

2023-2024年度工作报告

汇报人：马秋梅

职 称：高级工程师

2024-11-22

提纲

一、岗位职责

二、本人本年度科研任务及完成情况

1. 研究任务完成情况
2. 本人研究成果
3. 公共服务
4. 其它贡献

三、下年度工作计划



■ BESIII ■ JUNO ■ Daya Bay ■ CEPC

一、岗位职责



- BOSS发布
- BOSS升级
- 离线刻度框架
- BESIII在超级计算机上的应用
- 负责中央数据库系统的运行、管理和维护
- 数据管理系统
(bookkeeping)



- 负责离线软件开发应用平台管理、BOSS发布
- 离线刻度框架
- BESIII在超级计算机上的应用
- 负责中央数据库系统的运行、管理和维护
- 数据管理系统
(bookkeeping)



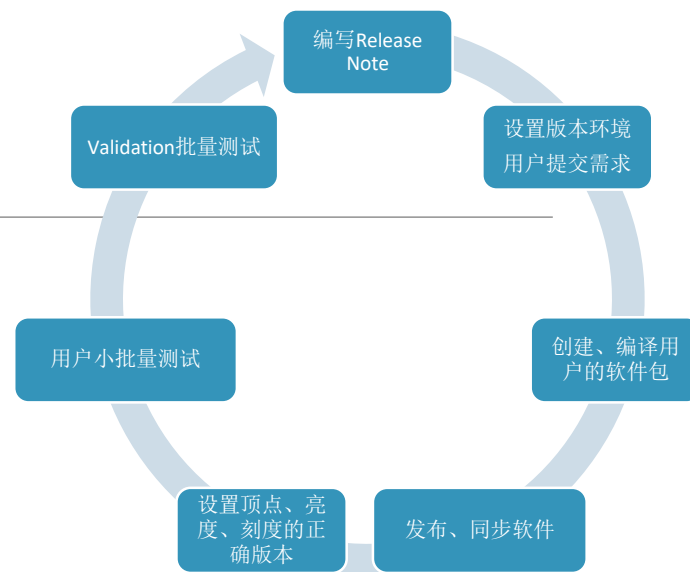
- 网络服务器运行管理
- 中央数据库系统运行管理
- 数据质量监控服务器



CEPC数据库系统管理

1.1 BOSS软件发布

- 7.1.1版本：
 - ✓ Round16 $\psi(3770)$ 数据产生
- 7.1.2版本：
 - ✓ $20fb^{-1} \psi(3770)$ 数据产生
 - 世界最大近阈值粲介子样本
 - 基金委重大项目“粲夸克衰变中标准模型的精确检验”中期考核指标
 - ✓ Round17 3780, 3768, 3800~3890扫描和 χ_{c2} 数据
- 7.1.3版本：
 - ✓ R-Scan数据重建



