

第二十届全国科学计算与信息化会议

Report of Contributions

Contribution ID: 1

Type: 分会报告

基于分布式数据流平台的中子谱仪通用化数据处理框架

Monday, 10 July 2023 15:15 (15 minutes)

中国散裂中子源的二期工程项目中，使用了一套基于数据流的通用化框架作为所有束线谱仪的技术方案。这套全新规划的设计根植于中子谱仪的运行模式和数据特征，内容上涵盖了从底层硬件规范到上层应用拓扑的各个环节，旨在解决传统中子谱仪由于系统耦合复杂造成的低效状况，同时提供灵活多样的部署配置和便捷高效的运维管理。针对未来中子束流通量将大幅提升的趋势，新设计利用业界优秀开源工具，结合对硬件接口的规范和软件适配，实现了基于分布式消息网络的数据流处理平台，并进一步扩展 EPICS 控制、监测报警与显示服务等标准化接口，在确保系统稳定性和冗余性的基础上实现了全链条的数据处理服务，并达到了极高吞吐压力的承载性能。新的通用化设计框架大幅提升了中子谱仪的建设和运行效率，同时亦可有效降低部署和运行维护阶段的人力成本，其将是推动材料分析等相关学科和应用产业进一步繁荣发展的重要因素。

Summary

Primary authors: 滕海云, UNKNOWN (高能所); 邱, 勇翔 (中国科学院高能物理研究所)

Co-authors: Prof. SUN ZHIJIA; Mr JIAN ZHUANG, jian (高能所); Dr CAO, ping (USTC, HeFei)

Presenter: 邱, 勇翔 (中国科学院高能物理研究所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 2

Type: 分会报告

基于 OMAT 的一平台多中心监控方案设计与实现

Monday, 10 July 2023 16:45 (15 minutes)

高能物理计算领域一平台多中心方案，旨在整合分布在各个地域的高能物理计算、存储资源，形成逻辑统一的大规模集群，通过统一管理、运维、资源调配作业分发，降低多计算中心运维成本，优化计算、存储资源配置，实现硬件资源和业务需求的高效对接，从而缓解高能物理日益增长的科学计算需求。

为更好的应对各个地域的站点资源运维管理，通过对高能所计算集群开放运维分析平台(OMAT)进行重构，设计实现了支持异地实时采集分析的一平台多中心监控设计方案。

监控范围层面，该方案不断拓展采集手段已支持节点性能、系统服务、存储性能、日志分析、动力环境、网络流量、作业调度和资源使用情况等领域的数据采集和分析展现。运维告警层面，该方案设计实现了统一告警平台，支持监控数据基于业务告警策略实时分析和告警，同时支持异地数据漏采实时告警，保证跨地域数据采集的真实性、可靠性和完整性。监控数据可视化层面，该方案基于数据可视化实际需求，设计实现了面向特定应用需求的数据缓存层用于存储关键的指标型监控数据，使用极小的存储空间，实现监控平台业务层数据和核心层数据的解耦合，保障监控数据的持久化。

该监控方案现已应用于高能物理一平台多中心的日常运维，较好的实现高能所站点和其他异地站点的日常运维监控需求。

Primary author: Mr HU QINGBAO, 胡 (高能所)

Co-authors: Mr JIANG XIAOWEI, Xiaowei (高能所); Dr CHENG, Yaodong (IHEP); Mr WEI, Zheng (Sun Yat-sen University); Ms 石, 京燕 (高能所)

Presenter: Mr HU QINGBAO, 胡 (高能所)

Session Classification: 科学计算技术与平台

Track Classification: 科学计算技术与平台

Contribution ID: 8

Type: 分会报告

高能物理容器镜像管理服务

Monday, 10 July 2023 16:30 (15 minutes)

随着容器应用范围越来越广泛普遍，成为高能物理实验中不可缺少的一部分，随之对容器镜像的管理要求越来越高，安全也更加严格。本文主要介绍实现高能物理容器镜像管理的方法、架构、安全以及典型的应用场景等。

Summary

Primary author: Mr ZHENG, Wei (IHEP)

Co-authors: Ms 石, 京燕 (高能所); Mr HU QINGBAO, 胡 (高能所)

Presenter: Mr ZHENG, Wei (IHEP)

Session Classification: 科学计算技术与平台

Track Classification: 科学计算技术与平台

Contribution ID: 12

Type: 分会报告

海洋科学大数据汇交与共享管理系统建设

Wednesday, 12 July 2023 16:45 (15 minutes)

科研数据的价值在大数据时代是重者恒重，特别是对于以调查观测为基础的海洋科学，科研观测数据尤为重要。而科研人员掌握的大量原始调查观测数据，一般只存放在个人电脑或移动存储设备中，造成了数据“各自为战”的局面。如何帮助科研人员科学管理分散的数据，让数据活起来，更加充分的发挥数据效能，创造更多的价值，是现在科研数据管理的重要工作之一。

为更好支撑海洋科学研究，推进海洋科学数据共享，设计研发了海洋科学大数据汇交与共享管理系统。系统采用 B/S 模式，按照数据层、服务层、应用层三层架构设计，统一部署，并将应用服务器和数据服务器分离。提供海洋科学数据的在线化、流程化、自助式的汇交、管理和共享服务以及计算资源服务，支持科研人员与数据中心共同建设、管理、运维数据，实现多源数据融合汇聚，集中存储，统一标准，规范管理，实现数据的四统一，即统一编码、统一管理、统一存储、统一发布。汇聚多渠道数据资源，包括科研项目数据汇交管理、论文关联数据汇交、DOI/CSTR 注册服务以及相应审核流程管理。统一数据资源编码和标准，设计标准化的元数据格式，涵盖了元数据信息、联系信息、摘要和详细描述、扩展信息、相关文献、使用声明以及实体数据、数据图和数据说明文件等。提供数据集/数据产品的统一展示与发布，包括中英文两种方式的数据展示、检索和下载数据。实现数据分级分类共享，公开共享的数据提供直接下载，申请共享的数据提供申请下载，经审批流程审批后可以将数据共享给申请人。实体数据集中存储管理，在到数据服务器，保障数据安全，统一管理和检索。

系统具有如下特色：

- (1) 设置标准的元数据存储格式，可下载元数据模板，通过文件导入实现数据快速入库存储。
- (2) 科研人员通过“个人中心”可自助管理和维护各自汇交的数据、论文、成果等信息，并可根据需要更新数据和数据信息，查看数据浏览、下载情况，掌握数据流向。
- (3) 统一展示和发布“专题数据产品”，为每个数据集/数据产品设置独立的中英文网页，全面展示数据信息的同时，可以方便地查询下载数据。
- (4) 提供微信小程序展示数据产品，更便捷。
- (5) 科研人员提交 DOI/CSTR 申请后，经审核通过可以自动生成 DOI/CSTR 注册码并发布链接地址。
- (6) 系统管理模块，提供网站用户管理、权限管理和日志信息等，保证数据安全。

Summary

海洋科学大数据汇交与共享管理系统采用 B/S 模式，按照数据层、服务层、应用层三层架构设计，统一部署，并将应用服务器和数据服务器分离。提供海洋科学数据的在线化、流程化、自助式的汇交、管理和共享服务以及计算资源服务，支持科研人员与数据中心共同建设、管理、运维数据，实现多源数据融合汇聚，集中存储，统一标准，规范管理，实现数据的四统一，即统一编码、统一管理、统一存储、统一发布。

Primary author: 任, 焕萍 (grace)

Co-author: 郑, 双强 (中国科学院海洋研究所)

Presenter: 任, 焕萍 (grace)

Session Classification: 科研信息化管理与系统

Track Classification: 科学数据存储与管理技术

Contribution ID: 16

Type: 分会报告

Back-n 多用途时间投影室

Monday, 10 July 2023 14:45 (15 minutes)

Back-n 白光中子源依托中国散裂中子源建设, 利用质子打靶产生的反冲中子进行核数据测量、探测器标定、材料辐照和单粒子效应等方面的研究。在 Back-n 一期建设, 研制了基本的实验谱仪/探测器, 如: 多层裂变电离室探测器用来测量裂变截面和全截面; C6D6 谱仪用来测量中子俘获截面; 基于 $\Delta E-E$ 的阵列探测器用来测量中子诱发轻带电粒子。其中轻带电粒子出射实验采用的方式是在真空靶室内布局一组 $\Delta E-E$ 探测器, 测量轻带电粒子能量并进行粒子鉴别。其主要特点是粒子鉴别方法简单。其不足之处在于: $\Delta E-E$ 鉴别方法有最低鉴别阈值, 无法鉴别低能出射带电粒子; 探测器个数受空间限制覆盖角度小、探测器效率低; 整个系统必须位于真空靶室内, 系统结构复杂受限制较大; 测量的精度收到探测器颗粒度限制, 无法精确测量出射粒子角度。核裂变截面测量探测器为快响应多层裂变电离室, 这样优点在于结构简单。缺点在于信号小, 受干扰较大; 没有粒子鉴别功能, 本底 α 影响下精度难以提高; 对靶的均匀性要求较高等。在束剖面监测装置使用的是 Micromegas 探测器, 由于探测器不具备顶点识别功能, 使用重心算法造成性能限制, 无法做到更好的位置分辨。以上这些探测器在每次测量中都需要进行真空管道拆装, 探测器调试的等过程, 浪费了大量宝贵的束流时间。与此同时, 每种探测器都有些不足之处, 制约了高精度数据测量, 亟需改进。

在未来基于 Back-n 白光中子源的高精度数据测量中, 新型高效的探测器设备是重中之重。一些基于高密度闪烁体及半导体阵列的 γ 射线探测器预研已经开展。在裂变碎片和带电粒子探测器的研究上, Back-n 合作组目前着眼于使用更具有前沿特点的高性能气体探测器——时间投影室。

时间投影室 (TPC) 探测器是在 1978 年由 Nygren 从屏栅电离室 (GIC) 改进而来的。通过测量带电粒子电离出的电子漂移时间结合二维位置灵敏读出, TPC 探测器可以重建出非常好的 3D 粒子径迹, 通过粒子能量损失和径迹长度测量, 它可以给出元素粒子分辨甚至可以给出轻带电粒子的同位素的分辨。TPC 可以覆盖几乎所有粒子的出射角度, 并保持较低的物质质量, 更适应在较低能量的核反应产物测量。对比传统探测器, TPC 具有以下几个明显优势: 具备较强的带电粒子鉴别能力; 可以测量能量极低的次级反应产物; 可以覆盖 4π 立体角, 且具备微分截面测量能力。TPC 的应用会将现有的核数据测量精度大幅度提高, 从而达到核工程及核物理科学需求。

Back-n 多用途时间投影室可以作为裂变截面、带电粒子截面、白光中子能谱及束斑的高精度测量的重要手段。在未来, 还可以应用在高 γ 本底的中子辐射场及高精度中子成像实验中, 成为核数据测量、中子探测的有力工具。

Summary

Primary author: Dr 樊, 瑞睿 (高能所)

Co-author: 孙艳坤

Presenter: Dr 樊, 瑞睿 (高能所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 17

Type: 分会报告

反角白光中子源 ΔE - ΔE -E 轻粒子探测望远镜探测阵列研制及应用

Monday, 10 July 2023 14:30 (15 minutes)

中子诱发轻带电粒子出射 (n,lcp) 反应的截面测量, 特别是双微分截面的测量有重要意义, 但由于中子源及测量技术的限制, 目前该能区的数据缺乏测量。随着中国散裂中子源 (CSNS) 的建成, 反角白光中子源 (Back-n) 于 2018 年投入运行。白光中子源可以提供 eV 至 MeV 宽能区的中子束流, 为 (n, lcp) 研究提供了良好的研究平台。为了进行 (n, lcp) 反应研究及反应截面测量, 需要对反应产生的轻带电粒子产物进行粒子鉴别。除了使用 ΔE -E 方法进行核素鉴别外, 在白光中子源建设了一套 ΔE - ΔE -E 轻粒子探测望远镜探测阵列。 ΔE - ΔE -E 望远镜系统分为硅阵列探测器、 ΔE -E 望远镜系统、 ΔE - ΔE -E 望远镜系统三个阶段完成, 并且可以用来开展不同物理目标的实验测量。硅阵列探测器由 15 块 500 μ m 厚的 PIN 型硅探测器组成, 每块硅探测器的有效探测面积是 20mm $\times$ 25mm。每块硅探测器距离靶的中心的距离是 20cm, 15 块硅探测器可以覆盖 18 到 160 度的粒子出射角。探测器的信号经由 Mesytec 公司生产的 MSI-8 前放放大处理后进入白光中子源的共用电子学进行波形数字化及存储。共用电子学基于 3U 的 PXIe, 波形数字化的采样率为 1G Hz, 采样精度为 12bit。使用四组份 α 源 (^{234}U : 4.775 MeV; ^{239}Pu : 5.155 MeV; ^{238}Pu : 5.499 MeV; ^{244}Cm : 5.805 MeV) 进行探测器测试及能量刻度, 得到每块探测器的能量刻度系数。对于 5.499MeV α 源, 硅探测器的能量分辨率好于 1%。利用低气压多丝正比室探测器与硅探测器, 以及硅探测器与 CsI(Tl) 探测器组成的 ΔE -E 望远镜系统, 在束实验获得了 0.5-100MeV 的质子的 ΔE -E 鉴别结果。

使用硅阵列探测器及 ΔE -E 望远镜系统, 在白光中子源开展了 $^6\text{Li}(\text{n}, \text{t})\alpha$ 、 $^1\text{H}(\text{n}, \text{n})^1\text{H}$ 以及 $^{10}\text{B}(\text{n}, \alpha)^7\text{Li}$ 实验研究。在实验汇总, 利用硅阵列探测器输出的波形, 使用平滑滤波等波形处理方法, 提取信号的幅度及时间信息。由波形的时间信息可以得到相对应的中子的飞行时间, 进而得到中子的能量。在每块硅探测器的幅度-飞行时间二维谱中, 可以看到反应产物的粒子鉴别带。在 $^6\text{Li}(\text{n}, \text{t})\alpha$ 反应中, 得到了 t 和 α 的鉴别带; 在 $^{10}\text{B}(\text{n}, \alpha)^7\text{Li}$ 反应中, 得到了 ^7Li 和 α 的鉴别带; 而在 $^1\text{H}(\text{n}, \text{n})^1\text{H}$ 反应中, 清晰地得到了质子的鉴别带, 以及能量大于 6.5MeV 的质子穿透硅探测器的特征。硅阵列探测器中的每块探测器均得到了良好的核素鉴别效果, 意味着测量到了反应的截面角分布。在数据分析中, 在 eV-MeV 能区取出 80 个能点, 进行单个能点的截面角分布分析, 得到能点的微分截面分布。同时, 通过勒让德多项式拟合, 得到总截面。在 $^6\text{Li}(\text{n}, \text{t})\alpha$ 、 $^{10}\text{B}(\text{n}, \alpha)^7\text{Li}$ 与 $^1\text{H}(\text{n}, \text{n})^1\text{H}$ 研究工作中, 得到了系统的截面结果; 经过与 ENDF/B-VIII.0, JEFF-3.3 等数据库中的已知部分结果对比, 本工作在已有数据部分与数据库中的数据符合很好, 同时也弥补了数据库中的数据缺失。

ΔE - ΔE -E 轻粒子探测望远镜探测阵列的整体方案于 2019 年通过专家评审, 探测器于 2020 年组装完成, 望远镜阵列内有 16 个低气压多丝正比室探测器, 16 块硅探测器以及 16 块 CsI(Tl) 探测器, 每个望远镜阵列腔体内有接线 100 余根。2020 年 7 月初在 Back-n 开展了 48 小时的测试实验, 目前正在计划利用 ΔE - ΔE -E 轻粒子探测望远镜探测阵列开展物理实验。

此外, 在 Back-n 也开展了硅探测器的脉冲形状甄别等相关研究工作, 目前取得了不错的进展。

Primary author: Dr JIANG, Wei (IHEP)**Presenter:** Dr JIANG, Wei (IHEP)**Session Classification:** 核电子学与探测技术**Track Classification:** 核电子学与探测技术

Contribution ID: 20

Type: 分会报告

高气压 ^3He 的 MWPC 在散裂中子源上的应用

Monday, 10 July 2023 14:00 (15 minutes)

