

PR #26485 完整报告

sgl-project/sglang

[UnifiedTree]: Update Unified Radix Cache README

合并时间: 2026-05-27 21:50

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/26485>

执行摘要

- 一句话: 更新 Unified Radix Cache README
- 推荐动作: 建议合并, 文档更新有助于代码维护。可忽略 review 建议中的格式问题, 因为不影响理解。

功能与动机

README 文档需要同步最新的代码变更, 包括组件类型枚举、Full 组件的 eviction 方式等, 避免文档与实现脱节。

实现拆解

修改 `python/sglang/srt/mem_cache/unified_cache_components/README.md` 文件:

1. 更新组件映射: 将 `{name → TreeComponent}` 改为 `{ComponentType → TreeComponent}`, `lru_lists` 同理。
2. 更新 `UnifiedTreeNode` 示例: 使用 `ComponentType.FULL` 等枚举访问。
3. 更新 `UnifiedLRUList` 描述: 说明 Auxiliary 组件使用 LRU, Full 组件使用设备 / 主机叶子集。
4. 更新 `release_component_lock` 钩子描述格式, 保持与 `acquire_component_lock` 一致 (使用冒号和句号)。

关键文件:

- `python/sglang/srt/mem_cache/unified_cache_components/README.md` (模块文档; 类别 docs; 类型 documentation; 符号 `create_unified_radix_cache`): 唯一修改的文件, 更新文档以匹配近期架构变更

关键符号: `create_unified_radix_cache`

评论区精华

仅有一条来自 `gemini-code-assist[bot]` 的评论, 建议 `release_component_lock` 描述中组件名后使用冒号并以句号结尾, 以保持与 `acquire_component_lock` 一致。该建议未被采纳 (PR 已合并, 但未应用该建议)。

- 格式一致性: `release_component_lock` 描述 (style): 未采纳, PR 已合并。

风险与影响

- 风险：该 PR 仅为文档更新，不涉及代码变更，因此无技术风险。
- 影响：对用户：文档更准确，开发者能正确理解组件架构。 对系统：无影响。 对团队：需确保后续代码变更同步更新文档。
- 风险标记：暂无

关联脉络

- PR #26387 Support KV events for UnifiedRadixCache: 该 PR 是 UnifiedRadixCache 的近期变更，本 PR 为其更新文档