



GOPS2017
Shanghai



GOPS

全球运维大会

2017

上海站

指导单位:  信息中心联盟
SHANGHAI CENTER FOR INFORMATION TECHNOLOGY

主办单位:  信息技术协会
SHANGHAI INFORMATION TECHNOLOGY ASSOCIATION

 网络运维联盟
NETWORK OPERATING ALLIANCE

大会时间: 2017年11月17日-18日

大会地点: 上海光大会展中心国际大酒店 (上海徐汇区漕宝路67号)





GOPS2017
Shanghai

机器学习云

Spectrum

叶璐



GOPS2017
Shanghai

目录

- 1 ▶ 深度学习的兴起
- 2 ▶ 深度学习在Qunar的应用
- 3 ▶ GPU云化解决方案
- 4 ▶ GPU云使用简介



GOPS2017
Shanghai

- 深度学习的兴起



017
hai



AlphaGo



DeepMind



5月27日下午消息，围棋峰会今日进入最后一天，柯洁九段中盘告负，总比分 0 : 3 败于 AlphaGo。比赛中，柯洁在局面不利是长时间离开，回来后又泪洒现场。赛后柯洁一度哽咽称：它太完美我很痛苦，看不到任何胜利的希望。



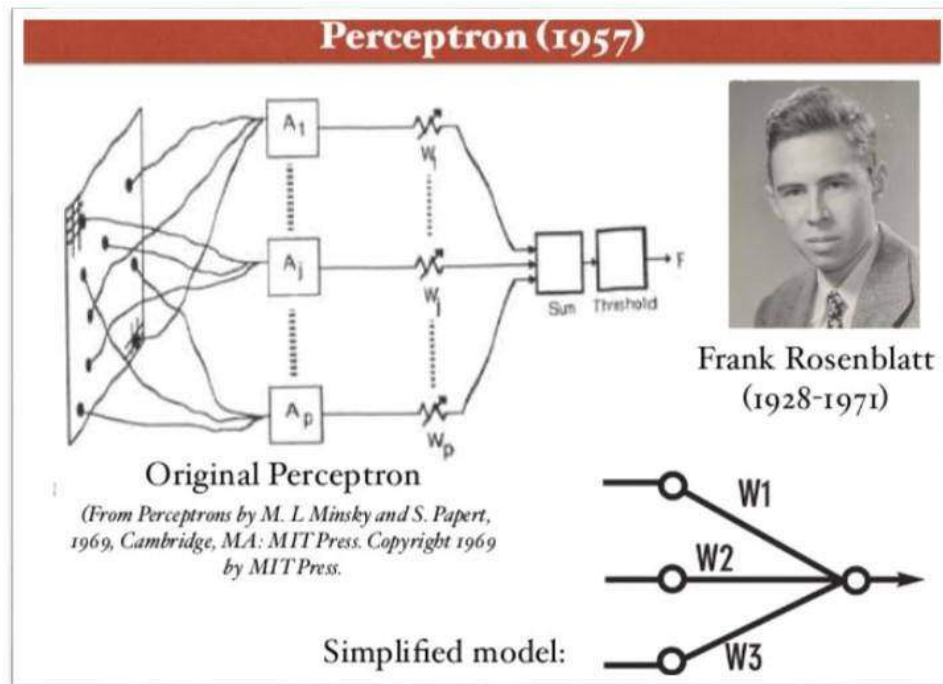
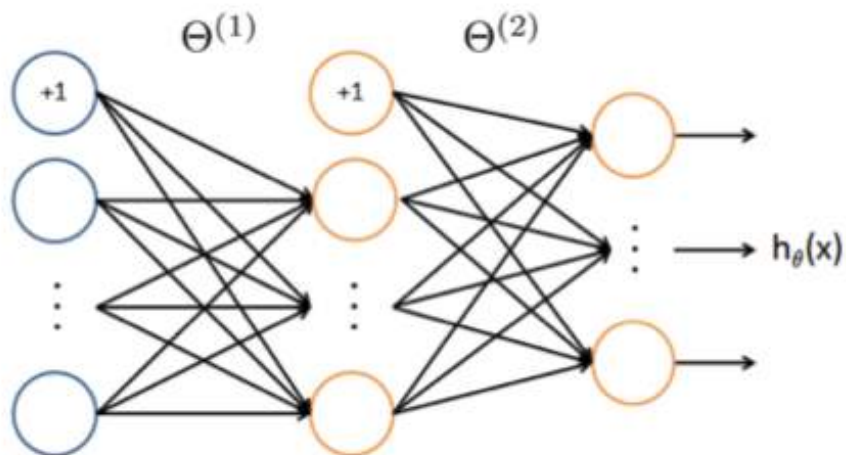
GOPS2017
Shanghai



什么是深度学习



GOPS2017
Shanghai



奠基人



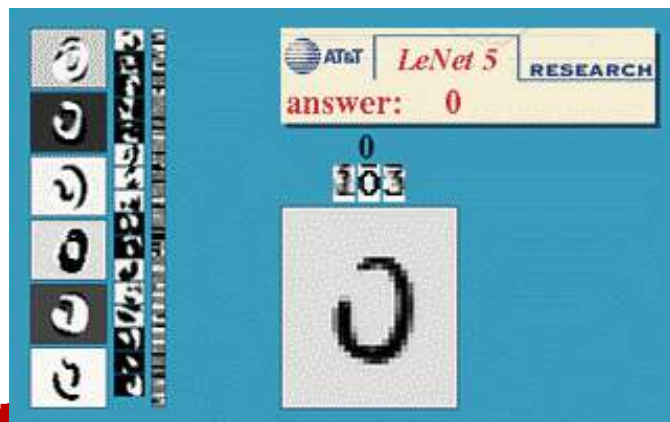
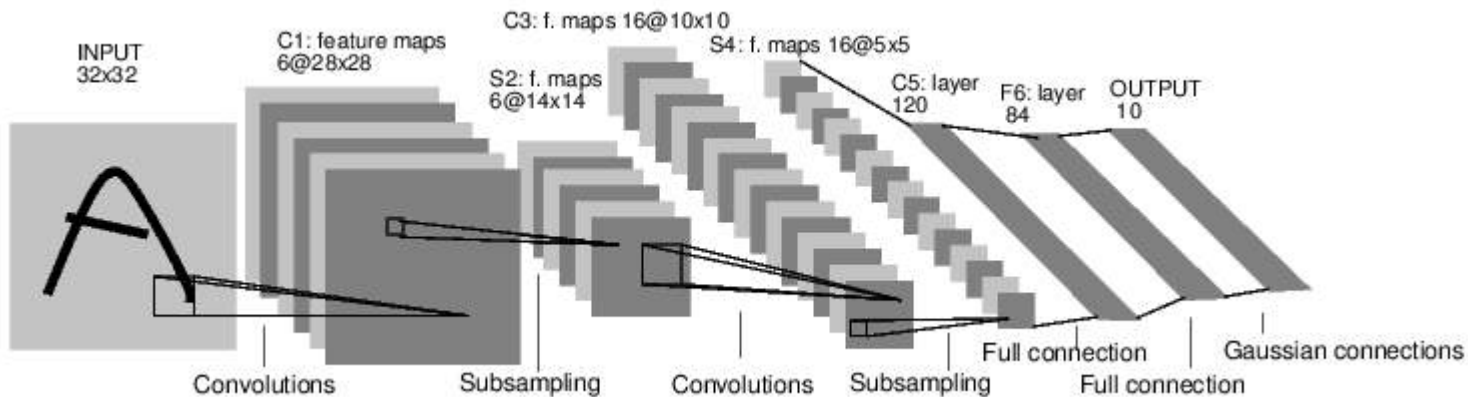
GOPS2017
Shanghai



