Яндекс

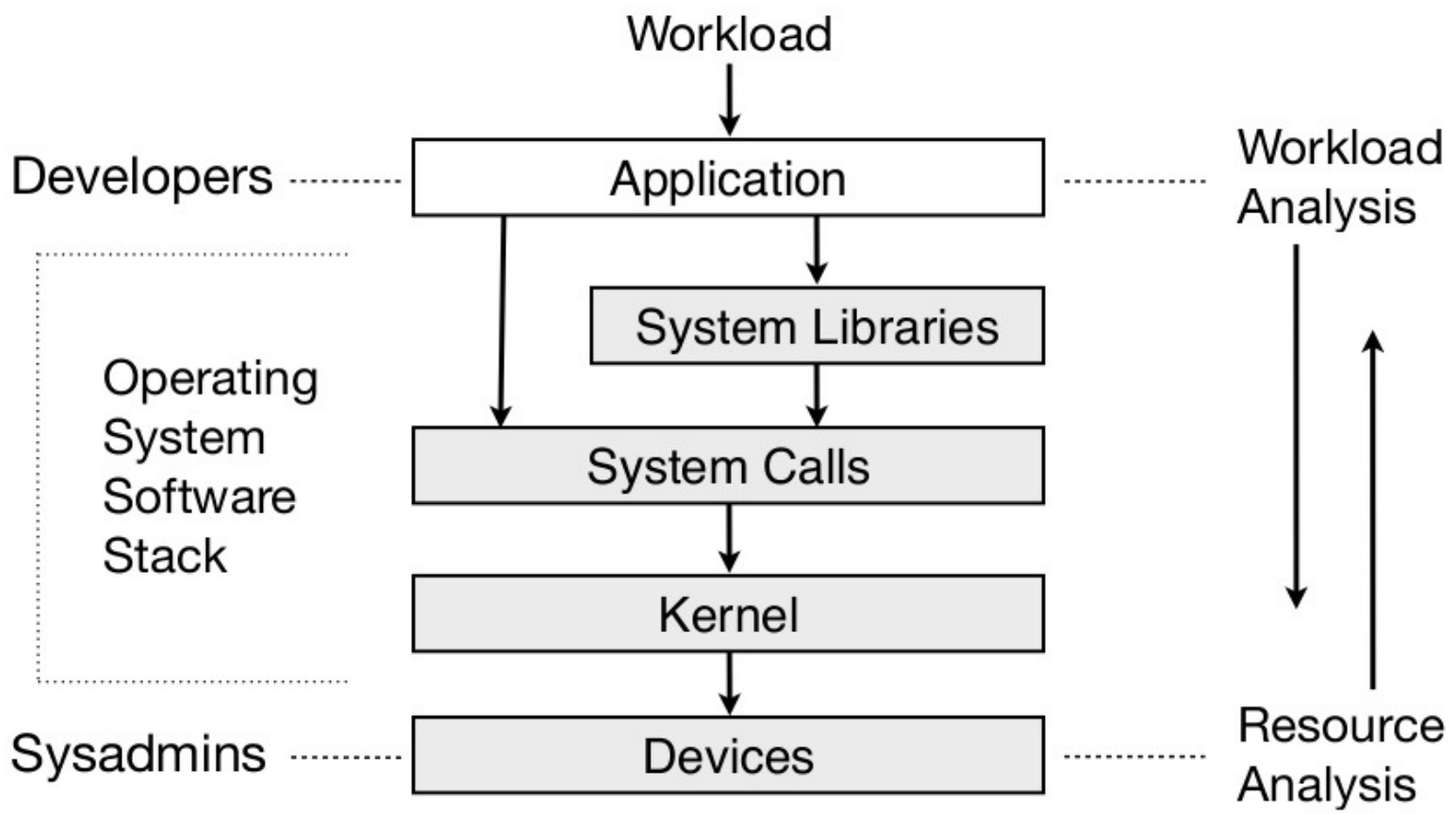
Бутылочное горлышко

Епихин Михаил, Яндекс

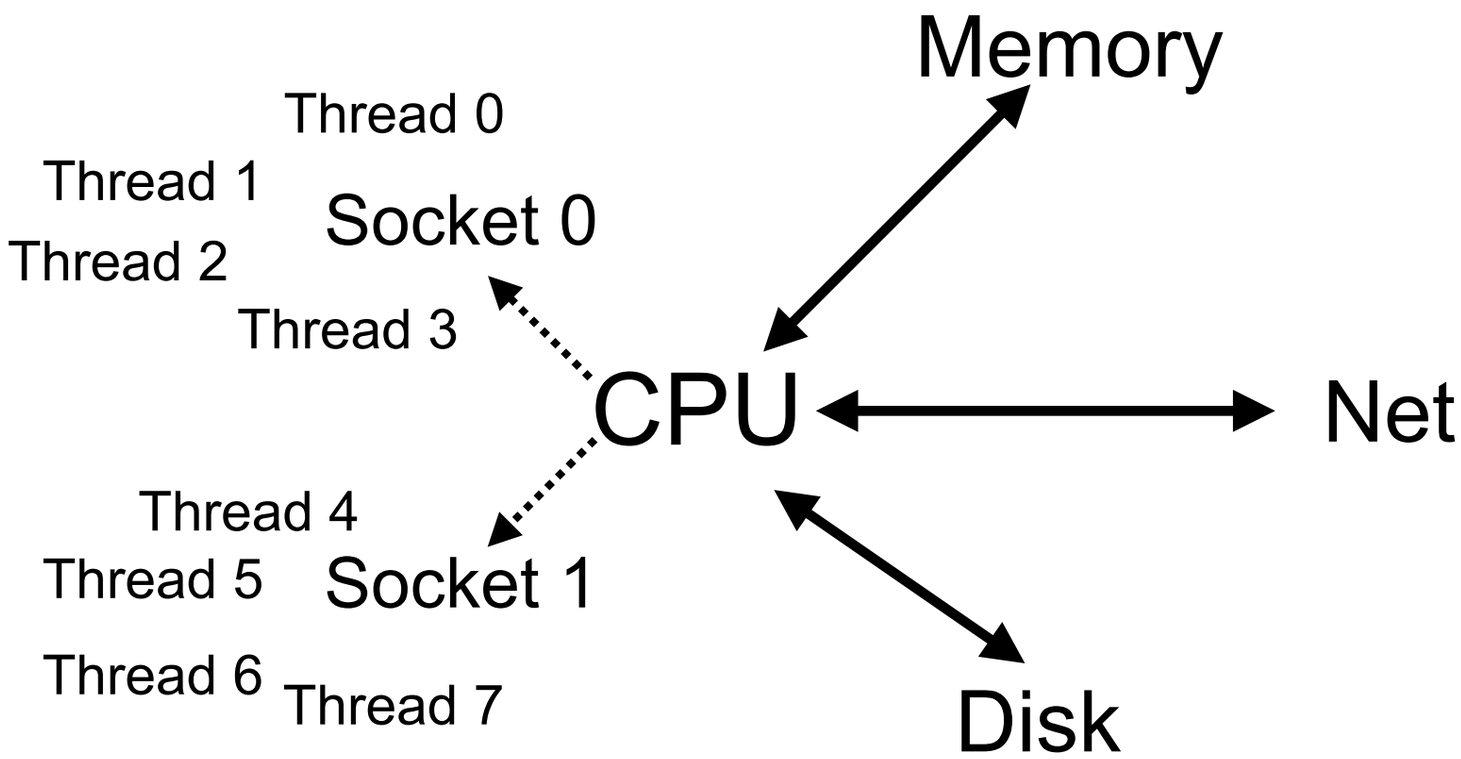
Узкое место

От всех узких мест невозможно избавиться. Узкое место можно только расширить или перенести.

