

外接设备透传 使用教程

产品版本 : ZStack 3.3.0

文档版本 : V3.3.0

版权声明

版权所有©上海云轴信息科技有限公司 2019。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标说明

ZStack商标和其他云轴商标均为上海云轴信息科技有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受上海云轴公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，上海云轴公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目录

版权声明.....	1
1 GPU透传.....	1
1.1 介绍.....	1
1.2 前提.....	1
1.3 添加物理机.....	1
1.4 确认GPU设备启用.....	4
1.5 云主机加载GPU设备.....	10
1.6 典型应用场景.....	16
1.6.1 3D渲染.....	16
1.6.2 人工智能.....	16
1.6.3 云游戏.....	17
1.6.4 VDI.....	18
2 USB透传.....	20
2.1 介绍.....	20
2.2 前提.....	20
2.3 添加物理机.....	20
2.4 查看USB设备详情.....	23
2.5 云主机加载USB设备.....	26
3 块设备透传.....	29
3.1 介绍.....	29
3.2 前提.....	29
3.3 添加物理机.....	29
3.4 云主机加载块设备.....	32
3.5 典型应用场景.....	36
3.5.1 微软故障转移集群.....	36
3.5.2 Oracle RAC.....	37
术语表.....	38

1 GPU透传

1.1 介绍

ZStack支持GPU透传功能，物理机GPU设备可直接透传到云主机，让云主机享有物理机强劲的GPU并行计算能力，不仅适用于3D渲染、高清转解码场景，还适用于诸多高性能计算（HPC）场景，如机器学习、医疗成像、石油勘探数据分析、比特币挖掘等具有大量密集运算特点的场景。

1.2 前提

- 本教程假定用户已安装最新版本ZStack，并部署完成创建云主机必要的资源。
具体方式请参考《[用户手册](#)》安装部署章节。
- 本章节将从添加物理机的步骤开始，详细介绍GPU透传功能的使用方法。

1.3 添加物理机

操作步骤

1. 登录ZStack

使用Chrome浏览器或Firefox浏览器进入ZStack管理界面（http://your_machine_ip:5000/），默认用户名和密码为：`admin/password`。

如图 1: [ZStack登录界面](#)所示：

图 1: ZStack登录界面



2. 添加物理机

ZStack支持IP、IP范围和模板导入方式添加物理机，详情请参见《[用户手册](#)》添加物理机章节。

本教程以IP方式为例添加物理机。

在ZStack私有云主菜单，点击**硬件设施 > 物理机**，进入**物理机**界面，点击**添加物理机**，在弹出的**添加物理机**界面，可参考以下示例输入相应内容：

- **添加方式**：选择手动添加方式
- **名称**：设置物理机名称
- **简介**：可选项，可留空不填
- **集群**：选择物理机所在的集群
- **添加物理机IP**：选择IP方式
- **物理机IP**：输入物理机IP
- **扫描物理机IOMMU设置**：勾选该选项



注：

- 如果用户在创建物理机过程中未勾选**扫描物理机IOMMU设置**，则需在创建物理机后手动开启内核IOMMU；
- 需确保全局设置中**PCI设备IOMMU开关**为True；
- 需确保BIOS中**Intel VT-d或AMD IOMMU**开启；
- 扫描后，需重启物理机以使得IOMMU配置在内核生效。

- **关闭Intel EPT硬件辅助**：选择关闭或开启Intel EPT硬件辅助功能



注:

- 默认不勾选，表示开启Intel EPT硬件辅助功能；
 - 若因服务器CPU型号过旧，导致无法创建云主机或者创建的云主机不能显示控制台界面，可勾选此项，关闭Intel EPT硬件辅助功能；
 - 可在物理机详情页修改Intel EPT硬件辅助功能的开启/关闭状态；
 - 该功能仅针对Intel CPU生效。
- **SSH端口**：设置SSH端口，默认为22
 - **用户名**：默认为root用户，也可输入普通用户
 - **密码**：输入物理机对应的用户密码
 - **添加更多物理机**：点击+按钮，继续添加其他物理机。默认会采用相同的用户名、密码；如果用户名、密码不一致，可自定义输入

如图 2: 添加物理机所示，点击**确定**按钮，完成物理机添加。

图 2: 添加物理机

确定 取消

添加物理机

添加方式 * ?

手动添加 模版导入

名称 * ?

HOST

简介

集群 *

Cluster-1 ⊖

添加物理机IP *

IP IP 范围

物理机IP *

172.20.14.32

扫描物理机IOMMU设置 ?

关闭Intel EPT硬件辅助

SSH端口 *

22

用户名 *

root

密码 *

.....

添加更多物理机 ?

+

1.4 确认GPU设备启用

操作步骤

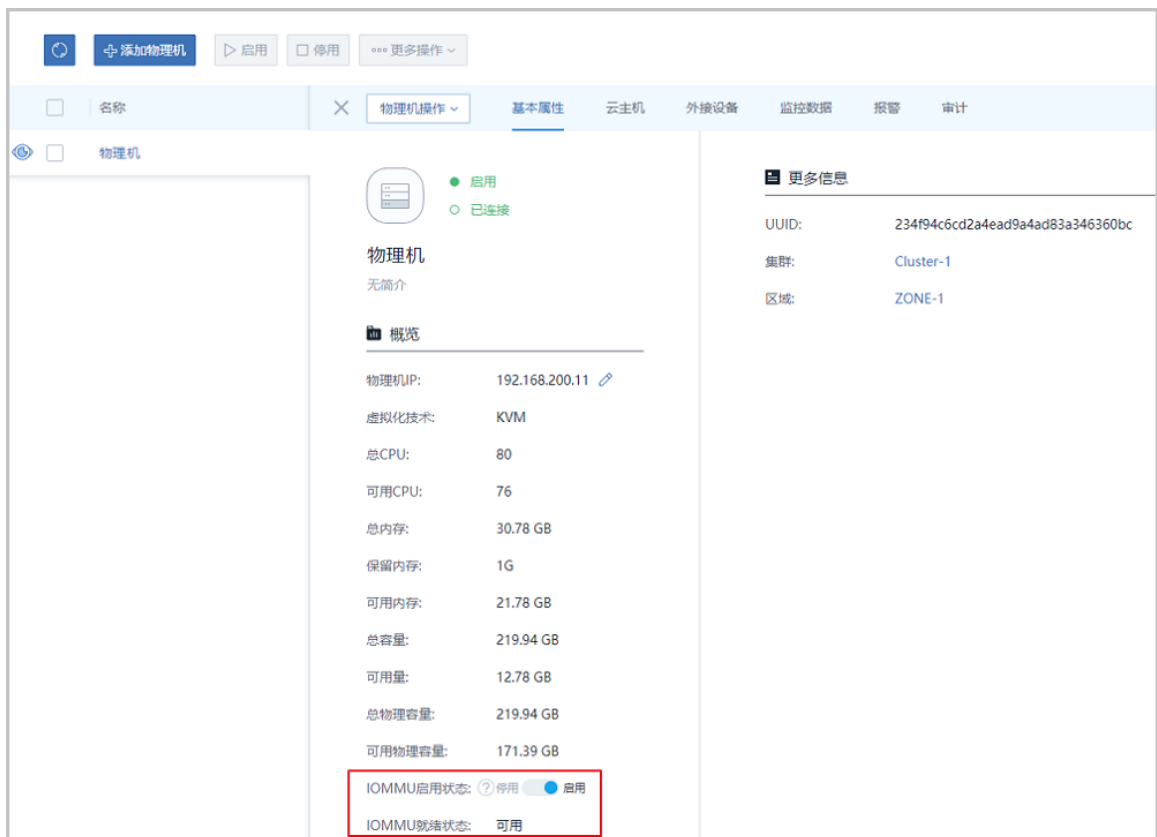
1. 确认IOMMU的启用状态和就绪状态

在**物理机**界面，点击物理机名称进入**物理机详情**页面，在**基本属性**子页面，请确认：

- IOMMU启用状态：**启用**
- IOMMU就绪状态：**可用**

如图 3: [确认IOMMU启用和就绪状态](#)所示：

图 3: 确认IOMMU启用和就绪状态



1. 要使用GPU透传功能，必须确保两点：

- IOMMU开启（包括内核IOMMU开启，数据库中IOMMU状态为启用）
- BIOS中Intel VT-d或AMD IOMMU选项开启

2. IOMMU启用状态：

- **启用**：开关划至启用，表示IOMMU开启
- **停用**：开关划至停用，表示IOMMU关闭



注：

目前ZStack的**IOMMU启用状态**开关特性：

- 首次将**IOMMU启用状态**开关划至**启用**，内核IOMMU开启，需重启服务器生效，同时数据库中IOMMU状态为启用。
- 将**IOMMU启用状态**开关划至**停用**，内核IOMMU开启记录依然存在，但数据库中IOMMU状态变为停用，此时，GPU透传功能关闭。
- 目前ZStack的**IOMMU启用状态**开关，其**停用**仅支持修改IOMMU在管理节点的状态标签，不支持内核IOMMU参数的删除操作。

