arXiv:2611.28851v1 [physics.ins-det] 21 Nov 2026

费米实验室缪子 $g-2$实验中的数据处理、生产与模拟

D. P. Aguillard[32], T. Albahri[29], D. Allspach[7], J. Ananala[7], K. Badgley[7], S. Baeßler[34], L. Bailey[26], E. Barlas-Yucel[27,a], T. Barrett[6], E. Barzi[7], F. Bedeschi[20], M. Berz[16], M. Bhattacharya[7], H. P.Binney[35], P.Bloom[17], J. Bono[7], E. Bottalico[29], T. Bowcock[29], S. Braun[35], M. Bressler[31,a], G. Cantatore[12,b], R. M. Carey[2], B. C. K. Casey[7], D. Cauz[25,c], R. Chakraborty[28], A. Chapelain[6], S. Chappa[7], S. Charity[29], C. Chen[21,20,d], M. Cheng[27], R. Chislett[26], Z. Chu[20,d], T. E. Chupp[32], C. Claessens[35], F. Confortini[9,e], M. E. Convery[7], S. Corrodi[1], L. Cotrozzi[29], J. D. Crnkovic[7], S. Dabagov[8,f], P. T. Debevec[27], S. Di Falco[10], G. Di Sciascio[11], S. Donati[10,g], B. Drendel[7], A. Driutti[10,28], M. Eads[18], A. Edmonds[2,36], J. Esquivel[7], M. Farooq[32], R. Fatemi[28], K. Ferrab[29], C. Ferrari[10,h], M. Fertl[14], A. T. Fienberg[35], A. Fioretti[10,h], D. Flay[31], S. B. Foster[28,2,i], H. Friedsam[7], N. S. Froemming[18,j], C. Gabbanini[10,h], I. Gaines[7], S. Ganguly[7], J. George[31,k], L. K. Gibbons[6], A. Gioiosa[24,l], K. L. Giovanetti[13], P. Girotti[10,m], W. Gohn[28], L. Goodenough[7], T. Gorringe[28], J. Grange[32], S. Grant[1,26], F. Gray[19,n], S. Haciomeroglu[5,o], T. Halewood-Leagas[29], D. Hampai[8], F. Han[28], J. Hempstead[35], D. W. Hertzog[35], G. Hesketh[26], E. Hess[10,p], A. Hibbert[29], Z. Hodge[35], S. Y. Hoh[21,20,q], K. W. Hong[34], R. Hong[1,28], T. Hu[21,20,d], Y. Hu[20,d], M. Iacovacci[9,e], M. Incagli[10], S. Israel[2,31], P. Kammel[35], M. Kargiantoulakis[7], M. Karuza[12,r], J. Kaspar[35], D. Kawall[31], L. Kelton[28,22], A. Keshavarzi[26], D. S. Kessler[31], K. S. Khaw[21,20,d], Z. Khechadoorian[6], B. Kiburg[7], M. Kiburg[7,17], O. Kim[33,n], N. Kinnaird[2], E. Kraegeloh[32], J. LaBounty[35], K. R. Labe[6], M. Lancaster[30], S. Lee[5], B. Li[20,r], D. Li[20,s], L. Li[20,d], I. Logashenko[4,t], A. Lorente Campos[28], Z. Lu[20,d], A. Luc`a[7], G. Lukicov[26], A. Lusiani[10,u], A. L. Lyon[7], B. MacCoy[35], R. Madrak[7], K. Makino[16], S. Mastroianni[9], R. McCarthy[2,v], J. P. Miller[2], S. Miozzi[11], B. Mitra[33,w], J. P. Morgan[7], W. M. Morse[3], J. Mott[7], A. Nath[9,e], J. K. Ng[21,20,d], H. Nguyen[7], Y. Oksuzian[1], Z. Omarov[15,5], W. Osar[6], R. Osofsky[35], S. Park[5], G. Pauletta[†25,c], J. Peck[28], G. M. Piacentino[24,l], R. N. Pilato[29], K. T. Pitts[27,x], B. Plaster[28], N. Pohlman[18], C. C. Polly[7], D. Po˘cani´c[34], J. Price[29], B. Quinn[33], M. U. H. Qureshi[14], G. Rakness[7], S. Ramachandran[1,k], E. Ramberg[†7], R. Reimann[14,y], B. L. Roberts[2], D. L. Rubin[6], M. Sakurai[26], L. Santi[†25,c], C. Schlesier[27,z], A. Schreckenberger[7], Y. K. Semertzidis[5,15,aa], A. K. Soha[7], M. Sorbara[11,ab], J. Stapleton[7], D. Still[7], C. Stoughton[7], D. Stratakis[7], D. St¨ockinger[23], H. E. Swanson[35], G. Sweetmore[30], D. A. Sweigart[6], M. J. Syphers[18], Y. Takeuchi[21,20,d], D. A. Tarazona[6], T. Teubner[29], A. E. Tewsley-Booth[28,32], V. Tishchenko[3], N. H. Tran[2], W. Turner[29], E. Valetov[16,ac], D. Vasilkova[29], G. Venanzoni[29,ad], T. Walton[7], A. Weisskopf[16], L. Welty-Rieger[7], P. Winter[1], Y. Wu[1,ae], B. Yu[33], M. Yucel[7], E. Zaid[29], Y. Zeng[21,20,d], C. Zhang[29]

（缪子 $g-2$ 合作组）

[1]阿贡国家实验室，莱蒙特，伊利诺伊州，美国。[2]波士顿大学，波士顿，马萨诸塞州，美国。[3]布鲁克海文国家实验室，厄普顿，纽约州，美国。[4]布德克核物理研究所，新西伯利亚，俄罗斯。[5]轴子与精密物理中心（CAPP）/基础科学研究院（IBS）。[6]康奈尔大学，伊萨卡，纽约州，美国。[7]费米国家加速器实验室，巴达维亚，伊利诺伊州，美国。

[8]意大利国家核物理研究院弗拉斯卡蒂国家实验室，弗拉斯卡蒂，意大利。[9]意大利国家核物理研究院那不勒斯分部，那不勒斯，意大利。[10]意大利国家核物理研究院比萨分部，比萨，意大利。[11]意大利国家核物理研究院罗马第二大学分部，罗马，意大利。[12]意大利国家核物理研究院的里雅斯特分部，的里雅斯特，意大利。[13]詹姆斯·麦迪逊大学物理与天文学系，哈里森堡，弗吉尼亚州，美国。[14]物理研究所和PRISMA卓越集群++，美因茨约翰内斯·古腾堡大学，美因茨，德国。[15]韩国科学技术院（KAIST）。[16]密歇根州立大学，东兰辛，密歇根州，美国。[17]北中央学院，内珀维尔，伊利诺伊州，美国。[18]北伊利诺伊大学，迪卡尔布，伊利诺伊州，美国。[19]里吉斯大学，丹佛，科罗拉多州，美国。[20]上海交通大学物理与天文学学院，上海，中国。[21]上海交通大学李政道研究所，上海，中国。[22] 三一大学物理与天文学系，圣安东尼奥，得克萨斯州，美国。[23] 核与粒子物理研究所，德累斯顿工业大学，德累斯顿，德国。[24]莫利塞大学，坎波巴索，意大利。[25]乌迪内大学，乌迪内，意大利。[26]伦敦大学学院物理与天文学系，伦敦，英国。[27]伊利诺伊大学厄巴纳-香槟分校，厄巴纳，伊利诺伊州，美国。[28]肯塔基大学，列克星敦，肯塔基州，美国。[29]利物浦大学，利物浦，英国。[30]曼彻斯特大学物理与天文学系，曼彻斯特，英国。[31]马萨诸塞大学物理系，阿默斯特，马萨诸塞州，美国。[32]密歇根大学，安娜堡，密歇根州，美国。[33]密西西比大学，大学城，密西西比州，美国。[34]弗吉尼亚大学，夏洛茨维尔，弗吉尼亚州，美国。[35]华盛顿大学，西雅图，华盛顿州，美国。[36]纽约城市大学约克学院，牙买加，纽约州，美国。

摘要

本文详细介绍了费米实验室缪子 g–2 实验的计算策略。该实验通过处理超过 10 拍字节的探测器和模拟数据，以 127 ppb 的精度测量了缪子的反常磁矩。本文探讨了处理此类大型数据集时，在数据生产、数据处理和数据分析方面遇到的计算与运行挑战。通过采用费米实验室开发的软件基础设施和生产工具、高效利用资源，以及改进由合作组开发的重建工作流，这些挑战得以克服。在实验寿命期间，克服这些挑战使实验的数据吞吐能力得到提升，分析周转时间缩短了一个数量级。

---

[a]现任职于美国伊利诺伊州巴达维亚的费米国家加速器实验室。[b]另任职于意大利的里雅斯特大学。[c]另任职于意大利乌迪内的意大利国家核物理研究院乌迪内附属组的里雅斯特分部。[d]另任职于中国上海的暗物质物理国家重点实验室；另任职于中国上海的粒子天体物理与宇宙学重点实验室（教育部）；另任职于中国上海市粒子物理与宇宙学重点实验室。[e]另任职于意大利那不勒斯大学。[f]另任职于俄罗斯莫斯科的列别捷夫物理研究所；另任职于俄罗斯莫斯科的国立核能研究大学MEPhI。[g]另任职于意大利比萨大学。[h]另任职于意大利比萨的国家光学研究所—国家研究委员会。[i]现任职于美国马萨诸塞州阿默斯特的阿默斯特学院。

