

MOVING JAVA FORWARD

ORACLE°

Intro: Performance Engineering

Сергей КуксенкоJava SE Performance
sergey.kuksenko@oracle.com

Алексей ШипилёвJava SE Performance
aleksey.shipilev@oracle.com
@shipilev





San Francisco

September 30-October 4, 2012

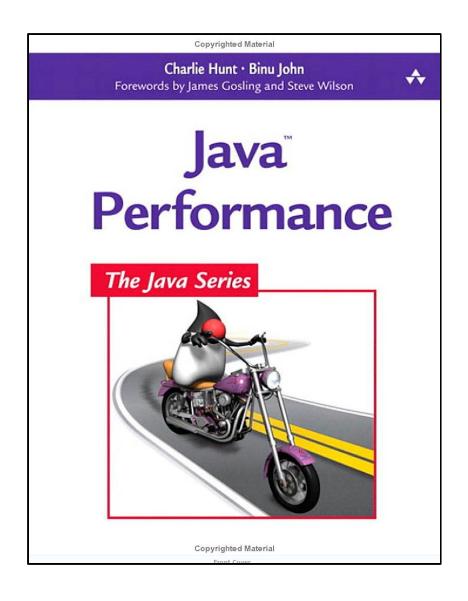


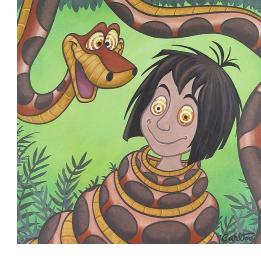


The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.













Performance Engineering





Performance Engineering

абстрактно и отлично об отличиях в абстракциях

Computer Science → Software Engineering

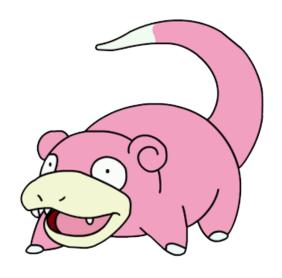
- Строим приложения по функциональным требованиям
- В большой степени абстрактно, в "идеальном мире"
 - Теоретически неограниченная свобода искусство!
 - Можно строить воздушные зАмки
- Рассуждения при помощи формальных методов

Software Performance Engineering

- "Real world strikes back!"
- Исследуем взаимодействие софта с железом на типичных данных
 - Производительность уже нельзя оценить
 - Производительность можно только измерить
- Естественно-научные методы
 - Основываемся на эмпирических данных







У меня всё медленно! Ваш первый шаг?





Performance Engineering

первый шаг

• Классические ошибки первого шага

- "я *вижу*, что метод foo() реализован неэффективно"
- "по профилю видно, что метод bar() самый горячий и занимает 5%"
- _ *"по-моему,* у нас тормозит БД, и необходимо перейти с DB_x на DB_y"





Performance Engineering

первый шаг

Классические ошибки первого шага

- "я *вижу*, что метод foo() реализован неэффективно"
- "по профилю видно, что метод bar() самый горячий и занимает 5%"
- _ *"по-моему,* у нас тормозит БД, и необходимо перейти с DB_x на DB_y"

Правильный первый шаг:

- Выбрать метрику
 - ops/sec, transactions/sec
 - время исполнения
 - время отклика
- Убедиться в корректности метрики
 - релевантна (учитывает реальный сценарий работы приложения)
 - повторяема

Цель – улучшение метрики!









Throughput, bandwidth

Throughput, Bandwidth

- Количество работы, выполненное системой за единицу времени
- Принимает разные формы:
 - MB/sec
 - ops/sec, transactions/sec
 - FPS (frames per second, frags per second)
 - MIPS
 - FLOPS





Latency, Response Time

Время...

- ...работы
 - Execution time: общее время исполнения
- ...отклика
 - Latency: время отдельной операции
 - Response time: задержка между стимулом и реакцией
- ...запуска
 - Startup time: время до начала работы
 - Time to performance: время до начала хорошей работы
- etc.





