

# Использование технологии VDI для работы с большими сеточными моделями

П. И. Каледа, В. Д. Никоноров

АО «НИКИЭТ»

Для работы с "большими" сеточными моделями необходимо, чтобы рабочая станция пользователя обладала большим объемом оперативной памяти и высокопроизводительной графической картой. Кроме того, мы выдвинули требование непосредственного подключения к файловой системе суперкомпьютера, чтобы избежать копирования данных между рабочей станцией для сеточных моделей и суперкомпьютером. При помощи технологии виртуализации рабочих мест (VDI) можно обеспечить удаленную работу с виртуальными рабочими местами, удовлетворяющими этим требованиям. В настоящей работе представлены описание стенда для тестирования данного подхода и результаты испытаний.

## 1. Введение

При выполнении задач численного моделирования качество и размер расчетной сеточной модели непосредственно влияют на точность получаемых результатов. Размеры сеточных моделей постоянно увеличиваются в связи с повышающимися требованиями к точности. Современные размеры сеточных моделей таковы, что при использовании сеточных пре/постпроцессоров, работающих с общей памятью (ANSYS Meshing / ANSYS Icem CFD и аналогов), объема оперативной памяти и графических ресурсов, доступных обычно на персональных компьютерах, становится недостаточно. Поэтому используются дорогостоящие специализированные рабочие станции с высокопроизводительными графическими картами. Но стоимость таких станций делает нецелесообразной установку их каждому пользователю, работающему с большими сеточными моделями.

С другой стороны, при использовании в работе с пре/постпроцессором данных, полученных в результате расчетов на суперкомпьютере, необходимо иметь непосредственное подключение к системе хранения данных суперкомпьютера (в данной работе использовалась файловая система Lustre [1]), чтобы избежать копирования данных между суперкомпьютером и машиной, на которой происходит работа с сеточными моделями.

В совокупности эти обстоятельства – высокая стоимость рабочей станции с достаточным количеством ресурсов и необходимость подключения к системе хранения данных суперкомпьютера – приводят к идее централизации ресурсов, требующихся для организации рабочих мест. Такой подход предусматривает наличие рядом с суперкомпьютером нескольких рабочих станций, подключенных к системе хранения данных суперкомпьютера, на которых пользователи могут осуществлять решение своих задач удаленно. При этом рабочие станции не закрепляются за конкретными пользователями, а являют собой некий общедоступный разделяемый ресурс.

## 2. Описание проблемы

При организации удаленного доступа к приложениям с повышенными требованиями к графической подсистеме основной проблемой становится доставка рабочего стола, на котором работает приложение, до пользователя. В случае, когда удаленное приложение использует 3D-графику, оно, как правило, не может полноценно воспользоваться функционалом GPU. Например, обработка 3D-команд происходит на центральном процессоре сервера (процессор может выполнять эту задачу, но делает это не так эффективно, как GPU), или приложение передает 3D-команды на клиентское устройство [2] (количество информации, которую приложению необходимо при этом передавать для отображения картинки на экран, становится

настолько внушительным, что отзывчивость этого удаленного рабочего стола будет низкой или очень низкой, какой бы протокол для удаленного доступа к рабочему столу ни использовался).

### **3. Подходы к решению проблемы**

Для решения проблемы низкой отзывчивости необходимо исключить обработку 3D-команд на центральном процессоре и передачу 3D-команд от приложения через локальную сеть к дисплею пользователя. Это делается путем прямой обработки 3D-команд локальным GPU сервера и последующей отсылки на клиентское устройство уже сформированного на GPU растрового изображения.

В ходе исследования рынка рассмотрены открытая технология, реализующая этот функционал (связка перехватчика 3D-команд VirtualGL и инструмента для удаленного доступа TurboVNC) и проприетарные решения: VMware Horizon View и Citrix XenDesktop. Два последних решения не просто предоставляют средства доставки пользователю удаленного рабочего стола, а сочетают такие средства с виртуализацией рабочих мест (VDI).

