

PR #2395 完整报告

radixark/miles

docs: drop the Platforms section and refresh the hardware list

合并时间: 2026-08-12 07:07

原文链接: <http://prhub.com.cn/radixark/miles/pull/2395>

执行摘要

- 一句话: 删除重复的 Platforms 文档, 统一硬件清单到安装页
- 推荐动作: 可以将本 PR 作为“文档单一事实来源 (SSOT) 整理”的范例阅读, 重点看 docs.json 中重定向配置的写法、硬件表按代际与 CI 保障级别分档的思路。无需精读被删除页面的细节; 若团队仍依赖 ROCm 自建镜像或特征矩阵信息, 建议后续在安装指南或高级主题中补回这些内容。

功能与动机

PR body 明确说明: "Two places describing the same thing is two places to keep in sync, and they had already drifted (Platforms said B100, Installation said B100 / B200, the welcome page said something else again)." 即三个页面描述同一份硬件状态却彼此不一致, 删除重复来源是为了让文档站只有一处需要维护的硬件支持说明。

实现拆解

1. 删除重复页面: 移除 docs/platforms/index.md (含 per-GPU 特征矩阵)、docs/platforms/nvidia.md (NVIDIA 镜像、NCCL/NVLink 排障、环境变量) 和 docs/platforms/amd.md (ROCm 镜像与从源码构建配方), 共约 231 行。
2. 调整站点配置: 在 docs/docs.json 中删除 "Platforms" 导航 tab, 并在 redirects 数组新增三个永久重定向, 将 /platforms/index、/platforms/nvidia、/platforms/amd 统一指向 /getting-started/installation, 保证旧书签与站外链接不失效。
3. 刷新首页硬件列表: docs/index.md 的 Supported hardware 段落中, NVIDIA 行移除不受支持的 B100、补入 B300; 首段“Verified on multiple hardware generations”同步补充 A100; 原指向 /platforms/index 的链接改为指向安装指南的 #hardware-requirements 锚点。
4. 收口安装页硬件表: docs/getting-started/installation.md 中指向 /platforms/index 的链接改为本页自引用锚点, 并将硬件状态表按代际重排: GB300/GB200/B300/B200 为 Production, H200/H100 为 Production (CI guarded), A100 保持 Supported, AMD 行不变, 使该表成为全站权威来源。
5. 链接与配置校验: PR 自述 docs.json 可解析、79 个导航条目全部解析到文件、所有重定向目标存在、两个 #hardware-requirements 锚点有效、无死链。

关键文件:

- docs/docs.json (模块 导航配置; 类别 config; 类型 configuration) : 站点导航与重定向的核心配置: 删除 Platforms tab, 并新增三个永久重定向, 是本次文档目录结构调整的契约所在。
- docs/platforms/index.md (模块 硬件文档; 类别 docs; 类型 deletion) : 被删除的 Platforms 总览页, 包含唯一的 per-GPU 特征矩阵 (BF16/FP8/INT4 QAT/ 推测解码等), 删除后大部分内容没有迁移到别处。
- docs/platforms/nvidia.md (模块 NVIDIA 说明; 类别 docs; 类型 deletion) : 被删除的 NVIDIA 硬件页, 包含镜像内容说明、NCCL/NVLink 排障、环境变量等大量实操细节, 是删除文件中信息最密集的一份。
- docs/platforms/amd.md (模块 AMD 说明; 类别 docs; 类型 deletion) : 被删除的 AMD/ROCm 平台页, 包含 ROCm 容器镜像与从源码构建配方 (Dockerfile.rocm_MI350-5), 该构建配方删除后没有替代归宿。
- docs/index.md (模块 文档首页; 类别 docs; 类型 documentation) : 首页硬件列表刷新, NVIDIA 行去掉 B100、补入 B300, 并把 Platforms 链接改为安装指南锚点, 是本次对外可见度最高的改动。
- docs/getting-started/installation.md (模块 安装指南; 类别 docs; 类型 documentation) : 安装页承接了原 Platforms 的硬件状态职责, 成为全站硬件支持状态的唯一权威来源, 并自引用 #hardware-requirements 锚点。