Throughput vs. Response Time

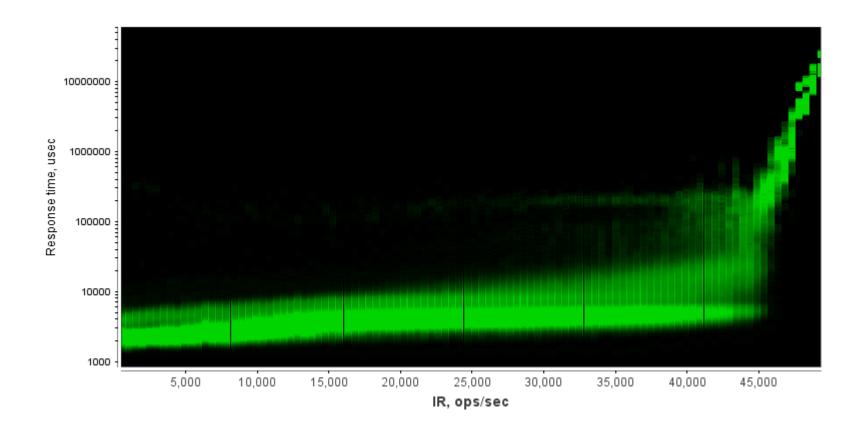
- Производительность композитная метрика:
 - throughput vs RT → throughput & RT
- Оказывается, что проще улучшить throughput:
 - DDR2 PC2-3200: 3200 Mb/sec, CL 4ns
 - DDR2 PC2-6400: 6400 Mb/sec, CL 5ns
- В большинстве систем, RT тесно связан с throughput
 - _ Из пункта А в пункт Б один автобус может перевезти сотню человек
 - Достаточно много автобусов перевезут миллион человек
 - Несколько больших автопоездов перевезут миллион человек
 - ...но в обоих случаях будут дикие очереди на погрузку и выгрузку :)





Throughput vs. Response Time

• Обычно RT быстро растёт с throughput

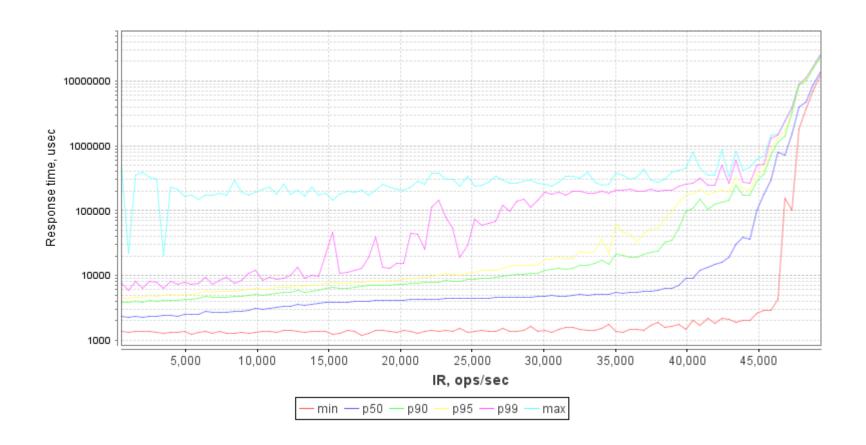






Throughput vs. Response Time

• Обычно RT быстро растёт с throughput







Прочие

• Прочие метрики:

- потребление ресурсов
 - память, сеть, etc.
- отказоустойчивость
 - MTBF, MTTR
- потребляемая и отводимая мощность
 - performance per watt
 - TDP
- etc.









Утилизация

• Утилизация – на сколько ресурс занят?

$$Utilization = \frac{ResourceBusyTime}{TotalTime}$$

• Idle – на сколько ресурс свободен?

$$IdleTime = 1 - Utilization$$

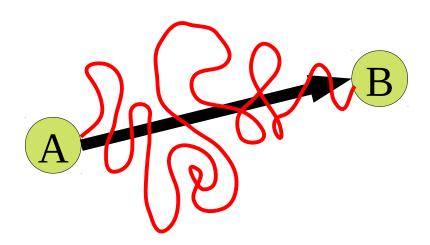




Эффективность

• Эффективность (efficiency)

- Оценка КПД для времени: часть общего времени, потраченная на выполнение полезной работы
- Субъективно, невозможно строго вычислить
 - Высокая утилизация CPU не означает хорошую эффективность







SpeedUp

• "А в N раз быстрее В" означает

$$SpeedUp = \frac{time(B)}{time(A)} = \frac{throughput(A)}{throughput(B)}$$





%Boost

• "А на n% быстрее В" означает:

$$SpeedUp = 1 + \frac{n}{100\%}$$

$$Boost\% = (SpeedUp - 1) * 100\%$$

$$Boost\% = \frac{time(B) - time(A)}{time(A)}$$

$$Boost\% = \frac{throughput(A) - throughput(B)}{throughput(B)}$$





Закон Амдала

- Допустим, есть две независимые части:
 - Часть А занимает 70% времени, ускорябельна в 2 раза.
 - Часть В занимает 30% времени, ускорябельна в 6 раз
 - Где профит больше?