3. 如果用户将**IOMMU启用状态**开关划至**启用**，但发现**IOMMU就绪状态**为**不可用**：

- 此时页面右下角会弹出操作助手，如图 4: 弹出操作助手所示
- 此时内核IOMMU已开启，需重启服务器生效，同时，用户需进入BIOS开启Intel VT-d或AMD IOMMU选项

图 4: 弹出操作助手

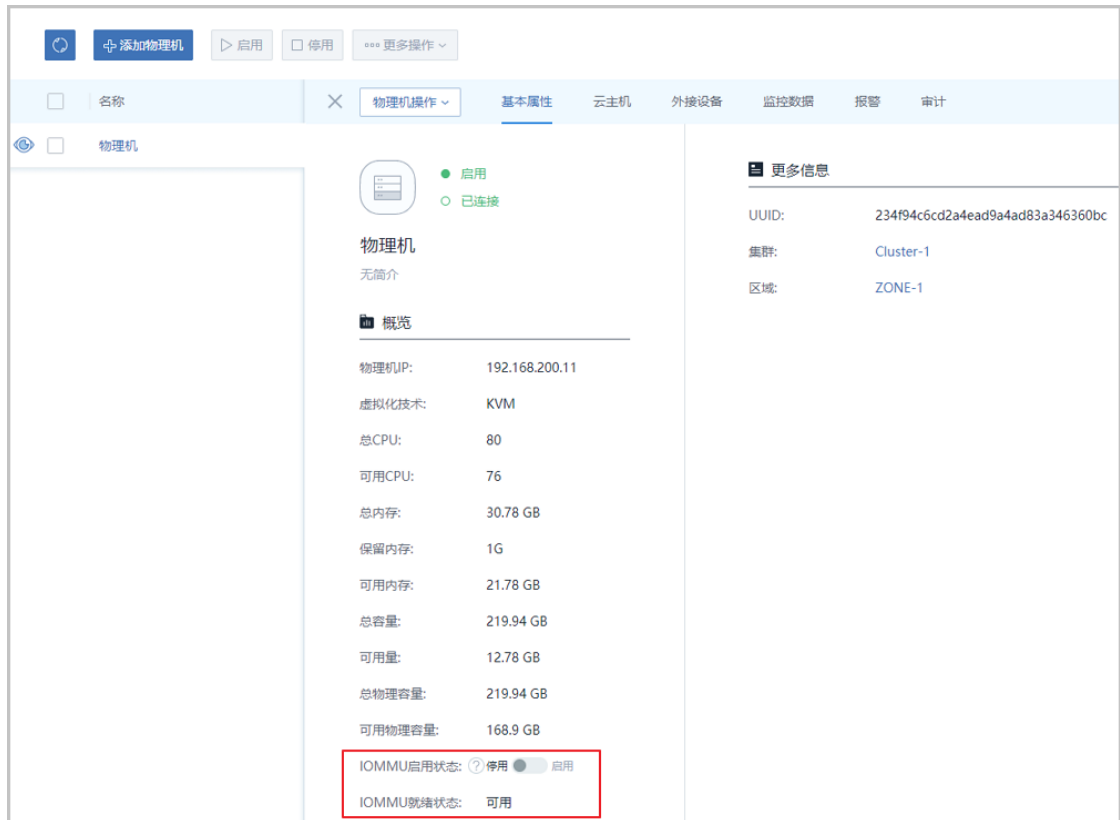


4. IOMMU就绪状态：

- 只有当 " 1. IOMMU开启；2. BIOS中Intel VT-d或AMD IOMMU选项开启 " 两点同时具备，IOMMU就绪状态才为**可用**，才能使用GPU透传功能。
5. 当IOMMU就绪状态已经为**可用**，用户希望暂时关闭该物理机的GPU透传功能，可将IOMMU启用状态的开关划至**停用**。

如图 5: IOMMU停用所示：

图 5: IOMMU停用



此时：

- 该物理机GPU设备已不能再次加载到云主机。
- 对于已加载该物理机GPU设备的云主机，透传不受影响，但GPU设备一旦卸载就无法再次加载。
- 用户希望再次启用该物理机的GPU透传功能，只需将IOMMU启用状态的开关划至**启用**即可。

2. 查看物理机的GPU设备详情

点击物理机名称，在**物理机详情**页面，点击**外接设备 > GPU设备**，进入**GPU设备**页面，展示了该物理机GPU设备的设备名、设备地址、类型、启用状态、就绪状态、加载到云主机的情况等，如图 6: GPU设备详情所示：

图 6: GPU设备详情



注:

- GPU设备需同时具备 "1. 启用状态为启用；2. 就绪状态为就绪" 两点，才能加载到云主机。
- GPU设备未加载到任何云主机，云主机状态为**未加载**；GPU设备加载到某个云主机，云主机状态将显示该云主机名，同时GPU设备**就绪状态为已加载**，如[图 7: GPU设备加载到某个云主机](#)所示。关于云主机加载GPU设备，详情见下节。
- 一旦GPU设备加载到某个云主机，该GPU设备就为该云主机独享。
- 云主机只能加载所在物理机的GPU设备，不支持跨物理机GPU设备的加载。

图 7: GPU设备加载到某个云主机



3. GPU设备支持的操作

- 启用：

启用某个GPU设备，该设备的**启用状态变为启用**，如果此时**就绪状态为就绪**，表示云主机可加载该设备。

- 停用：

停用某个GPU设备，该设备的**启用状态**变为**停用**，表示该设备暂时不可被加载到云主机。
- 全局共享：

将已加载的GPU设备全局共享给普通账户。
- 全局召回：

将共享给普通账户GPU设备全局召回。
- 删除：
 - 如果物理机异常掉电/断网，修复后重启物理机，重连过程中自动扫描GPU设备，由于GPU设备的BDF号很可能已更改，GPU设备状态在数据库中被刷新为**不可用**，此时用户可删除**就绪状态为不可用的GPU设备**。
 - 如果GPU设备从当前插槽拔出，重启物理机，重连过程中自动扫描GPU设备未找到，GPU设备状态在数据库中被刷新为**不可用**，此时用户可删除**就绪状态为不可用的GPU设备**。

如图 8: GPU设备支持的操作所示：

图 8: GPU设备支持的操作



4. 物理机所在集群的**集群详情页**，点击**外接设备 > GPU设备**，进入**GPU设备**页面，可查看GPU设备详情，如图 9: 集群详情页支持查看GPU设备详情所示：

图 9: 集群详情页支持查看GPU设备详情



后续操作

确认GPU设备启用后，接下来需将GPU设备加载到云主机。

1.5 云主机加载GPU设备

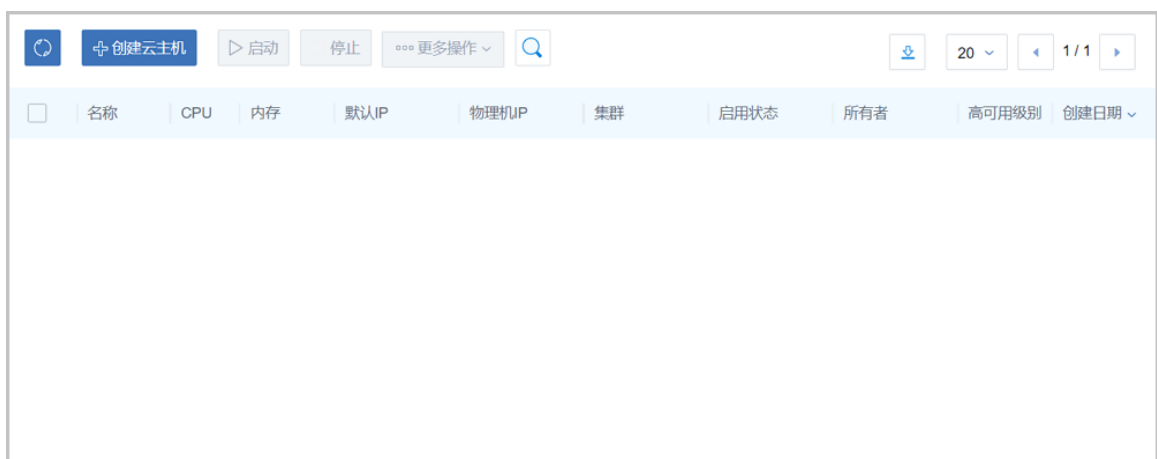
操作步骤

1. 创建云主机

a) 进入云主机界面

在ZStack私有云主菜单，点击云资源池 > 云主机，进入云主机界面，如图 10: 云主机界面所示：

图 10: 云主机界面



b) 点击**创建云主机**按钮，弹出**创建云主机**界面，如图 11: 创建云主机所示，可参考以下示例输入相应内容：

- **名称**：设置云主机名称

- **简介**：可选项，可留空不填
- **计算规格**：选择云主机的计算规格
- **镜像**：选择云主机的镜像
- **网络**：选择创建云主机的网络
- **高级-GPU设备**：可选项，创建云主机时支持在高级操作中选择GPU设备；也可先创建云主机，然后在配置信息中加载GPU设备

**注:**

- 需提前在物理机BIOS中开启Intel VT-d或AMD IOMMU，且在物理机内核开启IOMMU支持，确保物理机可正常使用GPU设备透传功能。
- 支持加载多个不同类型的GPU设备到云主机
- 不能跨物理机加载GPU设备到云主机

图 11: 创建云主机

确定
取消

创建云主机

添加方式

单个 多个

名称 *

VM

简介

计算规格 *

InstanceOffering-1
⊖

镜像 *

Image-1
⊖

网络

网络地址类型 * ?

