

УДК 519.6

**ОБЛАЧНЫЕ ВЫЧИСЛЕНИЯ В ОБРАЗОВАНИИ, НАУКЕ И ГОССЕКТОРЕ**

В.П. Иванников

*Институт системного программирования РАН*

Россия, 109004, Москва, ул. А. Солженицына, 25

E-mail: [ivan@ispras.ru](mailto:ivan@ispras.ru)

Рассматривается новое воплощение старой идеи 60х годов на принципиально новой технологической основе как, в телекоммуникациях, так и в программном обеспечении. Гибкость предоставления ресурсов дает возможность резко снижать капитальные вложения в локальную инфраструктуру. Возможность создания web-ориентированных лабораторий (хабов) в конкретных предметных областях предоставляет принципиально новые возможности по передаче знаний: лекции, семинары, лабораторные работы и др. Ниже в докладе описаны возможности программы “Универсальный кластер” и проекта OpenCirrus.

**CLOUDY CALCULATIONS IN FORMATION, THE SCIENCE AND THE STATE SECTOR** / V.P. Ivannikov (Institute for System Programming, ISP RAS, st. A. Solzhenitsyn 25, Moscow, 109004, Russia). The new embodiment of old idea 60th years based on essentially new technological basis as, in telecommunications, and in the software is considered. Flexibility of granting of resources gives the chance to reduce sharply capital investments in a local infrastructure. Possibility of creation of the web-focused laboratories (hubs) in concrete subject domains gives essentially new possibilities on transfer of knowledge: lectures, seminars, laboratory works, etc. Further in the report are described program possibilities “Universal cluster” and project OpenCirrus.

Идейно облачные вычисления близки к центрам коллективного пользования 60х годов прошлого столетия, т.е. к реализации компьютерного обслуживания на уровне обычной коммунальной услуги (сервиса).

Концептуально облачные вычисления декларируют:

- Все есть сервис (ХааS)
  - АааS: приложения как сервис
  - РааS: платформа как сервис
  - СааS: программное обеспечение как сервис
  - ДааS: данные как сервис
  - ИааS: инфраструктура как сервис
  - НааS: оборудование как сервис

Однако старая идея воплощается на принципиально новой технологической основе, как в телекоммуникациях, так и в программном обеспечении.

И как любая коммунальная услуга обладает свойствами:

- масштабируемости
- оплаты по реальному использованию (pay-as-you-go).

В настоящее время многие компании предлагают «Облачные» решения. Ожидаемый рост рынка облачных вычислений к 2015 г. до 200 млрд. долларов.

Создание чрезвычайно крупномасштабных центров обработки данных (ЦОД), в которых осуществляются облачные вычисления, обеспечивают десятикратное снижение стоимости (использование систем построенных из компонент общего назначения, дешевые помещения, масштаб и др.), резко снижают капитальные вложения в локальную инфраструктуру клиента.

