

## ВОПРОСЫ ПОСТРОЕНИЯ ЭФФЕКТИВНЫХ СИСТЕМ ТЕХНИЧЕСКОГО КОМПЬЮТИНГА:

## ОТКРЫТЫЙ МОДУЛЬНЫЙ ПОДХОД ИЛИ ЧАСТНОФИРМЕННЫЕ РЕШЕНИЯ

НСКФ 2013 Переславль Залесский Александр Буравлев aburavlev@aq.ru





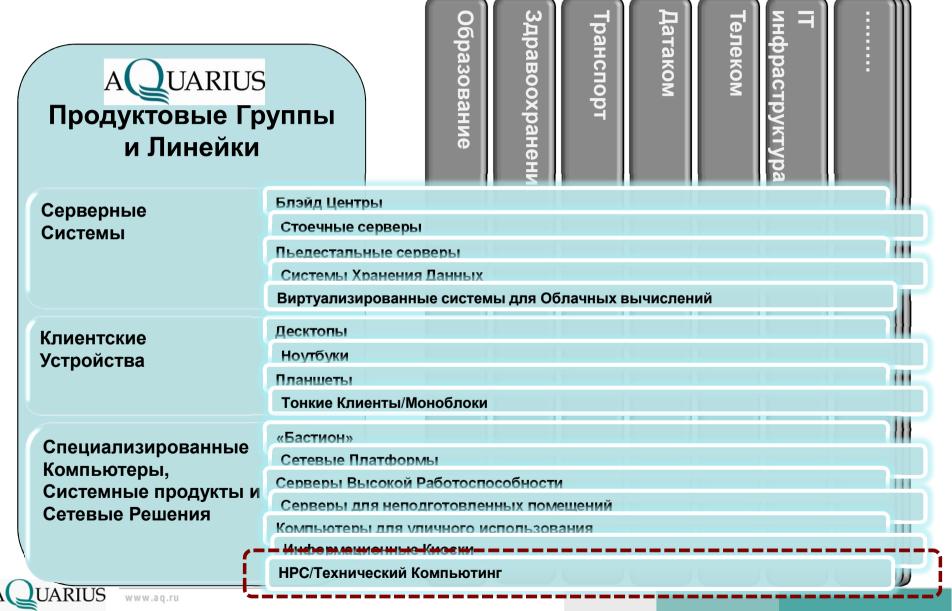
#### Миссия Компании

Компания Аквариус создает наиболее <u>современные</u> компьютерные решения <u>оптимизированные</u> для нужд рынков и конкретных заказчиков





#### Решения Для Вертикальных Рынков





## Терминология

A supercomputer is a <u>computer</u> at the frontline of contemporary processing capacity – particularly speed of calculation. (Redirected from <u>High-performance computing</u>)

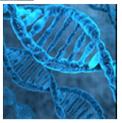
Technical computing is the application of the mathematical and computational principles of scientific computing to solve practical problems of industrial interest.

ВЫСОКОПРОИЗВОДИТЕЛЬНЫЙ КОМПЬЮТИНГ (НРС)	ТЕХНИЧЕСКИЙ КОМПЬЮТИНГ (НРТС)	
Для достижения рекорда (ов)	Для решения конкретных прикладных задач	
Широкий набор задач	Ограниченный набор задач	
Научные и прикладные задачи	Прикладные задачи	
Широкий круг пользователей	Один или несколько пользователей	
Средний-большой размер и мощность	Малый-средний размер и мощность	
Инвестиционный вклад	Средство производства	

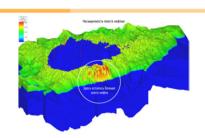




## Эффективность: Категории











ВЫСОКОПРОИЗВОДИТЕЛЬНЫЙ КОМПЬЮТИНГ (НРС)	ТЕХНИЧЕСКИЙ КОМПЬЮТИНГ (НРТС)
ВЫЧИСЛИТЕЛЬНАЯ ЭФФЕКТИВНОСТЬ: - Flops, - Flops/m3	ВЫЧИСЛИТЕЛЬНАЯ ЭФФЕКТИВНОСТЬ: - Скорость обработки приложения
ЭНЕРГОЭФФЕКТИВНОСТЬ: - Flops/Вт, - PUE (редко)	ЭНЕРГОЭФФЕКТИВНОСТЬ: - PUE
ЭФФЕКТИВНОСТЬ САРЕХ: - Экспертная оценка, - Flops/\$	ЭФФЕКТИВНОСТЬ САРЕХ: - Скорость обработки приложения/\$
ЭФФЕКТИВНОСТЬ OPEX: - Энергопотребление + Поддержка	ЭФФЕКТИВНОСТЬ ОРЕХ: - Энергопотребление + Поддержка