CNN



GOPS2017
Shanghai



Watson



GOPS2017
Shanghai



Google无人驾驶



GOPS2017
Shanghai



Google

IBM®

Microsoft

facebook

Baidu 百度

Alibaba Group



GOPS2017
Shanghai

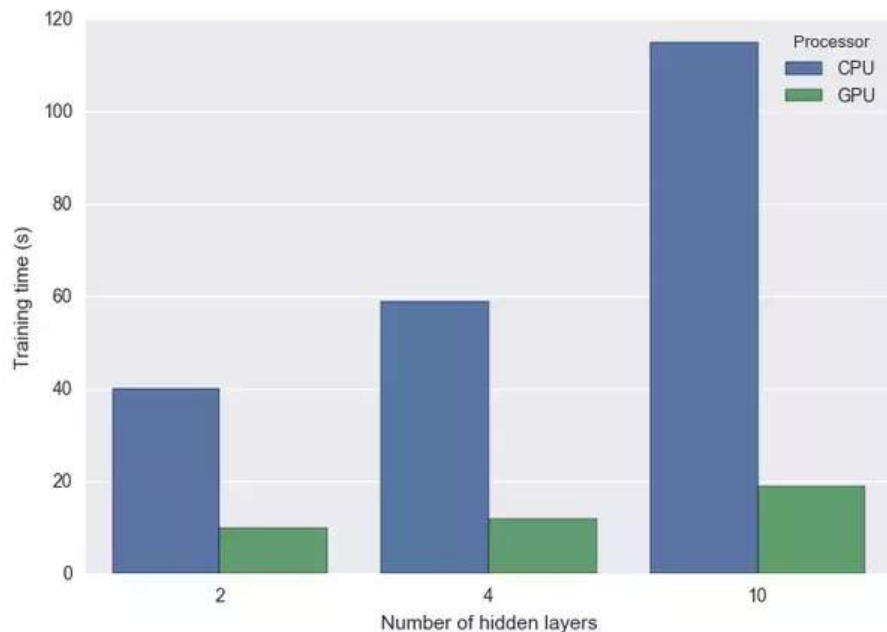
兴起的原因

- 大数据
- GPU等计算资源成本减低
- 开源工具的普及

深度学习加速器-GPU



GOPS2017
Shanghai



Nvidia在股市上表现抢眼



GOPS2017
Shanghai

NVIDIA Corporation

NASDAQ: NVDA - Jun 28, 7:59 PM EDT

151.75 USD ↑ 5.17 (3.53%)

After-hours: 152.50 ↑ 0.49%

1 day

5 day

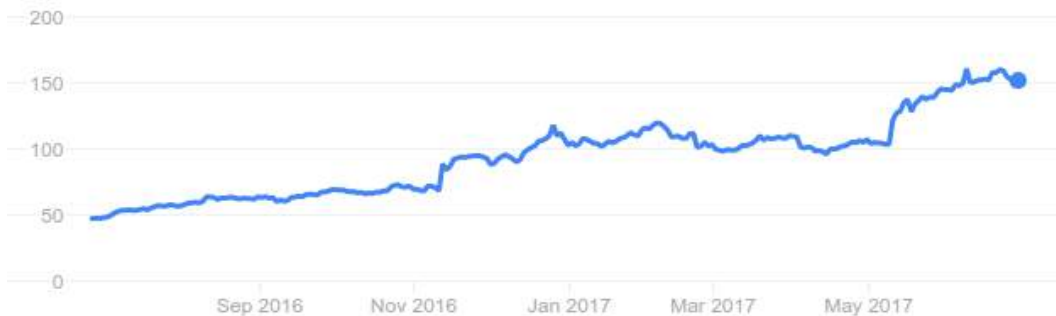
1 month

3 month

1 year

5 year

max



Open 149.32
High 151.94
Low 145.75

Mkt cap 86.99B
P/E ratio 51.38
Div yield 0.37%



GOPS2017
Shanghai

- 深度学习在Qunar的应用



GOPS2017
Shanghai

深度学习在Qunar的应用

- 机票验证码识别等
- 看图写诗-小诗机
- 智能客服
- 拿去花用户信用评级
- 根据历史访问记录推荐酒店
- 计算不同酒店房型的价格系数



GOPS2017
Shanghai

多个team的共享GPU资源

谁把老子跑了三天的任务杀了



采购周期长



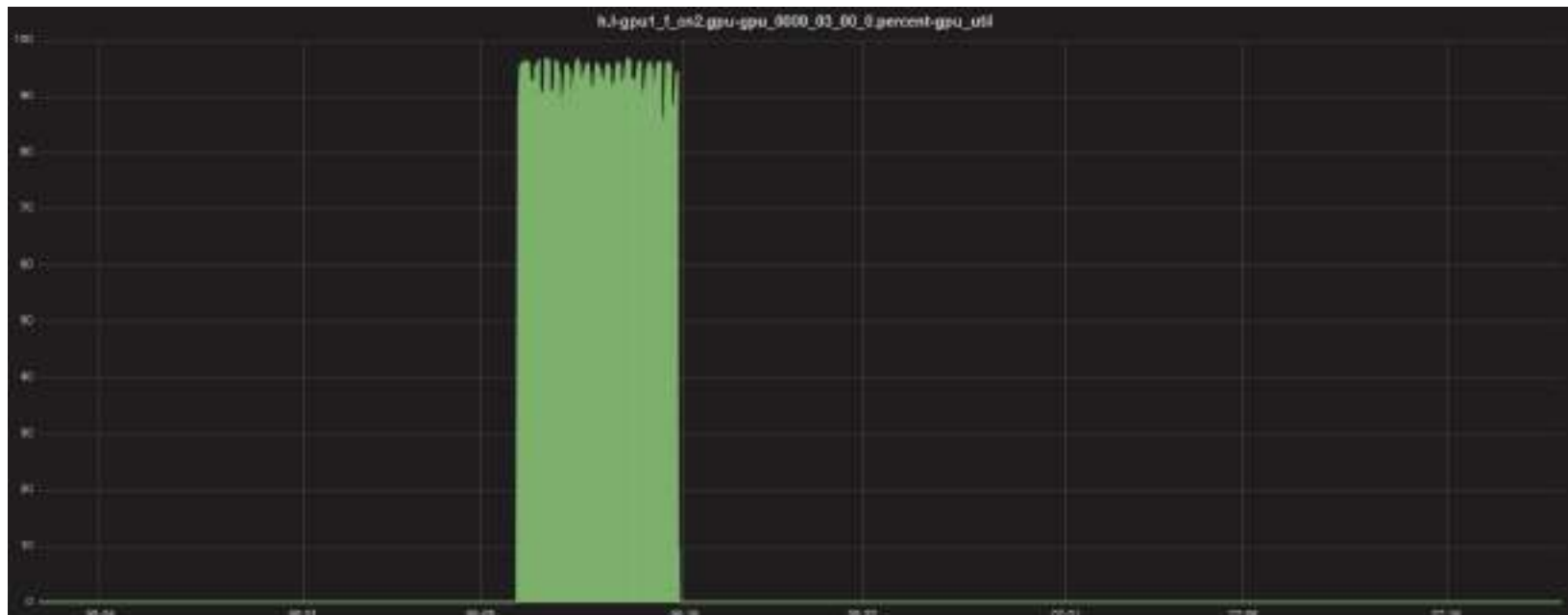
GOPS2017
Shanghai



资源利用率低



GOPS2017
Shanghai





GOPS2017
Shanghai

GPU使用资源的问题

- 环境无隔离
- 采购周期长
- 资源利用率低
- 各种工具的环境部署成本高



GOPS2017
Shanghai

我们该怎么办？

我们的宗旨



GOPS2017
Shanghai

- 降低资源使用的门槛
- 提高资源利用率



GOPS2017
Shanghai

GPU云第一期的目标

- GPU资源云化，支持业务线同学快捷新建机器学习应用秒建秒删，一键释放GPU资源
- 统一GPU资源申请和管理等入口到Portal，降低业务线同学接入和学习成本
- 环境隔离
- 保证训练数据在分布式环境下的持久化和可靠性
- 支持Tensorflow全工具链



GOPS2017
Shanghai

为什么是Tensorflow?

