

28 месеца scalability

(какво научих от vbox7.com)

Защо 28 месеца ?

Преди и сега

- 30 Юли 2006
 - посещения: 23
 - уникални: 13
 - импресии: 576
- вчера
 - посещения: 631 887
 - уникални: 293 434
 - импресии: 13 676 540

Пускай да върви



Началото

- Двама човека
- Лято, соса сола и пици
- Работа след работа
- Два месеца

Резултатът?

VBOX7
.com

гук можеш да споделиш твоите видео файлове
сайта в момента се доработва (т.е. beta) възможни са разни сътресения
и появата на страни съобщения. Ако много ви е страх драснете 2 реда на: bugs@vbox7.com

ВИДЕО: [ГЛЕДАЙ](#) / [КАЧИ](#)

Търси

на фокус



[The Spice Boys](#)

Какво става в момента ?

 **martines** гледа клипче 

 **krex** гледа клипче 

 **vanko_tu** гледа клипче 

 **vanko_tu** гледа клипче 

 **krex** гледа клипче 

Гледай:
>> [Нови](#)
>> [Популярни](#)
>> [Коментирани](#)

онлайн зяпачи
(последни 20 от 20)

 [vanko tu](#)

 [krex](#)

 [messi](#)

 [evelin](#)

 [darki4a](#)

 [fbi668](#)

 [anasteisha_sz](#)

 [gan](#)

 [halipaf](#)

 [lampi](#)

 [melon](#)

 [dini](#)

 [martines](#)

 [bloody_merry](#)

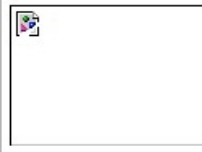
 [vikitooo](#)

 [kristianvi](#)

Новите парчоци



[DMC - VCD 99 Efnet Di Battle Скреч](#)
времетраене: 01:40мин.
гледания: 1
качил: [jevels1](#)



[Roots Manuva - Witness The Fitness](#)
времетраене: 04:41мин.
гледания: 4
качил: [jevels1](#)



[Смешните животинки](#)
времетраене: 02:10мин.
гледания: 22
качил: [karakoniul](#)



[Rol X - Хайде на море](#)
времетраене: 04:17мин.
гледания: 20
качил: [karakoniul](#)

Актуалните днес



[Фнимавайте на](#)



[Митко лудия :\)](#)



[Пиян Дядо](#)



[Опааа изненада](#)

- „Тук можеш да споделиш твоите видео файлове. Сайта в момента се доработва (т.е. beta) възможни са разни сътресения и появата на странни съобщения. Ако много ви е страх драснете 2 реда на: bugs@vbox7.com“

Основни решения

- Прост линеен код
- PHP4
- без frameworks
- без Smarty и ...
- „ просто да работи“
- UTF8
- MyISAM

Набиране на скорост

- Малко хора - малко проблеми
- Apache benchmark не е достатъчен
- Бъгове, бъгове и пак бъгове
- и малко фичъри
- Никой не мисли за скалируемост, по-важна е скоростта на разработка
- Липса на време и организация

Растеж

- октомври ~ 40 000
 - ноември ~ 50 000
 - декември ~ 70 000
 - януари ~ 100 000
-
- извод: фиксвайте бъгове, скалирането после

Първите пикове

- проблемите слаби, но проследими
- с какво да гледаме
 - EXPLAIN SELECT
 - Microslow patch (Percona)
 - Slow log analyzer
- какво да търсим
 - бавни заявки
 - много големи кръпки за ремонт

Първи изводи

- броячи и заключвания в базата
- анализ в кои полета се пише често
- анализ кои таблици четем най-често
- MyISAM vs InnoDB
- mysql_query hook
- извод: „Без пренаписване няма да мине“ :(

Още желязо

- най-лесното решение
- сървърите струват пари
- искат място
- евтините сървъри правят проблеми

Кой яде RAM?

- mysql
- статичен контент (картинки, видео ...)
- конвертиране картинки
- зареждане излишно големи масиви от данни в паметта (LIMIT 0,30)

Кой яде HDD?

- логове
- upload (ако правим допълнителни местения)
- user generated content
- backup скриптове

Кой яде CPU?

