按模型与数据量试跑GPU并规划租期和存储

完整操作教程 · 参考固定 · 往返恢复验收

固定模型与代表性POD5做单GPU试跑，把识别、传输、校验和工作盘分开预算。

运行边界：示例代码需按安装版本和真实项目输入复核；未提供真实生物数据，语法或小样本检查不代表科学结果和性能验收。

选择Dorado识别服务器时，先固定识别任务与模型，再测代表性POD5的实际吞吐。fast、hac、sup以及增加修饰识别会改变资源和方法要求，不能把便宜模型的速度直接用于另一个目标。下面测一组单独整理的完整pilot POD5文件；不是用文件大小换算通用GPU型号。

把可比较条件写清楚

依当前Dorado2.1.1支持表核对系统、驱动与设备，并用实际程序试跑确认。GPU能够被nvidia-smi看到只是第一关。pilot最好包含本项目典型读长与信号分布，另用较大一批检验稳定性；只取最小文件，启动与模型加载开销会夸大单位时间。不同读长下每秒read数也不能直接比较。

固定条件	记录值	避免的偏差
方法与模型	simplex/修饰、完整模型名及版本	不同任务混做速度对比
数据	pilot清单、输入字节、实际读数/碱基	只比较POD5压缩体积
设备	可见GPU、驱动、其他占用	单卡试跑后直接承诺多卡线性
存储	本地或网络盘、输出路径	缓存或共享盘差异被算成GPU性能

测一次完整pilot，并保留主机与GPU记录

需安装GNU time、samtools和nvidia-smi，模型已下载并验证。替换两个路径，在Linux Bash中运行。示例只请求当前可见第一块GPU；GPU资源应已分配给本任务。time记录进程主机资源，不是GPU显存统计；另存的设备快照也不是完整峰值，需结合平台监控。

```
set -euo pipefail
: "${MODEL:?■■■■■■■■■■}"
: "${PILOT:?■■■■■■■■POD5■■}"
test -d "$MODEL"; test -d "$PILOT"
mkdir dorado-benchmark1
cd dorado-benchmark1
dorado --version > version.txt
nvidia-smi > gpu-before.txt
du -sb "$PILOT" > input-bytes.txt
/usr/bin/time -v -o resources.txt dorado basecaller \
  "$MODEL" "$PILOT" --recursive --device cuda:0 \
  > pilot.bam 2> basecaller.log
samtools quickcheck -u pilot.bam
samtools view -c pilot.bam > records.txt
du -sb pilot.bam > output-bytes.txt
nvidia-smi > gpu-after.txt
```

把试跑结果变成时间与盘预算

填写下表中的实测值，输入、输出统一用字节或同一种单位。耗时至少分为模型下载、传入、识别、校验和传出；下载完成的热启动不能被称为全部准备时间。若用输入字节比例外推，只把它当本批同条件的粗估，并对第二批核对，不保证不同化学、模型或读长可复用。

预算项	本项目填写	计算/验证
识别时间	pilot耗时____；全量/pilot比例____	同条件粗估耗时=二者乘积，再核对较大批次
工作盘	POD5____；BAM____；模型____；失败文件____	加上同时保留的数据和约定余量
占用周期	准备____；识别____；验收____；回传____	按总周期排期，不只算核心识别时间
并发	单任务显存____；实际可用____	先单任务稳定，再小批次并发验证

低利用率和显存不足怎么分开处理

GPU利用率低且网络读取等待明显，先把获准数据放到本地工作盘复测；不要马上换更大GPU。CUDA OOM时查同卡其他进程与Dorado自动batch size，按官方建议从原批量逐步降低并记录实际值。改变hac/sup会改变方法配置，不能为让测量好看而偷偷降级。

验收除退出码和BAM可读外，还要核对模型、输出量、read完整性与项目所需质量指标。分批结果合并须先确认样本归属、头部和重复read处理；不要直接cat BAM。保留一次失败重跑和结果回传的时间，满足这些条件后再以实际缺口决定GPU数量、盘容量与使用周期。

模型下载产生的文件也计入持久盘；更换候选模型时分别登记，不把多个测试目录遗忘在容量预算之外。

试跑前先确认输入代表性和模型身份

POD5目录需要先整理成明确的输入清单：文件数量、文件大小、样本或flow cell归属、生成批次、是否包含重复文件以及所处存储位置。不同批次的数据若读长、化学版本或信号特征不同，应至少选择一个代表样本分别试跑。试跑数据太小会被模型加载时间主导，太大又可能耗尽临时盘；建议先做短试跑排除安装问题，再挑中位或偏大样本测完整流程。所有文件必须有授权和完整性确认，不能把网络盘上仍在写入的POD5作为稳定基准。

记录Dorado版本、模型名称及路径、模型是否包含修饰识别、设备参数、驱动版本和输入目录校验信息。fast、hac、sup代表不同准确度/速度取舍，修饰识别又会增加相应分析负担。项目应先定义准确度与输出需求，再挑模型，不以一个无关模型跑得快来推算正式任务。固定参数后执行pilot，任何模型变更都另建结果目录和基准记录。

把等待时间拆成能定位的阶段

总租期不应只看basecaller运行时间。至少记录数据上传、文件校验、模型准备、试跑、正式识别、BAM检查、结果迁出和失败重跑的时间。若数据已在目标磁盘且模型缓存完成，测得的是热启动分析耗时；新项目通常还要计入上传和模型下载。把每阶段的起止时间写入记录，可避免网络传输慢却误判GPU不足，也能提前排定数据交接窗口。

