



Cloud Computing

Концепция, технология, экономика

Лобанов Игорь

СибирьЭнерго-Биллинг

Инфраструктура

- Прикладной код (wordpress)
- Платформа (Apache+PHP+MySQL)
- Операционная система (Linux)
- Сервер (hardware)
- Место в стойке и сеть

Услуги хостинга

- Shared Hosting (Apache+PHP+MySQL):
 - \$3-5
- VPS (Linux)
 - \$10-20
- Dedicated Server
 - \$50-100
- Colocation
 - \$200-400

Cloud Computing

- Парадигма (основной принцип) построения информационных систем
- Динамическая масштабируемость
- Тарификация pay as you go

Направления

- Приложение как услуга
 - Software as a Service – SaaS
- Инфраструктура как услуга
 - Infrastructure as a Service – IaaS
- Платформа как услуга
 - Platform as a Service – PaaS

Приложение как услуга

- Хостер устанавливает приложение самостоятельно
- Заказчик настраивает приложение под себя
- Заказчик не знает о других заказчиках
- Хостер гарантирует качество в широких пределах нагрузки

Инфраструктура как услуга

- Три кита базовой инфраструктуры
 - вычислительная мощность
 - сетевое подключение
 - долговременное хранение данных
- Заказчик пользуется инфраструктурой хостера как услугой

Amazon Simple Storage System

- долговременное хранение
- `http://<name>.s3.amazonaws.com/<key>`
- HTTP API
 - GET – чтение данных
 - POST – запись данных
- стоимость *pay as you go*
 - \$0.15 – хранение 1 Гб в месяц
 - \$0.10 – входящий трафик
 - \$0.17 – исходящий трафик

Amazon Cloudfront

- Content Delivery Network
- доставка файлов по HTTP с ближайшего из 14 ЦОДов Amazon
- до 1 Гб/сек и выше
- источник данных – S3
- СТОИМОСТЬ
 - ~\$0.22 – 1 Гб исходящего трафика

Amazon Elastic Compute Cloud

- вычислительная мощность
- VPS-хостинг on-demand по произвольному образу диска
- тарификация в «машиночасах»
 - от \$0.10 за 1x1.0-1.2 ГГц, 1.7 Гб, 32
 - до \$0.80 за 8x2.5-2.9 ГГц, 7 Гб, 64

Amazon EC2, кейсы

- Нагрузочная ферма
 - 20 машин x 3 часа x \$0.10 = \$6
- Вспомогательный сервер на рабочее время
 - 9 часов x 22 дня x \$0.10 = \$20
- Бесперебойная работа
 - без резервирования \$75
 - резервирование \$43 (в пересчёте на год)

Amazon EC2, продукты 1/2

- Block Store – СХД (SAN)
 - свободно монтируемое блочное устройство
 - snapshots
 - \$0.10 – хранение 1 Гб
 - \$0.10 – 1 млн. операций с диском
- Multiple Locations
 - datacenter failover

Amazon EC2, продукты 2/2

- **Elastic IP Address**
 - выделенный внешний IP
 - \$0.10 за 1 Гб трафика
- **Elastic Load Balancer**
 - транспортный уровень
 - \$0.025 за 1 час работы
 - \$0.008 за 1 Гб трафика
- **Virtual Private Cloud**
 - VPN-подключение

Растяжимая архитектура

- Система измеряет нагрузку
- **Человек** сравнивает результаты с эталоном
- **Человек** заказывает дополнительное оборудование
- Система измеряет нагрузку
- **Система** сравнивает результаты с эталоном
- **Система** заказывает дополнительное оборудование

Платформа как услуга

- Google App Engine
 - shared hosting для высоконагруженных систем
 - Java или Python
 - Google BigTable

GAE, что можно?

- распределённый кэш a-la memcached
- исходящие HTTP- и SMTP-запросы
- логирование
- cron
- долговременное хранение данных в Google BigTable
 - не mysql и вообще не SQL

GAЕ, чего нельзя?

