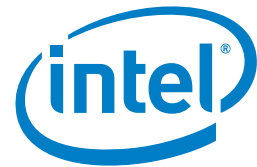


成功案例
英特尔®至强™处理器E5系列
Apache Hadoop®软件的英特尔®发行版
交通
大数据分析



利用大数据分析改进交通管理

杭州诚道科技有限公司通过运行于英特尔®至强™处理器平台上的Apache Hadoop®软件的英特尔®发行版显著提高交通管理能力



“得益于Apache Hadoop®软件的英特尔®发行版的优势，海量图像和视频数据不但实现了可靠和高性能的存储，而且还能被大量的使用者快速地访问和使用。浙江省某市可保存的历史违法数据从3个月延长到24个月，从24亿条过车数据中完成机动车的号牌精确查询和行车轨迹查询，仅需不到1秒的时间。”

陈海清
杭州诚道科技有限公司

浙江省某市经济发展迅猛，地方交通越来越繁忙，机动车辆不断增加，地方政府需要好的方法对过往车辆进行监控管理，从而提供更好的交通服务。地方交通部门采用了数据驱动的方法，在市内重要检查点安装了上千台数字监控设备，这些设备7×24小时不间断捕获图像和视频数据，每月数据量达TB级。现在，当地交通部门面临着如何通过有效利用这些不断增加的交通信息数据该改进交通管理的挑战。

挑战

- **集中管理交通数据。**集中访问分散存储在不同的支队数据中心的图像或视频等交通数据，和道路交通管理设施、装备和应用系统等。
- **优化海量数据利用。**提供尽可能长时间段的车辆监控数据为市公安局、刑侦、经侦部门人员及一线民警等提供信息支撑服务。
- **改善交通。**提高对各种交通突发事件的应急调度能力，依据历史数据预测交通或突发事件的趋势。

解决方案

- **部署统一的数据中心。**22台基于英特尔®至强™处理器E5系列的服务器、198T存储空间形成浙江省某市的交通数据中心，对数字交通信息实现集中存储。
- **部署Apache Hadoop®软件的英特尔®发行版。**利用Hadoop 分布式文件系统（HDFS）和Apache HBase®实现基础过车结构化数据的永久存储，以及最近24个月的交通违法图像数据，实时对数据检索，并可随时无缝扩容。
- **部署诚道重点车辆动态监管系统。**挥英特尔®开放数据分析平台的优势，实现海量的数据的挖掘和分析。

技术成就

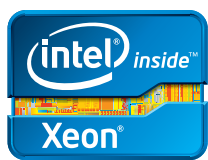
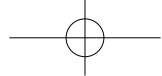
- **大数据存储能力。**Apache Hadoop®软件的英特尔®发行版提供了高度容错性和高吞吐量的海量数据存储解决方案，可为海量信息提供可靠存储，并无缝容量扩充。
- **实现强大I/O处理功能。**英特尔®至强™处理器E5系列增强了系统I/O处理能力，单服务器允许平均每秒250次500KB尺寸图片同步传输或2000次异步并发存储。
- **高性能Hbase数据库。**Apache Hadoop®软件的英特尔®发行版满足了车辆监控系统复杂的数据查询需求，不到1秒的时间即可得到从24亿条过车数据中的机动车号牌查询出的精确结果和行车轨迹。

商业价值

- **提升交通案件侦破能力。**机动车违法图像信息在系统的保存周期从3个月延长到24个月，交通警察等部门可根据车辆的颜色、车型、号牌等信息实时查询其历史行为、行车路线和车辆营运公司、驾驶人等关联信息。
- **增强交通警察对机动车辆的监管能力。**交警可以从24亿条过车数据中轻松检索被监测机动车的号牌精确查询和行车轨迹。
- **便捷利用关联车辆的分析数据。**针对24亿条实际过车数据进行两卡点、多卡点的伴随车辆和碰撞车辆的复杂分析，查询耗时仅为10秒左右。