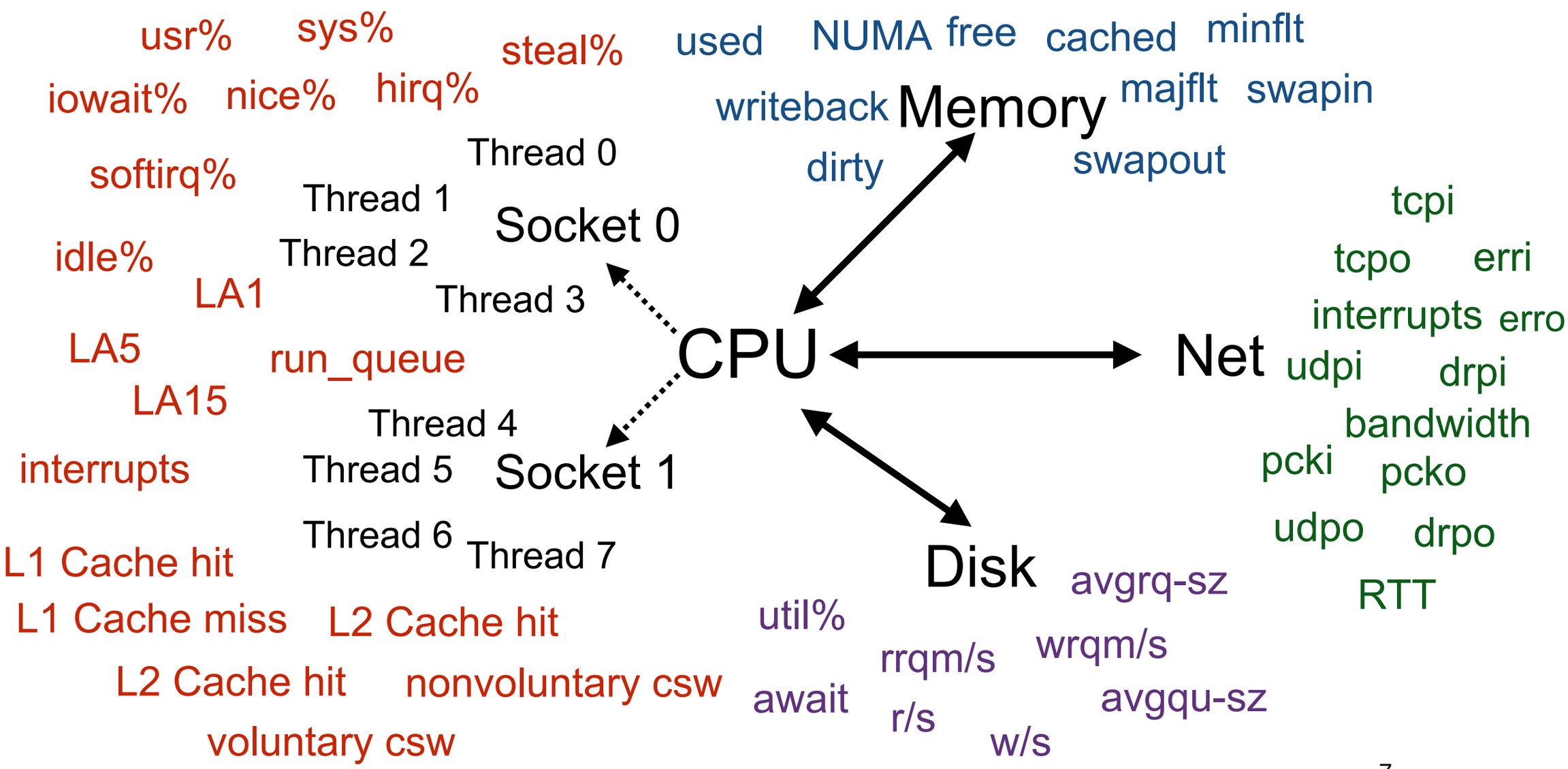
Виды анализа



Анализ ресурсов



Анализ ресурсов



USE Method

Utilization (Утилизация)

USE Method

Utilization (Утилизация)

Saturation (Насыщение)

USE Method

Utilization (Утилизация)

Saturation (Насыщение)

Errors (Ошибки)

USE Method: CPU

Utilization us%, sy%, io%, si%, st%, id%

Saturation run_queue

Errors —

USE Method: CPU & Cores

```
Cpu0  :  3.0%us,  0.3%sy,  0.0%ni,  3.0%id,  0.0%wa,  0.3%hi, 93.4%si
Cpu1  : 39.3%us, 13.5%sy,  0.0%ni, 45.9%id,  0.0%wa,  0.0%hi,  1.3%si
Cpu2  : 37.6%us, 14.2%sy,  0.0%ni, 46.2%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu3  : 36.4%us, 14.2%sy,  0.0%ni, 48.0%id,  0.0%wa,  0.0%hi,  1.3%si
Cpu4  : 31.8%us, 13.6%sy,  0.0%ni, 52.6%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu5  : 30.2%us, 12.0%sy,  0.0%ni, 56.1%id,  0.0%wa,  0.0%hi,  1.7%si
Cpu6  : 29.7%us, 14.2%sy,  0.0%ni, 54.1%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu7  : 35.0%us, 13.2%sy,  0.0%ni, 50.5%id,  0.0%wa,  0.0%hi,  1.3%si
```

