PROFILING AND DEBUGGING Erlang Systems

Roberto Aloi - Martin Kjellin









M.Sc in Computer Engineering Working with Erlang technologies since 2007 Senior Consultant and Trainer at Erlang Solutions Sicilian of origin, based in Stockholm

M.Sc in Computer Science, almost Working with Erlang Technologies since 2005 Developer at Klarna



Contents

IntroMindset & what to look forToolsUseful tools & their usageWar StoryCase study of a well-known Erlang systemConclusions & QAand maybe mild criticism?

Profiling and Debugging?

Not enough to determine where a system spends time - why it takes time needs to be answered too Shared toolset for both profiling and debugging

You find - you fix!

Profiling mindset

Measure, don't guess Trust your measurements Measure before optimising Know what you're measuring

Measure!

What are we looking for?

OS Level

- Memory Usage
- Disk IO & IO Wait
- · CPU usage
- Network utilisation

Erlang VM

- Message queue
- · Reductions
- (garbage collections, stack traces, excessive bif usage ...)

Rarely that simple... or?

How do we find it?

OTP Tools fprof, cprof, eprof, observer, dbg, trace bifs 3rd party Erlang Tools redbug, recon, dtop, erlgrind OS Tools htop, iostat, glances, iftop, sar, strace and many, many more

http://www.erlang.org/doc/efficiency_guide/profiling.html https://code.google.com/p/eper/wiki/redbug

A misbehaving system



Database write times increases Number of handled database writes decreases

Health check with Observer - load charts



Health check with Observer - Process overview

00			N	Observer -	riak@fred-test-a	app1.interna	l.machi	nes	
File View	Trace	Node	es Help						
System	Load C	harts	Applications	Processes	Table Viewer	Trace Over	rview		
Pid		Name	e or Initial Func		Reds	Memory	MsgQ	Current Function	_
<9648.36	<06.0>	proc_l	ib:init_p/5		2907832	89176	0	gen_fsm:loop/7	
<9648.40	19.0>	proc_l	ib:init_p/5		2904657	89176	0	gen_fsm:loop/7	=
<9648.52	84.0>	proc_l	ib:init_p/5		2903657	89176	0	gen_fsm:loop/7	
<9648.23	84.0>	proc_l	ib:init_p/5		2903242	89176	0	gen_fsm:loop/7	
<9648.29	84.0>	proc_l	ib:init_p/5		2902583	89240	0	gen_fsm:loop/7	
<9648.30	53.0>	proc_l	ib:init_p/5		2901069	55648	0	gen_fsm:loop/7	
<9648.47	33.0>	proc_l	ib:init_p/5		2895723	89096	0	gen_fsm:loop/7	
<9648.39	55.0>	proc_l	ib:init_p/5		2889978	89176	0	gen_fsm:loop/7	
<9648.52	<0.00	proc_l	ib:init_p/5		2883386	89176	0	gen_fsm:loop/7	
<9648.45	95.0>	proc_l	ib:init_p/5		2880107	89176	0	gen_fsm:loop/7	
<9648.48	<0.68	proc_l	ib:init_p/5		2875386	55728	0	gen_fsm:loop/7	
<9648.20	96.0>	proc_l	ib:init_p/5		2871258	89176	0	gen_fsm:loop/7	
<9648.11	35.0>	proc_l	ib:init_p/5		2870136	89176	0	gen_fsm:loop/7	
<9648.44	75.0>	proc_l	ib:init_p/5		2860145	89320	0	gen_fsm:loop/7	
<9648.40	84.0>	proc_l	ib:init_p/5		2858853	89176	0	gen_fsm:loop/7	
<9648.27	10.0>	proc_l	ib:init_p/5		2856568	89176	0	gen_fsm:loop/7	
<9648.50	75.0>	proc_l	ib:init_p/5		2856019	89176	0	gen_fsm:loop/7	
<9648.32	57.0>	proc_l	ib:init_p/5		2851645	89176	0	gen_fsm:loop/7	
<9648.21	64.0>	proc_l	ib:init_p/5		2851390	89096	0	gen_fsm:loop/7	
<9648.22	32.0>	proc_l	ib:init_p/5		2851363	89176	0	gen_fsm:loop/7	
<9648.51	43.0>	proc_l	ib:init_p/5		2850250	89320	0	gen_fsm:loop/7	
<9648.23	<01.0>	proc_l	ib:init_p/5		2847788	89176	0	gen_fsm:loop/7	*
<									
Observer -	riak@fr	ed-tes	st-app1.intern	al.machines					

Inspect info from Observer closer with redbug

(user@db-node)1> redbug:start("gen_fsm:loop").

23:54:04 <{db_core_vnode,init,1}> {gen_fsm,loop,

[<0.192.0>,<0.5143.0>,active, {state, 1370157784997721485815954530671515330927436759040, db_kv_vnode, {state, 1370157784997721485815954530671515330927436759040, db_kv_multi_backend, {state, [{<<"memory_multi">>, db_kv_memory_multi">>, db_kv_memory_backend, {state,103678989,103547916, undefined,undefined,0,undefined}},