70 секунд 30 секунд





Закон Амдала

• Допустим, есть две независимые части:

- Часть A занимает 70% времени, ускорябельна в 2 раза
- Часть В занимает 30% времени, ускорябельна в 6 раз
- Где профит больше?







Закон Амдала

• Закон Амдала

$$Part(A) = \frac{A}{A+B}$$

$$SpeedUp = \frac{1}{(1 - Part(A)) + \frac{Part(A)}{SpeedUp(A)}}$$

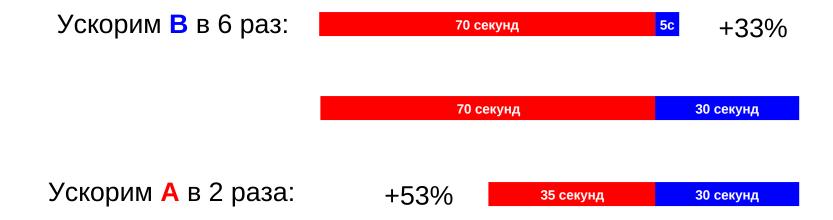




Закон Амдала

• Допустим, есть две независимые части:

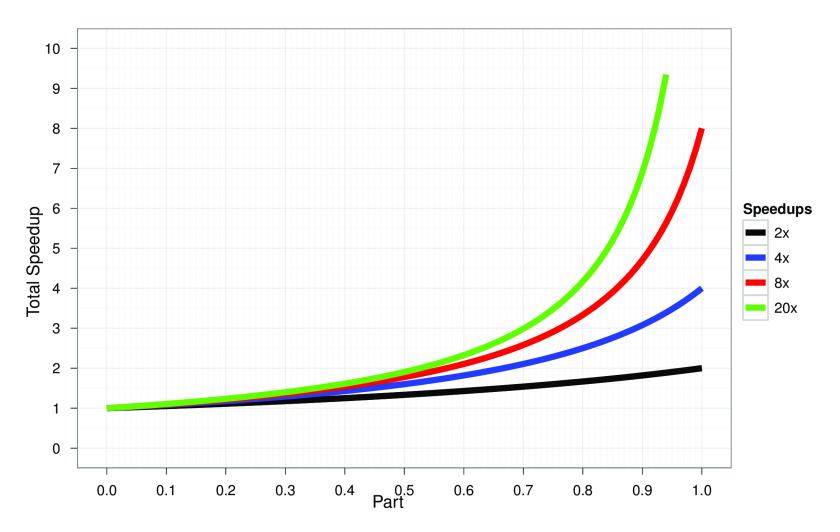
- Часть А занимает 70% времени, ускорябельна в 2 раза.
- Часть В занимает 30% времени, ускорябельна в 6 раз
- Где профит больше?







Закон Амдала







Когда закон Амдала может не работать

Composability

- Предположим, есть функциональные блоки А и В
- Есть ли разница при последовательном (...) и параллельном (||) исполнении?

• Общая функциональность:

- Functionality(A ... B) = Functionality(A || B)
- "Black Box": поведение одинаково

• Общая производительность:

- Performance(A ... B) ? Performance(A || B)
- Про это сказать толком ничего нельзя
 - А и В соревнуются за аппаратные ресурсы
 - Возможно, > (эффективный размер кеша меньше)
 - Возможно, < (два потока на НТ машине)
 - Возможно, = (нет конфликтов)





Scalability

• Масштабируемость (scalability)

 Способность системы увеличивать производительность при добавлении ресурсов

• Ресурсы:

- CPU

~ processor scaling

– тактовая частота

~ frequency scaling

- память

~ memory scaling

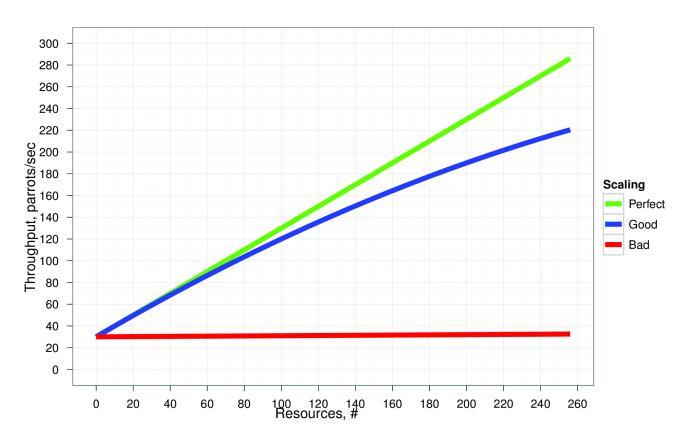
etc.