IPv4
IPv6
双栈

三层网络 *

L3-扁平网络
⊖

默认网络
设置网卡

⊕

高级 ▾

数据云盘规格

亲和组

CPU绑定 ?

vCPU : pCPU

+ 添加更多

集群

主存储

物理机

虚拟光驱

+ 创建更多光驱

GPU设备 ?

Advanced Micro Devices, Inc. [AMD/ATI], D... -

2. 云主机加载GPU设备

如果创建云主机时已添加GPU设备，可通过该方法继续加载或更换GPU设备。

a) 进入云主机详情的配置信息子页面。

在云主机界面点击云主机名称，显示云主机详情页面，点击配置信息进入配置信息子页面。

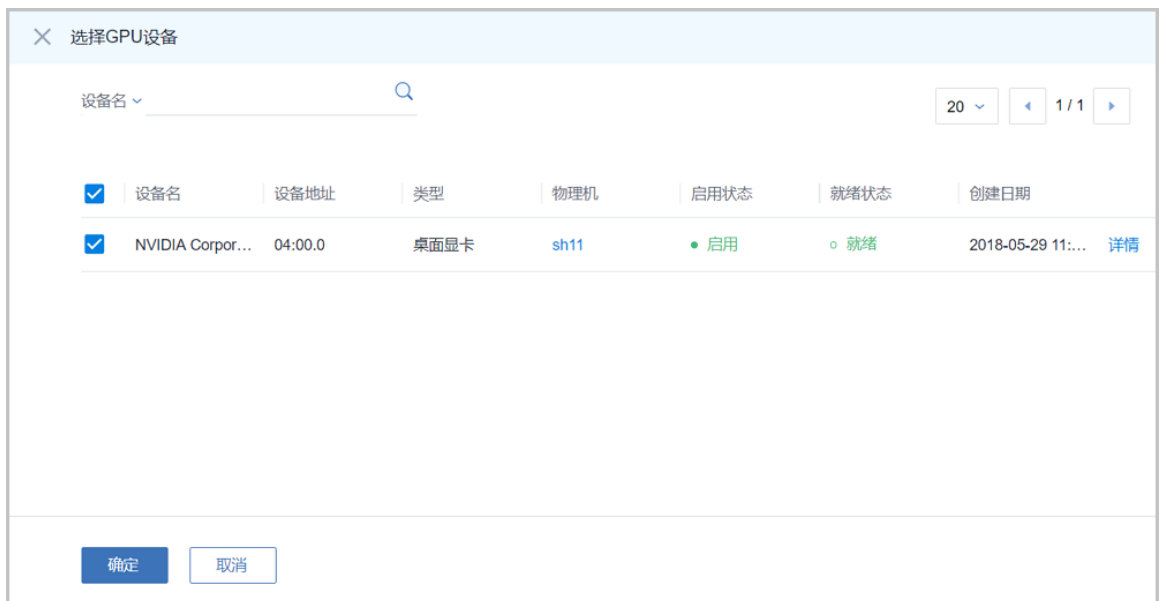
在GPU设备处，点击操作，在下拉菜单中选择加载，如图 12: 配置信息所示：

图 12: 配置信息



b) 弹出**选择GPU设备**界面如所示，选择要加载到云主机VM上的GPU设备，如图 13: [选择要加载到云主机VM上的GPU设备](#)所示：

图 13: 选择要加载到云主机VM上的GPU设备



c) 云主机成功加载GPU设备，如图 14: [云主机成功加载GPU设备](#)所示：

图 14: 云主机成功加载GPU设备



3. 安装显卡驱动

加载GPU设备后，需要安装对应的驱动程序。请在AMD官网或NVIDIA官网获取官方驱动程序。

- Linux支持AMD的计算卡、游戏卡、专业卡。Linux自带社区驱动，如需支持计算加速和显示加速功能。请[点击这里](#)安装官方驱动。
- Linux支持NVIDIA的计算卡、游戏卡、专业卡。Linux自带社区驱动，如需支持计算加速和显示加速功能。请[点击这里](#)安装官方驱动。
- Windows支持AMD的计算卡、游戏卡、专业卡和NVIDIA的计算卡。请[点击这里](#)根据显卡类型和Windows操作系统版本下载合适的显卡驱动。
- Windows仅支持NVIDIA的计算卡。请[点击这里](#)根据显卡类型和Windows操作系统版本下载合适的显卡驱动。

后续操作



注:

用户可自行设置是否允许云主机热插拔GPU设备，设置方法：

进入**设置 > 全局设置 > 高级设置**，设置**PCI设备热插拔开关**，默认为true。

若热插拔时出现硬件兼容性错误，或不支持该硬件设备时，可以设置为false，关闭此功能。

至此，GPU透传功能已成功开启。

1.6 典型应用场景

GPU透传功能通过云主机透传物理机强劲的GPU计算能力，可适用于3D渲染、人工智能、云游戏、VDI等典型应用场景。

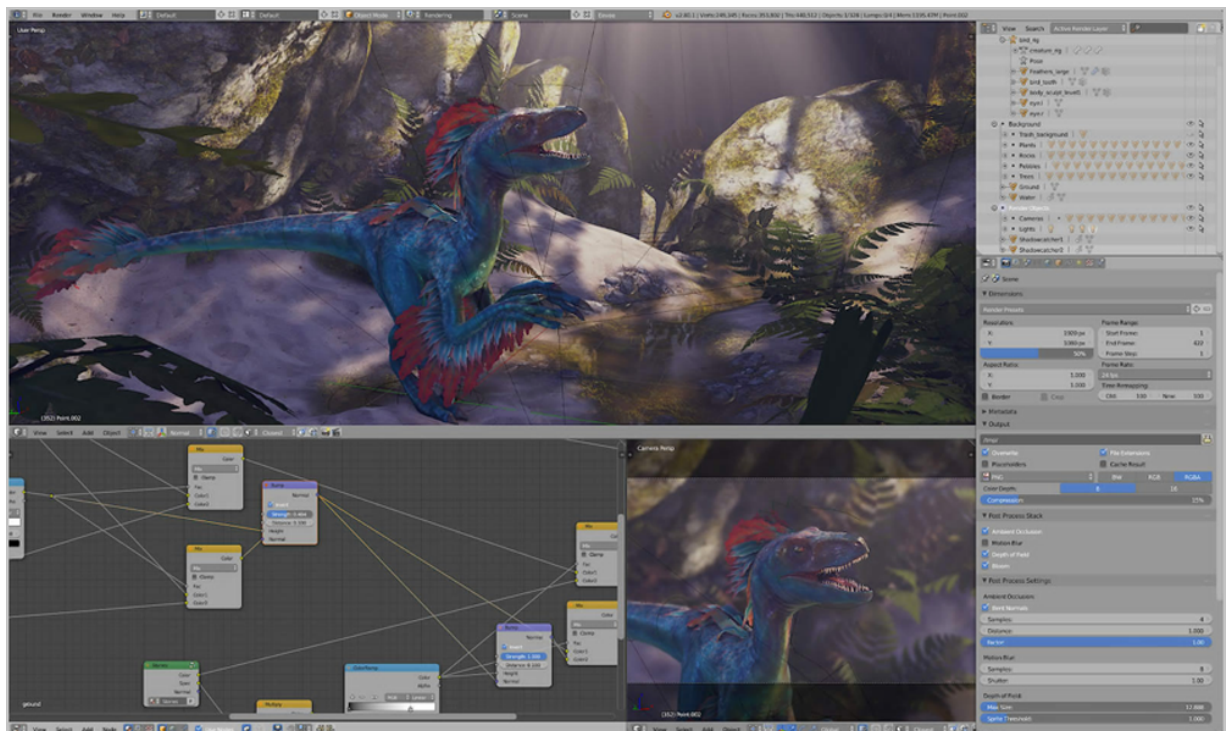
1.6.1 3D渲染

三维计算机图形的预渲染 (Pre-rendering、Offline rendering) 或实时渲染 (Real-time rendering、Online rendering) 速度都很缓慢。预渲染常用于电影制作，要求很高的计算强度，需要大量的服务器提供运算能力；实时渲染常用于三维视频游戏，通常依靠图形处理器 (GPU) 完成这个过程。