- запись в файлы на сервере
- фоновые потоки
- длительная обработка запросов
- потоковая отдача

GAE, стоимостная модель

- бесплатные квоты
 - 1,3 млн. запросов в день
 - 10 Гб исходящего и входящего трафика в день
 - 46 часов процессорного времени в день
 - 10 млн. обращений к BigTable в день
 - 7 тыс. обращений к API email
 - ...
- за деньги квоты расширяются

Риски

- vendor lock-in
 - отраслевые стандарты только зарождаются
 - нехватка хотя бы схожих линеек
- трудности с развёртыванием
 - см. ситуация с GAE sandbox
- новое направление

Кейс 1, задача

Интернет-ресурс средней посещаемости:

- 10,000 уникальных посетителей в сутки
- 500,000 запросов в сутки
- 2 Тб. исходящего трафика в месяц при 20 Гб входящего (1:100)
- 50 Гб база данных, храним 10 снимков резервного копирования
- Для обслуживания пиковой суточной нагрузки используется кластер из 2 веб-серверов и 1 сервер под СУБД
- Оба сервера отображаются на один внешний IP-адрес с помощью балансировщика нагрузки

Кейс 1, ХОСТИНГ

- 1U, Dual Xeon 2.8Ghz, 2Gb RAM, 2x72Gb SCSI HDD - 10,000 руб./мес. x 2 + 2,500 руб. установка
- Балансировщик нагрузки -- размещение в стойке, 1U - 3,250 руб./мес. + 2,000 руб. установка
 - не считая стоимости покупки, около 15,000 руб.
- **Итого:** 22,000 руб. вложений и 23,250 руб./мес. эксплуатационных расходов (280 тыс. руб./год.)
 - не считая расходов на резервное копирование

Кейс 1, in cloud

- 2 High-CPU Medium Instance 1.7 GB of memory, 5 EC2 Compute Units (2 ядра по 2,7 ГГц)
 - годовое резервирование ($\$455 \times 2 = 27$ тыс. руб.)
 - часы работы $\$0,06 \times 2$ экземпляра $\times 24$ часа $\times 30$ дней = 1300 руб./мес.
- Трафик: $(\$0.01 + \$0.17) * 2,000$ Гб = 11 тыс. руб./мес.
- Резервное копирование на S3: 500 Гб $* \$0.15 = 2,3$ тыс. руб./мес.
- **Итого:** 14,3 тыс. руб./мес. + 27 тыс. руб./год (198 тыс. руб./год)

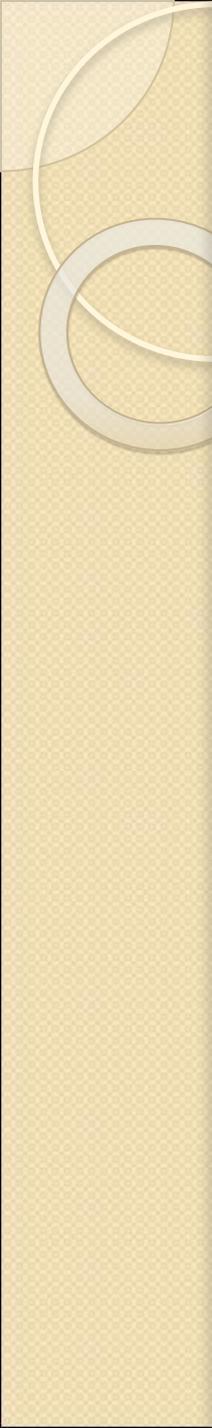
Кейс 2, задача

Популярный региональный интернет ресурс (НГС):

- 140 тыс. уникальных посетителей в сутки
- 26 млн. запросов в сутки
- 20 Тб исходящий трафик в месяц
- 1 Тб база данных
- Пиковую нагрузку держит распределённая система:
- Веб-кластер: 8 серверов x 4 ядра
- СУБД-кластер: 5 серверов x 4 ядра
- Основной и резервный балансировщик нагрузки

Кейс 2, in cloud

- 24 High-CPU Medium Instance 1.7 GB of memory, 5 EC2 Compute Units (2 ядра по 2,7 ГГц) - соблюдаем соотношение ядер
 - годовое резервирование ($\$455 \times 24 = 330$ тыс. руб.)
 - часы работы $\$0,06 \times 24$ экземпляра $\times 24$ часа $\times 30$ дней = 15,6 тыс. руб./год.
- Трафик: $\$0,01 \times 20,000 + \$0,17 \times 10,000 + \$0,13 \times 10,000 = 96$ тыс. руб./мес.
- Резервное копирование на S3: $1,000 \times \$0.15 = 4,5$ тыс. руб./мес.
- **Итого:** 116 тыс. руб./мес. + 330 тыс. руб./год.



Спасибо за внимание

Лобанов Игорь

СибирьЭнерго-Биллинг

<http://javatoday.ru>