[j]现任美国得克萨斯州圣安东尼奥三一大学物理与天文学系。[k]现任印度班加罗尔联盟大学。[l]亦任职于意大利国家核物理研究院罗马第二大学分部，罗马，意大利。[m]现任意大利国家核物理研究院弗拉斯卡蒂国家实验室，弗拉斯卡蒂，意大利。[n]现任美国华盛顿州西雅图华盛顿大学。[o]现任伊斯坦布尔耶迪特佩大学，伊斯坦布尔，T¨urkiye。[p]亦任职于克罗地亚里耶卡大学，里耶卡，克罗地亚。[q]现任马来西亚雪邦厦门大学马来西亚分校物理系，雪兰莪州，马来西亚。[r]亦任职于中国浙江省杭州之江实验室。[s]亦任职于中国广东省深圳技术大学，深圳，广东省，中国。[t]亦任职于俄罗斯新西伯利亚国立大学，新西伯利亚，俄罗斯。[u]亦任职于意大利比萨高等师范学校，比萨，意大利。[v]亦任职于美国马萨诸塞州波士顿东北大学。[w]现任美国伊利诺伊州埃文斯顿西北大学。[x]现任美国弗吉尼亚州布莱克斯堡弗吉尼亚理工学院暨州立大学。[y]现任德国多特蒙德多特蒙德工业大学。[z]现任美国纽约州伊萨卡康奈尔大学。[aa]现任美国纽约州石溪创新解决方案研发有限责任公司。[ab]亦任职于意大利罗马罗马第二大学。[ac]亦任职于美国夏威夷州檀香山夏威夷大学马诺阿分校。[ad]亦任职于意大利比萨意大利国家核物理研究院比萨分部。[ae]现任美国马萨诸塞州波士顿波士顿大学。

[†] 已故。

1 引言

缪子 $g-2$ 实验在费米实验室开展了一项为期多年的粒子物理学研究计划[1, 2]。在束流数据采集期间，缪子 $g-2$ 实验每天积累约 15,000 个文件，每个文件大小为 2 GB。在为期六年的数据采集过程中，我们收集了约 10PB 的数据，相当于约 10^7 个文件。此外，物理模拟产生了约 2PB 的数据。生产的职责是对所有采集到的数据进行物理重建，并筛选出质量良好的数据。重建过程必须具有高度可重复性，并能够运行多种重建算法，以便开展多项独立分析，用于交叉检查和系统不确定度估算。每个文件至少需要处理两次：首先，预生产阶段确定依赖于数据的校准参数；随后，完整生产阶段执行重建。在整个处理流程中，持续评估数据质量并验证重建性能。鉴于所有数据都必须经过重建和分析，重建过程的速度和效率至关重要，也构成了一项重大的计算挑战。

本文介绍了该合作组如何成功完成这项计算任务。第2 节将概述缪子 $g-2$计算及所采用的整体计算策略。第3 节将介绍费米实验室提供、供多个实验使用的硬件和软件框架。第4、5 和6 节将介绍如何利用这些框架工具实现缪子 $g-2$ 产生过程中的重建、模拟和分析目标。最后，第7 节将讨论实验期间计算方面所做的改进。第8 节总结全文并分享经验教训。

2 缪子 $g-2$ 计算策略

缪子 $g-2$ 计算基础设施的主要目标是支持分析工作流，从已处理数据中提取物理结果。对量能器数据进行分析，以提取μ子磁异常；对径迹探测器数据进行分析，以了解可能影响结果的束流动力学效应。装置的描述 $g-2$ 可参见文献 [1] 及其中的参考文献。提取出的物理结果必须精确、稳健且可复现，同时还要高效利用资源（包括计算资源和人员）。该 $g-2$ 计算模型涵盖两个生产工作流：实验数据处理（第 3-5节）和蒙特卡罗（MC）模拟（第 6节）。

实验数据处理从探测器原始数据开始，并将其整理为对应于μ子填充或磁场测量时段的离散事例，每个事例都作为一个处理单元。数据采集后，原始数据首先存储在费米实验室的共享计算基础设施上，并归档至磁带存储。该文件传输系统必须能够可靠地应对意外删除文件、文件重复等故障。

随后，存储的数据经过多个生产阶段的处理，逐步将探测器原始波形转换为经过校准、可用于物理分析的数据产品（第 4节和第 5节）。处理首先通过多种独立的重建算法进行最小化重建，以提取校准常数；随后进行完整的探测器重建，并应用这些校准，最终生成分析工作流所用的高级数据产品。这需要一个事件处理框架和条件数据库，以便处理实验整个寿命期间校准常数的变化。

分析所追求的精度要求尽可能降低统计误差和系统误差。要降低统计误差，就必须分析每个探测到的正电子；要评估系统误差，则必须采用多种技术重复进行完整分析。为确保物理结果稳健，我们采用了四种彼此独立的量能器数据重建方法 [3–6]，将量能器模块的原始波形数据转换为沉积能量和时间。径迹探测器站的数据也进行了重建。为了确保这一过程可复现，我们

需要一个作业提交系统，能够向网格计算提交作业，并自动恢复失败的作业，以确保每次迭代都能成功处理所有数据。

缪子 $g-2$ 计算模型必须优化有限的计算资源和人力资源的使用。有限的人力资源大力推动了对通用基础设施服务的使用；这些服务由费米实验室计算专家开发，旨在供多个实验使用（第3节）。此外，数据处理被组织为一系列生产步骤，使实验所有成员都能方便地访问数据的不同阶段；全体合作组成员轮流承担“离线生产值班”，根据实验生产专家的指导管理数据生产的各个阶段。整个过程均有完整记录，并向所有合作者详细说明；代码也经过版本控制，以确保可复现性。这意味着值班人员使用的所有系统都必须便于用户操作。资源限制促使人们使用网格计算资源进行处理，并根据每年对总体计算需求的汇报，在多个实验之间高效共享生产和存储资源。

在实验数据采集期间，我们经历了六个不同的运行周期，大致每年一个。每个后续运行周期都包含更完善的实验条件、更多的事例、更精进的分析技术，以及额外的校准和灵敏度测量，从而持续降低统计误差和系统误差。在实验寿命期间，生产策略不断演变，以应对数据量增长一个数量级的情况，并维持及时的出版进度。为了充分利用数据量的增长，我们需要提高效率：这通过更高效地利用资源、加强生产活动的管理和组织、在整个合作组内就筛选后的数据集的定义达成一致（使后期数据集中更多的事例能够在线保存，同时所需磁盘空间大致保持不变），以及在最后一年的数据采集中近乎实时地运行数据预生产得以实现，从而使我们能够比以往运行批次早得多地启动完整生产。

3 计算基础设施

3.1 费米实验室计算环境与服务

费米实验室的缪子 $g-2$ 实验需要一个稳健且可扩展的计算基础设施，以高效处理大量探测器数据和模拟数据。该基础设施支持完整的数据生命周期，从原始数据采集和归档存储，到校准常数提取、重建、验证、筛选和最终分析。

缪子 g–2 实验的计算模型基于费米实验室开发的基础设施和生产工具构建，其中包括 FabrIc for Frontier Experiments (FIFE) [7]。FIFE 提供了一个通用服务层，供多个费米实验室实验共用，从而减少重复工作，并为存储、数据处理、网格提交、监控和生产管理提供一致的接口。

基础设施的最底层由 dCache [8] 磁盘池、长期磁带存储和数据库服务组成。本节中的池指由磁盘支持的暂存区或缓存区，而非处理算法。**GM2 写入池**会暂存新采集的 MIDAS 原始数据文件，这些文件以 `.mid` 为扩展名，之后再归档到磁带。**GM2 读取池**用于暂存从磁带调回的原始文件，以供大规模生产使用；而**通用读写池**则存放重建和模拟输出，供后续处理和分析使用。条件、校准、模拟和数据质量信息存储在数据库服务中，并在重建和验证期间供系统访问。

缪子 $g-2$ 实验采用了分布式计算模型，以提供所需的处理能力。该模型利用 FermiGrid、开放科学网格（OSG） [9,10]，和美国国家能源研究科学计算中心（NERSC）的资源，能够根据生产和模拟的需求动态分配 CPU 时隙。工作流的可扩展性源于缪子 $g-2$ 离线作业的并行特性：每项作业处理一组独立的输入文件，通常

使用单个 CPU 核心，内存占用低于每个时隙的限制，并生成相应的一组输出文件。这种结构使数千项作业能够并发运行，前提是同时满足存储和输出吞吐量要求。

在这些底层基础设施服务之上，实验使用 SAM（通过元数据顺序访问） [11] 来定义数据集、跟踪文件位置，并提供元数据驱动的文件访问。强度前沿数据处理（IFDH）工具 [12]将输入文件暂存到工作节点，而文件传输服务（FTS）则根据配置的规则编目并归档输出文件。生产活动通过 FIFE 作业提交工具和 POMS 工具提交、监控并恢复 [13]。

工作节点上运行的数据处理层是 *art* 框架 [14]。在缪子 $g-2$ 工作流中，*art* 执行实验特定的重建、校准、验证和模拟模块，并将输出存储为 ROOT 文件，其中包含带类型的数据产品和来源信息。生成的生产文件通常称为 *art*-ROOT 文件，或在工作流示意图中称为`.art` 文件。ROOT 提供底层对象输入/输出功能，也是标准的轻量级分析环境。生产和筛选之后，选定的重建量会转换为独立的`.root`文件，其中包含树和直方图，更便于开展不需要完整 *art* 环境的高层物理分析。

图 1 概述了缪子 $g-2$计算基础设施。图2 通过展示费米实验室开发的服务如何被该实验使用，对其进行了补充。两幅图共同将实验的计算工作流阶段与底层存储系统、元数据服务和作业管理工具联系起来。

3.1.1 数据处理框架

采用 *art* 数据处理框架 [14] 以逐事件方式处理离线数据和模拟数据；对于大多数 $g-2$ 数据，一个事件对应一次 µ 子填充。*art* 框架由库和可在运行时配置的主程序组成，全部使用 C++ 语言编写。框架与用户代码之间有明确的界限。前者由专门从事粒子物理学基础设施软件开发的费米实验室软件工程师开发和维护。后者由缪子 $g-2$ 物理学家开发，用于定义分析模块，以协调一致的工作流处理实验数据和模拟数据。

除了事例处理能力之外，使用 *art* 框架还通过“数据产品”这一概念，促进了系统化的数据管理。在 *art* 术语中，数据产品是由重建、校准、验证或模拟模块生成的输出对象，例如拟合脉冲、重建的量能器簇，或径迹探测器中的重建径迹。这些对象与用于标识生成模块和处理过程的元数据一同存储，从而保留每个处理阶段的溯源信息。这种结构使物理学家能够追踪重建量的生成过程，包括每一步所应用的变换、校准和筛选。此类可追溯性对于验证数据完整性、比较独立的重建方法，以及提高从这些事例中获得的物理测量的准确性至关重要。