阶段	应记录的量	诊断方向
上传/读取	输入字节、文件数、实际读速、挂载点	网络盘或校验成为瓶颈时先处理I/O
识别	碱基数或read数、模型、每小时进度	相同模型和样本条件下比较吞吐
显存与GPU	设备占用采样、其他进程、错误日志	低利用率先查读取等待，OOM再查batch
输出/验收	BAM大小、索引、quickcheck、统计	确认输出完整且样本对应正确
迁出与清理	结果字节、校验时间、保留路径	为副本和传输中间态预留空间

按实际瓶颈处理低利用率和显存错误

如果GPU利用率经常很低，先确认Dorado正在读取文件、输入盘是否繁忙、CPU是否被限制，以及设备是否被其他用户占用。可以将同一授权pilot复制到本地NVMe工作盘，在其他参数不变时复测；如果吞吐明显改善，说明I/O路径可能是主要限制。不要同时换GPU、改模型和改盘，否则试跑无法回答问题。利用率截图只反映采样时刻，应保存一段时间序列或平台监控。

若出现CUDA out of memory，先排查其他GPU进程、可见设备映射和同卡并发。根据官方当前版本说明逐步调整batch size，并记录修改前后值、显存峰值与输出质量。更换模型不是通用的显存修复方法，因为它可能改变分析目标。一次成功的小文件运行也不证明长时间任务稳定：使用较大pilot观察温度、错误、速度波动及工作盘增长，并确认失败后可从可行位置重跑。

预算容量时区分临时盘、输出和保留

工作盘计划应包含输入POD5、模型文件、临时输出、BAM、索引、日志、分批中间结果和传输中的副本。若基线POD5必须保留用于复算，就不能把它计为任务完成后可立即删除。实际输出大小应从pilot测得，且BAM压缩与读长、模型和输出设置有关，不要使用一个固定比例作为保证。为每个样本记账，按同时运行样本数加总峰值；不同文件系统的剩余空间不能简单相加。

验收时先检查退出码、日志中的完成信息和BAM可读性，再核对记录数、read ID抽样、样本来源和必要的质量指标。samtools quickcheck只能做基础完整性检查，不会证明碱基识别正确，也不能替代下游质量分析。上传、计算与迁出各自保存SHA256，核对源和目标一致后再按项目保留策略处理原始文件。满足科学和交付条件后，才根据pilot和完整批次差异来确定所需GPU、磁盘和使用周期。

区分容量估算和实际分析验收

资源试跑的主要问题是“给定输入和模型能否在当前额度内稳定完成”，不能用它替代测序质量评估。Dorado输出后，需要按项目下游要求查看Q值分布、reads数量、read长度、产出碱基数和样本归属等指标；若有修饰检测，还要验证对应标签和工具链支持情况。不同研究目的对准确度与产量要求不一样，模型选择由分析问题决定，资源测试只能在方法已确定后比较配置。

对完整批次，可以在运行计划里把样本按输入量分层：小、中、大三档，先分别做代表性pilot；然后按计划并发跑一个短批次。检查吞吐是否随文件量近似稳定，长任务是否出现GPU重置、读盘停顿或磁盘增长异常。若不同档位的单位字节耗时差别很大，回看读长、信号分布和存储吞吐，不要用平均速度掩盖最慢批次。为正式排期保留重跑和验收时间，并说明这个估计只适用于已测试模型、化学版本和设备组合。

最后做一份交付包清单：原始输入及校验和、模型和软件版本、完整命令、设备信息、运行日志、BAM及其索引、quickcheck结果、read统计、输出校验和、迁出位置以及清理计划。需要复算时，应保留足够的原始信号和模型信息；若计划删除POD5，先与课题负责人确认后续是否还需重新识别或复核。数据保留和服务器租期安排应依据复算需求及院校数据管理政策，而不是只看此次输出是否已经下载。

将试跑范围写进最终资源建议

资源建议应说明测试了哪种化学、模型、Dorado版本、POD5大小范围、GPU型号与并发数。未覆盖的修饰类型、超大样本或多卡任务明确标为待测。对输入读取和结果回传分别给出实测速度区间及测试位置，不能承诺校园网络和用户本地网络具有相同速度。完成一批任务后用真实账目校正预估：比较预计与实际GPU小时、磁盘峰值、上传下载时间和返工次数，后续同类实验再逐步形成团队自己的基准。这样资源选择有可追溯证据，而不是依赖单次演示。

给课题组的最终说明写清假设

汇报时用“本次测试在某版本、某模型、某类输入和某设备上完成”描述结论，同时列出未测范围。若真实样本规模或GPU型号不同，应重新做小批验证。估算耗时给出组成和数据来源，说明是否含上传、下载、校验、模型准备和返工时间。磁盘建议说明哪些数据需同时保留、是否包含备份副本及清理时间点。条件透明，课题组才能决定是否需要更多资源或更长周期，也能在后续实验变化时快速更新计划。

同一批次中若有失败文件，先按清单隔离并记录对应样本，不要把未完成文件与成功输出直接合并。重跑结束后通过read ID或样本标签确认来源，检查重复输出处理规则。资源试跑与结果验收分别记录，可以让算力选择和测序分析质量都得到清晰说明。

参考资料

配套练习包

本PDF对应的ZIP包含本文代码块、主题练习模板或合成数据、README和SHA256校验文件。请从同一教程页面下载练习包；先核对校验值，再在独立目录复现。

参考资料

Dorado 2.1.1平台与设备选择 · 来源资料

Dorado显存与读写排错 · 来源资料

Dorado模型选择及固定路径 · 来源资料

GNU time运行资源记录 · 来源资料

samtools quickcheck检查范围 · 来源资料

操作建议：先在合成小样本上验证流程，再依据项目参考、标签合同、存储策略与获配资源验收真实数据。