# Performance Profiling Tools Tutorial

CS-255 Systems Programming Lab

Giannis Malliotakis – <u>jmal@csd.uoc.gr</u> Giorgos Xanthakis – gxanth@csd.uoc.gr

### Motivation

#### Let's say you've written a program

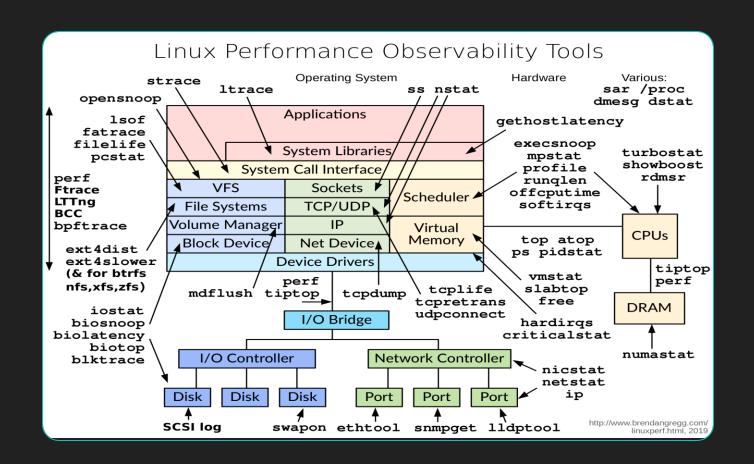
- It looks to be working
- O But the performance is bad
- You want to find the underlying cause efficiently
- O How do you do it?

#### Another example:

- Your server/machine is underperforming
- You want to check resource usage and running programs
- O How do you do it? (in Linux)

## A roadmap of available tools

OMany tools available, for different system components

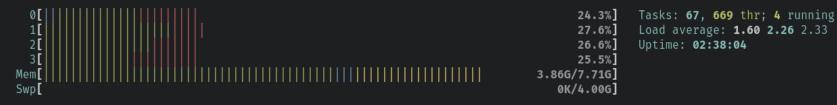


#### top

- Display Linux processes
- Continuous output, until pressing 'q'
- Lots of functionality, press 'h' to display help

top - 11:29:08 up 90 days, 38 min, 1 user, load average: 0.03, 0.08, 0.04
Tasks: 135 total, 1 running, 134 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.0 us, 0.2 sy, 0.0 ni, 99.5 id, 0.3 wa, 0.0 hi, 0.0 si, 0.0 st
KiB Mem: 4033100 total, 2060884 used, 1972216 free, 271464 buffers
KiB Swap: 3905532 total, 2680 used, 3902852 free. 1224392 cached Mem