中子散射技术是研究物质结构和动力学性质的理想探针，已被广泛应用于诸多领域的研究中，与 X 射线散射技术相比，中子散射技术具有穿透能力极强、轻元素敏感、同位素分辨以及磁结构微观分析等不可替代的优势。随着中国散裂中子源的正式启用，中子探测器作为中子散射实验中的关键设备之一，需要具有高计数率、高探测效率、高位置分辨、灵敏面积大、 n/γ 抑制比好等性能。高气压多丝正比室 (MWPC) 中子探测器，是采用 ^3He 气体作为中子探测媒介，其高探测效率、高位置分辨、良好的二维读出能力是其优势，也正是散裂中子源需要的中子探测器类型。

基于 ^3He 气体的高气压多丝正比室 (MWPC)，整个过程自主研发，包括高气压室体的设计、研制，内部关键部件 MWPC 的制作测试、充气平台的设计研制、气密性、保压测试等均达到预期的要求。探测器有效面积 $200\text{mm} \times 200\text{mm}$ ，为提高探测器的探测效率，气体比分采用 $6\text{ atm}^3\text{He} + 2.5\text{ atm}^3\text{C}_3\text{H}_8$ ，样机在散裂中子源 #20 束流线进行性能测试，测试结果显示，热中子探测效率达到 64.8%，位置分辨在 $1.4\text{mm} \times 1.3\text{mm}$ ，探测器结构与性能均达到设计指标，完成了整个探测器的验收，也进一步证明该探测器的设计合理，目前探测器能达到的指标如下：

- 有效面积 $200\text{mm} \times 200\text{mm}$
- 阳极丝丝间距 2mm
- 气体比分 ($6\text{ atm}^3\text{He} + (2.5\text{ atm})^3\text{C}_3\text{H}_8$)
- 探测效率 64.8% (@ $2\text{\AA}$)
- 位置分辨 (FWHM) $1.4\text{mm} \times 1.3\text{mm}$
- 计数率大于 100kcps

目前，高气压 MWPC 中子探测器作为 CSNS 反射谱仪的主探测器之一，安装在谱仪反射臂上，已完成了探测器有效计数率的升级、散射室内电机噪声影响的处理、散射室内强 γ 本底的屏蔽等，也对探测器做了进一步的优化处理，工作状态达到最优。在测试样品反射光过程中，随着反射光角度增大，反射光会变弱，当该样品的反射光角度 $R_z = -3.2$ 时，反射光计数率约 0.23cps，比较弱的反射光，但在探测器上的二维成像很明显，在位置谱上也能看到明显的反射光峰位。

高气压 MWPC 中子探测器在研制过程中，克服了很多的困难，包括镀金钨丝的布局，核心部件的选择，读出方式的优化，工作条件的判选，现场环境的处理等等，深入了解谱仪对探测器的需求，取得了很好的经验。目前，探测器在反射谱仪上使用期间，性能非常稳定，为用户实验提供了很大的支持，也取得了很多不多的成果。

Primary author: Mrs WANG, Yanfeng (IHEP)**Presenter:** Mrs WANG, Yanfeng (IHEP)**Session Classification:** 核电子学与探测技术**Track Classification:** 核电子学与探测技术

Contribution ID: 21

Type: 分会报告

散裂中子源上闪烁体中子探测器研制进展

Monday, 10 July 2023 14:15 (15 minutes)

作为研究物质结构和动力学性质的理想探针，中子散射技术已在凝聚态物理、化学、生命科学、材料科学等多学科领域的研究中被广泛采用。先进的中子源是中子科学研究的基础。作为发展中国家拥有的第一台散裂中子源，中国散裂中子源（CSNS）项目于 2008 年开始在广东东莞动工，2018 年 11 月顺利打靶并出束，未来其脉冲中子通量将位居世界前列，可以很好满足我国在多学科领域内对散射中子的强劲需求。

位置灵敏型中子探测器是中子散射实验中的关键设备，通过入射中子与探测器内核素的相互作用来实现中子探测。由于 ^3He 气体具有探测效率高、 n/γ 抑制比好等优点，截至 2007 年国际上 70% 以上的中子探测器都是以 ^3He 气体作为中子探测材料。但随着 ^3He 气体资源短缺、价格飞涨， ^3He 气体探测器已经无法满足中子探测的需求。基于 $^6\text{LiF}/\text{ZnS}(\text{Ag})$ 闪烁体和波移光纤结构的大面积位敏中子探测器，具有高的中子探测效率、高位置分辨率和可大面积拼接等优点，可以很好的满足中国散裂中子源（CSNS）上多个中子谱仪对大面积中子探测的需求。

CSNS 一期建成的三台谱仪中，通用衍射谱仪（General Purpose Power Diffractometer, GPPD）主探测器采用自主研发的大面积闪烁体探测器阵列（NSDA），该探测器单元由双层闪烁屏、波移光纤阵列、多阳极光电倍增管及其读出电子学构成，实现了 $4\text{mm} \times 4\text{mm}$ 的位置分辨和好于 45% 的热中子探测效率，各项参数指标满足谱仪的物理需求。GPPD 谱仪散射室内的探测器阵列目前由 48 台探测器单元构成，覆盖了底角、90 度和高角区域 6m^2 的空间。该探测器阵列在 2017 年底已经完成现场安装及调试。2018 年底第一次谱仪在束调试时，探测器阵列就配合谱仪，在线得到标准硅粉末样品的衍射数据，并得到的 0.184% 的谱仪分辨率。

基于闪烁体探测器在工程上的成功应用，散裂中子源项目目前在建的 7 台合作谱仪中，工程材料谱仪和成像谱仪也计划采用自主研制的闪烁体探测器，但两个谱仪在中子探测效率和位置分辨上，对探测器提出更高的要求。基于此，项目组研发了基于硅光电倍增管读出的高探测效率型闪烁体中子探测器，该探测器利用斜入射结构的闪烁屏实现 $1\text{\AA}$ 波长中子 40% 以上的探测效率，位置分辨为 3mm （50~200mm 可调）。该探测器计划 2021 年初完成在谱仪上的安装。

Primary author: Mrs 唐, 彬 (高能所)**Presenter:** Mrs 唐, 彬 (高能所)**Session Classification:** 核电子学与探测技术**Track Classification:** 核电子学与探测技术

Contribution ID: 34

Type: 分会报告

GECAM 卫星溴化镧晶体性能研究

Monday, 10 July 2023 15:00 (15 minutes)

引力波暴高能电磁对应体全天监测器 (Gravitational wave high-energy Electromagnetic Counterpart All-sky Monitor, 简称 GECAM) 卫星是专门探测引力波高能电磁对应体 (简称引力波伽马暴) 的小型空间科学探测项目。GECAM 项目由两颗微小卫星组成, 两颗卫星运行于相同轨道面内, 且轨道相位相反, 从而两颗卫星形成对全天的完整覆盖监测, 每颗卫星的有效载荷包括 25 个伽马射线探测器 (GRD)。GRD 设计使用 3 英寸直接的溴化镧晶体作为伽马射线探测灵敏材料, 后端耦合硅光电倍增管进行读出。

溴化镧 (LaBr₃) 闪烁晶体具有高光产额、高密度以及发光衰减时间短的优点, 同时溴化镧晶体极易潮解需要封装使用。使用放射源对 GECAM 卫星所有的溴化镧晶体进行了性能测试, 并根据晶体的掺杂类型进行了分类, 对 65 个交付的溴化镧晶体本底类型进行了初步分类, 包括以下 4 个类型, 1. 低本底型, 2. 中本底类型 3. 中本底类型 2, 4. 高本底。

对溴化镧晶体的辐照性能进行了测试研究。测试结果表明, 溴化镧晶体的性能符合设计要求, 能量分辨 $<6\%$ @ 662keV, 满足设计使用要求。在晶体的抗辐方面, 使用 Co60 放射源对晶体进行了辐照性能测试, 从测试结果看, 160krad 累积辐照后, 经过较长时间的退激发, 溴化镧晶体对 662keV (用 Cs137 源进行测试) 的试验前后峰值变化小于 20%, 符合空间环境条件下 (在轨三年) 的使用要求。不同晶体的性能衰退显示出不同的结果, 表明溴化镧晶体不同组成成分及掺杂比例, 对晶体抗辐照性能有不同的影响。从辐照试验前后的测试看, 溴化镧晶体辐照后, 晶体短期内 (几天至十几天内) 存在活化激发的问题, 辐照后晶体自发荧光, 主要会导致低能端 x 射线探测存在问题。在轨运行后, 由于晶体平均每天的辐照剂量约 150rad, 是目前辐照试验最小辐照剂量的 1/60, 因此一般情况下在轨不存在晶体活化激发的问题, 但仍然需要控制晶体的组分及掺杂等晶体生长工艺, 对项目使用的晶体进行抽样辐照检测, 符合空间应用要求。

Primary author: 安, 正华 (高能所)

Co-authors: ZHANG DALI; Mr 孙, 希磊 (IHEP); 李, 新乔 (高能所)

Presenter: 安, 正华 (高能所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 35

Type: 分会报告

一种适用于高通道数和高事例率下的高能物理实验软件触发算法

Tuesday, 11 July 2023 14:30 (15 minutes)

在 高能物理实验中，数据获取系统（DAQ，Data Acquisition）是获取物理数据的重要一环。实验中所有数据，包括探测器产生的物理信号数据和运行控制的监控数据，都可能需要汇集在 DAQ 系统中。随着高能物理实验规模的扩大，实验中使用的探测器和相应的电子学通道数目都在增加，DAQ 系统的数据处理压力也不断增大。同时，为了增加实验效率，增大事例率，越来越多的实验开始考虑 triggerless 的实验装置。无触发系统移除了物理硬件实现的触发系统，直接在数据获取系统中使用软件和触发算法的方式实现触发系统筛选满足物理条件的好事例，压缩本底事例的功能。这无疑进一步增加了数据获取系统的数据处理压力，降低了数据处理的实时性，增加了数据处理的时间和中途故障丢失的风险。故此，对 triggerless 的实验装置来说，好的软件触发算法将极大的提升数据获取系统的数据处理时间，能够更快的将数据存盘而降低有效物理数据丢失的风险。同时，系统资源使用少的算法，也能降低整个 DAQ 服务器节点的使用数目，降低实验的硬件成本。

本文介绍了目前无触发数据获取系统（以 LHAASO 为例）常用的 N-hit 软件触发算法，分析了算法的时间复杂度。研究了 N-hit 算法在处理大通道数和高事例率场景下性能下降的原因，并提出了一种在大通道数和高事例率场景下仍有优势的 Fill-table 算法。Fill-table 算法的核心思想是通过使用一个触发表，以空间换时间的方式节省下 N-hit 算法需要对数据排序的时间，从而更加高效的处理大通道和高事例率情况下的软件触发速度。不过面对通道数少，事例率低的条件，Fill-table 算法的优势并不明显。除此之外，Fill-table 算法解决了 N-hit 算法不易于并行计算的缺点，可以通过通道级别的并行计算减少算法处理的总时间，对于有高实时性要求的实验系统很有帮助。

Primary authors: Mr LI, Fei (IHEP, CAS, China); Mr 顾国皓, minhao (高能所); 彭, 宇 (中科院高能物理研究所); Dr 季, 筱璐 (高能所)

Co-author: 罗, 代斌 (中国科学院高能物理研究所)

Presenter: 彭, 宇 (中科院高能物理研究所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 38

Type: 分会报告

网络边界安全威胁 IP 阻断技术

Wednesday, 12 July 2023 15:45 (15 minutes)

伴随 Internet 网络的高速发展和不断扩大的网络应用，网络安全威胁屡见不鲜，只要是对外网开放服务的系统，每天都可能承受着成千上万次的尝试攻击，这些攻击可以造成网络性能下降，影响业务系统正常工作，甚至会成功地入侵，给系统造成破坏或损失。如何快速阻断和减少这类攻击，成为网络安全管理的一个重要问题。本文结合高能所网络安全管理的应用讲述了基于边界的安全威胁 IP 封锁方案和相关技术。

通过在网络数据通讯边界设备（防火墙或路由器）上配置访问策略，动态封锁或解封存在安全威胁的 IP 通讯；

本系统部署在高能所网络运行，通过接收来自 SOC 威胁监测系统，Bro 入侵检测平台以及人工监测或情报收集发现的黑 IP 资源，快速封堵存在安全威胁的 IP 地址，给来自网络的黑客攻击增加了阻力。封锁过程中收集积累了大量 IP 地址，可作为网络安全威胁的情报资源。

基于边界路由器的有害 IP 地址封锁是局域网络安全的重要部分，利用静态路由表的封锁技术应用简单可靠，与在防火墙上封锁 IP 地址相比，性能开销要小的多，以高能所使用的华为路由器为例，可以封锁百万数量级的 IP 地址。

Primary author: Mr 安, 德海 (高能所)

Presenter: Mr 安, 德海 (高能所)

Session Classification: 科研信息化管理与系统

Track Classification: 科研信息化管理与系统

Contribution ID: 39

Type: 分会报告

The environmental monitoring system of instrument based on EPICS

Wednesday, 12 July 2023 14:00 (15 minutes)

中国散裂中子源（CSNS）是利用中子研究物质微观结构和运动的重要科学设施，主要由质子加速器、中子靶站和中子散射谱仪等三大部分组成。中子散射谱仪是中子散射的实验装置，通过探测器电子学等精密设备获取实验数据从而分析得出实验结果，保证设备良好的运行状态对于实验结果至关重要。精密设备对运行环境要求较高，为保证设备的运行状态需要实时监测其运行环境。

本文设计并实现了基于 EPICS 的谱仪运行环境监测系统，能够在实验过程中实时监测散射室和机房的设备运行环境。将谱仪运行环境监测接入散裂中子源靶站谱仪的全局控制系统中，可及时避免环境变化对设备运行的影响，保证实验结果的准确性。

基于 EPICS 的谱仪运行环境监测系统结构如图 1 所示，多个环境监测传感器通过 RS485 总线连接至环境监测盒，环境监测盒读取传感器数据，然后将其以 PV 量的形式传递至 EPICS 网络中，环境信息经过处理后以曲线形式展示其实时变化，并将数据保存至数据库，用于分析长时间内整体环境的变化情况。环境数据可选择的显示方式包括：谱仪本地控制室界面显示和远程综合报警平台显示。

图 1 环境监测系统结构图

环境监测盒内部结构如图 2 所示，用于接收传感器信息，并通过将环境数据作为 PV 量发送至 EPICS 网络，同时显示在前面板的显示屏上。

图 2 环境监测盒外观

环境监测系统部署的重点在于传感器的布点及走线，需要根据被测环境的实际情况结合工程需求进行设计。该系统可选的环境监测内容包括：温度、湿度、氧气含量、烟雾、灰尘、噪音等。目前该系统已部署应用在多物理谱仪上，以此为例对环境监测系统的实现进行简要说明。

多物理谱仪的散射室和机房属于密闭空间，除了要保证各类精密设备运行在合适的温湿度环境下，还需要考虑到工作人员在散射室正门和顶部换样门操作时的安全问题，故需要布置温度、湿度、氧含量传感器对环境进行监测。图 3 是多物理谱仪的传感器布点及走线示意图。

图 3 多物理谱仪的传感器布点及走线示意图

传感器通过弱电桥架走线接入机房，再从静电地板下接入机柜中的环境监测盒。通过铸铝接线盒将 RVV4*0.5 规格电缆连接并用螺丝固定于墙面，墙面使用 PVC 线槽固定电缆至弱电桥架，电缆末端通过环形公制连接器接入环境监测盒。

环境监测系统部署完成后，环境信息进入靶站实验全局控制系统，实现本地显示报警、远程显示报警并将数据保存至靶站数据库。图 4 是多物理谱仪环境监测系统的远程显示。

图 4 多物理谱仪环境监测显示

目前基于 EPICS 的谱仪运行环境监测系统已经成功应用在中子源的多物理谱仪上，实现了该谱仪运行环境的远程实时监测及环境异常报警。下一步计划将该环境监测系统推广使用到散裂中子源其他谱仪及靶站大厅中，实现对靶站谱仪的全局环境监测，保证设备的良好运行环境。

Summary

本文设计并实现了基于 EPICS 的谱仪运行环境监测系统，能够在实验过程中实时监测散射室和机房的设备运行环境。将谱仪运行环境监测接入散裂中子源靶站谱仪的全局控制系统中，可及时避免环境变化对设备运行的影响，保证实验结果的准确性。

Primary author: Mrs 王, 晓庄 (中国科学院高能物理研究所东莞研究部)

Presenter: Mrs 王, 晓庄 (中国科学院高能物理研究所东莞研究部)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 43

Type: 分会报告

基于 REST API 的 slurm workbench 原型系统开发

Monday, 10 July 2023 16:15 (15 minutes)

Slurm 自 20.02 版本提供了 REST API, 可实现以 HTTP 协议方式与 slurm daemons 交互。slurm workbench 基于 Slurm REST API 开发, 可为管理员提供集群管理信息系统, 也可为普通用户提供可视化集群界面。