关键符号: 未识别

关键源码片段

docs/docs.json

站点导航与重定向的核心配置: 删除 Platforms tab, 并新增三个永久重定向, 是本次文档目录结构调整的契约所在。

```
// 节选 docs/docs.json 的 redirects 数组 (为便于阅读添加注释, 实际文件为严格 JSON 不含注释) { "redirects": [ // 既有重定向: DeepSeek V3 页面替换为 V3.2 时遗留
{"source":"/models/deepseek/deepseek","destination":"/models/deepseek/deepseek-v3-2","permanent":true}, // 本次新增: 旧 Platforms 三个路径统一指向安装指南, 保证站外链接与旧书签不失效
{"source":"/platforms/index","destination":"/getting-started/installation","permanent":true}, {"source":"/platforms/nvidia","destination":"/getting-started/installation","permanent":true}, {"source":"/platforms/amd","destination":"/getting-started/installation","permanent":true} ] }
重定向的语义: 被删除页面不再有独立内容, 统一收敛到安装指南; permanent 设为 true 表明这是长期结构变化而非临时迁移。
```

评论区精华

该 PR 无 review 评论, 唯一审核动作是 yueming-yuan 的 APPROVED (空评论)。值得记录的是 PR body 中作者自提的问题: per-GPU 特征矩阵中除 FP8 部分外 (已存在于 advanced/fp8-low-precision.md), 其余内容 (如 R3、P2P 权重传输、Megatron CP、确定性推理支持情况) 以及 ROCm 从源码构建配方 (docker build -f Dockerfile.rocm_MI350-5) 没有其他归宿, 作者询问是否应并入安装页, 最终合并时未并入。

- 暂无高价值评论线程

风险与影响

- 风险:

1. 内容丢失风险: docs/platforms/index.md 的 per-GPU 特征矩阵 (BF16/FP8 GEMM/INT4 QAT/ 推测解码等) 除 FP8 外未迁移到任何页面; amd.md 的 ROCm 从源码构建配方 (Dockerfile.rocm_MI350-5) 也被直接删除, 依赖该配方自建镜像的 AMD 用户将失去指引。
2. 重定向依赖: 三个旧路径全部依赖 docs/docs.json 的 redirects 配置, 若发布时配置解析失败或未生效, 外部链接会直接 404; PR 虽声称已校验, 但没有 CI 或测试证据。
3. 残留双源漂移: 硬件支持状态在 docs/index.md 和 docs/getting-started/installation.md 仍各有一份, 数目虽从三个减到两个, 但未来仍可能再次漂移。 - 影响: 对用户: 文档站点导航减少一个 tab, 硬件支持与容器镜像信息统一在安装指南查看, 旧书签与站外 /platforms/* 链接依赖永久重定向继续可用。对系统: 纯文档变更, 不涉及代码、CI、镜像或运行逻辑, 风险面限于文档站点本身。对团队: 文档维护点减少一处, 信息权威来源收敛到 installation.md, 但被删除内容若确有外部依赖, 需要靠重定向与后续补充来兜底。 - 风险标记: 删除内容无替代归宿, 重定向依赖配置正确性, 内容存在外部链接依赖

关联脉络

- PR #2391 docs: replace DeepSeek V3/R1 page with a DeepSeek-V3.2 recipe: 同样采用“页面替换 + docs.json 永久重定向”的模式, 是本次 redirects 用法直接参照的先例。
- PR #2374 docs: split FAQ out of Resources and link the blog to LMSYS: 同样调整 docs.json 的导航 tab 结构, 说明该文件是文档站结构的核心契约, 近期被反复演进。
- PR #2400 docs: fix SEO gaps across the docs site: 与本次改动同一配置文件 docs.json, 且涉及页面索引与可见性, 属于文档站结构维护的同一脉络。
- PR #2271 docs: refresh the homepage supported-models table: 与本次首页硬件列表刷新同属首页信息维护节奏, 体现 supported-* 列表的持续治理。