

中国科学院高能物理研究所

*Institute of High Energy Physics  
Chinese Academy of Sciences*



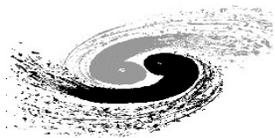
高能所计算中心

IHEP Computing Center

# LHAASO 模拟作业从X86到ARM计算集群的移植

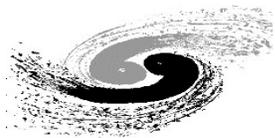
程垚松

计算中心, 中国科学院高能物理研究所

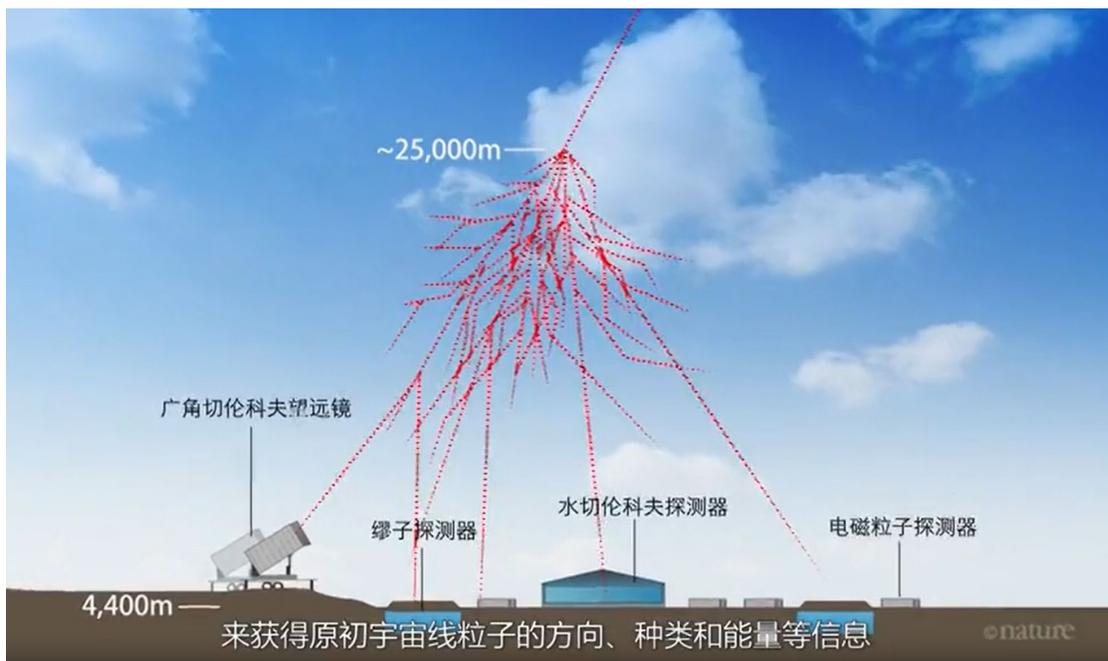


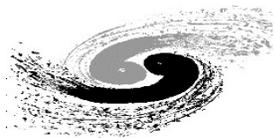
- LHAASO 实验与事例模拟
- ARM 架构与计算集群
- 软件移植
- 性能评估
- 总结与展望





- LHAASO: 高海拔宇宙线观测站
  - 水切伦科夫探测器阵列 (WCDA)
  - 一平方公里探测器阵列 (KM2A)
  - 广域切伦科夫望远镜阵列 (WFCTA)





- 蒙特卡洛模拟事例产生:

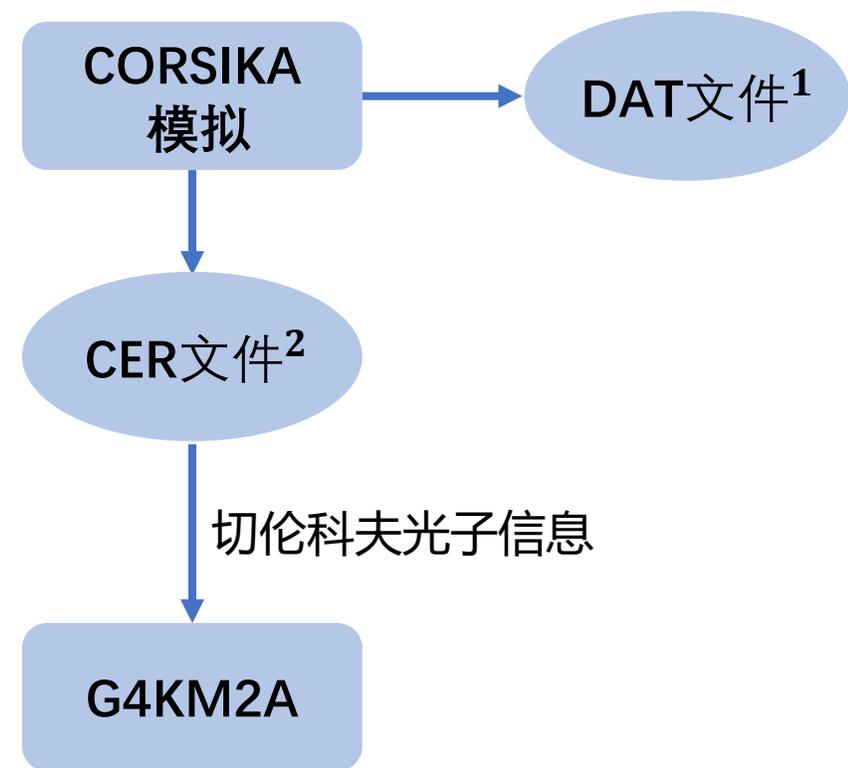
- 第1部分: 使用CORSIKA软件模拟原初粒子在大气中产生的广延大气簇射
- 第2部分: 使用G4KM2A软件模拟LHAASO的各个探测器运行

