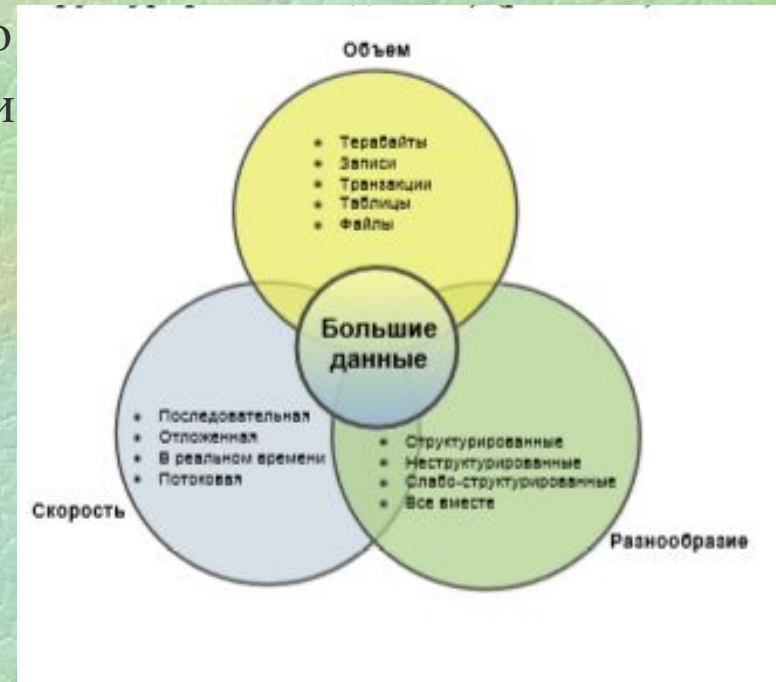


Big Data

- «большие данные» (англ. Big Data) – это серия инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов и значительного многообразия для получения необходимых результатов обработки.
- характеристики для больших данных - «три V»:
- объём (англ. volume, в смысле величины физического объёма),
- скорость (англ. Velocity, в смыслах как скорости прироста, так и необходимости высокоскоростной обработки и получения результатов),
- многообразие (англ. variety, в смысле возможности одновременной обработки различных типов структурированных и неструктурированных данных)



Технологии облачных вычислений



Современные тенденции

- По прогнозам в International Data Corporation (IDC) (<http://www.idc.com>), объем сгенерированных данных в 2012 г. составил 2,8 зеттабайта и прогнозируется к 2020 г. увеличение объема до 40 зеттабайт, что превосходит прежние прогнозы на 14% (в 1 ЗБ содержится около 1 млрд Гб, 1 зеттабайт приравнивают к объемам информации, которые могут храниться в 50 библиотеках Конгресса США).
- Значение облачных вычислений в управлении «большими данными» (Big Data): количество серверов в мире должно вырасти в 10 раз, объем данных, управляемых напрямую корпоративными центрами обработки данных (ЦОД), - в 14 раз.
- По оценкам IDC, к 2020 г. облачными вычислениями будет затронуто почти 40% данных. При этом, последующие несколько лет радикально изменятся типы хранящихся в облаках данных.

Облачные вычисления (Cloud Computing)

- Облачные вычисления – это программно-аппаратное обеспечение, доступное пользователю через Интернет в виде сервиса, позволяющего использовать удобный веб-интерфейс для удаленного доступа к выделенным ресурсам (вычислительным ресурсам, программам и данным).
- Термин «cloud computing» был впервые использован еще в 1993 г. Эриком Шмидтом для обозначения сервисов, дистанционно поддерживающих различные данные и приложения, размещенные на удаленных серверах.

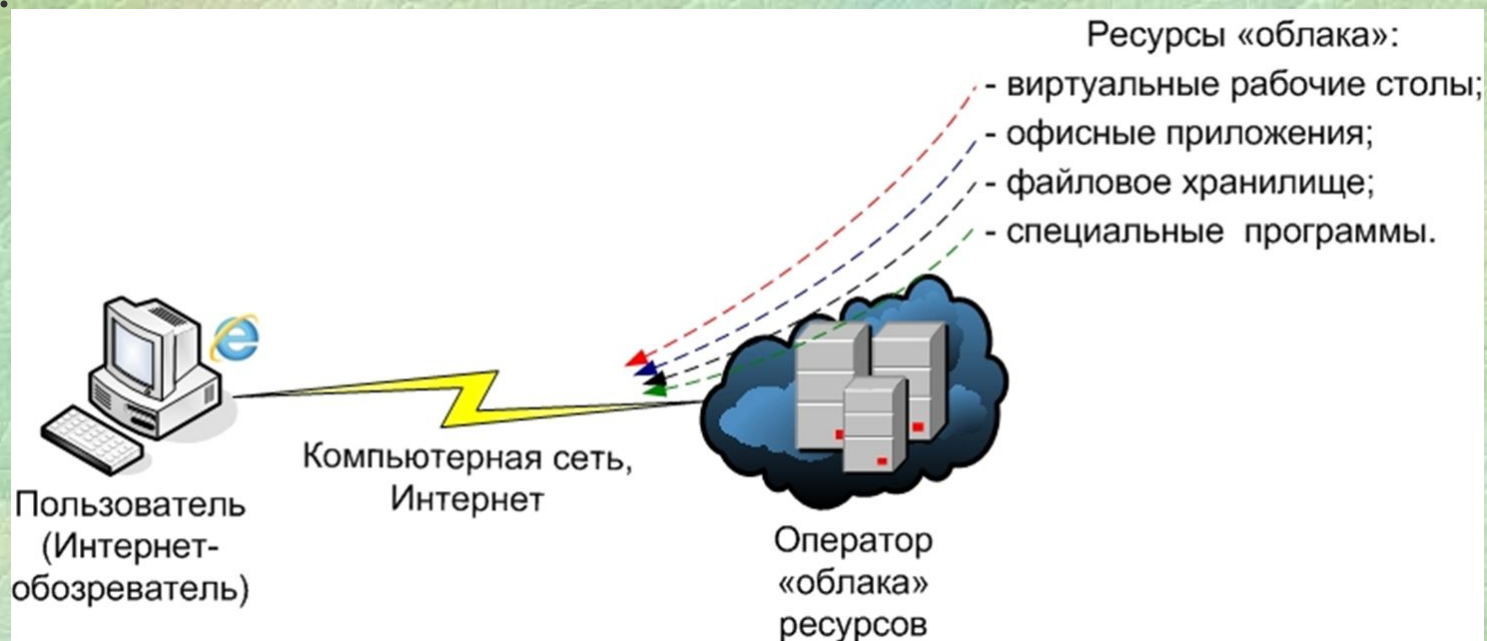
Синтез технологий и подходов приведшие к облачным вычислениям