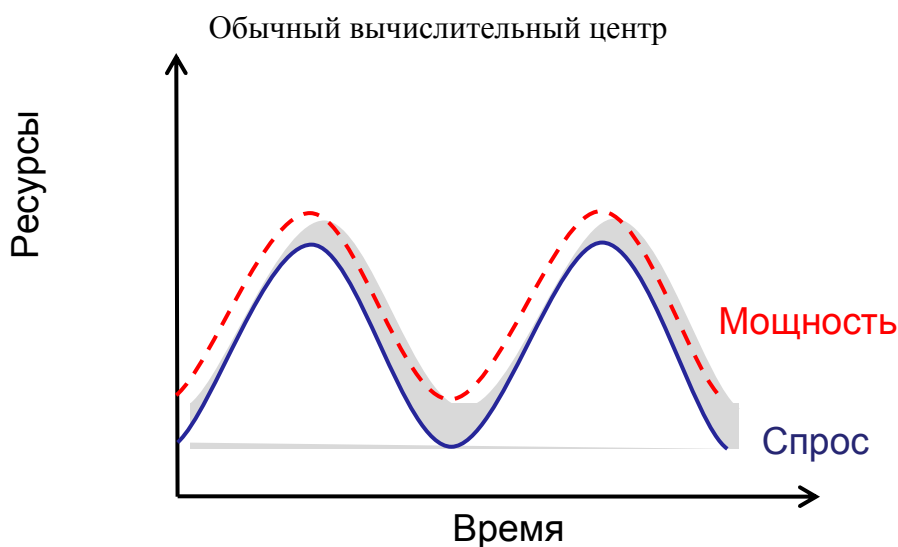
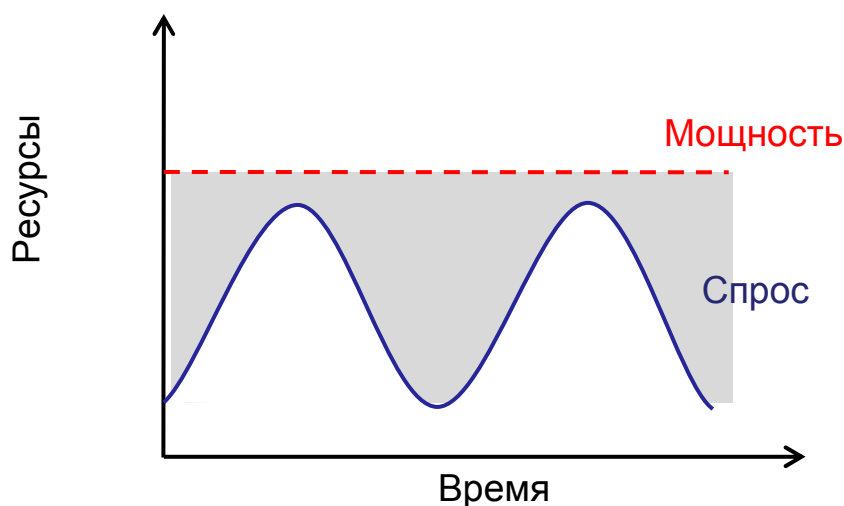
Кроме того технологии облачных вычислений опираются на:

- Всеобъемлющий широкополосный Интернет
- Эффективную виртуализацию (зависимость программы от платформы существенно ослаблена)
- Стандартный стек системного ПО
- Распространение свободного ПО

Некоторые примеры применений:

- Конвертирование большого количества файлов из одного формата в другой (пакетная обработка) *Washington post*: 17.5 тыс. стр. документации – 1500 серверчасов – 200 EC2
- Обработка запросов в Google (MapReduce) несколько тысяч запросов в секунду, каждый запрос – 20000 серверов
- Перенос в «облако» приложений, выполняемых на ПК *Matlab*
- Доступ к прикладным пакетам, рассчитанным на высокопроизводительные вычисления *NanoHub*

Гибкость предоставления ресурсов может обеспечить беспрецедентную экономию – минимум неиспользуемых ресурсов, включая энергопотребление, в условиях краткосрочных пиковых нагрузок.



Облачный ЦОД



Неиспользуемые ресурсы

### «Облачные вычисления» в науке и образовании

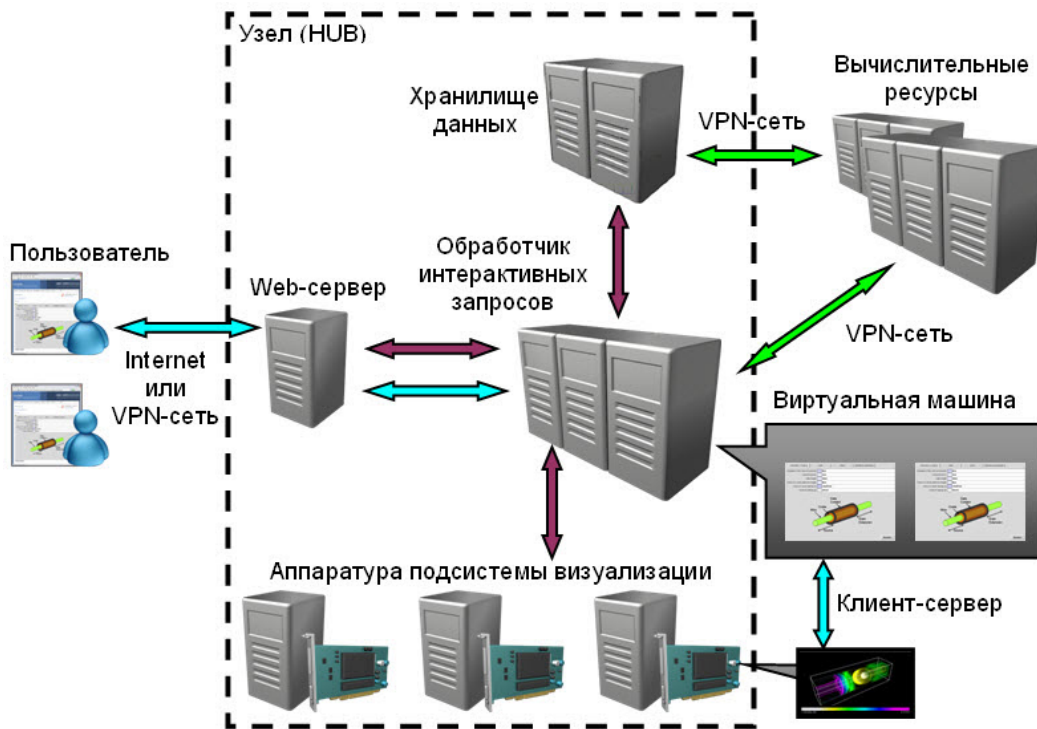
Возможность создания web-ориентированных лабораторий (хабов) в конкретных предметных областях (объединение современных концепций web 2.0 с возможностью доступа к прикладным моделям):