USE Method: CPU & Cores

```
Cpu0  :  3.0%us,  0.3%sy,  0.0%ni,  3.0%id,  0.0%wa,  0.3%hi, 93.4%si
Cpu1  : 39.3%us, 13.5%sy,  0.0%ni, 45.9%id,  0.0%wa,  0.0%hi,  1.3%si
Cpu2  : 37.6%us, 14.2%sy,  0.0%ni, 46.2%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu3  : 36.4%us, 14.2%sy,  0.0%ni, 48.0%id,  0.0%wa,  0.0%hi,  1.3%si
Cpu4  : 31.8%us, 13.6%sy,  0.0%ni, 52.6%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu5  : 30.2%us, 12.0%sy,  0.0%ni, 56.1%id,  0.0%wa,  0.0%hi,  1.7%si
Cpu6  : 29.7%us, 14.2%sy,  0.0%ni, 54.1%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu7  : 35.0%us, 13.2%sy,  0.0%ni, 50.5%id,  0.0%wa,  0.0%hi,  1.3%si
```

idle 50%

USE Method: CPU & Cores

bottleneck

```
Cpu0  :  3.0%us,  0.3%sy,  0.0%ni,  3.0%id,  0.0%wa,  0.3%hi, 93.4%si
Cpu1  : 39.3%us, 13.5%sy,  0.0%ni, 45.9%id,  0.0%wa,  0.0%hi,  1.3%si
Cpu2  : 37.6%us, 14.2%sy,  0.0%ni, 46.2%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu3  : 36.4%us, 14.2%sy,  0.0%ni, 48.0%id,  0.0%wa,  0.0%hi,  1.3%si
Cpu4  : 31.8%us, 13.6%sy,  0.0%ni, 52.6%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu5  : 30.2%us, 12.0%sy,  0.0%ni, 56.1%id,  0.0%wa,  0.0%hi,  1.7%si
Cpu6  : 29.7%us, 14.2%sy,  0.0%ni, 54.1%id,  0.0%wa,  0.0%hi,  2.0%si
Cpu7  : 35.0%us, 13.2%sy,  0.0%ni, 50.5%id,  0.0%wa,  0.0%hi,  1.3%si
```

USE Method: CPU & Cores

```
~$ cat /proc/interrupts | tr -s ' ' | grep eth0-TxRx
91: 2868134299 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-0
92: 2868146311 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-1
93: 2868133640 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-2
94: 2868134872 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-3
95: 2868144236 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-4
96: 2868135354 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-5
97: 2868134226 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-6
98: 2868134787 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-7
```

USE Method: CPU & Cores

```
~$ cat /proc/interrupts | tr -s ' ' | grep eth0-TxRx
91: 2868134299 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-0
92: 2868146311 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-1
93: 2868133640 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-2
94: 2868134872 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-3
95: 2868144236 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-4
96: 2868135354 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-5
97: 2868134226 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-6
98: 2868134787 0 0 0 0 0 0 0 IR-PCI-MSI-edge eth0-TxRx-7
```

USE Method: Memory

Utilization free, used, cached

Saturation swapin, swapout, page fault

Errors malloc() errors

USE Method: Network

Utilization bandwidth, pcki, pcko

Saturation overruns, dropped

Errors retransmits

USE Method: Disk

Utilization bandwidth, util, r/s, w/s

Saturation iowait, await, avgqu-sz

Errors —

USE Method

Utilization (Утилизация)

Saturation (Насыщение)

Errors (Ошибки)

Resource: Lock

Utilization time_in_critical_section/s

Saturation queue_size

Errors ?

mongostat

```
1.1.1.1:27017 … db:13.0% 0 0|0 0|0 741k 657k 64 db PRI
1.1.1.1:27017 … local:149.1% 0 0|0 10|0 336b 3k 3 db SEC
```

bottleneck?

mongostat

```
1.1.1.1:27017 … db:13.0% 0 0|0 0|0 741k 657k 64 db PRI

1.1.1.1:27017 … local:149.1% 0 0|0 10|0 336b 3k 3 db SEC
```

NO!

