



Межведомственный суперкомпьютерный центр РАН

**Шабанов Б.М.
Баранов А.В.
Киселёв Е.А.**

**Овсянников А.П.
Аладышев О.С.
Жуков Я.О.**

**Методы управления параллельными
заданиями суперкомпьютера,
требующими развёртывания
отдельных программных платформ и
виртуализации сетей**



Облачные сервисы для HPC

Частное облако

Публичное/Корпоративное
облако



**Инфраструктура
как сервис**

выделение
аппаратных ресурсов
по требованию



Платформа как сервис

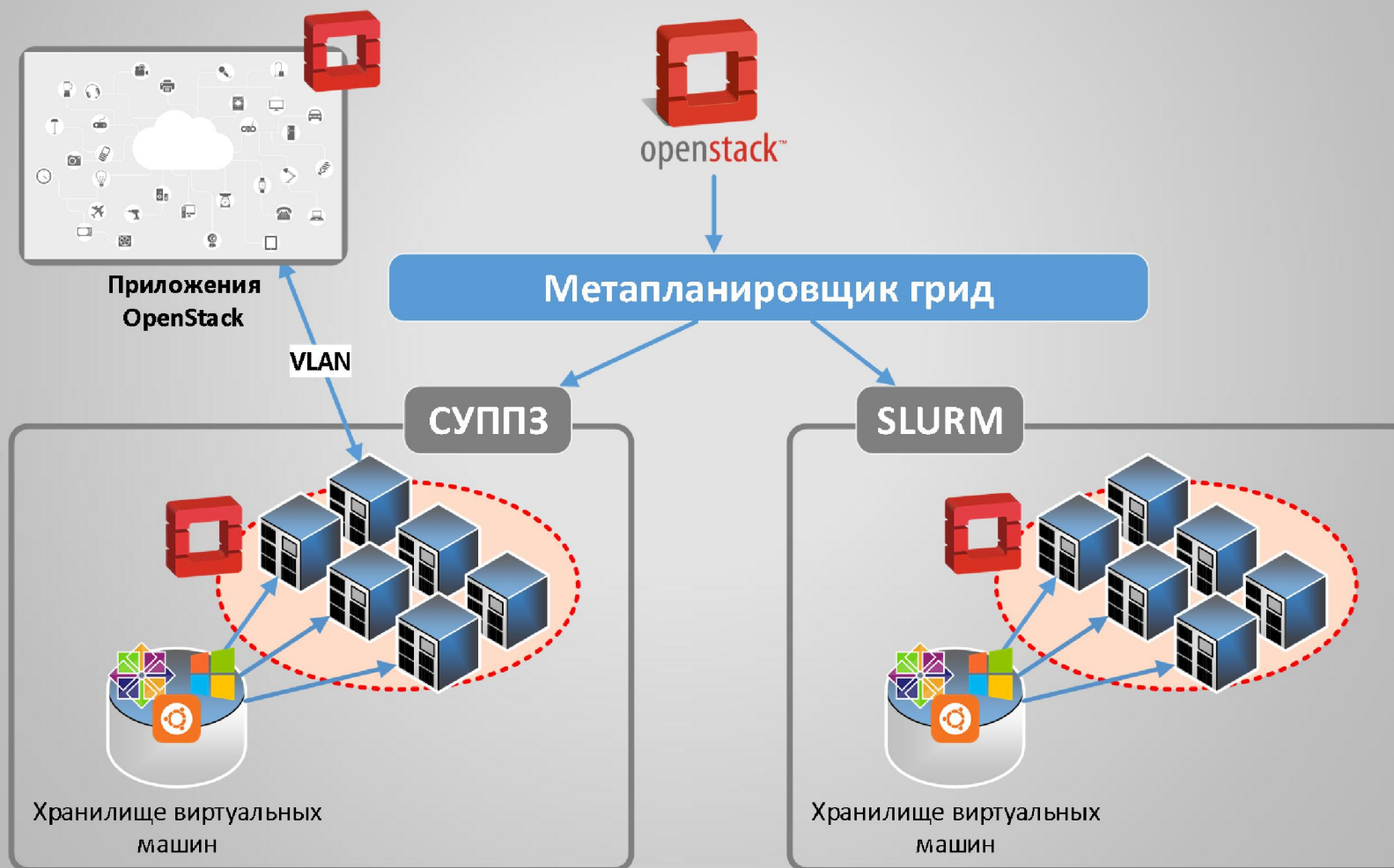
эластичное выделение
ресурсов заданной
конфигурации по требованию



