PR #2813 完整报告

radixark/miles

feat(ci): add user allowlists for command tiers

合并时间：2026-08-31 11:33

原文链接：http://prhub.com.cn/radixark/miles/pull/2813

执行摘要

- 一句话：CI 命令网关新增用户 ID 白名单，种子授权 4 位贡献者
- 推荐动作：值得精读，尤其适合管理开源仓库 CI 命令权限的团队。两个设计决策值得借鉴：①授权身份绑定 GitHub 稳定数字 ID 而非 login，避免用户名变更导致权限漂移；②每个 tier 独立白名单，且 /clear-labels 这类涉及 CI 策略的操作始终要求实时 write/admin，体现最小权限原则。建议后续补强 v4 的 users 条目 schema 校验与缺键错误测试，弥补删除 _validate_user_ids 后的校验空档。

功能与动机

PR body 明确说明动机：'The label-command tier and prior-contributor rerun tier currently have fixed repository-permission and author-association gates. Maintainers need to grant a small number of contributors one of those existing capabilities without granting repository write access or broadening the other tier.' 即现有网关只依赖仓库权限与作者关联两扇门，维护者希望对少数贡献者精确授予某一层命令能力，而不授予仓库 write 权限、也不扩大其他 tier 的门槛。

实现拆解

1. 策略 schema 升级 v3→v4：.github/workflows/policies/comment-command-access.json 将 user_ids: [] 替换为 users: [{"id": ..., "login": ...}] 对象数组；add_label_access 与 prior_contributor_access 各持独立 users 白名单，repo_write_access 保持无白名单。种子 4 位贡献者（zyzshishui、nanjiangwill、xiuhu17、zianglih）到 label 层。
2. loader 改造：.github/workflows/scripts/comment_ci_command.py 的 load_policy 删除原 _validate_user_ids（正整数数组校验），改为内联 frozenset(user["id"] for user in raw_group["users"])；版本校验从 3 改为 4；expected_keys 按组区分，users 键同时允许出现在两个可定制 tier。命令注册表中 rerun_failed_ci / run_test_file 的对应标志位从 False 改为 True，使 prior-contributor 白名单参与这两个 rerun 命令的授权判断；clear_labels 维持 repo_write_access 严格门禁。
3. 文档契约：docs/ci/05-command-identity.md 重写 Access groups 段落，明确 users[].id 是唯一授权身份、login 可过期、维护者需保证唯一正整数 id、外部动态策略超出范围。
4. 测试配套：tests/ci/test/test_comment_ci_command.py 新增 8 个测试，覆盖双白名单独立加载、login 仅展示不影响授权、tier 隔离（prior-contributor 白名单不授予 label 命令、

repo-write 命令不能用带白名单的组）、用户 ID 免权限查询放行并直接分发文件 run 等场景，共 284 个测试通过。

关键文件：

- .github/workflows/scripts/comment_ci_command.py（模块 命令网关；类别 infra；类型 infrastructure；符号 load_policy, _validate_user_ids）：评论命令网关授权核心：load_policy 从 schema v3 升级到 v4，新增 users 白名单解析路径，并调整命令注册表使两个 rerun 命令接入 prior-contributor 白名单。
- tests/ci/test/test_comment_ci_command.py（模块 命令网关；类别 test；类型 test-coverage；符号 test_policy_parser_loads_independent_users_for_both_customizable_tiers, test_user_login_is_display_only_for_authorization, test_non_label_commands_cannot_use_a_group_with_explicit_user_ids, test_repo_write_command_cannot_use_a_group_with_explicit_user_ids）：本次变更的主要验证载体：284 个测试覆盖双白名单独立加载、login 仅展示、tier 隔离、默认门禁不变，以及用户 ID 免权限查询直接分发文件 run。
- .github/workflows/policies/comment-command-access.json（模块 权限策略；类别 infra；类型 infrastructure）：实际落地配置：schema v4 + 种子 4 位贡献者的 id/login 对，prior_contributor_access 补空 users 数组，repo_write_access 无白名单。
- docs/ci/05-command-identity.md（模块 CI 文档；类别 docs；类型 documentation）：定义维护契约：users[].id 为唯一正整数、login 可过期仅展示、外部动态策略超出范围，并同步各 tier 准入条件说明。

