

2014年3月21日星期五

HUAWEI ENTERPRISE ICT SOLUTIONS **A BETTER WAY**

GPU直通图形工作站解决方案亮点

enterprise.huawei.com

HUAWEI TECHNOLOGIES CO., LTD.



Content

1

VDI图形工作站面临的挑战及解决方案

2

相关案例

机车制造设计业务流程

HUAWEI ENTERPRISE ICT SOLUTIONS A BETTER WAY

设计研发

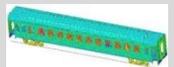
零部件设计



转向架设计



车头设计



车体设计

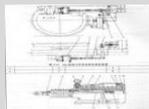
.....

集成总装

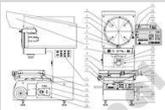


总装设计

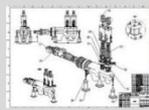
投影二维



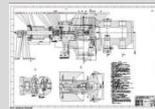
转向架投影



车头投影



车体投影



总装投影

.....

零件打开、拖动、
旋转、缩放、增删

模型打开、拖动、旋转、
缩放、零件增删

模型打开、拖动、旋转、
缩放、投影二维



生产加工

VDI图形工作站面临的挑战



挑战一：中大型模型文件打开慢

使用虚拟图站打开中大型模型文件耗时长

模型	物理图站	客户可接受标准	虚拟图站
	规格：w3670 2.7G CPU 6core, 8GB内存, AMD FirePro V5800 2G显存		规格：E5-2680 2.7G CPU 4core, 12GB内存, NVIDIA Q2000 1GB显存
	打开时长		
小型模型（930个零件）	48秒	60秒	56秒
中型模型（3327个零件）	1分06秒	1分23秒	1分40秒
大型模型（6302个零件）	3分12秒	4分	4分13秒

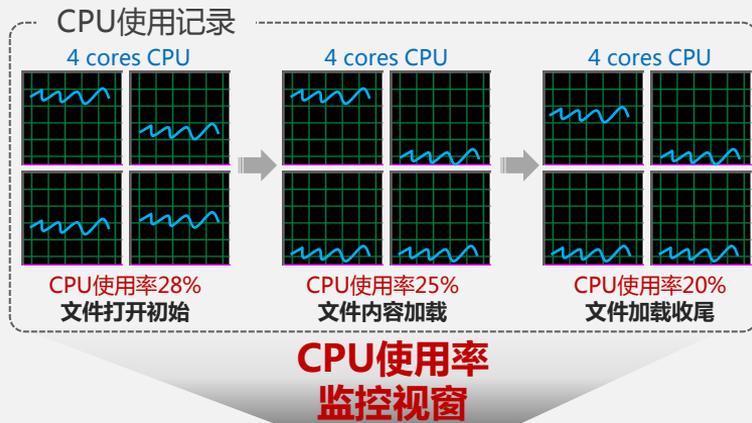
结论：中型、大型模型的文件打开速度，无法达标客户可接受标准。

使用虚拟图站打开文件技术指标分析

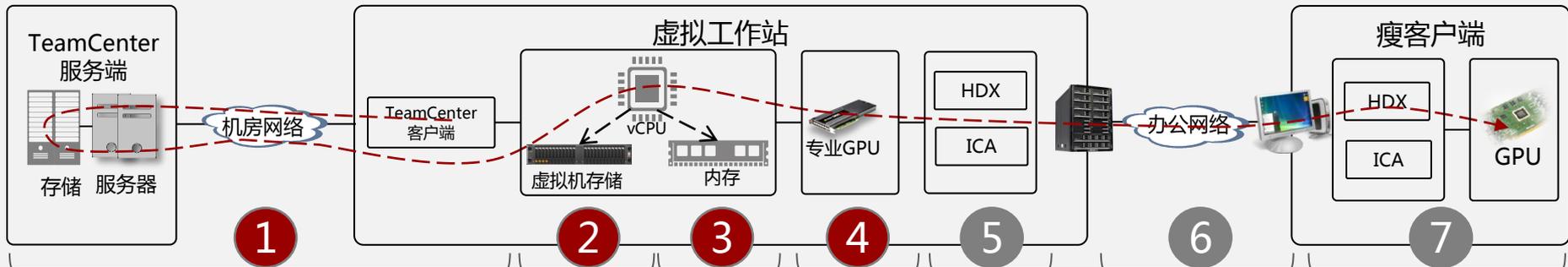
模型	虚拟图站规格： E5-2680 2.7G CPU 4core, 12GB内存, NVIDIA Q2000 1GB显存	
	打开时长	CPU/内存
小型模型（930个零件）	56秒	20%~28%/2.53~2.63G
中型模型（3327个零件）	1分40秒	20%~28%/3.0~3.12G
大型模型（6302个零件）	4分13秒	20%~28%/2.95~3.24G

数据分析结论：

- 1、打开文件操作时，采用CPU多核处理；
- 2、在整个打开过程中，CPU单核处理所占比例较大，且该单核负荷已经达到100%。



模型文件打开慢关键因素分析



1. 数据获取

- TeamCenter请求模型文件
- TeamCenter服务器从对应的存储获取模型相关数据
- TeamCenter服务器将获取的数据传送给虚拟工作站，虚拟工作站CPU4个核并行接收数据