3.1.2 数据库应用

缪子 $g-2$ 实验使用了一套费米实验室条件数据库系统，该系统最初为 MINERvA [17], 开发，用于管理数据校准和模拟常数。 $g-2$ 条件数据库按照模式组织常数，这些模式规定了常数的数据类型和用途。常数被组织到各个文件夹中，每个文件夹用于特定的校准或配置目的，例如探测器增益常数、计时偏移量或磁场谐波。在每个文件夹中，数据按时间分段存储为条目，每个条目都标记有“有效时间区间”（IOV），用于定义常数适用的时间范围或运行周期。尽管各个 IOV 的数值有所不同，但底层数据库模式——即用于描述常数的结构和字段定义——在所有 IOV 中都保持一致。这种统一的模式确保同一软件接口无需修改，即可访问和解读不同有效区间的数据。此外，常数还可以

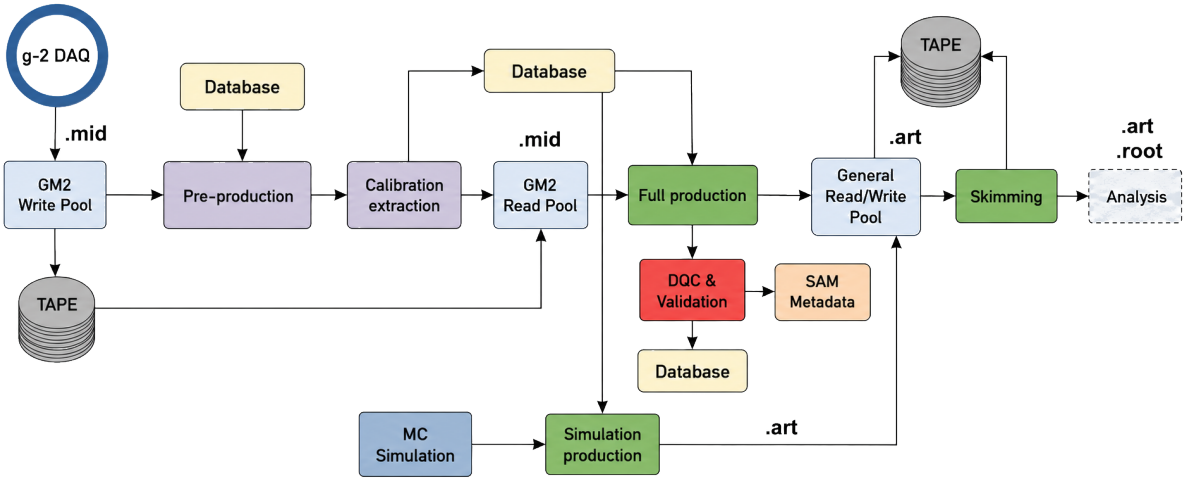

图 1 缪子 g–2 计算基础设施概览。在实验数据路径（上方分支）中，数据采集系统将原始 `.mid` 文件写入 GM2 写入池这一临时磁盘缓冲区，并将文件归档到磁带。预生产从数据库中获取配置和校准信息，并执行初始处理流程，为校准常数提取做准备。提取出的常数写入中央条件数据库。在完整生产中，从磁带调回原始文件并写入 GM2 读取池，然后使用该数据库中的常数进行重建。完整生产生成的 `.art` 文件写入通用读写池这一共享磁盘暂存区，并归档到磁带。同时，数据质量控制与验证将质量结果存储在专用数据库中，并更新 SAM 元数据目录。筛选生成用于分析的精简 `.art` 或 `.root` 数据集，筛选后的输出文件根据需要归档。在蒙特卡罗路径（下方分支）中，蒙特卡罗模拟向模拟生产提供生成事件；模拟生产使用中央数据库中的条件，并将 `.art` 文件写入通用读写池；此后，模拟样本和实验样本共享相同的存储、筛选和分析路径。蓝色框表示数据源和磁盘存储池；淡紫色框表示预处理和校准；绿色框表示生产或数据精简阶段；红色表示数据质量验证；黄色表示数据库服务；橙色表示 SAM 元数据目录；灰色圆柱表示长期磁带存储；白色虚线框表示用户分析阶段。 `.mid` 扩展名标识原始 MIDAS 文件， `.art` 表示包含 *art* 数据产品和溯源信息的 *art*-ROOT生产文件， `.root` 表示 ROOT 分析文件。

添加标签以便进一步分类。常数按照“通道”进行系统排列，每个通道由一个从 1 开始的长整数标识，代表一个特定的探测器元件，例如量热器晶体。对于适用于多个通道的全局值，则使用单独的单通道文件夹，以优化数据库存储。

读取条件数据库时，定制工具会验证有效时间区间，从而避免检索到错误的常数。生产工作流在每个子运行开始时，使用 *art* 中的“数据产品生成服务”从数据库读取常数。随后，这些常数被转入一个数据产品，供分析模块使用，从而绕过直接访问数据库。这种方法充当了内部缓存机制，确保同一有效时间区间的常数不会被重复检索，并通过减少重复的常数数据库查询，提高了安全性和效率。

3.2 数据处理与管理

数据处理是连接文件位置、元数据、传输服务和网格计算执行的层。来自缪子 $g-2$ 数据采集系统的原始数据最初存储在 **GM2 写入池** 中，这是一个专用磁盘存储区，容量约为 60 TB，旨在快速导入数据。在数据转移到长期磁带存储之前，这种临时存储为早期处理步骤提供了快速访问。

经过初始存储和归档后，数据进入预生产处理阶段，在此期间提取了校准常数。校准完成后，数据被暂存或移至专用存储区**GM2 读取池**，该存储区经过优化，可将原始文件传输至全生产任务。随后，通过数据质量控制（DQC）流程验证已处理数据，并将其记录在实验数据库中。

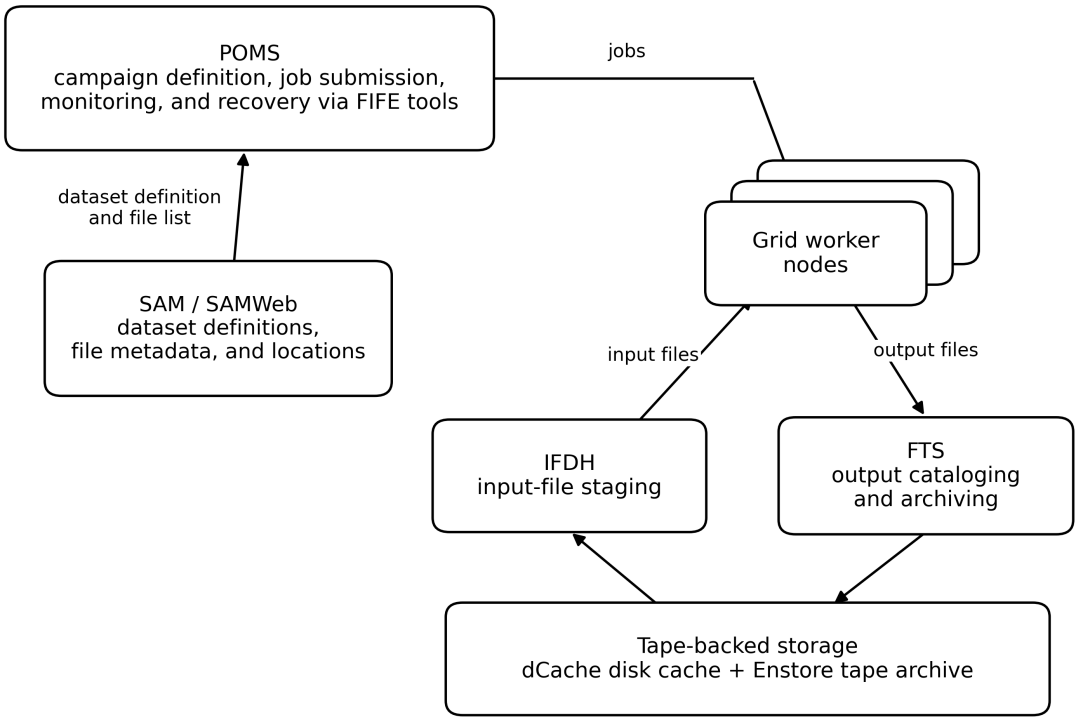

图 2 费米实验室开发的基础设施示意图，以及该基础设施在缪子 $g-2$ 实验中的应用。缪子 $g-2$ 合作组在 POMS 中配置计算活动，SAM 则提供数据集定义、文件元数据和文件位置。POMS 使用 FIFE 作业提交工具，在网格工作节点上提交、调度和监控作业。IFDH 将输入文件从磁带存储分阶段传输到工作节点。输出文件交由 FTS 编目，并通过 dCache [8] 和 Enstore [15] 存储系统进行归档，同时在 SAM 中登记相应的元数据。请注意，费米实验室于 [16] 2025年5月将 Enstore 系统替换为 CERN 磁带档案系统（CTA）。

生产成功完成后，数据可在**通用读写池**中访问。这是费米实验室大多数实验共用的磁盘存储区。在此阶段，还会进行筛选等额外处理，以提取相关事件子集。最终的数据产品以`.art`和`.root`格式存储，随后用于物理分析。输出文件被整理在费米实验室*pnfs*存储命名空间中，该命名空间提供了一个类文件系统接口，用于访问由实验集中管理的磁盘和磁带存储区。实际操作中，生产工作流将每个*pnfs*目录中的文件数量限制在约 5000 个。这是操作层面的限制，而非数据完整性限制：超过此数量可能导致目录列表和元数据操作变慢或不可靠，影响生产工作流，因此文件会分散存储在多个目录中。

缪子 $g-2$ 实验生成了约 9 PB 的原始数据，重建数据的大小也大致相同。此外，模拟数据约占 2–3 PB。鉴于数据量庞大，且很大一部分存储在磁带上，因此需要高效的数据处理和管理机制。SAM 提供了元数据目录，将数据集定义与磁盘和磁带存储中的物理文件位置关联起来。用户和生产作业通过 SAMWeb 接口查询元数据、定义数据集并获取文件传送信息，无需了解每个文件的物理存储位置。

