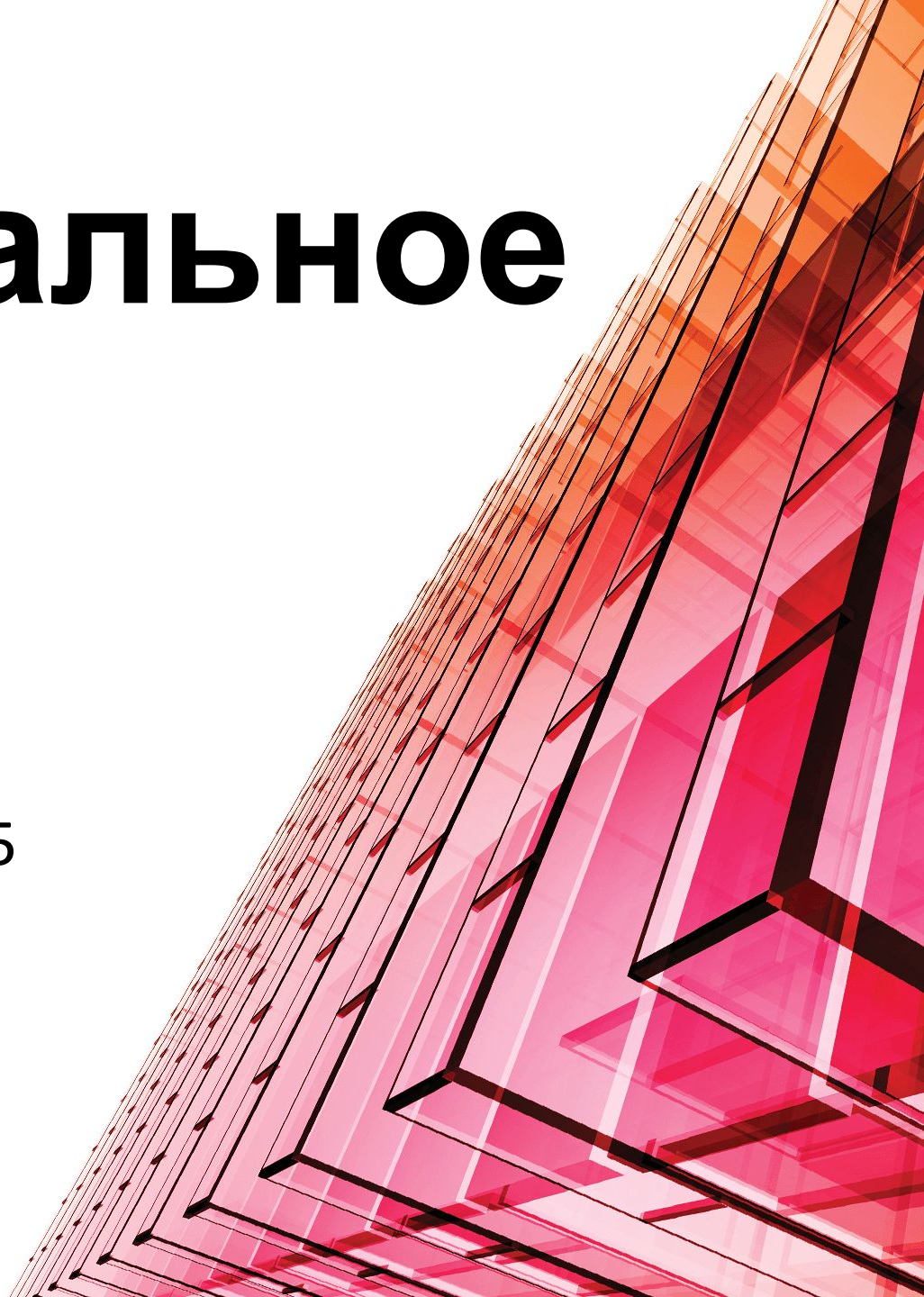
Криминальное чтиво

Андрей Аксёнов, Sphinx

v.2.0, HighLoad, Мск, 2015



HighLoad++

PULP
MINION

ENTERING
BORING

Зачем вообще доклад

- Я уже говорил всякие общие слова “как именно хорошо бы уметь”
- Хочу сказать всякие более конкретные “что именно надо бы изучать” (мегазамах)
- ВУЗы типично канают слабо, поэтому первый драфт учебного плана жахнем тут
  - Плюс, возможно, собеседования!!!
  - Тот раз в пять лет, где сложный алгоритм!!!