1分32秒 ✓耗时点

2. 写入磁盘

- 写磁盘的I/O操作排队等待
- CPU4个核的并发处理

3. 读入内存

- CPU单核
- 100%负荷
- 将文件数据读入内存

1分22秒 ✓耗时点

4. 模型显示处理

- GPU图形渲染，放入显卡缓冲区

5. 图像压缩

6. 数据传输

100毫秒

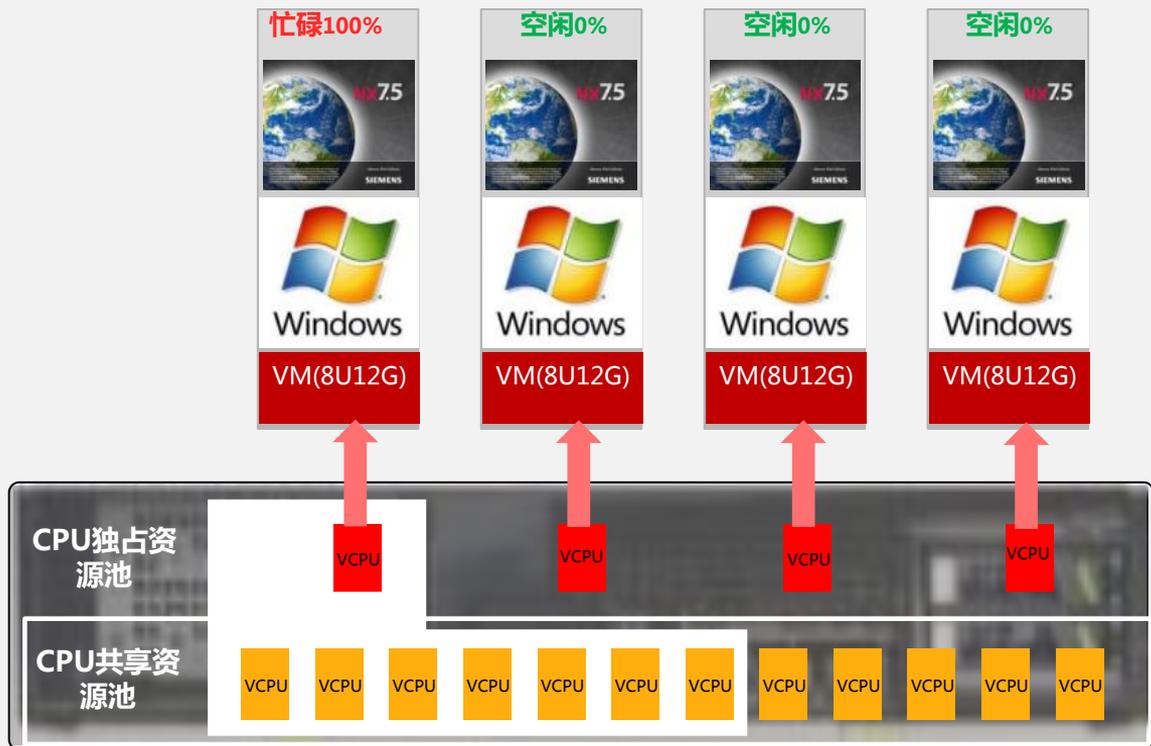
7. 显示模型

耗时分析：I/O并发量大，CPU核数少

耗时分析：内存读入量大，CPU抢占费时长

打开大型模型耗时分析

解决之道：增加虚拟机共享内核+独占CPU单核提升文件打开速度



解决方案设计点

□增加虚拟机共享内核，多核并发处理I/O请求

- 虚拟机CPU配置至少8核
- 12G内存

□虚拟图站独占CPU单个内核，缩减CPU抢占耗时

- 单台虚拟机独占CPU单个核，提高资源调度效率，文件打开耗时可缩短25%

合理优化虚拟机CPU配置，加快文件打开速率

优化前

模型	物理图站	客户可接受标准	虚拟图站
	规格：w3670 2.7G CPU 6core, 8GB内存, AMD FirePro V5800 2G显存		规格：E5-2680 2.7G CPU 4core, 12GB内存, NVIDIA Q2000 1GB显存
打开时长			
小型模型 (930个零件)	48秒	60秒	56秒
中型模型 (3327个零件)	1分06秒	1分23秒	1分40秒
大型模型 (6302个零件)	3分12秒	4分	4分13秒

优化策略

- 1、虚拟图站CPU共享内核配置8核
- 2、虚拟图站独占CPU单个核

优化后

模型	物理图站	虚拟图站
	规格：w3670 2.7G CPU 6core, 8GB内存, AMD FirePro V5800 2G显存	规格：E5-2680 2.7G CPU 8core, 12GB内存, NVIDIA Q2000 1GB显存
打开时长		
小型模型 (930个零件)	48秒	49秒
中型模型 (3327个零件)	1分06秒	1分11秒
大型模型 (6302个零件)	3分12秒	2分44秒

优化效果

- 1、优化后的虚拟图站与原配置虚拟图站相比，打开模型文件时长平均缩短**25%**
- 2、优化后的虚拟图站打开大型模型文件的时长**赶超物理图站**，比物理图站缩短**15%**