SAM 最初由费米实验室开发，并由 D0 和 CDF 实验使用 [18, 19]。之后，它经过现代化改造，采用轻量级接口 [11]。SAM 使用实验专用元数据为文件位置建立索引并进行管理，从而支持通过元数据驱动的查询访问文件，并允许定义数据集。在生产过程中，多个工作节点可以连接到 SAM 项目；该项目由一组符合特定元数据查询条件的文件组成，SAM 通过 URL 链接提供访问。使用 IFDH 工具 [12], ，工作节点通过

这些 URL 链接获取文件。这种设置通过抽象数据获取和传输协议的复杂性，简化了外部网格上的作业执行，确保能够无缝访问所需数据，而无需详细了解其来源或所涉及的传输协议。

3.2.1 数据检索与预暂存

SAM 框架能够高效地将文件分发到多个工作节点。然而，访问存储在磁带上的大量数据会显著减慢数据检索速度。为缓解这一问题，传统的预暂存会在网格作业需要使用文件之前，将文件从磁带存储传输到磁盘存储，通常以数据集级别为单位。当数据集中的文件分散在许多磁带上，或每盘磁带只需少量文件时，这种方法效率低下。频繁挂载和卸载磁带会造成大量开销，因为每次访问都需要装载磁带、定位、读取，然后卸载。频繁的机械操作不仅会延长数据检索过程，还会加剧磁带硬件的磨损，从而进一步降低小规模、离散数据访问操作的效率。鉴于费米实验室的磁带存储系统由多个用户共享且吞吐量有限，因此必须谨慎规划预暂存，以支持顺畅的离线数据处理。

在最后几次生产活动中，缪子$g-2$ 实验主要使用**迁移模式** 进行预暂存。与传统的按数据集进行预暂存不同，迁移模式在磁带级别运行，将选定磁带上的所有文件传输到磁盘。当选定磁带上的大多数文件都需要时，这种方式可减少重复的磁带操作并提高吞吐量。其代价是磁盘空间利用效率较低：如果只需要磁带上的少量文件，迁移模式仍可能暂存当前并不需要的文件。不过，对于大型生产活动，这种取舍是有利的，因为它能最大限度地减少磁带挂载，并提高整体数据处理速率。

图 3 展示了 Run-5 全量生产数据集映射到磁带卷上的情况。这说明了数据集级预暂存为何效率低下：单个数据集可能分布在多盘磁带上，而单盘磁带也可能包含来自多个数据集的文件。这种结构促使人们采用迁移模式，用于相关磁带上的大部分内容最终都需要使用的活动。

3.2.2 文件传输

FTS专为磁带系统与磁盘位置之间的文件传输而设计。它会监控选定目录（通常是在作业提交期间通过 POMS 工具 [13] 指定为输出目录的目录）中是否有新文件。随后，它会根据配置文件中预先设定的规则，对发现的文件进行编目并存档到指定位置。此外，FTS 支持可提取实验专用元数据的自定义插件。这些元数据随后用于在 SAM 系统中注册相关文件。

3.2.3 自动化与工作流优化

缪子 $g-2$ 离线处理和模拟处理主要使用 POMS 工具实现了自动化。POMS 提供网络服务，可根据用户请求自动向分布式资源提交作业。此外，它还支持监控、失败提交的恢复、调试和记录保存。POMS 支持执行多个相互依赖的阶段：只有在初级阶段完成后，次级阶段才会启动，并将初级阶段的输出作为其输入。这一功能已在 $g-2$ 实验中得到广泛应用，实现了数据处理各阶段之间的无缝衔接。这种自动化显著减少了人工干预，使生产团队能够专注于高级优化，而不是执行例行任务。

此外，专门的 POMS 用户配置文件“production-shifter”使非专家合作者能够协助生产团队。该配置文件授予其访问部分命令的权限，使这些合作者能够根据需要监控并启动处理批次。这一改进简化了工作流，提高了$g-2$ 实验中数据处理的效率。

4 离线生产工作流

实验数据的生产分为两个主要阶段。第一阶段

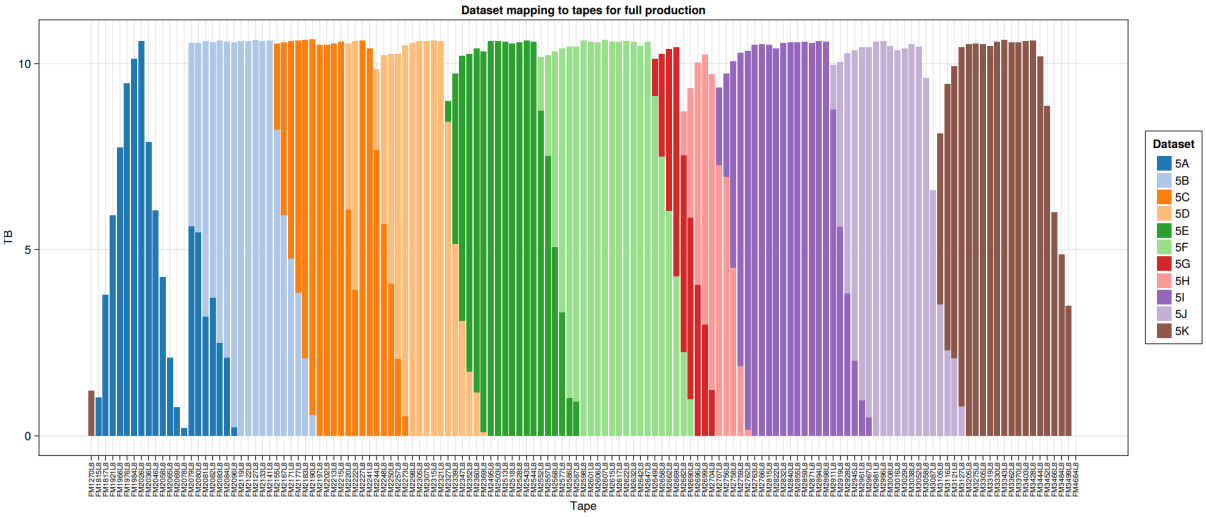

图 3 Run-5 完整生产数据集映射到磁带卷的示例。每个竖条代表一个磁带卷，堆叠的颜色表示该磁带上存储的各数据集的数据量。该图说明了为何按数据集进行预暂存可能效率低下：单个数据集可能分布在多盘磁带上，而单盘磁带也可能包含来自多个数据集的文件。当大部分内容都需要时，迁移模式通过暂存完整的所选磁带卷，减少重复的磁带挂载。</陪

（“预生产”）准备数据集，并对量热器数据进行最简重建，以提取校准常数。第二阶段（“生产”）通过对量热器和径迹探测器数据进行完整重建，并应用第一阶段确定的校准常数，生成数据集。两个阶段都从原始数据开始，并纳入数据质量控制（DQC）测试，以确保数据符合正常运行条件。生产工作流分为 7 个步骤，如图4所示：预生产、校准常数提取和填充期间 *DQC*、数据库上传、测试生产、测试 *DQC* 和验证、完整生产和子运行数据质量控制。为最大限度地利用资源并缩短总体生产时间，每个步骤在任意时刻至少处理一个数据集。每个步骤都为每个输入文件生成一个输出文件，以便轻松识别并重新处理失败的文件。

4.1 准备数据集

准备数据集的第一步是解压原始二进制文件，并对量能器数据进行最基本的重建。在此阶段，仅进行了信号脉冲拟合和能量校准。脉冲拟合所用的波形模板以及能量校准常数此前已在每个运行周期开始时，通过针对激光和束流数据的专用工作流提取。

第二步是分析重建数据，以提取增益校准常数并确定各次 μ子填充的质量。如果所有子系统的数据都处于标称范围内，则该次 μ子填充可用于分析。这一步高度自动化，并分为彼此独立的任务：每个子系统对应一个任务。这些任务并行运行，生成包含树和直方图的 JSON 文件或 ROOT 文件。校准专家使用这些文件核验结果，然后将结果上传到数据库。

这两个步骤均通过 POMS 运行；第一步完成后，POMS 会自动执行第二步。每次网格提交都会生成由 $O(1000)$ 个文件组成的数据集切片。根据网格节点的可用情况，每隔 4 至 6 小时进行一次网格提交。对于第一次提交中未能完成的文件，会进行最后一次恢复处理轮次。POMS 会自动处理这次重新提交。

这一准备阶段的最后一步，是确认提取出的校准常数正确无误，并将其上传到条件数据库。校准专家手动运行了验证程序并执行了数据库上传。

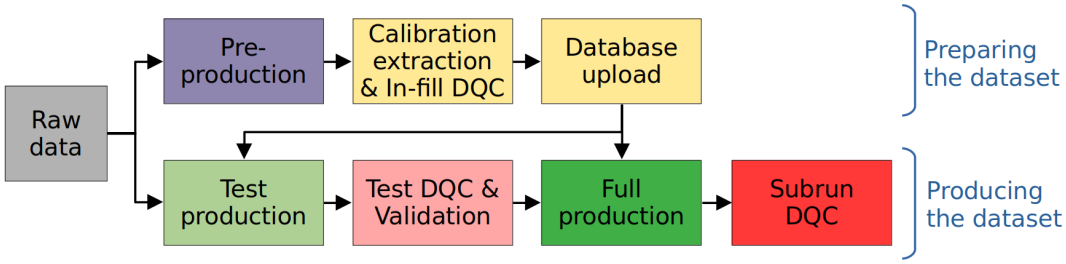

图 4 缪子 $g-2$ 实验的离线生产工作流。

校准常数和质量标志存储在数据表和状态表中。前者包含常数本身，后者则包含元数据和激活标志。最后，每张表都标注了一个有效区间（IOV），说明这些常数是从哪些运行批次中提取的。表格首先上传到开发数据库；验证确认上传过程中没有出错后，再将其克隆到生产数据库。最后，对于未能通过预生产流程处理或被发现不含束流数据的任何原始文件，都会通过更新其 SAM 元数据将其标记为此类文件，并且不会进入完整生产流程。