现在由于GPU的高速发展，已经有相当多的3D渲染是在GPU服务器集群中完成。结合ZStack的GPU透传功能，在性能损失极低的情况下 (5%以内) 同时可获得集中高效的集群管理功能，再配合智能监控软件以及ZStack自带的计费功能，可以形成一整套更便捷高效的渲染农场方案。

如图 15: 3D渲染所示：

图 15: 3D渲染



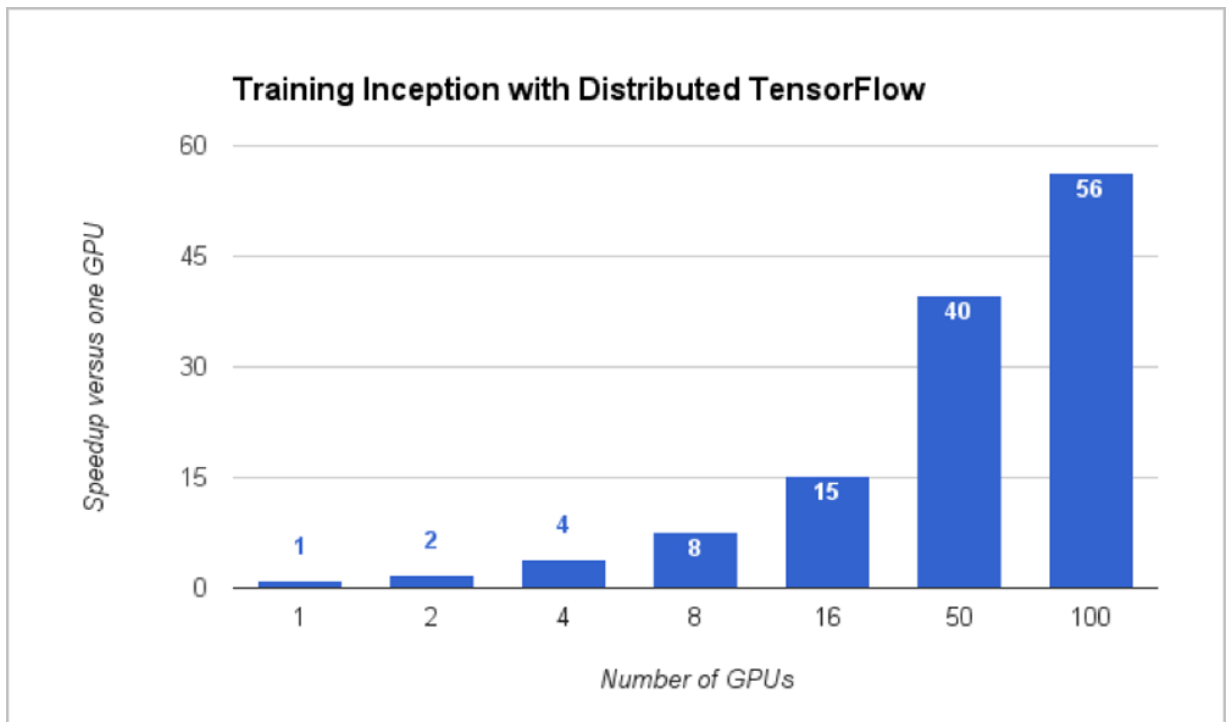
1.6.2 人工智能

GPU的计算能力可以应用于深度学习。自Google推出神经网络工具TensorFlow后，许多科研机构以及企业应用都日渐明显偏向使用GPU作为基础设施。

以规格较高的NVIDIA P100显卡为例，通过ZStack的GPU透传功能，将其透传至云主机后，性能测试结果显示，几乎与标称完全一致，能够充分满足大规模模型训练对基础设施的要求。

如图 16: 人工智能所示：

图 16: 人工智能



1.6.3 云游戏

随着宽带网络的发展，以及移动终端设备的普及，将游戏计算至于云端，客户端仅仅负责显示与控制的游戏模式也悄然开始流行。云端服务器上渲染3D游戏，即时为每一帧进行编码，将结果以流的形式传输至任何接驳有线或无线网络的设备。

这种云游戏模式，可以借助GPU以及服务器CPU能力，通过ZStack的GPU透传功能，为游戏创造隔离性最佳的虚拟环境，从而保证计算与渲染的流畅度，为用户提供更好的游戏体验。

如图 17: 云游戏所示：

图 17: 云游戏



1.6.4 VDI

GPU一直是VDI（桌面云）中非常重要的设备，它不仅能够改善桌面视觉体验，同时在特殊的应用程序中承担主力计算角色，从而完全代替传统PC图站，让用户在更为安全的环境中进行3D设计。

通过ZStack的GPU透传功能，以及配合RDP、PCoIP等协议，可充分利用显卡能力，比如3D设计、游戏等流畅运行，提供更逼近本地物理机的用户体验。

如图 18: VDI所示：

图 18: VDI



2 USB透传

2.1 介绍

ZStack支持USB透传功能，物理机USB设备可直接透传至该物理机上所运行的云主机，从而让云主机能够直接使用物理机上的USB设备。

2.2 前提

- 在此教程中，假定用户已安装最新版本ZStack，并部署完成创建云主机必要的资源。
具体方式请参考《[用户手册](#)》安装部署章节。
- 本章节将从添加物理机的步骤开始，详细介绍USB透传功能的使用方法。



注：USB透传功能需要ZStack管理员权限才可以操作。

2.3 添加物理机

操作步骤

1. 登录ZStack

使用Chrome浏览器或FireFox浏览器进入ZStack管理界面（http://your_machine_ip:5000/），默认用户名和密码为：admin/password，如[登录界面](#)所示：

图 19: 登录界面



2. 添加物理机

ZStack支持IP、IP范围和模板导入方式添加物理机，详情请参见《[用户手册](#)》添加物理机章节。本教程以IP方式为例添加物理机。

在ZStack私有云主菜单，点击**硬件设施** > **物理机**，进入**物理机**界面，点击**添加物理机**，在弹出的**添加物理机**界面，可参考以下示例输入相应内容：

- **添加方式**：选择手动添加方式
- **名称**：设置物理机名称
- **简介**：可选项，可留空不填
- **集群**：选择物理机所在的集群
- **添加物理机IP**：选择IP方式
- **物理机IP**：输入物理机IP
- **扫描物理机IOMMU设置**：不勾选。USB透传功能无需IOMMU支持
- **关闭Intel EPT硬件辅助**：选择关闭或开启Intel EPT硬件辅助功能



注：

- 默认不勾选，表示开启Intel EPT硬件辅助功能；
- 若因服务器CPU型号过旧，导致无法创建云主机或者创建的云主机不能显示控制台界面，可勾选此项，关闭Intel EPT硬件辅助功能；
- 可在物理机详情页修改Intel EPT硬件辅助功能的开启/关闭状态；
- 该功能仅针对Intel CPU生效。
- **SSH端口**：设置SSH端口，默认为22
- **用户名**：默认为root用户，也可输入普通用户
- **密码**：输入物理机对应的用户密码
- **添加更多物理机**：点击+按钮，继续添加其他物理机。默认会采用相同的用户名、密码；如果用户名、密码不一致，可自定义输入

如图 20: 添加物理机所示，点击**确定**按钮，完成物理机添加。所示：

图 20: 添加物理机

确定 **取消**

添加物理机

添加方式 * ?

手动添加 模版导入

名称 * ?

HOST

简介

集群 *

Cluster-1 ⊖



添加物理机IP *

IP IP 范围

物理机IP *

172.20.14.32

扫描物理机IOMMU设置 ?

关闭Intel EPT硬件辅助

SSH端口 *

22

用户名 *

root

密码 *

.....

添加更多物理机 ?

+

2.4 查看USB设备详情

操作步骤

1. 查看物理机的USB设备详情

点击物理机名称，进入**物理机详情**页面，点击**外接设备 > USB设备**，进入**USB设备**页面，展示了该物理机中可用的USB设备列表，包括：设备名、生产商、类型、启用状态、USB版本、加载到云主机的情况等，如图 21: USB设备详情所示：

图 21: USB设备详情

<input type="checkbox"/>	设备名	生产商	类型	启用状态	USB版本	云主机	创建日期
<input type="checkbox"/>	SanDisk-00...	SanDisk	Cruzer Blade	• 启用	2.10	未加载	2017-12-21...
<input type="checkbox"/>	American ...	American ...	Virtual Key...	• 启用	1.10	未加载	2017-12-21...