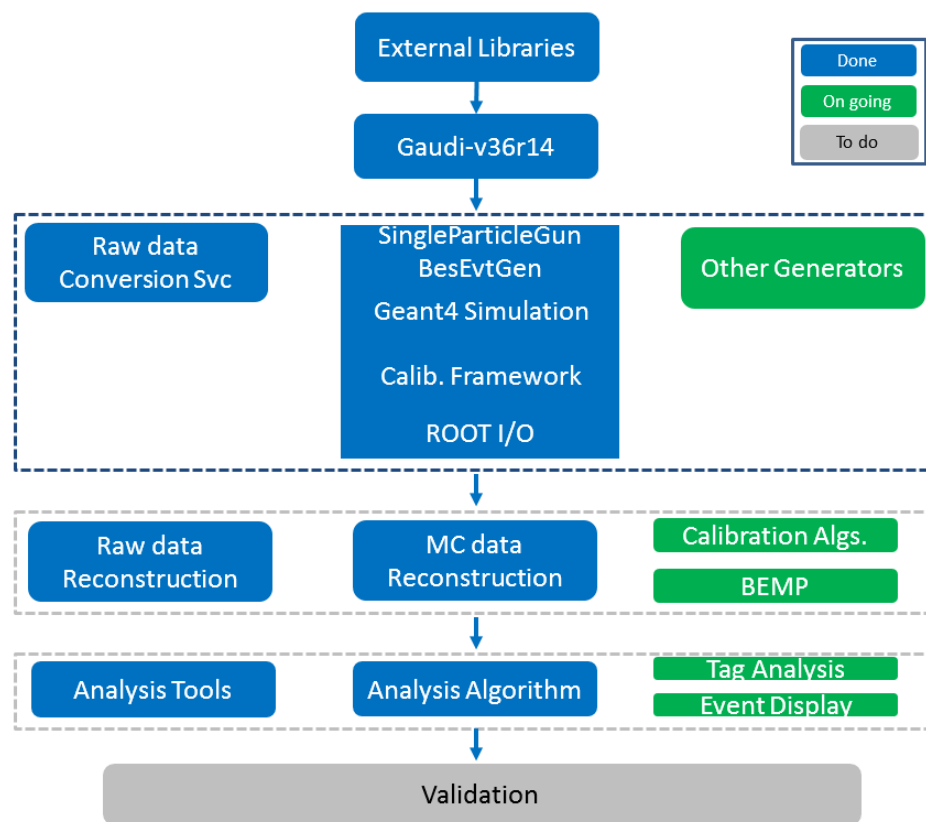
1.2 BOSS升级

- 升级原因：
 - 新操作系统不支持原有软件
 - 不支持原有的软件构建方式
 - 升级后优势：
 - ✓ 编译、运行速度提升
 - ✓ 新性能
 - ✓ 新工具的使用
- 最大规模、最复杂的升级

名称	7.1.3	7.2.0
操作系统	CentOS7	Alma9
编译器	GCC-4.9.3	GCC-13.1.0
外部库	LCG_84	LCG_106
Gaudi	Gaudi-v27r1	Gaudi-v38r2
版本控制	CVS	Git
配置管理	CMT	CMake
发布流程	手动	CI Git结合自动

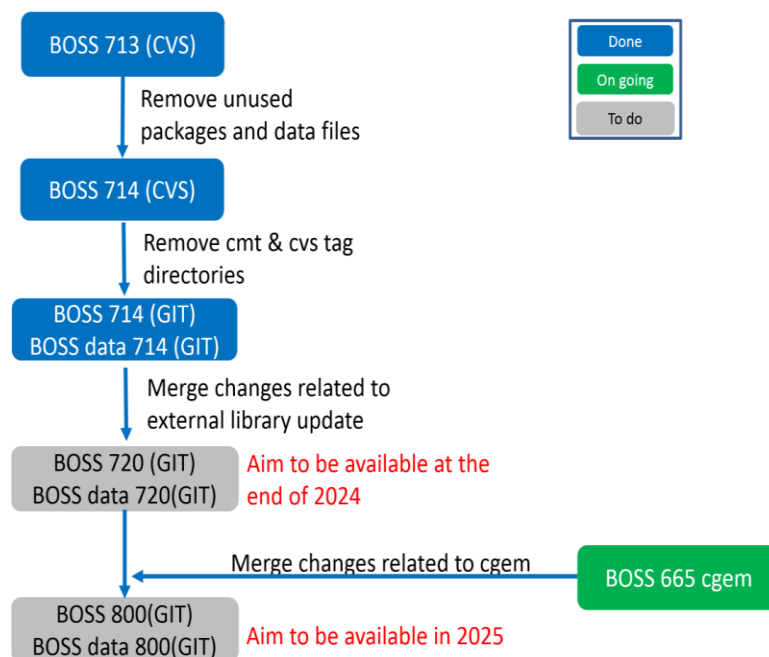
进展

- 已部署GCC、外部库、Gaudi到cvmfs
- 已部署7.2.0-pre1到cvmfs
- 已Validation, 与7.1.3结果基本一致
- BOSS源代码已部署到Git, 用户可以浏览、更新:
<https://code.ihep.ac.cn/boss>
 1. BOSS_Source
 2. BOSS_Gen
 3. BOSS_Data
- CI Git的节点已向计算中心申请



计划

- 7.2.0-pre1版本下测试代码
- 发布7.2.0-pre2、pre3等
- 7.1.3与7.2.0进行大批量validation
- 发布正式版本：7.2.0
- 升级前后编译、运行速度比较
- cgem软件加入BOSS
- 编写手册、培训用户、用户反馈