Решения с использованием виртуализации рабочих мест представляются более гибкими и эффективными в плане распределения ресурсов между пользователями, интеграции с существующей информационной инфраструктурой предприятия, организации управления доступом пользователей к рабочим местам и выбора программного обеспечения, используемого на удаленных рабочих станциях. Решение связки VirtualGL+TurboVNC не имеет встроенных средств разделения ресурсов между несколькими пользователями и не имеет готовых средств интеграции с системой централизованной авторизации, используемой в ОАО НИКИЭТ (доменная сеть под управлением Microsoft Active Directory). Решения, совмещающие доставку приложений или рабочих столов пользователю с виртуализацией рабочих мест, как правило, не используют собственную подсистему аутентификации, а интегрируются в инфраструктуру предприятия, например, Microsoft Active Directory

### **4. Выбор подхода**

Решения по виртуализации рабочих мест (VDI) предлагают многие компании: VMware [3], Citrix [4], Microsoft [5], RedHat [6] и другие, но классическая виртуализация рабочих мест, использующаяся уже достаточно давно, обычно не предполагает интенсивную нагрузку на графическую подсистему. Так, решение от RedHat вообще не поддерживает 3D-графику, а решение от Microsoft плохо подходит для приложений, которые используют графический интерфейс OpenGL, так как ориентировано на использование интерфейса DirectX. Citrix XenServer официально не поддерживает работу с сетью стандарта InfiniBand (которая используется в том числе для доступа к данным), и, следовательно, не способен обеспечить многопользовательский высокоскоростной доступ к файловому хранилищу суперкомпьютера.

Итак, остается решение на базе VMware vSphere с доставкой при помощи VMware Horizon View или Citrix XenDesktop. В связи со сложностями, возникшими при настройке взаимодействия между XenDesktop и гипервизором vSphere, выбор был сделан в пользу Horizon View. Виртуализация рабочих мест на платформе vSphere позволяет получить максимум от графического процессора NVIDIA и в то же время обеспечить быстрый доступ к файловому хранилищу суперкомпьютера. Также отметим, что данное решение интегрируется с Active Directory, что существенно упрощает задачу управления пользовательскими учетными записями.

### **5. Описание эксперимента**

#### **5.1. Аппаратное обеспечение для работы с графикой**

С точки зрения плотности размещения виртуальных рабочих мест, специализированные решения для виртуализации NVIDIA GRID K1 и K2 имеют преимущество перед классическими картами линейки Quadro для рабочих станций, поскольку карты GRID имеют на борту

несколько (4 для GRID K1 и 2 для GRID K2) графических процессоров, управление каждым из которых можно передать отдельной виртуальной машине.

## 5.2. Общая схема протестированного решения

Принципиальная схема протестированного решения изображена на рис. 1. Виртуальное рабочее место доставляется пользователю по сети предприятия. Графический процессор NVIDIA предоставляется в монопольное распоряжение виртуальной машины (технология VMDirectPath), чтобы 3D-приложение могло использовать его напрямую.