Scalability (вариант 1)

Добавляем ресурсы → работаем быстрее

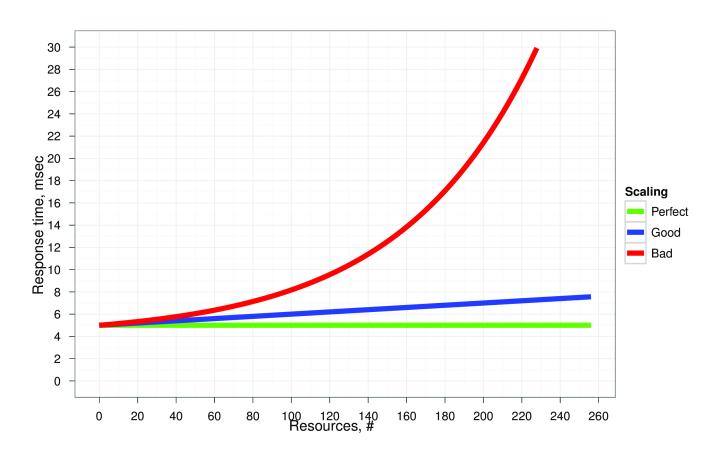






Scalability (вариант 2)

Добавляем ресурсы и нагрузку → работаем так же







Процессорно-специфичные метрики

- Clocks Per Instruction, CPI
 - Описывает параллелизм команд (чем меньше, тем лучше)
 - Обратная величина: Instructions Per Clock, IPC
- Path Length
 - количество инструкций на пути исполнения
- Время на выполнение:

TimeToExecute = P*CPI*Path Length





Методология





Узнай себя!

• «Я вижу, что метод foo() реализован плохо, перепишем и посмотрим, что изменится»





Узнай себя!

- *«Я вижу*, что метод foo() реализован плохо, перепишем и посмотрим, что изменится»
 - …а метод не используется вообще.
 - ...или используется, но занимает пару миллисекунд
 - ...или используется, но дело не в этом.
- Не самый плохой вариант, если таких методов мало, и изменения очень быстрые





Узнай себя!

• «По профилю видно, что метод bar() – самый горячий и занимает целых 5%, надо его убить»





Узнай себя!

- «По профилю видно, что метод bar() самый горячий и занимает целых 5%, надо его убить»
 - ...а оказывается, что это 5% от всего приложения, которое занимает 6.25% CPU 16-процессорной системы
 - ...или это метод, помогающий всем остальным быть быстрее
 - ...или он действительно проблемный, но *дело не в этом*.





Классические ошибки второго шага

Узнай себя!

• «Полносистемная профилировка показывает, что у нас тормозит база данных, и нужно срочно перейти с СУБД, на СУБД,»





Классические ошибки второго шага

Узнай себя!

- «Полносистемная профилировка показывает, что у нас тормозит база данных, и нужно срочно перейти с СУБД, на СУБД,»
 - ...а вдруг оказывается, что диски слабоваты
 - ...или оказывается, что злые админы зашейпили сеть
 - ...или просто БД уже давно никто не «пылесосил»
- Без шуток, мы наблюдали спонтанные переезды с Х на Ү
 - ...и обратно, немного погодя. А потом опять с X на Y.





Как ускорить приложение?

- К.О. сообщает:
 - "Нужно что-то где-то как-то изменить!"
- Что?

• Где?

Как?





Как ускорить приложение?

- К.О. сообщает:
 - "Нужно что-то где-то как-то изменить!"
- Что мешает работать быстрее?

• Где это находится?

• Как это исправить?





Как ускорить приложение?

- К.О. сообщает:
 - "Нужно что-то где-то как-то изменить!"
- Что мешает работать быстрее?
 - Используем голову и monitoring tools
- Где это находится?
 - Используем голову и profiling tools
- Как это исправить?
 - Используем голову и прямые руки





Классический "нисходящий" метод поиска узких мест

• Уровень системы

- Сеть
- Диск
- Операционная система
- Процессор/память

• Уровень приложения

- Блокировки, синхронизация
- Execution Threads
- API
- Алгоритмические проблемы

• Микроархитектурный уровень

- Code/data alignment
- Cache optimizations
- Processor stalls
- Branch prediction





"нисходящий" метод поиска узких мест

• Уровень системы

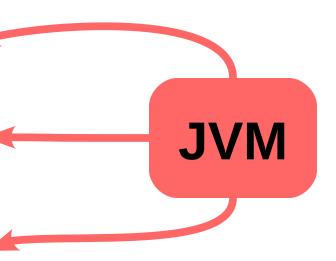
- Сеть
- Диск
- Операционная система
- Процессор/память