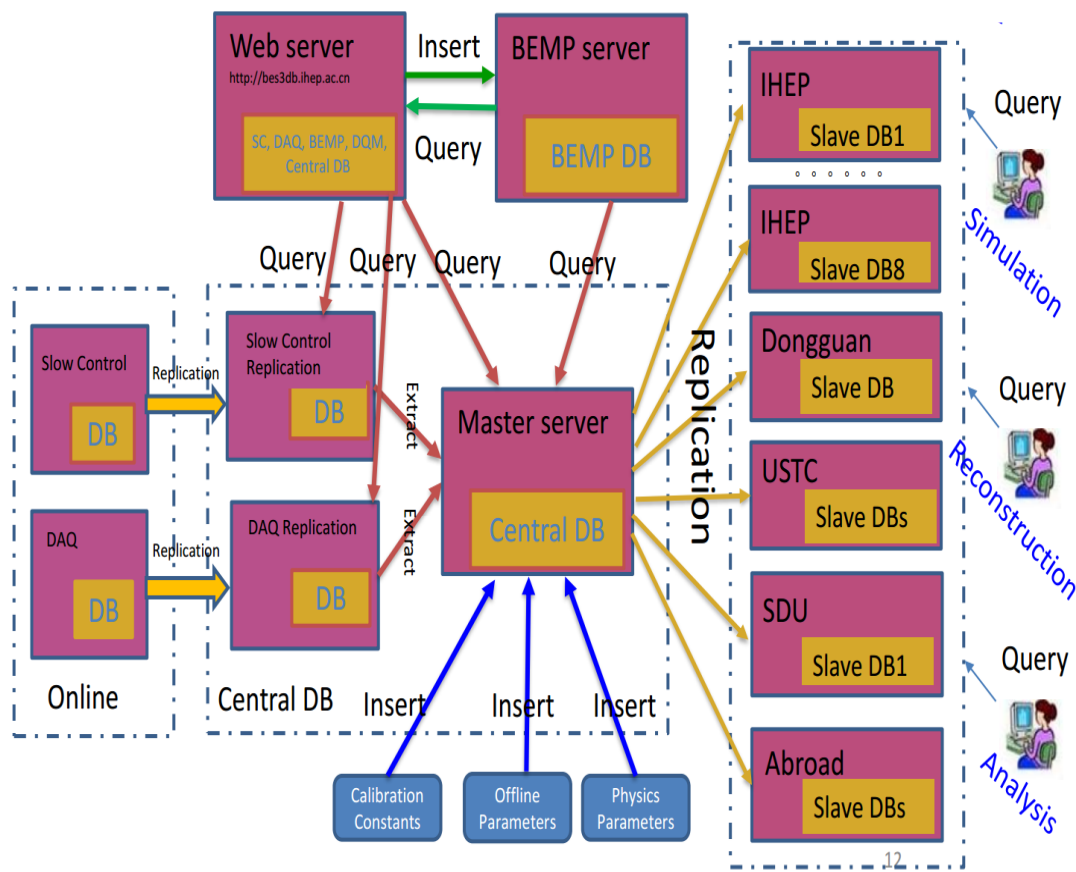


To make source code repository as small as possible:

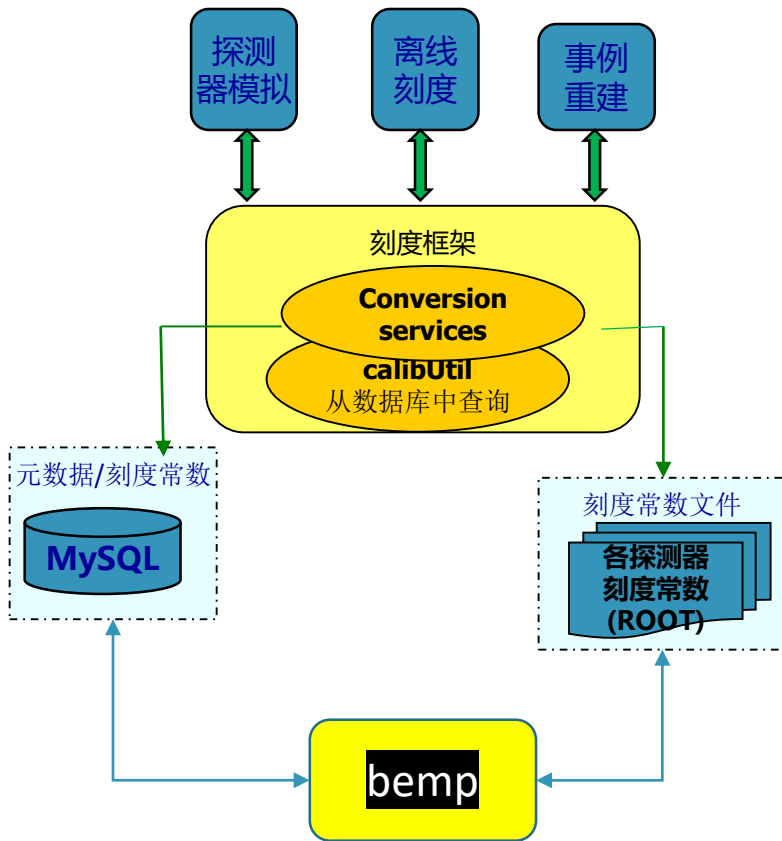
- (1) Remove unused packages and data files, validated with BOSS 713 and 714
- (2) Divide into two repositories: source code and large data files

1.3 BESIII中央数据库管理系统

- 有效地存储、管理和检索数据
 - ✓ 刻度常数
 - ✓ 顶点信息
 - ✓ 离线、在线亮度
 - ✓ Beam能量
 - ✓ 在线信息
 - ✓ 慢控信息
 - ✓ 离线数据库共**55**个数据表，约**1,000,000**条记录
 - ✓ 在线、慢控共**52**个数据表，约**3,138,000**条记录
- 极大提高数据的**可用性、可靠性和安全性**
 - ✓ 为高统计量和高质量的实验数据奠定基础
 - ✓ 为物理目标实现奠定基础



1.4 刻度框架的管理：保证所有刻度常数可以正常产生、读取



it	Calibration	Releases
	Calibration Tasks	
	Est ToF	Calibration
	Estime Calibration	
	InjSigInterval	
	InjSigTime	
	OffEvtFilter	
	MDC Calibration	
	MDC Alignment	
	MDC Data Constant	
	Mdc Tuning	
	TOF Calibration	
	TOF QE Correction	
	TOF Sim Calibration	

离线事例过滤刻度常数列表

Search

Boss Version

run from run to

status

event type

Search

Duplication

New Boss Version * New Status *

Duplicate Selected Duplicate All

添加 返回

共找到 2,507 条。 [首页/上一页] 1, 2, 3, 4, 5, 6, 7, 8 [下一页/尾页]

编号	起始运行号	结束运行号	文件名称	状态	BOSS 版本	刻度常数版本	事例类型	创建时间
557	65208	65208	OfflineEvtFilter_0065208_0065208_706.root	OK	7.0.6	1		2020-12-22
558	65209	65209	OfflineEvtFilter_0065209_0065209_706.root	OK	7.0.6	1		2020-12-22
559	65210	65211	OfflineEvtFilter_0065210_0065211_706.root	OK	7.0.6	1		2020-12-22