Primary author: Ms 杜然, Ran (高能所)

Presenter: Ms 杜然, Ran (高能所)

Session Classification: 科学计算技术与平台

Track Classification: 科学计算技术与平台

Contribution ID: 51

Type: 分会报告

高带宽高通量数据中心网络的流量分析方法研究

Tuesday, 11 July 2023 16:15 (15 minutes)

中国科学院高能物理所计算中心，承担了高能物理海量数据的处理工作。本地计算系统是一个批作业系统。该计算系统采用 HTCondor 实现计算资源的调度和管理、采用 EOS/Lustre 实现数据存储和访问。数据中心网络是采用机框式双核心加 10G/25GToR 构成星型拓扑的大二层网络。骨干链路聚合带宽为 8*100G。在高能所数据中心网络环境中，批作业在骨干链路上产生的流量峰值已经达到了 350Gbps。高带宽、高通量数据中心网络中的超大流量给流量获取和分析带来了挑战，传统的基于 SNMP、gRPC 等技术实现的流量监测，只能基于网络设备的端口计数器，进行总体流量的监控，无法对流量构成进行分析和研究。基于镜像、sflow 等方法实现数据采集，虽然可以实现对流量构成的分析，但是要实现全拓扑的流量监控，设备能力、数据分析能力面临很大的挑战。为实现全拓扑流量采集和分析，掌握数据中心网络的流量构成、分布和运行特征，本文提出一种实现方法，该方法通过在操作系统层面部署软件采集每个 socket 的流量；通过网络扫描发现主机的流量路径；并将二者相结合实现从主机、应用、网络拓扑以及时序等多维度的分析网络流量的构成、分布和运行动态，揭示高能物理计算系统在数据中心网络中产生的流量的运行规律，为数据中心网络的运行和优化提供决策依据。该方法解决了高带宽高通量数据中心网络的全拓扑流量获取的难题，并且能够在一定粒度上实现了对流量构成的分析和基于应用的流量运行动态展示。

Primary author: Mr 崔涛, Tao (高能所)**Co-author:** Ms 曾, 珊 (高能所)**Presenter:** Mr 崔涛, Tao (高能所)**Session Classification:** 科学数据存储与管理**Track Classification:** 科学计算技术与平台

Contribution ID: 56

Type: 分会报告

备份系统在高能所集群中的应用

Tuesday, 11 July 2023 16:45 (15 minutes)

高能物理实验每年产生大量的数据，其中有些特别重要的数据，比如用户处理数据的作业脚本，数据分析后的结果和记录数据信息的 mysql 数据库等都需要定期进行备份，以避免发生软硬件故障时导致的损失。高能物理实验数据量的快速增长也意味着重要数据的数据量也越来越大，从之前的 KB、MB 到现在的 GB、TB 数据级。为了有效完成这些重要数据的备份和恢复，经过调研和对比，将开源备份软件 Restic 加入到高能所集群的备份系统中，用于解决 TB 级别的大容量文件目录的备份。基于 Restic 开发了策略初始化、定期备份、并行备份、镜像文件校验、日志采集、用户自主恢复数据等功能，有效地减少备份总时间和备份文件占用空间，提高了数据备份和恢复的效率。

Summary

Primary author: Mrs YAO, qiuling (ihep)**Presenter:** Mrs YAO, qiuling (ihep)**Session Classification:** 科学数据存储与管理**Track Classification:** 科学数据存储与管理技术

Contribution ID: 59

Type: 分会报告

CEPC 上基于 DeepSets 模型的喷注标记算法研究

Monday, 10 July 2023 14:15 (15 minutes)

背景介绍 2012 年在大型强子对撞机 (LHC) 上发现希格斯 (Higgs) 玻色子是粒子物理学界的大事件, 该发现不仅补全了标准模型缺失的最后一角, 而且开启了粒子物理探索的新篇章。由于 WZH 三种玻色子 (尤其是希格斯玻色子) 与新物理现象及新物理规律的联系十分密切, 因此精确测量他们的性质, 是探索新物理现象和新物理规律的关键手段。强子喷注是 WZH 三种玻色子最主要的衰变末态, 因此喷注的重建 (Clustering) 和标记 (Tagging) 算法对实现 WZH 三种玻色子的精确测量至关重要。本文主要介绍喷注的标记算法。传统喷注标记算法有两种, 第一种是基于选择条件和人工变量, 第二种是基于传统机器学习算法, 比如决策树 (Boost Decision Tree, BDT) 和神经网络 (MLP)。近年来, 随着深度学习的发展及其性能的进一步提升, 越来越多的深度学习算法应用于粒子物理实验中, 例如依旧基于人工提取变量方法的 DNN 算法、基于图像识别的 CNN 算法和序列处理的 RNN 算法, 以及本文在正负电子对撞机 (CEPC) 上采取的基于 DeepSets 模型的神经网络方法对夸克喷注标记进行分析。

数据、模型和性能本研究采用了新提出的、基于 DeepSets 的能量流深度学习算法以及 Z 玻色子强衰变末态产生的喷注模拟数据集, 尝试对重味夸克喷注的标记进行研究。其中 Z 玻色子衰变为 $b\bar{b}$, $c\bar{c}$, $s\bar{s}$, $u\bar{u}$, $d\bar{d}$ 等 5 种不同强子末态。所有数据通过 Whizard 1.9.5 产生, 再用 Pythia 进行强子化模拟, 最后再采取 CEPC CDR 的基准探测器对蒙特卡洛样本进行探测器模拟和重建。DeepSets 模型有以下三个特点: 第一, 喷注中末态粒子的顺序交换不影响喷注的性质, 因此 DeepSets 模型必须考虑粒子顺序交换不变性这一对称性; 第二, 在顺序交换不变性的基础上采用多个的过滤器充分提取喷注的关键物理特征; 第三, 把提取出的特征作为 DNN 的输入进行标记。以上的数据集和模型在采用 GPU 进行充分训练后, 与传统的决策树和 XGboost 算法相比, 其性能提高约 6%, 使平均精度从 80% 提高到 85% 左右。

Summary

讨论和展望基于 DeepSets 模型的初步研究对夸克喷注标记给出了乐观的结果, 下一步作者将进行两方面工作: 一是对数据信息进行进一步发掘, 以及对模型的超参数进行深入优化, 期望得到性能的进一步提高; 二是尝试对完整事例进行标记, 预期能对 WZH 三种玻色子性质的精确测量有较大作用, 例如使得 WZH 三种玻色子的衰变分支比测量有显著的精度提高。

Primary author: 廖, 立波 (wuzhou univercity)

Co-author: Dr LI GANG (EPD.IHEP), Gang (高能所)

Presenter: 廖, 立波 (wuzhou univercity)

Session Classification: 人工智能与应用

Track Classification: 人工智能与应用

Contribution ID: 62

Type: 分会报告

Progresses of Data Science in SHINE Project

Monday, 10 July 2023 16:00 (15 minutes)

The rapid development of X-ray free electron laser (XFEL) facilities has opened a new paradigm of X-ray measurement. They can deliver ultra-short and coherent x-ray pulses with extremely high brilliance, thus enabling the observation of the physical and chemical behaviors in condensed matters and biomaterials, at high spatial and temporal resolution simultaneously. Shanghai High repetition rate XFEL and Extreme light facility (SHINE) has started its construction in April 2018 and may become one of most sophisticated XFEL facilities in the world.

The advanced detectors for XFEL allow an extremely high throughput of data rates and volumes. Big data has been identified as one of the major challenges in XFEL experiments, e.g. serial femtosecond crystallography and single-particle imaging. Dedicated data center and DAQ system shall be developed to meet the requirement of data acquisition, real-time analysis, and big data management of various instruments. A RD; data platform has already been built for the high throughput data processing based on high performance cluster as well as cloud computing. Algorithms and softwares of data acquisition, data management and analysis for tackling XFEL big data are also under development. In this presentation, most recent progresses of SHINE data science will be reported as well as envision of next decade.

Summary

Primary authors: Prof. FAN, Haiwei (ShanghaiTech University); Prof. HUAI, Ping (ShanghaiTech University)

Co-authors: Dr YIN, Congcong (Shanghai Advanced Research Institute CAS); Ms LEI, Lei (ShanghaiTech University); Dr SHI, Wujun (ShanghaiTech University); Dr ZHANG, Xiaofeng (ShanghaiTech University); Dr YIN, Yaru (ShanghaiTech University); Dr LI, Zhengheng (ShanghaiTech University)

Presenter: Prof. HUAI, Ping (ShanghaiTech University)

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 63

Type: not specified

硬 X 射线自由电子激光的电子档案库系统

Monday, 10 July 2023 16:45 (15 minutes)

1 引言

随着云计算、大数据、移动网络等信息技术的飞速发展和广泛应用，科研和管理信息化程度不断加深，充分借助信息技术，尤其是数字网络技术对档案内容信息以及相关衍生信息资源进行有效的组织、管理、存储和利用，已经成为新时期档案工作突破传统、创新发展的必经之路。2021 年起正式实施的《中华人民共和国档案法》，明确要求推动档案信息化建设，开辟档案管理现代化新路径。同时，肯定了电子档案的效力与凭证价值，对电子档案管理信息系统、数字档案馆、档案信息资源共享服务平台建设提出了更高要求。

硬 X 射线自由电子激光 (Shanghai High repetition rate XFEL and Extreme light facility, SHINE) 项目是“十三五”国家重大科技基础设施建设规划优先启动项目，是上海建设张江综合性国家科学中心的核心项目，总投资近百亿。该项目档案具有责任主体多、专业门类广、时间跨度大、载体复杂等特点，在项目建设过程已有电子文档、音像文件等大量的电子文件产生。单纯的纸质文件保管已无法覆盖项目建设全部内容，且利用不方便，还易导致有价值文件的流失，亟需建设电子档案管理系统。

SHINE 电子档案库遵守国家有关电子档案管理系统规定，依据先进、实用、安全、发展的原则，第一阶段开发任务设置了档案收集、档案整理、档案保存、档案利用、档案鉴定与处置、档案统计以及系统管理等基本功能。以期满足 SHINE 建设、调试与运行期间全部电子档案的存储需求，推动电子档案科学管理。

本文结合档案的安全性对 SHINE 电子档案库软硬件部署进行了整体设计，已建成的档案库运行情况良好。

2 系统设计

SHINE 电子档案库系统总体上基于集中式部署分布式利用的原则，主要由电子档案库后端管理软件、电子档案库前端管理软件、服务器、主存储、在线备份存储、光盘离线备份存储等组成。

电子档案库后端管理软件实现电子档案的统一采集，具有数据共享和在线编辑功能，共享后的文件交由各总体档案员按档案归档规范与要求，核查文件真实性、完整性、有效性。该后端管理软件基于 owncloud 技术搭建文件共享平台，实现多个单位的文件实时共享和在线编辑。

电子档案库前端软件主要由南京轩恩软件开发有限公司基于国家颁布的档案管理主要法规指导下进行开发，主要包括电子档案标准入库、档案整理、档案管理、档案利用和档案长期保存。

为系统数据安全考虑，前后端系统都具备每日备份机制，后端系统还具备每三个月离线备份机制。

3 电子档案库后端系统

3.1 后端系统架构

电子档案库后端系统基于 owncloud 技术开发，包括运行 owncloud 的主服务器和主存储、用于自动备份的备份服务器和备份存储、以及能够连接手机、电脑、高拍仪等设备的电子文件输入端。电子档案库后端系统具体部署如下图所示，整个系统服务器硬件位于上科大图信中心机房，客户端依赖于各用户具体位置，在上科大可直接接入，在上科大外需先接入上科大 VPN 后再接入。电子文件可通过手机、电脑直接传入电子档案库后端系统，针对已有纸质档案，档案管理人员可调用本地的高拍仪对纸质档案进行扫描，在扫描的同时可进行旋转、缩放、纠偏、合并等处理后上传到服务器。

3.2 后端软件目录结构及权限

为实现档案安全采集，设计了以下目录结构：第一级目录为 SHINE 电子档案，SHINE 项目档案员可查看 SHINE 采集的所有电子档案；第二级目录为总体级档案，分为 PM（综合管理）电子档案、AC（加速器总体）电子档案、BE（束线站总体）电子档案、EL（极端光物理总体）电子档案和 CE（建安总体）电子档案，各总体档案员只拥有各总体档案的管理权限；第三级目录为 SHINE 各个工程设计人员，各工程设计人员只可查看自己输出的电子档案。

4 电子档案库前端系统

4.1 前端系统架构

电子档案库前端系统由应用和数据库服务器及主存储、在线备份服务器及备份存储、离线刻录光盘以及运行在服务器硬件上的电子档案库前端软件组成。为系统性能考虑，将应用服务器和数据库服务器二合一，集中部署在 SHINE 工程办服务器上，主存储可用空间为 44TB，最高可扩容至 216TB；在线备份服务器及备份存储部署在上科大图信中心机房，与其他系统共用 144TB 备份存储空间。

4.2 权限管理

为了保证档案安全，不易被篡改、删减，将电子档案管理员权限分为 4 个层级，分别为 Admin、软件配置管理员、用户配置管理员和档案移除管理员。Admin 拥有所有的权限，Admin 账户用户名和密码封存保管，紧急场景再使用。档案员按管理文件范围分别项目档案员和各分总体档案员，项目工程人员只具备查看电子文件的权限。

a) Admin

拥有最高权限，所有档案收集/整理/管理/利用、档案目录调整及配置、用户增加/删除/权限配置

b) 软件配置管理员

档案目录调整及配置

c) 用户配置管理员

用户增加/删除/权限配置

d) 档案移除管理员

档案管理阶段的档案移除和删除权限

e) 项目档案员

所有档案的编辑/利用，档案收集/整理阶段的档案移除和删除权限

f) 各总体档案员

相应总体档案的编辑/利用，档案收集/整理阶段的档案移除和删除权限

g) 项目工程人员

查看已申请权限的电子档案

Summary

硬 X 射线自由电子激光的电子档案库系统第一阶段已完成系统设计、模块功能开发、软硬件搭建与测试、域名固化和部分纸质档案扫描件、音像文件导入编目，具备基本的统一档案入口、档案按规入库、档案整理、在线档案检索与利用、智能统计数据展示和数据的多重备份，系统运行良好，当前已存储约 5G 的电子档案文件。

SHINE 电子档案库下一阶段计划在用户的借阅利用、与其他信息化管理系统的挂接、档案的安全监控与防范升级、统一认证、库房智能化管理等功能进行设计开发，远期可进一步实现电子档案库前端系统的移动端查阅、档案编研、风险预警等功能。

Primary authors: Prof. HUAI, Ping (Shanghai Advanced Research Institute, Chinese Academy of Sciences); Dr 尹, 聪聪 (Shanghai Advanced Research Institute, Chinese Academy of Sciences)

Co-authors: Ms ZHU, Jie (Shanghai Advanced Research Institute, Chinese Academy of Sciences); Ms LEI, Lei (ShanghaiTech University); Dr SHI, Wujun (ShanghaiTech University); Ms XUE, Youfei (ShanghaiTech University)

Presenter: Dr 尹, 聪聪 (Shanghai Advanced Research Institute, Chinese Academy of Sciences)

Session Classification: 先进光源数据与软件

Contribution ID: 72

Type: 分会报告

面向人工智能模型的异构数据融合和容器化 workflow 系统

Tuesday, 11 July 2023 16:15 (15 minutes)

针对人工智能模型应用场景下，数据多源化和异构化导致的数据融合效率较低，各种模型间难以流程化地组合运行等问题，提出了面向人工智能模型的异构数据融合和容器化 workflow 系统。该系统主要包括异构数据融与持续化、容器化模型部署与运行、工作流定制等 3 个部分。通过综合运用异构数据融、容器化模型部署、工作流定制等相关技术，面向人工智能模型的异构数据融合和容器化 workflow 系统，具有人工智能应用场景的普适性，包括工业图像识别、医疗影像和生态数据分析等，能够极大的提升了不同人工智能模型间组合运行的效率。本文以陆地生态系统碳循环预测过程为例，通过容器化技术构建深度学习和数据同化等模型，在无需人工干预的情况下，使用工作流引擎和定时任务调度器可以实现每天自动化地预测未来 7 天的气象数据和生态数据，验证了本系统的可用性和易用性。

Primary author: 万, 萌 (cnic)**Presenter:** 万, 萌 (cnic)**Session Classification:** 数据处理软件与分析方法**Track Classification:** 科学计算技术与平台