- CORSIKA:

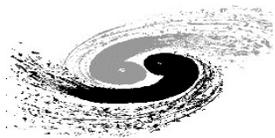
- 主要的模拟过程包括电磁成分的模拟, 高低能强子相互作用的模拟, 切伦科夫光子传播等

- G4KM2A:

- 模拟不同探测器对不同粒子的响应过程



1. DAT文件: 存储次级粒子信息
2. CER文件: 存储切伦科夫光子信息的文件



# 什么是ARM?



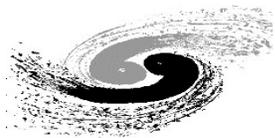
- ARM

- 一系列用于计算机处理器的精简指令集（RISC）架构
- 由ARM公司开发并将其授权给其他公司开发产品，如片上系统（SoC）和模块上系统（SoM）设计
- 与其他芯片相比成本低、功耗小、散热更低

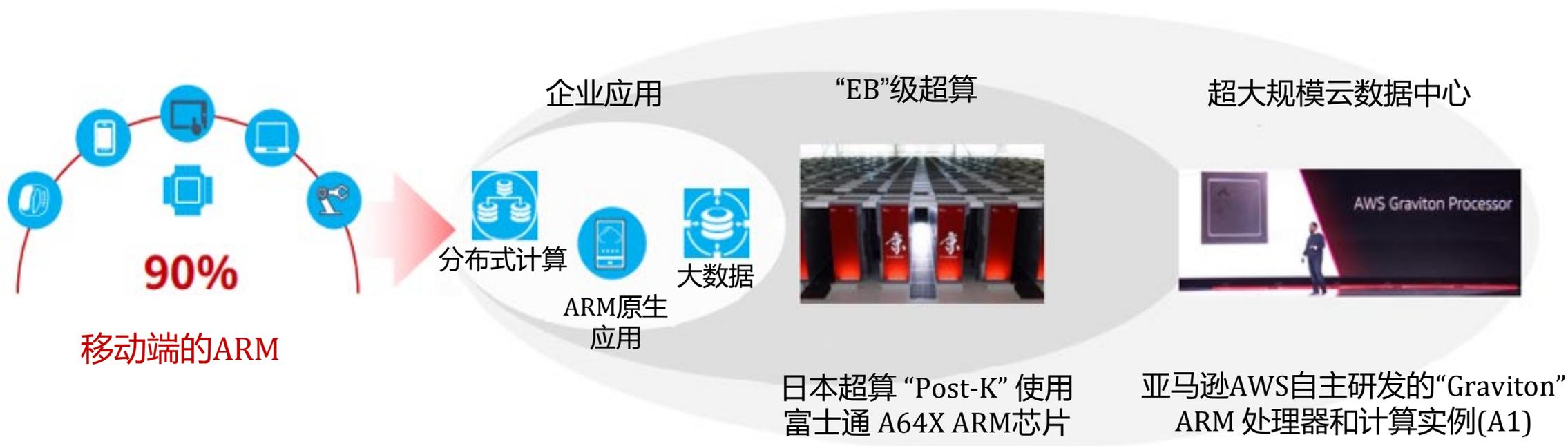
- AArch64

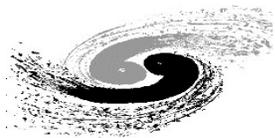
- ARM架构系列的64位扩展





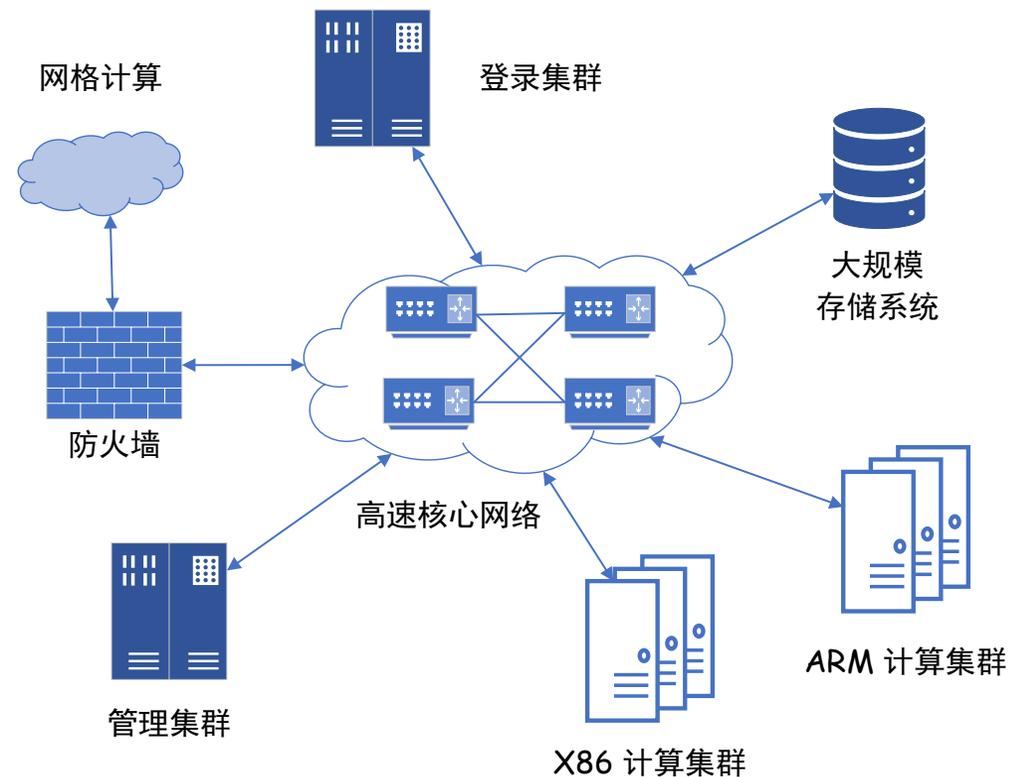
- ARM架构越来越多的应用于大型数据中心



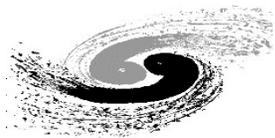


## • 高能物理大规模ARM计算集群

- 100个 ARM 节点, 9600 CPU 核
- 5.6 PB 文件存储& 1.6 PB 块存储
- 100 Gb RoCE v2 网络互联

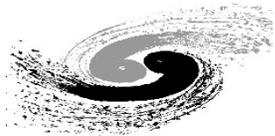


高能物理计算环境及其组成部分

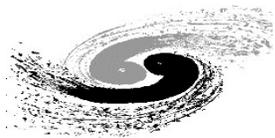


- Kunpeng 920 aarch64 ARM v8.2 , 32/48/64 core CPU
  - 最大主频率3.0 GHz
  - 8\*DDR4 controller 2933MT/s(max)
  - 支持 2\*100 Gb RoCE 网卡
- 计算服务器
  - 单节点包含96核 (48\*2), 内存 256 GB
  - 支持 Atlas 300 AI 加速卡



