



CERN 利用 Quantum StorNext 文件系统加快研究和发现

收集数十亿个粒子之间的碰撞数据需要一个快速、可靠、可扩展的 IT 基础结构。欧洲原子能研究机构 CERN 部署了高性能的工作流程和智能归档解决方案 Quantum StorNext 来帮助管理其 ALICE 实验生成的数据。ALICE 是“一种大型离子对撞实验”，它收集大量的数据并依靠 StorNext 来提供高速共享工作流程操作。他们创建了共享数据池，以支持更快地处理文件，并让科学区更快地了解和探索新的想法和发现。

世界领先的粒子物理学实验室

CERN—欧洲原子能研究机构—是世界领先的粒子物理学实验室。该实验室成立于 1954 年，总部设在瑞士日内瓦，是欧洲最早的合营机构之一，现有 20 个成员国。CERN 最近建立了大型强子对撞机 (LHC) — 一种粒子加速器。在 LHC 中，超高能量质子将与质子碰撞，重离子（比如铅的原子核）将与重离子碰撞。这将使科学家能够进一步渗透到物质的结构中，并在“宇宙大爆炸”之后的几亿分之一秒内重建宇宙中的普遍条件。

分配给 LHC 的实验之一就是 ALICE。这是世界上专门以无限小规模研究物质物理学的最大实验之一。该实验在 CERN 进行，来自 30 个国家的 1,000 多名物理学家、工程师和技术人员参与了这场国际协作。他们一起解决了基础物理学中最新的难题之一：重新计算物质的起源。ALICE 投入运行后，粒子碰撞将生成通过检测器收集的大量数据。

所有这些皆需要一个可靠的数据获取、选择、传输、存储和处理系统来管理每秒生成的数十亿字节数据。过去几年间，ALICE 团队一直在开发一种尖端的信息管理系统，以帮助捕捉、记录和分析数据。为了全面了解这些难题，系统需要相当于每秒 1 GB 的性能，而且在一次实验的至少一个月时间里必须昼夜保持这个速度。这相当于在该月期间将累积超过 1 PB 的数据。

所有这些数据全部驻留在 CERN 主计算中心的大规模存储系统上。问题在于 ALICE 大型强子对撞机实验将在距离这个计算中心 3 公里之外的地方进行。项目经理 Pierre Vande Vyvre 阐述了团队所面临的技术挑战。

“我们需要在实验地点和大规模存储站点之间进行数据的缓冲，”他说，“ALICE 数据获取系统包含足够的存储设备，能在数据迁移到计算中心之前支持数小时的自主数据获取。我们把数据存储到磁盘，然后脱机传送到主计算中心。”

选择最合适的 CFS

该存储设备在数十个集群节点之间共享，但是这些计算机上使用的 Linux 操作系统的当前版本不包含可在节点之间共享的真实文件系统。数据获取 (DAQ) 系统经理 Ulrich Fuchs 一直密切参与了实施工作。他说：“我们需要能够快速轻松地访问所有主机上的数据，而不必通过缓慢的网络连接传递数据。Linux NFS 是一种可能性，但它在集合带宽以及客户端节点数量的可扩展性方面显示了其局限性。”

因此，Vande Vyvre、Fuchs 以及 CERN 团队的其他成员调查了可用于集群文件系统 (CFS) 的不同选项。对 CFS 的主要要求是最大的集合带宽性能 (ALICE 需要高达 3 GB/s 的读写流量



“ALICE 数据获取系统包含足够的存储设备，能在数据迁移到计算中心之前支持数小时的自主数据获取。我们把数据存储到磁盘，然后脱机传送到主计算中心。”

Pierre Vande Vyvre
CERN 项目经理

主要优点

- ☛ 在降低存储成本的同时实现连续的数据访问
- ☛ 在评估的所有方案中提供最高的文件共享性能
- ☛ 使用户能够访问多个主机的数据，从而加快研究分析
- ☛ 支持广泛的操作系统、平台和存储，改进协作
- ☛ 保护用户的选择自由性，并消除对任何单一供应商的依赖性
- ☛ 基础结构以最高的性能运行，利用 Affinity 把数据指向特定的主磁盘

解决方案概述

- ☛ Quantum StorNext
- ☛ 20 个 InfoTrend 光纤通道 4 GB/s 磁盘阵列
- ☛ 60 台运行 Linux 的服务器
- ☛ QLogic 光纤通道适配器和交换机

性能)；最小的硬件设备占用空间 (ALICE 数据获取室只有 60 平方米)；管理大量客户端的可扩展性 (最多 100 个用户节点)；重要的是 CFS 与基础硬件之间的独立性。