1.5 BES III 数据管理系统 (BEMP)

- 确保原始数据从磁带**及时、完整、正确地**拷贝到硬盘上

run: 3814个

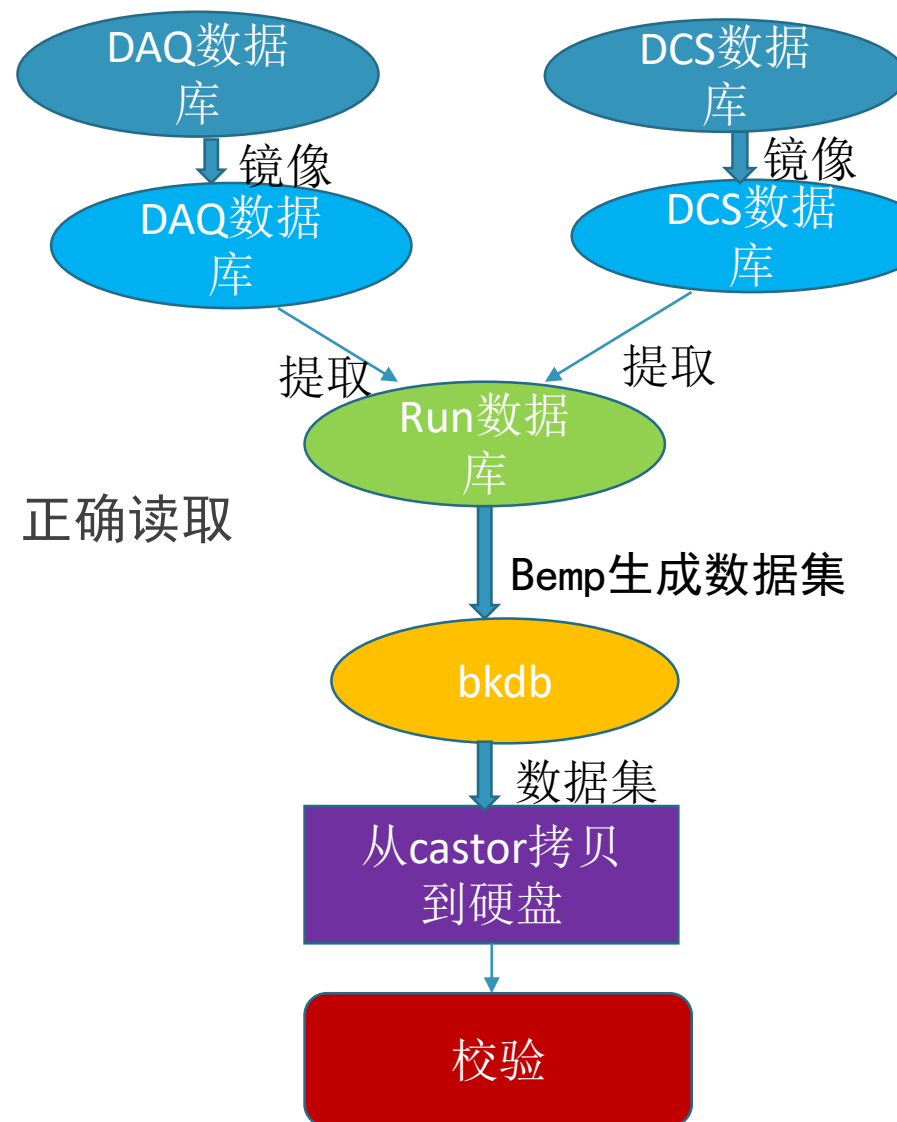
文件: 152801个

- 确保刻度常数及时正确地写入数据库、正确读取

- 修改刻度常数状态;
- 刻度常数复制;
- 查询、验证数据库中数据的正确性。

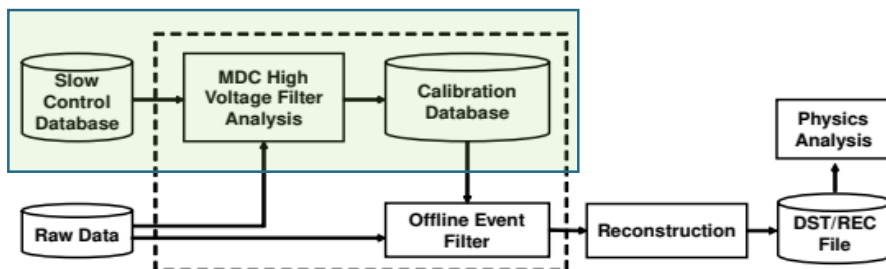
- 用户管理

- 功能完善和更新

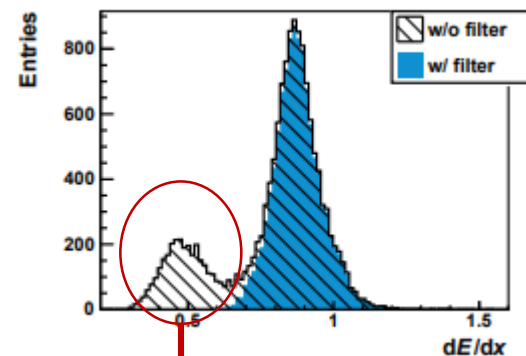


1.6 MDC高压异常数据过滤

- MDC高压异常导致dE/dx异常，影响粒子鉴别（PID）性能
- 建立有效过滤方法，去除异常事例，确保数据质量
 - 读取慢控制数据库MDC高压信息，与事例时间匹配
 - 生成的过滤时间窗存入刻度数据库
- 提供关键技术支撑
 - 慢控制数据库的访问和匹配
 - 刻度常数的存储与访问