Tensorflow Community vs Others



GOPS2017
Shanghai

ARM



Quantipar

AIRBUS
DEFENCE & SPACE

CIST

CEVA

Google

Movidius



JD.COM 京东

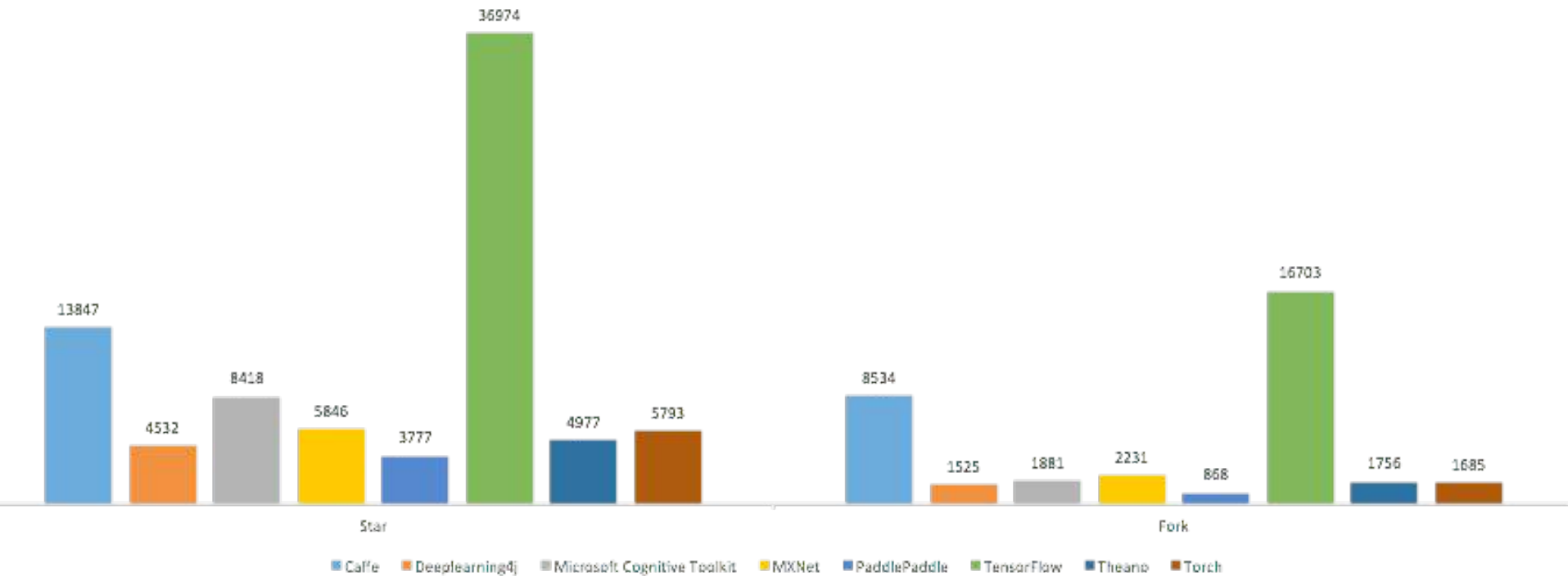


DeepMind

Tensorflow Community vs others



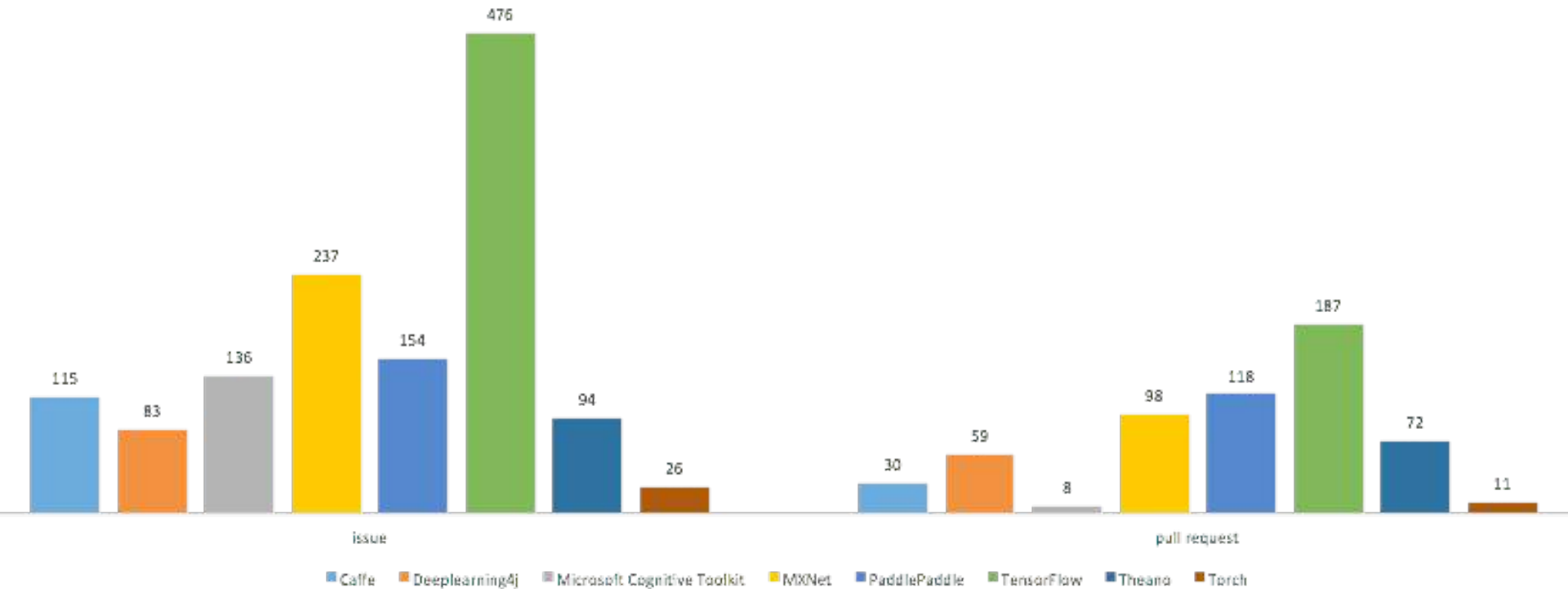
GOPS2017
Shanghai



Tensorflow Community vs others



GOPS2017
Shanghai



Mnist (149)



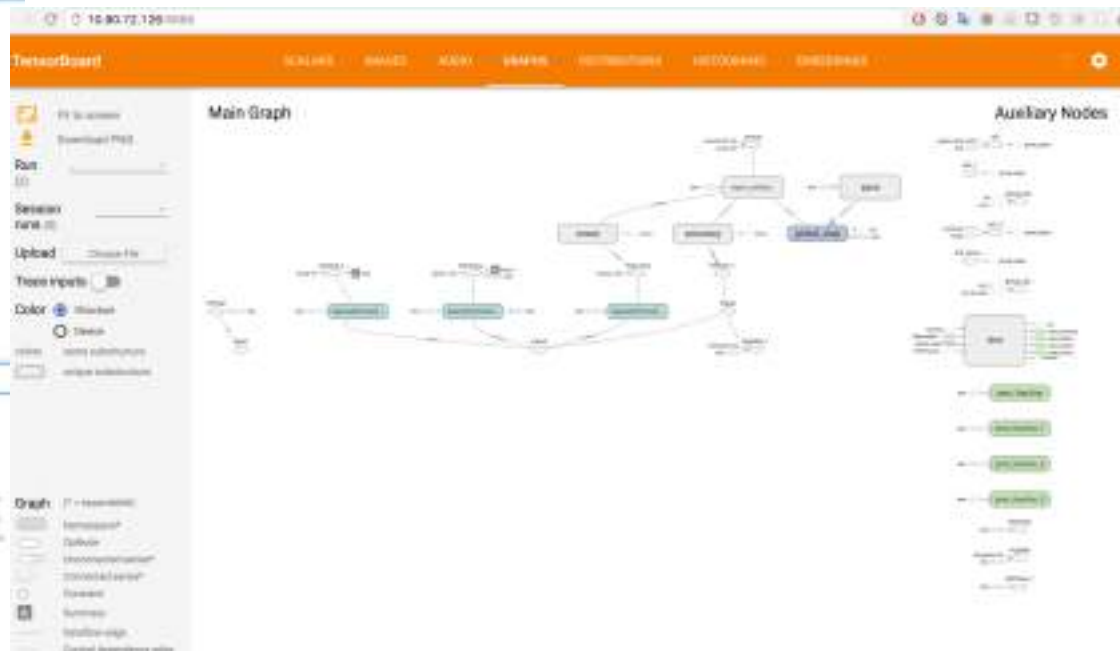
GOPS2017
Shanghai

```
# Hidden 1
with tf.name_scope('hidden1'):
    weights = tf.Variable(
        tf.truncated_normal([IMAGE_PIXELS, hidden1_units],
                            stddev=1.0 / math.sqrt(float(IMAGE_PIXELS))),
        name='weights')
    biases = tf.Variable(tf.zeros([hidden1_units]),
                        name='biases')
    hidden1 = tf.nn.relu(tf.matmul(images, weights) + biases)
# Hidden 2
with tf.name_scope('hidden2'):
    weights = tf.Variable(