PID	USER	PR	NI	VIRT	RES	SHR	S		%MEM	TIME+	COMMAND
22628	jmal	20	0	5388	2464		R	0.3	0.1	0:00.02	top
	root	20	0	5788	3780	2652	S	0.0	0.1	1:11.14	systemd
2	root	20	0	0	0	0	S	0.0	0.0		kthreadd
	root	20	0	0	0	0	S	0.0	0.0		ksoftirqd/0
	root	0	-20	0	0	0	s	0.0	0.0		kworker/0:0H
	root	20	0	0	0	0	S	0.0	0.0		rcu_sched
	root	20	0	0	0	0	S	0.0	0.0	0:00.00	
9	root	rt	0	Ø	0	0	s	0.0	0.0		migration/0
	root	rt	0	Ø	0	0	S	0.0	0.0		watchdog/0
	root	rt	0	0	0	0	s	0.0	0.0		watchdog/1
	root	rt	0	0	0	0	s	0.0	0.0		migration/1
	root	20	0	Ø	0	0	S	0.0	0.0		ksoftirqd/1
	root	0	-20	0	0	0	S	0.0	0.0		kworker/1:0H
	root	rt	0	0	0	0	s	0.0	0.0		watchdog/2
	root	rt	0	Ø	0	0	S	0.0	0.0		migration/2
	root	20	0	0	0	0	S	0.0	0.0		ksoftirqd/2
	root	0	-20	0	0	0	s	0.0	0.0		kworker/2:0H
	root	rt	0	0	0	0	s	0.0	0.0		watchdog/3
	root	rt	0	0	0	0	s	0.0	0.0		migration/3
	root	20	0	0	0	0	s	0.0	0.0		ksoftirqd/3
	root	0	-20	0	0	0	S	0.0	0.0		kworker/3:0H
	root	0	-20	0	0	0	s	0.0	0.0		khelper
	root	20	0	0	0	0	s	0.0	0.0		kdevtmpfs
	root	0	-20	0	0	0	s	0.0	0.0	0:00.00	netns
	root	20	0	0	0	0	s	0.0	0.0		khungtaskd
	root	0	-20	0	0	0	s	0.0	0.0		writeback
	root	25	5	0	0	0	S	0.0	0.0	0:00.00	
	root	39	19	0	0	0	s	0.0	0.0		khugepaged
	root		-20	0	0	0	S	0.0	0.0	0:00.00	
	root	0	-20	0	0	0	S	0.0	0.0		kintegrityd
	root		-20	0	0	0	S	0.0	0.0	0:00.00	
	root			0	0	0	S	0.0	0.0		kblockd
	root	20	0	0	0	0	S	0.0	0.0	1:22.81	kswapd0
	root	0	-20	0	0	0	S	0.0	0.0	0:00.00	
	root	20	0	0	0	0	S	0.0	0.0		fsnotify_mark
	root		-20	0	0	0	S	0.0	0.0		kthrotld
	root	0	-20	0	Ø	0	s	0.0	0.0		ipv6_addrconf
	root		-20	0	Ø	Ø	S	0.0	0.0		deferwq
	root		-20	0	0	Ø	s	0.0	0.0		acpi_thermal_pm
93	root	20	0	0	Ø	Ø	s	0.0	0.0	0:00.00	khubd
	root	0	-20	0	0	Ø	s	0.0	0.0	0:00.00	kpsmoused
	root	0	-20	0	Ø	Ø	s	0.0	0.0		ata_sff
	root	20	0	0	0	Ø	s	0.0	0.0	0:00.00	scsi_eh_0



PID USER	PRI	NI VIRT	RES	SHR S	CPU%	MEM%▽	TIME+	Command	
5916 mallias	20	0 3424M	761M	211M S	4.6	9.6	12:04.36	usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	parentBuildID 202104:
5920 mallias	20	0 3424M	761M	211M S	1.3	9.6	1:02.40	usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	parentBuildID 202104:
5921 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5922 mallias	20							usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5923 mallias	20							usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5924 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5925 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5926 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5927 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5929 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5930 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5931 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5932 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5933 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
5934 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6044 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6050 mallias								usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6051 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6052 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6060 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6086 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6087 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6088 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6108 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6125 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6126 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6127 mallias								/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	
6128 mallias	39	19 3424M						/usr/lib/firefox/firefox -contentproc -childID 18 -isForBrowser -prefsLen 10391 -prefMapSize 252953 -	

#### htop

- O An interactive Linux process viewer
- O Similar to top, but with better UI, some extra features (mouse support!)

```
jmal@thanatos ~ % vmstat 1
     procs
                  buff cache
                              si
                                       bi
                                             bo
                                               in
                                                     cs us sy id wa st
   b
      swpd
            free
                                  S0
      7156 1514700 447452 1673396
                                     0
                                                    0
      7156 1514568 447452 1673396
                                                   40
                                                       81
                                          0
      7156 1514568 447452 1673396
                                                   36
                                                       84
                                                               100
   0
      7156 1514784 447460 1673388
                                     0
                                               16
                                                  187
                                                      532
   0
      7156 1514784 447460 1673388
                                     0
                                                   27
                                                       63
                                                               100
   0
      7156 1514520 447460 1673396
                                                   30
                                                       78
                                                          0
                                                               100
   0
                                     0
                                          0
                                                             0
```

#### vmstat

- Virtual Memory Stat
- View system memory and processor statistics
- Configurable output frequency, count

```
jmal@thanatos ~ % free -m
                                      free
                                               shared
                                                         buffers
                                                                      cached
             total
                          used
Mem:
              3938
                          2462
                                     1476
                                                   87
                                                              437
                                                                        1635
-/+ buffers/cache:
                           388
                                     3550
Swap:
                                     3807
              3813
                             6
jmal@thanatos ~ % ■
```