挑战二：图形计算耗时长

使用虚拟图站投影二维操作耗时长

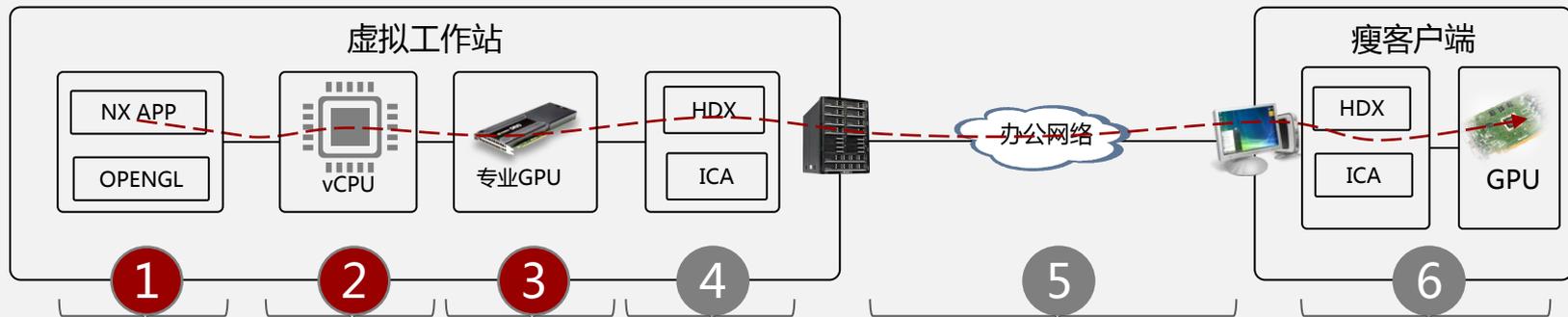
模型	物理图站	客户可接受标准	虚拟图站
	规格：X5600 2.8GHz CPU 12core, 24GB内存, NVIDIA Q2000 1GB显存		规格：E5-2630 2.7G CPU 8core, 16GB内存, NVIDIA Q2000 1GB显存
投影二维时长			
小型模型（930个零件）	30秒	38秒	49秒
中型模型（3327个零件）	10分	12分30秒	12分40秒
结论： 投影二维操作的耗时，无法达标客户可接受标准。			

使用虚拟图站投影二维技术指标分析

模型	规格：E5-2680 2.7G CPU 4core, 12GB内存, NVIDIA Q2000 1GB显存, 千兆到桌面	
	时长	CPU/内存
小型模型（930个零件）	49秒	25%/2.64G
中型模型（3327个零件）	12分40秒	25%/2.76G
数据分析结论：		
1、在整个投影二维过程中，CPU单核处理所占比例较大，且该单核负荷已经达到100%。		