Contribution ID: 76

Type: 大会报告

天文数据中心三十四年回顾与展望

Tuesday, 11 July 2023 09:30 (30 minutes)

Summary

Presenter: 崔, 辰州 (University of Chinese Academy of Sciences)

Session Classification: 大会报告

Contribution ID: 77

Type: 大会报告

高能物理科学数据共享与服务

Tuesday, 11 July 2023 10:50 (30 minutes)

Summary

Presenter: 陈, 刚 (IHEP)

Session Classification: 大会报告

Contribution ID: 83

Type: 大会报告

粒子物理实验硅像素探测器研究现状与发展趋势

Wednesday, 12 July 2023 08:30 (30 minutes)

Summary

Presenter: Dr 张, 亮 (Shandong University, CHINA)

Session Classification: 大会报告

Contribution ID: 90

Type: 大会报告

数字经济高质量发展中数据安全思考

Tuesday, 11 July 2023 08:30 (30 minutes)

Summary

Presenter: 廖, 方宇

Session Classification: 大会报告

Contribution ID: 91

Type: 大会报告

超快粒子辐射探测器关键技术研究

Wednesday, 12 July 2023 09:00 (30 minutes)

Summary

Presenter: 钱, 森 (高能所)

Session Classification: 大会报告

Contribution ID: 92

Type: 大会报告

高精度白兔同步网络及其在 LHAASO 的应用

Wednesday, 12 July 2023 10:45 (30 minutes)

Summary

Presenter: Prof. 龚, 光华

Session Classification: 大会报告

Contribution ID: 93

Type: 大会报告

合肥先进光源控制系统设计

Wednesday, 12 July 2023 09:30 (30 minutes)

Summary

Presenter: 刘, 功发 (高能所)

Session Classification: 大会报告

Contribution ID: 95

Type: 大会报告

X 射线探测和成像技术及其在文物研究中的应用

Wednesday, 12 July 2023 11:15 (30 minutes)

Summary

Presenter: 魏, 存峰 (IHEP)

Session Classification: 大会报告

Contribution ID: 96

Type: 分会报告

空间项目科学应用平台分析

Tuesday, 11 July 2023 16:00 (15 minutes)

随着我国空间项目的不断增多，科学应用部分的工作任务不断扩展，接口界面也不断增多，需要梳理各项功能和 workflows，在此基础上为建立比较综合的空间应用平台开展前期规划。本报告对空间应用平台进行了需求分析，提出集中建设方案提供讨论。

Summary

随着我国空间项目的不断增多，科学应用部分的工作任务不断扩展，接口界面也不断增多，需要梳理各项功能和 workflows，在此基础上为建立比较综合的空间应用平台开展前期规划。本报告将对高能所参与的不同类型的空间项目进行整理，对空间应用平台的规划进行需求分析，提出集中建设方案提供讨论。

Primary author: SONG, Liming (IHEP,CAS)**Presenter:** SONG, Liming (IHEP,CAS)**Session Classification:** 数据处理软件与分析方法**Track Classification:** 科学计算技术与平台

Contribution ID: 97

Type: 分会报告

基于 IHEP 的多园区异地统一身份认证系统设计

Wednesday, 12 July 2023 16:00 (15 minutes)

中国科学院高能物理研究所是中国从事高能物理研究、先进加速器物理与技术研发与利用、先进射线技术与应用的综合性研究基地。

High Energy Institute(IHEP) 的 SSO (Sing sign on) 系统用户超过 2.2 万, AFS (calculation cluster) 用户 3.2 万, Web 应用超过 150 个, 客户端应用超过 10 个。随着高能研究所的发展, 国际合作越来越频繁, 高能研究所 SSO 系统应运而生。

IHEP 的 SSO 系统联通了高能所所有应用和 AFS 用户帐户。并实现了中国认证联盟 CARSI 与国际认证联合会 EduGain 的接入, 不仅实现了异地多园区统一账户管理, 还逐步实现了国内高校与国际组织的认证。

Summary

Primary author: LUOQI, UNKNOWN (高能所)

Co-authors: 胡, 皓 (高能所); 王丽, Li (高能所); 齐, 法制 (高能所)

Presenter: LUOQI, UNKNOWN (高能所)

Session Classification: 科研信息化管理与系统

Track Classification: 科研信息化管理与系统

Contribution ID: 98

Type: 分会报告

HEPS 数据管理关键技术实现与进展

Monday, 10 July 2023 16:30 (15 minutes)

高能同步辐射光源（High Energy Photon Source, HEPS）是国内建设的首个第四代同步辐射光源，为我国材料科学、化学工程、能源环境、生物医学等众多领域提供了先进的实验平台。一期建设的十五条线站在运行期间预计每天总共会产生 200TB 至 500TB 的原始数据。HEPS 所产生的数据具有数据量大、通量高、结构复杂等特点。因此，为了高效地利用数据，需要对这些海量数据从数据获取、存储、传输、利用到共享的全生命周期过程进行跟踪和管理。本文将着重介绍当前 HEPS 数据管理的关键技术实现以及相关进展，并讨论后续的发展方向。

Summary

Primary authors: 王, 浩帆 (高能所); 胡, 皓 (高能所); LUOQI, UNKNOWN (高能所); ZHUANG BO, UNKNOWN (高能所); 齐, 法制 (高能所)

Presenter: 王, 浩帆 (高能所)

Session Classification: 先进光源数据与软件

Track Classification: 科学数据存储与管理技术

Contribution ID: 99

Type: 分会报告

基于机器学习的同步辐射 X 射线小角散射数据预处理和挖掘

Monday, 10 July 2023 14:30 (15 minutes)

同步辐射 X 射线散射技术是研究高分子薄膜材料多尺度结构演变的有力工具。其中，时间分辨原位小角 X 射线散射实验将产生大量 SAXS 数据。通过分析 SAXS 散射花样，研究人员可以获得大量硬弹性 iPP 薄膜在拉伸过程中的结构演变信息。然而，传统手动数据处理方式的带宽有限，无法适应现代高帧率检测器不断增加的数据密度和数量，这可能会导致大数据中包含的关键信息丢失。因此，迫切需要一种能够对大规模 SAXS 数据集进行预处理并快速探索参数空间的新方法。在这项工作中，我们应用几种机器学习方法来处理硬弹性 iPP 膜的原位 SAXS 数据集。结果表明，VAE 和 cVAE 模型能够提取 SAXS 花样中隐含的结构信息，并将其映射到低维空间中。SAXS 散射花样在二维和一维潜空间中的表示揭示了 iPP 膜的关键特征，如微观结构相似性、结构变化率和结构演化路径，从而可以为研究人员进一步指出数据分析的方向。为了建立加工-结构关系图谱，我们发展了一种混合 VAE 和多层感知器（MLP）神经网络，其中 MLP 用于将加工参数映射到潜空间坐标，VAE 解码器用于生成给定加工参数下的 iPP 薄膜 SAXS 散射花样。为了测试模型的鲁棒性，我们进行了验证实验，并在连续温度-应变空间中生成了硬弹性 iPP 膜的 SAXS 散射花样，以帮助研究人员更全面地了解结构演变，并指导高目标性实验。这项工作中开发的机器学习方法可以用于大型 SAXS 数据集的快速预处理，及加工-结构关系的映射。此外，该工作所发展的方法框架是通用的，因此未来可扩展到其他需要高通量数据分析的材料体系。

Summary

With the rapid development of the synchrotron radiation X-ray characterization techniques, the preprocessing of large small-angle X-ray scattering (SAXS) datasets and the data mining become urgent requirements for researchers. In this work, we apply the variational autoencoder (VAE) and the conditional variational autoencoder (cVAE) to visualize a large SAXS dataset of hard-elastic isotactic polypropylene (iPP) films in 2- and 1-dimensional latent spaces. The low-dimensional representations enable us to capture key features of the dataset rapidly, such as the similarity among SAXS patterns and the structural evolution trends. The preprocessing of the dataset points out the further direction of data analysis so that researchers can focus on the most valued regions in the dataset. Then, we develop a hybrid VAE-multilayer perceptron (MLP) neural network to realize the processing-structure mapping of iPP films. The robustness of the hybrid VAE-MLP network is verified. Finally, SAXS patterns in the temperature-strain space are generated, which allows us to explore the processing parameter space not involved by previous experiments. These capabilities indicate that the developed machine-learning methods are valuable artificial intelligence toolset to assist in the preprocessing of large-scale SAXS datasets and the establishment of comprehensive processing-structure relationship of hard-elastic iPP films.

Primary authors: 赵, 晨皓 (中国科学技术大学); 昱, 万程 (中国科学技术大学); 李, 良彬 (中国科学技术大学)

Presenter: 赵, 晨皓 (中国科学技术大学)

Session Classification: 人工智能与应用

Track Classification: 人工智能与应用

Contribution ID: 101

Type: 大会报告

超导量子计算及融合计算

Monday, 10 July 2023 08:45 (30 minutes)

近几年，量子计算已经从实验室内的科研工作部分转向了产业应用。如今已经可以通过互联网渠道直接使用到最先进型号的量子计算机。

本文从量子计算的发展讲起，并对现阶段量子计算机的组成，尤其是超导量子计算的实现进行介绍，并展示互联网上的量子计算机的使用方式以及经典与量子计算的融合计算形式。最终，展望未来的量子计算形式，以及量子计算实现过程的设备需求、与现今各学科的交叉等。

Summary

Primary author: 梁, 福田 (中国科学技术大学)

Presenter: 梁, 福田 (中国科学技术大学)

Session Classification: 大会报告

Contribution ID: 102

Type: 分会报告

基于集群的分布式计算平台

Monday, 10 July 2023 17:00 (15 minutes)

高能所计算集群长期为高能物理实验提供计算服务，拥有着大量的实验用户。随着实验规模的不断扩大以及用户量的增加，集群现有队列排队情况愈发严重。

为了解决本地集群资源紧张，排队久的情况，东莞大科学数据中心提供了 18000 个 CPU 核用于拓展高能所集群规模。考虑到高能所实验用户长期保持着使用集群提交计算作业的习惯，采用网格计算的方式很难在各个实验得以推广，因此本文设计并实现了基于集群的分布式计算。

该方式首先将异地站点资源加入到高能所集群资源池中，进行统一的作业调度。为了保证用户长期以来的使用集群提交作业的习惯，该方式会对提交到异地站点的作业脚本进行分析和处理，从而将用户作业中使用到的软件和数据等文件传输到异地站点的计算节点上，另外对脚本中使用到的路径进行相应的修改，从而保证作业的正确性。对用户而言，保证了之前提交作业的方式，用户无需进行任何修改即可将作业提交到远程站点执行。对平台而言，基于集群的分布式计算的模式便于拓展本地集群的规模，无论是异地站点、商业云等资源都可以通过这样的形式来拓展本地集群的规模。

目前，异地站点的资源已经纳入到高能所本地集群提供给 BES、LHAASO、HERD 等实验使用。

Summary

Primary authors: 郭超奇, UNKNOWN; JIANG XIAOWEI, Xiaowei (高能所); 石, 京燕 (高能所); 杜然, Ran (高能所); 齐, 法制 (高能所)

Presenter: 郭超奇, UNKNOWN

Session Classification: 科学计算技术与平台

Track Classification: 科学计算技术与平台

Contribution ID: 103

Type: 分会报告

JUNO 探测器监控系统在大规模设备数量下的研究

Wednesday, 12 July 2023 14:30 (15 minutes)

随着江门中微子实验的不断建设，实验规模不断扩大，对探测器监控系统 (DCS) 提出了更高的控制需求。采用了分布式、多线程的处理方案。为了研究 DCS 在分布式·多线程方案下，与电子学联合测试的性能表现，进行了测试、分析。本文分析了基于 EPICS 框架的 IOC 程序 (input output control) 与 GCU (Global Control Unit) 数量之间的关系，以及其初始化的时间复杂度。研究了 PV(Process variable) 队列的过大或过小导致 IOC 性能下降的原因以及使用 ControlHub 和 IPBus UDP 两种不同方案对 IOC 程序运行效率的影响。以此，提出了多台服务器、多个 IOC 多个端口同时运行的方案，优化了服务器资源分配方案，以及每个 IOC 连接设备数量，提升了 DCS 工作效率。为江门中微子实验的最终运行，提出了可靠有效的大规模运行方案和量化参考。

Summary

Primary author: Mr 黎, 晃 (中国科学院高能物理研究所)

Co-author: 叶, 梅 (高能所)

Presenter: Mr 黎, 晃 (中国科学院高能物理研究所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 104

Type: 分会报告

散裂中子源科学计算平台建设及应用

Monday, 10 July 2023 16:00 (15 minutes)

介绍散裂中子源科学计算平台建设、运行及后续规划。

Summary

Primary authors: 侯, 丰尧 (中国科学院高能物理研究所); Ms 王, 丽 (中国科学院高能物理研究所); Mr 洪, 剑书 (中国科学院高能物理研究所); Mr 袁, 千金 (曙光信息产业有限公司)

Presenter: 侯, 丰尧 (中国科学院高能物理研究所)

Session Classification: 科学计算技术与平台

Track Classification: 科学计算技术与平台

Contribution ID: 105

Type: 分会报告

基于 Python 的大型物理实验控制软件框架研究与部署测试

Tuesday, 11 July 2023 16:30 (15 minutes)

基于 Python 的大型物理实验控制软件框架研究与部署测试

夏守腾 1,2,3, 张银鸿 1,2*, 钱森 1,2, 杨洁 3

(1. 中国科学院高能物理研究所, 北京 100049; 2. 核探测与核电子学国家重点实验室, 北京 100049; 3. 郑州大学, 郑州 450001)

不同大型物理实验设备的软件实现需求存在明显差异。传统的控制软件是针对特定实验设备而开发的, 其代码模块结构紧密, 因此难以复用和扩展。为解决这个问题, 本文研究并设计了一个通用框架, 可用于不同类型实验设备的控制软件, 以便快速构建控制软件以满足不同需求。

该框架是一个基于 Python 的、开源的慢控制系统软件框架, 可用于所有粒子物理实验的设备控制。该框架包括多个功能模块, 如显示、数据查询、警报、日志和远程监视等。其最主要的数据采集、数据监控以及报警功能之间的数据通信是采用 MQTT (Message Queuing Telemetry Transport) 协议, 不仅降低了代码耦合度, 也确保了实时和可靠的数据传输。基于 MVC (Module、View、Control) 架构, 利用 PyQt5 开发了用户界面, 实现了用户界面与后端逻辑处理之间的分离, 提高了各模块的灵活性和可移植性。本研究基于面向对象的编程方法保证了每个功能模块的独立性和可移植性, 并且统一的配置文件格式使它们可以在各种实验中轻松地通过简单地组合模块来适应需求的变化和创建新的应用程序, 适用于物理实验设备控制软件的更新换代。本研究中开发的控制软件框架已经成功部署在了 SY5527 与 VME 设备上, 并都经过了 Docker 测试, 同时结合 K3S 对 Docker 进行管理, 解决了框架在大型物理实验中大数据量下的负载均衡, 便于用户对软件的管理与维护。

关键词: 慢控制系统; MQTT; PYQT5; Docker

Summary

Primary authors: 夏, 守腾; 钱, 森 (中国科学院高能物理研究所); 张, 银鸿 (中国科学院高能物理研究所)

Presenter: 夏, 守腾

Session Classification: 数据处理软件与分析方法

Contribution ID: 106

Type: 分会报告

LHAASO 模拟作业从 X86 到 ARM 计算集群的移植

Wednesday, 12 July 2023 15:45 (15 minutes)

随着许多大型高能物理实验的推进，需要处理和分析的数据量显著增加，对计算资源的需求飞速增长。例如高海拔宇宙线观测站（LHAASO），它是一个复合型阵列，主要由平方公里阵列（KM2A）、水切伦科夫探测器阵列（WCDA）和广角切伦科夫望远镜阵列（WFCTA）这三个子阵列组成。自 2020 年实验开始运行以来，它们的模拟作业一直在 Intel X86 集群上运行，由于 CPU 资源有限，仅生产了第一阶段计划数据的一部分。因此有必要探索新的并拓展现有的计算服务设备。2021 年，在中国广东省东莞市建成了包含 10k CPU 核心的 ARM 计算集群，我们在此之上构建了一个基于 ARM 架构的应用程序生态系统，以支持高能物理离线数据处理。我们的主要工作包括：将基于 KM2A、WCDA 和 WFCTA 实验的离线软件移植到 ARM 机器上运行，在 ARM 集群中制定数据传输和作业调度策略，并评估在 Intel X86 和 ARM 集群中的性能和功耗。结果表明，LHAASO 模拟作业在 ARM 计算集群可以正确运行；基于 Intel X86 架构的 CPU 单核性能优于 ARM CPU，但是对于多核架构的整个服务器来说，ARM 服务器性能更好。