#### free

- Display system memory statistics
- Configurable output unit, report duration

# mpstat

- Processor statistics
- Various processor time metrics
- Configurable CPU set to display stats

mallias@Arche					l 05/18/2021	_;	k86_64_	(	4 CPU)		
01:08:43 PM	CPU	%usr	%nice	%sys	%iowait	%irq	%soft	%steal	%guest	%gnice	%idle
01:08:44 PM	all	7.97	0.00	1.80	0.00	0.77	0.00	0.00	0.00	0.00	89.46
01:08:44 PM	0	4.17	0.00	1.04	0.00	1.04	0.00	0.00	0.00	0.00	93.75
01:08:44 PM	1	5.21	0.00	3.12	0.00	0.00	0.00	0.00	0.00	0.00	91.67
01:08:44 PM	2	15.31	0.00	1.02	0.00	1.02	0.00	0.00	0.00	0.00	82.65
01:08:44 PM		7.07	0.00	2.02	0.00	1.01	0.00	0.00	0.00	0.00	89.90
Average:	CPU	%usr	%nice	%sys	%iowait	%irq	%soft	%steal	%guest	%gnice	%idle
Average:	all	7.97	0.00	1.80	0.00	0.77	0.00	0.00	0.00	0.00	89.46
Average:	0	4.17	0.00	1.04	0.00	1.04	0.00	0.00	0.00	0.00	93.75
Average:	1	5.21	0.00	3.12	0.00	0.00	0.00	0.00	0.00	0.00	91.67
Average:	2	15.31	0.00	1.02	0.00	1.02	0.00	0.00	0.00	0.00	82.65
Average:		7.07	0.00	2.02	0.00	1.01	0.00	0.00	0.00	0.00	89.90
mallias@Arch	Corsai	ir ~ %									

#### iostat

- I/O related statistics
- Displayed per device
- Configurable report frequency/duration

	~ % iostat -> 0-514.16.1.el		4 (sith5.	cluster.	ics.forth.	.gr) 05	5/18/2021	_x86	_64_	(32	CPU)			
	Guser %nice 0.35 0.00	%system 0.04	%iowait 0.01	%steal 0.00	%idle 99.60									
Device: nvme0n1 sda sdb	rrqm/s 0.00 0.00 0.00	wrqm/s 0.00 0.01 0.00	r/s 0.00 0.24 1.27	w/s 0.00 0.21 0.01	rMB/s 0.00 0.00 0.16	wMB/s 0.00 0.00 0.00	avgrq-sz 77.96 27.08 252.28	avgqu-sz 0.00 0.01 0.01	await 0.10 11.22 4.76	r_await 0.10 7.09 4.35	w_await 0.00 16.08 66.81	svctm 0.10 6.35 2.22	%util 0.00 0.28 0.28	
	Guser %nice 0.00 0.00	%system 0.00	%iowait 0.00	%steal 0.00	%idle 100.00									
Device: nvme0n1 sda sdb	rrqm/s 0.00 0.00 0.00	wrqm/s 0.00 0.00 0.00	r/s 0.00 0.00 0.00	w/s 0.00 0.00 0.00	rMB/s 0.00 0.00 0.00	wMB/s 0.00 0.00 0.00	avgrq-sz 0.00 0.00 0.00	avgqu-sz 0.00 0.00 0.00	await 0.00 0.00 0.00	r_await 0.00 0.00 0.00	w_await 0.00 0.00 0.00	svctm 0.00 0.00 0.00	%util 0.00 0.00 0.00	

## **Profiling**

- Program = Algorithms + Data Structures
- Both can be debugged to avoid errors
- O But how to optimize efficiently?
- Profiling: dynamic program performance analysis
- O Program analysis can examine many things:
  - 1. Time spent in functions (stalls/delays)
  - 2. Call paths and function call frequencies
  - 3. Memory Consumption

## gprof

- GNU Profiler (Not a debugger!)
- Compile (and link) your program with -pg
- Run your program as you would normally
- Once your program exits there should be a *gmon.out* file
- O Run gprof <gpre>gprof options> <your executable>gmon.out to perform profiling
- Common gprof options:
  - 1. -p: flat profile, shows the time your program spent executing each function
  - 2. -q: call graph analysis, view function calls in a tree-like manner

