

Пожидаев Михаил Сергеевич

Предстоящая эпоха: от настольных систем к распределённым вычислениям

1. Современные тенденции на рынке ПК.
2. Роль параллельных вычислений в новых условиях.
3. Теоретические аспекты новых подходов к вычислениям.

Разделение понятия “настольная система”



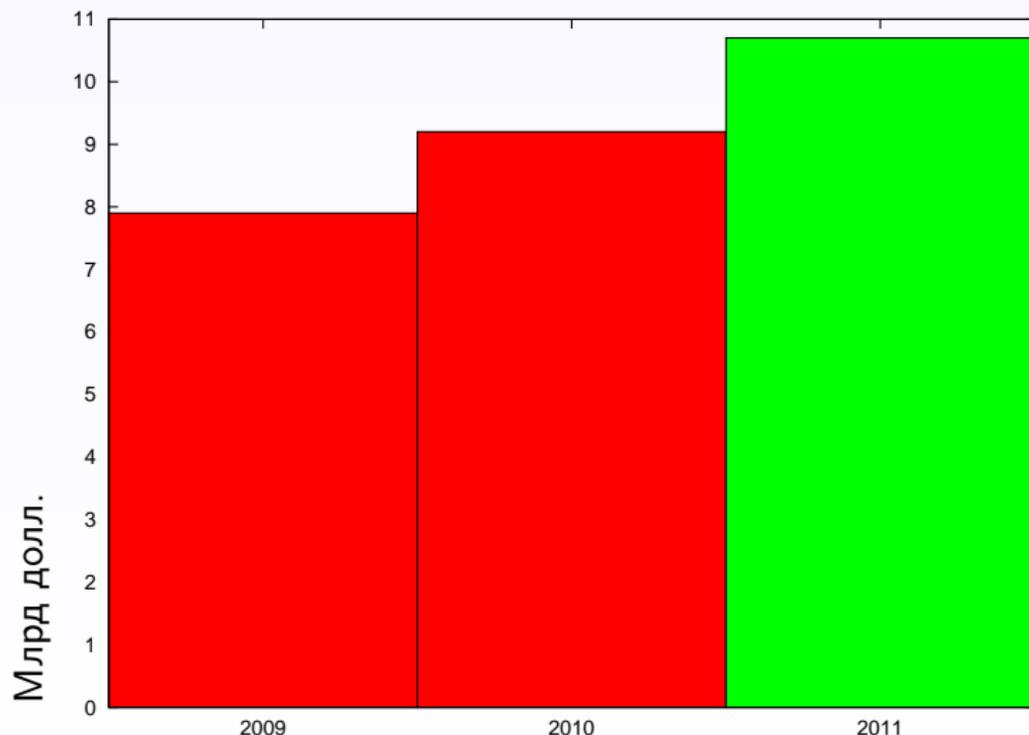
“Заря нового дня”

Мнение главного программного архитектора *Microsoft* Рэя Оззи:

“...мир движется от сложных персональных компьютеров и локального ПО к бесчисленному количеству облачных сервисов, доступ к которым предоставят устройства, по сложности сопоставимые с бытовой техникой...”

По оценке аналитиков *iSuppli* в 2013 г. мировые продажи настольных систем впервые окажутся ниже продаж мультимедийных планшетов. Более того, продажи смартфонов превысят совокупные продажи всех типов компьютеров. Со временем планшеты заменят ноутбуки. С расширением функциональности мобильных операционных систем для планшетов и развитием графического интерфейса данные устройства превратятся в полноценные рабочие инструменты, которых нужно будет всего лишь снабдить беспроводной клавиатурой.

Объёмы рынка Software as a Service (SaaS)



По данным *Gartner* мировой рынок SaaS растет на 15% в год.

Сервисы к 2015 г. — приоритетная модель приобретения ИТ

Исследование *IBM Tech Trends Survey*:

- ▶ 91% опрошенных ИТ-профессионалов считают, что в следующие пять лет сервисы станут приоритетной моделью приобретения организациями информационных технологий по сравнению с традиционным развертыванием решений непосредственно на предприятии.
- ▶ 63% респондентов признают, что им не хватает отраслевых знаний для сохранения своей конкурентоспособности.