Определение облачных вычислений, предложенное

Национальным институтом науки и технологий в 2011г.

- **Облачные вычисления** - модель предоставления возможности повсеместного и удобного сетевого доступа по требованию к пулу разделяемых **конфигурируемых вычислительных ресурсов** (например, сетям, серверам, средствам хранения, приложениям и сервисам), которые могут оперативно предоставляться и освобождаться при минимальном усилии управления или взаимодействии с провайдером (поставщиком).



Основные характеристики облачных

Характеристики облачных вычислений (Cloud Computing)

Сетевая доступность
Ubiquitous network
access

Облачные услуги должны быть доступны с любого устройства в любой точке мира в любое время

Эластичность
Elasticity

Услуги могут быть предоставлены, расширены, сужены в любой момент времени, как правило, в автоматическом режиме

Измеряемые сервисы
Metered use

Оплачивается только та вычислительная мощность (объём хранимых данных, количество пользователей), которую пользователь действительно использует

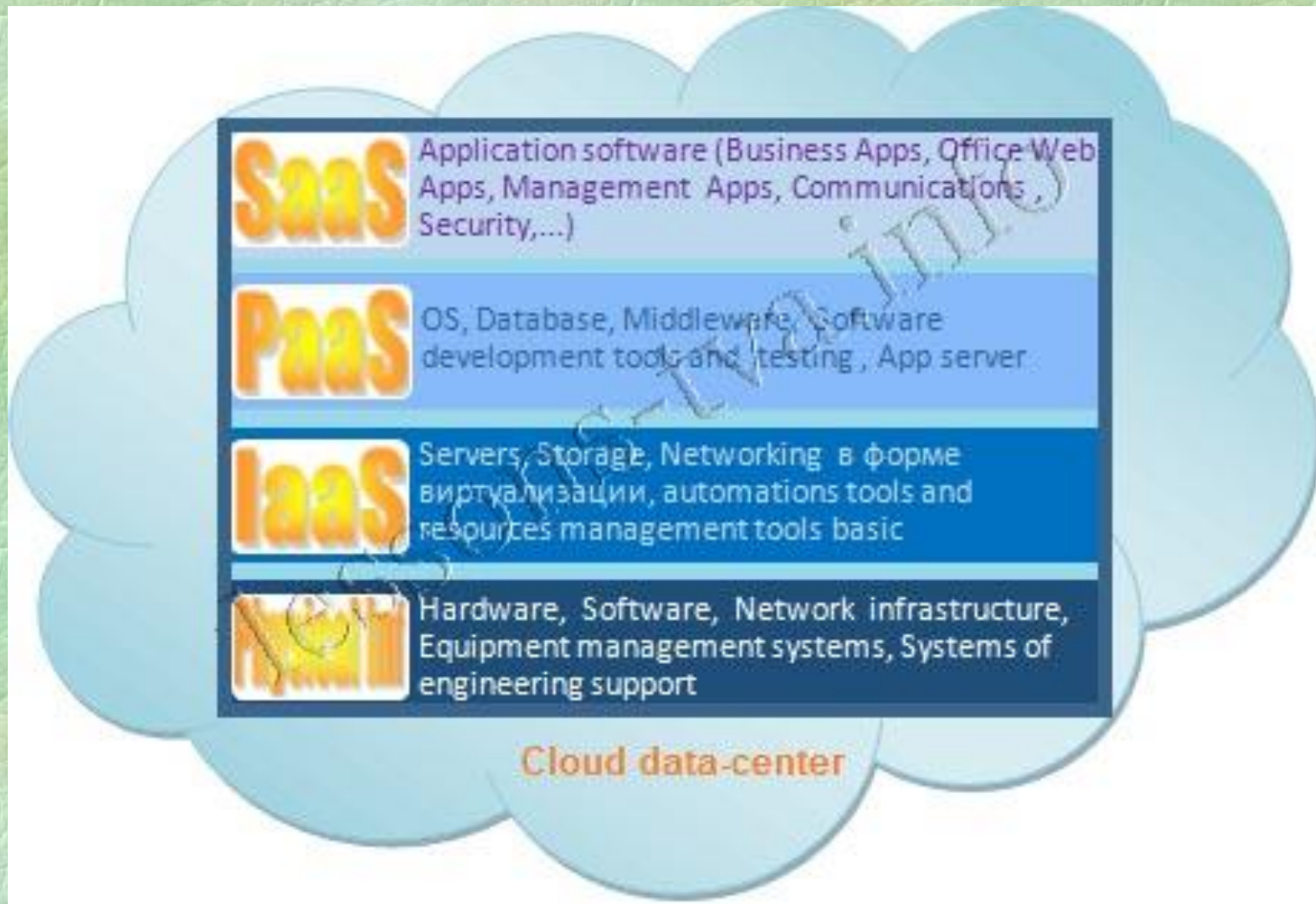
Самообслуживание по
требованию
On-demand self service