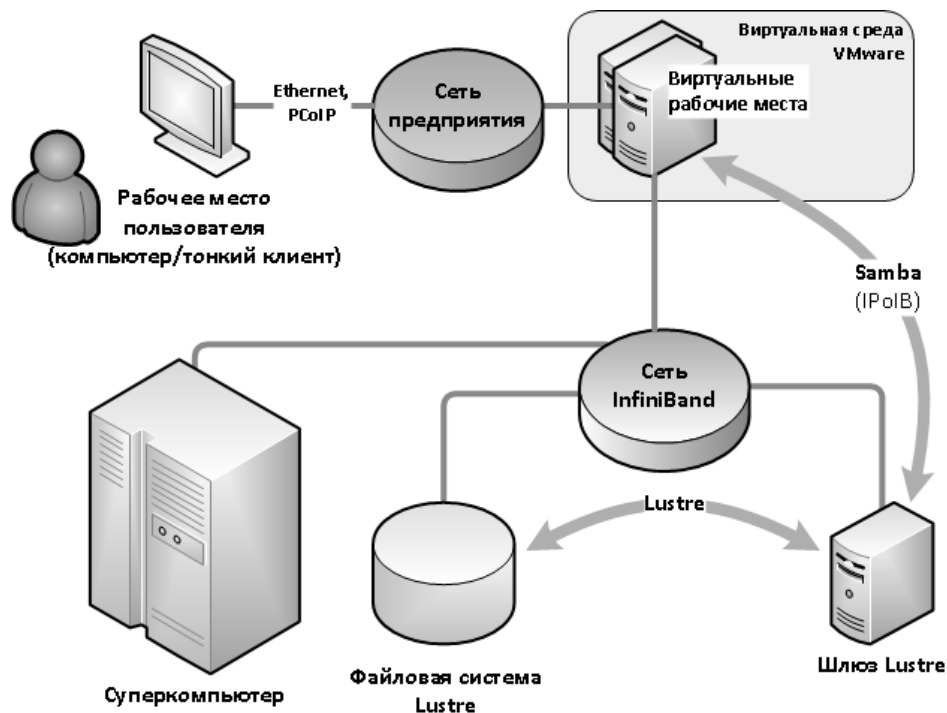


Рис. 1. Принципиальная схема решения

## 5.3. Подключение виртуальных рабочих мест к файловому хранилищу суперкомпьютера

Следующий вопрос – подключение к системе хранения данных суперкомпьютера, которая построена на основе параллельной файловой системы Lustre. Важно, что и VMware Horizon View, и Citrix XenDesktop поддерживают доставку лишь рабочих мест под управлением Windows. В связи со сложностями, возникающими при настройке доступа windows-пользователей к linux-объектам (пользователям рабочих станций нужно иметь доступ к своим папкам в ФС Lustre), наиболее простым подходом к решению задачи предоставления доступа представляется использование «шлюза». Под шлюзом в данном случае подразумевается промежуточный Linux-сервер, который подключается к Lustre-хранилищу и реэкспортирует ресурс, к которому он получает доступ, виртуальной машине по протоколу CIFS. Подключение шлюза к системе хранения данных суперкомпьютера организуется посредством программного Lustre-клиента аналогично узлам кластера.

Возможны различные варианты организации шлюза:

- Виртуальный шлюз Lustre. В эту виртуальную машину средствами гипервизора передается InfiniBand-адаптер, и с гостевой операционной системы работает с ним напрямую. Виртуальным рабочим местам данные предоставляются по протоколу CIFS через внутреннюю сеть гипервизора.
- Физический шлюз Lustre и доступ к CIFS-ресурсам по каналу InfiniBand. Гипервизор VMware ESXi может предоставлять ресурсы InfiniBand-адаптера виртуальным

машинам посредством виртуального сетевого адаптера, который гостевая операционная система видит как десятигигабитный Ethernet-адаптер.

- Физический шлюз Lustre и доступ к CIFS-ресурсам по каналу Ethernet (10 Гбит/с).

Третий вариант имеет смысл в том случае, если на предприятии уже имеется сетевая инфраструктура, построенная с использованием стандарта 10G Ethernet. В нашем случае такой инфраструктуры не было, но присутствовала InfiniBand-инфраструктура, поэтому выбор был сделан в пользу варианта 2.

#### 5.4. Описание аппаратной части

Экспериментальный стенд был построен на серверной платформе HP ProLiant DL 380p Gen8, его конфигурация указана в табл. 1.

Таблица 1. Состав сервера HP ProLiant DL 380p Gen8

Процессор	2 процессора Intel Xeon E5-2630v2 6C 2.60 15MB
Оперативная память	320GB(20x16GB) DDR3-1600MHz
Графическая карта	1 карта NVIDIA GRID K2 (2 графических процессора, аналогичных Quadro K5000)
Диски	4 диска по 300GB, 15000 RPM
Адаптер InfiniBand	HP Infiniband QDR/Ethernet 10Gb 2-port 544FLR-QSFP Adapter

#### 5.5. Описание виртуальных машин, развернутых на стенде

На стенде были развернуты следующие машины:

- Два виртуальных рабочих места с ОС Windows 7 64bit. Именно они доставлялись пользователям, и на них осуществлялось построение сеточных моделей. **Характеристики рабочего места vDesktop-1:** 6 виртуальных процессоров, графический процессор NVIDIA GRID K2; ~192 ГБ оперативной памяти (это максимум оперативной памяти, который поддерживается Windows 7; для Windows 8 ограничение по оперативной памяти составляет 512 ГБ; для Windows Server 2012 — 4 ТБ [7]). **Характеристики рабочего места vDesktop-2:** 6 виртуальных процессоров, ~100 ГБ оперативной памяти, графический процессор NVIDIA GRID K2.
- VMware vCenter Server – сервер управления виртуальной средой VMware vSphere.
- VMware Horizon View Connection Server – диспетчер подключений к виртуальным рабочим местам.
- Контроллер домена – необходим для авторизации пользователей и хранения учетных записей виртуальных машин.
- Маршрутизатор – виртуальная машина, необходимая для маршрутизации пакетов между различными сетями стенда.

#### 5.6. Подключение виртуальных машин к сети предприятия и высокоскоростной сети суперкомпьютера

На стенде были созданы виртуальные сетевые коммутаторы для следующих TCP/IP-сетей:

- Внутренняя сеть VMware Horizon View (vSwitch0). К ней были подключены виртуальные рабочие места и все инфраструктурные сервера.

- Сеть предприятия (vSwitch1). К ней был подключен VMware Horizon View Connection Server, к которому пользователи обращались для подключения к виртуальным рабочим местам.
- Сеть Lustre (vSwitch2). Использовалась для подключения виртуальных рабочих мест к шлюзу Lustre.
- Внутренняя сеть суперкомпьютера (vSwitch3). Через эту сеть осуществлялось управление физическим сервером.

Упрощенная схема сети TCP/IP представлена на рис. 2.

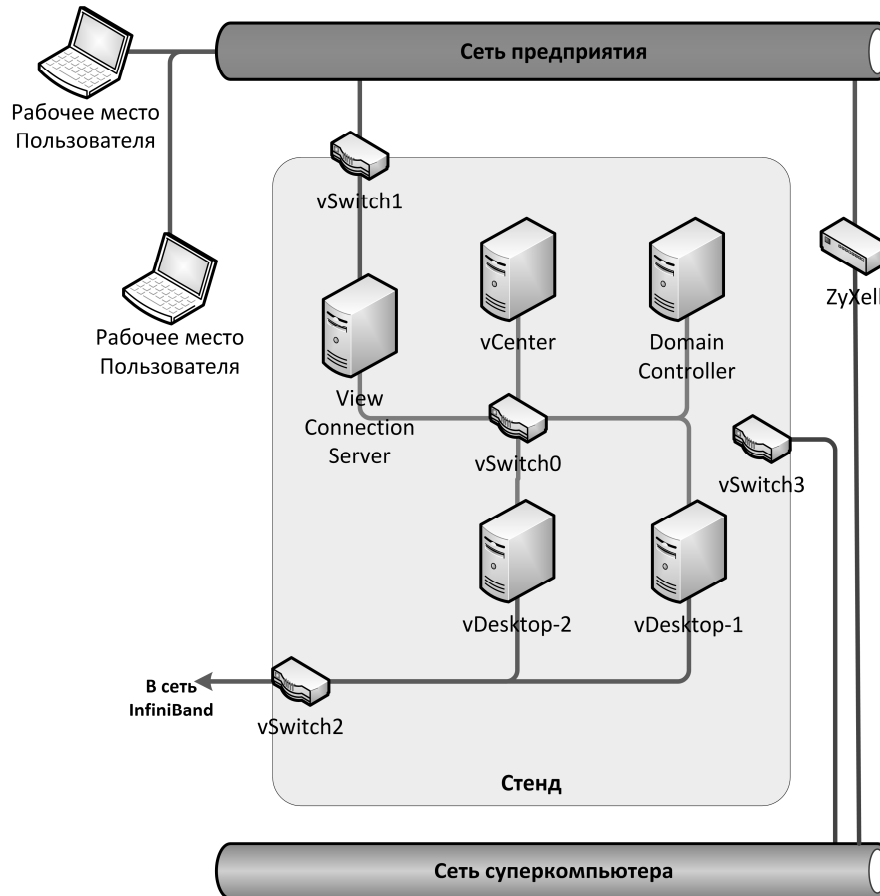
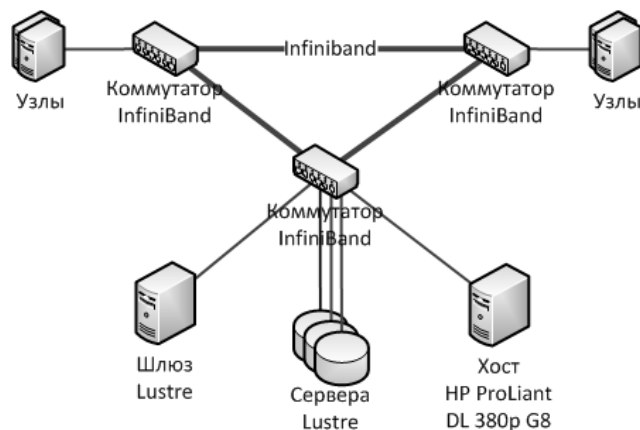