4.2 生成数据集

提取校准常数并确定高质量文件后，数据集即可进行完整生产。该阶段再次从原始数据开始，对所有探测器执行完整重建，并应用在数据准备阶段提取的校准常数。完整生产涉及量能器数据、径迹探测器数据和辅助探测器的重建。对同一份量能器数据执行了四种不同的重建。

在开始生产整个数据集之前，先对其中一个子集进行了测试生产。首先处理了均匀分布在整个数据集中的 ~10% 文件，以检测潜在的重建问题。随后对已生产的文件执行 DQC 步骤，以确保没有需要进一步调查的错误或异常。这一验证步骤在识别软件错误和意外的实验条件方面非常有效，也避免了浪费网格计算资源去处理最终会被丢弃或重新处理的数据。

DQC 步骤由值班人员执行。值班人员检查了两份 PDF 手册文件：一份包含增益验证直方图，另一份包含重建验证直方图，以及每个文件中重建正电子数量等量的时间序列图。值班人员将 PDF 手册文件与参考手册进行比对。增益手册包含每个量能器晶体的基于激光的增益校准信息、每个量能器在增益校正前后的正电子能谱，以及整个数据集的平均正电子能量。重建手册将不同重建得到的正电子能量谱和时间谱绘制在一起，以确保各重建之间的一致性。值班人员检查时间序列图，以确保各量保持稳定。图 5 中给出了一个时间序列图示例。

测试 生产中的任何重建异常或时间序列不稳定性解决后，便在数据集剩余部分恢复完整生产。完整生产完成后，再次执行 DQC 步骤，并将结果上传至 Postgres 数据库。该数据库为每项 DQC 检查设有一张表，每行代表一个文件。最后，提取通过所有 DQC 类别检查的文件列表，并更新 SAM 中原始 MIDAS 文件的元数据标记。随后利用这些元数据标记定义交付给分析人员的最终数据集。

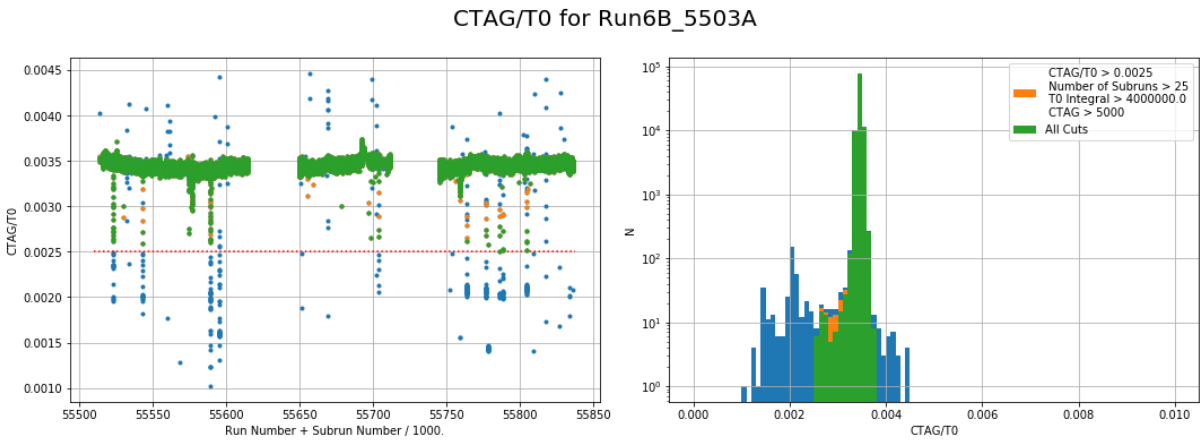

图5 DQC 步骤中生成的示例时间线，显示了归一化到μ子束流强度的重建正电子（“CTAGs”）数量，分别以直方图（右）和文件编号的函数（左）表示。左侧图中的红色虚线对应于文件通过此项检查所需的最小值。蓝色点表示未通过此项检查的文件，橙色点表示通过此项检查但未通过其他检查的文件，绿色点表示通过 DQC 步骤中所有检查的文件。只有绿色点所代表的文件才会提供给分析人员进行分析。

5 分析计算工作流

第4 节所述数据生产的输出是*art-root* 文件，其中包含以*art* 数据产品形式呈现的重建数据。虽然这种文件格式便于保留重建步骤的历史记录及相关质量筛选信息，但也给分析人员带来了一些挑战：

•大量文件：由于每个数据文件都是单独处理的，从 Run-1 到 Run-6，通过所有数据质量检查的文件总数约为360万。分析人员需要分析每一个文件，才能实现实验的精度目标。如此庞大的文件数量带来了挑战。例如，如果分析人员的网格作业失败，他们需要确定哪些文件需要重新处理。

• 数据集大小非常大：每个输出文件都包含量热器和径迹探测器数据的完整重建链，以及辅助探测器数据。完整的重建链使分析人员能够获取部分重建的事例，以开展系统学研究。然而，这使得 Run-1 至 Run-6 的重建数据总大小达到近 9 拍字节。

• 可用性：重建数据的大小过大，无法始终将其暂存在磁盘上。分析团队需要频繁、按需访问数据，以便迭代其分析算法并开展系统学研究。Run-1（0.3 拍字节）和 Run-2/3（1.7 拍字节）重建数据的大小足够小，因此可以暂存在费米实验室通用缓存池中。然而，Run-4 至 Run-6 的分析（6.6 拍字节）则无法采用这种方式。

• 可移植性：在*art-root* 文件上进行任何分析都需要*art* 框架。它功能强大，但学习曲线陡峭。许多分析人员更愿意使用标准ROOT 文件处理数据，因为这类文件可以在个人电脑上轻松处理和检查。

为了解决这些问题，大多数分析人员会将感兴趣的信息提取到一组ROOT 文件中，其中包含树和直方图，并基于这些文件开展分析。随着数据集从 Run-1 增长到 Run-6，制作团队逐渐负责为分析人员创建这些精简数据的“分析文件”。

5.1 art 筛选

虽然ROOT 文件被证明是分析人员的高效解决方案，但有时仍需访问完整的离线文件，以便重新运行完整的分析

程序或执行新的分析。如前所述，Run-4/5/6 数据集过于庞大，无法存储在磁盘上并随时访问。解决方案是创建精简版的art文件，称为“art-skims”，方法是移除典型分析中不需要的所有数据产品。

为此，只保留未经校准和完全校准的量能器击中，从而将大小缩减了五倍。若舍弃未经校准的击中，还能进一步缩减大小，但保留这些击中可确保能够应用自定义增益校准来开展系统学研究。因此，存储 Run-4/5/6 所需的磁盘缓存从 6.6 拍字节降至 1.3 拍字节。

为减少文件总数，精简后的 $art\text{-}skims$ 也经过了合并处理。每个最终文件包含 10 个属于同一运行编号的随机文件。我们将合并的文件记录为元数据。

当分析人员需要永久缓存的数据集之外的生产数据时，他们会向生产经理提交请求，由后者按需预先暂存所需数据。根据可用资源情况，这些操作可能需要数小时到数天才能完成。

5.2 分析文件

实验接近尾声时，生产团队管理着五种不同类型的分析文件：两种用于径迹探测器数据，三种用于量能器数据。这些文件包含分析人员进行分析所需的 ROOT 树或 ROOT 直方图。对于所有这些文件，任务都通过 POMS 提交，输出也使用与离线生产相同的工具分配元数据。

生产始终通过为每个输入文件生成一个输出文件来进行。因此，需要额外执行一步合并输出文件，以减少分析人员需要处理的文件总数。由于分析文件始终是ROOT文件，因此使用内置的 `hadd` 工具将它们合并。

合并过程分为两到三步，每一步都会将文件数量减少几个数量级。合并程序通过 POMS 在 FermiGrid 上提交运行，或在本地计算机上运行。考虑到一个典型的数据集包含 100,000 个文件，第一步会在网格计算上将文件合并为每组 100 个。这样，任务就可以分配到 1000 个核心上并行运行。最后一步是将 1000 个中间文件合并为一个文件，这一步在核心数较多的服务器上本地完成。这样，单个 `hadd` 任务就可以以多线程方式运行。中间文件存储在暂存磁盘上，最终结果则存储在磁带或持久化磁盘上。

5.3 分析人员的计算资源

缪子 $g-2$ 数据的分析由分析人员在费米实验室拥有的多台服务器上，或参与机构本地拥有的计算资源上完成。费米实验室提供的合作组共享资源包括八台虚拟机和一台高性能工作站。每台虚拟机配备四个 CPU 线程和 12 GB 内存。每位用户均获得 400 GB 磁盘空间，用于存储个人数据。该工作站配备 48 个 CPU 线程、187 GB 内存和 7 TB 本地磁盘空间。正如第5.2节所述，该工作站用于在生产期间运行本地文件合并作业。这些资源的特点是拥有大量 CPU 线程、较大的内存，以及由大容量磁盘和网络附加存储（NAS）服务器组合而成的数据存储空间。

CPU 密集型分析作业在 FermiGrid 和 OSG 上运行，使用与缪子 $g-2$ 生产相同的共享资源。除为生产预留的网格配额外，始终还留有一定的资源余量供分析人员运行作业。提交脚本与正式生产所用的相同，并提供故障恢复例程。它们通常通过虚拟机手动执行，但少数情况下会使用 POMS 来运行分析作业。

机构自有的 NAS 服务器在存放用于分析的筛选后数据集方面非常实用，但当 CPU 密集型作业需要处理存储在这些服务器上的数据时，它们就会成为瓶颈。网格节点只能通过输入/输出访问费米实验室的中央磁盘缓存，因此，网格作业所需的任何 NAS 数据都必须先临时复制到相应节点上的临时缓存文件夹。由于 NAS 服务器不在速度更快的 FermiGrid 网络内，复制速度最高仅为 10 Gbps；网络繁忙时，实际速度可能低至 100 MB/s。典型的分析数据集大小约为 10 TB，因此所需的复制过程可能需要几天。

本节介绍的工作流方案使所有分析人员都能高效开展分析，尽管他们的需求各不相同。他们使用数量较少的文件进行分析，并可随时访问数据，无需预先从磁带调取数据。他们还可以使用生产网格提交脚本来满足自身的数据处理需求，并使用专用计算机分析数据。这些流程最初是为 a_μ 分析开发的，但目前也用于该实验的电偶极矩、CPT 洛伦兹破缺和暗物质分析。