- mysql
- всякакви конвертори (да не са на web)
- лошо написан PHP код

Какво къде да upgrade

- първо RAM на MySQL, после повече ядра
- RAM не винаги е съвместим (среднощни изненади - размер, скорости)
- HDD подмяната да се планира предварително
 - „ Ще дойдат след 2 седмици ...“
 - „ Ами нямаме точно такива дискове ...“
- Българските доставчици се скалират слабо

Хардуерни проблеми

- дисковете обичат да изгарят в събота вечерта
- хубав логически райд е по-добър от евтин хардуерен
- SMART статус на дисковете
- RAID статус на контролера
- Изключвайте write-back кеша
- след 5-тия сървър вече трябва планиране

Сървърна недвижимост

- колко сървъри събира един шкаф
- стандартни размери по типове
 - 1U - web, db
 - 2U - малки storage машини, backup (8 hdd)
 - 4U - големи storage машини (24 диска)
- шкафовете се скалират много трудно
- хубавите switch-ове са скъпи

Септември 2007 - 350 000

- цикличност при ползването на интернет
 - септември - юли
 - юли - септември
- дневни цикли
 - „на работа“ - 9:00 до 17:00
(обедната почивка се усеща)
 - „В къщи“ - 20:00 до 23:30
 - много зависи от аудиторията

„Още желязо“ не работи

- MySQL
 - брой lock-нати insert-и
 - show status like 'Table_locks_waited' (бр.чакащи lock)
 - show status like 'Table_locks_immediate' (бр.получили веднага)
 - брой конекции в секунда
- PHP/Web
 - top -d1 и load average (графики, strace)
 - брой http процеси (apache benchmark)

Бързи ремонти - php

- memcached
- eaccelerator
- без frameworks
- без template engines

Бързи ремонти - mysql

- разделяне web/mysql
- master - slave репликация
 - slave конекцията без права да пише
 - много внимателно с INSERT INTO ... SELECT
 - skip error клаузи в slaves
- МОНИТОРИНГ И СТАТИСТИКА
 - NAGIOS
 - CACTI

Бързи ремонти - web

- статичния контент не се сервира от Apache
 - lighttpd
 - nginx
- внимателно с keepalive
- брой процеси и максимална памет на процес
- кукитата на отделен дял в паметта
- дребния статичен контент от паметта

Основен ремонт - база данни

- TEXT/BLOB в отделни таблици
- не пишете в силно натоварени таблици
 - броячите в memory таблица + cron за периодичен update на голямата
- денормализирането помага
- EXPLAIN всяка заявка
- ровете slow log-a
- FULLTEXT отделна таблица / сървър

Шест сървъра съм заклала

- изключвайте функционалност
- mysql
 - допълнителни slave
 - max connections 300-400
 - дърпайте LAN, преди захранването
- СТАТИЧЕН КОНТЕНТ
 - I/O scheduler магии
 - keepalive = off

Vbox7 - трети сезон

- Кога простите подходи са изчерпани
 - insert-ите минават критичната граница
 - базата е по-голяма от RAM-а
 - > 25 000 online
- Как да го забележим?
 - application wide db query hook + статистика
 - ровене на Mysql логове (mysqlda)
 - гледане на show process list при натоварен сайт
 - САСТІ отрязани графики

Как да се борим с това?

- web (haproxy + memcached)
- php - обикновено не е проблем
- mysql
 - sharding
 - mysql replication trees
 - хибридни подходи
 - NDB cluster + replication trees
 - master-master
 - специализирани сървъри

Как да се борим с това?

- search
 - sphinx
 - не работи storage engine режима нормално
 - phpAPI модула е стабилен и работи добре
- трафик
 - повече LANки на сървър (bond-инга яде CPU)
 - кеширащи проксита (varnish/squid)
 - проксита по доставчиците

Скалиране на желязото

- един доставчик
- стандартизиране на машините (шаси, захр.)
- стандартни компоненти м/у различните типове сървъри (RAID, памет, дискове)
- KVM или KVM платки (за хора с дебели инвеститори)

Мониторинг, статистика и backup

- nagios + posthook SMS
- cacti
- backup
 - web - tar + cron (не забравяйте да чистите)
 - db - mysqldump, бинарен snapshot
 - storage
 - много файлове - rsync
 - on the fly (на 2 места едновременно се записват)
 - разчитайте на RAIDa

Хората се скалират трудно

- откъде да търсим хора
 - ГОТОВИ НЯМА
 - да знаят
 - основи на програмирането
 - основи на SQL
 - „да не си мислят, че са фанали сървъра за вентилатора“
- организация на работата
 - SVN
 - staging сървър
 - големите промени = големи рискове
 - тайната формула (накрая на презентацията)
 - добър програмист спестява сървъри

Парите се скалират трудно

- по-малко програмисти, но по-добри
- пренаписването пести пари
- freelancer-ите също (ако са добри)
- маркови сървъри от Ebay
- нови HDD
- не ви трябва офис

- Тайната формула е:

- Тайната формула е:
- $\text{Scalability} = \text{stats} + \text{RTFM} + \text{Google}$