Рис. 2. Схема сети Ethernet

Схема сети InfiniBand (высокоскоростной коммуникационной сети суперкомпьютера) приведена на рис. 3. Сеть включает в себя три коммутатора, к каждому из них подключены двадцать узлов кластера, в том числе серверы Lustre. Между собой каждые два коммутатора соединены избыточными связями (эти соединения обозначены на рисунке более жирными линиями).



**Рис. 3.** Схема сети InfiniBand

Шлюз подключен к коммутатору, обслуживающему серверы Lustre.

## 6. Результаты тестирования

В тестировании принимали участие представители различных подразделений АО НИКИЭТ.

В качестве окончательного оборудования на клиентской стороне можно использовать как программные клиенты, так и тонкие/нулевые клиенты. В нашем тестировании использовались программные клиенты VMware Horizon View Client и тонкие клиенты HP t610 Plus (E4T96AA).

### 6.1. Общие тесты производительности

Была проведена серия тестов виртуального рабочего места vDesktop-1. Для этого использовалось ПО SPECwrc; результаты приведены в табл. 2.

**Таблица 2.** Ускорение в сравнении с рабочими станциями

Тест	HP Z230 Intel Xeon E3-1280 v3 3.6 GHz, 16 ГБ оперативной памяти, графическая карта Quadro K4000	HP Z420 Intel Xeon E5-1660 v2 3.7GHz, 32 ГБ оперативной памяти, графическая карта Quadro K5000
Entertainment (Maya, HD Video, Tessellation):	1.1x	0.8x
Financial (использование финансовых программ: стат. обработка больших баз данных)	1.2x	0.9x
CAD (SolidWorks, SiemensNX, Catia, Showcase)	1.5x	1.0x
CAE (расчеты CFD)	1.7x	1.15x
Energy (в основном расчеты на CPU)	2.1x	1.3x
Уравнения Пуассона	4.0x	2.7x
LifeSciences (расчеты на CPU сложных математических задач, постобработка больших баз данных)	1.5x	1.0x
General (7zip, компилирование и	2.3x	1.9x