В России в исследовании приняло участие более 650 респондентов.

Почему это происходит?

В основе лежит стремление к снижению расходов на ИТ.

1. Администраторы занимаются обслуживанием только data-центра, не затрачивая усилия на поддержку рабочих мест.
2. Одновременное решение целого комплекса задач: резервное копирование, защита от вирусов, обновление ПО и пр.
3. Низкий уровень расходов на ИТ, необходимых для начала работы новой компании.
4. Сервисы удобнее продавать.

В конкурентной борьбе участвуют все крупные компании

1. *Apple* — лидер в области мобильной техники.
2. *Google* — лидер в области сервисов.
3. *Microsoft* включилась в борьбу с запозданием.

Особенности будущей Windows 8

1. Windows будет поддерживать платформу ARM (официально).
2. Использование схемы распространения приложений через единый магазин в Интернете (неофициально).
3. Использование нового плиточного интерфейса Mosh взамен традиционных окон (неофициально).

По данным *NetApplications* доля Windows стала снижаться. При условии признания планшетных компьютеров персональными, доля Windows скоро упадёт ниже 90%.

Услуги предоставления корпоративных инструментов компании Salesforce

1. Sales Cloud — система управления взаимодействием с клиентами.
2. Service Cloud — инструмент для поддержки работы call-центров.
3. Data Cloud — инструмент для управления корпоративной информацией.
4. Collaboration Cloud — среда для общего доступа к рабочей информации между сотрудниками.
5. Custom Cloud — средство для реализации индивидуальных возможностей.

82400 клиентов по всему миру.

Сервисы требуют нового подхода к параллельным вычислениям

Традиционные способы параллельных вычислений не подходят для реализации коммерческих приложений.

1. Оборудование должно быть дешёвым и массовым.
2. Разработка приложений должна быть простой и быстрой.
3. Сервис не может быть остановлен на обслуживание.

“The data center as a computer”

“...a key challenge for software architects is to build cluster infrastructure and services that hide most of this complexity from application developers...”

1. Кластер для разработчиков приложений представлен как единый “персональный компьютер”.
2. Постоянная готовность к отказу оборудования и дешёвые средства устранения неполадок..

Основные компоненты параллельной платформы Google

1. MapReduce — вычислительная стратегия.
2. Google File System — распределённая файловая система.
3. BigTable — распределённая СУБД типа “ключ-значение”.

Apache Hadoop — свободная реализация платформы Google

Все материалы и исходные тексты опубликованы.
Крупнейшее внедрение — кластер Yahoo! , содержащий 4 000 узлов.

1. Стоечные серверы или лезвия.
2. 100-мегабитная или гигабитная сеть.
3. Неразделяемые устройства хранения.