关键符号：load_policy, _validate_user_ids（移除）, test_policy_parser_loads_independent_users_for_both_customizable_tiers, test_user_login_is_display_only_for_authorization, test_prior_contributor_access_user_id_preflight_does_not_require_repository_permission, test_prior_contributor_access_user_id_dispatches_a_file_run

关键源码片段

.github/workflows/scripts/comment_ci_command.py

评论命令网关授权核心：load_policy 从 schema v3 升级到 v4，新增 users 白名单解析路径，并调整命令注册表使两个 rerun 命令接入 prior-contributor 白名单。

```
def load_policy(path):
    """加载并校验评论命令权限策略，当前仅接受 schema v4。"""
    raw = load_json(path)
    if not isinstance(raw, dict) or set(raw) != {"version", "groups", "commands"}:
        raise CommentCommandError("policy must contain only version, groups, and commands")
    if type(raw["version"]) is not int or raw["version"] != 4:
        raise CommentCommandError("policy version must be 4")

    groups = {}
    for name, raw_group in raw["groups"].items():
        # add_label_access 与 prior_contributor_access 各自持有独立的 users 白名单；
        # repo_write_access 不允许白名单，/clear-labels 永远要求实时 write/admin。
```

```
        expected_keys = {"repository_permissions"}
        if name in {"add_label_access", "prior_contributor_access"}:
            expected_keys.add("users")
        if name == "prior_contributor_access":
            expected_keys.add("author_associations")
        if not isinstance(raw_group, dict) or set(raw_group) != expected_keys:
            raise CommentCommandError(f"invalid policy fields for groups.{name}")

        groups[name] = {
            "repository_permissions": _validate_permissions(
                f"groups.{name}.repository_permissions",
                raw_group["repository_permissions"],
            ),
            # 授权身份只取稳定数字 id；login 是展示注解，不参与匹配。
            # 注意：若配置项缺 id 键会在此抛 KeyError，属于 fail-closed 还是
            # 500 取决于上层异常处理——这是新 schema 校验弱化后需要留意的点。
            "user_ids": (
                frozenset(user["id"] for user in raw_group["users"])
                if "users" in raw_group
                else frozenset()
            ),
            "author_associations": (
                _validate_author_associations(
                    f"groups.{name}.author_associations",
                    raw_group["author_associations"],
                )
                if "author_associations" in raw_group
                else frozenset()
            ),
        }

    # 后续继续解析 commands 段，将每个命令绑定到已构建的 group 引用，
    # 并校验命令只能使用注册表内的 group（默认拒绝）。
    ...
```

tests/ci/test/test_comment_ci_command.py

本次变更的主要验证载体：284 个测试覆盖双白名单独立加载、login 仅展示、tier 隔离、默认门禁不变，以及用户 ID 免权限查询直接分发文件 run。

```
@pytest.mark.parametrize(
    ("group", "body"),
    [
        ("add_label_access", "/run-ci-short"),
        ("prior_contributor_access", RUN_FILE_BODY),
    ],
)
def test_user_login_is_display_only_for_authorization(tmp_path, group, body):
    """验证授权只依赖稳定数字 id，login 过期也不会改变访问结果。"""
    raw = raw_policy()
```

```
    # 故意写入已过期的 stale-login，模拟用户改名后策略文件未同步
    raw["groups"][group]["users"] = [{"id": ACTOR_ID, "login": "stale-login"}]
    path = tmp_path / "policy.json"
    path.write_text(json.dumps(raw))
    # 构造一个无任何仓库权限、且作者关联为 FIRST_TIMER 的请求
    api = FakeAPI(pull(), permission="none")

    result = HANDLER.authorize_policy(
        event(body=body, author_association="FIRST_TIMER"),
        HANDLER.load_policy(path),
        api,
    )

    # 放行：返回 (pull_number, actor_id, request_type)，且全程未查实时权限
    assert result[0:2] == (123, ACTOR_ID)
    assert api.calls == []
```