6 模拟

缪子 $g-2$ 实验使用多个模拟软件包，对储存环中的μ子束进行了模拟。这些模拟用于优化实验设计，并了解束流动力学对最终分析的影响。计算量最大的模拟涉及将μ子注入储存区域，并在单次持续 700μs 的填充过程中，让它们绕储存环运行 4700 圈，径迹总长超过 200 km。要在模拟中准确跟踪单个μ子走过这么长的距离，对跟踪工具提出了很高要求，也需要使用辛积分器，在整个路径中严格保持能量守恒。根据所用的模拟软件包，这些任务可能需要很长的 CPU 运行时间，导致每个在整个填充过程中留存的μ子都需要 ~500秒的运行时间。综合这些因素，无法生成或存储与实验期间 $\sim 10^{11}$ 采集的μ子衰变数据样本数量相当的模拟样本。相反，我们针对要解决的问题优化并部署了各种工具和技术。本节将简要介绍各类模拟软件包，说明每种工具的主要用途，并记录生成模拟样本所需的计算资源。关于μ子园区、储存环组件、模拟软件包及相关束流动力学效应的介绍，见首篇缪子 $g-2$ 束流动力学论文 [20]。

6.1 模拟软件包

在该实验中，质子撞击铁镍铬靶后产生的 π 介子发生衰变，由此生成 μ子。该过程使用两个软件包进行模拟：MARS[21][22] 和G4Beamline[23]。采用两套软件包进行模拟，既能检验模型实现，也能验证各程序所依据的物理原理。随后，两套软件包的模拟结果均与费米实验室测试束流数据进行了基准比对。从锂聚焦透镜下游出口射出的带正电 π 介子、μ子、正电子和质子被保存下来，以便进一步在传输环的 G4Beamline模型和末级聚焦的 BMAD[24] 模型中进行径迹追踪 [1]。这些程序组合的主要目标是生成高度极化的 μ^+ 分布，作为模拟 μ子注入和储存环内存储的软件包的输入。

储存环中存在的随时间变化的电磁场由三个软件包建模：BMAD、COSY INFINITY[25, 26] 和 GM2Ringsim [27][1]。后者基于 Geant4 [28–30]。此外，BMAD和GM2Ringsim 实现了注入模型，其中包括束流离开最终聚焦束线并进入储存环时的方向。

图 6 展示了缪子 $g-2$ 实验的模拟数据生成阶段，以及用于生成这些数据的程序。蓝色方块表示为各阶段提供支持的软件包，绿色箭头表示各阶段产生、随后输入到新阶段的缪子分布。COSY INFINITY 未包含注入模型

因此，输入缪子分布取自 BMAD 或 GM2Ringsim在注入器出口处的数据。进入储存环后，每个软件包都能跟踪 μ子在储存环各处特定时刻的位置、动量和自旋。GM2Ringsim 软件包还可以跟踪缪子衰变产物以及材料相互作用产生的粒子。尽管这三个软件包都使用相同的场图作为输入，但它们在缪子跟踪的执行方式上存在重要差异。例如，GM2Ringsim是唯一能够全面考虑材料效应的软件包，而 COSYINFINITY 则能够采用完全的辛积分。这些工具各自的优势互为补充，为提取出的物理可观测量提供了至关重要的一致性检验。

6.2 模拟生产与工作流

6.2.1 MARS靶站模拟

MARS 研究最初在 RHIC ATLAS计算设施运行，之后则在 Fer- miGrid 上运行。研究使用MARS软件包，在数天内生成了$\sim 10^9$ 打靶质子（POT）。请注意，这一数量比实验期间每次 μ子填充的典型 POT 少三个数量级；后续采用G4Beamline 实现，生成了$\sim 10^{12}$ POT 样本。MARS 的存储需求很低，1B POT 样本仅需 0.2 GB 磁盘空间。输出以与其他工具兼容的格式存储，可供其读取。

6.2.2 G4beamline靶站与粒子传输束线模拟

使用 G4beamline Docker 镜像在 NERSC 和 OSG 上运行。根据计算资源的可用情况，在 OSG 上模拟 $\sim 10^{12}$打靶质子数需要数天至一周。生成注入传输环位置处的 μ子和 π介子粒子分布，每 $\sim 10^{12}$ 打靶质子数约需 1.8 GB。

6.2.3 BMAD

BMAD 模拟在多个站点运行。对于传输环和束线模拟，BMAD在本地笔记本电脑或本地工作站（配备 16 核、32 线程处理器）上编译并运行。通常，3×10^{12} 打靶质子数的运行在 24 小时内完成，转换为 G4beamline的 BLTrackFile2格式后需要 $\sim$ 300 MB 的存储空间。

BMAD注入线和储存环模拟通常在康奈尔大学的 CLASSE计算网格上运行。在该网格上的典型运行时间约为 12 小时，输出大小约为 1 GB，并存储在本地一对 Tier-1专用磁盘上。模拟输出以文本文件形式存储，数据则通过自定义 Fortran 程序和 gnuplot 的某种组合进行处理。

6.2.4 COSY INFINITY

COSYINFINITY 生产任务在密歇根州立大学计算集群和 NERSC 的 Cori超级计算机上运行。束流动力学校正和磁场卷积每种实验设置约需 2 天：一天用于准备非线性传输映射，另一天用于束流跟踪。映射和束流分布存储约需 1 TB 磁盘空间。

6.2.5 GM2Ringsim

该GM2Ringsim 软件包通常以两种不同模式运行。第一种称为束流枪模式，会注入一组 μ 子，并在储存环中追踪其运动一段有限时间。粒子与探测器的相互作用，以及其位置、动量和自旋，都会记录在虚拟跟踪平面中。这些平面是布置在储存环内侧战略位置的虚拟探测器。虚拟跟踪平面利用了 GEANT4 [28–30], 的平行世界功能，可让用户提取粒子在特定位置的信息，而不受材料相互作用的影响。通常还会人为延长粒子循环运动的寿命，以提高效率，例如研究径向和垂直位置以及贝塔振荡随时间的变化。通常，每组

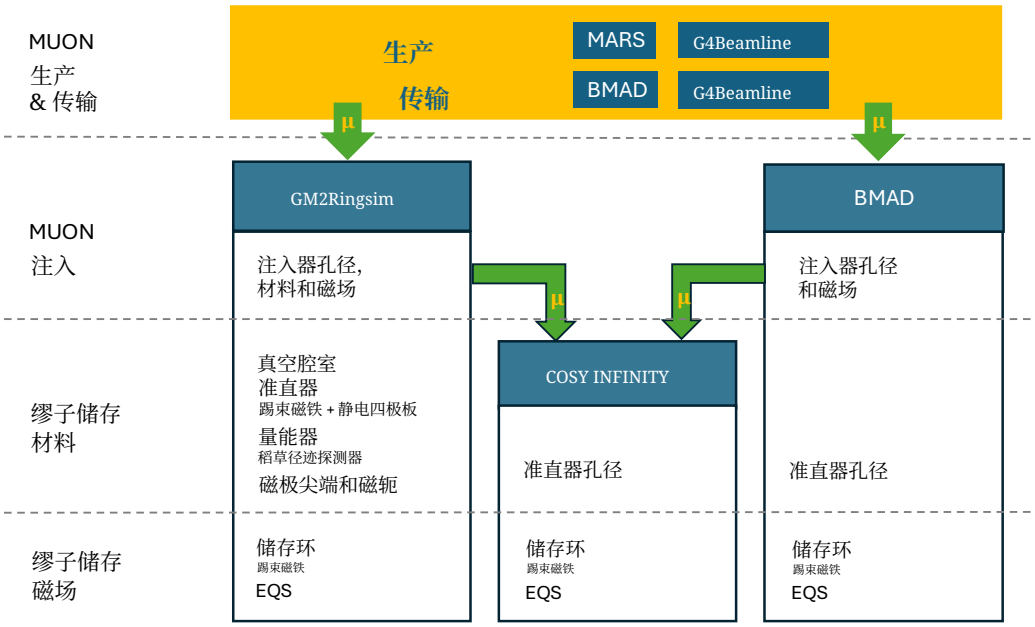

图6 该示意图展示了缪子 g–2 模拟基础设施以及第6节中讨论的软件包之间的关系。可以使用 `MARS` 或 `G4Beamline` 来模拟靶处的粒子产生，随后通过组合使用 `BMAD` 和 `G4Beamline`将粒子传输经过束线和传输环。这些软件包的输出是 M5束线末端的 μ子相空间分布，进而作为通过注入器注入储存环的输入，用于 `BMAD` 和 `GM2Ringsim` 储存环实现方式。 `COSY INFINITY` 软件包以在注入器末端计算得到的 μ子分布为输入，来自 `GM2Ringsim` 或 `BMAD`。

实验条件都会生成一个样本，其中包含 300 万个事件，且关闭了衰变。与其他模拟软件包不同，`GM2Ringsim`束枪模拟数据的生产紧随实验数据的生产进行。作业通过 POMS 提交至 FermiGrid 和 OSG。模拟输出以 art-ROOT 文件的形式存储。输出文件被传回费米实验室，然后通过 FTS 写入磁带。数据集记录在 SAM 数据集访问系统中。在缪子$g-2$实验的第二篇论文中（运行批次 $2+3$），为研究束流动力学修正并估算系统误差，共生产了 30 个模拟数据集。每个数据集耗时约三天完成，使用了费米实验室分配给 $g-2$ 模拟生产团队的 2000 个计算槽。每个数据集的存储输出约为 5 TB。模拟生产概要见表1。