QR/QW 0/0

mongostat

```
1.1.1.1:27017 … db:13.0% 0 0|0 0|0 741k 657k 64 db PRI

1.1.1.1:27017 … local:149.1% 0 0|0 10|0 336b 3k 3 db SEC
```

AR/AW 10/0

<твой> resource

Thread Pool или Connection Pool

Utilization *used_objects / pool_capacity*

Saturation *queue_size*

Errors *queue_overflow*

Если в процессе есть узкое место, то большинство задач будут стоять перед ним.

Thread State Analysis

D uninterruptible sleep (usually IO)

R running or runnable (on run queue)

S interruptible sleep (waiting for an event to complete)

. . .

poor man’s profiler

```
gdb -ex "set pagination 0" -ex "thread
apply all bt” \

  --batch -p $(pgrep process) > 1.td
```

TSA Method & GDB

```
~/research $ grep "#0" 1.td | sort | uniq -c

    2 #0  0x00007fa7298d2cd3 in epoll_wait ()

   17 #0  0x00007fa729b8843c in pthread_cond_wait ()

   13 #0  0x00007fa729b8b054 in __lll_lock_wait ()

    1 #0  0x00007fa729b8c2a5 in sigwait ()

    1 #0  bn_sqr4x_mont () at x86_64-mont.s:1227
```

TSA Method & GDB

```
~/research $ grep "#0" 1.td | sort | uniq -c | magic

S   2 #0  0x00007fa7298d2cd3 in epoll_wait ()

S  17 #0  0x00007fa729b8843c in pthread_cond_wait ()

S  13 #0  0x00007fa729b8b054 in __lll_lock_wait ()

S   1 #0  0x00007fa729b8c2a5 in sigwait ()

R   1 #0  bn_sqr4x_mont () at x86_64-mont.s:1227
```

TSA Method & GDB

```
~/research $ grep "#0" 4.td | sort | uniq -c | magic

S     1 #0  0x00007f51ce80cd3a in _int_free ()

R     1 #0  0x00007f51ce81dba3 in memcpy ()

S     7 #0  0x00007f51ce87ccd3 in epoll_wait ()

S    17 #0  0x00007f51ceb3243c in pthread_cond_wait()

S     1 #0  0x00007f51ceb35a2d in recvmsg ()

S     1 #0  0x00007f51ceb362a5 in sigwait ()

R     7 #0  bn_sqr4x_mont ()
```

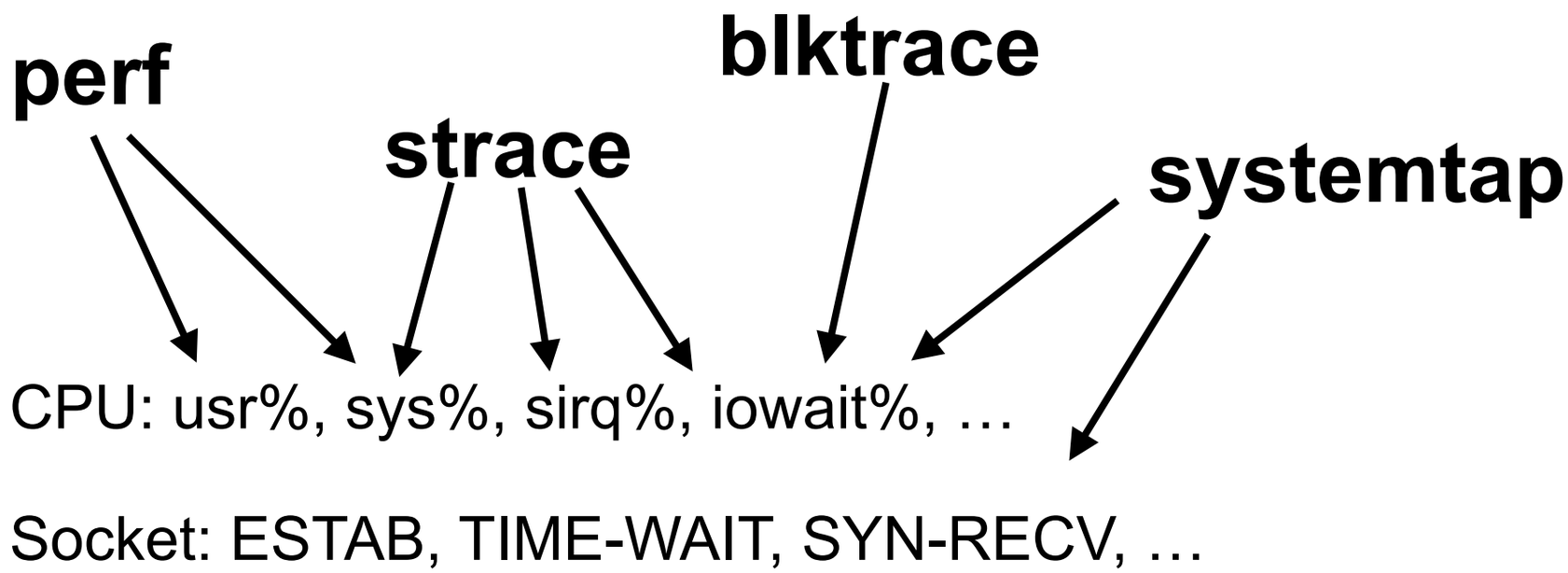
TSA Method & GDB

```
~/research $ grep "#0" 4.td | sort | uniq -c
S     1 #0  0x00007f51ce80cd3a in
R     1 #0  0x00007f51ce81dba3 in
S     7 #0  0x00007f51ce87ccd3 in epoll_wait ()
S    17 #0  0x00007f51ceb3243c in pthread_cond_wait()
S     1 #0  0x00007f51ceb35a2d in
S     1 #0  0x00007f51ceb362a5 in sigwait ()
R     7 #0
```