Потребитель самостоятельно определяет и изменяет вычислительные потребности без взаимодействия с представителем поставщика услуг

Объединение
ресурсов
Resource pooling

Поставщик объединяет свои вычислительные ресурсы для обслуживания большого числа потребителей, используя принцип множественной аренды. Ресурсы динамически распределяются в соответствии с потребностями пользователей.

Модели развертывания облачных вычислений



Визуальная модель определения облачных вычислений (Cloud Computing)

Характеристики

Широкополосный
сетевой доступ

Эластичность

Измеряемые
сервисы

Самообслуживание
по требованию

Объединение ресурсов

Модель поставки

Программное
обеспечение как
сервис (SaaS)

Платформа как
сервис (PaaS)

Инфраструктура
как сервис (IaaS)

Модель развертывания

Обществен-
ные

Приватные

Гибридные

Групповые



MyShared

5/3/4

Первый слой облачных услуг - IaaS

(инфраструктура, Infrastructure-as-a-Service; IaaS)

- IaaS - это предоставление пользователю компьютерной и сетевой инфраструктуры (servers, storage, networking) и их обслуживание как услуги в форме виртуализации, т.е. **виртуальной инфраструктуры**. Другими словами, на базе физической инфраструктуры дата-центров или ЦОД поставщик (провайдер) создает виртуальную инфраструктуру, которую предоставляет пользователям как сервис. Средства виртуализации позволяют преобразовать физическую инфраструктуру data centers в виртуальную и таким образом создать первый слой облачных услуг - IaaS.

Что такое виртуализация? Технология виртуализации ресурсов позволяет физическое оборудование (серверы, хранилища данных, сети передачи данных) разделить между пользователями на несколько частей, которые используются ими для выполнения текущих задач. Например, на одном физическом сервере можно запустить сотни виртуальных серверов, а пользователю для решения задач выделить время доступа к ним. Реализация виртуализации может быть осуществлена как на программном уровне, так и на аппаратном.

Первый слой облачных услуг - IaaS (инфраструктура)

- Кроме виртуализации для создания IaaS используется автоматизация, которая обеспечивает динамическое распределение ресурсов без участия персонала поставщика услуг, т.е. система автоматически может добавлять или уменьшать количество виртуальных серверов, дисковое пространство для хранения данных, или изменять сетевую пропускную способность каналов связи. **Виртуализация и автоматизация обеспечивают эффективность использования вычислительных ресурсов и снижение стоимость аренды облачной услуги IaaS.**
- Как правило, IaaS предоставляется в аренду (предоставляется IaaS-сервис на условиях аутсорсинга) корпоративным пользователям. То есть пользователи получают интегрированные ресурсы для создания своей собственной вычислительной инфраструктуры. В этом случае пользователь должен сам установить и настроить ОС и необходимые программы для выполнения производственных задач или для разработки приложений.
-

Первый слой облачных услуг - IaaS (инфраструктура)

- *Инфраструктура как сервис* - это решение корпоративного уровня для предприятий разного масштаба. Инфраструктура может быть размещена как в центре обработки данных предприятия, так и во внешнем дата-центре.
- Услуги IaaS предназначены для создания и использования защищенных частных, публичных и гибридных облачных сред.
- Кроме того, к IaaS-услугам облачных вычислений относится облачный хостинг (Cloud Hosting). *Облачный хостинг* - это хостинг, который может обеспечить динамическое распределение ресурсов, обладает возможностью автоматического масштабирования ресурсов и имеет повышенную отказоустойчивость.

Первый слой облачных услуг - IaaS (инфраструктура)

- На облачных хостингах, размещают облачные сайты – это современные облачные приложения. В облачных сайтах (облачных приложениях) данные хранятся в облачных БД, серверные приложения сайтов хранятся, и выполняется на облачных виртуальных серверах, а клиентская часть сайта выполняется в браузере пользователя.
- Кроме того, для облачных вычислений используются облачные базы данных, т.е. базы данных, которые запускаются на платформах облачных вычислений.
В качестве облачных баз данных применяются как SQL-ориентированные модели, так и модели данных NoSQL.
- Основные IaaS Solution/Vendor: **Amazon** Web Services/Amazon, **IBM** SmartCloud/IBM, SoftLayer IaaS/IBM, Azure Virtual Machines/**Microsoft**, **Google** Compute Engine/Google, **HP** Cloud/HP, **EMC**/EMC Corporation, **Oracle** Cloud Infrastructure Services/Oracle.