- интерактивный доступ к инструментам моделирования;
- поддержка распределенной разработки (система контроля версий, инструмент управления проектами и отслеживания ошибок);
- механизмы добавления новых ресурсов;
- информационные ресурсы (wiki, презентации и др.);
- поддержка пользователей;
- визуализация результатов и др.

Принципиально новые возможности для исследователей по организации доступа, разработке и распространению прикладных моделей (следствие возможность создания сообществ профессионалов в специализированных областях, стандартизация используемого инструментария, форматов хранения данных и др.).

Принципиально новые возможности по передаче знаний: лекции, семинары (практические занятия), лабораторные работы и др.

Приведем общую структуру организации «хаба».



Упрощенно стек программного обеспечения выглядит следующим образом.

Интерактивные инструменты моделирования	Механизм интеграции ресурсов	Средства поддержки пользователей	Информационные Wiki разделы, блоги
Поддержка разработки новых инструментов моделирования	Документация, интерактивные курсы	Социальные сети пользователей	

Linux, Apache, LDAP, PHP, Joomla, MySQL, Sendna (ИСП РАН), Xen, Hadoop, VNC, Rapture Toolkit,

Globus, Condor-G, gLite, ...C

## Программа « Университетский кластер»

Программа учреждена 4 сентября 2008 года Российской академией наук (ИСП РАН и МСЦ РАН), компаниями НР и «Синтерра»

Цель:

- повышение уровня компетенций в параллельных и распределенных вычислениях в образовательной и научно-исследовательской деятельности
- создание сообщества специалистов использующих и разрабатывающих современные технологии
- передача знаний и технологий в Российскую индустрию (энергетика, машиностроение, транспорт, связь и пр.).

Для достижения целей Программы решаются следующие задачи:

- построение, развитие и поддержка вычислительной инфраструктуры;
- создание и развертывание на базе вычислительной инфраструктуры сервисов различных уровней (в модели «облачных вычислений»);
- развертывание на базе вычислительной инфраструктуры испытательных стендов, на которых можно будет осуществлять проверку эффективности, разработку и доводку, новых концепций и парадигм программирования, новых информационных технологий
- создание учебных планов, учебных программ и средств поддержки учебных курсов
- создание и развертывание предметно-ориентированных научно-исследовательских web-лабораторий («хабов»).

Современные аппаратные технологии, современные программные технологии, современные сетевые технологии, компетенция ведущих научных центров, пользователи технологий параллельных и распределенных вычислений создают Единую инфраструктуру исследований и разработок в области параллельных и распределенных вычислений.

Сейчас в программе участвуют более 70 университетов и научных учреждений.



## Проект OpenCirrus

Дальнейшее развитие программы Университетский кластер связывается с развитием проекта Open Cirrus. OpenCirrus был основан компаниями HP, Intel и Yahoo.

Цель – создание открытого испытательного стенда на базе распределенных центров обработки данных, который призван поддержать разработчиков, как прикладных, так и системных программных средств в новой инновационной области «облачных вычислений»

Российская Академия наук, в составе ИСП РАН, МСЦ РАН и РИЦ «Курчатовский институт», стала первой (июнь 2009) в Восточной Европе и седьмой в мире организацией, присоединившейся к программе OpenCirrus, став одним из семи «центров компетенции» (Center of Excellence, COE).