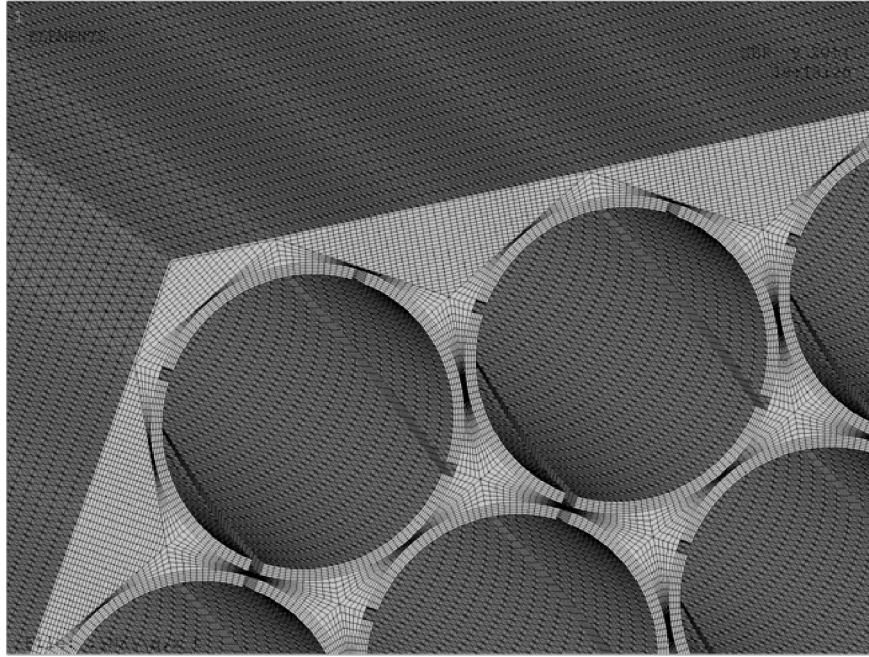
Из таблицы можно сделать следующие выводы:

- Виртуальная машина при решении реальных инженерных задач сопоставима по производительности с рабочей станцией уровня HP Z420
- В некоторых задачах с повышенными требованиями к CPU, таких как решение уравнений Пуассона, производительность виртуальной машины оказалась существенно выше, чем производительность Z420. Причина в том, что кэш процессора E5-2630 v2 в сервере VDI много больше кэша E3-1280 v3 и E5-1660 v2, установленных на рабочих станциях Z230 и Z420 соответственно.
- Производительность графики соизмерима с производительностью графики на аппаратных решениях, падения производительности за счет виртуализации нет или она незначительна (~10 %).
- Производительность CPU на реальных задачах также достаточна. Влияние системы виртуализации незаметно (E5-1660v2 быстрее E5-2630v2 на 30 %, а на результатах расчетов тестовых задач такой большой разницы нет).
- Наблюдается некоторая задержка в передаче видеопотока от виртуальной машины к пользователю. Возможно, «обычного» тонкого клиента для передачи изображения высокого разрешения недостаточно. Необходимо протестировать клиенты с аппаратной поддержкой протокола PCoIP.

## 6.2. Работа с сеточными моделями

Представителем расчетного подразделения были произведены тесты по использованию виртуального рабочего места vDesktop-2 для разработки сеточных CAE моделей ANSYS CFX. В качестве расчетных моделей использовались условия экспериментов по проливкам тепловыделяющей сборки из 19 стержней в ИТ СО РАН и продувкам сборки из 37 стержней в МГТУ на кафедре Э7. Размеры сеточных моделей составляли 10...100 млн элементов, размер занимаемой оперативной памяти – 1 млн элементов  $\approx$  1 ГБ.

На физической рабочей станции HP Z620 (2 процессора Intel Xeon E5 2640 v2, 48 ГБ оперативной памяти) комфортно работать, когда занимаемый сеточной моделью объем оперативной памяти не превышает 40 ГБ. Напротив, работа на виртуальной машине позволяла весьма эффективно снять пики загрузки операций с сеточными моделями, требующим до 80 ГБ оперативной памяти. В результате расчетов на моделях в десятки миллионов ячеек были получены сведения о механизмах межъячеистого обмена массой теплоносителя и их количественные характеристики. Полученные результаты хорошо согласуются с результатами экспериментов.