ПО как сервис

предоставление услуг по
расчетам с помощью
прикладных программных
пакетов

Вариант организации облачных вычислений на базе МСЦ РАН





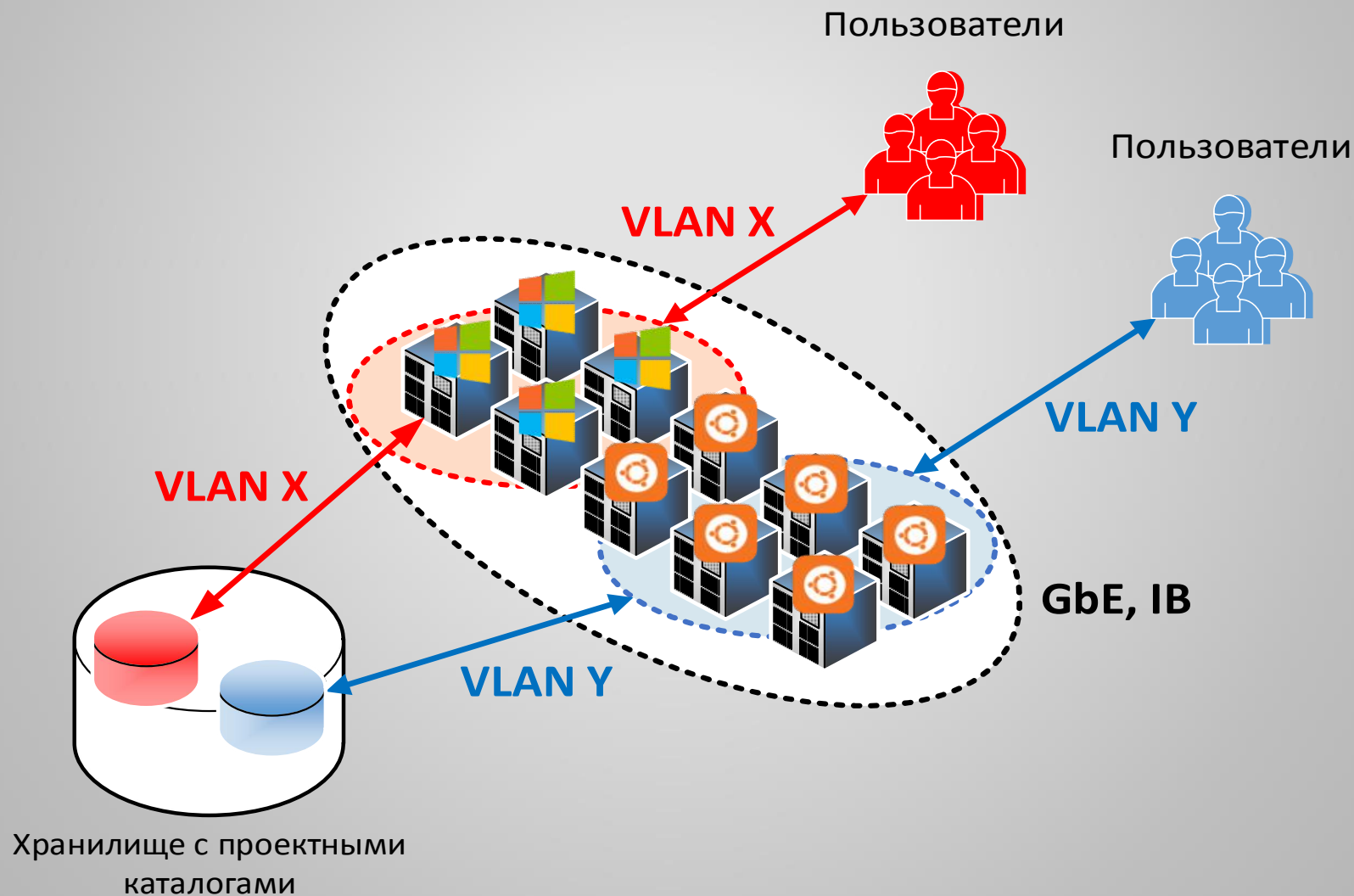
Макет облачной среды для высокопроизводительных приложений