## gprof output

mallias@ArchCorsair ~/Documents/HY255/EX04 % gprof -p -b sudoku gmon.out Flat profile:											
reac pr	reac province.										
Each sample counts as 0.01 seconds.											
% с	umulative	self		self	total						
time	seconds	seconds	calls	ms/call	ms/call	name					
66.73	0.18	0.18	4841	0.04	0.05	sudoku_init_choices					
7.41	0.20	0.02	2087055	0.00	0.00	grid_clear_choice					
7.41	0.22	0.02	456892	0.00	0.00	grid_choice_is_valid					
7.41	0.24	0.02	4841	0.00	0.00	sudoku_try_next					
7.41	0.26	0.02	4780	0.00	0.00	sudoku_is_solvable					
3.71	0.27	0.01	4842	0.00	0.00	sudoku_is_correct					
0.00	0.27	0.00	1647725	0.00	0.00	grid_read_choice					
0.00	0.27		1209593	0.00	0.00	grid_read_value					
0.00	0.27	0.00	742673		0.00	grid_read_count					
0.00	0.27	0.00	229593	0.00	0.00	grid_clear_count					
0.00	0.27	0.00	123310	0.00	0.00	grid_remove_choice					
0.00	0.27	0.00	4471	0.00	0.00	grid_update_value					
0.00	0.27	0.00	4390		0.00	sudoku_eliminate_choice					
0.00	0.27	0.00	4390	0.00	0.00	sudoku_update_choice					
0.00	0.27	0.00	729	0.00	0.00	grid_set_choice					
0.00	0.27	0.00	454	0.00	0.00	grid_clear_unique					
0.00	0.27	0.00	81	0.00	0.00	grid_set_count					
0.00	0.27	0.00	2	0.00	0.00	sudoku_print					
0.00	0.27	0.00	1	0.00	0.00	grid_read_unique					
0.00	0.27	0.00	1	0.00	0.00	grid_set_unique					
0.00	0.27	0.00	1	0.00	0.00	sudoku_read					
0.00	0.27	0.00	1	0.00	0.00	sudoku_solution_is_unique					
0.00	0.27	0.00	1	0.00	270.27	sudoku_solve					

mallias@ArchCorsair ~/Documents/HY255/EX04 % gprof -b -q sudoku gmon.out Call graph

granularity: each sample hit covers 2 byte(s) for 3.70% of 0.27 seconds

index	% time	self	childre	en called	name <spontaneous></spontaneous>
[1]	100.0	0.00	0.27		main [1]
[1]	100.0	0.00	0.27	1/1	sudoku solve [2]
		0.00	0.00	1/4842	sudoku is correct [8]
		0.00	0.00		sudoku_rint [20]
		0.00	0.00		sudoku read [23]
		0.00	0.00		sudoku_solution_is_unique [24]
				4840	sudoku_solve [2]
		0.00	0.27	1/1	main [1]
[2]	100.0	0.00	0.27		sudoku_solve [2]
		0.18	0.04		sudoku_init_choices [3]
		0.02	0.00	4841/4841	sudoku_try_next [6]
		0.02	0.00		sudoku_is_solvable [7]
		0.01	0.00		sudoku_is_correct [8]
		0.00	0.00	4390/4390	sudoku_update_choice [9]
		0.00	0.00		sudoku_eliminate_choice [11]
		0.00	0.00	4780/123310	grid_remove_choice [10]
		0.00	0.00	454/454	grid_clear_unique [18]
				4840	sudoku_solve [2]
		0.18	0.04	4841/4841	 sudoku_solve [2]
[3]	81.3	0.18	0.04	4841	sudoku_init_choices [3]
		0.02	0.00	456892/456892	grid_choice_is_valid [5]
		0.02	0.00	2029123/2087055	grid_clear_choice [4]
		0.00	0.00	1502262/1647725	grid_read_choice [12]
		0.00	0.00	392121/1209593	grid_read_value [13]
		0.00	0.00	225203/229593	grid_clear_count [15]
		0.00		18422/2087055	
		0.00	0.00		
		0.02		2029123/2087055	
[4]	7.4	0.02 	0.00	2087055 	grid_clear_choice [4] 
		0.02	0.00		sudoku_init_choices [3]
[5]	7.4	0.02	0.00	456892	grid_choice_is_valid [5]
		0.02	0.00	4841/4841	 sudoku_solve [2]
[6]	7.4	0.02	0.00		sudoku_try_next [6]
		0.00	0.00	662645/742673	grid_read_count [14]