Summary

Primary author: CHENG, Yaosong**Co-authors:** 毕, 玉江; 石, 京燕 (高能所)**Presenter:** 毕, 玉江**Session Classification:** 科学计算技术与平台**Track Classification:** 科学计算技术与平台

Contribution ID: 107

Type: 分会报告

TAO DAQ 系统的设计和实现

Wednesday, 12 July 2023 14:45 (15 minutes)

台山中微子实验 (Taishan Antineutrino Observatory, TAO) 是江门中微子实验 (Jiangmen Underground Neutrino Observatory, JUNO) 的卫星实验, 其主要物理目标是为 JUNO 实验提供高精度的反应堆中微子参考能谱, 并测量反应堆中微子能谱的精细结构。基于上述目标, 实验采用约 2.8 吨掺钎低温液体闪烁体, 全覆盖约 10m² 的硅光电倍增管以实现高能量分辨率。实验中包括中心探测器, 水切伦科夫探测器以及顶部塑料闪烁体探测器三种类型探测器。对于 TAO 数据获取系统, 需要处理的总输入数据带宽在 Gbps 量级, 输出带宽必须小于 100Mbps, 数据需要进行软件触发、压缩等实时处理, 因此需要设计高效的在线实时数据处理。本文基于 JUNO DAQ 的框架进行开发, 根据其实际需求对 TAO 数据获取系统进行定制化设计和实现, 系统主要分为数据流和在线软件两部分。数据流软件负责获取来自三个探测器电子学和触发系统的原始数据, 并进行数据组装、处理和存储等操作。在线软件负责提供配置服务, 运行控制, 进程管理和消息日志等辅助功能。本文将详细介绍 TAO 数据获取系统的设计和实现。

Summary

Primary author: 张, 水涵 (高能所)**Co-author:** 季, 筱璐 (高能所)**Presenter:** 张, 水涵 (高能所)**Session Classification:** 核电子学与探测技术**Track Classification:** 核电子学与探测技术

Contribution ID: 108

Type: 分会报告

面向高能物理的磁带库迁移系统

Tuesday, 11 July 2023 17:00 (15 minutes)

磁带存储是目前高能物理数据长期保存的存储技术之一。随着磁带制造技术以及高能物理存储架构的不断发展，在高能物理中长期使用的 Castor 技术已经被新的磁带库存储系统 CTA 取代。为了满足不同实验对数据备份的需要，针对 CTA 系统，设计开发了通用的高能物理磁带库传输系统 CTS。本系统支持 LHAASO、BESIII、JUNO 等实验的不同应用场景，结合 EOS 系统的 WFE，自动化数据的归档和迁出，提高了数据迁移效率，有效保证了数据存储安全。

Summary

Primary authors: 毕, 玉江 (IHEP); 姚秋玲, qiuling (高能所); 程, 国松 (高能所)

Presenter: 毕, 玉江 (IHEP)

Session Classification: 科学数据存储与管理

Track Classification: 科学数据存储与管理技术

Contribution ID: 109

Type: 分会报告

大型科学实验运行故障诊断系统设计及技术研究

Wednesday, 12 July 2023 16:30 (15 minutes)

像高海拔宇宙线观测站（LHAASO）、江门中微子实验（JUNO）的大型科学实验为了观测宇宙线或中微子等物理对象，以及伽马射电暴或超新星爆发等物理现象需要系统不间断运行。为了保障实验的长期稳定运行，需要在系统出现故障的时候，实验运行人员快速分析故障原因，修复故障并恢复实验运行。这些大型科学实验一般由数量庞大的探测器及电子学系统、复杂的交换机网络、多达百台计算集群组成的数据获取系统、离线存储系统组成。为了快速检测故障并对故障原因进行分析，设计了大型科学实验运行故障诊断系统。该系统通过收集来自探测器、电子学、数据获取系统的运行状态信息，实时检测链路状态，在发生故障时检测并快速分析故障原因，并及时提供给实验运行人员。此外，还可以按需分析指定时间段的历史数据，生成数据报表。

该系统的设计要求包含高吞吐、实时性、拓展性和可靠性。为了满足这些要求，该系统采用了分布式架构，将数据收集、数据处理、数据存储和数据展示等功能分别部署在不同的节点上。该系统使用消息数据流平台 Redpanda 作为消息中间件，使用高性能分布式数据库 Cassandra 作为存储仓库，通过流处理将来自实验系统的大量数据压缩并实时保存至仓库，通过批处理分析数据并利用自动化手段自动运行给出故障原因。该系统构建了一个科学物理实验故障的因果树，来提高故障诊断的准确性和效率。

目前已完成该系统的技术验证和原型实现，并在 LHAASO 中得到了应用。该系统可以快速、准确地诊断运行故障，为实验运行人员提供有效的解决方案，提高实验的稳定性和效率，为科学研究提供有力的支持。

通讯作者：张航畅，高能物理研究所，15938638566，zhanghc@ihe

Summary

Primary authors: ZHANG 张, Hangchang 航畅 (IHEP); 顾珏皓, minhao (高能所)

Presenter: ZHANG 张, Hangchang 航畅 (IHEP)

Session Classification: 科学计算技术与平台

Track Classification: 核电子学与探测技术

Contribution ID: 110

Type: 分会报告

高亮度正电子束团湮没数据采集与处理软件设计

Wednesday, 12 July 2023 15:00 (15 minutes)

随着正电子存储及控制技术的发展和产生高亮度、超快正电子束团成为可能，为实现正电子湮没寿命瞬态测量方法提供了实验条件。正电子束团在纳秒甚至百皮秒内湮没后产生巨量湮没伽马光子的探测，要求数据采集系统在短时间内对其所携带的湮没数据信息进行快速收集和提取。

本文基于 SCA 技术的 DRS4 (Dumino Ring Sampler version 4) 阵列式波形采样电子学，设计了正电子束团湮没数据采集与处理软件。其中，通过 UDP 数据传输、LabVIEW 自动多线程、队列缓存和 TDMS 文件存取等技术实现了高速波形采样数据的传输和存储；应用误差补偿、正弦波过零和全局校正等方法，设计出适用于多通道波形采样数据的校正功能程序，从而实现了波形采样数据的精确重建。

经实验测试：该数据采集软件性能稳定、操作方便，实时处理并存储数据的最高速度稳定在 43.7 MB/s，可以满足正电子束团湮没测量中高通量数据处理的需求。同时，在信号处理方面，该软件可实现较好时间分辨：电子学单通道的固有时间分辨优于 22.6 ps；耦合“LYSO+SiPM”探测器后，单通道时间分辨率可达到百 ps 以内，满足正电子束团湮没测量中高时间分辨的需求。

Summary

Primary author: Dr FUYAN, Liu

Co-authors: Mr 卞忠伟; Dr 况鹏; Dr 王英杰; Dr 王冕; Prof. 章志明; Prof. 曹兴忠; Prof. 王宝义

Presenter: Dr FUYAN, Liu

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 111

Type: 分会报告

高能所数据存储系统现状及规划

Tuesday, 11 July 2023 16:00 (15 minutes)

高能所数据存储系统是一个庞大的分布式存储系统，旨在满足高能所各个领域的数据存储需求。这些领域包括粒子物理和天体物理、空间项目、多学科交叉（CSNS/HEPS）、理论物理和加速器设计等领域。为了支持这些需求，我们运行了多个分布式存储系统，其中包括主要的物理数据存储系统 Lustre 和 EOS，以及通用文件系统、专用 HPC 集群、对象存储系统、内容分发文件系统和用户的云盘存储。目前，这些系统的总容量已超过 100PB。

在过去的十多年里，Lustre 和 EOS 作为高能所主要的物理数据存储系统，提供了 100GB/s 的数据访问带宽。为了适应变化的需求，我们引入了最新的磁带库管理软件 CTA，已全面替代现有的 Castor 系统。我们还采用 Restic 作为数据备份系统，以实现关键数据的高效备份。为了满足用户对 HOME 目录的需求，我们正在探索新的商业文件系统解决方案。为了满足公共云服务平台的存储需求，我们提供了 CEPH 存储集群。同时，我们通过 CVMFS 为高能所的所有实验提供软件存储库。

目前，高能所存储系统采用了“一平台多中心”架构，以北京和东莞两大数据中心为主，辅以稻城、江门等多个分数据中心，形成了分布式存储架构。本文将重点介绍去年高能所数据存储的经验和服务情况，并讨论即将发生的变化、未来的改进和面临的挑战。

Summary

Primary author: LI, Haibo (高能所)

Presenter: LI, Haibo (高能所)

Session Classification: 科学数据存储与管理

Track Classification: 科学数据存储与管理技术

Contribution ID: 112

Type: 分会报告

基于 Rucio 的高能物理网格数据管理的研究和应用

Tuesday, 11 July 2023 16:30 (15 minutes)

Rucio 是一个可由用户自定义策略以实现组织、管理和访问大规模科学数据功能的软件框架。整合来自不同存储技术和网络连接技术的数据成一个整体，Rucio 可以管理分布式的、多存储类型的数据站点。相比于传统网格数据管理系统，Rucio 提供了一些高级特性如：分布式数据恢复、自适应的数据复制以及高度量化、模块化和可扩展性的数据管理。Rucio 最先在高能物理实验 ATLAS 上开发和应用，后来也逐渐应用在其他高能物理实验的科学数据管理中。

在使用网格框架分布式计算系统的国内高能物理实验中，JUNO、HERD、CEPC 等实验都已经使用或有计划使用 Rucio 系统对实验数据进行管理。我们基于不同类型的实验对数据管理的不同需求，设计并开发不同的基于 Rucio 系统数据管理解决方案，实现了分布式数据统一命名、数据增删改查基础管理功能、多站点数据副本管理、原始数据分发管理、实验软件数据管理接口嵌入等功能或相关服务。

Summary

本报告将会从介绍 Rucio 的特性以及在数据管理中的优势开始，结合 JUNO 和 HERD 等实验的需求，介绍基于 Rucio 系统数据管理解决方案，最后介绍 Rucio 系统在数据管理和数据传输等方面测试和性能情况。

Primary author: ZHANG, Xuantong (IHEP)**Co-author:** ZHANG, Xiaomei**Presenter:** ZHANG, Xuantong (IHEP)**Session Classification:** 科学数据存储与管理**Track Classification:** 科学数据存储与管理技术

Contribution ID: 113

Type: 分会报告

大装置和科学数据中心网络安全保障平台总体设计与实现

Wednesday, 12 July 2023 16:15 (15 minutes)

国家重大科技基础设施是为探索未知世界、发现自然规律、实现技术变革提供极限研究手段的大型复杂科学研究装置（系统），是突破科学前沿、解决经济社会发展和国家安全重大科技问题的物质技术基础。由于大多数重大科技基础设施必须对国内外用户高度开放，使得依托重大科技基础设施开展的科研活动、应用研发等工作面临着包括工业控制安全、互联网安全以及数据安全等在内的多方面安全威胁。为了应对日益严重的网络安全威胁，保障重大科技基础设施及其科研活动的网络安全，建设面向重大科技基础设施的网络安全技术平台已经迫在眉睫。

本报告介绍大装置和科学数据中心网络安全保障平台总体设计与实现，该平台架构包括：安全数据采集层、安全数据预处理层、安全数据存储层、安全数据分析层、安全数据应用层和威胁情报层。安全数据采集层实现复杂多源异构安全类数据的采集，包含网络流量数据、日志数据、脆弱性数据以及资产数据等；安全数据预处理层实现数据的实时处理，包括安全数据解析、富化等；安全数据存储层利用大数据存储、关系数据库等平台实现多源异构安全类数据的存储；安全数据分析基于多源异构安全类数据，利用规则分析、统计分析、机器学习、深度学习、图学习等智能检测算法实现威胁建模、监测和发现；安全数据应用层利用可视化技术、自动化工作流程等方式，实现态势感知、事件处置响应、资产和漏洞管理等能力；威胁情报层通过自动采集、人工事件调查等方法 and 途径，实现威胁情报收集、应用和共享。以上的不同逻辑层之间通过联动协作，形成了安全威胁建模、安全信息感知、安全事件分析、安全事件预警及安全事件应急处置的一体化的网络安全技术平台，有效保障依托大装置和科学数据中心开展的科研活动、应用研发等业务可持续发展。

Summary

Primary author: Dr 王, 佳荣

Co-authors: Dr 颜, 田; Mr 安, 德海; Mr 郭, 超奇; Ms 刘, 国怡; Mr 许, 应衡; Mr 王, 嘉瑞; Mr 刘, 禹; Dr 齐, 法制

Presenter: Dr 王, 佳荣

Session Classification: 科研信息化管理与系统

Track Classification: 科研信息化管理与系统

Contribution ID: 114

Type: 大会报告

高性能数据同化算法实现

Monday, 10 July 2023 09:15 (30 minutes)

Summary

Presenter: 谭, 光明

Session Classification: 大会报告

Contribution ID: 117

Type: 分会报告

Disruption prediction on EAST with different wall conditions based on a multi-scale deep hybrid neural network

Monday, 10 July 2023 14:00 (15 minutes)

Plasma disruption is a very dangerous event for future tokamaks and fusion reactors. Therefore, predicting disruption is crucial for ensuring the safety and performance of reactors. In this study, the features of two deep learning algorithms are integrated to establish a multi-scale hybrid network disruption predictor. Firstly, 43 diagnostic signals are extracted by a convolutional neural network (CNN), and the time information is learned by a long short-term memory network. The predictor is trained and tested on a database containing ~ 104 non-all-metal wall discharges. Its area under the receiver operator characteristic curve (AUC), which is a common performance metric for deep learning algorithms, reaches 0.97, and the true positive rate is ~ 95.3%, while the false positive rate is ~ 8%. Since EAST was upgraded in 2020, the wall condition has been upgraded from non-all-metal to all-metal. To examine the robustness of the predictor, the EAST disruption predictor is migrated to the all-metal wall experiment for the first time. It is again tested under the all-metal wall experimental data, and its warning performance decreases significantly, with an AUC of only 0.79. To improve the robustness and sensitivity of the predictor against disruptions, the convolutional attention mechanism is introduced into the CNN. After training and testing with the same data set, the warning performance for the all-metal wall data is improved, with the AUC value increasing to 0.84. To further improve the robustness of the predictor, the t-SNE algorithm is employed to explore the difference of the sample distribution before and after EAST upgradation, and the transfer learning algorithm is used to reduce this difference. By applying transfer learning to a small amount of discharges from the all-metal wall data, the warning performance in the all-metal wall experiment is further improved, and the AUC value increases to 0.93.

