



Суперкомпьютер по требованию  
Решения виртуализации и оркестрации для HPC-облаков

# Облачные вычисления для наукоемких предприятий

## Задачи моделирования естественных процессов



Геофизика, геологоразведка



Биоинформатика, геномика



Фармацевтика



Аэро- и гидродинамика



Microsoft  
Azure

**НЕ ПОДХОДЯТ**



amazon  
web services

SOFTLAYER®  
an IBM Company

Необходима инфраструктура с

- с сетью малой латентностью Infiniband **и/или**
- быстрым разделяемым файловым хранилищем

# Проблема

## Традиционная схема

- Построение крупными компаниями собственных вычислительных кластеров
- Аренда в крупных научных суперкомпьютерных центрах

## Недостатки

- Высокая стоимость оборудования и обслуживания
- Трудоёмкость настройки и запуска оборудования
- Разовый или кратковременный доступ нерентабелен из-за длительности предварительных согласований
- Проблема остановки и возобновления задачи
- Проблема переноса настроенного окружения при смене и апгрейде суперкомпьютеров
- Сложность обеспечения конфигураций софта, удобных всем пользователям

## Облачные вычисления

- Общего назначения:  
Rescale, Nimbix, Bull, Penguin on demand, т.д.
- Ориентированные на индустрии:  
Illumina BaseSpace, DNANexus, EasyGenomics, т.д.

## Недостатки

- Не все облачные вендоры поддерживают сеть малой латентности
- Все сервисы имеют хостинг за рубежом, что накладывает ограничения по передаче данных

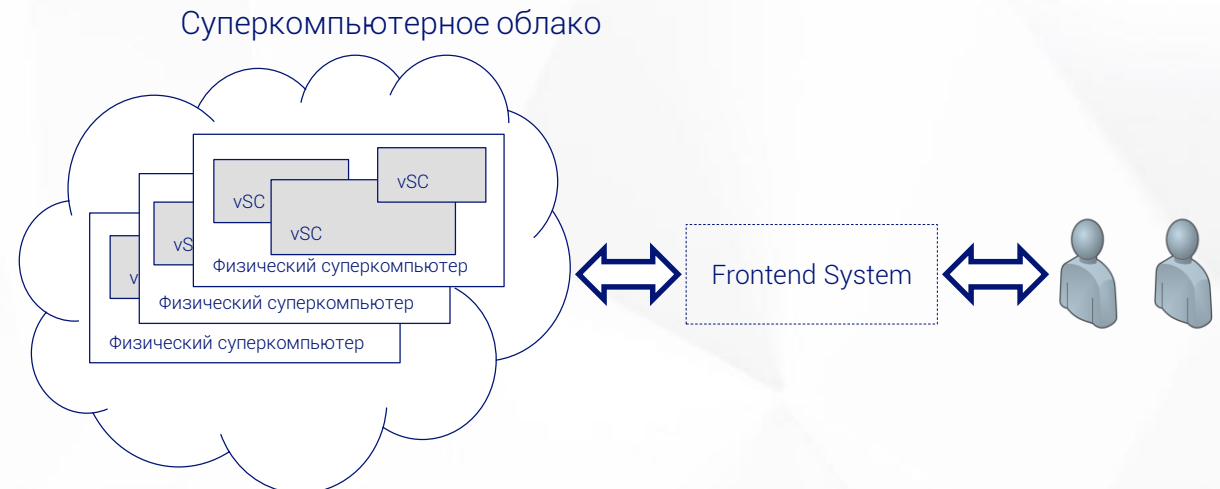
# Суть подхода

## Реализуемые технические требования

- Каждая группа пользователей видит только свой vSC с полным набором необходимого ПО и возможностью гибкой настройки: чистая среда
- Масштабируемость среды, гибкое квотирование
- Манипулирование образами, как отдельными сущностями
- Предустановка софта
- Безопасность: L2 изоляция сети, виртуализация KVM, архитектура на базе OpenStack
- Надёжность хранения данных: репликация данных на уровне Ceph
- vSC сохраняет среду исполнения программ и пользовательский интерфейс реальных современных кластеров, что ускорит внедрение и позволит использовать привычное ПО
- vSC привносит минимальные потери производительности по сравнению с обычным кластером

## Основные технологии

- Система полной виртуализации **KVM**
- Система управления ресурсами **OpenStack**
- Распределенное хранилище **Ceph**
- Кластерная файловая система **GFS2**



# Преимущества подхода

Разделяемое файловое хранилище Ceph:

- распределенная файловая система
- почти линейное масштабирование
- высокая производительность

Архитектура на основе облачного решения OpenStack

Безопасность вычислений:

- L2 изоляция сети
- виртуализация KVM

Надежность хранения данных -  
репликация данных на уровне Ceph

Минимальные накладные расходы от виртуализации за счет тонких настроек (NUMA tuning, vCPU pinning, huge pages)