        tf.truncated_normal([hidden1_units, hidden2_units],
                            stddev=1.0 / math.sqrt(float(hidden1_units))),
        name='weights')
    biases = tf.Variable(tf.zeros([hidden2_units]),
                        name='biases')
    hidden2 = tf.nn.relu(tf.matmul(hidden1, weights) + biases)
# Linear
with tf.name_scope('softmax_linear'):
    weights = tf.Variable(
        tf.truncated_normal([hidden2_units, NUM_CLASSES],
                            stddev=1.0 / math.sqrt(float(hidden2_units))),
        name='weights')
    biases = tf.Variable(tf.zeros([NUM_CLASSES]),
                        name='biases')
    logits = tf.matmul(hidden2, weights) + biases
return logits
```


• Tensorboard



GOPS2017
Shanghai





GOPS2017
Shanghai

计算资源云化选型

Mesos VS Kubernetes



GOPS2017
Shanghai

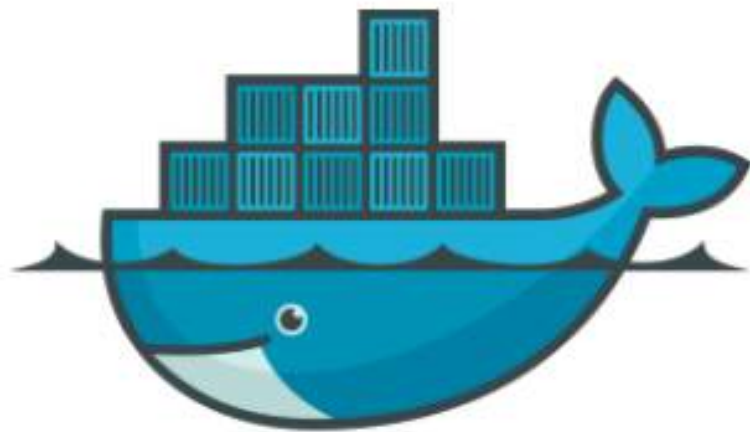


Why Kubernetes?