Summary

Primary authors: 陈, 大龙 (ipp); 郭, 笔豪 (中国科学院合肥物质科学研究院); 沈, 国 (ipp)

Presenter: 郭, 笔豪 (中国科学院合肥物质科学研究院)

Session Classification: 人工智能与应用

Track Classification: 人工智能与应用

Contribution ID: 118

Type: 大会报告

开幕式

Monday, 10 July 2023 08:30 (15 minutes)

Primary author: 赵海兴, 青海师范大学副校长

Presenter: 赵海兴, 青海师范大学副校长

Session Classification: 大会报告

Contribution ID: 119

Type: 大会报告

新型信息化基础平台

Monday, 10 July 2023 10:45 (30 minutes)

Summary

Presenter: 齐, 法制 (高能所)

Session Classification: 大会报告

Contribution ID: 120

Type: **not specified**

AI 大模型的原理和进展

Monday, 10 July 2023 09:45 (30 minutes)

Summary

Presenter: 张, 正德 (中科院高能物理研究所)

Session Classification: 大会报告

Contribution ID: 121

Type: **not specified**

机器学习和量子机器学习和在高能量前沿对撞机物理上的应用

Monday, 10 July 2023 11:15 (30 minutes)

Summary

Presenter: 方, 亚泉 (高能所)

Session Classification: 大会报告

Contribution ID: 122

Type: 分会报告

光源实验中虚拟计算系统的设计和应用

Tuesday, 11 July 2023 14:45 (15 minutes)

同步辐射光源是最先进的大科学装置和实验技术之一，用于研究物质的结构和性质。光源实验会产生大量实验数据，用于多样化的科学分析，同时需要大量的计算和存储资源来支持实验的进行和数据处理。为了满足同步辐射光源的实验需求，需要使用先进的计算和存储技术提供高性能、高可靠性和高扩展性的服务。云计算以高性能计算能力、灵活性和弹性扩展性、高可靠性和高可用性及高安全性等优势成为光源等大科学装置的最佳选择。云计算平台可以提供虚拟化和容器化技术，帮助大科学装置实现资源的灵活配置和管理。容器技术性能损失小、启动时间短。虚拟机之间是完全隔离的，安全性更高，可以运行多种操作系统。本文重点介绍虚拟化技术在光源实验中的设计和应用，包括虚拟云桌面系统、虚拟计算系统、虚拟 PCI 设备等，重点介绍虚拟云计算平台的体系结构、服务模式、认证系统、PCI 设备的虚拟化和资源混合调度策略的设计。最后，给出了虚拟云桌面系统在光源实验中的实际应用，展示了虚拟化技术的优越性和在同步辐射光源领域的良好应用前景。

Summary

Primary authors: 徐, 吉平 (高能所); HU QINGBAO, 胡 (高能所); 李, 亚康

Presenter: 徐, 吉平 (高能所)

Session Classification: 先进光源数据与软件

Contribution ID: 123

Type: 分会报告

用于 GEM 探测器的自研前端多通道 ASIC 性能测试

Tuesday, 11 July 2023 14:00 (15 minutes)

用于 GEM 探测器的自研前端多通道 ASIC 性能测试

曾莉欣 1,2,3, 周健荣 1,2,3, 周晓娟 1,2,3, 朱林 1,2,3, 赵豫斌 1,2,3, 陈少佳 1,2,3, 刘洪斌 1,2,3, 殷伟刚 1,2,3, 于 莉 1,2,3, 骆 宏 1,2,3, 孙志嘉 1,2,3

(1. 中国科学院高能物理研究所, 北京 100049; 2. 散裂中子源科学中心, 广东 东莞 523803; 3. 核探测与核电子学国家重点实验室, 北京 100049)

摘要: 随着中子通量的提升, 最高计数率可达 10MHz/mm² 的 GEM(Gas Electron Multiplier) 中子探测器有望替代传统 ³He 气体探测器成为未来中子实验装置主要探测器之一, 近些年得到大力发展研究, 并已经作为束流监测器成功应用到 CSNS(China Spallation Neutron Source)。当前与其匹配的前端读出电子学芯片主要采用的是德国海德堡大学研制的商用 CIPix ASIC 芯片, 该芯片集成 64 通道电荷灵敏前置放大器、成型放大器以及数字比较器, 可实现多通道测量, 但单通道最高读出计数率仅能达到 330kHz, 因而限制了 GEM 探测器的发展。为了提升整体系统性能, 实现 GEM 探测器高计数率和多通道测量的需求, 自研了一款 32 通道前端读出 ASIC 芯片, 目前已经完成第一版原型设计, 为了评估原型芯片的相关性能以及推进后期性能改进, 专门针对自研芯片的功能设计了一款前端读出子板, 该板集成了两片自研前端芯片, 可实现 64 路通道读出, 采用一片四通道高性能 DAC(Digital Analogue Converter), 单通道为 16 路输出提供比较阈值, 并通过 FPGA(Field - Programmable Gate Array) 实现四通道 DAC 的配置, 同时为了减少输出线路数量, 将接收到的 64 路数字信号, 实现 8: 1 的通道复用功能, 转为 8 路差分输出, 传输给后端数字处理板。为了与商用 ASIC 芯片进行了对比实验测试, 评估芯片性能, 完成了电子学测试系统的搭建。测试结果表明: 与商用 ASIC 相比, 自研 ASIC 的线性动态范围更大, 灵活性更强。

Summary

核电子学与探测器技术

Primary author: 曾, 莉欣

Presenter: 曾, 莉欣

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 124

Type: 大会报告

基础学科数据中心

Tuesday, 11 July 2023 11:20 (30 minutes)

Summary

Presenter: 胡, 良霖

Session Classification: 大会报告

Contribution ID: 125

Type: **not specified**

开放科学时代的空间科学数据治理与应用——国家空间 科学数据中心实践及服务

Tuesday, 11 July 2023 09:00 (30 minutes)

Summary

Presenters: 许, 琦 (中国科学院国家空间科学中心); 邹, 自明 (中科院空间中心)

Session Classification: 大会报告

Contribution ID: 127

Type: 分会报告

GECAM 卫星科学数据发布和警报信息提取技术

Wednesday, 12 July 2023 16:45 (15 minutes)

引力波暴高能电磁对应体全天监测器 (Gravitational wave high-energy Electromagnetic Counterpart All-sky Monitor, GECAM) 卫星是专门针对引力波伽马暴 (即引力波高能电磁对应体) 的研究机遇而提出的, 通过双星在地球两侧共轭星座布局, 在轨监测引力波伽马暴等高能辐射现象, 破解宇宙中致密天体剧烈并合之谜。同时探测快速射电暴、特殊伽马暴 (超长暴、软暴)、磁星爆发、以及太阳耀斑和地球伽马闪等空间高能辐射现象, 进一步理解它们的爆发机制。GECAM 卫星是我国首颗能够快速下传及发布天文警报的卫星, 已经开始发布科学数据。本报告介绍 GECAM 卫星科学数据发布和人工智能技术在警报信息提取中的应用。

Summary

Primary authors: ZHANG, Hongmei (高能所); 王, 文帅 (高能所)

Presenter: 王, 文帅 (高能所)

Session Classification: 科学计算技术与平台

Track Classification: 人工智能与应用

Contribution ID: 128

Type: 分会报告

面向先进光源的数据处理 workflow 管理系统

Monday, 10 July 2023 16:15 (15 minutes)

新一代先进光源是满足国家重大战略需求、助力基础科研突破的先进实验平台，如在建的高能同步辐射光源（High Energy Photon Source, HEPS）。这类大科学装置的特点是实验线站众多（HEPS 共 90 多条）、学科领域丰富、实验分析方法复杂，从而导致数据处理面临诸多挑战：高通量多模态的数据、灵活多样的学科方法学和高度差异化的分析流程。长久以来，国内外一直缺少通用的、用户友好的、功能完善的实验数据处理全流程管理系统。高能所计算中心 HEPS 软件组将实现一套自主可控的可视化通用管理系统，解决如下关键问题：线站科学家、实验用户和方法学开发人员如何将数据处理方法快速共享和应用到实验分析之中；科研人员如何灵活定制和监控复杂多样的数据处理流程；实验分析全过程如何批量化应用到同类实验及结果复现。此次报告将汇报此项目的背景、意义和进展，以及未来的计划安排。

Summary

Primary authors: Dr 孙, 浩凯 (高能所); 胡, 誉 (高能所)

Co-authors: 刘, 建利; Dr 符, 世园 (高能所); Dr 王, 磊 (高能所); Dr 刘, 锐 (高能所); Dr 王, 爽 (高能所); Ms 杨, 阳 (高能所); 李, 海波 (高能所计算中心); 齐, 法制 (高能所)

Presenter: Dr 孙, 浩凯 (高能所)

Session Classification: 先进光源数据与软件

Contribution ID: 129

Type: 分会报告

通过机器学习理解 X 射线吸收谱：实验结果到理论计算的自动匹配

Tuesday, 11 July 2023 14:15 (15 minutes)

通过机器学习理解 X 射线吸收谱：实验结果到理论计算的自动匹配

王可心，卢项尉，王天阳，陈留国，刘功发，陈凯 *

中国科学技术大学国家同步辐射实验室

Email: kaichen2021@ustc.edu.cn

过去十年，材料信息学显著加快了新材料性能的分析。加速相关材料发现的关键因素之一是高通量实验中的动态数据分析，引发了对材料特性快速而准确的自动估计的需求。X 射线吸收光谱 (XAS) 是一种广泛使用的材料表征技术，用于确定氧化态、配位环境和其他局部原子结构信息。XAS 分析依赖于测量光谱与可靠参考光谱的比较。然而，现有的 XAS 光谱数据库在可用参考光谱的数量和化学覆盖范围方面都非常有限 [1]。在参考光谱的基础上，实现实验结果到理论计算的自动匹配的关键要求是相似性度量函数的选择，以及上述函数在实验结果中可能存在的峰移、峰展宽和噪声等情况下的适用性 [2]。我们利用 *Quanty*——一种量子多体计算程序计算了 Co 元素两种价态 (2+, 3+) 在不同晶体场劈裂下的 L-edge 吸收谱作为参考谱 (Fig. 1a)，通过 Pearson 函数计算相似度可以在空间中准确区分元素的价态及晶体场信息 (Fig. 1b)。在此基础上，实现了来自于不同样品（价态和自旋态）实验吸收谱和理论计算结果 (Fig. 1c)，以及更复杂的具有混合价态 (Fig. 1d) 以及混合自旋态 (Fig. 1e) 的实验吸收谱和理论计算的自动匹配。该方法可帮助同步辐射吸收谱学用户快速而准确理解实验数据，加速新材料物性的研究和发现。

Fig 1: 实验结果到理论计算的自动匹配。(a) Co 元素两种价态 (2+, 3+) 在不同晶体场劈裂下的 L-edge 吸收谱 (b) Pearson 函数计算相似度的空间分布。不同样品 (c)，混合价态 (d) 以及混合自旋态 (e) 的实验吸收谱和理论计算结果的自动匹配。

[1]: Chen Zheng et al., npj Computational Materials: 4, No: 12 (2018).

[2]: Yuta Suzuki et al., npj Computational Materials: 5, No: 39 (2019).

Summary

Primary author: CHEN, kai (University of Science and Technology of China)

Presenter: CHEN, kai (University of Science and Technology of China)

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 130

Type: 分会报告

基于机器学习在正电子谱学分析方法的研究

Tuesday, 11 July 2023 16:45 (15 minutes)

正电子谱学是一门在材料的探索与表征方面的成熟技术，广泛应用于材料科学、化学、生物医学等领域。正电子湮没谱学基本实验方法有正电子湮没寿命测量、湮没辐射角关联测量、湮没辐射多普勒展宽测量、双探头 CDB 测量以及慢正电子束技术等。本研究提出了基于机器学习的方法对正电子谱学进行分析。机器学习中广泛应用的神经网络是一种模拟人脑神经系统的计算模型，由多个神经元和它们之间的连接组成。每个节点接收输入信号并产生输出信号，这些信号通过连接传递到其他节点。通过对输入数据进行多次迭代学习，神经网络能够自动学习和提取特征，并在未知数据上进行预测和分类。基于神经网络的机器学习方法可以用于反解正电子寿命谱，通过大规模的网络事例数训练，神经网络能够自动给出未知寿命谱的各项参数，且解谱精度仅取决于神经网络的数量和网络学习时间。另外，神经网络还能用于对连续晶体阵列探测器的正电子湮没角关联测量进行定位研究和对多普勒展宽能谱中的噪声进行降噪处理。除了神经网络之外，随机森林和多项式朴素贝叶斯等机器学习的方法也可以用于多普勒展宽能谱的谱学数据处理中，减少探测器和电子学系统等带来的噪声，提高其分辨率。综上所述，机器学习在正电子谱学分析中的应用为该领域带来了新的发展机遇。这些研究成果为正电子谱学的进一步发展提供了有益的参考，并促进了相关领域的研究和应用。

Summary

Primary authors: 徐, 佳一; 况, 鹏 (中国科学院高能物理研究所); 刘, 福雁 (中国科学院高能物理研究所); 张, 鹏 (中国科学院高能物理研究所); 王, 海英 (中国地质大学 (北京)); 曹, 兴忠 (中国科学院高能物理研究所); 王, 宝义 (中国科学院高能物理研究所)

Presenter: 徐, 佳一

Session Classification: 数据处理软件与分析方法

Track Classification: 数据处理软件与分析方法

Contribution ID: 131

Type: 分会报告

JUNO-DAQ 远程监控系统的设计

Tuesday, 11 July 2023 14:45 (15 minutes)

JUNO 江门中微子实验是一项位于中国广东省江门市的大型中微子探测器项目。该项目旨在通过探测中微子的性质和行为，深入研究中微子的物理学和天文学问题，包括中微子的质量、振荡、CP 对称性破缺等。江门中微子实验的运行离不开探测器，电子学，数据获取，物理分析等环节。

其中数据获取系统除了需要完成基本的取数功能以外，还必须对各个探测器以及其读出电子学的工作状态进行监测，快速识别和告知实验中可能的异常，值班人员可以随时随地监控 DAQ 系统的运行状态，及时处理运行过程中出现的问题。但当前各类数据信息分布在不同的平台，信息显示不全面，出现故障发现问题并不及时，因此为了减轻值班人员工作压力，提高远程工作效率，需要建立 JUNO-DAQ 的远程监控系统。

充分了解 JUNO 项目的实际情况，从功能性和非功能性两个方面完成系统的设计。在功能性上，系统应实现为数据获取系统的在线软件的各个服务提供图形界面接口；显示数据流的各类曲线图和直方图；实现各探测器运行状态的 3D 显示；能够支持动态监测参数的变化；支持自动化检查并发布报警。非功能上，系统应通过分角色登录的方式保障安全问题；在性能上满足可操作性，可靠性和耦合性，面向值班人员提供友好的软件界面。

本文主要对 JUNO 数据获取系统远程监控系统的设计进行介绍，目前系统正处于开发阶段。

Summary

Primary authors: 吴, 寅慧 (高能物理研究所); LI, Fei (IHEP, CAS, Beijing, China)

Presenter: 吴, 寅慧 (高能物理研究所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 132

Type: 分会报告

高能所科研管理信息化数据接口服务平台设计与应用

Wednesday, 12 July 2023 16:30 (15 minutes)

本报告旨在介绍高能所科研管理信息化数据接口服务平台的设计和应用情况。该平台旨在提供高效、可靠和安全的数据接口服务，以满足高能所科研管理的需求。通过该平台的设计与应用，预计可以提升高能所科研管理的效率和质量，促进科研成果的共享与交流，推动高能所科研工作的创新与发展。

Summary

Primary authors: ZHANG, rui; 孙智慧, zhihui (IHEP)

Presenter: ZHANG, rui

Session Classification: 科研信息化管理与系统

Contribution ID: 133

Type: 分会报告

面向先进光源软件开发过程的 CI/CD

Tuesday, 11 July 2023 15:00 (15 minutes)

高能同步辐射光源 HEPS 是基础科学和工程科学等领域原创性、突破性创新研究的重要支撑平台。HEPS 具有低水平的发射度、极高的空间相干性和高亮度等特点，每天产生海量数据且具有高通量、大容量等特点。面对以上特点，传统的数据处理软件和分析方法已经不能满足 HEPS 数据处理需求，许多科学实验需要在实验过程中提供实时分析和快速反馈，以提供决策指导和修正实验过程。而目前传统学科软件的功能固化、数据处理流程单一；新开发的软件开发周期长，需要学科工作人员和开发人员反复沟通。为了适应 HEPS 实验的工作强度和改变传统软件开发困境，需要引入一套“持续开发、持续交付、持续部署”（CI/CD）的软件开发实践流程，以实现软件开发过程的自动化和智能化。CI/CD 核心是通过引入自动化手段来提高软件交付效率，使得学科工作人员更关注科学算法本身而不是开发流程，软件开发人员更关注软件开发过程而省去大量的重复性编译、交付、部署等工作，缩短学科软件的开发时间，提高实验效率。此次报告将汇报 CI/CD 的背景介绍、部署流程和开发进展，以及未来的计划安排。

Summary

Primary authors: 刘, 建利; 胡, 誉 (高能所); 孙, 浩凯 (高能所); HU QINGBAO, 胡 (高能所); 齐, 法制 (高能所)

Presenter: 刘, 建利

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 134

Type: 分会报告

JUNO 数据获取系统高可用在线软件设计与研制

Tuesday, 11 July 2023 15:15 (15 minutes)

江门中微子实验（Jiangmen Underground Neutrino Observatory, JUNO）位于阳江核电站和台山核站的中垂线上，实验使用的中心探测器位于地下 700 米，内部装有 2 万吨的液体闪烁体，能量精度到达 3%。JUNO 将通过实验获得的数据确定中微子质量顺序和中微子混合参数，并在其它交叉领域展开研究，为研究超出标准模型的物理学提供了机会。探测超新星中微子事件是 JUNO 实验的物理目标之一，由于超新星中微子事件 30~40 年才会发生一次，超新星中微子事件到来时数据获取（Data Acquisition, DAQ）系统在线软件稳定运行对获取超新星中微子数据十分重要，这对 JUNO 在线软件的可用性提出了较高要求。