## Going a level deeper: perf

- o perf is a **robust** Linux profiler
- Can monitor lots of system events (use perf list to check them out)
- Configurable monitoring frequency
- Useful mainly for the Linux kernel (perf\_events), but can also be used in userspace.
- O Userspace command interface, use perf <perf\_command> <perf\_options> <your command>

## Some useful perf commands

- o perf stat: obtain event counts
- o perf record: record events for later reporting
- O perf report: event breakdown
- O perf top: live event count

#### Sample perf stat output

```
mallias@ArchCorsair ~/Documents/HY255/EX04 % perf stat ./sudoku < puzzles/hard.txt
800000000
003600000
070090200
050007000
000045700
000100030
001000068
008500010
090000400
8 1 2 7 5 3 6 4 9
9 4 3 6 8 2 1 7 5
6 7 5 4 9 1 2 8 3
1 5 4 2 3 7 8 9 6
287169534
5 2 1 9 7 4 3 6 8
4 3 8 5 2 6 9 1 7
7 9 6 3 1 8 4 5 2
Has a unique solution: 0
 Performance counter stats for './sudoku':
                                                0.998 CPUs utilized
          630.81 msec task-clock:u
                     context-switches:u
                                            # 0.000 K/sec
                     cpu-migrations:u
                                            # 0.000 K/sec
                     page-faults:u
             245
                                            # 0.388 K/sec
                     cycles:u
                                            # 3.552 GHz
    2,240,681,990
    1,854,715,553
                                                0.83 insn per cycle
                     instructions:u
      321,735,923
                                            # 510.034 M/sec
                     branches:u
       2,293,988
                                                0.71% of all branches
                     branch-misses:u
      0.631850229 seconds time elapsed
      0.630962000 seconds user
      0.000000000 seconds sys
```

#### perf record and report

```
mallias@ArchCorsair ~/Documents/HY255/EX04 % perf record -F 99 -g ./sudoku < puzzles/hard.txt
800000000
003600000
070090200
050007000
000045700
000100030
001000068
008500010
090000400
8 1 2 7 5 3 6 4 9
9 4 3 6 8 2 1 7 5
6 7 5 4 9 1 2 8 3
1 5 4 2 3 7 8 9 6
3 6 9 8 4 5 7 2 1
287169534
5 2 1 9 7 4 3 6 8
4 3 8 5 2 6 9 1 7
7 9 6 3 1 8 4 5 2
[ perf record: Woken up 1 times to write data ]
```

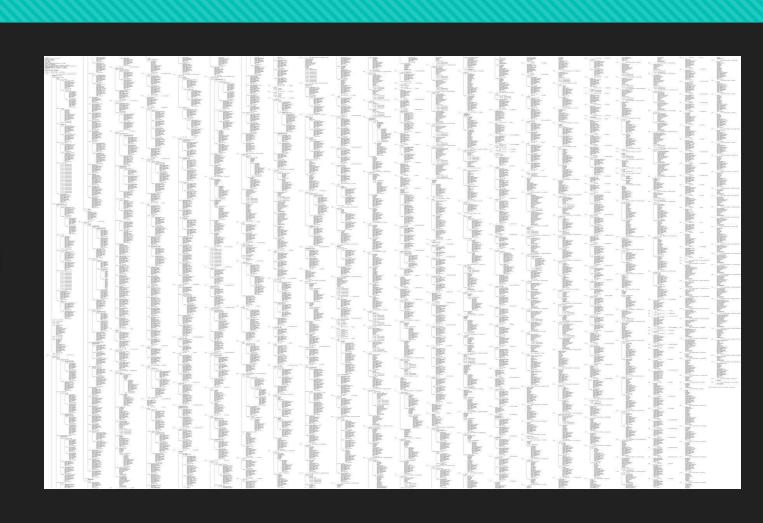
[ perf record: Captured and wrote 0.029 MB perf.data (97 samples) ]