.github/workflows/policies/comment-command-access.json

实际落地配置：schema v4 + 种子 4 位贡献者的 id/login 对，prior_contributor_access 补空 users 数组，repo_write_access 无白名单。

```
{
  "version": 4, // schema v3 -> v4：user_ids 数组升级为 users 对象数组
  "groups": {
    "add_label_access": {
      "repository_permissions": ["write", "admin"],
      "users": [ // 仅 label 命令使用的白名单，种子 4 位贡献者
        {"id": 82826991, "login": "zyzshishui"},
        {"id": 59716405, "login": "nanjiangwill"},
        {"id": 101526713, "login": "xiuhu17"},
        {"id": 106564213, "login": "zianglih"}
      ]
    },
    "repo_write_access": { // /clear-labels 专属：永远不允许白名单
      "repository_permissions": ["write", "admin"]
    },
    "prior_contributor_access": {
      "repository_permissions": ["write", "admin"],
      "users": [], // rerun 命令白名单，当前未种子，未来按需加入
      "author_associations": ["OWNER", "MEMBER", "COLLABORATOR", "CONTRIBUTOR"]
    }
  }
}
```

评论区精华

本 PR 没有内联 review 评论，claude[bot] 仅提示该仓库配置为手动 review；两位维护者 yushengsu-thu 与 yueming-yuan 直接 APPROVED，未留下具体意见。实质设计权衡写在 PR body 的 Design Notes 中：策略文件拥有 tier 例外，网关绑定事件作者的稳定数字 ID、

选择命令组、在 users[].id 命中或默认门禁通过时放行；login 不参与授权。决策结论是 schema v4 将 checked-in JSON 视为可信配置，外部动态策略明确 out of scope。

- 暂无高价值评论线程

风险与影响

- 风险：
  1. 校验弱化：删除 _validate_user_ids 后，v4 不再校验 users 条目中的 id 是否为正整数、是否重复，也不再校验 login 非空；若配置缺 id 键会直接抛 KeyError，依赖上层异常处理决定是 fail-closed 还是 500，建议后续补充显式 schema 校验（如 Pydantic）及缺省错误测试。
  2. 权限面扩大：prior_contributor_access 新增用户 ID 授权路径，使 rerun_failed_ci 与 run_test_file 可被白名单用户触发；虽然 rerun 只重放已授权运行，但仍消耗 runner 时间，需要维持白名单纪律。
  3. 网关全局生效：comment_ci_command.py 是所有 PR 评论命令的统一入口，schema 变更影响每次授权判断；好在默认拒绝与 fail-closed 语义保留，旧 v3 配置会被版本校验拒绝（policy version must be 4），本 PR 已同步升级仓库内配置。
  4. login 展示语义：文档允许 login 与 GitHub 实际登录名不一致，用户改名后策略文件看起来可能“过期”，无安全影响但需在维护文档中说明。 - 影响：对维护者：从此可以只授予某个贡献者 label 命令或 rerun 命令权限而不给仓库 write，细化 CI 命令的授权粒度。对贡献者：种子名单中的 4 人获得 run-ci-* 标签操作权，未来可通过编辑 comment-command-access.json 扩展。对系统：网关的授权路径从“仓库权限 + 作者关联”扩展为“仓库权限 + 作者关联 + 显式用户 ID 白名单”，/clear-labels 仍保持最高门禁。对团队：CI 权限治理从二值权限走向细粒度白名单，为后续更多命令层级（如文档命令、自定义 label）铺路。 - 风险标记：CI 权限模型变更，schema v4 校验弱化，权限面扩大，网关全局生效

关联脉络

- PR #2720 fix(ci): grant pull-requests write to comment-gateway feedback jobs: 同一 comment gateway 的能力进化（评论反馈权限），与本 PR 共同完善评论命令网关的运维与授权模型。
- PR #2773 ci: default a file run to the PR's own image: /rerun-test 文件 run 的执行逻辑相关，本 PR 为 run_test_file 新增用户 ID 白名单授权路径。