К настоящему времени в рамках программы Университетский кластер реализованы базовые сетевые службы VPN «Университетский кластер» (*служба доменных имен DNS, централизованная авторизация, каталог ресурсов*)

Разворачиваются базовые сервисы:

**«рабочее место»** разработчика, обеспечивающего, в том числе, совместную разработку распределенных коллективов

**«виртуальная аудитория»** - возможности по проведению лекций, лабораторных работ в режиме «он-лайн».

Сервисы стенда в рамках проекта OpenCirrus:

Physical Resource Set (*Tycoon*), Elastic Compute (*Tashi*).

Группа сервисов, обеспечивающих работу с большими распределенными массивами данных (*Hadoop*).

Уже получено более 20 заявок на использование этих сервисов.

ИСП РАН совместно с РИЦ «Курчатовский институт» и компанией HP, на базе открытого пакета OpenFOAM реализует сервис:

*CFD Compute* – решение задач механики сплошной среды.

Для обеспечения полного цикла решения задач в рамках сервиса *CFD Compute*, а также других задач инженерного анализа реализуются сервисы:

**Scientific Visualisation** на базе пакета *ParaView*;

**CAD Compute** (инструмент построения расчетных сеток) на базе открытого пакета **SALOME**.

- Проведены 3 конференции, в том числе «Облачные вычисления. Образование. Исследования. Разработки»

- В июне 2011 будет проводится международная конференция OpenCirrus в Москве

- Планируется подключение к Программе образовательных и научно-исследовательских организаций из Белоруссии, Казахстана, Украины и Польши

- На базе инфраструктуры Программы разрабатываются предметно-ориентированные научно-исследовательских Web-лаборатории.

### **Примеры внедрения «Облачных» решений**

- Nebula – «облачная» платформа NASA
- RACE – частное облако для DISA (Defence Information Systems Agency)
- BBC США – заказали и подписали контракт с IBM на разработку защищенной инфраструктуры облачных вычислений, способной поддерживать оборонительную и разведывательную сеть
- Panasonic – предоставление сервисов на основе IBM cloud, для эффективного взаимодействия с поставщиками
- Муниципалитет города Los Angeles переводит свою IT-инфраструктуру в облако, в частности, электронную почту в *Gmail*
- Муниципалитет города Miami совместно с Microsoft разработал систему регистрации и отображения на карте неаварийных ситуаций (*Microsoft Windows Azure*).

### **Правительственные инициативы по «Облачным» решениям**

- G-Cloud – Правительственное облако Великобритании, которое опирается на инициативу: «*Deliver on Open Source, Open Standards and Reuse Strategy*»
- Kasumigaseki Cloud – правительственное облако Японии, которое, в том числе, используется для реализации элементов электронного правительства
- Federal Cloud Computing Initiative (США) – различные аспекты применения облачных вычислений в государственных учреждениях и бизнесе
- Европейское агентство по охране окружающей среды (ЕЕА) разработало платформу Eye On Earth, которая позволяет собирать информацию о большом количестве климатических и экологических факторов и отображать их на карте.

### **NEBULA Cloud Computing Platform (<http://nebula.nasa.gov/>)**

Nebula (туманность) – это проект который разрабатывается в Исследовательском центре Эймс а (NASA ) целью которого является интеграция компонент свободного ПО в единую инфраструктуру обеспечивающую высококачественный вычислительные сервисы по предоставлению мощностей, хранению данных и сетевых подключений. Nebula в настоящее время используется в открытых образовательных и исследовательских проектах.

В основе проекта лежат открытое ПО и предлагаются следующие сервисы:

- Infrastructure as a Service
- Platform as a Service
- Software as a Service

### **Свободное ПО и «Облачные вычисления»**

Одно из основных направлений развития :

- Стандартный стек системного ПО
- Распространение свободного ПО: Linux, Xen, Tashi, Hadoop, VNC, десятки прикладных пакетов и др.

Существующий уровень свободного ПО дает возможность организации «облачных вычислений» на всех уровнях:

- Nebula – «облачная» платформа NASA реализована на основе компонент из свободного ПО
- Компания Yahoo! объявила, что в 2011 г. вся используемая ее платформа будет иметь статус свободного ПО.