## Вычислительная Эффективность







## Энергоэффективность

5 лет срок службы

100% загрузка

СОВОКУПНОЕ ЭНЕРГОПОТРЕБЛЕНИЕ
1 Вт \*100% загрузка\* 24ч \* 365 дней \*5 лет \* = 44 КВт ч

\$0.1/КВт ч - тариф для промышленности

10% - стоимость капитала в России

 Цена приобретения электроэнергии промышленными предприятиями

 2400
 2400

 2200
 2000

 1800
 1600

 1400
 1200

 1200
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 1200

 1000
 <td

Источник: Росстат

СОВОКУПНАЯ СТОИМОСТЬ 1 ВТ ~ 2.7.....4.4 \$/Вт

БОРЬБА ЗА ЭНЕРГОЭФФЕКИВНОСТЬ ИМЕЕТ СМЫСЛ Если инвестиции < 3...5 \$ за 1 Вт сэкономленный (@ \$0.1/КВтч)





## Повышение Энергоэффективности (1)

#### ІТ ОБОРУДОВАНИЕ

Выбор Эффективных ИП

Тип∖Нагрузка	10%	20%	50%	100%
80 PLUS Silver	-	85%	89%	85%
80 PLUS Gold	-	88%	92%	88%
80 PLUS Platinum	PLUS'	90%	94%	91%
80 PLUS Titanium	90%	94%	96%	91%

40 узлов 0.3КВт, 5 лет эксплуатации ~ \$5256 (@\$0.1/КВт)

Эффективный режим работы ИП => на пике эффективности

Питание узлов AC и DC => DC backup, DC 400 B, ....

Резервирование 1+1 => N+1 (Обобщение узлов)

Питание

## Повышение Энергоэффективности (2)

#### ІТ ОБОРУДОВАНИЕ

#### Обобщение вентиляторов:

Конструктив	Вентиляторы на узел	Стойка 40 Узлов
1U Сервер	8*12 Вт = 72 Вт	72 Bt * 40 = 2880 Bt
Сервер в составе обобщенного шасси	16.8 Вт	672 Вт

40 узлов, 5 лет эксплуатации - \$9320 (@\$0.1/КВт)

Горячий жидкий теплоноситель : ~ 5 Вт/узел Требует соответствующего дизайна датацентра !



## Повышение Энергоэффективности (3)

#### Датацентр

HPTC, в основном, «въезжает» в существующий датацентр

#### Повышение Т выходящего воздуха:

- повышение Т эксплуатации
- отсутствие перемешивания горячего с холодным

#### Специализированные системы охлаждения:

- Встроенный воздушно-жидкостной теплообменник
  - Внутристоечный
  - Охлаждаемые двери
- Встроенные кондиционеры
- Контейнерные системы и др.







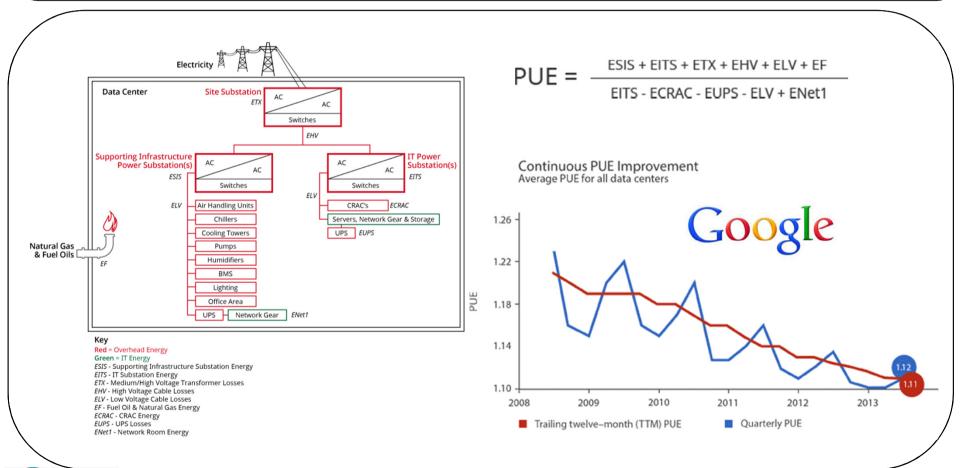






## PUE = Total Datacenter Power IT Power

#### **Нет стандартизированного определения PUE**







#### Промежуточные выводы

Аппаратная Конфигурация HTPC подбирается для оптимального решения задачи и определяется алгоритмом расчета

Оптимизация OPEX - Технологии охлаждение воздухом Поставщиков Интернет Сервисов позволяет получить отличные значения энергоэффективности

- Обобщение системы воздушного охлаждения
- Оптимизация питания

Оптимизация САРЕХ - Расчет инвестиций {Энергоэкономия, Bт}х{%Загрузки}х{Срок Эксплуатации,ч}х{Стоимость Электроэнергии \$/КВт ч} / {стоимость инвестиций}