**注:**

- USB设备没有统一的命名规范，管理员需要通过生产商、类型、USB版本等信息建立起与实际USB设备之间的关联。支持修改设备名，进一步提高USB设备的辨识度。
- 当一个USB3.0设备插入一个USB2.0口时，对外表现为USB2.0设备，即**USB版本显示2.00**。
- USB设备未加载到任何云主机，云主机状态为**未加载**；USB设备加载到某个云主机，云主机状态将显示该云主机名，如图 22: [USB设备加载到某个云主机](#)所示。关于云主机加载USB设备，详情见下节。
- 一旦USB设备加载到某个云主机，该USB设备就为该云主机独享。
- 云主机处于**运行中**状态时：只能选择当前所在物理机的可用USB设备，不支持跨物理机USB设备的加载。
- 云主机处于**已停止**状态时：
 - 如果此前未挂载USB设备，可从所在集群的全部可用USB设备中选择；如果该云主机加载多个USB设备，需确保所有USB设备处于同一物理机上；
 - 如果此前已加载USB设备，只能从该USB设备所在物理机选择其它可用USB设备；
 - 总原则：一台云主机加载的所有USB设备只能处于同一台物理机上。

**注:**

云主机加载USB设备，可能会影响其调度结果：该云主机只能在USB设备所在物理机上运行，如果物理机没有足够的资源，可能导致云主机无法开机！

图 22: USB设备加载到某个云主机



2. USB设备支持的操作

- **修改设备名**：修改设备名，进一步提高USB设备的辨识度。
- **启用**：启用某个USB设备，该设备的**启用状态**变为**启用**，表示云主机可加载该设备。
- **停用**：停用某个USB设备，该设备的**启用状态**变为**停用**，表示该设备暂时不可被加载到云主机。停用操作可用于过滤掉某些永远不会被加载的USB设备。
- **加载**：将某个USB设备加载到云主机。
- **卸载**：将某个USB设备从云主机上卸载。

如图 23: [USB设备支持的操作](#)所示：

图 23: USB设备支持的操作



3. 点击**硬件设施 > 集群**，在**集群**页面点击集群名称，进入物理机所在集群的**集群详情页**，点击**外接设备 > USB设备**，进入**USB设备**页面，可查看USB设备详情，如图 24: [集群详情页支持查看USB设备详情](#)所示：

图 24: 集群详情页支持查看USB设备详情



2.5 云主机加载USB设备

操作步骤

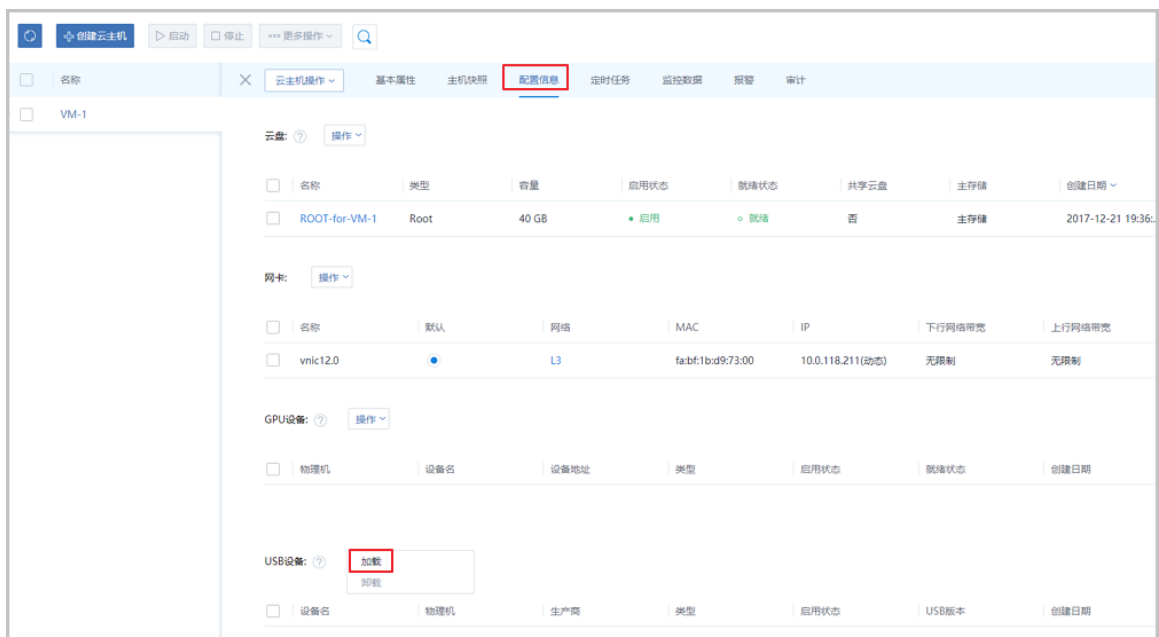
1. 云主机加载USB设备

- a) 在**云主机**界面，点击云主机名称，进入**云主机详情**的**配置信息**子页面。

在**云主机**界面点击云主机名称进入**云主机详情**页面，点击**配置信息**进入**配置信息**子页面。

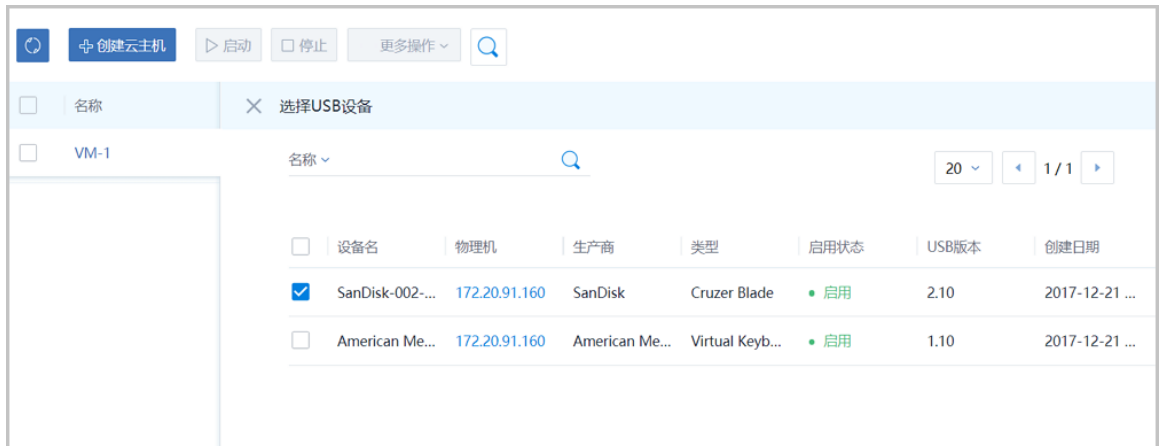
在USB设备处，点击**操作**，在下拉菜单中选择**加载**，如图 25: [配置信息](#)所示：

图 25: 配置信息



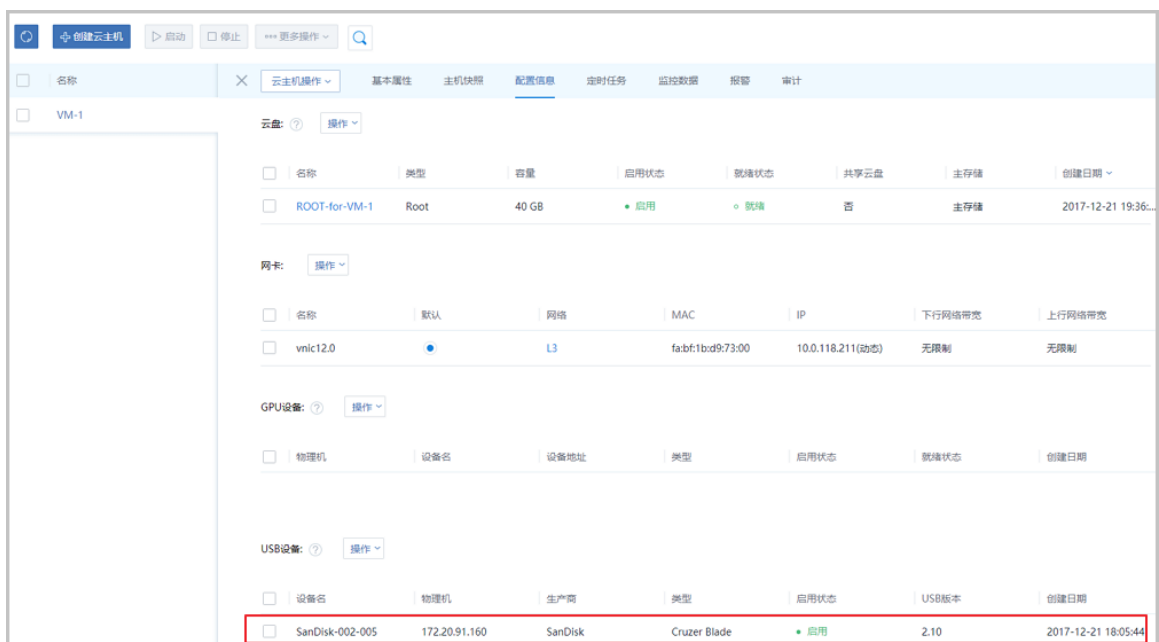
- b) 弹出**选择USB设备**界面，选择要加载到云主机VM上的USB设备，如图 26: 加载USB设备所示：