Второй слой - PaaS (программная платформа, Platform as a Service)

- **PaaS** предлагает разработчикам **ПО средства разработки, тестирования, развертывания и поддержки различных приложений**. Кроме того пользователю предоставляются инструменты администрирования и управления. В основном PaaS используется для разработки и размещения web-приложений (например, связанных распределенных приложений - SaaS mashup, облачных сайтов и т.д.).
- **OS** - сетевая операционная система (Unix-системы, включая Ubuntu Server, BSD/OS Family, Solaris/SunOS и т.д. или Windows Server),
- **Database** - система управления базой данных СУБД (MySQL, Microsoft SQL, SQL Database, PostgreSQL, Oracle и др.),
- **Middleware** - программное обеспечение среднего слоя или связывающее (промежуточное) программное обеспечение, которое предназначено для обеспечения взаимодействия между различными приложениями, системами и компонентами,
- **Software development tools and testing** - инструментальное программное обеспечения для разработки веб-приложений и их тестирования (среда разработки ПО: программные фреймворки, библиотеки и т.д. для создания веб-приложений на языках программирования: Python, Java, PHP, Ruby, JS для Node.js и т.д.),
- **App server** - сервер приложений для разработки, тестирования, отладки и работы веб-приложений.

Второй слой - PaaS (программная платформа)

Основные PaaS Вендоры:

- AWS Elastic Beanstalk/**Amazon**,
- **IBM** Bluemix/IBM (облачная платформа IBM Bluemix предлагает единую среду решений и широкий набор языков и фреймворков для построения приложений),
- **Microsoft** Asure/Microsoft,
- **Google** App Engine/Google,
- Salesforce1 Platform Cloud application development/Salesforce объединяет Force.com, Heroku и ExactTarget в одну семью облачных сервисов и предоставляет инструменты для разработки различных приложений.
- Heroku/Salesforce,
- Oracle Cloud Platform Services/Oracle,
- OpenShift/Red Hat,
- Cloud Foundry/VMware,



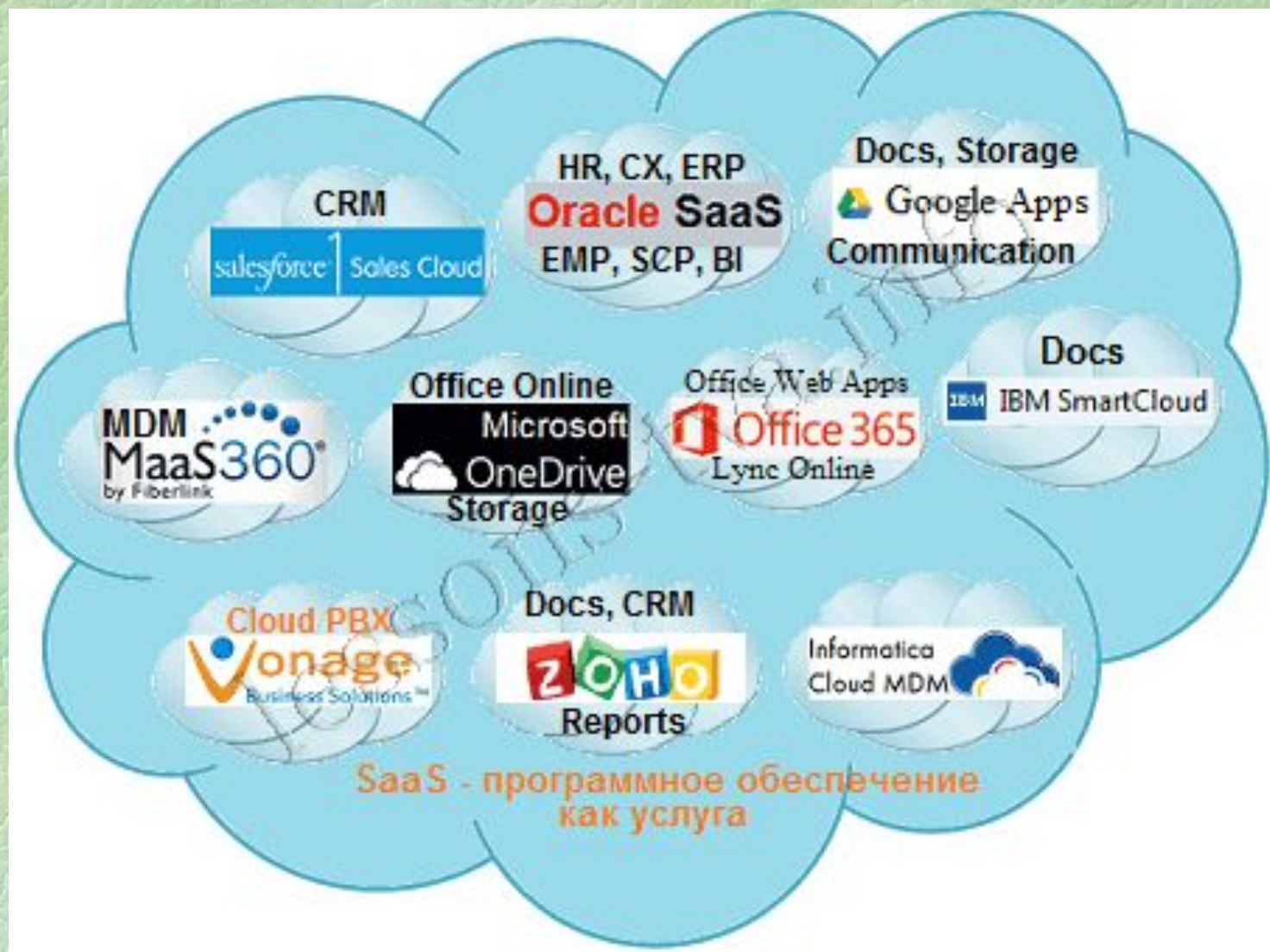
Третий слой - SaaS (облачное приложение, software as a service — программное обеспечение как услуга)