Виртуальная коммуникационная среда

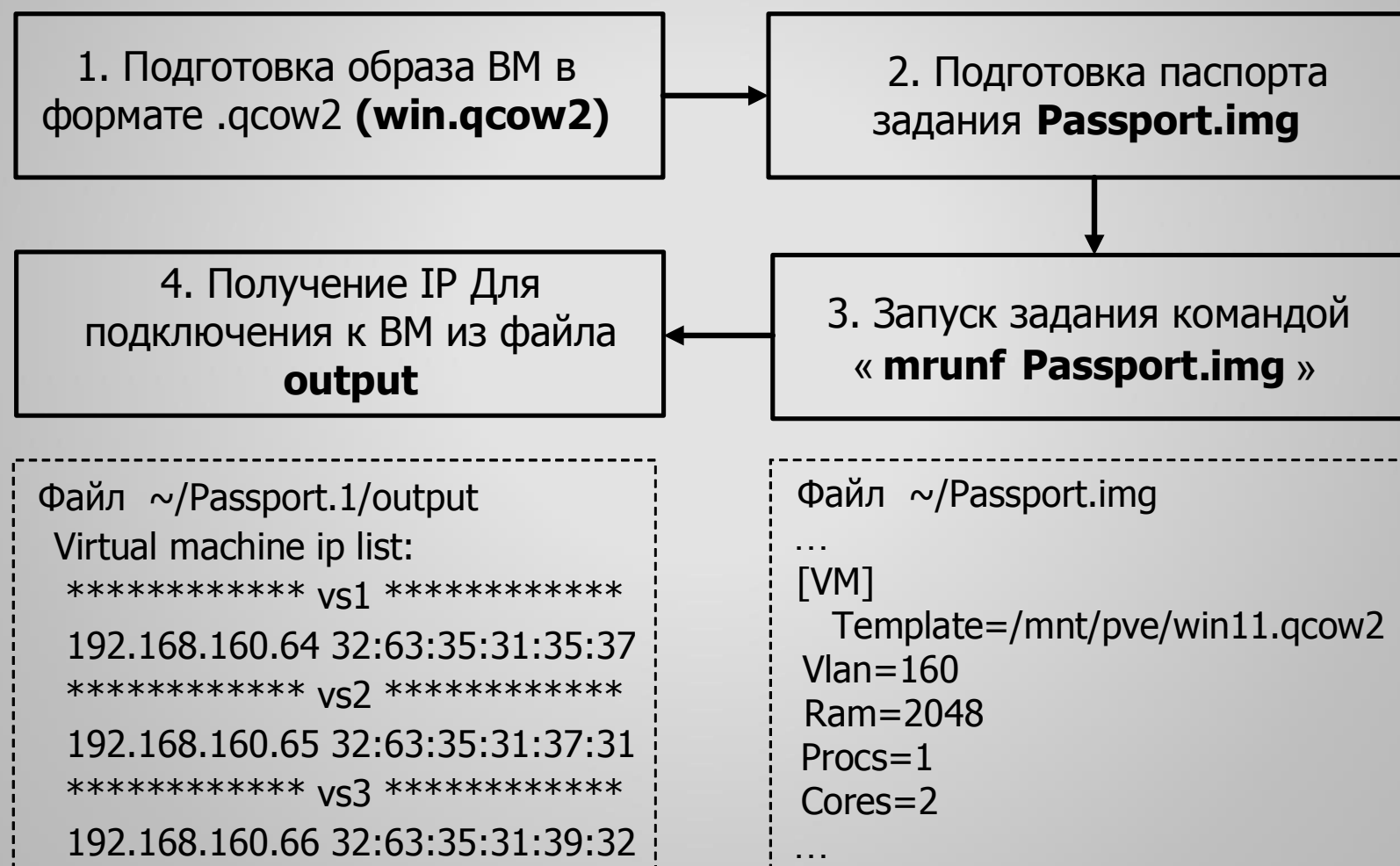
5



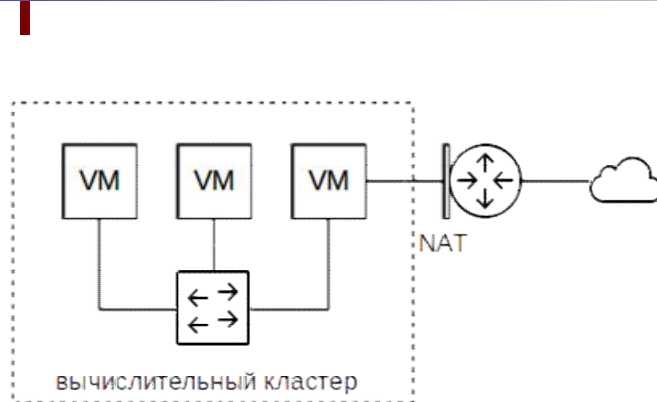


Процесс подготовки и запуска задания через СУППЗ и Proxmox

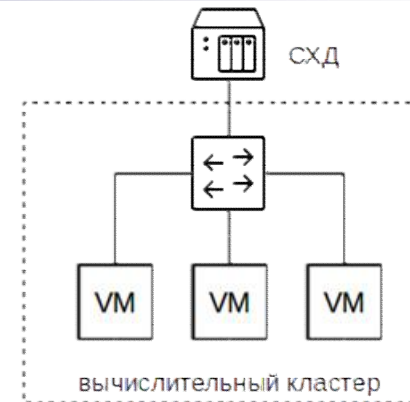
6



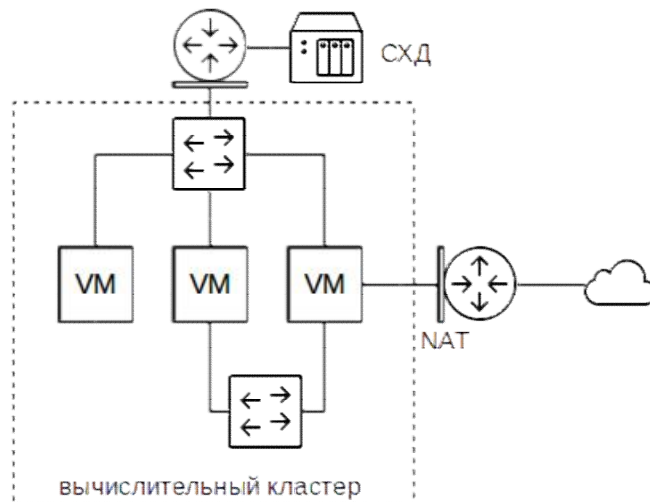
Возможные сценарии сетевого взаимодействия виртуальных машин



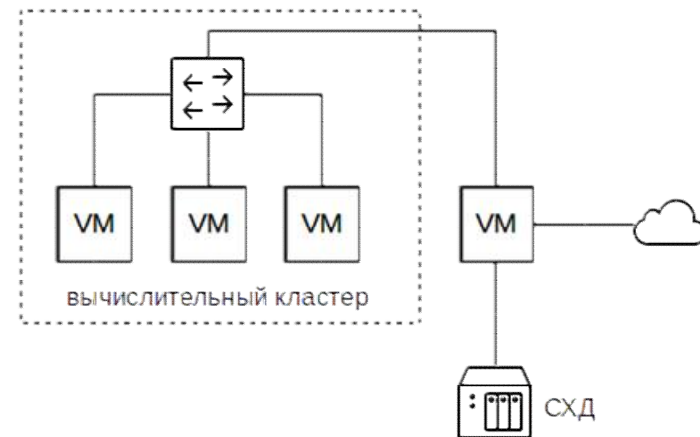
а) Динамически запускаемые виртуальные машины с внешним доступом к ним



б) Динамически запускаемые виртуальные машины с доступом к системе хранения



в) Динамически запускаемые виртуальные машины, взаимодействующие по нескольким сетям