**Рис. 4.** Пример сеточной модели тепловыделяющей сборки из 37 стержней – 80 млн ячеек

### **6.3. Задачи рендеринга**

В процессе тестирования виртуального десктопа vDesktop-2 для задач рендеринга в программных продуктах Adobe After Effects и Adobe 3DS Max были получены следующие результаты:

- Рендеринг 3D модели из Adobe 3DS Max на виртуальном десктопе VDI был выполнен за 1 ч 12 мин. Для просчета аналогичной модели на рабочей станции HP Z230 (Intel Core i7-4770, 32 ГБ оперативной памяти) было затрачено 43 мин.
- Рендеринг видеоролика из программного продукта Adobe After Effects на VDI был выполнен за 10 ч 32 мин. Для рендеринга аналогичного видеоролика на рабочей станции HP Z620 (2 процессора Intel Xeon E5-2640, 48 ГБ оперативной памяти) было затрачено 17 ч 08 мин.

Исходя из вышеизложенного, можно констатировать приемлемое время рендеринга видеороликов на виртуальных машинах VDI и, следовательно, применимость данного решения для этих задач.

## **7. Заключение**

Описанный в данной статье стенд обеспечивает двум пользователям одновременную работу с большими сеточными моделями. Размер моделей ограничивается размером оперативной памяти: для 192 ГБ оперативной памяти (машина vDesktop-1) потолок составляет около 200 млн контрольных объемов; для 100 ГБ оперативной памяти (машина vDesktop-2) потолок составляет около 100 млн контрольных объемов (данные оценки справедливы для Ansys Mechanical APDL).

Важно, что виртуализация позволяет легко перераспределить оперативную память между виртуальными рабочими местами. Это крайне полезно, когда на разных этапах расчетов используются сеточные модели разных размеров с разными требованиями к оперативной памяти.

В завершение нужно сказать также несколько слов о масштабировании. Для данного подхода главными характеристиками оборудования являются объем оперативной памяти и количество графических процессоров. Оба эти параметра можно увеличить: в настоящий



момент на рынке доступны сервера с объемом оперативной памяти до 1.5 Тб, а также решения, допускающие установку нескольких карт NVIDIA GRID. Оборудование такого уровня позволило бы увеличить объем оперативной памяти виртуальному рабочему месту и/или количество виртуальных рабочих мест, которые можно разместить на одном физическом сервере. Кроме того, решение допускает горизонтальное масштабирование за счет увеличения числа физических серверов.

Авторы благодарят Hewlett-Packard Russia за оборудование, предоставленное для тестирования, компании ANSYS, Inc и КАДФЕМ Си-Ай-Эс за предоставление временных лицензий на программное обеспечение ANSYS, а также сотрудников АО НИКИЭТ Огнерубова Д. А., Мариничева Д. В. и Куликова А. А. за помощь в тестировании решения.

## Литература

1. Сайт проекта Lustre  
URL: [http://wiki.lustre.org/index.php/Main\\_Page](http://wiki.lustre.org/index.php/Main_Page) (дата обращения 13.11.2014).
2. Сайт проекта VirtualGL, документация  
URL: <http://svn.code.sf.net/p/virtualgl/code/trunk/doc/index.html#hd003> (дата обращения 31.10.2014)
3. Сайт компании VMware, страница о Horizon  
URL: <http://www.vmware.com/ru/products/horizon-view/> (дата обращения 28.05.2014)
4. Сайт компании Citrix, страница о XenDesktop  
URL: <http://www.citrix.ru/products/xendesktop/overview.html> (дата обращения 28.05.2014)
5. Сайт компании Microsoft, страница о VDI  
URL: <http://www.microsoft.com/ru-ru/windows/enterprise/products-and-technologies/virtualization/vdi.aspx> (дата обращения 28.05.2014)
6. Сайт компании RedHat, страница о виртуализации  
URL: <http://www.redhat.com/products/cloud-computing/virtualization/> (дата обращения 28.05.2014)
7. Сайт компании Microsoft, информация о максимальном размере ОЗУ:  
URL: <http://msdn.microsoft.com/en-us/library/windows/desktop/aa366778%28v=vs.85%29.aspx> (дата обращения 28.05.2014)
8. Сайт корпорации NVIDIA, страница о технологии GRID  
<http://www.nvidia.com/object/nvidia-grid.html> (дата обращения 13.02.2015)