- По схеме SaaS поставляются следующие типы облачных приложений и их обслуживание: Business Apps, Office Web Apps, Management Apps, Communications, Security и др.
- Наиболее востребованными облачными приложениями являются: **CRM** (система управления взаимоотношениями с клиентами), **HRM** (система по работе с персоналом, т.е. с кадрами), **ERP** (система планирования ресурсов предприятия, например 1С), офисные приложения, средства коммуникаций и др. Salesforce.com является крупнейшим в мире поставщиком облачных приложений CRM.
- Под средствами коммуникаций подразумевается электронная переписка (например, Gmail), аудио и видео чаты (например, Microsoft Lync Online), Cloud PBX или облачная АТС (например, виртуальная АТС Манго-Офис), облачный сервис MDM (Mobile Device Management - управление мобильным устройством). Облачный сервис MDM предназначен для работы с корпоративными системами при помощи мобильных устройств.
-

Третий слой - SaaS (облачное приложение)

- Основные приложения:
- **Business Apps** (CMR, FRM, IBM® B2B Cloud Services, Axway Cloud B2B, amoCRM SaaS сервис для B2B, Google Apps for Business),
- **Business Intelligence** (PowerBI в составе Office 365/Microsoft, Oracle Business Intelligence Managed Cloud Service, Anaplan/ADE Professional Solutions, Brand Analytics),
- **Office Web Apps** (Google Docs, Office Online/Microsoft OneDrive, Office Web Apps/Microsoft, Zoho Docs, IBM SmartCloud Docs и т.д.),
- **Management Apps** (ERP/АРЕНДА 1С, HRM, SCM, MRP),
- **Communications** (Gmail, Google Hangouts, Microsoft Lync Online, Cloud PBX или облачная АТС, MDM),
- **Security** (Panda Cloud Email Protection, Panda Cloud Internet Protection, McAfee SaaS Email Protection & Continuity, Comfortway Mobile Security и т.д.), и др.
- **Collaboration and Multi-tenant** (Google Docs, Google Sites, Microsoft Office Online, Office 365, Office Web Apps).

Третий слой - SaaS (облачное приложение)



Границы управляемости облачными сервисами



MyShared

Облачные вычисления

1. Хранение данных в качестве услуги (storage-as-a-service)
2. База данных в качестве услуги (database-as-a-service)
3. Информация в качестве услуги (information-as-a-service)
4. Процесс в качестве услуги (process-as-a-service)
5. Приложение в качестве услуги (application-as-a-service = software-as-a-service)
6. Платформа в качестве услуги (platform-as-a-service)
7. Интеграция в качестве услуги (integration-as-a-service)
8. ИТ в качестве услуги (IT-as-a-service)
9. Аппаратная инфраструктура в качестве услуги (Hardware Infrastructure-as-a-service)
10. Программная инфраструктура в качестве услуги (Software Infrastructure-as-a-service)
11. Безопасность в качестве услуги (security-as-a-service)
12. Управление в качестве услуги (management/governance-as-a-service)
13. Тестирование в качестве услуги (testing-as-a-service)
14. Унифицированные коммуникации в качестве услуги (UC-as-a-service)
15. Коммуникации в качестве услуги (CaaS)
16. Инфраструктура в качестве услуги (infrastructure-as-a-service)
17. Восстановление после аварии в качестве услуги (Disaster Recovery-as-a-service)