- Container Orchestration
- Infrastructure-agnostic, e.g. handling GPU resources
- Storage services object storage/block storage, like S3/RBD and so on
- Community support using Tensorflow on top of kubernetes
- Applications and DevOps tooling “couple” to infrastructure



GOPS2017
Shanghai





GOPS2017
Shanghai

架构简述



GOPS2017
Shanghai

如何完成应用定制

OPS Namespace

Kubernetes Deployment



Containerized TF PS

Kubernetes Pod



Containerized TF Worker

Kubernetes Pod



Containerized TF Worker

Kubernetes Pod



Containerized TF Worker

Flight Namespace

Kubernetes App



Containerized TF

Kubernetes App



Containerized TF

Server 1

GPU 1

/dev/nvidia0

GPU 2

/dev/nvidia1

Server 2

GPU 0

/dev/nvidia0

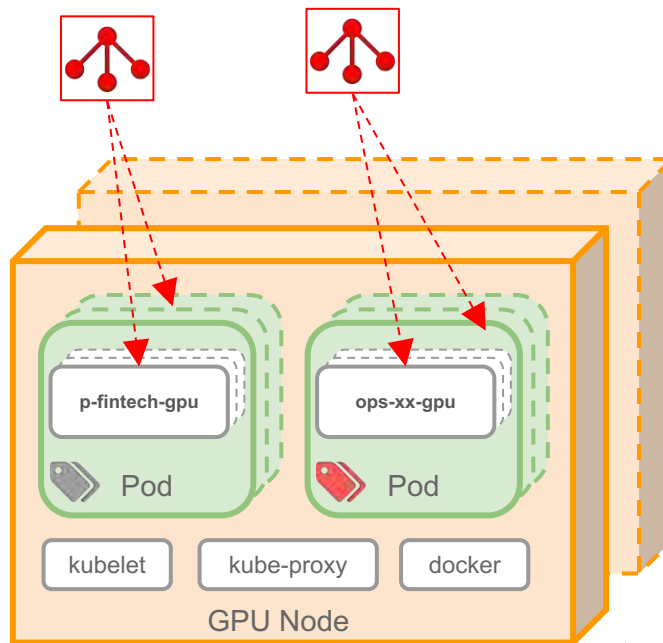
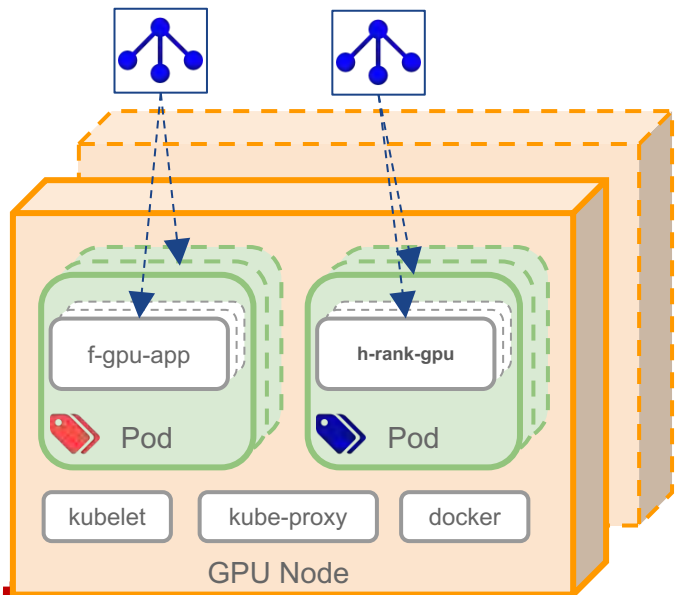
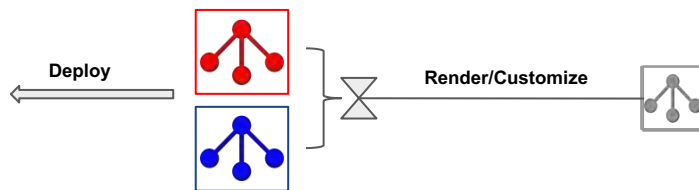
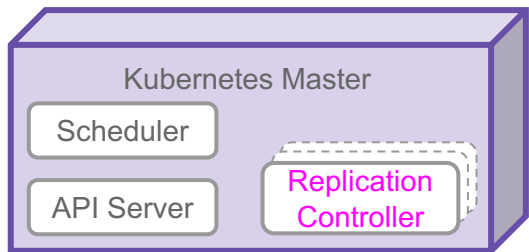
GPU 1

/dev/nvidia1



GOPS2017
Shanghai

Kubernetes Cluster



= Service

= Labels



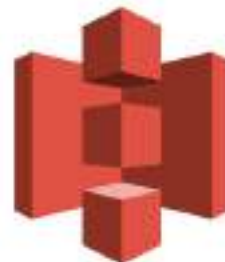
GOPS2017
Shanghai

数据存在哪？

RBD + S3 (Based on Ceph)



GOPS2017
Shanghai





GOPS2017
Shanghai

Jupyter-让算法工程师轻松上手

Jupyter



GOMS2017
Shanghai

Jupyter 3_mnist_from_scratch [username@hostname]

```
File Edit View Insert Cell Help Python 3.5.2
```

```
In [5]: # Let's collect the whole image to 28 bit floats and rescale
# the values to be centered around 0, between [-0.5, 0.5]
#
# We again plot the image and histogram to check that we
# haven't mangled the data
scaled = image.astype('float32')
scaled = (scaled - 255) / 255
_, (ax1, ax2) = plt.subplots(2, 1)
ax1.imshow(scaled, cmap=plt.cm.gray)
ax2.hist(scaled, bins=50, range=[-0.5, 0.5])
```

Great -- we've retained the correct image data while properly rescaling to the range [-0.5, 0.5].

Reading the labels

Let's next unpack the test label data. The format here is similar: a magic number followed by a count followed by the labels as uint8 values. In more detail:

[offset]	[type]	[value]	[description]
8000	32 bit integer	8x00000001(2048)	magic number (MSB first)
8004	32 bit integer	10000	number of items
8008	unsigned byte	??	label