第二种模式称为气枪模式，它完全跳过注入，转而使 μ子充满储存环；这些 μ子会立即衰变为正电子，随后正电子沿着.....进行径迹追踪电磁场，以及进入用于重建正电子能量和缪子衰变位置的量热器和稻草管径迹探测器。该模式能够高效地进行模拟生产，用于研究衰变正电子与探测器的相互作用，例如接受度效应以及它们与自旋和动量之间的任何相关性。由于正电子在离开储存环前行进的距离不到半圈，采用气枪模式时，追踪每个事例所需的时间会缩短几个数量级。不过，接受度研究所需的保真度远高于使用束流电子枪的束流动力学研究。为研究相位接受度校正及分析中的系统误差，我们生成了一个包含 10 B 个事例的气枪模拟样本。样本在 NERSC 上生成，选择该平台是因为它具有很高的吞吐量。作业从 FermiGrid 和 NERSC 两个平台启动。在一台拥有 68 个核心的 NERSC KNL 节点上，测试作业耗时 100 分钟，生成了 5.4 M 个事例。输出再次存储为 *art* 数据产品和 ROOT文件，总计 0.8 PB。整个过程花了数月

才完成全部 100 亿个事例。数据在生成期间暂时存储于 NERSC 的 Cori 系统上，之后通过专用的 FTS 协议传输至 FermiGrid。

7 时间线与改进

缪子 $g-2$实验在费米实验室于2018年至2023年采集了物理数据。数据采集结束时，生产流程已能够随着数据采集同步处理数据集。为了达到这一水平，我们多年来进行了许多改进，本节将对此进行概述。

Run-1 之后，离线重建起步缓慢，因为软件仍在最终完善中，同时数据也被用于改进流程。因此，Run-1 数据集被多次重新处理——仅 Run-1A 就在达到可供发表的形式之前被处理了 11 次。此外，离线生产工作流尚未针对高效处理大量数据进行优化。生产周转时间，即从束流数据采集结束到生产完成所经过的时间，如表2所示，为 1.5 年。

随后，从 Run-1 到 Run-2/3，已采集的统计量和数据大小增加了五倍；从 Run-2/3 到 Run-4/5/6，又额外增加了四倍。该实验在费米实验室计算部门的支持下，在 Run-1 之后消除了数据处理积压，并通过持续提高生产流程的效率、生产数据速率能力和生产管理水平，跟上了不断增长的数据量。

Run-1 重建完成并准备好进行分析后，研究人员为 Run-2 生产开发了一个多阶段生产工作流，详见第 4 节。与此同时，也实施了第 3 节所述的生产管理工具 POMS 和条件数据库。在数据库实施之前，Run-1 的校准常数存储在文本文件中。这些软件和工作流改造的实施耗时较长，进一步延迟了 Run-2 和 Run-3 数据集的生产。不过，新流程实现了对各项流程的进一步优化、并行化和自动化，从而使整体生产速率提高了一个数量级。

尽管 Run-4、Run-5 和 Run-6 的规模大得多，三者合计贡献了总统计量和数据大小的 ∼75%，其生产周期却持续缩短，到 Run-6 时最终缩短至 2.5 个月。图 7 展示了历年来的数据生产时间线。

最具影响力的改进包括：

- **计算资源**。
  + FermiGrid 中为 g–2 预留的槽位数从 2500 个增加到 5700 个，每个槽位对应 1 个 CPU线程和 2 GB 内存。这使我们能够并行运行更多作业。
  + 生产作业的内存占用减少到 2 GB 以下，因此每个作业只占用一个网格槽位。这需要优化生产代码以及 art 作业配置。
  + 每次网格提交处理的文件数从 5000 个增加到 20000 个，每个作业处理的文件数从 2 个增加到 8 个。通过缩短处理连续两批各含 5000 个文件的数据片段之间的间隔时间，这提高了网格占用率。事实上，作业处理任何特定文件所需的时间，很大程度上取决于存储的 μ 子数量以及作业最终运行所在的具体中央处理器型号。因此，有些作业可能提前结束，而另一些作业在处理任何一批 5000 个文件时，耗时可能是前者的两倍。增大数据片段的大小有助于充分利用速度更快的节点，并提高整体效率。
- **数据处理**。
  +条件数据库采用了针对多个并发请求优化的 HTTPS 接口，使数据校准和模拟常数的管理更加简便。
  + 第 3 节中介绍的“迁移模式”，将预暂存速度提高了 3 至 5 倍。通过加快从磁带到磁盘存储的数据传输，有效解决了离线生产中最大的瓶颈之一。
  +最优数据集大小确定为 100,000 个文件。这项调整利用了更大的 20,000 个文件处理批次，并减少了

运行周期	Gun	事件数（百万）	数据量（拍字节）
运行1	gas	10,800	0.8
运行1	beam	1330	2.21
第2轮运行	beam	1066	1.78
Run 3	beam	78	0.12
第4轮运行	beam	4609	0.69
第5/6轮运行	beam	122	0.17

表1 各运行时期模拟数据集大小表。请注意，第5轮运行和第6轮运行采用相同的实验配置，因此生成了一个共用的模拟数据集。

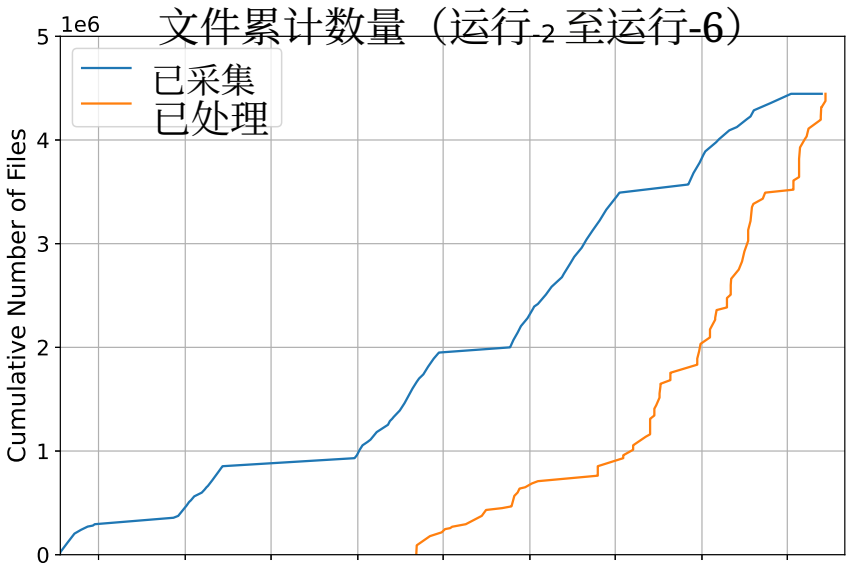

图7 从Run-2到Run-6期间，由数据采集系统（蓝色）已采集的原始数据文件累计数量，以及通过离线生产工作流处理并用于最终分析的文件累计数量（橙色）。曲线之间的垂直间距表示特定时间的处理积压，而实验结束时两条曲线趋于重合，表明生产已追上数据采集进度。

每个数据集都必须重复执行的步骤数量，例如恢复和验证例程。原则上，更大的数据集（例如总计 200k 个文件）效率会更高。然而，某些流程（如直方图合并）将超出可用的本地资源。

+ 在第6轮数据采集期间，预生产阶段自动运行。具体而言，数据采集系统产生的新原始文件会先存储在临时磁盘缓存池中，等待复制到磁带。 $g-2$ 数据采集速度约为 21 TB/天，缓存池容量为 114 TB，因此新获取的原始文件会在缓存中保存五天，然后为新传入的文件腾出空间。每天，自动作业都会处理 $\sim 10^4$原始文件。少数未能通过这一自动化生产步骤的文件（通常为 ~5%）会被暂存，并每周手动重新处理一次。这一自动化流程显著加快了预生产阶段，使 Run-6 的完整生产能够在数据采集结束后迅速完成。

- **管理和人力资源。**
  +完整生产步骤末尾的验证程序以及测试生产的引入，极大地帮助发现了校准错误和算法缺陷。这显著减少了为了达到最终生产而需要反复生产某个

数据集的次数。

+ 整个合作组都参与了2021年11月Run 3生产期间的离线生产值班。这些值班计入合作者的值班积分。为做好值班准备，需要记录完整的生产流程，这也促成了核心流程的正式化以及不必要步骤的清理。尽管这一过程耗时较长，但它极大地促进了后续的工作流优化。值班人员排班确保生产始终能够以最大速率运行，不受少数生产专家可用时间的限制。

+ 费米实验室提供了技术和资源支持。基础设施专家团队逐项解决了基础设施瓶颈，其中最突出的是原始数据缓存的可用性问题，该问题通过专用缓存池和新的预暂存模式得到解决。在需求高峰期，系统获得了计算优先级，以满足发表截止日期的要求。

得益于这些以及其他优化，正如表2所示，完整生产的总体速率在Run-2和Run-3生产之间提高了一倍，并在Run-3和Run-4生产之间再次提高了一倍。最后一行所示的稍慢的Run-6速率表明，生产速率足以跟上实验数据采集速率。

8 结论

我们介绍了开发和运行缪子 $g-2$ 实验计算基础设施的经验教训，这些经验或许能为未来采集拍字节规模数据的中型实验提供借鉴。

- **软件与计算**

  + 拍字节规模的数据可用性一直是缪子 $g-2$ 生产的主要瓶颈之一。为简化将文件从磁带暂存到磁盘的流程，我们投入了大量精力。我们建立了专用缓存池和基于磁带的暂存例程，并配合由合作组开发的数据暂存计划和监控工具。

  + 事实证明，维护多样且可扩展的计算资源至关重要。在负载较高时，将场外资源用作 OpenScienceGrid 网格节点非常有帮助。部分模拟任务在美国国家能源研究科学计算中心（NERSC）的高性能计算（HPC）设施上执行，径迹探测器重建任务则转移至英国的计算资源。

  +软件准备就绪至关重要，其中包括稳健的校准基础设施，这关系到能否及时开展数据处理。μ子 g-2 实验起初缺少用于管理大量校准常数的条件数据库；后来实施该数据库时，需要花费数月进行重构。尽管使用真实数据进行一些微调不可避免，但核心基础设施应在数据采集开始前就绪。

- **管理**

  + 软件验证对于避免重新处理和闲置时间至关重要。*art* 代码具有高度模块化的结构，适用于高能物理重建，使不同人员能够分别处理重建链中的各个部分。然而，早期开发人员之间缺乏沟通，也没有明确的代码审查流程，常常导致细微错误未被发现便进入生产阶段，直到分析阶段才被发现。在生产流程中增加验证步骤，开展物理合理性检查和校准交叉检查，提高了发现错误的效率，并大幅减少了所需的重新处理量。建立更完善的代码审查系统会大有裨益。