Di mu事例dE/dx分布



高压异常导致有效过滤掉

1.7 BES III 软件在超级计算机上的应用

袁野 胡碧莹（中山大学）

张晓梅 陈璟锟（中山大学）

张瑶 唐健（中山大学）

目的：寻找更多的计算资源，更好地满足用户对计算资源的需求

完成：

- ✓通过容器和CVMFS两种方式把BES III 实验离线软件成功部署到了超级计算机上，离线数据库已部署并且成功实现了自动同步。
- ✓引入master+slave的二级结构，同时引入并行I/O管理技术，改善了软件在超级计算机上的并行性能表现。
- ✓通过采用Python+MPI脚本封装的方式实现高能物理数据处理软件在超级计算机环境中的MPI大规模运行，**十万核**测试，加速效率超过**75%**，成功实现了**十万CPU核和十万事例量的并发模拟**。
- ✓为方便用户操作，编写了从高能物理研究所的服务器提交作业到天河二号上的简化接口。
- ✓研究了从广州天河二号超算中心往北京高能物理研究所的数据传输情况，同时对数据传输的有效性、完整性进行了验证，传输速率可达20MB/S以上，基本可以满足数据传输的需求。

结论：

通过本项目的研究，为其它高能物理实验在超级计算机上的运行提供了一个范例，由于高能物理软件的相似性，其它高能物理实验的软件以后同样可以在超级计算机上运行，以解决当前高能物理实验计算资源紧缺的问题。

- ✓发表文章（1篇）

完善

- 优化远程一键提交作业系统
- 作业运行监测系统自动化
- 实现高能物理数据处理软件CPU 并行、加速卡并行和异构并行
- 解决大规模运行时对数据库同时随机访问的拥堵问题
- 适配不同并行环境的容器技术
- 提高数据传输速率

2.1 江门实验网络服务

- ✓ MediaWiki (论坛)
- ✓ DocDB、ihepjuno、EngDocDB
- ✓ 邮件列表 (70多个)
- ✓ Elog的管理维护
- ✓ 安全培训软件开发
- ✓ 成员数据库管理、创建作者列表 (11个)
(74个单位、732人)

- ✓ JUNO合作组的SSO认证审批
- ✓ 备份策略
- ✓ Gallery
- ✓ JunoBIM
- ✓ 现场数据库
- ✓ 现场网络服务器

Region	Institutions	Students	Postdocs	Seniors	Engineers	Non-Authors	Total Members
Armenia	1	1	0	2	0	0	3
Belgium	1	1	2	1	1	0	5
Brazil	2	0	0	2	0	0	2
Chile	3	2	0	4	2	0	8
China	32	168	25	204	16	32	445
Czech Republic	1	1	0	3	0	0	4
Germany	6	33	4	8	1	0	46
Finland	1	0	1	2	0	0	3
France	5	5	3	13	21	3	45
Italy	8	13	10	42	3	3	71
Pakistan	1	0	0	5	0	0	5
Russia	3	14	1	22	13	0	50
Slovakia	1	0	1	1	0	1	3
Thailand	3	6	0	9	0	0	15
Taiwan, China	3	9	1	5	0	0	15
United Kingdom	1	2	0	1	0	0	3
USA	2	6	1	2	0	0	9
Sum	74	261	49	326	57	39	732