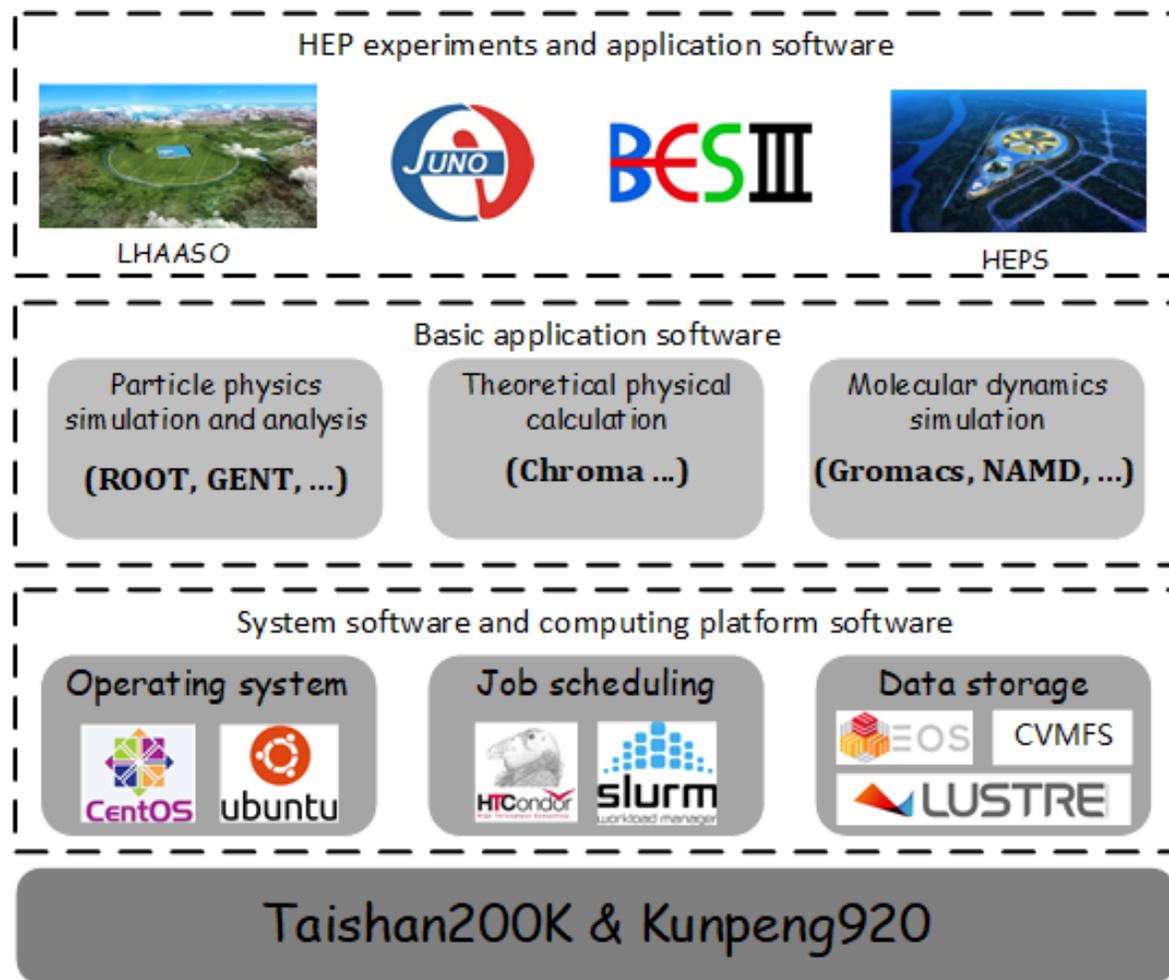


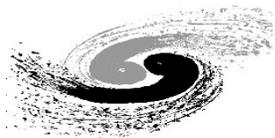
# 软件移植与评估



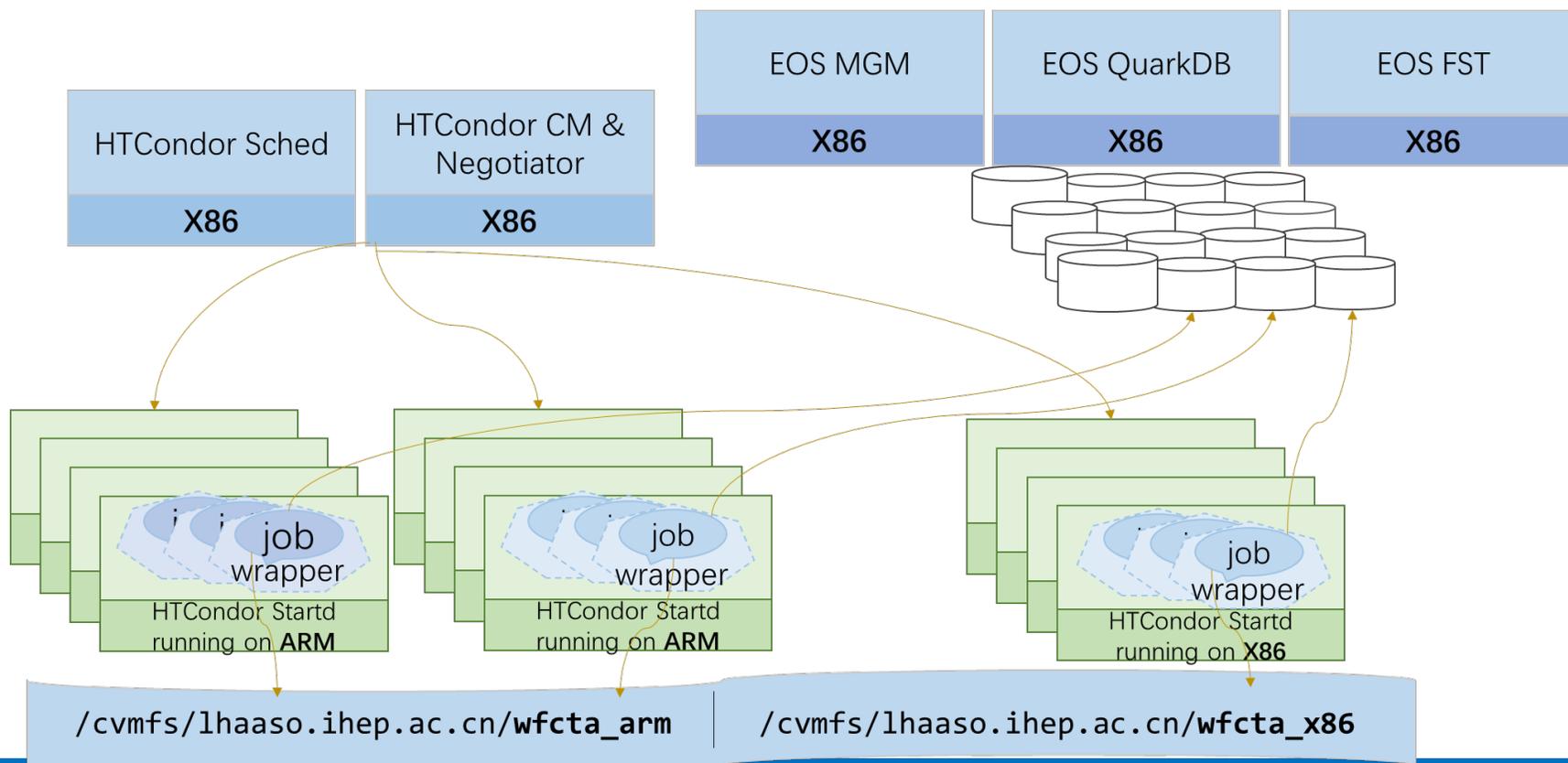
## • 高能物理数据处理系统

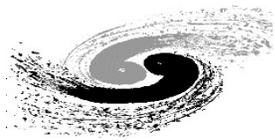
- 基于ARM架构
- 操作系统和计算平台软件
- 高能物理基础应用软件和框架
- 高能物理实验和实验分析软件





- 用户不需要关心作业将在哪种工作节点上运行
  - 调度器决定模拟作业的工作节点
  - 作业封装器打包用户作业脚本，并根据检测到的硬件架构选择软件版本





- ARM计算服务器

- KunPeng 920 CPU: aarch64 ARM v8.2
- Taishan 200K server: 96cores/node & 256 GB Memory

- CORSIKA:

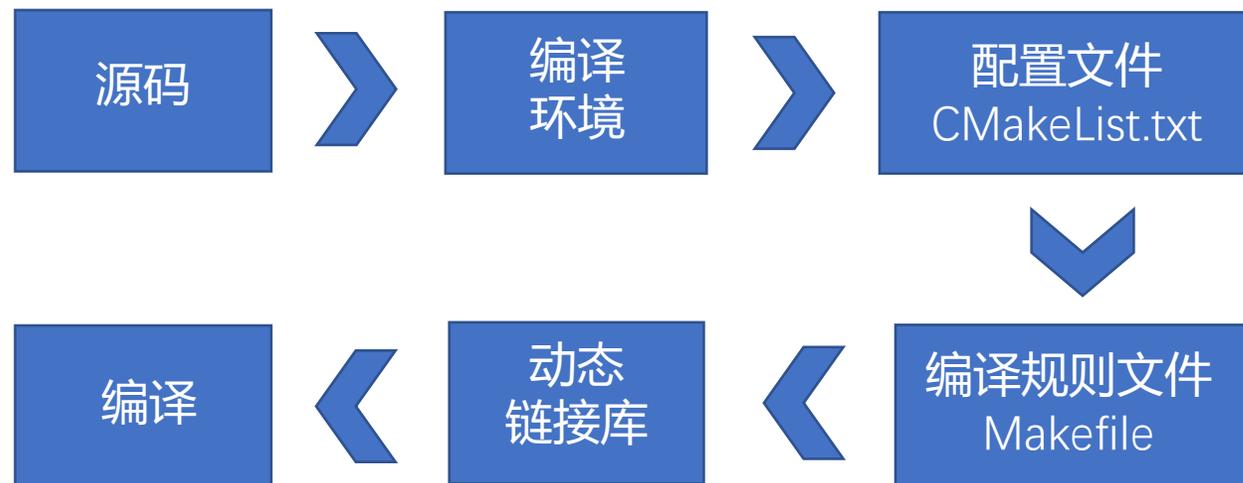
- Version: V77420
- Fluka: Fluka-CERN

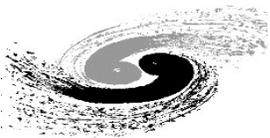
- G4KM2A

- Version: 10.1
- Geant4 Version: 10.4.2

- 其他系统软件

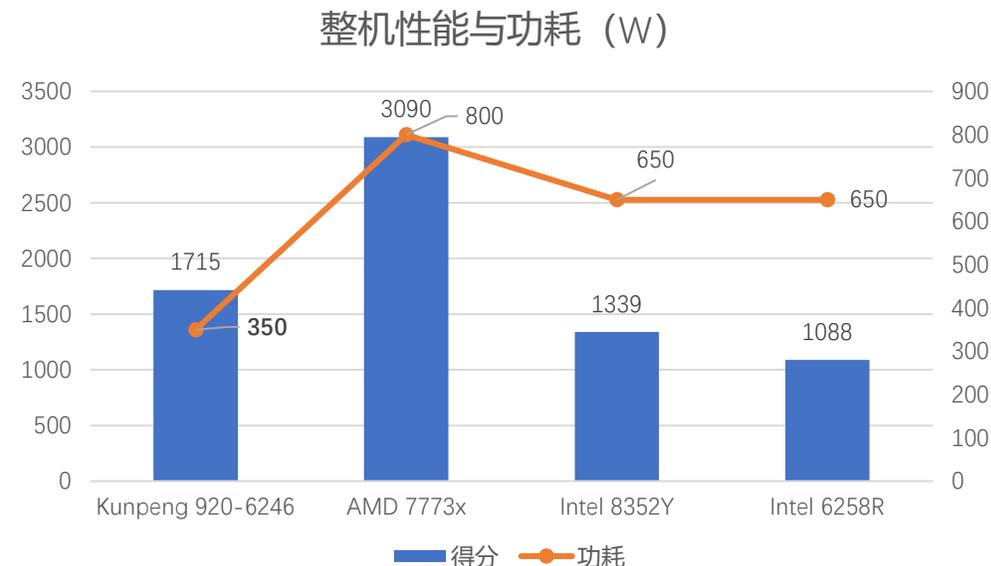
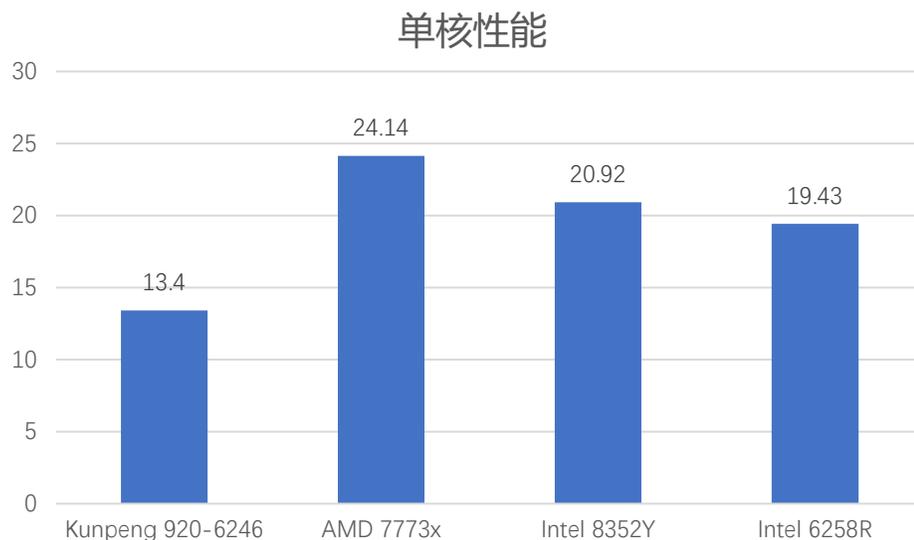
- HTCondor: 9.1.0
- EOS: 4.7.7
- CVMFS: 2.5.2
- ROOT 6.20 等

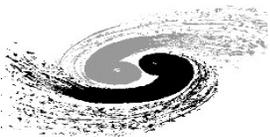




## • HS06 (HEP-SPEC06) 基准测试

- 在高性能物理中评估CPU性能的标准工具
- ARM : Kunpeng 920@2.6G Hz, 48 cores\*2 ~¥88000
- X86 : AMD EPYC 7773x@2.2G Hz, 64 cores\*2 ~¥83000
- Intel Xeon 8352Y@2.20GHz, 32 cores\*2 ~¥62000
- Intel Xeon 6258R@2.70GHz, 28 cores\*2 ~¥50000





## • 测试配置

- Corsika 模拟作业
- ARM : Kunpeng 920@2.6 GHz, 48\*2核 ~¥88000
- X86 : Intel 5218@2.3 GHz, 16\*2核 ~¥60000

