链滴

nvidia-docker2.0 GPU 隔离

作者：liuxin

原文链接：https://ld246.com/article/1511781062916

来源网站：链滴

通过环境变量来隔离

NVIDIA_VISIBLE_DEVICES

容器可以使用哪些GPU

- 0,1,2 GPU编号，多个逗号隔开
- all: 所有GPU，nvidia官方的镜像默认是这个选项
- none: 没有GPU，但是容器内部会映射GPU驱动
- empty: nvidia-container-runtime will have the same behavior as runc.

NVIDIA_DRIVER_CAPABILITIES

This option controls which driver libraries/binaries will be mounted inside the container.

- compute: required for CUDA and OpenCL applications,
- compat32: required for running 32-bit applications,
- graphics: required for running OpenGL and Vulkan applications,
- utility: required for using nvidia-smi and NVML,
- video: required for using the Video Codec SDK.

还有几个环境变量未做测试，可以点击进入GitHub查看。

几个感兴趣的问题可以参考下：

Is OpenGL supported?

No, OpenGL is not supported at the moment and there is no plan to support OpenGL+GLX in he near future.

OpenGL+EGL however will be supported and this issue will be updated accordingly.

If you are a NGC subscriber and require GLX for your workflow, please fill out a feature request for support consideration.

Do you support CUDA Multi Process Service (a.k.a. MPS)?

No, MPS is not supported at the moment. However we plan on supporting this feature in the uture, and this issue will be updated accordingly.

Do you support running a GPU-accelerated X server inside the contai er?

No, running a X server inside the container is not supported at the moment and there is no pl n to support it in the near future (see also OpenGL support).

Why is nvidia-smi inside the container not listing the running process s?

nvidia-smi and NVML are not compatible with PID namespaces.

We recommend monitoring your processes on the host or inside a container using --pid=host.

Can I limit the GPU resources (e.g. bandwidth, memory, CUDA cores) aken by a container?

No. Your only option is to set the GPU clocks at a lower frequency before starting the contain r.

What do I have to install in my container images?

Library dependencies vary from one application to another. In order to make things easier for developers, we provide a set of official images to base your images on.

Can I use the GPU during a container build (i.e. docker build)?

Yes, as long as you configure your Docker daemon to use the nvidia runtime as the default, y u will be able to have build-time GPU support. However, be aware that this can render your i ages non-portable (see also invalid device function).

The official CUDA images are too big, what do I do?

The devel image tags are large since the CUDA toolkit ships with many libraries, a compiler a d various command-line tools.

As a general rule of thumb, you shouldn’t ship your application with its build-time depende cies. We recommend to use multi-stage builds for this purpose. Your final container image sh uld use our runtime or base images.

As of CUDA 9.0 we now ship a base image tag which bundles the strict minimum of depende cies.

Do you support Kubernetes?

Since Kubernetes 1.8, the recommended way is to use our official device plugin. Note that this is still alpha support.

原文链接：nvidia-docker2.0 GPU 隔离