• Уровень приложения

- Блокировки, синхронизация
- Execution Threads
- API
- Алгоритмические проблемы

• Микроархитектурный уровень

- Code/data alignment
- Cache optimizations
- Processor stalls
- Branch prediction







"нисходящий" метод поиска узких мест (Java World)

• Уровень системы

- Сеть
- Диск
- Операционная система
- Процессор/память

• Уровень JVM

- Выбор JVM
- Heap/GC tuning
- JVM tuning

• Уровень приложения

- Блокировки, синхронизация
- Execution Threads
- API
- Алгоритмические проблемы





инструменты для анализа системы (monitoring tools)

	Solaris	Linux	Windows	Что смотрим
Сеть	netstat, dtrace, nicstat	netstat, nicstat	perfmon	количество соединений, объем трафика
Диск	iostat, dtrace	iostat	perfmon	количество обращений к диску, задержка
Память	vmstat, prstat, dtrace	vmstat, top	perfmon	подкачка страниц, размер памяти
Процессы	ps, vmstat, mpstat, prstat, dtrace	ps, vmstat, top	perfmon	количество нитей, состояние нитей, переключения контекста
Ядро ОС	mpstat, lockstat, plockstat, dtrace, intrstat, vmstat	vmstat	perfmon	kernel time, блокировки, системные вызовы, прерывания





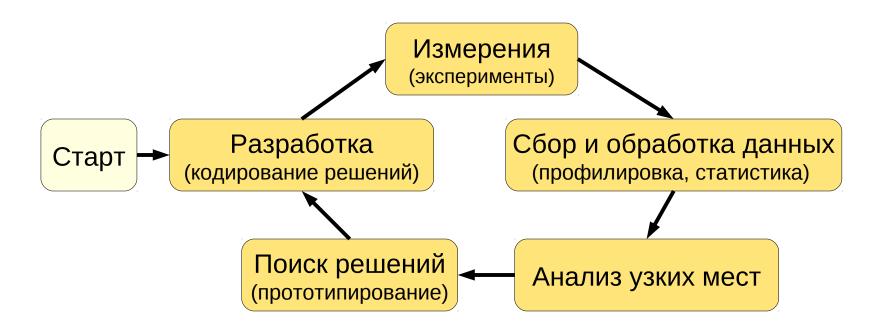
tools, tools, tools again, more tools

- VisualVM
 - http://visualvm.dev.java.net
- JRockit Mission Control
 - http://www.oracle.com/technetwork/middleware/jrockit/mission-control/index.html
- Sun Studio Analyzer
 - http://www.oracle.com/technetwork/server-storage/solarisstudio/overview/index.html
- NetBeans Profiler
 - http://www.netbeans.org
- DTrace
 - http://www.oracle.com/technetwork/systems/dtrace/dtrace/index.html
- Ещё могут быть полезны:
 - JProbe
 - OptimizeIt
 - YourKit





итеративный подход



Важно:

- Одно изменение за цикл!
- Документировать все изменения





НЕСМОТРЯ НА ДЕТАЛЬНЫЙ АНАЛИЗ ТЕКУЩЕЙ СИТУАЦИИ, Я ТАК И НЕ СМОГ СОСТАВИТЬ ЧЁТКОЕ ПРЕДСТАВЛЕНИЕ ОБ ОБСУЖДАЕМОЙ ПРОБЛЕМЕ В СИЛУ ВОЗНИКШЕГО КОНГИТИВНОГО ДИССОНАНСА.



Q/A







Backup





Tips and Tricks





CPU

• CPU, на что обратить внимание

- CPU utilization
 - User time
 - System (kernel) time
 - Idle time
- Context switching
 - Voluntary context switching
 - Involuntary context switching
- CPU scheduler run queue length