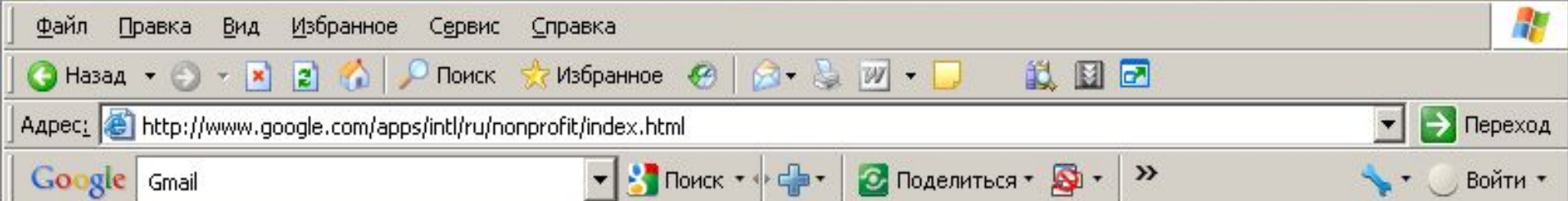


http://aws.amazon.com/ru/free/

The screenshot shows the AWS Free tier page in Russian. The browser address bar displays `aws.amazon.com/ru/free/`. The page layout includes a sidebar on the left with the 'Amazon Web Services' logo and navigation links. The main content area features a 'Сортировать по:' (Sort by) dropdown menu with options like 'Избранные продукты' (Selected products), 'Все' (All), 'Вычисления' (Compute), 'Хранилище и CDN' (Storage and CDN), 'База данных' (Database), 'Сервисы приложений' (Application services), 'Управление и администрирование' (Management and administration), 'Аналитика' (Analytics), 'Мобильные приложения' (Mobile applications), and 'приложения' (Applications). Below the dropdown, there are two main sections. The first section, titled 'Amazon', features an orange cube icon and text describing a service that is 'Масштабируемая, надежная и высокодоступная инфраструктура для хранения данных' (Scalable, reliable and highly available infrastructure for data storage). The second section, also titled 'Amazon', features a red cube icon and text describing a service that is 'Легко масштабируемая, надежная и высокодоступная инфраструктура для хранения данных' (Easily scalable, reliable and highly available infrastructure for data storage). To the right of the text, there are two green boxes with white text: '20 000 запросов Get' and '2000 запросов Put'. The sidebar on the left includes a section titled 'Уровень бесплатного пользования AWS' (AWS Free tier) and a section titled 'СВЯЗАННЫЕ ССЫЛКИ' (Related links) with links to 'Что такое облачные вычисления?' (What is cloud computing?), 'Начало работы с AWS' (Getting started with AWS), 'Продукты и сервисы AWS' (AWS products and services), and 'Размещение веб-сайта в AWS' (Hosting a website in AWS). At the bottom of the sidebar, there is a box that says 'Новые аккаунты AWS получают 12 месяцев доступа к...' (New AWS accounts get 12 months of access to...).

Бесплатные облачные сервисы

- **Google Apps.**
- Бесплатная версия включает Gmail, Google Calendar, Google Sites и Google Docs.
- Доступ к Google Apps открывается по ссылке www.google.com/apps.



Gmail, Календарь Google, Документы Google и другие службы в вашем распоряжении

[Начало работы](#)

Google Apps – набор онлайн-сервисов для совместной работы, общения и публикации информации. При этом вы получаете адреса электронной почты с доменным именем вашей организации (например, petrov@company.ru). С Google Apps некоммерческая организация сможет сократить административные расходы и получить доступ к новейшим технологиям Google. Чтобы начать работу с Google Apps нужно совсем немного времени, и это могут сделать даже те некоммерческие организации, в которых нет собственных ИТ-специалистов.



Gmail – почтовый ящик объемом Гб на один аккаунт, средства поиска почты и встроенный чат.



Календарь Google – назначайте и согласовывайте встречи, управляйте доступом к своему календарю.



Документы Google – создавайте документы и работайте над ними с коллегами без использования прикрепленных файлов.



Средства переноса электронной почты – переход на Google Apps не означает потерю имеющейся почты.



API расширения – возможности интеграции с имеющимися ИТ-системами или решениями других разработчиков.



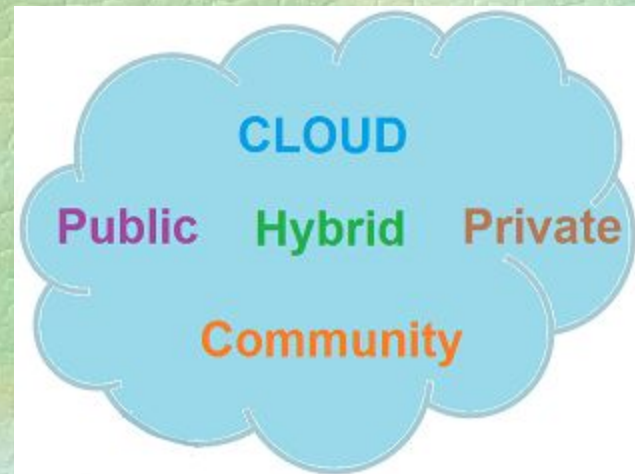
Постоянная круглосуточная поддержка клиентов – техническая поддержка по телефону и электронной почте для решения критических

проблем.

"Google Apps – это превосходное сочетание простоты и функциональности. Мы рекомендуем их тем организациям, которым нужны продуманные безотказно работающие средства с простым администрированием и поддержкой".

