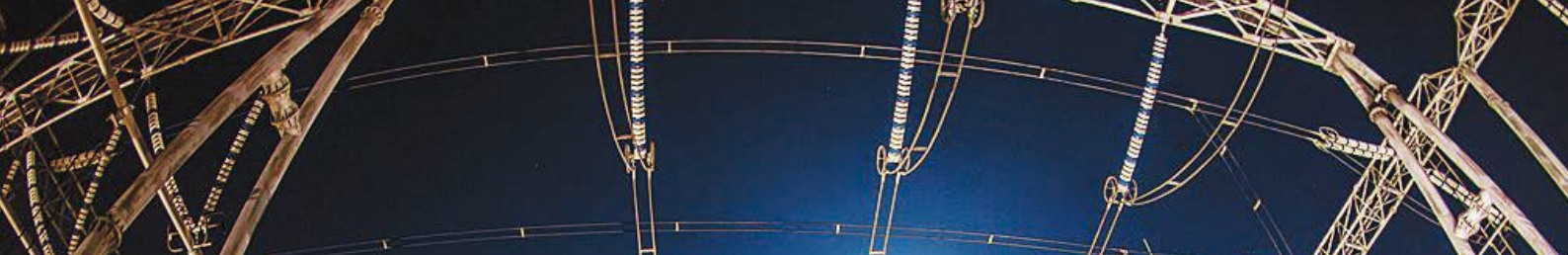


## 技术架构：

输变电设备展示大数据平台的规划满足企业级数据仓库的要求，以Hadoop平台技术为基础，以并行/分布式架构为主，结合Spark计算框架，满足数据存储以及不同计算场景对数据平台服务的要求。数据平台具备线性扩展能力，并提供完善存储策略、加载策略、安全策略、压缩策略、备份策略、灾难应对策略以及高效的运作管理。

该方案实现可管理、易管理的目标。数据平台可以是一个混合数据平台，可选的数据系统类型包括分布式文件系统、并行分析型数据库、关系数据库、内存数据库等，但数据系统类型的选择必须在充分的技术分析前提下进行，符合数据平台的整体规划的要求。





## 华南某市供电局输变电设备全景可视化大数据平台



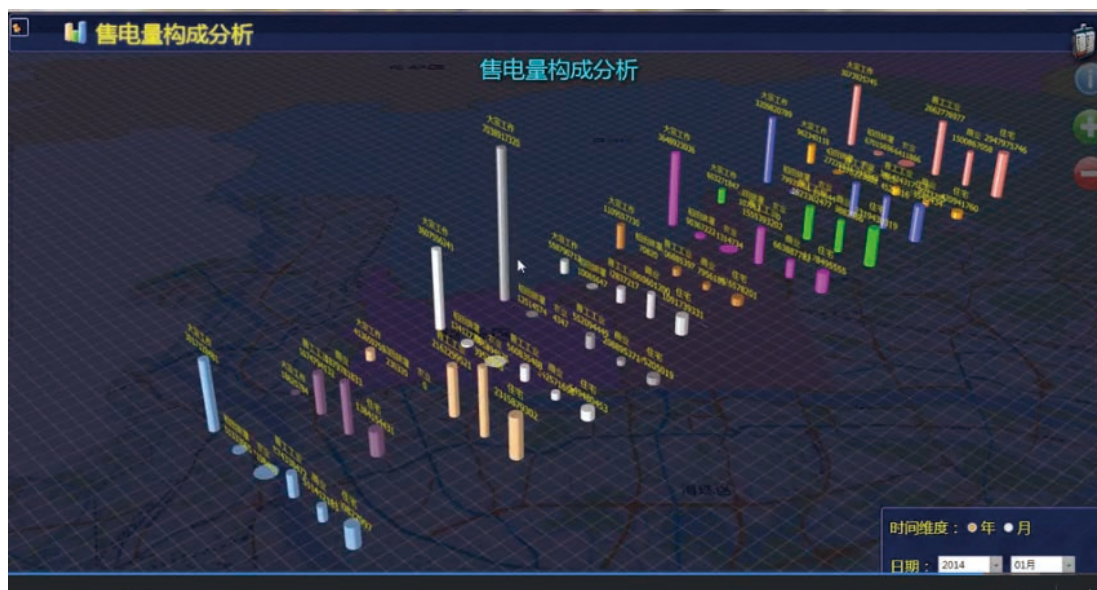
### 系统特点与优势：

本次的华南某市供电局全景可视化大数据平台具有几方面的特点与优势，使之完全区别于过去的旧有系统，更加符合当下的数据量与发展趋势。

首先，全景可视化大数据平台具有稳定的Spark计算引擎。面对企业级需求，企业目前应用开源Spark的主要困难在稳定性、可管理性和功能不够丰富上。开源Spark在稳定性上还有比较多的问题，在处理大数据量时可能无法运行结束或出现Out of memory，性能时快时慢，有时比Map/Reduce更慢，支持的SQL语法仍然非常有限，无法应用到复杂数据分析业务中。然而从星环的Transwarp Inceptor中，上述问题都已经解决并经过实践上线案例的考验，并且Inceptor对Spark进行了大量的改进，极大提高了Spark功能和性能的稳定性。Transwarp Inceptor能稳定的运行7\*24小时，并能在TB级规模数据上高效进行各种稳定的统计分析。

此外，新平台拥有当下最完整的SQL支持。星环SPARK产品实现了自己的SQL解析执行引擎，可以兼容SQL'99和HiveQL，自动识别语法，因此可以兼容现有的基于Hive开发的应用。同时支持比较完整的SQL'99的语法，以及PL/SQL的扩展，可以取代MPP实现数据仓库。星环Inceptor通过与LDAP相结合，实现可靠地安全策略与权限访问控制。目前最新的社区版Spark 1.0.1（包含在CDH5.1中）还不支持JDBC/ODBC，更没有安全控制，因此社区版SPARK产品在进行业务迁移时存在大量的复杂工作以及高度的不确定性。

与此同时，全景可视化大数据平台还支持Memory+SSD的混合存储架构，支持多数据源载入。并且在机器学习和数据挖掘能力以及交互式数据分析能力上有着更好的表现。而开放式的API/接口使用户的程序开发人员可以根据项目的实际需求，利用API编写自己的应用程序。用户的使用场景不限于SQL或者存储过程。



## 展望与规划：

华南某市供电局输变电设备全景可视化的大数据平台经过不断完善，目前运行情况良好，能够满足现有的业务需求和功能，但是在应用中，无论是技术还是业务，都在不断变化，也需要对平台的功能和技术进行不断完善和改进，以便更好地为该市客户服务和提高供电局的客户满意度。

星环信息科技（上海）有限公司

🏠 地址：上海市徐汇区桂平路481号18幢3层301室（漕河泾新兴技术开发区）

✉ 邮编：200233      ☎ 电话：4008 079 976

🌐 网址：[www.transwarp.io](http://www.transwarp.io)

**TRANSWARP**  
DATA HUB