按数据读写习惯配置工作盘与长期保存空间

规划Seurat归档空间时，不能只看一份RDS。项目可能同时保留原始输入、工作对象、检查点、最终结果及临时文件；若磁盘layer引用外部目录，那个目录也是对象可用性的一部分。先列文件生命周期和恢复目标，再估算同一时间真正并存的容量。列出文件及生命周期把原始输入、参考数据、工作对象、检查点、最终交付和

规划Seurat归档空间时，不能只看一份RDS。项目可能同时保留原始输入、工作对象、检查点、最终结果及临时文件；若磁盘layer引用外部目录，那个目录也是对象可用性的一部分。先列文件生命周期和恢复目标，再估算同一时间真正并存的容量。

列出文件及生命周期

把原始输入、参考数据、工作对象、检查点、最终交付和可重建缓存分开登记。记录每项大小、目录、产生步骤、保留期限和是否可重建。不要把中间文件自动当垃圾删除：先确认恢复流程用不到它，再由项目负责人决定清理。

类别	记录	迁移重点
输入/参考	版本、校验值与来源	重现所需文件完整
工作对象	RDS及外部layer目录	相对路径和权限
检查点	步骤、参数和可复用性	与脚本版本对应
最终结果	验收标准与交付目录	新位置能重新打开
缓存/临时	可重建依据与期限	清理前确认不再依赖

按同时存在的最大值估算

示例：输入200 GiB、工作对象60 GiB、检查点60 GiB、临时目录80 GiB、结果40 GiB，若这些确实同时存在且位于同一文件系统，则需约440 GiB，另留增长余量。若工作对象只是引用同一个输入目录，或目录位于不同挂载点，应按真实布局重新计算，避免重复计数或误把其他盘的余量加进来。

```
du -sh input objects checkpoints temporary results
df -h .
df -i .
```

空间估算还要覆盖文件系统配额、inode和写盘期间可能并存的新旧版本。压缩包或备份制作时，源与目标通常会同时占空间；外部磁盘矩阵也可能不在RDS目录所在挂载点。

迁移包要包含外部矩阵

普通saveRDS保存R对象；Seurat磁盘layer可能是路径引用。对BPCells等后端，应按SeuratObject当前的SaveSeuratRds说明决定move和relative行为，打包RDS与后端目录及必要依赖。复制后检查目录权限，不能只校验RDS文件存在。

```
library(SeuratObject)
obj <- LoadSeuratRds("bundle/project.rds")
print(Assays(obj))
for (a in Assays(obj)) {
  print(Layers(obj[[a]]))
  for (ly in Layers(obj[[a]])) {
    m <- LayerData(obj, assay=a, layer=ly)
    print(list(assay=a, layer=ly, dim=dim(m)))
  }
}
writeLines(capture.output(sessionInfo()),"bundle/session.txt")
```

在迁移目标启动新的R会话，执行重载和小范围矩阵读取，抽查样本标识、细胞数及必需layer。生成带相对路径的归档时，移动整套目录后再验收；绝对路径仍指向旧位置的对象不算迁移成功。

演练恢复，再决定容量

把一次恢复演练记为输入清单、目标路径、校验值、对象重载结果和失败记录。只有能从归档位置恢复到约定检查点，备份才满足该项目的恢复目标。原始数据、对象和结果涉及受限样本时，目录权限与共享范围也要纳入交接。

提交磁盘需求时提供并存文件清单、实际挂载点、最大空间、增长速度、并发写入和保留周期。这样才能区分容量不足、I/O等待、后端路径失效和权限问题，并按项目生命周期安排工作盘与归档空间。

写一份可检查的归档清单

每次交付记录相对路径、文件大小、修改时间、SHA256、文件用途与来源版本。生成清单后在目标位置重新计算哈希，比较文件是否完整；哈希一致只能证明字节相同，仍要启动R、加载对象并读外部layer，检查软件依赖和权限。清单应保存在归档目录旁，并另存于项目记录中。

```
find bundle -type f -print0 | sort -z | xargs -0 sha256sum > bundle/SHA256SUMS.txt
(cd restored_bundle && sha256sum -c SHA256SUMS.txt)
```

跨平台路径大小写、符号链接和共享权限也可能影响恢复。受限人类数据要确认访问控制和保留期限；迁移测试应使用独立目的目录，不要以覆盖方式同步唯一归档。

区分恢复目标和冷备份

经常重跑的工作对象适合可快速读取的工作盘；长期保留的原始数据和验收结果可能适合归档存储。两者读取频率和恢复时间目标不同，不能仅比较每GiB价格。压缩包与原目录在打包过程中会同时占空间，历史版本也要计入并存峰值；其它挂载点的空余空间不能合计为当前目录可用值。

参考资料

1. R saveRDS/readRDS官方手册

2. rsync官方手册

3. sha256sum手册（Debian发布）

参考资料

R saveRDS/readRDS官方手册 · 官方资料

rsync官方手册 · 官方资料

sha256sum手册（Debian发布） · 官方资料