Возможность пользователю передавать образы вычислительного кластера друг другу

Виртуализация позволяет реализовать технологию snapshoting

Возможность миграции vSC между физическими суперкомпьютерами

Легкая масштабируемость вычислительной среды

# Технологические вызовы

Построение кластера на виртуальных машинах

Минимизация накладных расходов от виртуализации

Виртуализации Infiniband в KVM

Технология разделяемого диска поверх распределённого хранилища

Оркестрация vSC и оборудования

vSC построены на CentOS 7.1, драйвер IB OFED 3.3, систему управления заданиями SGE 8.1.3, настроенный OpenMPI

Настроили NUMA расширение, vCPU pinning

Используем технологию SR-IOV

Настроили Ceph с двойным резервированием и сконфигурировали кластерную FS GFS2 внутри vSC

Настроили Openstack Mitaka и написали injection софт по настройке HPC окружения внутри vSC



# Результаты измерений

Построение кластера на виртуальных машинах

Минимизация накладных расходов от виртуализации

Виртуализации Infiniband в KVM

Технология разделяемого диска поверх распределённого хранилища

Оркестрация vSC и оборудования

Интеграция сложного ПО в разы упрощается:  
устанавливать нужно только 1 раз

Деградация по CPU не более 2%

Латентность IB VF – 1.09 мкс против 0.85 мкс без виртуализации

1.5-2 ГБ/сек – чтения, 350-450 МБ/сек - запись,  
10,000-12,000 IOPS, стоимость СХД: \$0.15 / 1ГБ

vSC deployment time – 10-15 минут (256 ядер)

# Развитие подхода

R&D: перевод узлов с KVM на LXC и Docker

Завершение R&D  
suspending/checkpointing/snapshotting

Управление образами

Различные подходы к предоставлению услуг

Предустановка софта

Минимизация накладных расходов

Возможность приостановки исполнения задачи / создания точки восстановления без остановки /  
Получение полного snapshot задачи и его миграция, выполнение из любой точки восстановления или snapshot'a

История образов, передача. Новые модели коммерческих услуг на этой базе

On-demand, on-spot, alltime

Возможность выбора предустановленных пакетов и аренда лицензий



# Наши решения

## Public HPC Cloud

облачный сервис, предоставляющий доступ к высокопроизводительным вычислительным ресурсам по модели PaaS и SaaS

- Автоматизированное предоставление через веб-интерфейс
- Оплата только за время счета
- Предустановленное счетное ПО
- Автоматический бухгалтерские отчеты для юр. лиц

## Private HPC Cloud

мы используем мощности клиента и наши знания, чтобы построить вычислительное облако внутри компании

- В разы выше SLA для внутренних заказчиков
- Унификация всей инфраструктуры в компании
- Утилизация мощностей больше в 2 раза
- Простое использование счётного ПО

# Всего в мире 30 млн ядер в HPC

## 30 млн ядер on demand с HPC HUB



### Win-Win стратегия

возможность заработать в 2 раза больше всем участникам рынка



Аренда только нужного объема мощностей на нужный срок (pay per use)



Гарантированное подключение за 15 минут (on demand)



Минимум ошибок за счет виртуализации и автоматизации



Клиентское ПО уже установлено, экономия времени на интеграцию

# Почему выгодно использовать облачные HPC-сервисы

## Покупка кластера на 1000 ядер

- Затраты на приобретение оборудования в лизинг в течение 5 лет – 50 000 000 руб
- Оплата размещения оборудования в ЦОД в течение 5 лет – 18 000 000 руб (300 000 руб/мес)
- Зарботная плата ИТ персонала 24 000 000 руб (400 000 руб/мес, включая налоги и отчисления)
- Затраты на ремонт оборудования, исходя из AFR 5% - 6 000 000 руб
- Реальная загрузка мощностей 30-50%, т.е. 18-30 мес счета за 5 лет
- ИТОГО: TCO за 5 лет – 98 000 000 руб

## vSC от HPC HUB

- Плата за аренду – 2 880 000 руб/мес (4 руб / ядро / час)
- Оплата только за счет – 18-30 мес
- ИТОГО: TCO за 5 лет – 51-86 млн. руб.

Дополнительные плюсы:

- нет штата сотрудников
- мгновенное выделение выч. мощностей
- Дополнительная экономия на аренде коммерческого ПО в облаке



Поставщик программно-аппаратных решений  
в сфере высокопроизводительных вычислений  
для решения ресурсоемких задач науки и бизнеса  
по модели вычислительных облаков

+7 495 774 1937  
[info@hpchub.net](mailto:info@hpchub.net)  
[www.hpchub.net](http://www.hpchub.net)