г) Динамически запускаемые виртуальные машины, связанные со статической виртуальной машиной



Настройка виртуального сетевого окружения задания

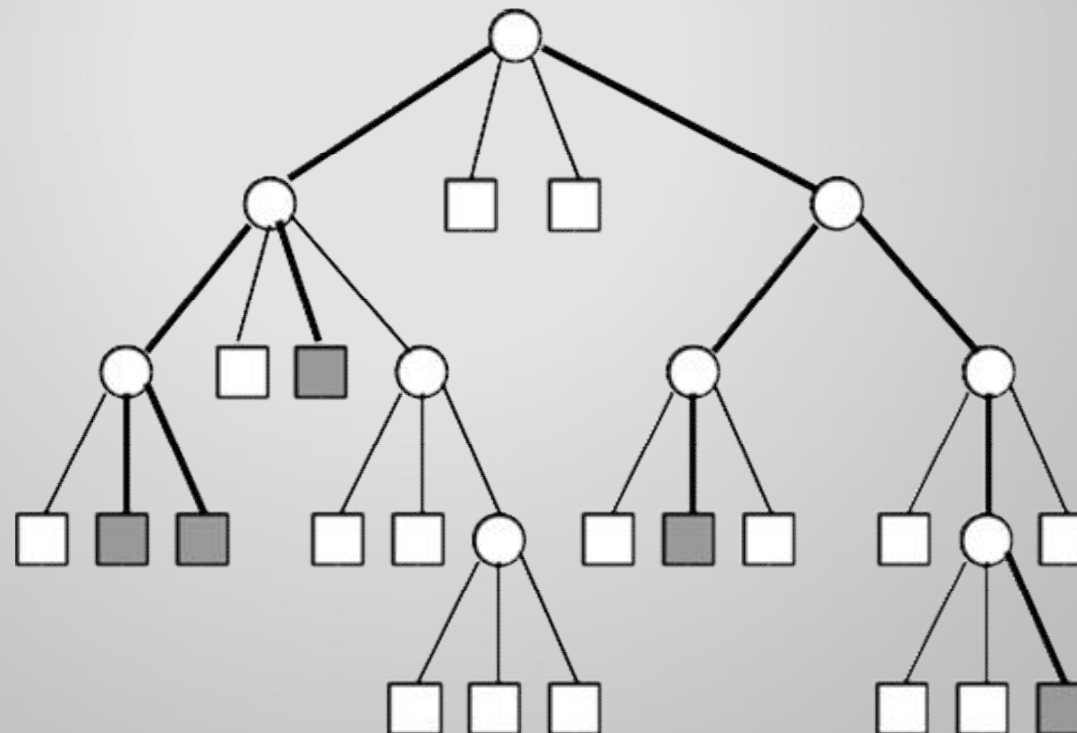
- задание информации о виртуальном сетевом окружении набора виртуальных машин в описании (паспорте) задания, запускаемого на кластере
- задание нумерации сетевых ресурсов виртуальных машин
- отображение виртуального сетевого окружения набора виртуальных машин задания на реальную физическую сетевую инфраструктуру вычислительного кластера
- настройка физической сетевой инфраструктуры перед запуском задания и обратные изменения по окончании задания
- передача информации о сетевой нумерации в виртуальные машины при развёртывании на кластере



Автоматическая настройка сетевой инфраструктуры

9

- Для сетей Ethernet вычислительного кластера при запуске задания:
 - строится подграф, объединяющий выделенные заданию вычислительные узлы





Автоматическая настройка сетевой инфраструктуры

10

- Для сетей Ethernet вычислительного кластера при запуске задания:
 - строится подграф, объединяющий выделенные заданию вычислительные узлы
 - настраиваются VLAN на коммутаторах в вершинах подграфа и портах, соответствующих соединяющим их ребрам (каналам)
 - настраиваются VLAN на портах коммутаторов, соответствующих ребрам подграфа (каналам) к используемым заданием вычислительным узлам
 - на узлах настраивается проброс VLAN в виртуальную машину
 - по завершении задания отменяется проброс VLAN в виртуальную машину на вычислительных узлах и удаляются VLAN на коммутаторах