CPU

vmstat

proc	procsmemory						swap		io		tem	cpu			-
r	b	swpd	free	buff	cache	si	S0	bi	bo	in	CS	us	sy i	d w	a
256	Θ	Θ	2945568	115784	556012	Θ	Θ	Θ	0	1396	1879	99	1	Θ	Θ
256	Θ	Θ	2946176	115792	556012	Θ	Θ	Θ	40	1170	1608	100	0	Θ	Θ
256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	0	1310	1818	100	Θ	Θ	Θ
256	Θ	Θ	2945600	115792	556012	Θ	Θ	Θ	0	1189	1609	100	1	Θ	Θ
256	Θ	Θ	2946176	115792	556012	Θ	Θ	Θ	0	1261	1755	100	1	Θ	Θ
256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	0	1149	1562	100	Θ	Θ	Θ
12	Θ	Θ	2945616	115800	556012	Θ	Θ	Θ	20	1445	2952	97	0	3	Θ
177	Θ	Θ	2946312	115800	556012	Θ	Θ	Θ	0	1358	2150	96	1	3	Θ
215	Θ	Θ	2945988	115800	556012	Θ	Θ	Θ	0	1374	2497	100	0	Θ	Θ
180	Θ	Θ	2946108	115800	556012	Θ	Θ	0	0	1141	1604	99	1	Θ	Θ
127	Θ	Θ	2946876	115800	556012	Θ	Θ	40	0	1271	1810	100	0	Θ	Θ
90	Θ	Θ	2947164	115808	556052	Θ	Θ	Θ	24	1283	1786	100	1	Θ	Θ





CPU

vmstat

proc	s -		memoi	ry		swa	p	io-		-syst	tem	<i></i>	-cpu)
r	b	swpd	free	buff	cache	si	S0	bi	bo	in	CS	us	sy i	d W	a
256	0	Θ	2945568	115784	556012	Θ	Θ	Θ	Θ	1396	1879	99	1	Θ	Θ
256	0	Θ	2946176	115792	556012	Θ	Θ	Θ	40	1170	1608	100	Θ	Θ	Θ
256	0	Θ	2945888	115792	556012	Θ	Θ	Θ	Θ	1310	1818	100	Θ	Θ	Θ
256	Θ	Θ	2945600	115792	556012	Θ	Θ	Θ	Θ	1189	1609	100	1	Θ	Θ
256	Θ	Θ	2946176	115792	556012	0	Θ	Θ	Θ	1261	1755	100	1	Θ	Θ
256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	Θ	1149	1562	100	Θ	Θ	Θ
12	Θ	Θ	2945616	115800	556012	0	Θ	Θ	20	1445	2952	97	Θ	3	Θ
177	Θ	Θ	2946312	115800	556012	0	Θ	Θ	Θ	1358	2150	96	1	3	Θ
215	Θ	Θ	2945988	115800	556012	Θ	Θ	Θ	Θ	1374	2497	100	Θ	Θ	Θ
180	Θ	Θ	2946108	115800	556012	Θ	Θ	Θ	Θ	1141	1604	99	1	Θ	Θ
127	Θ	Θ	2946876	115800	556012	Θ	Θ	40	Θ	1271	1810	100	Θ	Θ	Θ
90	0	Θ	2947164	115808	556052	Θ	Θ	Θ	24	1283	1786	100	1	Θ	Θ





CPU utilization

CPU

vmstat

cont	teyt	CW	tc	hes
COLL	LUNL	3 001	U	

procsmemory						SW8	ъ	iosys			temcpu				-
r	b	swpd	free	buff	cache	si	50	bi	bo	in	cs	us s	sy i	d w	a
256	Θ	0	2945568	115784	556012	Θ	Θ	Θ	Θ	1396	1879	99	1	Θ	Θ
256	Θ	0	2946176	115792	556012	Θ	Θ	Θ	40	1170	1608	100	Θ	Θ	Θ
256	Θ	0	2945888	115792	556012	Θ	Θ	Θ	Θ	1310	1818	100	Θ	Θ	Θ
256	Θ	Θ	2945600	115792	556012	Θ	Θ	Θ	Θ	1189	1609	100	1	Θ	Θ
256	Θ	Θ	2946176	115792	556012	Θ	Θ	Θ	Θ	1261	1755	100	1	Θ	Θ
256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	Θ	1149	1562	100	Θ	Θ	Θ
12	Θ	Θ	2945616	115800	556012	Θ	Θ	Θ	20	1445	2952	97	Θ	3	Θ
177	Θ	0	2946312	115800	556012	Θ	Θ	Θ	Θ	1358	2150	96	1	3	Θ
215	Θ	0	2945988	115800	556012	Θ	Θ	Θ	Θ	1374	2497	100	Θ	Θ	Θ
180	Θ	0	2946108	115800	556012	Θ	Θ	Θ	Θ	1141	1604	99	1	Θ	Θ
127	Θ	0	2946876	115800	556012	Θ	Θ	40	Θ	1271	1810	100	Θ	Θ	Θ
90	Θ	0	2947164	115808	556052	Θ	Θ	Θ	24	1283	1786	100	1	Θ	Θ
											\ /				





