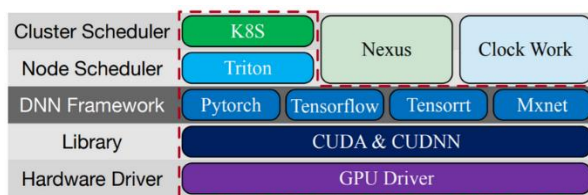


## GPU 调度架构.part2



除了依赖底层的时间片机制，还可以主动控制任务的发射时间进行调度。对于闭源框架需要对 CUDA API 进行拦截



CUDA 11.3 之后，如果我们想要截获 CUDA driver API，首先需要截获 cuGetProcAddress，再由它得到其他 API 的地址。截获的 API 通过一个中心化 server 守护进程(daemon) 进行管理。

## 确定性调度实验平台搭建

Jetson AGX Xavier

(ubuntu 18.04-RT 补丁, cuda 11.1, ROS2/ROS2-PiCAS/CyberRT )

- callback 性能测量（部署只包含通信和回调的 DAG 图，其中的任务用等价强度的 CPU/GPU 计算代替）
- 真实任务流测量（真实计算负载、memory 访问）