# Стоечные решения 2013 год «Aquarius HPC Rack»





#### Стоечное Решение "Aquarius HPTC Rack x86-Core"

## Решение с узлами на Intel Xeon E5-2600v2 серии

#### Производительность

- Теоретическая ~18..20 TFlops, двойная точность
- 41 TB RAM
- Количество узлов: 80, 42U

Узел: 2xIntel Xeon E5-2670v2 (12 ядер, 2.4 ГГц, 30 MBL3)), 512 ГБ RAM

- Интерконект обмена данных одинарный неблокирующий, FDR (56 Gb/s) InfiniBand, Mellanox
- Базовый Интерконнект GE
- Энергопотребление:
  - − Общее ~36 КВт (~0.5 Tflops/КВт)
  - Опция закрытой стойки IP55, 58dB
    - В-Ж теплообменник 1.5КВт max, 1 КВт typ.
    - Расход жидкости 4м³/ч max@18°C
- OS Linux (Red Hat Cluster Edition)
- Программная поддержка





#### Стоечное Решение "Aquarius HPC Rack x86- RAM"

## HPC Решение с узлами на Intel Xeon E5-4600 серии

#### Производительность

- Теоретическая 17 (26 в 2014г) TFlops,
- 50 T5 RAM
- Количество узлов: 50, 40U

Узел: 4xIntel Xeon E5-4650 (8 ядер, 2.7 ГГц, 20 MBL3), 1 ТБ RAM

- Интерконект обмена данных неблокирующий,
   Одинарный FDR (56 Gb/s) InfiniBand, Mellanox
- Базовый Интерконнект GE
- Энергопотребление:
  - Общее 44 КВт max. (0.4 (0.6) TFlops/КВт)
  - Опция закрытой стойки IP55, 58dB
    - В-Ж теплообменник 1.5КВт max, 1 КВт typ.
    - Расход жидкости 4м³/ч max@18°C
- OS Linux (Red Hat Cluster Edition)
- Программная поддержка



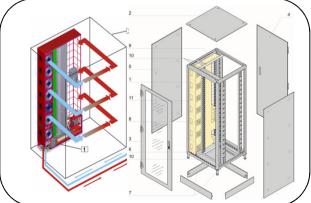














### Стоечное Решение "Aquarius HPTC Rack - Phi"

HPC Решение с узлами Intel Xeon E5v2 и Сопроцессингом Intel Xeon Phi

#### Производительность

- Теоретическая 72 TFlops, двойная точность
- Количество узлов: 18, 36U,

2xIntel Xeon E5v2 и 4xIntel Xeon Phi™ на узел

- Интерконнект обмена данных одинарный неблокирующий, FDR (56 Gb/s) InfiniBand,
- Базовый Интерконнект GE
- Энергопотребление:
  - Общее 30 КВт max. (2.4 TFlops/КВт)
  - Опция закрытой стойки IP55, 58dB
    - В-Ж теплообменник 1.5КВт max, 1 КВт typ.
    - Расход жидкости 4м³/ч max@18°C
- OS Linux (Red Hat Cluster Edition)
- Программная поддержка
  - C++, C, Fortran, Open CL







### Стоечное Решение "Aquarius HPTC Rack GPU"

## HPC Решение с узлами Intel Xeon E5v2 и Сопроцессингом NVIDIA TESLA K40

#### Производительность

- Теоретическая 103 TFlops, двойная точность
- Количество узлов: 18, 36U,

2xIntel Xeon E5v2 и 4 NVIDIA Tesla K40 GPU на узел

- Интерконнект обмена данных одинарный неблокирующий, FDR (56 Gb/s) InfiniBand,
- Базовый Интерконнект GE
- Энергопотребление:
  - − Общее ~28 КВт max. (3.7 TFlops/КВт)
  - Опция закрытой стойки IP55, 58dB
    - В-Ж теплообменник 1.5КВт max, 1 КВт typ.
    - Расход жидкости 4м³/ч max@18°C
- OS Linux (Red Hat Cluster Edition)
- Программная поддержка
  - C++, C, Fortran, Open CL







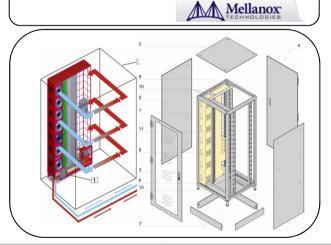
#### Стоечное Решение "Aquarius HPTC Rack Virtex"