Jupyter



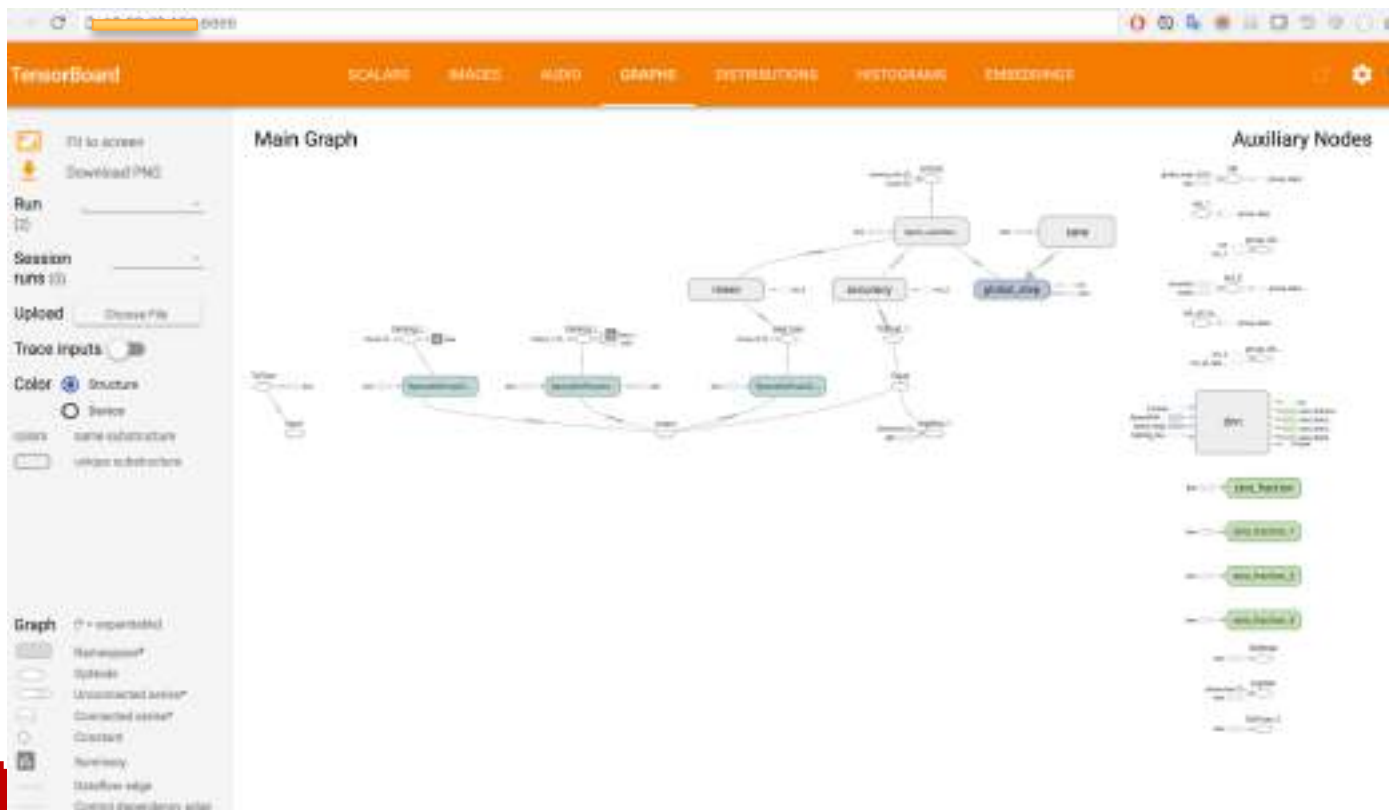
GOMS2017
Shanghai



• Tensorboard



GOPS2017
Shanghai

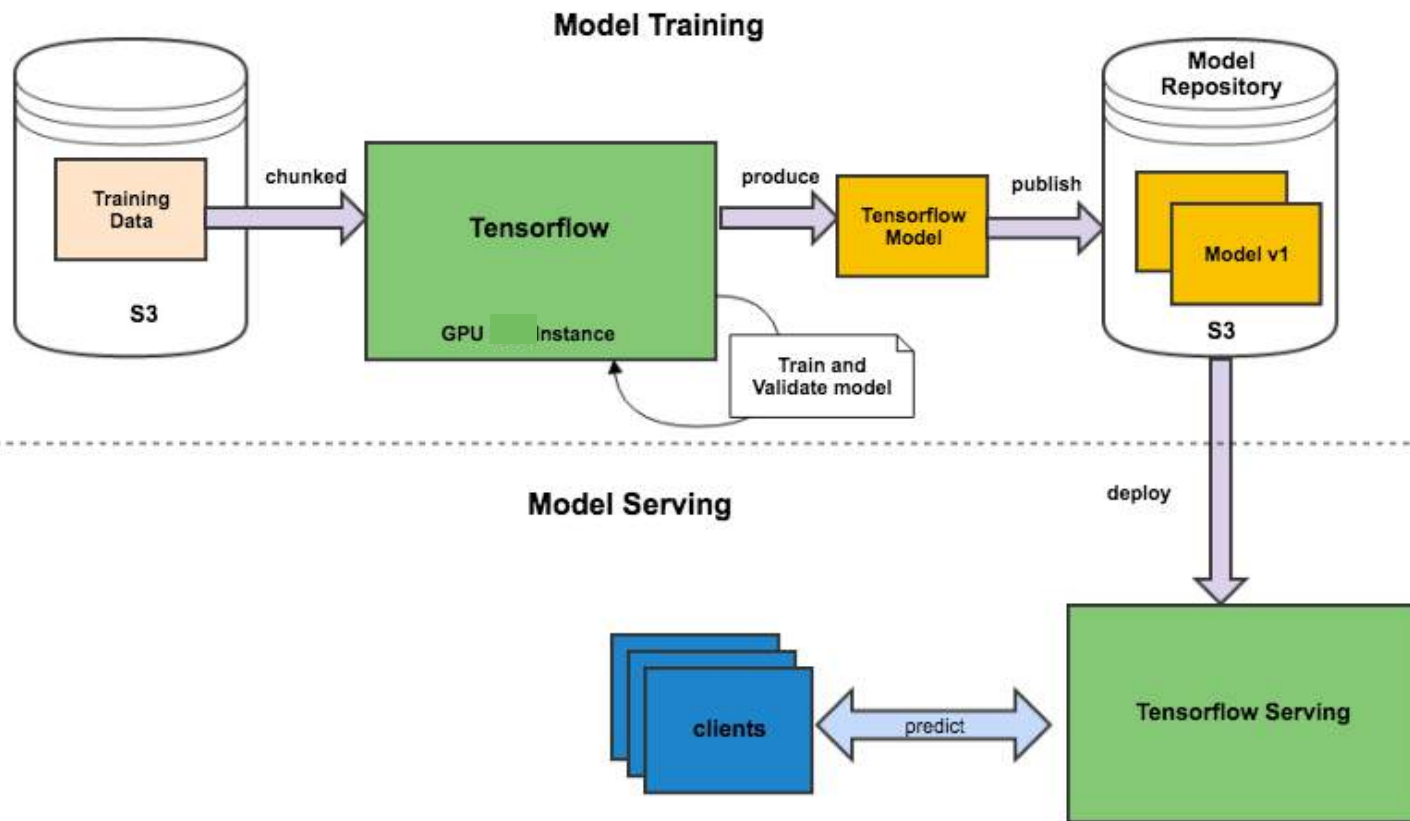




GOPS2017
Shanghai

Add-ons

- ✓ 基础环境固化，随时支持环境固化到post install里
- ✓ Qunar Model Registry & Serving 高可用
- ✓ Jupyter插件系统集成
- ✓ 分布式试用
- 基于events的GPU资源计费