图形计算耗时长关键因素分析



1.发起投影二维命令

- NX7.5调用OpenGL API进行投影二维操作

2.图形计算

- 单核CPU
- 100%负荷计算三维图像，生成二维图像对应数据

3.图形处理

- GPU图形渲染，生成二维位图，放入显卡缓冲区

4.图像压缩

5.图像传输

6.展示二维影像

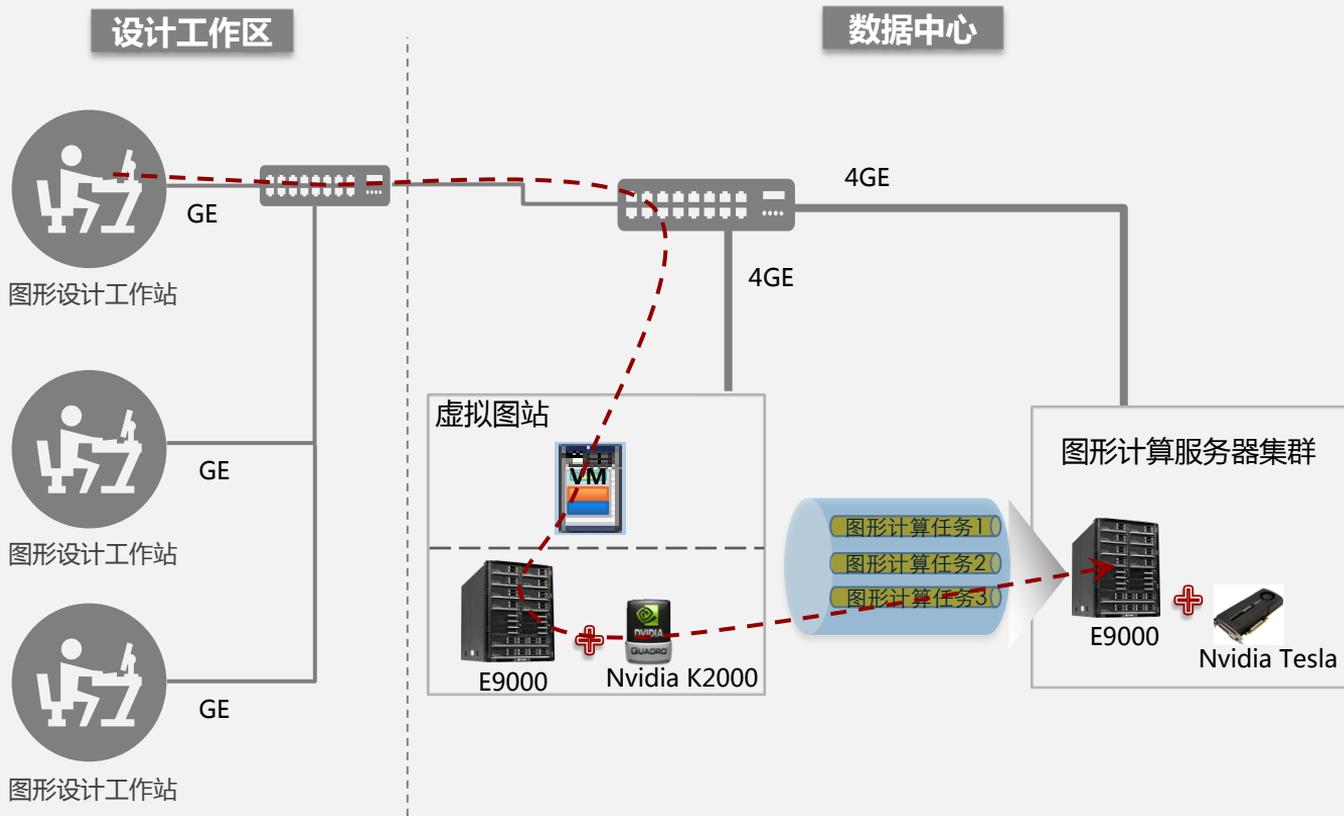
12分40秒 ✓耗时点

100毫秒

中型模型
投影二维
耗时分析

耗时分析：图形并发计算量大，纯CPU处理效率低

解决之道：GPU图形计算云化服务器加速图形计算业务



解决方案设计点

部署物理图形计算云化服务器，提升图形计算处理能力

- NVIDIA TESLA GPU图形计算加速卡直通服务器的方式，提供图形计算专属服务器集群
- 投影二维等图形计算操作被提交到后台图形计算专属服务器上排队处理，规避纯CPU处理图形计算效率低的问题

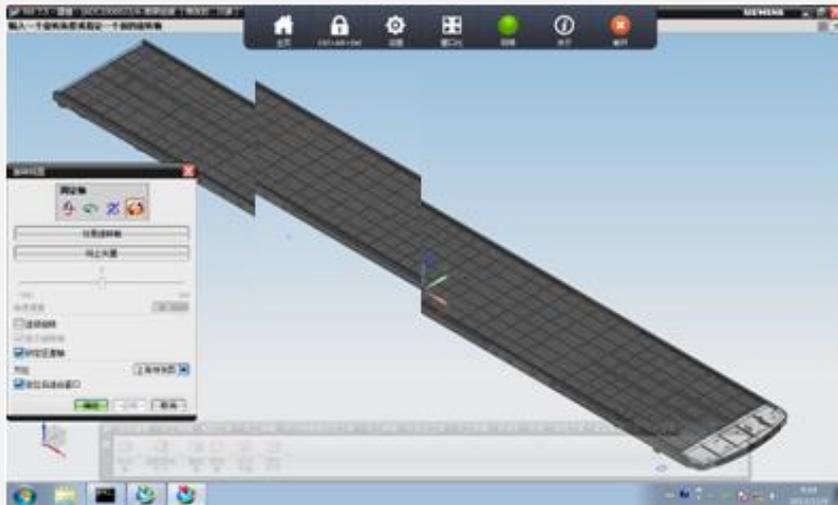
挑战三：操作不流畅

使用虚拟图站对模型进行旋转、缩放等操作时，画面不流畅

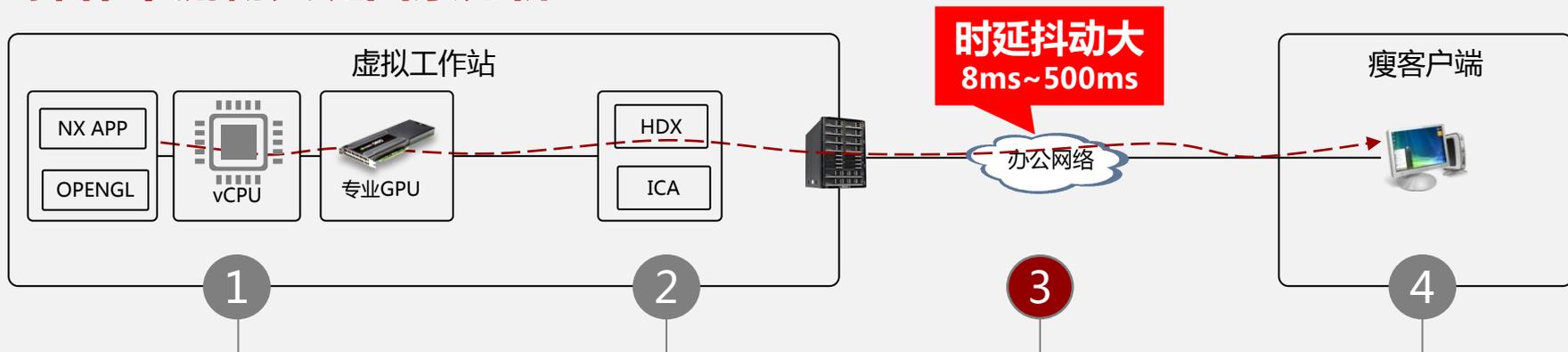
以自动旋转操作为例：

模型	流畅帧率	虚拟工作站	
		终端输出帧率	画面流畅度
小型	>15帧/秒	24帧/秒	流畅，偶有停滞
中型		10帧/秒	较流畅，偶有停滞
大型		8帧/秒	较流畅，偶有停滞

不流畅、明显停滞



操作不流畅关键因素分析



1. 图像生成

- Q2000显卡
- 小模型每秒处理约**30帧**画面
- 中型模型每秒约**8帧**画面
- 大型模型每秒约**5帧**画面
- K2000显卡
- 小模型每秒处理约**30帧**画面
- 中型模型每秒约**10帧**画面
- 大型模型每秒约**8帧**画面