在线软件要为 JUNO DAQ 系统提供重要功能支撑和公共软件库，如在线软件、文件系统、监控系统等，在线软件又可以分为配置管理、进程管理、运行控制和信息共享等功能。在线软件所有功能的设计和实现都要考虑软件可用性，降低不可用时间，为实验探测超新星中微子提供保障。本论文通过对高可用系统设计方法的研究，实现了微服务架构的在线软件，利用消息中间件和分布式协调服务隔离了 JUNO DAQ 系统中的数据流软件和在线软件，消除了在线软件异常对数据流软件的影响，完成了在线软件核心服务的开发和可用性研究。针对在线软件的基础设施软件，本论文完成了文件系统和监控系统的调研，通过高可用系统中常用设计方法完成了基础设施软件的高可用部署。完成在线软件的设计和部署后，本论文通过对 JUNO 在线软件建模，完成了对系统可用性的分析并给出了系统可用性指标，计算表明 JUNO 在线软件可以容忍每年 41 次故障，能够为 JUNO 获取实验数据提供有力支持。

本论文还对在线软件的功能进行了测试，测试结果表明监控系统能够完成对 DAQ 系统中基础设施软件、数据流软件和计算节点的全面监控，文件系统可以为 DAQ 用户提供文件的管理和存储。在线软件中的进程管理功能可以在 20 秒左右对数据流软件完成启动和停止控制；运行控制功能可以在 12 毫秒左右将命令信息分发给所有数据流进程；配置管理功能可以管理配置历史，在 78 毫秒左右将配置信息分发给所有数据流进程；消息报告服务可以处理数据流软件在百台节点产生的运行消息。

最后，为向数据流软件提供可选的高可用方案，本论文研究了高可用数据传递框架的性能和容错能力，测试结果表明，在单节点带宽 25 Gb/s 的集群中，基于发布订阅模型的高可用数据传递框架可以实现单节点 23 Gb/s 的数据吞吐，在单节点发生故障时可以保证数据流持续运行。

在线软件核心服务目前已经在 JUNO 实验现场部署，联合数据流软件完成了功能和可用性测试，为 JUNO 未来正式取数运行做出了重要铺垫。

Summary

Primary author: Dr 于, 泽众

Presenter: Dr 于, 泽众

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 135

Type: 分会报告

面向光源学科数据分析场景的交互式计算平台

Tuesday, 11 July 2023 14:30 (15 minutes)

光源学科数据分析需求和传统的高能物理计算存在较大差异，高能物理学科用户多采用命令行模式，登录计算集群 UI 节点，提交批处理作业进行数据分析，分析结果的时效性受作业排队时长影响较大。光源学科用户通常直接使用科学软件分析工具处理实验数据，快速获取分析结果调整后续分析策略，对软件操作的可交互性和分析结果的时效性要求较高。结合光源数据分析应用特点，HEPS 设计了面向光源科学分析场景的交互式计算平台，用户无需了解 ssh 等终端操作命令，直接通过浏览器随时随地接入计算集群环境，访问集群内部资源。同时，交互式计算平台使用 `python ui` 工具包，跳过命令行和代码编辑步骤，直接面向用户提供可交互的 UI 软件，用户通过点击按钮选取输入数据，即可基于集群计算资源快速完成数据分析，提升了用户对计算资源使用的便利性。

Summary

Primary authors: 胡, 庆宝 (高能所); 徐, 吉平 (高能所); LUOQI, UNKNOWN (高能所); 郭, 超奇 (高能所); 欧歌, Ge (高能所)

Presenter: 胡, 庆宝 (高能所)

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 136

Type: 分会报告

基于 GPU 的 N-hit 软件触发算法研究与实现

Wednesday, 12 July 2023 14:15 (15 minutes)

数据获取系统基于软件完成在线数据处理，具备灵活性和可处理复杂逻辑等优势。随着实验技术的不断发展，国内外当前建设和规划中的先进装置产生的数据带宽已经达到甚至超过 TBytes/s 的量级，对实验数据的实时处理提出极高的要求，因此亟需发展先进的数据获取在线处理技术。CPU/GPU 混合异构系统是当今先进的高性能计算平台，具有强劲的计算能力、高性价比、低能耗和易扩展等优势。

在在线处理中，软件触发是至关重要的一环。它依赖于预定义的物理事件特征，在实时模式下对数据进行分析 and 筛选，以剔除无效的背景数据、本底噪声和干扰数据等，从而保留实验所需的事件数据。

本文介绍了在无触发数据获取系统中常用的 N-hit 软件触发算法，研究了基于排序的 N-hit 算法，并利用 CUDA 动态并行特性对快速排序算法进行实现，通过与传统 CPU 快速排序进行比较，证明了动态并行排序算法在性能方面具有更高的优势。此外，基于 GPU 实现了一种在大通道数和高事例率场景下具有显著优势的触发算法，该算法的核心思想是使用一个触发表，将线程块映射到对应的通道，线程块中线程处理相应通道的数据，通过使用共享内存存储过程中的数据，充分发挥 GPU 高并发特性来加速触发表填表过程，从而提高在大通道和高事例率场景下的软件触发性能。

Summary

Primary authors: 江, 晓照 (高能所); 季, 筱璐 (高能所); 彭宇, UNKNOWN (高能所)

Presenter: 江, 晓照 (高能所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 138

Type: 分会报告

JUNO-DAQ 数据流读出模块的升级

Tuesday, 11 July 2023 14:15 (15 minutes)

江门中微子实验旨在测量中微子的质量顺序和精确测量中微子的振荡参数。该实验主要有探测器，电子学，数据获取，离线分析等系统的参与。读出模块是 JUNO DAQ 的数据流软件的第一级模块，主要功能包括前端电子学各通道的数据读出，按时间进行分片打包（预处理）并根据时间片对所有通道的数据进行一级组装。

原有的读出模块具备基本的取数功能，为满足实际应用需求，对其的升级主要包括性能优化和异常处理两方面的提升。性能优化方面，主要是修改了一级组装的判断逻辑，减少了各通道的遍历次数，提高了组装效率。异常处理方面，原先的程序缺少对异常的应对策略，发生异常后会导致程序卡死或退出。程序运行中遇到的主要异常包括异常数据和异常通道。对于异常数据，在预处理线程增加了数据格式的检查，对异常数据进行处理后定位至下一个正常数据。异常通道的处理的基本功能则是在一级组装中增加强制组装逻辑，当一个或多个通道长时间没有数据时，强制其它通道进行组装，确保数据流的正常运行。后期的进一步升级则考虑实现异常通道的动态删除和恢复，并提高读出模块在强制组装过程中运行的稳定性。

目前 JUNODAQ 数据流软件仍处在升级和测试过程中，模拟数据源测试的读出模块单节点最大读出性能约为 5.6GB/s。在实验现场完成了数次正式运行环境的测试中，新增的异常处理功能运行正常。

Summary

Primary author: 陈, 超 (高能所)

Co-authors: LI, Fei (IHEP, CAS, Beijing, China); ZENG, Tingxuan

Presenter: 陈, 超 (高能所)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 139

Type: 分会报告

面向格点量子色动力学的数据结构和矢量化方法

Wednesday, 12 July 2023 16:00 (15 minutes)

格点 QCD 是在粒子物理“标准模型”理论框架下，研究强相互作用力的一种第一性计算的非微扰方法，也是在低能标度下唯一有效的数值模拟方法。格点 QCD 将闵可夫斯基时空上的量子场论问题转变为欧几里得时空上的统计问题，将时间和空间等同对待，将四维时空切分为有限的四维晶格。格点 QCD 计算应用中的热点模块集中于循环边界条件的四维超立方格子中的一个稀疏线性系统求解。该线性系统表现为四维八点的协变差分运算，即在任意点上，是该点以及与之相邻的 8 个点到该点的 9 个 12×12 稠密复数矩阵计算。

本文介绍了一种基于结构网格数据结构的矢量化方案。在格点 QCD 计算过程中，涉及大量、同构、小型复矩阵的计算：大量是指要遍历所有网格点；同构是指每个网格点的数据结构和计算是相同的；小型复矩阵具体到格点 QCD 则为 $SU(3)$ 矩阵。在这些计算中有的矢量化难以实现或效果较差，有的则是对较宽的矢量单元难以适用。因此，面向这类具有复杂数据类型的结构化网格，本文提出一种新的数据类型——Tensor<Grid>，事项更高效的矢量加速，并且对矢量位宽具有更强的扩展性。该方法对其他类似的结构化网格问题同样具有帮助作用。

Summary

Primary authors: ZHANG, Ke-Long (中国科学院计算机网络信息中心); SHUN, Xu (中国科学院计算机网络信息中心)

Presenter: ZHANG, Ke-Long (中国科学院计算机网络信息中心)

Session Classification: 科学计算技术与平台

Track Classification: 科学计算技术与平台

Contribution ID: 141

Type: 分会报告

Hepsptycho 多 GPU 及 AI 并行软件开发进展

Tuesday, 11 July 2023 14:00 (15 minutes)

Ptychography 是一种相干衍射成像技术, 能够提供在 X 光衍射波长极限内样品的相位分辨率重建图像。在光学实验中, 使用小孔或者其它光学仪器将 X 光探针聚焦到待扫描样品上。样品每个预定位置的衍射图样由放置在样品 1m 距离 (或更远) 之后的探测器捕获。相邻两个扫描位置需要一定的重叠比例, 以此来保证被记录的实验数据能够多倍冗余, 为 ePIE 算法或者 DM 算法的重建提供足够多的信息。

针对 ePIE 算法和 DM 算法, 现有的开源软件并未提供 QT 界面及提供完备的工作流, 无法为用户提供友好的实验元数据参数的快速导入 GUI 及基于 CUDA C++ 的多 GPU 并行重建迭代过程。Hepsptycho 软件由高能物理研究所计算中心与多学科中心 B2 线站联合开发, 程序核心均由 CUDA C++ 实现, 以实现在 NVIDIA GPU 上的并行计算, 同时使用 MPICH 作为多 GPU 卡协同并行重建的框架, 实现了迭代收敛的计算时间随 GPU 增加而线性递减。其元数据由 HDF5 文件提供导入, 可以大大降低用户的软件学习成本。

在 ePIE 算法的基础上, 计算中心与 B2 线站初步改写了算法核心模块, 将其升级为快速稳定收敛的 mPIE 算法。

针对相干衍射实验数据量巨大而导致从硬盘读入缓慢的问题, Hepsptycho 借鉴开源软件的单进程内多线程读入经验而非多 MPI 进程并行读入经验, 大幅提高了 IO 的速度。

同时, 由计算中心及 B2 线站联合开发的 AI 快速重建模块 (名称暂定为 W1-Net 神经网络模块), 正在逐步集成到 Hepsptycho 软件中, 这为用户进行实际光学实验前的参数快速调试提供了强大的模拟支撑。

在国产化加速卡及软件框架应用方面, hepsptycho 正在联合中科曙光及百度的科学家, 尝试使用国产 DCU+ 百度 paddlepaddle 推理框架实现 AI 的训练及推理过程, 而非 NVIDIA GPU+ TensorFlow 的组合, 以实现“国产替代”的目标。

Summary

Primary author: Dr 磊, 王 (IHEP)

Co-authors: 洋洋, 穆 (IHEP); 成烨, 邢 (IHEP); 广才, 常 (IHEP); 家瑞, 胡 (IHEP); 京燕, 石 (IHEP); 法制, 齐 (IHEP); 誉, 胡 (IHEP); 建利, 刘 (IHEP); 浩凯, 孙 (IHEP); 世园, 符 (IHEP); 锐, 刘 (IHEP); 阳, 杨 (IHEP); 爽, 王 (IHEP)

Presenter: Dr 磊, 王 (IHEP)

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 142

Type: 分会报告

基于 CPU-FPGA 异构计算的可计算存储系统设计与实现

Wednesday, 12 July 2023 16:15 (15 minutes)

大数据和云计算技术的发展，使得“存-算分离”的计算模式广泛应用于大型数据中心。在计算节点和存储节点间频繁的数据搬运，占用了大量的访存带宽和网络带宽。当计算作业达到一定规模时，会遇到“内存墙”问题，以及各种网络问题、文件系统故障等，使整个计算系统处于不稳定状态。另外，由于存储器的访问速度低于处理器的运算速度，在访存带宽或网络带宽达到上限后，增加运算部件无法继续提升处理能力。本文基于 ARM CPU 和 FPGA 异构计算架构，设计并实现了可计算存储系统。利用 FPGA 可定制性强、并行度高的特点，将部分计算密集型任务和 I/O 密集型任务卸载到存储节点本地。通过本地的加速计算，缓解了节点间的数据搬运造成的 I/O 瓶颈和网络拥塞问题，提升了计算效率，增强了系统稳定性，同时降低了整体功耗。

Summary

Primary authors: LI, Haibo (高能所); GAO, Yu (IHEP); ZHANG, Minxing; Mr CHENG, Yaodong (IHEP); CHENG, Yaosong

Presenter: GAO, Yu (IHEP)

Session Classification: 科学计算技术与平台

Track Classification: 科学数据存储与管理技术

Contribution ID: 143

Type: 分会报告

基于 FPGA 的粒子物理图神经网络加速器的设计与实现

Monday, 10 July 2023 14:45 (15 minutes)

Jet tagging 是粒子物理实验中的一项关键分类任务。近年来，深度学习方法的引入使得 Jet tagging 分类任务的准确率大幅提高，其中以 ParticleNet 为代表的图神经网络在该任务中表现出色。若将该模型部署在大型对撞机上，需要设计低功耗、低延迟的实时处理系统，使用传统硬件难以满足应用需求。目前 FPGA 由于其低功耗、低延迟、硬件可编程等特性成为 AI 部署加速的前沿研究热点，相对于 CPU，FPGA 可以实现更好的并行操作，相对于 GPU 平台，使用 FPGA 可以提高计算效率并降低功耗。因此，将 ParticleNet 在 FPGA 上实现和加速，可以实现快速、低功耗的执行粒子物理学中的分类任务，从而减少经济成本并加快粒子物理中数据处理进程。本设计将基于 HLS 实现可重构架构，并将计算密集型任务卸载到 FPGA 上运算，以减少 ParticleNet 的算法延迟。

Summary

Primary author: 张, 玉涛**Presenter:** 张, 玉涛**Session Classification:** 人工智能与应用**Track Classification:** 人工智能与应用

Contribution ID: 144

Type: 分会报告

基于多特征融合和增量学习的未知流量识别技术

Monday, 10 July 2023 15:00 (15 minutes)

为了进行网络管理、安全监控并确保网络服务的质量，对互联网流量进行准确分类和识别至关重要。然而在现实世界中，许多不同种类的网络应用被定期开发和更新，产生大量在训练集流量类别之外的未知流量，这些未知的流量类型会对当前机器学习模型的准确性造成重大影响，降低流量分类的准确率。现有的未知流量分类算法无法优化流量特征，每次收集到新的流量数据时都需要对整个系统进行重新训练，导致识别效率低下，不适合实时流量检测。为解决上述问题，我们提出了一种基于多特征融合的增量学习方案来检测未知流量。该方法采用了多通道并行架构来提取流量的时间和空间特征。然后引入了 mRMR 算法对从每个通道提取的特征进行排序和融合，以克服加密流量特征冗余的问题。此外，我们结合基于密度比的聚类算法来识别未知的流量特征，并通过增量学习更新模型。分类器通过学习新获得的类知识，实现对已知和未知流量的实时分类。

我们使用公共数据集 ISCX-VPN-Tor 验证在不同场景下该模型对未知的加密隧道流量的检测能力，准确率为 86% 以上。此外，模型在入侵检测数据集 NSL-KDD 上达到了 90% 以上的准确率。该工作为识别未知网络流量提供了一种新颖的方法，有助于维护网络环境的安全和稳定。

Summary

Primary authors: 刘, 君怡 (IHEP); 王佳荣, UNKNOWN; YAN, TIAN (IHEP); 齐, 法制 (高能所); CHEN, Gang (IHEP)

Presenter: 刘, 君怡 (IHEP)

Session Classification: 人工智能与应用

Track Classification: 人工智能与应用

Contribution ID: 145

Type: 分会报告

面向 HEPS 的 IO 方法设计与优化

Wednesday, 12 July 2023 14:45 (15 minutes)

北京在建的同步辐射光源装置预计每天产生数百 TB 的数据量，每年的数据量达到 PB 量级，对 IO、存储和科学计算带来极大压力。实验过程中需要在线处理用于快速判断数据采集质量，目前从磁盘读取海量实验数据读取存在严重的 IO 瓶颈，因此 HEPS 亟需稳定高效的 IO 方法克服以上困难，首先分析光源下计算任务的读取模式，结合 HDF5 分块存储特性，减少数据跳读，结合并行异步策略加速读写，减少 IO 在计算过程中的占比；其次通过压缩的方式减少数据体积，为保证数据完整性，压缩采用无损的方式，引入压缩会带来额外的时间和资源消耗，而不同的数据压缩效果也有所不同，所以，以加速整个科学计算为目标，综合评价引入压缩的提升，自动触发压缩过程及压缩方法。因此本文拟通过以上方法优化 HDF5 在 HEPS 科学计算过程中的 IO 速度，加速科学结果产出。未来以流处理的方式可以规避海量数据落盘再读取导致的 IO 瓶颈问题，因此最后本文首先介绍了未来 HEPS 场景下 IO 方法的设计思路。