Автоматическая настройка сетевой инфраструктуры

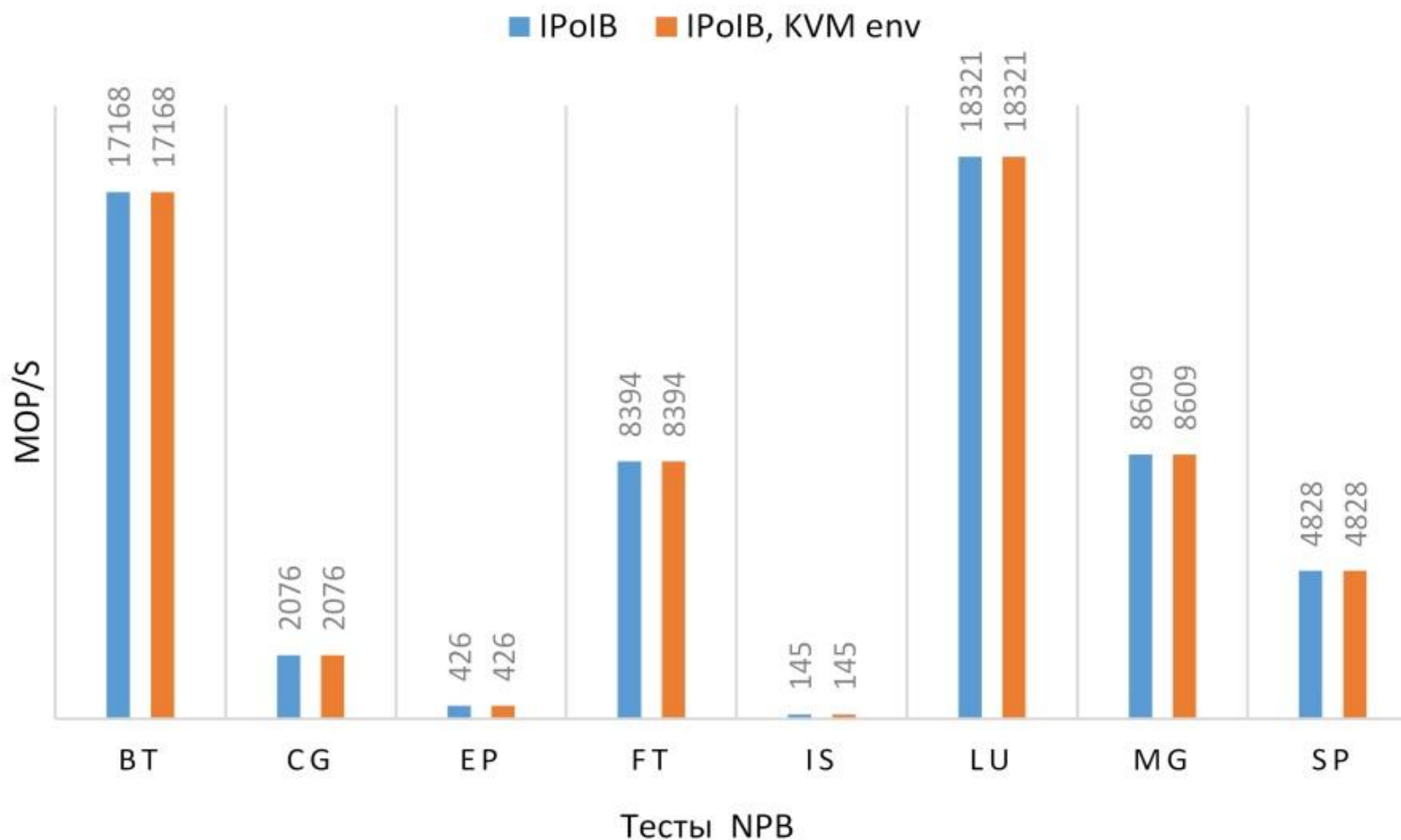
11

- Для сетей InfiniBand вычислительного кластера при запуске задания:
 - для выделенных заданию вычислительных узлов добавляется partition в файл */etc/opensm/partitions.conf* (PKey назначается динамически) и перезапускается subnet manager
 - если используется мост на основе коммутатора Mellanox между Ethernet и Infiniband, то дополнительно настраивается интерфейс **proxy-arp** между нужными PKey и VLAN
 - по завершении задания удаляются partition и интерфейс **proxy-arp**