2. 图像压缩

- 压缩处理，对于1080P以下的画面压缩处理时间小于**30ms**。设计最大帧频**24~25帧/S**。

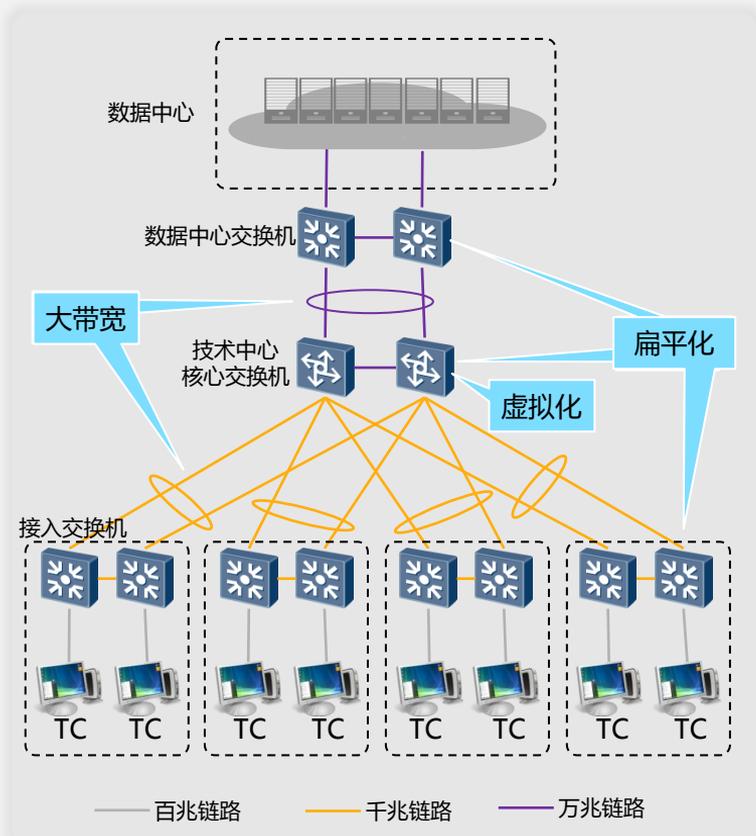
3. 图像传输

- 网络异常不断出现，导致画面卡顿。
- 时延抖动**8ms~500ms**。

4. 图像显示

- Q2000显卡
- 小模型每秒处理约**24帧**画面
- 中型模型每秒约**8帧**画面
- 大型模型每秒约**5帧**画面
- K2000显卡
- 小模型每秒处理约**24帧**画面
- 中型模型每秒约**10帧**画面
- 大型模型每秒约**8帧**画面

解决之道：减少网络跳点、提高带宽、简化网络架构

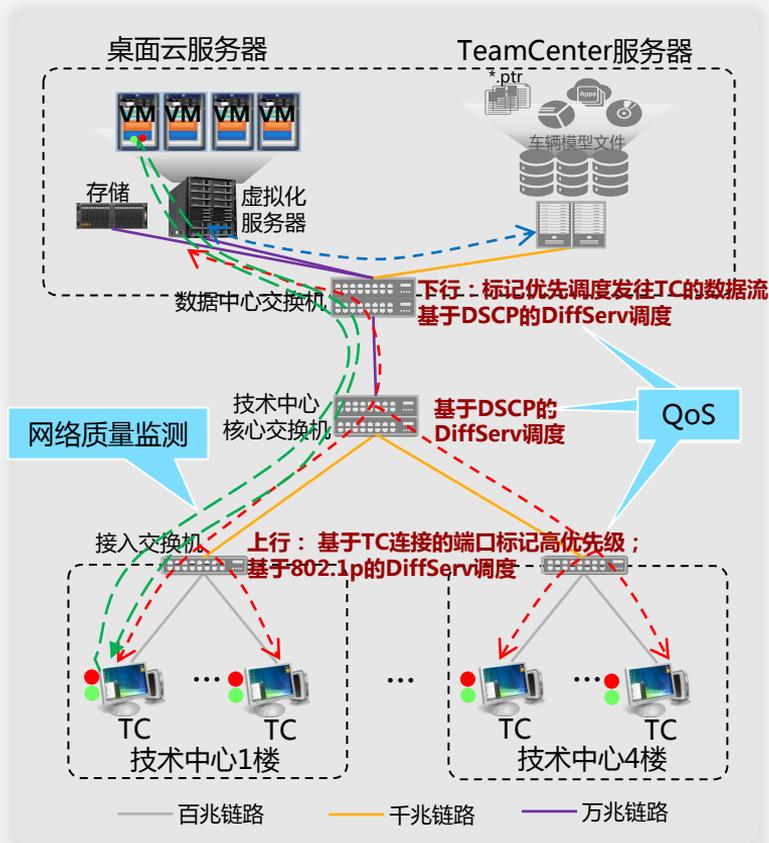


网络优化方案设计点

- **扁平化**：TC直接连接到汇聚交换机，汇聚交换机千兆堆叠，千兆双上行接入核心交换机，将TC到数据中心的网络节点减少到**3跳**，降低网络时延和抖动。
- **大带宽**：技术中心核心交换机通过**双万兆**链路 with 数据中心核心交换机互联，提高网络上行带宽，降低网络拥塞发生的概率。
- **虚拟化**：技术中心核心交换机采用双设备，启用**虚拟化**功能，形成一个逻辑上的独立实体，实现跨设备的链路捆绑，提高核心转发性能，同时简化网络拓扑，统一管理。

