

## 星环助力华数传媒实现Hadoop广电应用

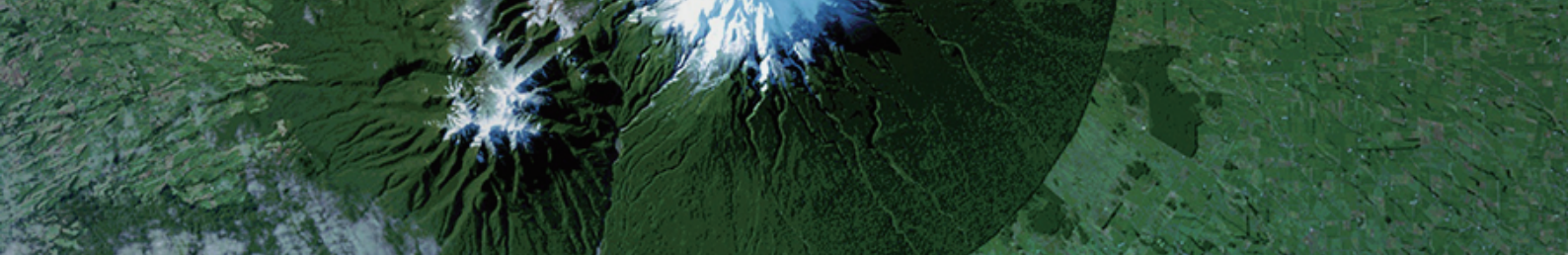
当下大数据之热使得技术界对Hadoop的话题热火朝天。但在日常工作中，企业往往还是遵循既有模式，对于Hadoop到底能否真正帮到企业的应用依然心存顾虑。Hadoop是不是很年轻？这个开源的事物是否符合公司业务级的严谨要求？有没有企业真的应用过？一系列问题萦绕人们心头。这可以理解，毕竟任何一个新生事物出来都要有一个接受过程。

对于Hadoop，其实这些都不是问题。专业人士都知道，Hadoop到现在已有15个年头，这对于一个实用技术的稳定发展已足够长久。事实上，虽然“大数据”一词才出来二三年，但它实际指称的海量的、多类型的数据现象早就有了，不但在互联网领域，更在工业、商业、通信、金融、传媒等存在久远。比如，生产线上巨量传感器数据的接收分析、通信系统全程全网的实时日志文件采集与分析、医疗系统密集数据采集与分析从而帮助快速的科学诊断……所有这些都需要新型的数据处理技术来支撑。Hadoop在这些领域突显了强大竞争力，并在国内外的相关实践中获得广泛应用。

在Hadoop还是开源的时候，众多非IT型企业不具备自我开发条件，由此限制了Hadoop在企业里的普及。但随着专注于企业级市场的Hadoop发行版的技术公司出现，这个问题迎刃而解。从此，各类型企业终于可以安心驾驭真正的企业化“大数据浪潮”了。

在中国，广电系统正经历着数字化浪潮的冲击，基于网络化的影视播放给传统广电运营商很大挑战。在此背景下，广电系统的生力军华数传媒敏锐地意识到，要想获得未来网络化传媒的生存与竞争优势，现在就必需向用户倾斜，以适应未来发展的数据基础架构为依托，才能打造“精准型”广电内容及传播运营商。

2013年，华数传媒的大数据系统完成了从无到有的基础建设，实现了基本应用。然而，华数大数据仍面临很多挑战：数据量增加带来的服务性能压力、数据分析无法满足高时效性业务、业务支撑功能无法满足复杂的商用需求、对网络和服



务器质量等数据分析仍为空白、等等。为此，华数传媒亟需解决方案以解决如下几个问题：

- 数据采集、存储和转发。通过大数据技术满足海量、多来源、多样性数据的存储、管理要求，支持平台硬件的线性扩展，并提供快速实时的数据分析结果，并迅速作用于业务。
- 个性化用户推荐。不仅限于数据本身的分析和决策价值，通过构建在大数据平台之上整合业务能力，为用户提供融合、个性化的内容服务。
- 从内容传输到内容制造。使用大数据挖掘技术提前先于观众知道他们需求，预知将受到追捧的电视。另外，还可通过观众对演员、情节、基调、类型等元数据的标签化，来了解受众偏好，从而进行分析观测，为后续的影视制作等内容开发做好准备。

作为众多广电运营商其中的重要一员，华数传媒的这些需求，也是广电运营商们的共同诉求。

针对华数传媒的需要，国内首家专业的企业级Hadoop发行版厂商星环科技给出了令人满意的答卷。星环科技基于自主研发的一站式Hadoop发行版Transwarp Data Hub (TDH)，综合运用了其中的TDH Hadoop、Inceptor分布式内存引擎、Hyperbase实时数据库等技术产品组件，为华数传媒提供了如下的综合方案：

#### 数字电视分析系统

TDH平台用以整合各个相关数据源数据，包括Portal、CA、CDN、SRM日志、用户使用浏览信息、AAA、BOSS结构化数据、用户基本信息、消费数据、用户上网流量数据、网管数据等。通过TDH的快速分布式数据查询引擎，实现海量数据的秒级查询。为用户提供智能推荐、实时榜单和新媒体指数分析。

- 提供基于全量数据的实时榜单。以时间（小时/天/周）、用户等维度，对点播节目、直播节目、节目类别、搜

索关键词等进行排名分析、同比环比分析、趋势分析等。地区风向标主要以城市和时间等维度分析点播排行、剧集排行、分类排行、热搜排行及用户数量的变化。另外，从时间、频道、影片类型、剧集等维度，根据在看数量、新增数量、结束观看数量、完整看完等分析用户走向。

- 新媒体指数分析。通过对用户行为分析获取很多的隐性指标，从侧面反映用户对业务的认可度、用户的使用行为习惯等。在此基础之上，TDH大数据分析可帮助华数传媒构建规范的指标分析和衡量体系，为业务运营提供强有力的指导。
- 智能推荐。运用星环科技大数据基础架构，通过对用户行为数据的采集分析，进行精准画像，使用智能推荐引擎，实现信息的个性化推荐（TV屏、手机、PC），个性化营销（个性化广告、丰富产品组合、市场分析）。基于可持续扩展和优化智能推荐算法，以及大数据带来的实时数据交互能力，为每一个用户量身定做的推荐节目极大提高了产品的到达率，增强了用户忠诚度。

#### 结语：

对于包括广电运营商在内的众多企业而言，精细化的用户需求掌控是未来发展不可变更的趋势。在此，Hadoop发行版系统再次证明了其业务价值，这是传统数据管理系统力所不能及的。星环科技是国内鲜有的掌握企业级Hadoop和Spark核心技术的高科技创业公司，在电信、金融、交通、政府等领域的落地应用拥有丰富经验。相信通过类似星环科技的Hadoop发行版系统，企业级大数据应用将更能顺应业务发展要求，企业也可更从容面对市场竞争。

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

**TRANSWARP**  
DATA HUB