HPC Решение с узлами на Intel Xeon E5 и сопроцессингом Xilinx Virtex-7 FPGA

#### Производительность

- Узел: 2U x86 сервер с подключенными по PCI-Е сопроцессорами FPGA (от 4 до 8).
- Количество сопроцессоров на узел: от 4 до 96 Xilinx Virtex 7V2000T (2 млн. лог.элементов)
- Интерконект обмена данных между узлами выбор заказчика: InfiniBand до FDR (56 Gb/s), Ethernet 1/10 Gb/s.
- Базовый Интерконнект GE
- Энергопотребление:
  - Общее от 20 до 60 КВт тах.
  - Опция закрытой стойки IP55, 58dB
    - В-Ж теплообменник 1.5КВт max, 1 КВт typ.
    - Расход жидкости 4м³/ч max@18°C
- OS Linux или Microsoft Windows Server









## **Aquarius HPTC Rack**

Решения Aquarius HPC Rack поддерживают самую широкую номенклатуру вычислительных ядер узлов:

- 1) x86,
- (intel) inside\*
- 2) Intel Phi
- 3) GPU и
- 4) FPGA







#### Решения Aquarius HPC включают аппаратные компоненты:

- 1) Интерконект с низкой задержкой
- 2) Высокоскоростные Подсистемы хранения
- 3) Энергоэффективное Охлаждение и Питание
- 4) Удаленный мониторинг и управление





#### Заключение

Обращайтесь в Аквариус для оптимизации аппаратной архитектуры компьютерного комплекса для НРТС.

Решения Aquarius HPTC базируются на коммерчески доступных компонентах и не содержат рисков частнофирменных решений

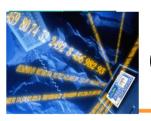
Решения Aquarius HPTC базируются на современных технологиях построения эффективных по САРЕХ и ОРЕХ компьютерных систем, используемых ведущими игроками мирового сегмента WEB 2.0





## Приложение





## Опция Охлаждения: Закрытая Стойка

- Охлаждение серверного и коммутационного оборудования внутри стойки обеспечивается пластинчатым вертикальным теплообменником вода-воздух и вентиляторами с переменной скоростью вращения, обеспечивающими установление вертикального воздушного ротора внутри стойки перемещающего тепло от процессоров и контроллеров к пластинам теплообменника.
- Целевые мощности (конструкции) теплообменников: 20, 40 и 60 КВт. Производительность кластера в составе 1 стойки на предыдущих слайдах указана для конструкции с целевой мощностью 40 КВт и может быть масштабирована до 20 или до 60 КВт. Фактические мощности охлаждения для каждой конструкции теплообменника зависят от температуры и потока охлаждающей воды, позволяя гибко конфигурировать и эксплуатировать систему.
- Презентованная подсистема охлаждения обеспечивает отвод тепла от стойки первичный контур охлаждения. Утилизация тепла горячей воды отводимой от стойки требует организации вторичного контура охлаждения и может быть произведено путем:
  - использования закрытого пассивного контура охлаждения,
  - использование открытого контура с градирней
  - утилизация тепла для подогрева контура горячей воды в организации
  - использования иного оборудования и схемы охлаждения.

Разработка и оптимизация контуров вторичного охлаждения должна быть привязана к конкретным условиям эксплуатации и возможности утилизации тепла на месте.





#### Важные Замечания

- □ Aquarius предлагает различные архитектуры HPTC. Выбор архитектуры HPTC должен осуществляться на основании задач, решения которых будет возложен о на компьютерную систему.
- □ Сравнительные результаты синтетических тестов Linpack для различных архитектур могут не коррелировать с результатами производительности на реальных задачах.
- □ Основой гранулярности Aquarius HPC является стойка 42U, укомплектованная узлами Aquarius HPTC Rack .
- □ Объединяя стойки в единый кластер с заданным интерконнектом обмена данными между узлами можно создавать кластерные решения с высокой общ ей производительностью Aquarius HPC Rack Cluster.
- □ Для высоко производительных систем Aquarius поддерживает топологию интер коннектов обмена данных между узлами Fat Tree и 3D Tor.





#### Подсистема Хранения Aquarius HPTC

Решение для задач, требующих самой высокой скорости чтения и записи при не больших объемах данных

64GB....2ТВ/узел хранения,

Объем хранения системы Масштабируется наращиванием

количества Узлов

28.5 GB/sec, (Мировой Рекорд)

Скорость постоянна и не зависит от степени

загрузки

IOPS 11.7 million, (Мировой Рекорд) постоянна

и не зависит от степени загрузки

5 µѕ чтение, 5 µѕ запись (Мировой

Время доступа Рекорд) постоянна и не зависит от

степени загрузки

Интерконнект связи

InfiniBand:

До 12 портов SDR, DDR, QDR

Управляющая сеть 10/100/1000 Ethernet, 2xRJ-45

Габариты,

Энергопотребление

4U, 1100Вт Максимум