降低时延和抖动、减少网络拥塞，为虚拟工作站获得更好的操作体验提供基础保证。

解决之道：部署QoS，优先转发虚拟机画面数据流



QoS方案设计点

□QoS规划、部署

- **优先级标记**：数据中心交换机标记VM到TC的下行数据，用于TC画面的实时显示；接入交换机标记TC上行控制数据流。
- **队列调度**：接入交换机基于802.1p的Diffserv调度，核心交换机和数据中心交换机基于DSCP的Diffserv调度，优先转发TC与虚拟机之间的数据队列。

□网络质量监测

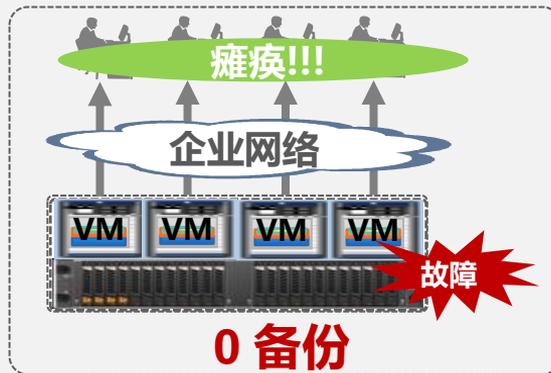
- **原理**：通过发送模拟业务数据流，实时监测网络延时及丢包率。
- **目的**：通过网络质量监测，及时发现网络拥塞段，为网络扩容、优化提供信息依据。

优先转发TC数据，实时监测网络服务质量，为虚拟工作站获得更好的操作体验提供服务保证。

挑战四：物理机故障，GPU直通VDI无法自动恢复

物理机缺乏合理备份

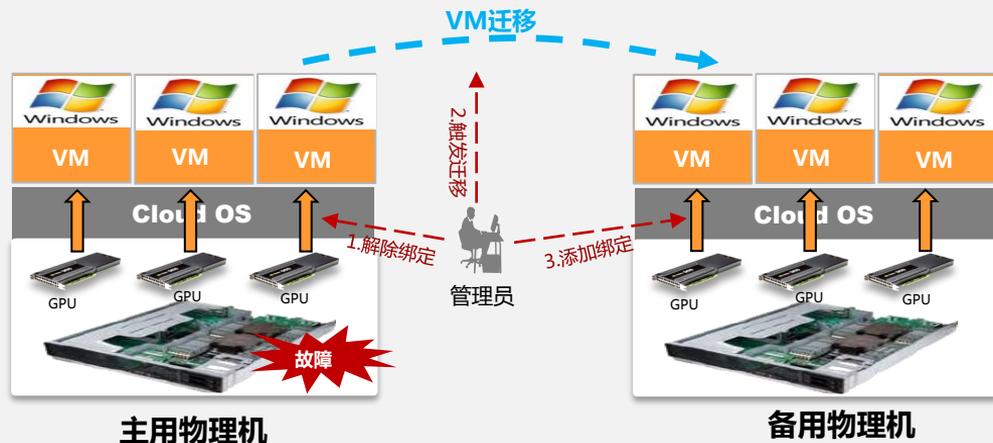
- 0备份-风险大：物理机故障导致4个VM用户无法使用，比传统网站故障（1个用户）影响范围更大；
- 1+1备份-性价比低：部署成本高（单刀片服务器10万¥以上），故障发生率低（<0.1%），备份利用率低。