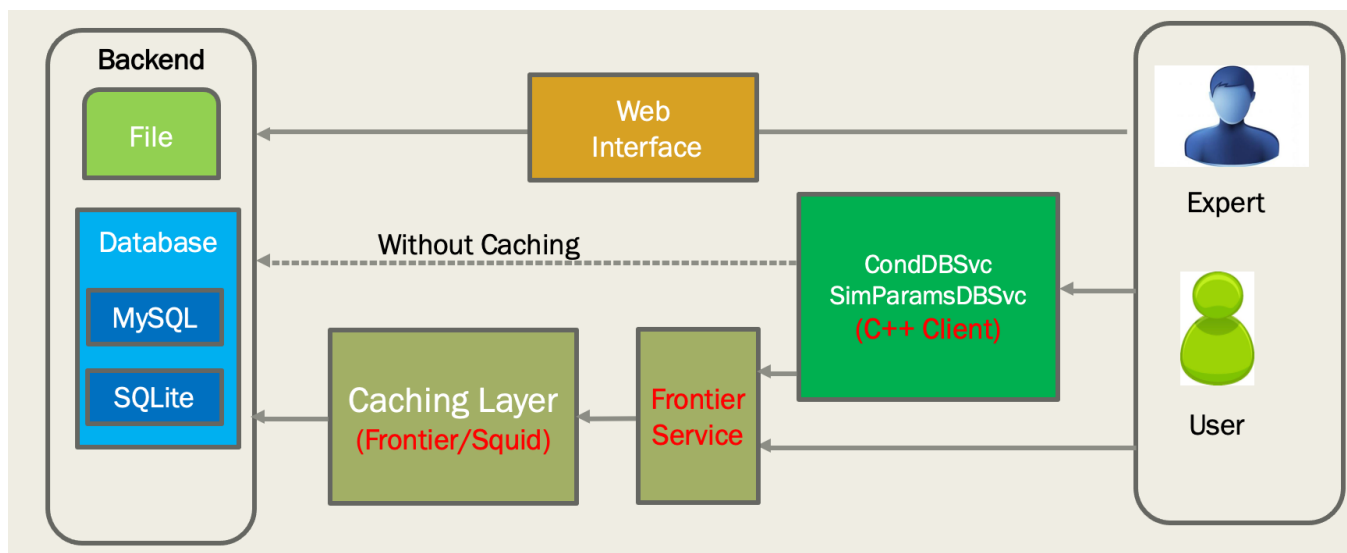
- [Ccepp-electroweak-juniors](#)
- [Ccepp-electroweak-seniors](#)
- [Ccepp-guanweihui](#)
- [Ccepp-hadrons-juniors](#)
- [Ccepp-hadrons-seniors](#)
- [Ccepp-lishihui](#)
- [Ccepp-neutrino-juniors](#)
- [Ccepp-neutrino-seniors](#)
- [Ccepp-quant](#)
- [Central_detector](#)
- [Dyb2+admin+wiki](#)
- [Ihep-pmt](#)
- [Interface to civil engineering](#)
- [Juno+admin+wiki](#)
- [Juno-centraldetector](#)
- [Juno-pmt](#)
- [Juno-pmt-cn](#)
- [Juno-wiki-admin](#)
- [Juno-zhihuibu](#)
- [Juno_cd](#)
- [Juno_cd_cn](#)
- [Juno_cdr_chinese](#)
- [Juno_cdr_english](#)
- [Juno_collaboration](#)
- [Juno_community](#)
- [Juno_community_china](#)
- [Juno_daq_dcs](#)
- [Juno_electronics](#)
- [Juno_ib](#)
- [Juno_ihep](#)
- [Juno_lowbkgcontrol](#)
- [Juno_observer](#)
- [Juno_pub_comm](#)
- [Juno_spm](#)
- [Juno_tb](#)
- [Juno_yellowbook](#)
- [Juno_youth](#)
- [Liquid_scintillator](#)
- [Mailman](#)
- [Offline](#)
- [Offline_chinese](#)
- [Physics_simulation](#)
- [Pmt](#)
- [Software_simulation](#)

2.2 江门实验中央数据库系统

离线数据库、Tao
数据库部署

在线数据库
镜像、抽取

慢控数据库



3. 大亚湾实验网络服务器

➤ 大亚湾网络服务器

- ✓ ELOG (值班日志)
- ✓ DocDB、EngDB、ProjectDB (文档管理)
- ✓ Trac、SVN、WebSVN、SVN用户的管理 (代码管理)
- ✓ Twiki (论坛)
- ✓ 图片、视频管理
- ✓ Maillist (邮件列表)
- ✓ 成员用户数据库管理 (成员管理)
- ✓ 备份策略

➤ dybdq:

- ODM
- DocDBs (5个)
 - ✓ IHEPDB
 - ✓ nEXOCNDocDB
 - ✓ nEXODocDB
 - ✓ nGdDocDB
 - ✓ ihepjun0

➤ dybdb、dcsdb

Internal > WebHome
Users | Groups | Changes | Index | Search | Go

Search in Daya Bay:

[DocDB](#) | [EngDB](#) | [ProjectDB](#) | [Subversion](#) | [Trac](#) | [Gallery](#) | [eLog](#) | [on-site](#) | [Off Shift](#) | [Data Quality Shift](#) | [Offline Data Monitor \(IHEP\)](#) | [Offline Data Monitor](#)

[Analysis](#) | [AnalysisCoordination](#)

What's new!

- [TechnicalPapers](#) - 2016-06-21 - 21:22 - [DavidJaffe](#)
- [Collab2016NTU](#) - 2016-06-20 - 14:57 - [BeiZhenHu](#)
- [DAQTroubleShooting](#) - 2016-06-18 - 10:53 - [KamBiuLuk](#)
- [ACUShiftSchedule](#) - 2016-06-16 - 12:56 - [WengqiangGu](#)
- [ConferencePresentations](#) - 2016-06-15 - 05:39 - [XinQian](#)
- [ShiftStandardHistograms](#) - 2016-06-14 - 16:03 - [QiangXiao](#)

Work Onsite

- [Run Operation / Shift](#)
- [Safety Training](#) -- Take it first before going onsite
- [People On Site](#) -- Register your travel plan here

PHPGRAPHY
root/ - 7 dir(s)

Welcome to Daya Bay Gallery

Instructions

You may upload your pictures with the following steps:

1. Click [admin](#) to log in with the common 'dayabay' user name
2. Select one directory or create a new one by clicking [admin](#).
3. Click [admin](#) | [Upload pictures/files](#).
4. Then select files from your computer to upload.

	Status Photos Jan 2011 4 file(s)
	video 2 dir(s)
	OnsiteMonitoring 6 dir(s)
	OnsiteMonitoring.20140331 6 dir(s)
	AD Camera Pictures 6 dir(s) - 1 file(s)

4.1 CEPC中央数据库

➤ CEPC数据库系统

- ✓ 性能优化及监控，故障诊断和排除
- ✓ 数据迁移，数据库备份、复制、镜像等容灾方案的制定和实施
- ✓ 数据库的安全管理、账户权限管理
- ✓ 确保数据库7*24安全、稳定运行

学术发展、学术交流、论文

会议和学术交流：

组织超算研讨会

参加BESIII软件组组长会并汇报工作进展

workshop、合作组会议

参加JUNO离线数据库讨论会，制定方案、部署方案

会议报告：

马秋梅，高能物理实验中央数据库管理系统，高能物理计算用户研讨会

发表论文：

1. Ze-Heng Zhang, **Qiu-Mei Ma**, *et al.*, [Offline filter of data with abnormal high voltage at BESIII drift chamber](#), *Journal of Instrumentation* **19** P060402
2. J.H. Zou, W.D. Li, **Q.M. Ma**, *et al.*, [Offline data processing system of the BESIII experiment](#), *Eur. Phys. J. C* **84**, 937(2024)

争取项目与经费

申请：高能物理大规模数据处理中的并行技术（未中）

参与：

- 1、COMET实验CDC软件发展
- 2、CGEM探测器数据处理关键技术
- 3、江门软件模拟与计算环境中科院先导专项

公共服务

- 1、BESIII 离线数据值班：2月/年，包括节假日。
- 2、BESIII 和江门的服务器采购
- 3、江门软件采购管理员和档案管理员
- 4、软件组工会联络人

下年度工作计划

BESIII实验

- 确保BESIII中央数据库系统正常运行
- 推动BOSS升级
- 确保BESIII数据管理系统正常运行

江门实验

- 实现江门中央数据库的数据同步机制；
- 配合分布式计算实现数据库的部署；
- 完善江门中央数据库管理系统的设计；
- 实现慢控到离线数据库系统的同步及数据抽取的自动化流程
- 完善江门网络服务器的功能。

大亚湾实验

- 大亚湾实验的网络和数据库服务的管理维护

CEPC实验

- CEPC中央数据库系统管理与维护

谢谢！