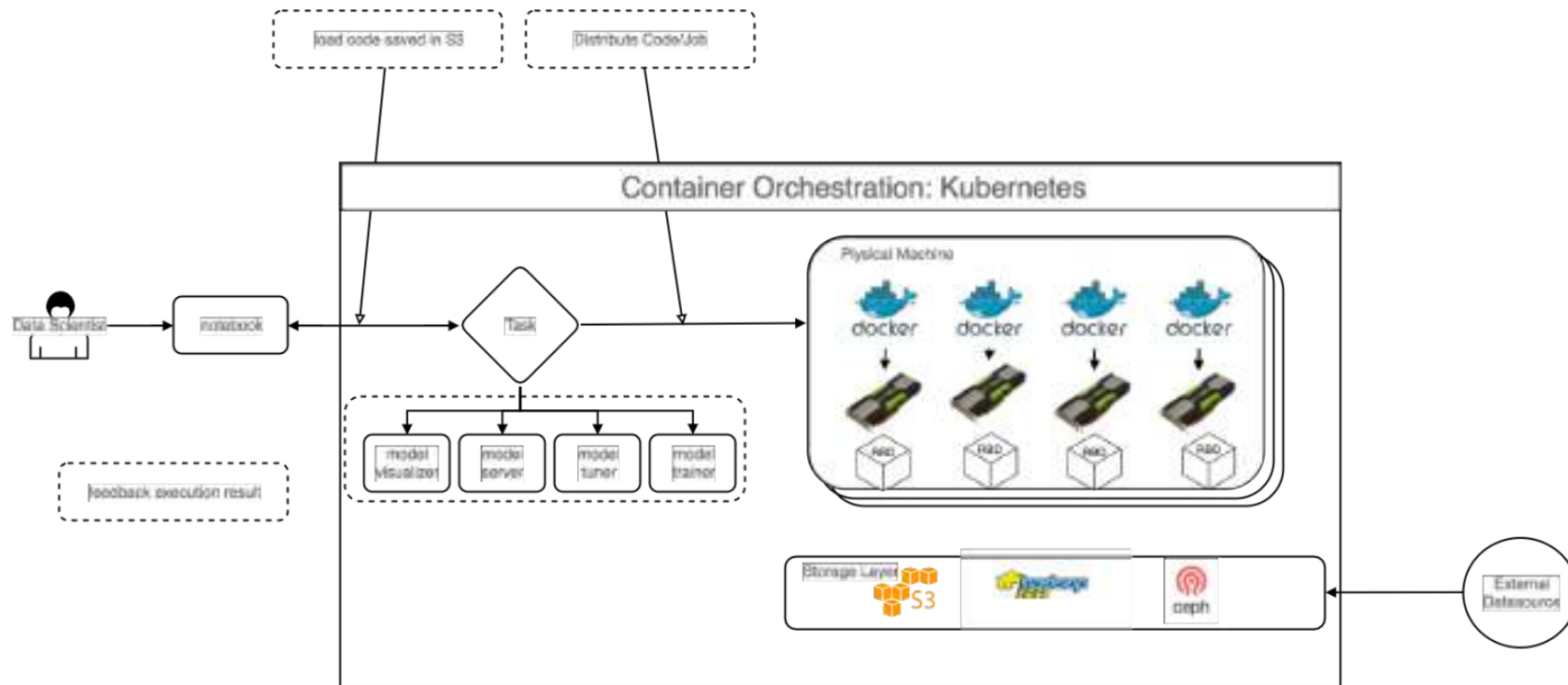
GOPS2017
Shanghai

一个日常的工作流

Workflow



GOPS2017
Shanghai



GPU云使用简介



The screenshot displays the 'GPUs' management console. On the left is a dark sidebar with navigation options: 'GPUs', 'Application', 'PUL', 'EXTEND', and 'ADD'. The main content area is titled 'Single Deploy instance' and contains an 'Applications' table. The table lists three applications: 'gpu_test_gpu', 'gpu_ML_gpu', and 'gpu_ML_gpu'. The third application is currently in a 'Pending' state, with a 'Cancel' button visible next to it. The interface also includes 'Return' and 'Search Instance' buttons at the top right.

Application	Create At	Application	Status	Action
gpu_test_gpu	null			Deploy
gpu_ML_gpu	null			Deploy
gpu_ML_gpu	2017-12-20 10:31	tensorflow	Pending	Cancel



GOPS2017
Shanghai

如何启动一个GPU环境

申请appcode



GOPS2017
Shanghai

The screenshot shows the GOPS portal interface. The top navigation bar includes 'Portal', '应用列表', '应用中心', '应用码', '应用发布', '权限管理', and 'ops'. The '应用中心' (Application Center) section is active, displaying a list of applications. A red box highlights the '应用码' (AppCode) button. The '应用码' (AppCode) section is also highlighted with a red box, showing a search bar and a list of application codes. The '主机环境' (Host Environment) section is visible on the right, showing details for the application 'qunarcorp.io.io_samba_b_log_collector'.

The screenshot shows the '新建Appcode' (New AppCode) form. The form fields are as follows:

- 资源归属应用节点 (归属应用节点): qunarcorp.ops.dev.spectrum
- 应用代号 (Appcode): npe_ix_test_gpu (highlighted with a red box)
- 应用名称 (默认与资源名称一致): npe_ix_test_gpu
- 负责人 (至少包含人与机器名称, 最多4个负责人): xiaoxu.li@yq
- 邮箱 (请提供非空邮箱地址, 最多1个邮箱, 最多10个): xiaoxu.li
- 备注 (最多100个字符, 最多100个字符, 无特殊限制): opdev@qunar.com

The screenshot shows the '新建Appcode' (New AppCode) form with the '资源类型' (Resource Type) section highlighted. The form fields are as follows:

- 资源类型* (单击选择, 若该应用上需要申请主机, 选择默认值 '主机' 即可): 主机 ES云 GPU云 (highlighted with a red box) 选择GPU云

Buttons: 取消, 上一步, 确定

• 创建一个应用



GOPS2017
Shanghai

The screenshot displays the GOPS2017 Shanghai web interface. On the left, a sidebar contains a navigation menu with 'GPU Applications' selected. The main content area is titled 'Single Deploy Instance' and shows a table of applications. Below the table is a 'Wizard: New Application' section with a 'Deploy' step selected.