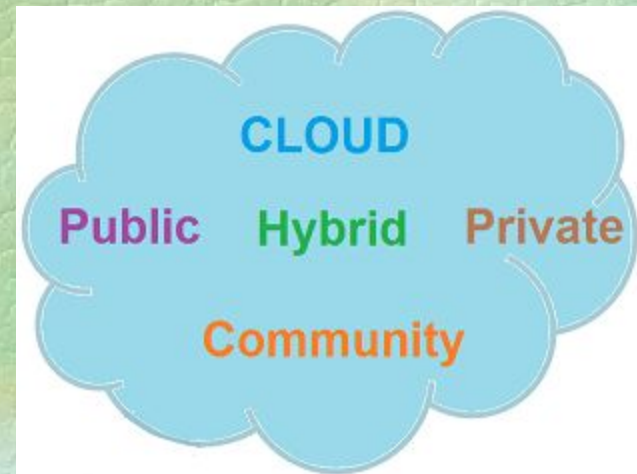
Модели развертывания облачных вычислений

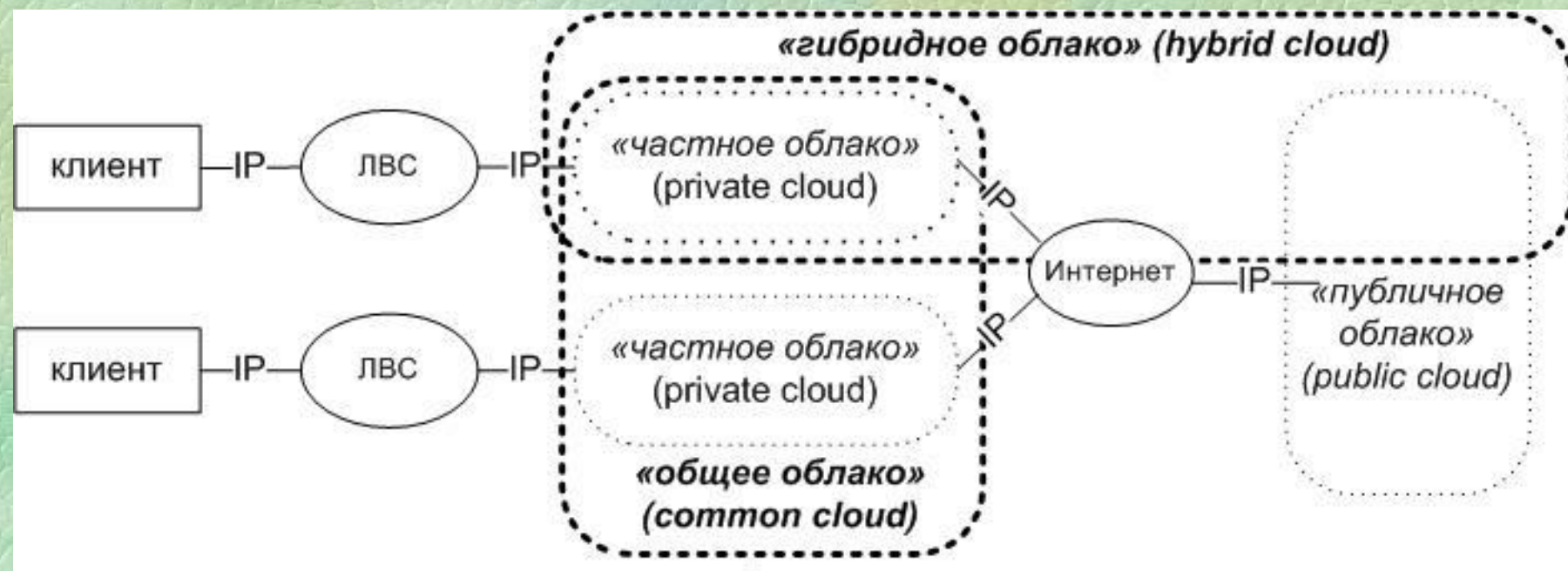
- **Private cloud** (частное облако) - инфраструктура, предназначенная для использования облачных вычислений в масштабе одной организации.
- **Community cloud** (облако сообщества) - облачная инфраструктура, которая предназначена для исключительного использования облачных вычислений определенным сообществом потребителей от организаций, которые решают общие проблемы .



Модели развертывания облачных вычислений

- **Public cloud** (публичное облако) - инфраструктура, предназначенная для свободного использования облачных вычислений широкой публикой.
- **Hybrid cloud** (гибридное облако) - это комбинация различных облачных инфраструктур (частных, публичных или сообществ), остающихся уникальными объектами, но связанных между собой стандартизованными или частными технологиями, которые обеспечивают возможность обмена данными и приложениями.





Перспективы

- Microsoft, Apple и Google именно с облачными технологиями связывают свое будущее.



Вход

Корпоративное облако СО РАН

[Облако СО РАН](#) | [Новости проекта](#) | [Ресурсы](#) | [Материалы семинаров](#) | [Участники проекта](#) | [Консолидированный заказ](#)

Сибирское отделение РАН

Microsoft

ООО «НПК Контакт»

Участники проекта

В пилотном проекте участвуют следующие институты и организации Сибирского отделения РАН:

- Институт вычислительных технологий СО РАН
- Институт нефтегазовой геологии и геофизики СО РАН им. А.А.Трофимука
- Институт геологии и минералогии СО РАН им В.С.Соболева
- Международный томографический центр СО РАН
- Институт автоматики и электрометрии СО РАН
- Институт цитологии и генетики СО РАН
- Конструкторско-технологический институт вычислительной техники СО РАН
- Институт неорганической химии СО РАН
- Институт теоретической и прикладной механики СО РАН
- Институт химической биологии и фундаментальной медицины СО РАН
- Президиум Кемеровского научного центра СО РАН
- Институт углехимии и химического материаловедения СО РАН
- Институт угля СО РАН
- Институт экологии человека СО РАН
- Институт водных и экологических проблем СО РАН
- Институт природных ресурсов, экологии и криологии СО РАН
- **Отделение ГПНТБ СО РАН**
- Центр новых медицинских технологий

Приглашаем институты и организации СО РАН присоединяться к проекту.