					(approx.): 3931367321
	Children			Shared Object	
t	97.49%			libc-2.33.so	[.]libc_start_main
+	97.49%	0.00%		sudoku	[.] main
	97.49%	0.91%	sudoku	sudoku	[.] sudoku_solve
	- 96.58% s		lve		
		ku_solve			
	sudol	ku_solve			
		ku_solve			
	sudol	ku_solve			
	sudol	ku_solve			
	sudol	ku_solve			
	sudol	ku_solve			
	sudol	ku_solve			
	sudol	ku_solve			
		ku_solve			
	+ 0.91%	_libc_sta	rt_main		
t	62.15%	47.25%		sudoku	[.] sudoku_init_choices
t	15.72%	12.06%		sudoku	[.] sudoku_try_next
t	13.17%	13.17%		sudoku	[.] sudoku_is_solvable
t		9.34%		libc-2.33.so	[.]mcount_internal
t	6.40%	6.40%	sudoku	sudoku	<pre>[.] grid_choice_is_valid</pre>
+	3.69%			sudoku	[.] sudoku_eliminate_choice
t	2.81%	1.88%	sudoku	sudoku	[.] grid_clear_choice
t	2.78%			sudoku	[.] grid_remove_choice
t	1.84%			sudoku	[.] sudoku_is_correct
+	1.58%	1.58%		libc-2.33.so	[.] _int_malloc
t	0.93%			sudoku	[.] grid_read_count
+	0.93%			libc-2.33.so	[.] malloc
+	0.92%			libc-2.33.so	[.] _mcount
	0.01%	0.01%		[unknown]	[k] 0xffffffffab601067
	0.01%	0.00%	sudoku	ld-2.33.so	[.] _start

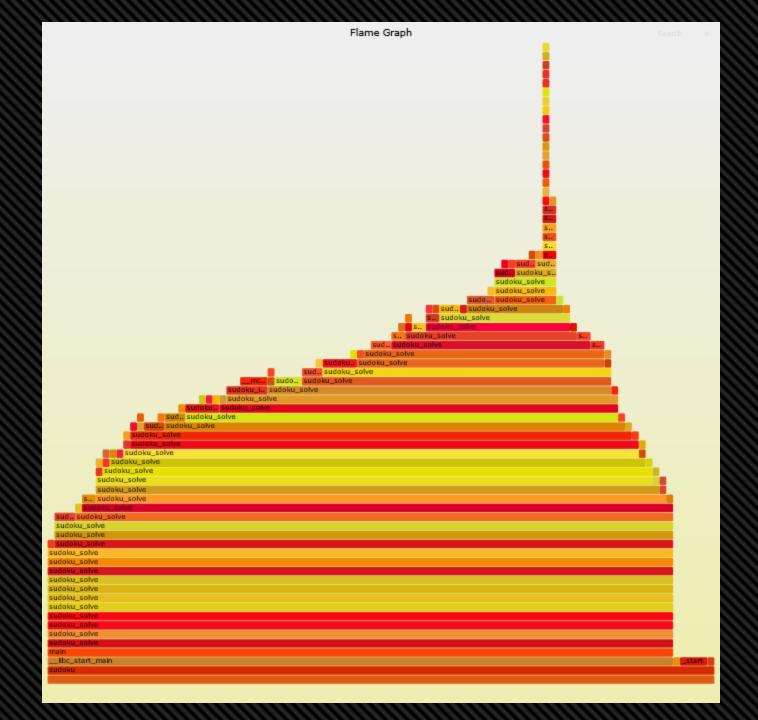
Tip: For memory address profiling, try: perf mem record / perf mem report

## Flame Graphs

- Perf output can be very long (especially for large programs)
- This is where flame graphs come in
- Visualize function call stacks (and other data) efficiently
- Interactive svg output



Sudoku solve sample CPU flame graph



### References

- Check the man pages for tool details/options (top, htop, vmstat, free, iostat, mpstat, perf)
- O Performance analysis in 60 seconds: <a href="https://netflixtechblog.com/linux-performance-analysis-in-60-000-milliseconds-accc10403c55">https://netflixtechblog.com/linux-performance-analysis-in-60-000-milliseconds-accc10403c55</a>
- gprof documentation: <a href="https://sourceware.org/binutils/docs/gprof/">https://sourceware.org/binutils/docs/gprof/</a>
- o gprof tutorial: <a href="https://linoxide.com/gprof-performance-analysis-programs/">https://linoxide.com/gprof-performance-analysis-programs/</a>
- o perf wiki: <a href="https://perf.wiki.kernel.org/index.php/Main\_Page">https://perf.wiki.kernel.org/index.php/Main\_Page</a>
- O Brendan Gregg's perf examples: <a href="http://brendangregg.com/perf.html">http://brendangregg.com/perf.html</a>
- Flame Graphs: <a href="http://brendangregg.com/flamegraphs.html">http://brendangregg.com/flamegraphs.html</a>