## • 单核心

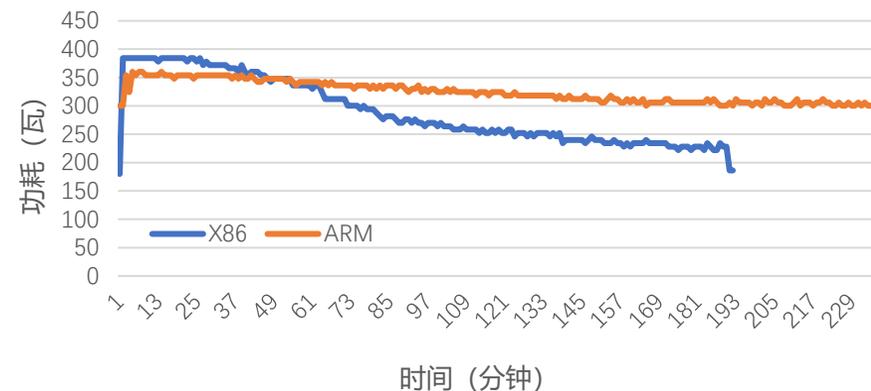
节点类型	作业数量	运行时间(m)	闲时功耗 (W)	运行功耗(W)
ARM	1	109	300	306(+6)
X86	1	90	180	240(+60)

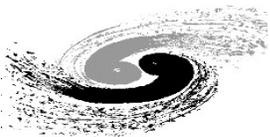
## • 整体服务器

节点类型	作业数量	总运行时间	单作业平均运行时间	耗电量 (瓦时)
ARM	96	4h6m	103.75m	1355.51
X86	32	3h20m	83.09m	967.36



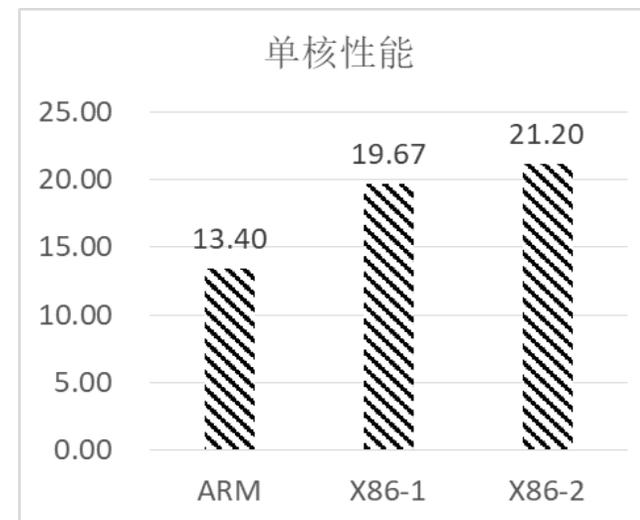
运行Corsika作业服务器功耗变化



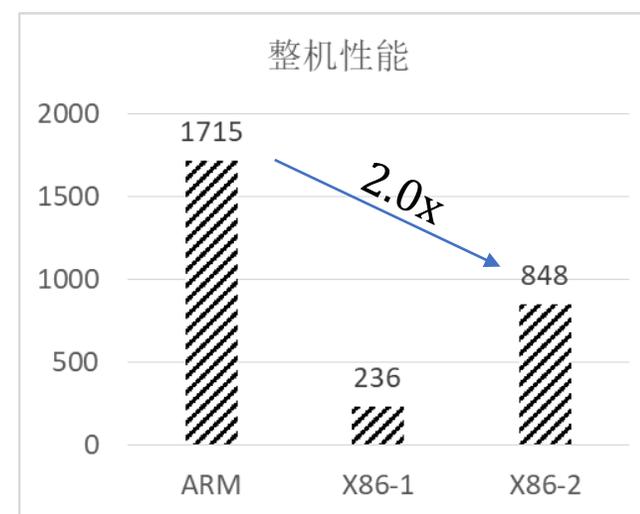
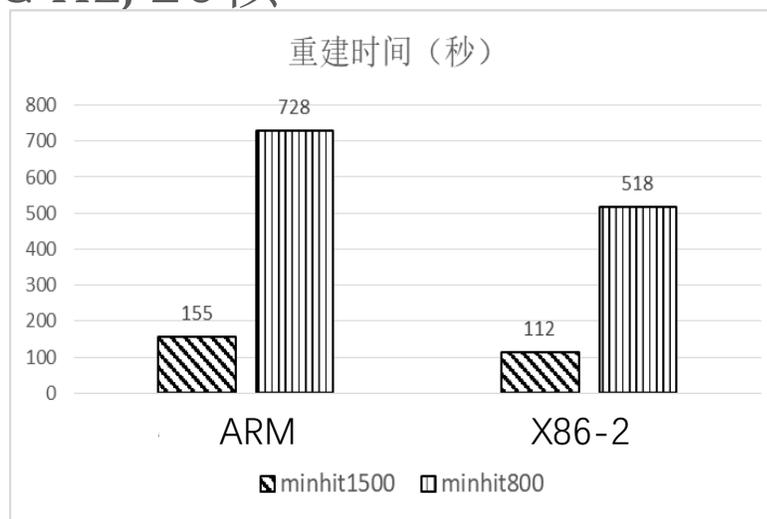


