PR #47010 完整报告

vllm-project/vllm

fix(security): prevent image decompression bomb OOM denial of service

合并时间：2026-06-30 17:39

原文链接：http://prhub.com.cn/vllm-project/vllm/pull/47010

执行摘要

- 一句话：添加图像像素限制防止解压炸弹 OOM
- 推荐动作：值得精读。设计上在加载前预检、环境变量控制、避免全局状态突变，是良好的安全实践，可推广到其他媒体类型。

功能与动机

恶意或意外的大尺寸压缩图像可膨胀消耗数 GB 内存，导致 OOM。本 PR 在解码前检查像素数，避免光栅内存分配，并提供可配置上限。

实现拆解

1. 环境变量注册：在 vllm/envs.py 中添加 VLLM_MAX_IMAGE_PIXELS 变量及其默认值（178,956,970，匹配 PIL 的 2x 解压炸弹阈值）。
2. 图像检查：在 vllm/multimodal/media/image.py 的 ImageMediaIO.load_bytes() 中，在 image.load() 前检查像素数，超出则抛 ValueError。
3. 视频检查：在 vllm/multimodal/video.py 中添加 _check_frame_pixel_limit() 函数，并在 opencv、PyNvVideoCodec、av 等解码路径中调用。
4. 移除全局绕过：在 nemotron_vl.py 和 nano_nemotron_vl.py 中删除 Image.MAX_IMAGE_PIXELS = None 的全局设置，避免绕过 PIL 内建检查。
5. 测试与文档：新增 3 个单元测试覆盖正常、超标、禁用场景；更新 docs/usage/security.md 说明。

关键文件：

- vllm/multimodal/media/image.py（模块 多模态；类别 source；类型 core-logic；符号 load_bytes）：核心图像加载逻辑，添加像素预检，拒绝超大图片
- vllm/multimodal/video.py（模块 多模态；类别 source；类型 core-logic；符号 _check_frame_pixel_limit）：视频帧像素检查，覆盖所有解码后端
- tests/multimodal/media/test_image.py（模块 测试；类别 test；类型 test-coverage；符号 test_image_pixel_limit_respected, test_image_pixel_limit_rejected, test_image_pixel_limit_disabled）：新增三个测试覆盖像素限制场景
- vllm/envs.py（模块 环境配置；类别 source；类型 configuration）：新增 VLLM_MAX_IMAGE_PIXELS 环境变量

- vllm/transformers_utils/processors/nano_nemotron_vl.py（模块 处理器；类别 source；类型 core-logic）：移除全局 Image.MAX_IMAGE_PIXELS = None 设置
- vllm/transformers_utils/processors/nemotron_vl.py（模块 处理器；类别 source；类型 core-logic）：移除全局 Image.MAX_IMAGE_PIXELS = None 设置
- docs/usage/security.md（模块 文档；类别 docs；类型 documentation）：新增像素限制配置说明

关键符号：ImageMediaIO.load_bytes, _check_frame_pixel_limit, test_image_pixel_limit_respected, test_image_pixel_limit_rejected, test_image_pixel_limit_disabled

关键源码片段

vllm/multimodal/media/image.py

核心图像加载逻辑，添加像素预检，拒绝超大图片

```
# vllm/multimodal/media/image.py

def load_bytes(self, data: bytes) -> MediaWithBytes[Image.Image]:
    try:
        image = Image.open(BytesIO(data))
        w, h = image.size
        max_pixels = envs.VLLM_MAX_IMAGE_PIXELS
        # 如果像素数超过限制，则拒绝加载
        if max_pixels > 0 and w * h > max_pixels:
            raise ValueError(
                f"Image dimensions {w}x{h} ({w * h} pixels) exceed "
                f"the maximum of {max_pixels} pixels. Set "
                f"VLLM_MAX_IMAGE_PIXELS to increase this limit."
            )
        image = normalize_image(image)
        image.load() # 此时才分配光栅内存
        image = self._convert_image_mode(image)
    except (OSError, Image.UnidentifiedImageError) as e:
        raise ValueError(f"Failed to load image: {e}") from e
    return MediaWithBytes(image, data)
```

vllm/multimodal/video.py

视频帧像素检查，覆盖所有解码后端

```
# vllm/multimodal/video.py

def _check_frame_pixel_limit(width: int, height: int) -> None:
    """Reject video frames exceeding VLLM_MAX_IMAGE_PIXELS before decoding."""
    max_pixels = envs.VLLM_MAX_IMAGE_PIXELS
    if max_pixels > 0 and width * height > max_pixels:
        raise ValueError(
            f"Video frame dimensions {width}x{height} "
```

```
            f"({width * height} pixels) exceed the maximum of "
            f"{max_pixels} pixels. Set VLLM_MAX_IMAGE_PIXELS to "
            f"increase this limit."
        )

# 在 opencv 后端调用示例
# decode_frames_opencv 中：
cap = cls.open_video_capture(data)
_check_frame_pixel_limit(
    int(cap.get(cv2.CAP_PROP_FRAME_WIDTH)),
    int(cap.get(cv2.CAP_PROP_FRAME_HEIGHT)),
)
```

评论区精华

1. Nemotron 处理器的全局设置：DarkLight1337 指出 Image.MAX_IMAGE_PIXELS = 300M 是误导性的且不尊重环境变量，建议所有模型统一使用默认值。作者最终完全移除了这些设置，恢复 PIL 内建检查。
2. 避免全局状态突变：DarkLight1337 建议不要在 image.py 中全局修改 Image.MAX_IMAGE_PIXELS，以避免导入 vLLM 影响 PIL 全局行为。作者采纳意见，改为在 load_bytes 中局部检查。

- Nemotron 处理器的全局 MAX_IMAGE_PIXELS 设置是否合适 (design): 接受建议，移除 Nemotron 处理器的全局设置，恢复 PIL 内建检查。
- 是否应全局修改 Image.MAX_IMAGE_PIXELS (design): 不全局修改 Image.MAX_IMAGE_PIXELS，仅在加载路径中检查。
- 视频加载器是否也应遵守像素限制 (question): 已实现，视频帧也会像素检查。

风险与影响

- 风险：影响所有图像 / 视频加载路径，但默认阈值较大（179M 像素，约 680MB RGB），极少有合法图片超过。性能影响极小（仅额外整数比较）。若用户依赖超高分辨率图片（如卫星图），可通过环境变量调整。
- 影响：对用户：默认提供安全防护，可灵活调节；对系统：减少异常 OOM 风险；对团队：统一媒体安全策略，为后续安全增强奠定基础。
- 风险标记：核心媒体加载路径变更，默认值可能影响极端用例

关联脉络

- 暂无明显关联 PR