

面向 100PB 级海量数据处理的存储系统：现状及未来规划

Thursday, 30 May 2019 16:40 (20 minutes)

高能物理计算是典型的数据密集型计算，预计在未来十年内，高能所计算中心将建成 100PB 级的海量存储系统，为 BESIII, JUNO, LHAASO, HEPS 等高能物理实验及大科学装置提供数据存储及 I/O 服务。报告首先将从 IT 系统的角度，介绍各实验数据处理和分析对存储的需求和挑战，主要包括访问性能、可扩展性、高可用和高可靠性、可管理性、分级存储及性价比、实验数据跨域访问以及数据长期保存等。然后报告将分别介绍 Lustre 和 EOS 两个主要的磁盘存储系统针对以上需求和挑战的功能设计，这些功能设计在高能所的应用现状和问题以及我们针对这些问题做的二次开发等。

特别地，我们将结合实例向软件开发人员介绍在程序性能调试过程中，存储系统能够提供的性能检测工具和手段。

之后，报告还将针对软件共享和长期保存、个人数据多设备统一视图等功能需求，介绍 AFS, CVMFS, IHEPBOX 等存储技术以及在高能所的部署应用实践。

最后，我们将介绍近期对“EOS 数据副本 + JBOD”和“Lustre+ 硬件数据冗余”两种架构的性能测试结果以及对未来存储架构的设计和考虑。

Primary author: Ms WANG LU, Lu (高能所)

Presenter: Ms WANG LU, Lu (高能所)

Session Classification: 数据与计算平台

Track Classification: 数据管理