Исполнители

В работах над проектом участвуют:

- Денис Косяков (ИНГТ СО РАН)
- Алексей Фомин (ИВТ СО РАН)
- Андрей Гавенко (ИВТ СО РАН)
- Александр Кузнецов (ИНГТ СО РАН)
- Владимир Яковлев (ИНГТ СО РАН)
- Владимир Мальцев (ИВТ СО РАН)

© Copyright 2011 Д.В.Косяков



DropBox. Если у вас под рукой нет флешки, на которую нужно быстро скинуть файлы, закиньте ваши файлы на виртуальную флешку DropBox.

На данном сервисе вам предложат скачать программу, после скачивания и установки которой, в папке «Мои документы» вашего компьютера будет создана синхронизируемая папка DropBox.

В данную папку можно скидывать любые файлы, после чего они автоматически сохраняются в Интернете.

Доступ к этим файлам вы сможете получить с любого компьютера, зайдя в свой аккаунт DropBox или в синхронизируемую папку на вашем компьютере. В DropBox вы получаете бесплатно 2 ГБ для ваших файлов и 50 ГБ за 99\$ в год. Вы можете увеличить объем пространства бесплатно, если по вашей ссылке в DropBox регистрируются ваши друзья, за каждого из которых вы будете получать по 250 МБ дополнительного пространства, вплоть до 8 ГБ.

Загрузить инструкцию можно по ссылке

<https://www.dropbox.com/s/hu32nymw3ud16z4/Bahota%20Dropbox.doc>

Проблемы

- Недостатками облачных технологий являются: необходимость постоянного соединения с сетью Интернет; ограничения в ПО и его настройках под собственные задачи; конфиденциальность данных хранимых на публичных «облаках»; безопасность – «облако» само по себе является достаточно надежной системой, однако пока нет 100% гарантии от проникновения на него злоумышленников и вирусов.
- Из-за значительного притока пользователей сервисов, использующих облачные вычисления (например, Flickr или Amazon), фиксируются ошибки, утечка и даже потеря информации.
- Риск этот, впрочем, вряд ли превышает опасность того, что пользователь сам упустит свои данные, не создав резервных копий.
- Тем не менее, несмотря на недостатки, многие эксперты придерживаются той точки зрения, что преимущества и удобства перевешивают возможные риски использования подобных сервисов.

Fog Computing

- Для работы технологий Интернета вещей можно использовать и туманные вычисления (Fog Computing). Туманные вычисления дополняют облачные вычисления и обеспечивают взаимодействие умных вещей между собой и облачными ЦОД в виде трехуровневой иерархической структуры.
- Верхний уровень занимают тысячи облачных ЦОД, предоставляющих ресурсы, необходимые для выполнения серьезных, например аналитических, программных приложений IoT.
- Средний уровень - располагаются десятки тысяч распределенных управляющих ЦОД, в которых содержится «интеллект» Fog Computing,
- Нижний уровень - находятся миллионы вычислительных устройств умных вещей.
- Fog Computing можно определить как в максимальной степени виртуализированную платформу, поддерживающую три основных типа сервисов, образующих межмашинные коммуникации M2M: вычисления, хранение и сеть. Задача Fog Computing заключается в обеспечении взаимодействия миллиардов устройств между собой и с облачными ЦОД.

Особенности.

- **Распределение вычислительной мощности и реальное время.** Значительные вычислительные ресурсы могут быть размещены на периферии Сети, причем не должно быть зависимости от координат того места, где находится устройство, и при этом работа в режиме реального времени предполагает низкий уровень задержек при обмене данными, к тому же в Fog Computing может произойти конвергенция двух существовавших долгое время автономно друг от друга систем — управления бизнесом и технологическими системами.
- **Географическое распределение компонентов.** Модель распределения сервисов в Fog Computing менее централизована, чем для облаков, а отдельные устройства могут быть связаны между собой потоками данных и предоставлять друг другу «тяжелые» сервисы.
- **Большой объем внешних данных.** Устройства, экипированные многочисленными сенсорами, могут в реальном времени генерировать гигантские объемы данных.
- **Сложная топология.** Миллионы географически распределенных узлов могут создавать разнообразные и не детерминированные заранее связи.
- **Мобильность и гетерогенность.** Мобильность устройств потребует использования альтернативных протоколов, например протокола маршрутизации LISP (Locator/ID Separation Protocol), который позволяет разделить функциональность IP-адресов на две части: идентификаторы хостов и локаторы маршрутизации. Концепция предусматривает установку туннельных маршрутизаторов, которые будут добавлять LISP-заголовки в информационные пакеты по мере их движения по сети.