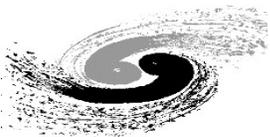
## • LHAASO-WCDA 事例重建测试

- 基于**ROOT**软件, 数据密集型计算
- 一个作业约1.1GB 包括 418, 816 个事例
- ARM: Kunpeng 920-6426@2.6G Hz, 48核
- X86-1: Intel E5-2620@2.0G Hz, 6核
- X86-2: Intel Gold 6230@2.1G Hz, 20核
- X86单核比ARM快40%
- ARM整机比X86快171%



HS06评分





# 准确性测试



- 使用V77420版本的CORSIKA, GHEISHA模型, 相同参数
- X86和ARM的运行结果对比完全一致

```
[yinlq@lxslc705 DATA]$ vi DAT008200.long
LONGITUDINAL DISTRIBUTION IN 201 SLANT STEPS OF 3. G/CM**2 FOR SHOWER 1
DEPTH GAMMAS POSITRONS ELECTRONS MU+ MU- HADRONS CHARGED NU
CLEI CHERENKOV
3.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.87852E+03
6.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 3.43058E+03
9.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 4.51305E+03
12.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 6.49557E+03
15.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 8.36225E+03
18.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 9.98005E+03
21.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.23217E+04
24.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.23217E+04
```

```
[wftamc@ln004 run082]$ vi DAT008200.long
LONGITUDINAL DISTRIBUTION IN 201 SLANT STEPS OF 3. G/CM**2 FOR SHOWER 1
DEPTH GAMMAS POSITRONS ELECTRONS MU+ MU- HADRONS CHARGED NU
CLEI CHERENKOV
3.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.87852E+03
6.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 3.43058E+03
9.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 4.51305E+03
12.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 6.49557E+03
15.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 8.36225E+03
18.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 9.98005E+03
21.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.23217E+04
24.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.23217E+04
```



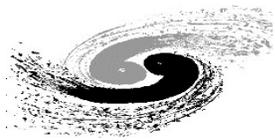
# 准确性测试



- 使用V77420版本的CORSIKA, FLUKA-CERN模型, 相同参数
- X86和ARM的运行结果不同, 可能与不同芯片浮点运算实现形式有关
- 准备大批量测试评估结果

```
[yinq@lxslc705 DATA]$ vi DAT008201.long
LONGITUDINAL DISTRIBUTION IN 201 SLANT STEPS OF 3. G/CM**2 FOR SHOWER 1
DEPTH GAMMAS POSITRONS ELECTRONS MU+ MU- HADRONS CHARGED NU
CLEI CHERENKOV
3.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.95017E+03
6.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 3.45858E+03
9.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 5.07139E+03
12.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 6.28789E+03
15.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 8.04014E+03
18.0 2.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 7.00000E+00 3.00000E+00 0.0000
0E+00 8.44670E+03
21.0 4.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 2.00000E+00 6.00000E+00 5.00000E+00 0.0000
0E+00 8.60475E+03
24.0 5.00000E+00 1.00000E+00 2.00000E+00 1.00000E+00 1.00000E+00 6.00000E+00 7.00000E+00 0.0000
0E+00 1.02565E+04
```

```
[wfctamc@ln004 run082]$ vi DAT008201.long
LONGITUDINAL DISTRIBUTION IN 201 SLANT STEPS OF 3. G/CM**2 FOR SHOWER 1
DEPTH GAMMAS POSITRONS ELECTRONS MU+ MU- HADRONS CHARGED NU
CLEI CHERENKOV
3.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 1.95017E+03
6.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 3.45858E+03
9.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 5.07139E+03
12.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 6.28789E+03
15.0 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 1.0000
0E+00 8.04014E+03
18.0 2.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 7.00000E+00 3.00000E+00 0.0000
0E+00 8.75959E+03
21.0 4.00000E+00 0.00000E+00 0.00000E+00 1.00000E+00 2.00000E+00 6.00000E+00 5.00000E+00 0.0000
0E+00 8.75959E+03
24.0 5.00000E+00 1.00000E+00 2.00000E+00 1.00000E+00 1.00000E+00 6.00000E+00 7.00000E+00 0.0000
0E+00 1.07156E+04
```



- 实现了ARM计算集群的建设和应用
- 高能物理软件的移植、测试与应用
  - CORSIKA、G4KM2A、Geant4、ROOT、EOS等
- ARM芯片与服务器性能评估
  - 与X86相比单核性能低，整体服务器表现更好
  - 生态仍需不断建设
- 评估更多、更新的芯片运行高能物理作业时的性能和功耗
- 进行大批量作业测试与生产
- 移植更多的高能物理软件并评估