在最初考虑的半打产品中，他们调查了三个：Quantum 的 StorNext from、Client File Systems 的 Lustre 和 TerraGrid 的 TerraScale。初步测试之后，由于不切实际的操作系统支持需求，Lustre 解决方案很快被排除在外。随后，他们在相同的硬件 (1 个 SAN 架构、6 个存储阵列和 9 个客户端节点) 上对剩下的 StorNext 和 Terrascale 文件系统进行了大量测试。目标性能至少为所需全局性能的五分之一，即写性能 300 MB/s、读性能 300 MB/s 和综合读写性能 600 MB/s。

在几个星期的时间里，他们对这两个解决方案进行了密集测试。最终结果是 Quantum StorNext 提供了明显高于 TerraScale 的性能。这可以用 StorNext 直接利用所用存储设备的光纤通道连接 (4 GB/s)，而 TerraScale 采用通过现有以太网连接 (1 Gbit/s) 来连接客户端的服务器这一事实中来说明。

在部署的初始阶段，ALICE 数据获取团队从一台服务器、一个客户端和一个磁盘阵列开始，随后增加到 20 个磁盘阵列和 60 个连接 16 端口交换机的客户端。Fuchs 对 StorNext 最强大的资产：Affinity 十分赞赏。这使得该团队能够通过写入 Affinity 的关联关系点把数据指向特定的主磁盘。他解释了该功能如此重要的原因：“如果我们把所有磁盘阵列放进一个全局文件系统中，就会让两台计算机连接起来，这样读写将并行进行。结果是磁盘会让大量的流量进出同时进行。我们的测试表明性能下降了 90%。Affinity 使我们可以把数据流量指向预先确定的磁盘，这意味着所有计算机将始终以最大性能运行。”

快速、高效的数据获取

尽管已经在实际环境下进行了高水平的试点和测试，但 ALICE 数据获取系统计划在今年晚些时候投入运行。将从计算机中的探测器接收数据的光学接受器用于生成数据，而且这使该团队可以测试数据获取速率。在 18 天的测试期间，总共生成和存储了 1 PB 的数据。在第四个昼夜，数据获取速率达到了必备的 1 GB/s。Vande Vyvre 说：“ALICE 数据获取系统以及 CERN 传输和存储网络基础结构已经准备就绪。”

StorNext 提供了 ALICE 要求的高速度、共享工作流程操作以及大规模多层归档。主机应用程序可以通过一个名称空间访问所有文件，不管文件位于何处，都能实现透明访问。

CERN 团队对 StorNext 的性能、平台独立性和可扩展性十分满意，甚至提名该软件参评由 ALICE 代表协作委员会评判的一项工业奖。Vande Vyvre 说：“由于其出色的设计和执行以及在预算内准时满足苛刻的异常要求的能力，Quantum StorNext 获得该奖项当之无愧。”

他总结说：“必须感谢参与 CERN ALICE 实验的后勤人员。他们承担了分析所需的大量数据。数据是 CERN 最宝贵的商品。Quantum StorNext 对于快速可靠地收集这些数据大有帮助，因而使得科学摄取能够了解和探索新的想法和发现。”

关于 CERN

CERN (欧洲原子能研究机构) 是世界最大的粒子物理学中心。CERN 实验室位于日内瓦附近，那里汇集了来自世界各地的科学家来研究物质的构建块以及把它们集中到一起的力。CERN 主要为它们提供加速器，这些加速器可以把粒子加速到接近光速，并有助于设计和建造使粒子可见的探测器。该实验室成立于 1954 年，是欧洲最早的合营机构之一，现有 20 个成员国。



欲获取详细信息，请访问 Quantum.com/cn

北京代表处

北京市朝阳区东三环北路霞光里18号佳程广场B座10层D单元
邮编：100027
电话：86.10.58299977
传真：86.10.58299978

上海代表处

上海市淮海中路93号大上海时代广场办公楼24层
邮编：200021
电话：86.21.63910175/76
传真：86.21.63910016

香港代表处

香港铜锣湾希慎道 33 号利园
宏利广场 45/F
电话：852.31802229
传真：852.31801717

关于 Quantum

Quantum Corp. (NYSE:QTM) 是专门提供备份、恢复和存档解决方案的全球领先的存储公司。通过结合针对性的存储专业技能、客户驱动的创新以及平台不相关的特点，Quantum 提供了全系列的磁盘、磁带、介质和软件解决方案，并有世界级的销售和服务机构做后盾。作为一家长期、可信的合作伙伴，Quantum 与广泛的经销商、OEM 以及其它供应商密切合作，以满足客户不断发展的数据保护需求。