CPU

vmstat

									/						
proc	s -		memor	ry		SW	ар	io-	>	-syst	tem		-cpu		-
r	b	swpd	free	buff	cache	si	SO	bi	bo	in	CS	us s	sy i	d w	a
256	Θ	0	2945568	115784	556012	0	Θ	Θ	0	1396	1879	99	1	Θ	Θ
256	Θ	0	2946176	115792	556012	Θ	Θ	Θ	40	1170	1608	100	Θ	Θ	Θ
256	Θ	0	2945888	115792	556012	0	Θ	Θ	0	1310	1818	100	Θ	Θ	Θ
256	Θ	0	2945600	115792	556012	0	Θ	Θ	0	1189	1609	100	1	Θ	Θ
256	Θ	0	2946176	115792	556012	0	Θ	Θ	0	1261	1755	100	1	Θ	Θ
256	Θ	0	2945888	115792	556012	0	Θ	Θ	0	1149	1562	100	Θ	Θ	Θ
12	Θ	0	2945616	115800	556012	0	Θ	Θ	20	1445	2952	97	Θ	3	Θ
177	Θ	0	2946312	115800	556012	0	Θ	Θ	0	1358	2150	96	1	3	Θ
215	Θ	0	2945988	115800	556012	0	Θ	Θ	0	1374	2497	100	Θ	Θ	Θ
180	Θ	Θ	2946108	115800	556012	Θ	Θ	Θ	Θ	1141	1604	99	1	Θ	Θ
127	Θ	Θ	2946876	115800	556012	Θ	Θ	40	0	1271	1810	100	Θ	Θ	Θ
90	Θ	Θ	2947164	115808	556052	Θ	0	Θ	24	1283	1786	100	1	Θ	Θ



I/O



CPU

vmstat

paging/swapping

procsmemory							ip) .	io-		-syst	tem		-cpu		-
r	b	swpd	free	buff	cache	Sİ	50	bi	bo	in	CS	us	sy i	d w	a
256	Θ	0	2945568	115784	556012	0	Θ	Θ	Θ	1396	1879	99	1	Θ	Θ
256	Θ	Θ	2946176	115792	556012	0	Θ	Θ	40	1170	1608	100	Θ	Θ	Θ
256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	Θ	1310	1818	100	Θ	Θ	Θ
256	Θ	Θ	2945600	115792	556012	Θ	Θ	Θ	Θ	1189	1609	100	1	Θ	Θ
256	Θ	Θ	2946176	115792	556012	Θ	Θ	Θ	Θ	1261	1755	100	1	Θ	Θ
256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	Θ	1149	1562	100	Θ	Θ	Θ
12	Θ	Θ	2945616	115800	556012	Θ	Θ	Θ	20	1445	2952	97	Θ	3	Θ
177	Θ	Θ	2946312	115800	556012	Θ	Θ	Θ	Θ	1358	2150	96	1	3	Θ
215	Θ	Θ	2945988	115800	556012	Θ	Θ	Θ	Θ	1374	2497	100	Θ	Θ	Θ
180	Θ	Θ	2946108	115800	556012	Θ	Θ	Θ	Θ	1141	1604	99	1	Θ	Θ
127	Θ	Θ	2946876	115800	556012	Θ	Θ	40	Θ	1271	1810	100	Θ	Θ	Θ
90	Θ	Θ	2947164	115808	556052	Θ	Θ	Θ	24	1283	1786	100	1	Θ	Θ





CPU

vmstat

CPU scheduler run queue length

1	1															
	/phocs							ap	10-		-syst	cem		-cpu		-
	r	b	swpd	free	buff	cache	si	50	bi	bo	in	CS	us	sy i	d w	a
-	256	0	Θ	2945568	115784	556012	0	Θ	Θ	Θ	1396	1879	99	1	Θ	Θ
1	256	Θ	Θ	2946176	115792	556012	Θ	Θ	Θ	40	1170	1608	100	Θ	Θ	Θ
1	256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	Θ	1310	1818	100	Θ	Θ	Θ
ı	256	Θ	Θ	2945600	115792	556012	Θ	Θ	Θ	Θ	1189	1609	100	1	Θ	Θ
ı	256	Θ	Θ	2946176	115792	556012	Θ	Θ	Θ	Θ	1261	1755	100	1	Θ	Θ
	256	Θ	Θ	2945888	115792	556012	Θ	Θ	Θ	Θ	1149	1562	100	Θ	Θ	Θ
ı	12	Θ	Θ	2945616	115800	556012	Θ	Θ	Θ	20	1445	2952	97	Θ	3	Θ
١	177	Θ	Θ	2946312	115800	556012	Θ	Θ	Θ	Θ	1358	2150	96	1	3	Θ
١	215	Θ	Θ	2945988	115800	556012	Θ	Θ	Θ	Θ	1374	2497	100	Θ	Θ	Θ
١	180	Θ	Θ	2946108	115800	556012	Θ	Θ	Θ	Θ	1141	1604	99	1	Θ	Θ
1	127	Θ	Θ	2946876	115800	556012	Θ	Θ	40	Θ	1271	1810	100	Θ	Θ	Θ
	90	Θ	Θ	2947164	115808	556052	Θ	Θ	Θ	24	1283	1786	100	1	Θ	Θ
	\ /															