Оценка влияния компонентов KVM и Proxmox VE на производительность

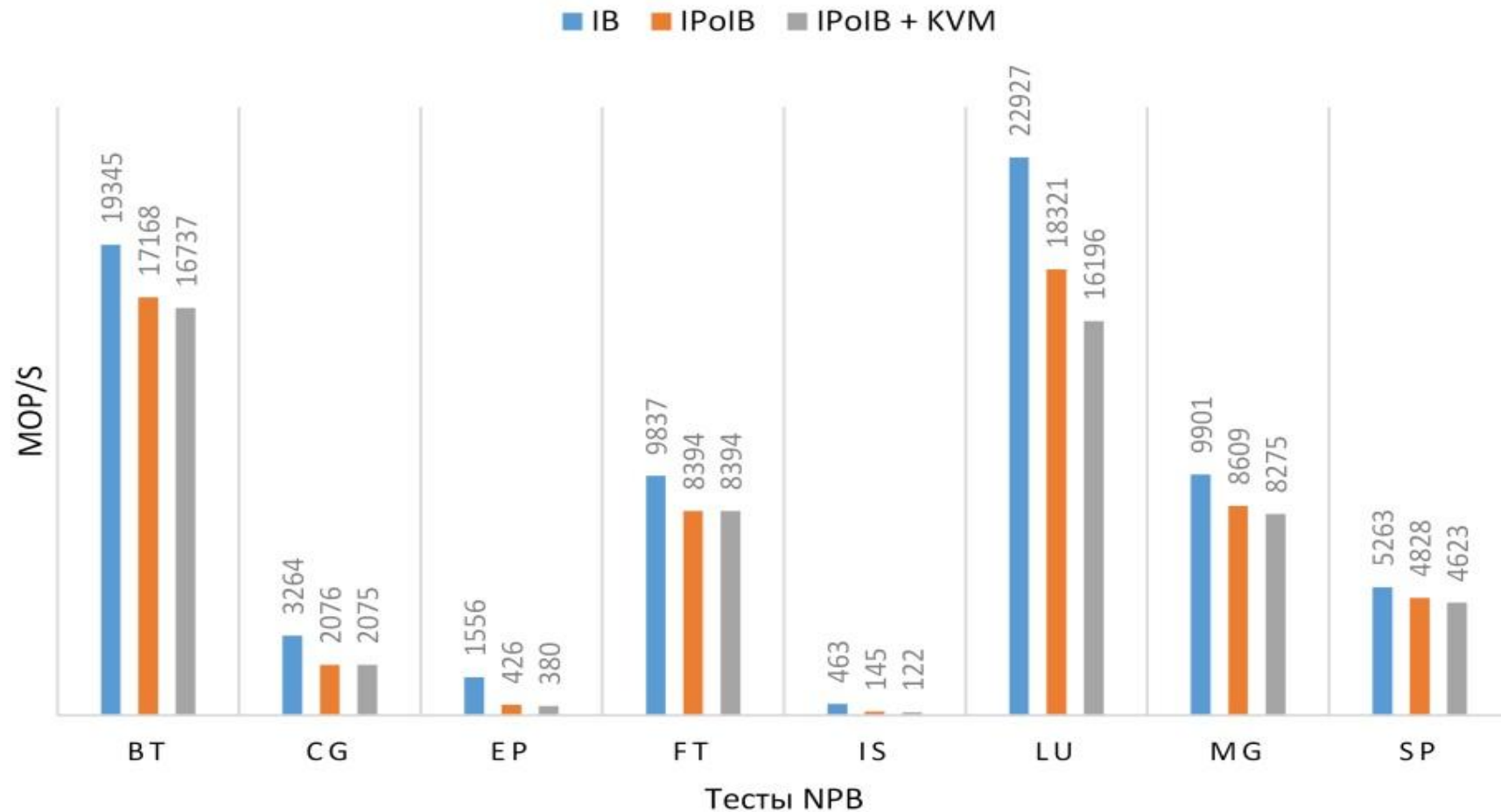
12





Оценка влияния виртуальной среды KVM на выполнение тестов NPВ