高速发展的城市面临交通管理挑战

浙江省某市经过几年交通信息化的发展，已经陆续接入了100多套智能监控卡口系统、300多套卡口式电子警察及500余路视频监控。这些数据采集设备获取的结构化数据，例如：时间、地点、车辆信息等集中存储在市交通支队数据中心，而图片和视频等半结构化数据存储在各县数据中心。



英特尔® 至强™

Apache Hadoop*软件的英特尔®发行版颠覆了传统的交通数据存储和处理架构。

交通支队利用这些数据实现了机动车非现场违法取证、行驶轨迹分析、道路通行实时状况分析、交通事故统计分析等业务。

大数据挑战

数据驱动的手段在一定程度上帮助交通管理部门提高了工作效率，改善了交通质量，然而交通数据不断增加，平均每月的数据量达10TB，而且由于图像或视频等数据被分散存储在不同的支队数据中心，跨数据中心使用这些数据非常不方便。同时有些道路交通管理设施、装备和应用系统等处于单独、孤立运行状态，作用比较单一。浙江省某市期望将这些分散的数据进行资源整合，充分发挥它们应有的作用。

浙江省某市目前可保存最近12个月的交通数据，历史数据显示，交通数据量在以每年60%的速度增长。在现有的数据存储方式下，可保存数据的周期不但会越来越短，同时面临数据可靠性的风险。该市需要延长交通数据的保存周期，以对尽可能长时间的数据进行查询和分析，为市公安治安、刑侦、经侦部门人员及一线民警等提供信息支撑服务，因为在很多情况下，机动车辆交通数据可为案件侦破等提供关键的证据和信息。

除了集中管理数据的需求，该市还面临着数据存储的挑战。一方面伴随城镇化发展的深入及城市化的扩大，浙江省某市道路监控的规模不断增大，信息化建设的不断推进及监控密集度要求的不断提高，监控点的建设逐年增加，这些不断增加的监控设备采集到的数据都需要及时的完成存储，另一方面新技术的不断出现，电子警察卡口功能的升级，例如：图像视频清晰度等因素，导致单个图像或视频数据比以往更大。这些都对存储性能提出了更高的要求。

此外，浙江省某市还面临着交通数据无法被充分利用的难题。首先依靠手工方式实现数据查询和统计分析的效率伴随数据量的增加变的越来越低下。其次，这些数据对城市的交通管理存在潜在的价值，该市希望将视频监控的应用范围从过去和目前以事后查看为主，转变为事前预警，从交通数据中挖掘出潜在的交通风险，为公安、交通等各行各业提供更为有效的业务信息支持。

使用Hadoop处理大数据

面对大数据挑战，该市和杭州诚道科技有限公司紧密合作。根据该市具体的需求，诚道科技为该市部署了基于Apache Hadoop*软件的英特尔®发行版和英特尔®至强™处理器E5系列的诚道重点车辆动态监管系统，通过集中的数据中心将全市卡口、电子警察、视频监控、流量检测设备、信号机、诱导设备等有效的连接起来。

浙江省某市实现了高吞吐量和高度容错性的海量数据存储目标。首先，服务器使用英特尔®至强™处理器E5系列，提供了强大的数据I/O能力，22台基于英特尔®至强™处理器E5系列的服务器、198T存储空间形成该市数据中心，经实际数据测试，单服务器可允许平均每秒250次500KB的全高清图图像图片提供存储，在异步传输模式下可以支持平均每秒2000次并发接入存储。换言之，即便浙江省某市全部监控点持续不断的有多台车辆经过，系统也能实时完成可靠的数据存储。

其次，Apache Hadoop*软件的英特尔®发行版通过多方面的优化保证数据的可靠性。数据在写入时被自动复制和分布到Hadoop服务器集群。数据在读取时自动进行数据校验，一旦发现数据校验错误将重新进行复制。英特尔®针对开源Hadoop的源代码做了大量的改进和优化，实现了更高的性能和可靠性，例如：通过改进分布式文件系统（HDFS）数据的分布和读取算法增强系统的I/O性能。