Summary

Primary authors: Dr 符, 世园 (中国科学院高能物理所计算中心); Dr 胡, 誉 (中国科学院高能物理所计算中心); 刘, 建利; 齐, 法制 (高能所); Dr 孙, 浩凯 (中国科学院高能物理所计算中心); 王, 磊; 刘, 锐

Presenter: Dr 符, 世园 (中国科学院高能物理所计算中心)

Session Classification: 先进光源数据与软件

Contribution ID: 146

Type: 分会报告

基于 Daisy 的天文卫星数据处理框架

Wednesday, 12 July 2023 15:00 (15 minutes)

随着计算机技术、卫星导航技术的迅速发展，通过卫星传感器产生的数据已经达到 TB 级甚至 PB 级。目前，高能所参与了多颗天文卫星的研制和建设，包括硬 X 射线调制望远镜 (Hard X-ray Modulation Telescope, HXMT)、引力波暴高能电磁对应体全天监测器卫星 (Gravitational wave high-energy Electromagnetic Counterpart All-sky Monitor, GECAM)、增强型 X 射线时变与偏振空间天文台 (enhanced X-ray Timing and Polarimetry mission, eXTP) 等。卫星数据的海量性特点给数据处理带来了巨大的挑战，卫星数据处理流程具有复杂性和专业性，往往涉及多个领域的科学知识，同时每个处理步骤也具有高度重复性，而传统人工处理的方式耗费大量的时间和人力成本，效率低下，降低了数据处理效率。为了解决以上问题，同时实现天文卫星数据处理流程的标准化和自动化，本研究提出了天文卫星科学数据处理软件框架。该框架基于 Daisy 数据处理软件框架进行扩展和衍生，核心组成部分包括数据引擎、算法、工作流引擎等。通过调研 HEASoft (High Energy Astrophysics Software) 框架的设计理念，在框架中增加卫星专用库、卫星基础库和第三方工具库。实验表明，本研究所提框架能够支持多个天文卫星数据处理工具、算法的集成和数据处理流程的规范化，满足不同应用场景的需求，具有重要的研究意义和实践价值。

Summary

Primary author: 王, 爽 (中国科学院高能物理研究所)

Co-authors: 张, 红梅 (高能所); 杨, 阳 (高能所); 胡, 誉 (高能所); 孙, 浩凯 (高能所); 符, 世园; 刘, 锐; 刘, 建利; 王, 磊

Presenter: 王, 爽 (中国科学院高能物理研究所)

Session Classification: 先进光源数据与软件

Contribution ID: 147

Type: 分会报告

基于深度学习的漂移室电离计数重建算法的研究

Tuesday, 11 July 2023 17:00 (15 minutes)

环形正负电子对撞机 (Circular Electron Positron Collider, CEPC) 的实验计划包括在质心能量 240 GeV 处运行以研究 Higgs 粒子, 以及在 Z-pole 和 WW 产生阈值处进行电弱精密测量和味物理研究等。良好的粒子鉴别 (PID) 对于味物理和喷注研究非常重要, CEPC 的物理目标也对粒子鉴别提出了很高的要求。粒子鉴别通常基于带电粒子穿过气体探测器介质时的电离行为。传统方法通过测量单位长度内的能量损失 (dE/dx) 来进行粒子鉴别, 而电离计数方法的目标是测量单位长度内的原初电离数 (dN/dx), 从而规避测量 dE/dx 的多项涨落和次级电离的影响。 dN/dx 的泊松性质使这种新测量方法更具统计显著性, 极大地减少了测量误差, 其理论分辨率可比 dE/dx 提高一倍以上。应用电离计数的漂移室已被提议作为 CEPC 的探测器候选者。

测量 dN/dx 需要确定漂移室单元感应电流波形中与原初电离相关联的峰的个数, 即所谓的电离计数。随着快速读出电子学及数据传输技术的发展, 电离计数方法的实现具备了技术基础, 其也成为了国际粒子鉴别探测技术的热门研究方向。电离计数重建算法是 dN/dx 测量的核心软件算法, 其主要挑战是在信号高堆叠和有噪声污染的情况下寻出电离峰, 并区分由原初电离和次级电离形成的两种电离峰。传统算法由于信息的低效使用而很难达到所需的效率, 而深度学习算法可以应用数据中的全部信息进行学习。在本研究中, 一个基于深度学习的分两步走的算法被用于电离计数重建。该算法包括一个基于循环神经网络 (RNN) 的寻峰模型和一个基于动态图卷积神经网络 (DGCNN) 的电离峰聚团模型。在对模拟样本的测试中, 本研究提出的深度学习重建算法在寻峰和聚团两个方面的性能均优于传统算法。

Summary

Primary authors: TIAN, Zhefei (Wuhan University); ZHAO, Guang (高能所); DONG, Mingyi (IHEP); ZHOU, Xiang (Wuhan University); 刘帅毅, UNKNOWN; SUN, Sheng-Sen (Institute of High Energy Physics); WU, Linghui (IHEP); XIN, Shuiting; ZHANG ZHENYU, Zhenyu (高能所)

Presenter: TIAN, Zhefei (Wuhan University)

Session Classification: 数据处理软件与分析方法

Track Classification: 数据处理软件与分析方法

Contribution ID: 148

Type: 分会报告

开源虚拟平台在大型加速器控制系统中的应用

Tuesday, 11 July 2023 15:00 (15 minutes)

米清茹陈广花缪海峰丁建国阎映炳
(中国科学院上海高等研究院光源科学中心, 上海 201204)

在当前中科院上海高等研究院光源科学中心已建和正在建设中的多个加速器工程中, 运行环境是控制系统的核心

Summary

Primary authors: 米, 清茹 (中科院上海高等研究院); 陈, 广花 (中国科学院上海应用物理研究所); 丁, 建国 (中科院上海应用物理所)

Presenter: 米, 清茹 (中科院上海高等研究院)

Session Classification: 核电子学与探测技术

Track Classification: 核电子学与探测技术

Contribution ID: 149

Type: 分会报告

HEPS 数据处理软件框架 Daisy 的进展及规划

Wednesday, 12 July 2023 14:15 (15 minutes)

随着第四代同步辐射光源的出现，以及高分辨探测器的发展，科学数据无论是在容量还是通量上都出现了数量级的增长。先进的数据处理和分析方法需要从底层技术进行革新，跟上数据增长的速度，并提供实时的快反馈和实验指导能力。同时随着新型先进实验方法的不断涌现，方法学算法及软件的发展需要底层软件框架的支持。为满足新一代光源实验的需求，HEPS 计算与通信系统开发了面向先进光源实验的数据处理软件框架 Daisy。本次报告将介绍 Daisy 项目的设计理念、开发进展及未来发展计划。

Summary

Primary authors: 胡, 誉 (高能所); 孙, 浩凯 (高能所); FU, Shiyuan; 刘, 建利; 刘, 锐; 王, 磊; 齐, 法制 (高能所)

Presenter: 胡, 誉 (高能所)

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 150

Type: 分会报告

上海光源同步辐射科学数据管理与分析系统

Wednesday, 12 July 2023 14:00 (15 minutes)

上海光源是中国大陆第一台中能第三代同步辐射光源，目前共有 27 条光束线 39 个实验站投入运行，至 2025 年将有 35 条光束线约 50 个实验站投入运行，已接待学界、产业界用户超过 3.5 万人。国家“十二五”重大科学基础设施建设项目上海光源线站工程配套建设的用户数据中心可提供超过 20PB 的统一存储空间，提供超 900Tflops 的多节点 CPU/GPU 混合集群算力支持。上海光源同步辐射科学数据管理分析系统依托上海光源用户数据中心基础设施，从全装置层面布局核心数据库系统设计研发、高通量数据快速分析系统研发、数据自动处理流水线研发，以实现海量实验数据的全生命周期高效组织管理、高通量数据在线处理以及计算资源的高效利用，为上海光源用户提供“一站式”的数据服务，实现科学数据的长期价值与战略价值，加速装置和用户的科学产出。

Summary

Primary author: 王, 春鹏 (中国科学院上海高等研究院)

Presenter: 王, 春鹏 (中国科学院上海高等研究院)

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 151

Type: 大会报告

RDTM 介绍

Wednesday, 12 July 2023 10:00 (15 minutes)

期刊介绍 PPT

Summary

Primary author: 郑, 文莉 (高能所)

Presenter: 边, 晓琳

Session Classification: 大会报告

Track Classification: 科研信息化管理与系统

Contribution ID: 152

Type: 分会报告

先进光源表征牵引的材料数字化研究平台规划与设计

Wednesday, 12 July 2023 14:30 (15 minutes)

摘要：伴随着新材料的成分、结构、维度等复杂性指数级的增加，传统的材料研发方法面临着巨大挑战，复杂材料体系的研究更是依赖于多平台、多方法的实验与理论协同探索与分析。而随着加速器、探测器以及实验方法学的发展，我国正处于先进光源的高速发展期，其中由上海科技大学负责或参与建设的软 X 射线自由电子激光用户装置和活细胞成像项目已进入运行阶段、硬 X 射线自由电子激光装置正在体系化加快推进，将为材料科学研究提供前所未有的原子级别高时空分辨能力和原位动态观测能力。同时，近年来人工智能在前交叉科学领域已取得许多颠覆性成果。随着 ChatGPT 类模型工具的纷纷涌现和生成式人工智能技术的快速发展，以 AI 为基础的新生产和研究模式已登上历史舞台。在材料研究过程中，人工智能在电子层面、原子/分子层面、微介观层面以及宏观层面都有望带来传统科研模式的变革和加速能力提升，具有前所未有的巨大潜力。但同时对于科学数据的多样性、质量和规模，以及不同领域人才的交叉协作均提出了更高的要求。

鉴于以上背景，为将光子大科学装置的尖端探测能力和 AI 技术充分融合到材料科学研究中，本文提出了一种 Data+AI 材料数字化融合创新平台，汇聚光子大科学装置实验数据、基于高精度 DFT、MD、MC 等多尺度模拟方法的理论大数据以及电镜中心、分析测试中心、微纳加工、红外芯片工艺加工等高价值数据资源，建立具有跨领域多源融合、安全共享、可信多方协作的全生命周期数据管理和共享协作系统，实现各类高质量科学大数据、学科领域方法学软件与 AI 算法融合能力平台的构建，满足多源材料大数据的分析挖掘需求，从而构建理论计算-AI 预测-实验验证的体系，形成复杂材料体系科学研究的新模式。

关键词：材料数字化；光子大科学装置；Data+AI 融合创新平台

Summary

Primary authors: 雷, 蕾 (上海科技大学); SHI, Wujun (ShanghaiTech University); ZHANG, Xiaofeng (ShanghaiTech University)

Co-author: 怀, 平 (Shanghaitech University)

Presenter: 雷, 蕾 (上海科技大学)

Session Classification: 先进光源数据与软件

Track Classification: 先进光源数据与软件

Contribution ID: 153

Type: **not specified**

基于 SEAWAN 的科学大数据广域传输和计算

Wednesday, 12 July 2023 11:45 (30 minutes)

Summary

Presenter: 吕, 国 (中科海网（成都）网络科技有限公司)

Session Classification: 大会报告

Contribution ID: 154

Type: **not specified**

HEPS 高通量多模态实验数据采集软件研制规划与进展

Monday, 10 July 2023 17:00 (15 minutes)

新一代同步辐射光源装置在 X 射线亮度与相干性、光学元件与探测器效率上获得了极大提升，所开展的成像实验向跨尺度、多模态、高通量与原位动态化表征模式转变，实验数据的年底通量将正式步入“Exascale”时代，对实验控制、数据采集、在线重建和特征分割等方面的软件技术与算法需求发生了革命性的变化。在建的北京同步辐射（HEPS）光束线软件团队在高通量多模态成像实验数据采集与分析中结合了大型软件框架、大数据技术、AI 工作流等方法的优势，以应对未来光源实验海量数据带来的挑战。

Summary

Primary authors: 张, 一; 张, 承龙; 毕, 筱雪; 黄, 颖科; 李, 鹏程

Presenter: 李, 鹏程

Session Classification: 先进光源数据与软件

Contribution ID: 155

Type: **not specified**

科研信息化联盟发展及本期主题介绍

Sunday, 9 July 2023 09:10 (10 minutes)

Presenter: 齐, 法制 (高能所)

Session Classification: 议题

Contribution ID: 156

Type: **not specified**

欢迎致辞

Sunday, 9 July 2023 09:00 (10 minutes)

Session Classification: 议题

Contribution ID: 157

Type: **not specified**

数据共享与安全

Sunday, 9 July 2023 14:50 (20 minutes)

Presenter: Prof. 张辉, 北京航空航天大学

Session Classification: 议题 2: 科学数据安全治理体系

Contribution ID: 158

Type: **not specified**

科学数据安全性与出境管理要求与实践

Sunday, 9 July 2023 14:30 (20 minutes)

Presenter: 汪洋 (中国科学院计算机网络信息中心)

Session Classification: 议题 2: 科学数据安全治理体系

Contribution ID: 159

Type: **not specified**

新技术赋能数据安全流转

Sunday, 9 July 2023 14:10 (20 minutes)

Presenter: 刘博, 绿盟科技首席创新官

Session Classification: 议题 2: 科学数据安全治理体系

Contribution ID: **160**

Type: **not specified**

大模型及天文科学数据处理思考

Sunday, 9 July 2023 09:20 (30 minutes)

Presenter: Prof. 崔辰州

Session Classification: 议题

Contribution ID: **161**

Type: **not specified**

高能物理科学数据处理大模型考虑

Sunday, 9 July 2023 09:50 (30 minutes)

Presenter: ZHANG (张正德), Zheng-De (中科院高能物理研究所)

Session Classification: 议题

Contribution ID: 162

Type: **not specified**

私域大模型助力前沿科技创新

Sunday, 9 July 2023 10:20 (30 minutes)

Presenter: 何宇锋, 新华三集团

Session Classification: 议题

Contribution ID: 163

Type: **not specified**

议题 1 讨论：大模型发展的共性问题、困难、合作

Sunday, 9 July 2023 11:10 (1h 20m)

Session Classification: 议题 1：科学数据处理大模型

Contribution ID: 164

Type: **not specified**

在大模型时代引发的安全思考

Monday, 10 July 2023 11:45 (20 minutes)

Presenter: 卢, 煜彬 (深信服)

Session Classification: 大会报告

Contribution ID: 165

Type: **not specified**

云桌面和一线通 IP 会议室

Tuesday, 11 July 2023 11:50 (20 minutes)

Presenter: 陈, 戎 (锐捷)

Session Classification: 大会报告

Contribution ID: 166

Type: **not specified**

为科研提供强劲算力——国产海光 DCU

Tuesday, 11 July 2023 10:00 (20 minutes)

Presenter: 曙光

Session Classification: 大会报告

Contribution ID: 167

Type: 分会报告

面向先进光源的高通量分布式并行计算框架的探索与设计

Tuesday, 11 July 2023 15:15 (15 minutes)

作为我国第四代同步辐射光源装置高能同步辐射光源 (High Energy Photon Source HEPS) 开展的实验将向高数据通量、多模态、超快频率、跨尺度形式转变所产生的年数据通量预计将迈入「Exascale」时代。另外复杂多样的数据实验导致不同线站数据产生速率差异较大。面对如此高通量、容量多样化的实验数据数据处理软件需要能够调用可伸缩的分布式异构计算资源提供不同规模的计算分析服务。

Summary

Primary authors: FU, Shiyuan; 刘, 建利; 王, 爽 (SYSU); 刘, 锐; 孙, 浩凯 (高能所); 王, 磊; 胡, 誉

Presenter: 刘, 锐

Session Classification: 先进光源数据与软件

Contribution ID: **168**

Type: **not specified**

安全工作组思考及建议

Sunday, 9 July 2023 13:30 (20 minutes)

Presenter: 齐法制

Session Classification: 议题 2: 科学数据安全治理体系

Contribution ID: 169

Type: **not specified**

科学数据安全治理体系思考

Sunday, 9 July 2023 13:50 (20 minutes)

Presenter: 王健, 农业科学数据中心

Session Classification: 议题 2: 科学数据安全治理体系