x12 throughput

250 handshakes/s → 3000 handshakes/s

sed “s/Thread/entity/g”



high sys time

```
~ $ sudo strace -f -c -p $(pgrep process) | magic
```

% time	seconds	usecs/call	calls	errors	syscall
48.18	0.005732	2	2560	2560	connect
15.03	0.001788	0	5147		close
10.02	0.001192	0	7106		epoll_wait
5.02	0.000597	0	2575		write
4.95	0.000589	0	13447		epoll_ctl

high sys time

```
~ $ sudo strace -f -c -p $(pgrep process) | magic
% time seconds    usecs/call calls  errors syscall
48.18  0.005732  2        2560   2560   connect
15.03  0.001788  0        5147          close
10.02  0.001192  0        7106          epoll_wait
5.02   0.000597  0        2575          write
4.95   0.000589  0        13447         epoll_ctl
```

high sys time

```
~ $ sudo strace -f -c -p $(pgrep process) | magic
```

% time	seconds	usecs/call	calls	errors	syscall
48.18	0.005732	2	2560	2560	connect
15.03	0.001788	0	5147		close
10.02	0.001192	0	7106		epoll_wait
5.02	0.000597	0	2575		write
4.95	0.000589	0	13447		epoll_ctl

socket statistics

```
schizophrenia@target:~$ ss -s

TCP:   3 (estab 67, closed 123, orphaned
0, synrecv 45, timewait 60143/0), ports 0
```

socket statistics

```
schizophrenia@target:~$ ss -s

TCP:   3 (estab 67, closed 123, orphaned
0, synrecv 45, timewait 60143/0), ports 0
```

socket statistics

```
schizophrenia@target:~$ ss -s

TCP:   3 (estab 67, closed 123, orphaned
0, synrecv 45, timewait 60143/0), ports 0
```

```
net.ipv4.tcp_tw_recycle = 0
  net.ipv4.tcp_tw_reuse = 0
```

Заключение

1. Управляйте ресурсами (USE Method)

Заключение

1. Управляйте ресурсами (USE Method)

2. Ищите узкие места (TSA Method)

Заключение

1. Управляйте ресурсами (USE Method)
2. Ищите узкие места (TSA Method)
3. Комбинируйте Workload & Resource

Заключение

1. Управляйте ресурсами (USE Method)
2. Ищите узкие места (TSA Method)
3. Комбинируйте Workload & Resource
4. Benchmark / Performance / Load testing

Спасибо! Вопросы?

Епихин Михаил

инженер по тестированию производительности

@SCH1Z0PHR3N1A

schizophrenia@yandex-team.ru