通过利用Apache Hadoop*软件的英特尔®发行版，该市基础过车数据实现了永久存储，可保存的历史交通违法图像数据从3个月延长到24个月，而且数据更加安全可靠。更重要的是，该系统允许在不宕机的情况下实现无缝容量扩充，满足了该市随时对系统进行扩容以保存更长历史数据的需求。

Apache Hadoop*软件的英特尔®发行版通过提供包括自动配置、管理和监控等工具，简化Hadoop集群的部署和管理。诚道重点车辆动态监管系统利用Apache Hadoop*软件的英特尔®发行版，帮助该市实现了对数据进行智能、高效的分析和处理，提高了车辆轨迹分析、套牌分析、车辆查缉布控等系统的查询性能及实时性，充分发掘数据背后的隐藏价值。

经验教训

- 运行在通用硬件上的Apache Hadoop*软件的英特尔®发行版可增强数据存储的可靠性和数据吞吐性能，是交通领域海量半结构化图像和视频等数据的理想存储解决方案。
- 英特尔®至强™处理器E5系列提供强大的运算能力和I/O处理性能，可以从容应对交通监控领域密集数据采集设备数据快速并行读写的挑战。
- Hadoop平台适合用于存储、挖掘和分析交通行业的图像和视频等半结构化数据。

Apache Hadoop*软件包含一个分布式计算框架：Mapreduce*用于处理结构化和半结构化的数据，诚道重点车辆动态监管系统利用该框架实现了海量的视频和过车图片的数据挖掘和检索，在交通信息数据中实现伴随车辆和碰撞车辆的高速数据挖掘分析。诚道对真实过车数据的测算显示，针对24亿条号牌实际数据两卡点、多卡点的过车信息的碰撞分析中，查询耗时仅为10秒左右。

此外，Apache HBase非关系型数据库解决了关系型数据库在处理海量数据时的理论和实现上的局限性。测试表明，从24亿条数据中精确查询机动车号牌和行车轨迹，仅需不到1秒的时间，即可得到查询结果。

诚道重点车辆动态监管系统还利用Hive*数据仓库技术，实现驾驶人信息、机动车信息、道路监控数据、事故信息、违法信息和公安盗抢监控信息的关联信息高效查询，以灵活简单方式实现了很多非常复杂的计算过程，这些分析产生的数据可用于决策指挥，实现对各种交通突发事件的调度处理。

面向未来，诚道公司将和英特尔®公司在大数据分析等领域继续合作，利用更新的技术为客户提供更智慧的智能交通管理解决方案。

找到适合于您公司的解决方案。请与您的英特尔代表联系，访问英特尔IT管理成功案例（www.intel.com/itcasestudies）或访问英特尔IT中心网站（www.intel.com.cn/itcenter）

本文档及其信息供英特尔客户的便利之用，“按现状”提供，不作任何明示或暗示的担保，包括任何可销售、适用于特定目的以及不违反知识产权的暗示保证。收到或拥有本文档不代表授予任何由此所描述、显示或包含的知识产权。英特尔®产品未计划用于医学、救生、生命维持、关键控制、安全系统或核设施应用。

用于性能测试的软件和工作负荷可能只针对英特尔微处理器进行了优化。SYSmark 和 MobileMark 等性能测试使用特定的计算机系统、组件、软件、操作和功能来测量性能。所有这些因素稍有变化都可能导致不同的结果。如果您有采购打算，应该参考其他信息和性能测试，以进行充分的评估，包括该产品与其他产品结合使用时的性能。

* 2013英特尔公司。保留所有权利。英特尔标识和英特尔至强是英特尔公司在美国和其他国家或地区的商标。

*文中涉及的其它名称及商标术语各自所有者资产。

0326/SHA/PMG/XX/PDF

328805-001 CN