Disclaimers

- Слов много, времени мало, продолжать трындеть можно ~~на afterparty~~ в будке!
- Некоторые будут нехорошие, дети до 16 и ханжи за 30, покиньте зал
- Обычно я стараюсь хоть одной вещи да научить, но сегодня – только запутать!!!

“Как именно хорошо бы уметь”

- Религия, тезисно
- Системы уже сложные – и будут сложнее
  - Commander Keen … Doom 1 … UE4
  - DOS … Windows 3.x … Windows 10
  - Lynx … IE4 … Chrome
  - CSV … MySQL … ???
  - index.html … LAMP guestbook … Google
- Проблемы в них тоже – и будут сложнее
- Отчизна требует героев!

RICHARD MILLE
SWISS MADE

Удивительный мир программирования

q.e.d.

- Про картинку #2 еще могут быть сомнения – система, на самом деле, не очень сложна!
- Про картинку #3 уже нет, совершенно очевидно, герои – необходимы…

- Заметьте, идеал – ***всегда*** картинка #1

Ну так каких ещё героев?!

- Религия вкратце: я за универсалов!
  - Иначе невозможно решать проблемы на стыке
  - Иначе невозможно придумывать решения
  - Широкий кругозор, не менее 3го размера!
- Антирелигия: универсалы чмо, спецы сила!
  - Универсалы всё знают одинаково плохо
  - Чем уже спец, тем приятнее эээ разработка
- Противоречия, на самом деле, (пока) нет

Чем мерить героизм [1/3]

- **Радиус туннельного зрения**
  - Можно уметь верстать HTML+CSS – и всё
  - Можно уметь пользовать LAMP стек – и всё
  - Или представлять устройство разного софта
  - Или разных уровней железа
  - Идеально “знать” вообще всё – невозможно
  - Однако представлять в общем – таки да
  - Как следствие, внезапно…

Чем мерить героизм [2/3]

- Радиус туннельного зрения
- **Понимание границ незнания**
  - “Культура это палимпсест”
  - Я не помню, как умножать кватернионы :(
  - Я не знаю, как писать на Angular.js или Django
  - Я не знаю, как устроен RTMP
  - Да и php programmerus vulgaris не отсобеседую
  - “Но есть нюанс…”

Чем мерить героизм [3/3]

- Радиус туннельного зрения
- Понимание границ незнания
- **Способность самостоятельно действовать**
  - Не надо “думать”, надо прыгать!
  - Думать иногда надо, но лучше бы пореже

Чем мерить героизм

- Радиус туннельного зрения
- Понимание границ незнания
- Способность самостоятельно действовать

- Плюс собеседования!!!
  - Вопрос на засыпку: а зачем вообще нужны собеседования?

Хочу всё знать! А что “всё”?

- Проверочные вопросы
  - “Я в целом понимаю, как устроен X”
  - “Я знаю ряд деталей и могу тюнить/чинить X”
  - “Я и сам могу сделать этот ваш X”
  - “Толкиен дурак, Сильм лажа, не так всё было”
  - X = {процессор 8086, самый первый тетрис, гастевуха штоп пасрать, TCP стек, всё ядро, …}
- Надо бы понимать все уровни!

Собственно, ключевые слова

- **Железо.** Обязательно CPU, RAM, диск, сеть. Бонусом про видео, звук, периферию, USB, PCI, IB/FC, про энергопотребление. Что есть, как внутри устроено, как снаружи работает, чего может.
  - Иногда ещё как влияет на софт: SSD vs HDD, GPU акселерация, тупо SIMD, виртуализация, итп.