CPU (симптомы)

- Высокая утилизация ядра ОС (system/kernel time)
 - I/O
 - конфликт ресурсов ОС
 - конфликт блокировок
- Низкая утилизация CPU (high idle time)
 - I/O (+ high kernel time)
 - конфликт ресурсов ОС (+ high kernel time)
 - конфликт блокировок (+ high kernel time)
 - слабая параллелизация приложения





CPU (симптомы)

•





CPU (симптомы)

- Высокая утилизация CPU (user time)
 - Неоптимальная архитектура приложения
 - Неправильное использование API
 - Неоптимальные настройки JVM (GC)
 - Неоптимизованный код приложения
 - (на этом месте можно переходить на нижний уровень)





CPU (симптомы)

- Voluntary context switching
 - I/O
 - Блокировки (lock contention)
- Involuntary context switching
- CPU scheduler run queue length
 - Рекомендуемое значение не больше 4*<кол-во СРU>
 - Что делать если больше?
 - Увеличить количество СРU
 - Уменьшить количество тредов
 - Копать глубже и оптимизировать каждый тред





выбор JVM

- Oracle (Sun) HotSpot
 - Client VM
 - Server VM
- Oracle (BEA) JRockit
- [эту JVM нам нельзя упоминать]
- [и эту тоже]
- [и эту]





выбор GC

Oracle (Sun) HotSpot

- -XX:+UseSerialGC (по-умолчанию)
- -XX:+ParallelGC
- -XX:+ParallelOldGC
- -XX:+ConcMarkSweepGC
- -XX:+UseG1GC

Oracle (BEA) Jrockit

- -Xgc:{single,gen}{con,par}{con,par}





выбор размера кучи

• Слишком мало?

- GC негде развернуться, видно по долгим и/или частым сборкам
- PermGen!

• Слишком много?

- Значит, у кого-то отобрали:
 - Кэши ФС
 - Стеки потоков
 - Direct-буфера
 - Нативные буфера





ТЮНИНГ

- Подавляющее большинство общеизвестных настроек выставлено по умолчанию
- Остальной тюнинг симптоматический:
 - Понимаем, что не работает
 - Понимаем/спрашиваем, как можно сделать быстрее
 - Печатаем -XX:+PrintFlagsFinal и находим нужный флаг
 - Гоняем функциональные тесты
 - Гоняем перформансные тесты
 - _ ???
 - PROFIT!





Уровень приложения

блокировки и синхронизация

- Конфликты на блокировках видны в профиле!
 - Даже банальный jstack может обозначить проблему
 - А если взять сотню jstack'ов подряд...
- Всегда видна только самая дикая блокировка
 - "Convoying": остальные прячутся за ней
- Лучше много маленьких, чем одна большая
 - "lock striping"
- Лучше lock-free, чем маленькая блокировка
 - Большая предсказуемость
 - Ещё лучше отсвечивается в профиле!





Уровень приложения

АРІ и алгоритмические проблемы

• Алгоритмическая сложность

- _ Не зря на собеседованиях спрашивают про сложность, ой не зря
- Компиляторы никогда не будут делать алгоритмические преобразования чужого кода
- _ Ваш супероптимизированный bogosort внутри останется bogosort'ом

Пользуйтесь правильным API

- Вам точно нужен List, а не Collection? Точно List, а не Set? Точно SortedSet, а не Set?
- _ Отходите от умолчаний, только когда очень нужно
- Exceptions

• Думаем заранее о concurrency

- Immutability
- _ Бьём себя по рукам за synchronized {}
- _ Бьём себя по рукам за Collections.synchronized...{}





JIT

для любопытных

Как получить ассемблерный код метода?

- Обычным дебаггером ;)
- JVMTI
- -XX:+PrintAssembly
 - http://wikis.sun.com/display/HotSpotInternals/PrintAssembly