图 26: 加载USB设备



- c) 云主机加载USB设备列表，如图 27: 云主机加载USB设备列表所示：

图 27: 云主机加载USB设备列表



2. 登入云主机控制台，执行lsusb命令，可查看已成功加载到云主机的USB设备。

图 28: 查看云主机加载的USB设备

```
[root@172-20-91-160 ~]# lsusb
Bus 002 Device 005: ID 0781:5567 SanDisk Corp. Cruzer Blade
Bus 002 Device 004: ID 046b:ff10 American Megatrends, Inc. Virtual Keyboard and Mouse
Bus 002 Device 003: ID 046b:ff01 American Megatrends, Inc.
Bus 002 Device 002: ID 8087:0024 Intel Corp. Integrated Rate Matching Hub
Bus 002 Device 001: ID 1d6b:0002 Linux Foundation 2.0 root hub
Bus 001 Device 002: ID 8087:0024 Intel Corp. Integrated Rate Matching Hub
Bus 001 Device 001: ID 1d6b:0002 Linux Foundation 2.0 root hub
```

**注:**

可使用yum install usbutils或apt install usbutils来安装lsusb命令。此时确保云主机可连接公网。

3. 安装USB3.0驱动

Windows XP、Windows 2003、Windows 2008和Windows 7系统，不支持USB 3.0设备的直接加载，需要手动安装驱动。

参考路径：

http://cdn.zstack.io/tools/renesas_usb30_21390.exe

双击文件完成安装。

**注:**

- 对于USB2.0设备，Linux和Windows云主机均可加载并识别；
- 对于USB3.0设备，Linux云主机可加载并识别；Windows云主机中，Windows XP、Windows 2003、Windows 2008和Windows 7系统需要手动安装，或将USB3.0设备插在USB2.0口中，以提高识别率。
- 云主机加载USB设备的数量有上限限制：
 - USB1.0设备：最多支持加载1个
 - USB2.0设备：最多支持加载6个
 - USB3.0设备：最多支持加载4个

后续操作

至此，USB透传功能介绍完毕。

3 块设备透传

3.1 介绍

ZStack支持块设备透传功能，包括FC存储块设备透传、iSCSI块设备透传。

- iSCSI块设备透传：添加iSCSI服务器后，通过同步数据，可在UI界面上实时展示其上所有块设备，支持块设备透传给云主机使用。
- FC存储块设备透传：部署FC存储后，支持在UI界面上实时展示FC存储与所有块设备，并支持块设备透传给云主机使用；

3.2 前提

- 使用最新版本ZStack安装包，为服务器安装系统，并部署完成创建云主机等必要的资源。
 - 所有服务器的操作系统版本需一致，即：全部基于c72 ISO安装或全部基于c74 ISO安装。
 - 以**管理节点模式**安装其中一台服务器作为管理节点。
 - 通过管理节点接管物理机，物理机自动部署安装LVM、Multipath等工具包。

具体方式请参考《[用户手册](#)》安装部署章节。

- 在物理机上部署FC存储或iSCSI存储，划分出一定容量大小的块设备（LUN），并通过在线扫描发现磁盘。
- 两个以上的链路的FC存储或iSCSI存储，可根据需要配置多路径访问方式。
- 本章节将从添加物理机的步骤开始，详细介绍块设备透传功能的使用方法。

3.3 添加物理机

操作步骤

1. 登录ZStack

使用Chrome浏览器或FireFox浏览器进入ZStack管理界面（http://your_machine_ip:5000/），默认用户名和密码为：admin/password，如[登录界面](#)所示：

图 29: 登录界面



2. 添加物理机

ZStack支持IP、IP范围和模板导入方式添加物理机，详情请参见《[用户手册](#)》添加物理机章节。本教程以IP方式为例添加物理机。

在ZStack私有云主菜单，点击**硬件设施** > **物理机**，进入**物理机**界面，点击**添加物理机**，在弹出的**添加物理机**界面，可参考以下示例输入相应内容：

- **添加方式**：选择手动添加方式
- **名称**：设置物理机名称
- **简介**：可选项，可留空不填
- **集群**：选择物理机所在的集群
- **添加物理机IP**：选择IP方式
- **物理机IP**：输入物理机IP
- **扫描物理机IOMMU设置**：不勾选。块设备透传功能无需IOMMU支持
- **关闭Intel EPT硬件辅助**：选择关闭或开启Intel EPT硬件辅助功能



注:

- 默认不勾选，表示开启Intel EPT硬件辅助功能；
- 若因服务器CPU型号过旧，导致无法创建云主机或者创建的云主机不能显示控制台界面，可勾选此项，关闭Intel EPT硬件辅助功能；
- 可在物理机详情页修改Intel EPT硬件辅助功能的开启/关闭状态；
- 该功能仅针对Intel CPU生效。

- **SSH端口**：设置SSH端口，默认为22
- **用户名**：默认为root用户，也可输入普通用户
- **密码**：输入物理机对应的用户密码
- **添加更多物理机**：点击+按钮，继续添加其他物理机。默认会采用相同的用户名、密码；如果用户名、密码不一致，可自定义输入

如图 30: 添加物理机所示，点击**确定**按钮，完成物理机添加。

图 30: 添加物理机

确定 取消

添加物理机

添加方式 *

手动添加 模版导入

名称 *

HOST

简介

集群 *

Cluster-1

添加物理机IP *

IP IP 范围

物理机IP *

172.20.14.32

扫描物理机IOMMU设置 ?

关闭Intel EPT硬件辅助

SSH端口 *

22

用户名 *

root

密码 *

.....

添加更多物理机 ?

+

3.4 云主机加载块设备

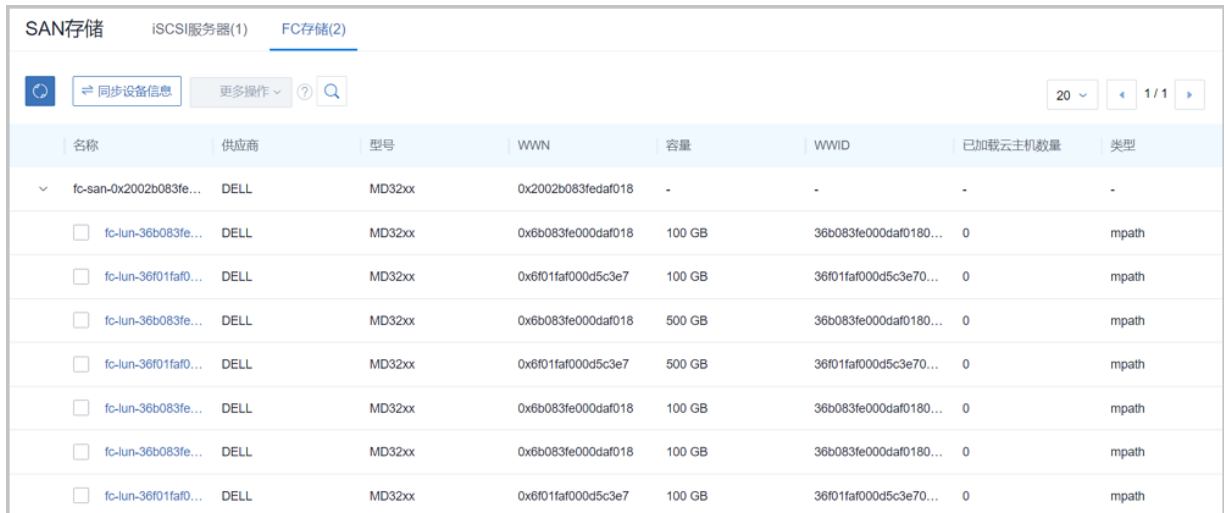
一个块设备可以同时加载到多个云主机，一个云主机可以挂载多个块设备。用户可通过以下页面为云主机加载/卸载块设备：

- FC存储页面：该页面支持将FC存储的块设备添加到云主机，详情请参考[FC存储页面加载块设备](#)
- iSCSI服务器页面：该页面支持将iSCSI存储的块设备添加到云主机，详情请参考[iSCSI/服务器页面加载块设备](#)
- 云主机详情页：该页面支持将FC存储的块设备、iSCSI存储的块设备添加到云主机，通过**类型**列进行区分所属的存储。详情请参考[云主机详情页加载块设备](#)
- 物理机详情页：该页面支持将FC存储的块设备、iSCSI存储的块设备添加到云主机，通过**类型**列进行区分所属的存储。详情请参考[物理机详情页加载块设备](#)

FC存储页面加载块设备

在ZStack私有云主菜单，点击**硬件设施** > **SAN存储**按钮，进入**SAN存储**界面的**FC存储**子页面，可查看FC存储及块存储上的块设备，如图 31: **FC存储**所示：