Собственно, ключевые слова

- **Софт L1. Ассемблер.** Инструкции, стеки, бранчи и цены, прерывания. Битики, байтики, форматы int/float, откуда берутся указатели. Частоты, латентности, кеши. Защищенный режим, переключения контекстов, ядра, треды, GPF vs TLB vs mwahaha. Про SIMD. Бонусом про Intel, ARM, muops, энергию, прочий ад.

Собственно, ключевые слова

- **Софт L2. OS.** Менеджмент ресурсов (память, диск, файлы, сокеты, треды, …), syscalls, IPC (shm, mmap, pipes, signals, …), ценники. Как что профайлить, трейсить, джентльменский набор vmstat, iostat, strace, oprofile, gperf итп. Виртуализация, гори она в аду. Бонусом устройство FS, IP/TCP/UDP.

Собственно, ключевые слова

- **Софт L3. Middleware.** Базы, OLTP, OLAP, DW, в широком смысле (+memcache итп). Веб, аппсервера. Бонусом обработка картинок, видео, звук, поиск, компзрение, сжатие, машобучение, bigdata, облачные хранилки и выполнялки (hadoop hadoop hadoop!) и т. п. как бы спецтемы и спецсофт.
  - Каждый пункт – отдельная большая наука, ага

Собственно, ключевые слова

- **Софт L4. Языки.** Про скрипты, нативный код, VM и байткод, JIT. Императивные, функциональные, типизированные и нет. Отладка тулзами, отладка логами, профайлинг. Статический и динамический анализ, рефакторинг, VCS, DVCS.

Экватор!

Ровно 0.5 списка ключевиков.

Щаз наконец будет и про код.

Собственно, ключевые слова

- **Софт L5. Базовый код: методы.** Базовые структуры данных: массивы, вектора, хеши, списки, деревья, очереди. Манипуляции с ними, комбинации их, родных. Базовые алгоритмы: сортировки, поиски, эээ, упс. Вычислительная сложность.
- Юнит-тесты, моки, ассерты, логи, внезапный отладчик.
- Комментарии, @#$

Собственно, ключевые слова

- **Софт L6. Код повыше: классы и потоки.** ООП, паттерны. ФП, хоть чуть. Стандартные техники: regexp, генераторы парсеров, прочий автоген, DSL.
- Потоки, файберы. Синхронизация (mutex, rwlock, прочий IPC), atomic, lockfree структуры. Races, deadlocks, отлов, борьба, недопущение. Внезапно снова ФП и пуризм.

Собственно, ключевые слова

- **Софт L7. Код сбоку: спецтемы.** Сжатие данных; работа с картинками, видео, звуком; прочий DSP; встроенные системы; интерпретаторы и компиляторы; 3D графика; компьютерное зрение; “гигазы варёза” и неизбежный MapReduce; машинное обучение; HPC; мобилы; …
  - Каждый пункт – отдельная наука опять!!!

Собственно, ключевые слова

- **Софт L8. Код еще выше: системы.** Модули, внешние и внутренние API, спеки, тесты (приемка, регрессии, перф, итп). Форматы, версии, совместимость API/данных. Беды распределенных систем, методы борьбы. Билды и беды с ними, прочий CI, поддержка деплоя.

Собственно, ключевые слова

- **Инфраструктура.** Билды и беды с ними, CI, деплой – теперь без кода. Бэкапы, восстановление. Логи, мониторинг, реакция на аварии, аналитика. Стресс-тесты, планирование емкости. Автоматизация этого всего, типа devops.

Preemptive FAQ

- **Q:** ...ать, это ж минима 37 томов читать.
  A: Наверное больше, плюс наколка: читать мало, надо стремиться их **уметь написать**.
- **Q:** ...уй и зачем же мне это всё, я херачу на PHP тупо, плюс у меня стоянка, капает кеш.
  A: Строго из личного интереса, мыгыгы, но может случайно появиться и бабло.
- **Q:** Да ты сам небось ни …ра не знаешь!

И для разнообразия о приятном

- Всё подряд и идеально (см. эксперт) знать вообще невозможно – но и не нужно
- Часть – вы знаете уже
- Часть – особо малоприменима
- Часть – лично неинтересна
- Порядок – может быть любой
- **Список неполон, но понятен – действуй!!!**

Читай книжки, например

- Но лучше не читай; можно зачитаться
- Оптимально позже; штоп проверить книжку
- Читай маны, читай код, читай гугель
- Сначала ПИШИ, потом уже всякое читай,
  не читать учимся
- Про подъем штанги еще сцуко почитай!!!