Appcode	Create At	Application	Status	Action
app_001_gpu	not			Deploy
app_011_gpu	not			Deploy
app_021_gpu	not			Deploy

The 'Deploy' wizard shows a progress bar with four steps: Step 1 (selected), Step 2, Step 3, and Step 4. Below the progress bar, there are input fields for 'Appcode' (containing 'GPU_001_gpu') and 'Application' (containing 'GPU Applications'). There are also buttons for 'GPU Applications' and 'GPU Applications', and a 'Go Back' button.



- 查看应用详情

Application ops_xx_gpu

ops_xx_gpu running

Metadata

App	Jupyter
Interface	Password stop refresh Jupyterhub使用密码
	Service refresh 浏览器
使用URL	
Access Key stop refresh 密钥	
Secret Key stop refresh 密钥 输入用户名	
Service refresh 浏览器	

Resource:

GPU Amount	1
Storage Size	100

Billing

GPU:	\$0.00
Storage:	\$0.00
Subtotal:	\$0.00

Refresh Disable



GOPS2017
Shanghai

如何上传数据？

S3 Usage client/web



GOPS2017
Shanghai

Name	Size	Last Modified
BUILD	2.52 KiB	Jun 22, 2017 11:42 AM
__init__.py	979 bytes	Jun 22, 2017 11:42 AM
fully_connected_feed.py	9.29 KiB	Jun 22, 2017 11:42 AM
mnist_deep.py	1.06 KiB	Jun 22, 2017 11:42 AM
mnist.py	5.17 KiB	Jun 22, 2017 11:42 AM
mnist_deep.py	5.27 KiB	Jun 22, 2017 11:42 AM
mnist_softmax.py	2.68 KiB	Jun 22, 2017 11:42 AM
mnist_softmax_ops.py	1.60 KiB	Jun 22, 2017 11:42 AM
mnist_with_summaries.py	8.19 KiB	Jun 22, 2017 11:42 AM

```
mc config host add ml s3url accesskey secretkey
```

```
# 查看目录
```

```
mc ls --recursive ml/theano
```

```
# 拷贝
```

```
mc cp --recursive ipynb ml/theano
```

```
example git:(master) mc config host list
2 ceph-cn :
2 ceph-test:
2 cn-beta :
2 gpu :
2 ml :
2 ts :
2

example git:(master) mc ls --recursive ml/examples
2 [2017-06-22 11:42:07 CST] 2.5KiB mnist/BUILD
2 [2017-06-22 11:42:07 CST] 979B mnist/__init__.py
2 [2017-06-22 11:42:07 CST] 9.3KiB mnist/fully_connected_feed.py
2 [2017-06-22 11:42:07 CST] 1.1KiB mnist/input_data.py
2 [2017-06-22 11:42:07 CST] 5.2KiB mnist/mnist.py
2 [2017-06-22 11:42:07 CST] 5.2KiB mnist/mnist_deep.py

root@l-gpu3:/notebooks# ls
examples lost+found test
root@l-gpu3:/notebooks# ls
mnist monitors udacity
root@l-gpu3:/notebooks#
```



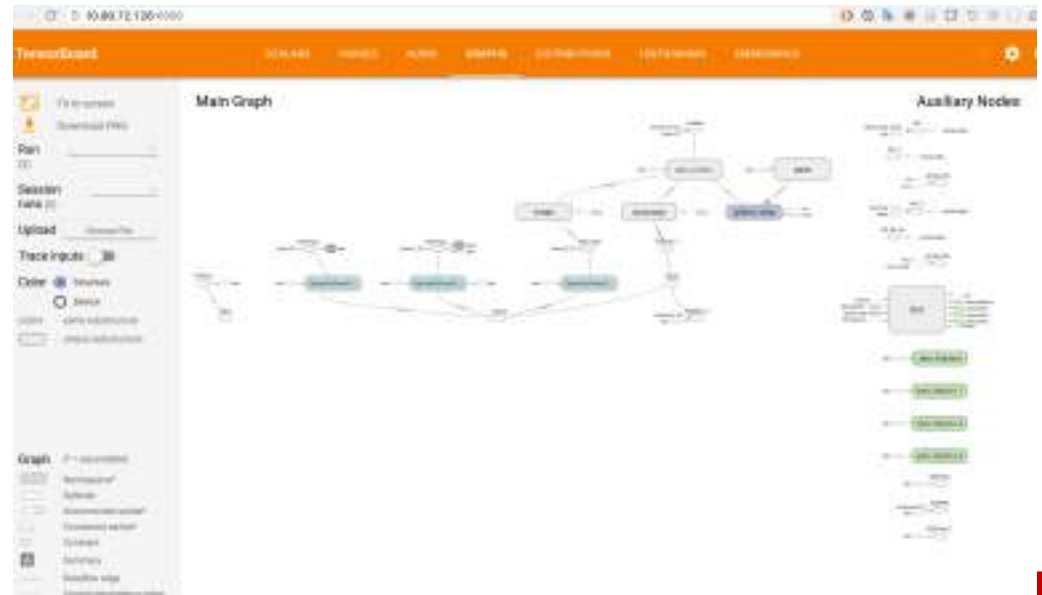
GOPS2017
Shanghai

启动TensorBoard

• Tensorboard



```
root@gu01:~/octocatool tensorboard --logdir /tmp/iris_model/ --port 8008
1 tensorflow/stream_executor/dso_loader.cc:158] successfully opened DSO library /lib64/libc.so.6 locally
2 tensorflow/stream_executor/dso_loader.cc:158] successfully opened DSO library /lib64/libc.so.6 locally
3 tensorflow/stream_executor/dso_loader.cc:158] successfully opened DSO library /lib64/libc.so.6 locally
4 tensorflow/stream_executor/dso_loader.cc:158] successfully opened DSO library /lib64/libc.so.6 locally
5 tensorflow/stream_executor/dso_loader.cc:158] successfully opened DSO library /lib64/libc.so.6 locally
6 tensorflow/stream_executor/dso_loader.cc:158] successfully opened DSO library /lib64/libc.so.6 locally
Starting TensorBoard on port 8008
(You can navigate to http://192.168.22.130:8008)
***WARNING: tensorflow found more than one watchdog event per run. Suppressing the watchdog with the newest event.
```





GOPS2017
Shanghai

如何停止/回收一个GPU环境？

应用详情页



GOPS2017
Shanghai

ops_xx_gpu ✓ running

Metadata

App	tensorflow
Supplier	
Password	4KRC2JhXFi
Service	browse
UI 使用手册	
Access Key	8E2PY03FkUz
SecretKey	sdU09aJem(tg)
Service	browse

Resource: Billing

Re-Deploy Disable

Status	Action
	Deploy
	Deploy
✓ running	Disable



GOPS2017
Shanghai

使用对比

Before:

- 协调GPU资源
- 安装环境, 环境污染(不同版本libs)
- 手动保存数据, 数据共享问题
- 数据大小受磁盘空间限制

After:

- 环境秒起秒删
- 环境隔离
- 对接S3, multiple access to data



GOPS2017
Shanghai

Q&A