图 31: FC存储

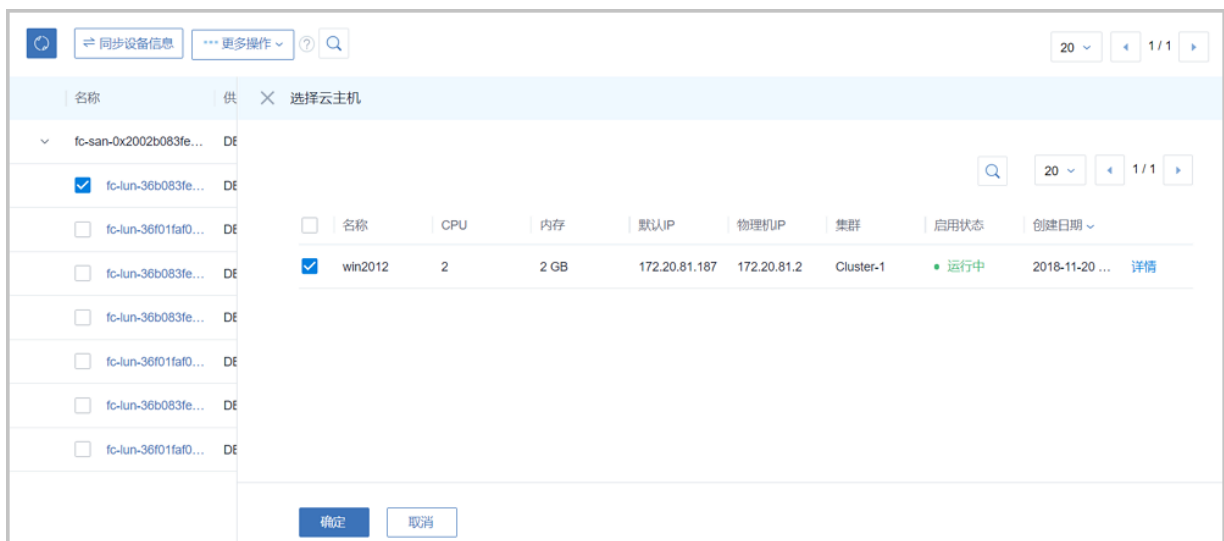


名称	供应商	型号	WWN	容量	WWID	已加载云主机数量	类型
fc-san-0x2002b083fe...	DELL	MD32xx	0x2002b083fedaf018	-	-	-	-
fc-lun-36b083fe...	DELL	MD32xx	0x6b083fe000daf018	100 GB	36b083fe000daf018...	0	mpath
fc-lun-36f01faf0...	DELL	MD32xx	0x6f01faf000d5c3e7	100 GB	36f01faf000d5c3e7...	0	mpath
fc-lun-36b083fe...	DELL	MD32xx	0x6b083fe000daf018	500 GB	36b083fe000daf018...	0	mpath
fc-lun-36f01faf0...	DELL	MD32xx	0x6f01faf000d5c3e7	500 GB	36f01faf000d5c3e7...	0	mpath
fc-lun-36b083fe...	DELL	MD32xx	0x6b083fe000daf018	100 GB	36b083fe000daf018...	0	mpath
fc-lun-36b083fe...	DELL	MD32xx	0x6b083fe000daf018	100 GB	36b083fe000daf018...	0	mpath
fc-lun-36f01faf0...	DELL	MD32xx	0x6f01faf000d5c3e7	100 GB	36f01faf000d5c3e7...	0	mpath

添加物理机后，如果物理机上部署了FC存储，ZStack会自动识别并展示到FC存储页面。点击**同步设备信息**，手动刷新列表信息。

选择一个块设备，点击**更多操作** > **加载云主机**按钮，在弹出的**选择云主机**页面，选择需要加载的云主机，点击**确定**按钮，完成块设备加载。如图 32: **加载云主机**所示：

图 32: 加载云主机



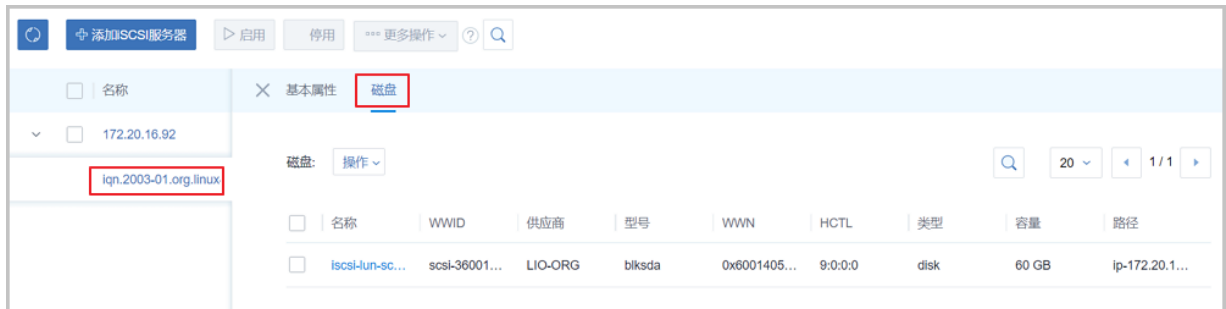
名称	CPU	内存	默认IP	物理机IP	集群	启用状态	创建日期
win2012	2	2 GB	172.20.81.187	172.20.81.2	Cluster-1	运行中	2018-11-20 ...

加载块设备后，点击**更多操作** > **卸载云主机**按钮，可将已加载的块设备从云主机卸载。

iSCSI服务器页面加载块设备

在ZStack私有云主菜单，点击**硬件设置** > **SAN存储**按钮，进入**SAN存储**界面的**iSCSI服务器**子页面，点击iSCSI Qualified名称，进入**磁盘**页面，如图 33: 磁盘子页面所示：

图 33: 磁盘子页面



在**磁盘**页面，选择一个iSCSI块设备，点击**操作** > **加载云主机**按钮，在弹出的**选择云主机**页面，选择需要加载的云主机，点击**确定**按钮，完成块设备加载。

加载块设备后，点击**操作** > **卸载云主机**按钮，可将已加载的块设备从云主机卸载。

云主机详情页加载块设备

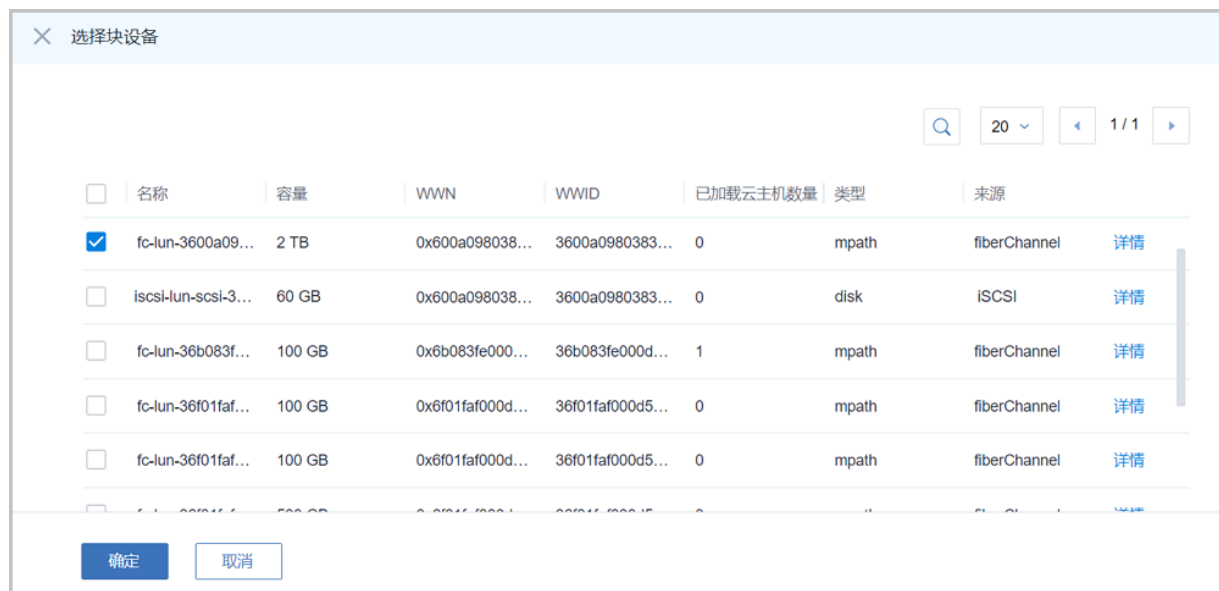
在ZStack私有云主菜单，点击**云资源池** > **云主机**按钮，进入**云主机**界面。点击需要加载块设备的云主机名称，选择**配置信息**子页面，如图 34: 配置信息页面所示：

图 34: 配置信息页面



点击块设备旁边的**操作** > **加载**按钮，进入**选择块设备**页面，选择透传的一个块设备，点击**确定**按钮，将块设备加载到云主机，如图 35: 选择块设备所示：