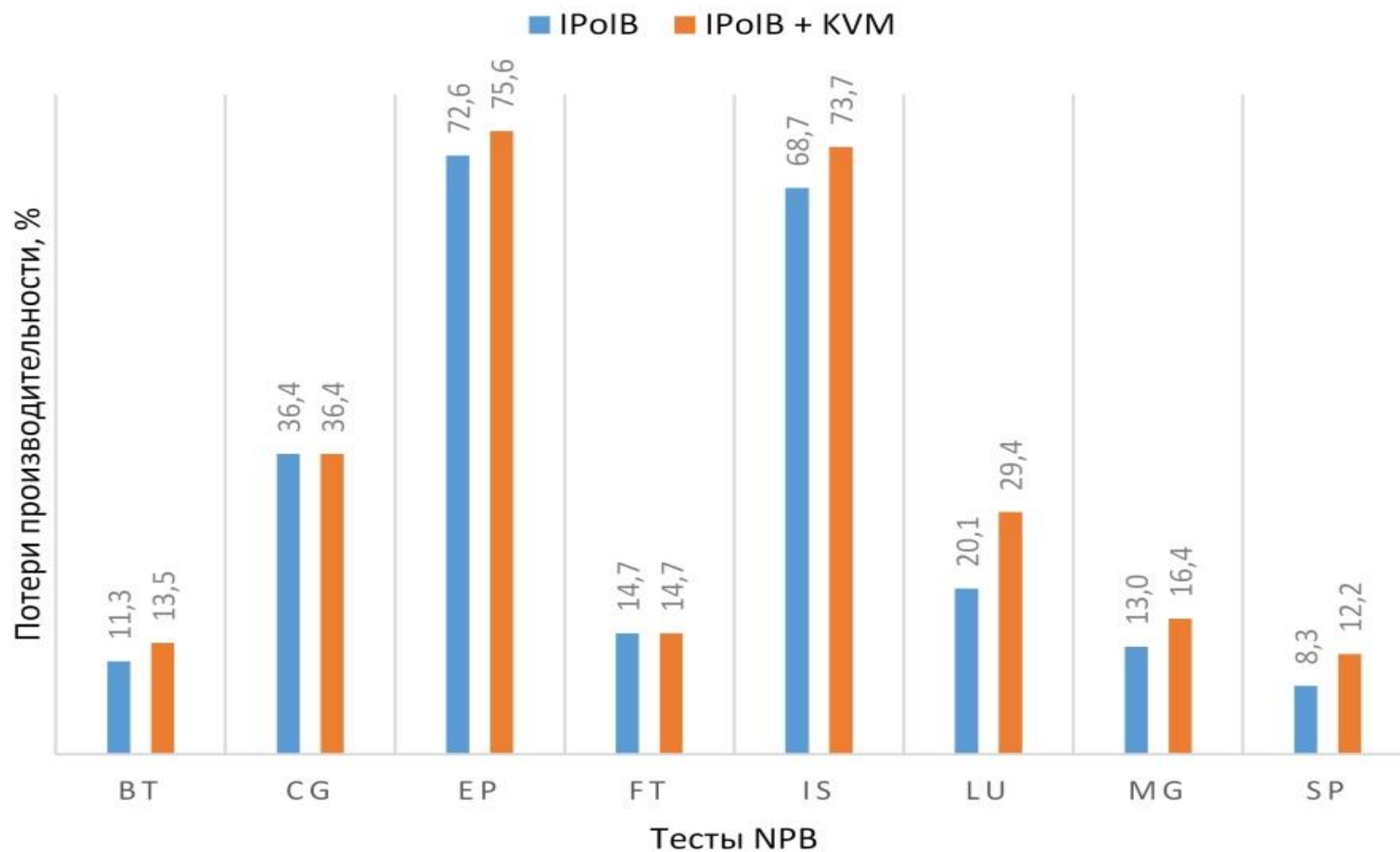
13





Потери производительности при использовании InfiniBand IPoIB и KVM

14





Спасибо за внимание!

Вопросы?