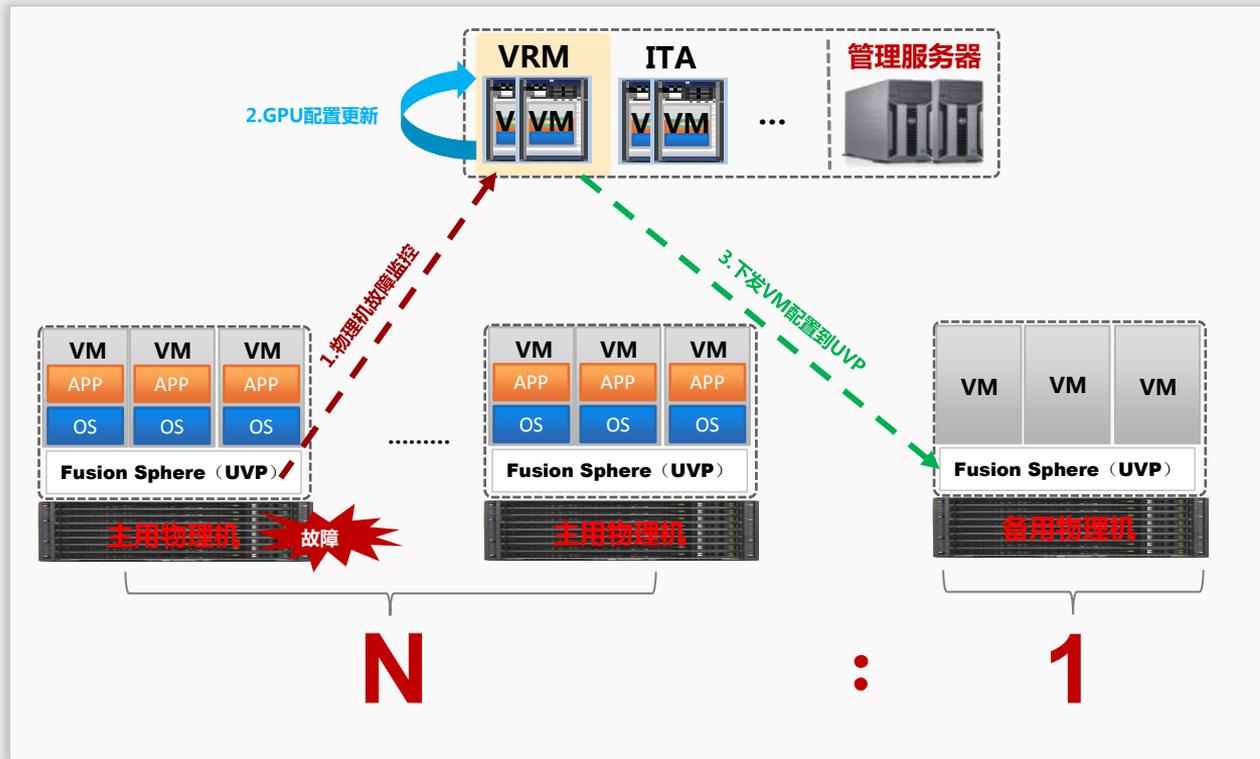
GPU直通VDI只能人工故障迁移

- VM和物理GPU绑定，故障迁移必须经过三个步骤：
 - VM和原物理GPU解绑；
 - VM迁移到新物理机；
 - VM和新物理GPU绑定。

目前1和3不支持自动化配置。



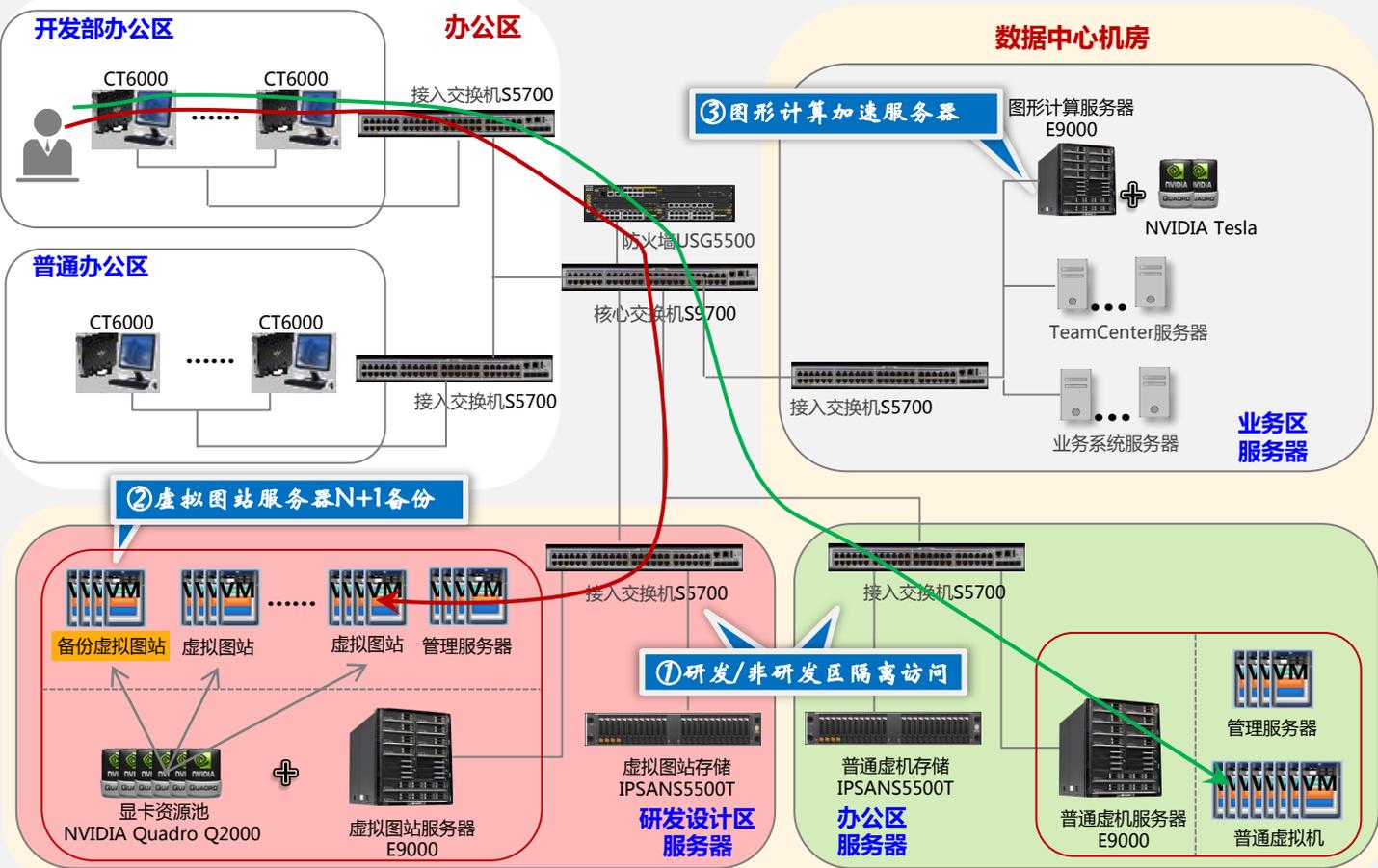
解决之道：GPU配置自动化，实现VM自动故障迁移



方案设计点

- **设计点1：物理机N+1备份配置**
 - ✓ 部署1台备用物理机，硬件上镜像配置，软件上不配置VM资源，平时处于待机状态。
- **设计点2：通过VRM自动更新GPU配置**
 - ✓ **GPU配置更新**：通过管理服务器VRM监控物理机状态，将故障VM的GPU配置修改为备用物理机上的GPU；
 - ✓ **VM配置下发**：VRM将VM新配置下发到UVP，后者进行VM资源调度。