Пиши код, например

- Начинай с простого, и быстрые итерации
- ВЕЗДЕ есть развлечения на полчаса, ну или максимум на вечер
- НИГДЕ нету невообразимой rocket science, всегда очень много можно спрототипить в очень мало строк

Пиши код, например

- L1, asm. Напиши Win32 hello world. Залей окошко gray(x xor y). А теперь на MMX. А теперь померь скорость.
- L2, OS. Побенчмаркай malloc(). Позырь strace, поудивляйся. Почитай ext4 какой.
- L3, middle. Напиши простенький calcd. А теперь простенький in-mem KV. А теперь чуть сложнее с диском, Btree, libevent.

Пиши код, например

- L4, lang. Выбери новый язык, напиши чего. Посмотри дизасм, в том числе байткода.
- L5, methods. Напиши хеш. Скрести его с чем. А теперь radixsort. А теперь SIMD.
- L6, classes/threads. Перепиши calcd на yacc. А теперь многопоточно. А теперь без сайд эффектов, just because fuck you.
- L7, specials. Напиши JPEG декодер. Напиши рейтрейсер. Затекстурируй

Пиши код, например

- Экспериментируй смело, чморить некому!!!
- Пробуй разное, но не распыляйся
- Фейлись много и с удовольствием, но чини
- Не забывай учиться, в третий раз… не пацан
- Доводи до конца, не работает => не сделал

Пиши код, например

- Самое главное.

- **Напиши себе свой личный список.**
- Мой неполный, кривой, biased и вообще.

- **Не жди никакого внезапного мега-толка.**
- Его не будет все равно, муахаха.

Сверхцель всего процесса

- Мысль про магию
- Байка про ниндзю
- Шутка про зарплату!!!

Вопросы?

(Не в церкви, не обманут.)

shodan.ru/ppt
shodan@sphinxsearch.com

Опа, бонус-трек!

(если вдруг останется время)

Про хайлоад

VAUXHALL
Currys

Есть ли специфика?

- А у вас какой хайлоад?
- Длинный или толстый?
  - Кривой? Это у всех, это ничего; “мы с мамой подумали, нам из него не стрелять...”
- Длинный == “много” “быстрых” запросов
- Толстый == “мало” “тяжелых” запросов
- **“Оптимизировать нельзя скейлить”**

Длинный highload

- “Оптимизировать нельзя! => Скейлить”
  - Но аккуратно тут, может быть иллюзией
- Распределенные системы в целом
  - Middleware в частности
- Железо, борьба со сбоями, боттлнеки

Толстый highload

- “Оптимизировать! => Нельзя скейлить”
  - Т.к. смысла нет, сначала хоть POC
- Локальная оптимизация в целом
  - Middleware в частности! mwahaha
- Железо, боттлнеки

Про собеседования

Но.. Но.. Но как же собеседования?

- И причем тут вообще все эти мегасписки!?
- Давайте поговорим про собеседования!
- Что вы хотите узнать? ☺️

Большая тайна собеседований!

- Вариантов реализации много!
  - Маленькая боевая задача
  - FizzBuzz, strrev(), strdup(), …
  - IndustrialRPN
  - “Интервью имени Баткина” (c) http://blog.gamedeff.com/?p=64 вроде
  - G, F, B, Y и их никогда не нужные загадки
- А сверхцель одна!!!
- А обобщенный критерий один!!!

Большая тайна собеседований!

- Тезис раз, всем в целом похер твой React, Objective-C или успехи на почве Clojure
- Тезис два, чем больше покрытие списка, тем больше общих тем (плюс лексикон)
- Тезис три, надо уметь писать простой код, сложный надо уметь не писать

- Плюс, собственно, можно ведь... готовиться

Sphinx

HighLoad++

Теперь совсем всё.

Упс.

shodan.ru/ppt

shodan@sphinxsearch.com