图 35: 选择块设备



加载块设备后，点击**操作** > **卸载**按钮，可卸载已加载的块设备。

物理机详情页加载块设备

在ZStack私有云主菜单，点击**硬件设施** > **物理机**按钮，进入**物理机**界面。点击存在块设备的物理机名称，选择**块设备**子页面，如图 36: 块设备页面所示：

图 36: 块设备页面



选择一个块设备，点击块设备旁边的**操作** > **加载云主机**按钮，进入**选择云主机**页面，选择需要加载的云主机，点击**确定**按钮，完成块设备加载。如图 37: 加载云主机所示：

图 37: 加载云主机



加载块设备后，点击块设备旁边的**操作** > **卸载云主机**按钮，可将已加载的块设备从云主机卸载。

至此，块设备透传功能介绍完毕。

3.5 典型应用场景

块设备透传可以使不同节点的应用可以对数据进行一致性访问，一般用于如下使用场景：

- 对数据块存储有共享要求的应用软件或者系统软件，例如：Oracle RAC、微软故障转移集群 (MSCS) 等
- 集群文件系统中，例如：OCFS2、GPFS、GFS2等

3.5.1 微软故障转移集群

故障转移群集是一组独立的计算机，这些计算机相互协作以提高群集角色（之前称为应用程序和服务）的可用性和可伸缩性。群集服务器（称为节点）通过物理电缆和软件连接。如果一个或多个群集节点出现故障，其他节点就会开始提供服务（该过程称为故障转移）。此外，会主动监视群集角色，以验证它们是否正常工作，如果不工作，则会重启这些角色或将其移动到其他节点。



注：微软故障转移集群目前需要云主机处于同一反亲和组中，即云主机不能位于同一台物理机中。

故障转移群集还提供群集共享卷 (CSV) 功能，该功能提供一致的分布式命名空间，群集角色可以使用这样的命名空间，从所有的节点访问共享存储。借助故障转移群集功能，用户将会在服务中体验到最低程度的中断。

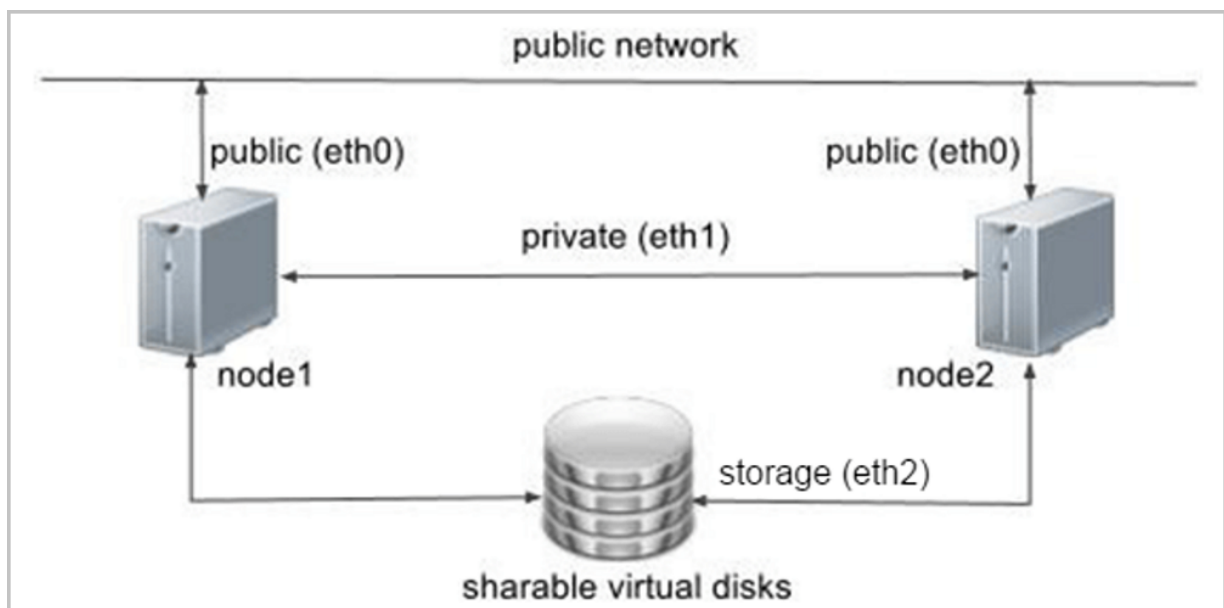
由于Windows故障转移集群对存储有较为特殊的需求，因而推荐使用FC透传或iSCSI透传的方式为云主机提供共享磁盘，或者使用独立的共享文件系统挂载至云主机提供集群仲裁与数据服务。详情请参考[微软官网文档](#)。

3.5.2 Oracle RAC

Oracle RAC (Oracle real application clusters) 最初旨在提供改进的数据库服务，经过多年的发展，如今它已立足于全面的高可用性 (HA) 体系基础上，该 HA 体系可用作数据库云系统的基础，也可以作为共享基础架构确保为数据中心中的所有应用程序提供高度的可用性、可伸缩性、灵活性和敏捷性。Oracle RAC 采用共享磁盘架构，因此卷管理和用于存储数据库数据的文件系统必须支持集群识别。Oracle 自动存储管理 (ASM) 是为 Oracle 数据库推荐的 (集群) 卷管理器。

由于ASM可以自动管理共享磁盘，因而用户可采用FC/iSCSI LUN透传的方式或者共享虚拟磁盘为RAC节点提供裸设备，交由RAC进行数据管理而无需额外部署共享文件系统。如图 38: 工作原理所示：

图 38: 工作原理



详细内容可参考[Oracle官网文档](#)。

术语表

区域 (Zone)

ZStack中最大的一个资源定义，包括集群、二层网络、主存储等资源。

集群 (Cluster)

一个集群是类似物理主机 (Host) 组成的逻辑组。在同一个集群中的物理主机必须安装相同的操作系统 (虚拟机管理程序, Hypervisor)，拥有相同的二层网络连接，可以访问相同的主存储。在实际的数据中心，一个集群通常对应一个机架 (Rack)。

管理节点 (Management Node)

安装系统的物理主机，提供UI管理、云平台部署功能。

计算节点 (Compute Node)

也称之为物理主机 (或物理机)，为云主机实例提供计算、网络、存储等资源的物理主机。

主存储 (Primary Storage)

用于存储云主机磁盘文件的存储服务器。支持本地存储、NFS、Ceph、Shared Mount Point等类型。

镜像服务器 (Backup Storage)

也称之为备份存储服务器，主要用于保存镜像模板文件。建议单独部署镜像服务器。

镜像仓库 (Image Store)

镜像服务器的一种类型，可以为正在运行的云主机快速创建镜像，高效管理云主机镜像的版本变迁以及发布，实现快速上传、下载镜像，镜像快照，以及导出镜像的操作。

云主机 (VM Instance)

运行在物理机上的虚拟机实例，具有独立的IP地址，可以访问公共网络，运行应用服务。

镜像 (Image)

云主机或云盘使用的镜像模板文件，镜像模板包括系统云盘镜像和数据云盘镜像。

云盘 (Volume)

云主机的数据盘，给云主机提供额外的存储空间，共享云盘可挂载到一个或多个云主机共同使用。

计算规格 (Instance Offering)

启动云主机涉及到的CPU数量、内存、网络设置等规格定义。

云盘规格 (Disk Offering)

创建云盘容量大小的规格定义。

二层网络 (L2 Network)

二层网络对应于一个二层广播域，进行二层相关的隔离。一般用物理网络的设备名称标识。

三层网络 (L3 Network)

云主机使用的网络配置，包括IP地址范围、网关、DNS等。

公有网络 (Public Network)

由因特网信息中心分配的公有IP地址或者可以连接到外部互联网的IP地址。

私有网络 (Private Network)

云主机连接和使用的内部网络。

L2NoVlanNetwork

物理主机的网络连接不采用Vlan设置。

L2VlanNetwork

物理主机节点的网络连接采用Vlan设置，Vlan需要在交换机端提前进行设置。

VXLAN网络池 (VXLAN Network Pool)

VXLAN网络中的 Underlay 网络，一个 VXLAN 网络池可以创建多个 VXLAN Overlay 网络 (即 VXLAN 网络) ，这些 Overlay 网络运行在同一组 Underlay 网络设施上。

VXLAN网络 (VXLAN)

使用 VXLAN 协议封装的二层网络，单个 VXLAN 网络需从属于一个大的 VXLAN 网络池，不同 VXLAN 网络间相互二层隔离。

云路由 (vRouter)

云路由通过定制的Linux云主机来实现的多种网络服务。

安全组 (Security Group)

针对云主机进行第三层网络的防火墙控制，对IP地址、网络包类型或网络包流向等可以设置不同的安全规则。

弹性IP (EIP)

公有网络接入到私有网络的IP地址。

快照 (Snapshot)

某一个时间点上某一个磁盘的数据备份。包括自动快照和手动快照两种类型。