虚拟图形工作站解决方案整体架构



方案关键点

- ①效率:** 数据中心安全隔离访问
数据中心研发区/非研发区安全隔离设计, 实现开发部研发人员可在办公位同时访问研发区和非研发区虚拟工作站, 提高办公**效率**
- ②可靠:** 虚拟图站服务器N+1备份
专门配置一台插入显卡的物理服务器作为备用机, 在虚拟图站物理服务器故障时可以及时恢复使用, 保证业务**可靠性**
- ③体验:** 图形计算加速云化服务器
图形计算加速卡与物理服务器组建图形计算专用服务器, 提升图形计算效率, 保障研发人员设计**体验**

Content

1

VDI图形工作站面临的挑战及解决方案

2

相关案例

株洲南车案例



南车株洲电力机车有限公司简介

- 中国南车株洲电力机车有限公司创建于1936年，国家特大型一类企业。
- 公司主营业务为干线铁路电力机车、动车组和城市轨道交通装备的研发、制造、维修、销售与售后服务。
- 公司本部占地2.25平方公里，现有员工10000余人，近2000名技术人员。

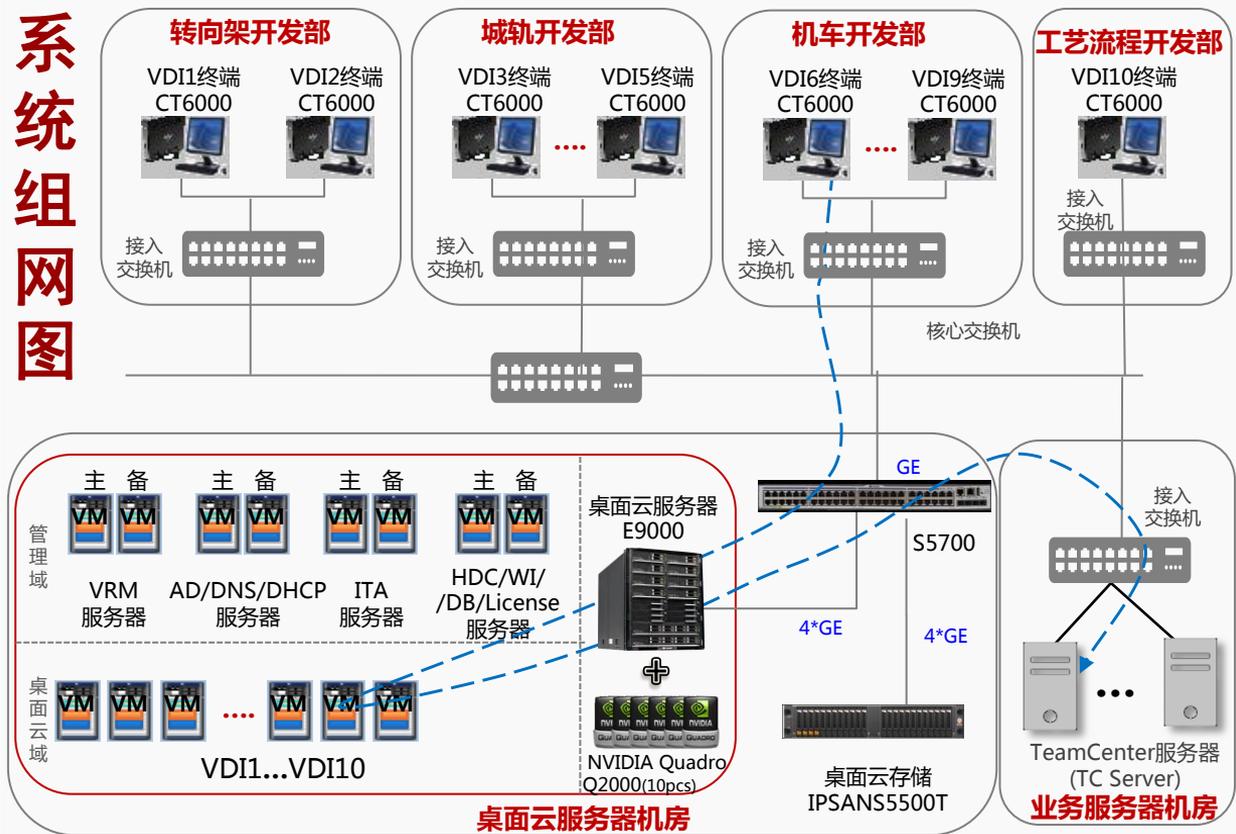
虚拟机配置需求

CPU：6core以上
内存：8GB以上
磁盘：C盘容量100G以上，总容量300G~500G

业务需求

GPU虚拟机与物理图形工作站做对比：
1、打开、投影二维操作时长差距控制在25%以内
2、旋转、缩放、增删、保存等操作流畅，时延控制在2-3秒以内

系统组网图





HUAWEI ENTERPRISE ICT SOLUTIONS **A BETTER WAY**

Copyright©2012 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.