  +离线计算是合作组范围内的一项工作；满负荷运行生产时，需要 10 个全职当量才能维持稳定运行。引入离线班次有助于确保生产持续运行。随着工作流程不断优化，进一步并行化和自动化这一流程，将值班人员人数从 8 人减少到 3 人。

  + 文档、说明和簿记至关重要。有时，五年前的文件因数据损坏或类似问题而不得不从头重建。相同的生产流程必须能够以相同的方式轻松复现，并由不同的人执行。

数据集	N°个文件	大小 [TB]	生产时间 [天]	速率 [文件/天]	运行结束后的天数
Run-1	258108	544	572	451	565
Run-2	267152	563	129	2071	750
Run-3	538622	1066	122	4415	794
Run-4	992005	1953	111	8937	550
Run-5	1453853	2838	141	10311	305
Run-6	748724	1760	106	7063	74

表2 每个聚合数据集的生产周转时间。除 Run-1 外（当时尚未设立预生产阶段），这里只考虑了完整生产处理时间。

POMS 是簿记和生产管理的优秀工具。

+ 承办实验室在基础设施和专业知识方面提供的支持，对于实验计算模型的初期开发和持续运行至关重要。合作组与实验室之间设立计算联络员，证明这一角色十分重要。发生重大基础设施瓶颈或服务中断时，此人协助组建危机应对团队。

本文所述的计算模型、整个合作组的参与以及承办实验室的支持，是缪子 $g-2$ 实验取得成功的基础。缪子 $g-2$ 合作组利用约 20 拍字节的原始数据、重建数据和模拟数据，以及超过 1.5 亿 CPU 小时的分布式计算，将约 1500 亿次记录在案的缪子衰变转化为对 μ子反常磁矩最精确的测定。

致谢

我们感谢费米实验室的管理层和工作人员（包括许多计算专业人员）对本实验的大力支持，也感谢我们的大学和国家实验室的工程师、技术人员及工场给予的巨大支持。

缪子 $g-2$ 实验在费米国家加速器实验室进行。该实验室是美国能源部科学办公室的高能物理用户设施。费米实验室由 Fermi Forward Discovery Group, LLC 管理，该公司依据合同编号 89243024CSC000002 开展工作。实验还获得了美国能源部科学办公室通过以下渠道提供的额外支持：高能物理、核物理和先进科学计算研究办公室；美国—日本高能物理科学技术合作计划；美国国家科学基金会；意大利国家核物理研究所；英国科学技术设施理事会和英国皇家学会（资助编号 URF\R1\231503）；国家自然科学基金委员会（资助编号 12475108、12305217、12075151）；韩国科学、信息通信技术和未来规划部（MSIP）、韩国国家研究基金会（NRF）以及基础科学研究院（项目代码 IBS-R017-D1）；德国研究联合会（DFG）通过卓越集群 PRISMA+ （EXC 2118/1，项目编号 39083149）和卓越集群 PRISMA++（EXC 2118/2，项目编号 39083149）；欧盟地平线 2020 研究与创新计划，通过玛丽·斯克沃多夫斯卡-居里资助协议（编号 101006726 和 734303）；欧盟 STRONG 2020 项目（资助协议编号 824093）；以及 Leverhulme Trust（LIP-2021-014）。

参考文献

[1]Aguillard，D.P.，等：关于正 μ 子反常磁矩测量的详细报告，精度达 0.20 ppm。物理评论**D110**(3)，032009（2024）https://doi.org/10.1103/PhysRevD.110.032009arXiv:2402.15410 [高能物理实验]

[2] D. P. Aguillard 等（缪子 $g-2$ 合作组）：将正 μ 子反常磁矩测量至 127 ppb。《物理评论快报》**135**，101802（2025）https://doi.org/10.1103/7clf-sm2v

[3] Fienberg，A.：测量 e989 μ子 g-2 实验中的进动频率。博士论文，华盛顿大学（2019）[4] Sweigart，D.：正 μ 子反常进动频率的测量。博士论文，康奈尔大学（2020）[5] Girotti，P.：在费米实验室的 μ子 g-2 实验中测量 μ子反常进动频率。博士论文，比萨大学（2023年）[6] Kelton，L.：用于测量费米实验室 μ子 g-2 实验反常进动频率的能量积分分析。博士论文，肯塔基大学（2022）[7]Herner，K.R.，等：费米实验室的 FIFE 项目：实验计算。PoS**ICHEP2016**，176（2017）https://doi.org/10.22323/1.282.0176[8]德国电子同步加速器研究所（DESY）、费米国家加速器实验室（FNAL）、北欧电子基础设施合作组织（NeIC）：dCache 软件。Zenodo（2026）。https://doi.org/10.5281/ZENODO.23085911。h
t
t
ps://zenodo.org/doi/10.5281/zenodo.23085911
[9] Pordes，R.，等：开放科学网格。载于：《物理学杂志：会议系列》78，卷78，第012057页（2007）。h
t
tps://doi.org/10.1088/1742-6596/78/1/012057

[10]Sfiligoi, I., Bradley, D.C., Holzman, B., Mhashilkar, P., Padhi,S., Wurthwein,F.: 使用 glideinwms 获取网格计算资源的试点方法。载于：2009 年 WRI 世界计算机科学与信息工程大会。第 2 届，第 2 卷，第 428–432 页（2009 年）。https://doi.org/10.1109/CSIE.2009.950

[11]Illingworth, R.A.: 现代及未来费米实验室实验的数据处理系统。物理学杂志：会议系列**513(3)**，032045（2014 年）https://doi.org/10.1088/1742-6596/513/3/032045

[12]Lyon, A.L., Mengel, M.W.：“数据处理的‘最后一公里’：费米实验室的 ifdh 工具。《物理学杂志：会议系列》**513**(3)，032068 (2014)https://doi.org/10.1088/1742-6596/513/3/032068

[13] Mengel, Marc, 等：费米实验室实验的生产运营管理系统（POMS）。EPJ 网络会议 **245**，03024（2020 年）https://doi.org/10.1051/epjconf/202024503024

[14]Green， C.， 等：art 框架。《物理学杂志：会议系列》 **396**(2)， 022020(2012) https://doi.org/10.1088/1742-6596/396/2/022020

[15]Bakken, J., Berman, E., Huang, C.-H., Moibenko, A., Petravick, D., Zalokar, M.: 《费米实验室数据存储系统的现状》。（2004）。费米国家加速器实验室（FNAL），巴达维亚，IL。https://www.osti.gov/biblio/15016939[16] Davis, M., Afonso, J., Bachmann, R., Bahyl, V., Vera, J.C., Leduc, J., Cort´es, P.O., Rademakers, F., Wardenær, L., Yurchenko, V.: 《欧洲核子研究中心以外的 CERN 磁带档案系统：面向高能物理的开源数据归档系统》。EPJ 网络会议 **295**，01048（2024）https://doi.org/10.1051/epjconf/202429501048
[17]Aliaga, L., *et al.*：《MINERvA 探测器的设计、校准与性能》。Nucl. Instrum. Methods P hys. Res. A**743**，130–159（2014）https://doi.org/10.1016/j.nima.2013.12.053[18] Terekhov, I.：《D0 的元计算》。Nucl. Instrum. Methods Phys. Res. A**502**，402–406（2003）https://doi.org/10.1016/S0168-9002(03)00452-2
[19]Stonjek, S., Baranovski, A., Kreymer, A., Lueking, L., Ratnikov, F., *et al.*：《为 CDF 实验部署 SAM》。载于：2004 年高能物理计算会议论文集，第 1052–1054 页（2005）。CERN-2005-002[20]Albahri, T., *et al.*：《束流动力学校正对费米实验室 Run-1 μ子反常磁矩测量的影响》。Phys. Rev. Accel. Beams

24（4），044002（2021）
htt
ps://doi.org/10.1103/PhysRevAccelBeams.24.044002arXiv:2104.
03240 [physics.acc-ph]

[21]Mokhov, N.V., James, C.C.：MARS代码系统用户指南，第15版（2016）（2017）
https://doi.org/10.2172/1462233

[22]Mokhov, N., 等：根据强度前沿需求开发 MARS15 代码。Prog. Nucl. Sci. Tech. 4, 496–501（2014）https://doi.org/10.15669/pnst.4.496 arXiv:1409.0033 [physics.acc-ph[23]Muons, Inc.: G4beamline（版本 3.04）（2017）。
h
ttps://
www.muonsinc.com/Website1/G4beamline[24] Sagan, D.：Bmad：一个相对论带电粒子模拟库。Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectro meters, Detectors and Associated Equipment 558(1), 356–359（2006）
https://doi.org/10.1016/j.nima.2005.11.001。《第八届国际计算加速器物理会议论文集》[25] Berz, M., Makino, K.：COSY INFINITY 及其在单通和多通系统中的应用。Microscopy 75(S1),15–23（2026）
https://doi.org/10.1093/jmicro/dfag016[26] Makino, K., Berz, M.：高阶像差及其误差的严格计算。Microscopy 75(S1),24–30（2026）
https://doi.org/10.1093/jmicro/dfag019[27]Albahri, T., 等：对费米实验室 Run-1 μ子反常磁矩测量的束流动力学修正。Phys. Rev. Accel. Beams 24, 044002（2021）
h
t
t
p
s://
doi.org/10.1103/PhysRevAccelBeams.24.044002[28] Allison, J., 等：Geant4 的最新进展。Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 835, 186–225（2016）
https://doi.org/10.1016/j.nima.2016.06.125

[29] Allison, J., 等：Geant4 的发展与应用。IEEE Transactions on Nuclear Science **53**(1), 270–278（2006）https://doi.org/10.1109/TNS.2006.869826

[30] Agostinelli, S, 等：Geant4——一款模拟工具包。《核仪器与方法物理研究 A 部分：加速器、能谱仪、探测器及相关设备》**506**（3），250–303（2003）
https://doi.org/10.1016/S0168-9002(03)01368-8