Вычислительная схема MapReduce

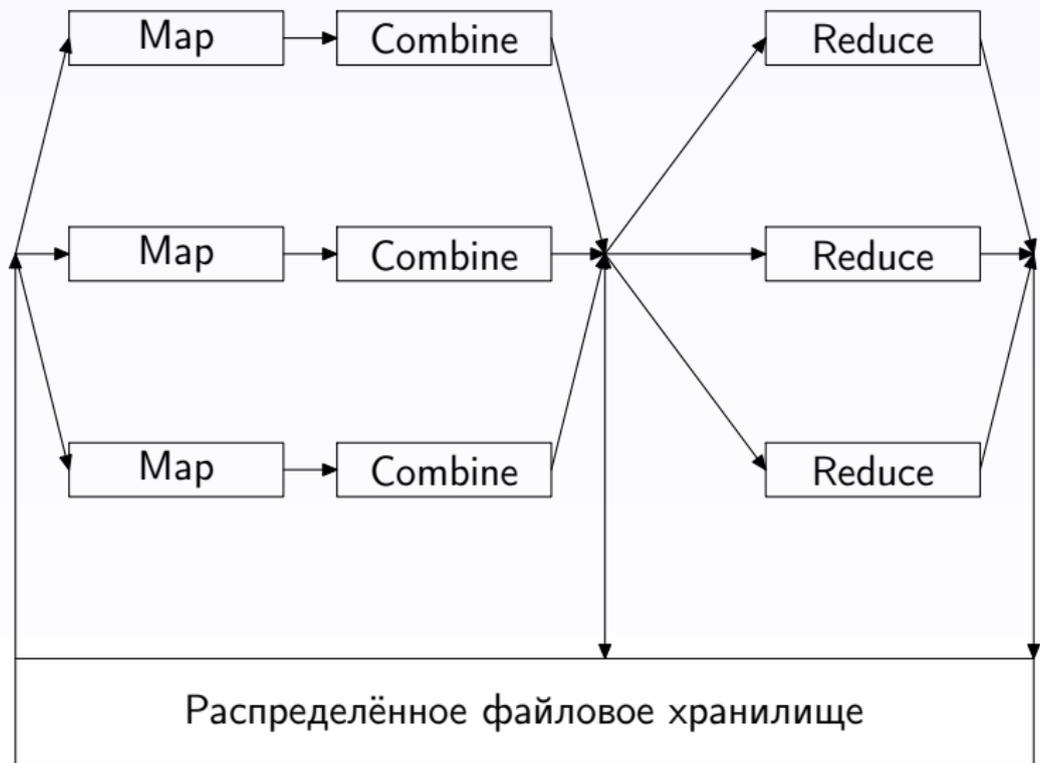
Базовая единица информации — пара “ключ-значение”.
Входные и выходные данные представляются в виде множеств пар “ключ-значение”.

1. Определяемая пользователем **map**-функция генерирует некоторое множество промежуточных пар “ключ-значение”. Требуется тщательно определить семантику ключей в промежуточных парах.
2. Общее множество промежуточных пар разбивается на группы с одинаковым значением ключа.
3. Для каждой группы вызывается определяемая пользователем **reduce**-функция, выполняющая агрегирование, результат которой включается в окончательный результат всей задачи.

Стандартные инструменты MapReduce

1. Данные загружаются из распределённой файловой системы при помощи загрузчика. Вычислительная платформа поставляется с набором стандартных загрузчиков.
2. После выполнения фазы **map** полученные результаты проходят предварительную локальную обработку на том же узле без возможности распараллеливания так называемым комбайнером (*combiner*), по назначению схожим с **reduce**-функцией.
3. Для выбора узла агрегирования используется механизм разделителей (*partitioner*).

Схема работы MapReduce



Алгоритм “PageRank”

Патент Стэнфордского университета от 4 сентября 2001 г.

Авторы: Сергей Брин и Ларри Пэйдж.

Алгоритм *PageRank* вычисляет для каждой вершины ориентированного графа v вероятность, с какой случайный маршрут на этом графе завершится в v .

Первоначальная формулировка опубликована как алгоритм ранжирования гипертекстовых документов.

Упрощённый вид PageRank для MapReduce

1. Задаётся ориентированный граф $G = (V, E)$. Каждая входная пара “ключ-значение” имеет вид $v \rightarrow W_v$. W_v — список смежных вершин для v .
2. Стадия **map**: для любой $w \in W_v$ посылается сообщение о вхождении дуги. Для v посылается сообщение о исходящей дуге. Сообщение имеет вид пары “вершина-тип сообщения”. Множество всех сообщений формирует промежуточный результат.
3. Стадия **reduce**: подсчитать для каждой вершины $v \in V$ количество сообщений обоих типов. Группировка множества промежуточных результатов по одинаковым вершинам в левой части пар выполняется автоматически реализацией MapReduce.

Хорошо исследованы:

1. Поиск в текстовых документах.
2. Базовые операции СУБД (сортировка и поиск).

Интерес к продолжению исследований

Ведутся работы по исследованию:

1. Алгоритмы на графах в MapReduce (с 2007 г., простейшие алгоритмы, поиск в ширину и пр.).
2. Геоинформационные алгоритмы в MapReduce.
3. Растровые алгоритмы в MapReduce

Спасибо за внимание!

E-mail: msp@altlinux.org

Тезисы и литература:

<http://www.marigostra.ru/materials/infTsu201103.html>