Code inspection showed this to be expected

System health check - htop

PID USER PRI NIT VIRT RES S CPUN HEHK TIME+ Command 19565 rick 20 0 35.26 20.20 227M S 933.16.0 857h /usr/lib64/rick/erts-5.9.1/bin/beam.smp -A 64 K true -P 256000 -S 64:64 -M 18719 vis 20 35.26 20.26 227M S 35.16.0 14065:09 /usr/lib64/rick/erts-5.9.1/bin/beam.smp -A 64 K true -P 256000 -S 64:64 M 18719 vis 20 35.26 20.26 227M R 22.9 16.0 1402:33 /usr/lib64/rick/erts-5.9.1/bin/beam.smp -A 64 K true -P 256000 -S 64:64 M 18706 vis 20 35.26 20.26 227M R 19.7 16.0 1402:25 /usr/lib64/rick/erts-5.9.1/bin/beam.smp -A 64 K true -P 256000 -S 64:64 M 18708 19.7 16.0 1405:14 <	1 2 4 5 6 7 8 Mem Swp			1111	26. 30. 27. 28. 22. 28. 21. 26.		9 [10 [11 [12 [13 [14 [15 [16 [25	28 24 29 17 28 22 24 254/129013 187/281	53] 23] 33] 33] 34] 34] 34] 34] 34] 34] 34] 3	17 [18 [20 [21 [22 [23 [23 [24 [Tasks: Load a Uptime	<pre>!!!! !!!! !!!! !!!! !!!! !!!! : 59, 359 werage: : 50 day</pre>	 9 thr; : 17 01 2 ys, 20:	23.4%] 36.9%] 28.4%] 23.1%] 48.1%] 25.3%] 16.6%] 26.6%] 19 running 5.52 14.68 38:00	25 26 27 28 29 30 31 32					12221214	768] 558] 178] 588] 588] 768] 768]
10555 1105 <t< th=""><th>PID</th><th>USER</th><th>PRI</th><th>NI</th><th>VIRT</th><th>RES</th><th>SHR</th><th>S CPU</th><th>% MEM%</th><th>TIME+</th><th>Comma</th><th>nd</th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></t<>	PID	USER	PRI	NI	VIRT	RES	SHR	S CPU	% MEM%	TIME+	Comma	nd										
10710 20 8 35.25 28.26 22/1 S 46.4 16.8 14145:2 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10719 20 35.26 20.26 22/1 R 22.2 16.8 14147:08 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10711 20 35.26 20.26 22/1 R 22.2 16.8 14147:08 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10711 20 35.26 20.26 22/1 R 22.2 16.8 14147:08 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10716 20 35.26 20.26 22/1 R 19.7 16.0 14102:25 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10708 20 35.26 20.26 22/1 R 19.7 16.0 14102:25 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10704 20 35.26 20.26 22/1 R 19.7 16.0 14102:25 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10704 20 35.26 20.26 22/1 R 19.7 16.0 1315:11 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10730 20 35.26 20.26 22/1 R 19.1 16.0 1315:11 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10732 0 35.26 20.26 22/1 R 19.1 16.0 14103:14 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10732 0 35.26 20.26 22/1 R 19.1 16.0 14105:14 /usr/lib64/ridk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:6	10505	riak	20	0	35.20	20.2G	227M	S 933	. 16.0	857h	/usr/	lib64/ric	ik/erts-5	5.9.1/b	in/beam.smp	-A 64	+ -K	true	-P 25	6000	-S	54:64 -W
10/17 20 6 35.26 20.16 22/11 S 3/5.16.0 14/40:33 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 K Lue -256000 -S 64:64 -W 10711 20 0 35.26 20.26 22711 R 22.9 16.0 14/47:08 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10716 20 0 2764H 2201 8.21.8 0.2 16/33.011/opt/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10704 20 0 35.26 20.26 22711 S 19.7 16.0 14/45133 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10704 20 0 35.26 20.26 22711 S 19.7 16.0 14/4512 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 W 10733	10710		20	0	35.2G	20.20	227M	S 46.	16.0	14h45:52	/usr/	lib64/ric	k/erts-5	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-5	54:64 -W
10/19 20 6 35.26 20.26 22/11 22.9 10.0 11472130 20.3 2711 22.9 26.46 27.11 22.9 26.47 22.9 26.47 22.9 26.47 22.9 26.47 22.9 26.47 22.9 26.47 22.9 26.47 27.11	10717		20	0	35.26	20.26	227M	\$ 37.	5 16.0	14105:09	/usr/	11664/ric	k/erts-5	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-5	54:64 -W
10/11 28 8 3.25 29.26 22/11 k 22.16 1414/195 /ust/1054/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10716 108 10 25.26 29.26 22/11 k 19.7 16.0 141402:25 /ust/1064/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10706 10 35.26 29.26 22/11 k 19.7 16.0 14146:33 /ust/1064/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10708 10 35.26 29.26 22/11 k 19.7 16.0 14151:14 /ust/1064/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10718 10 35.26 29.26 22/11 k 19.7 16.0 14151:14 /ust/1064/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10718 10 35.26 29.26 22/11 k 19.1 16.0 141401:14 /ust/1064/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10722 10 35.26 29.26 22/11 k 19.1 16.0 131465:12 /ust/1064/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10722 10 35.26 29.26 22/11 k 16.0 131465:12 /ust/1064/rick/erts-5.9.1/bi	10719		20	0	35.26	20.26	227M	R 22.9	9 16.0	1902:33	/usr/	11064/110	K/erts-5	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-5	54:64 -W
10049 100 <td< th=""><th>10711</th><th></th><th>20</th><th>0</th><th>35.26</th><th>20.26</th><th>227M</th><th>R 22.</th><th>2 16.0</th><th>14147:08</th><th>/usr/</th><th>11064/110</th><th>K/erts-5</th><th>5.9.1/b</th><th>in/beam.smp</th><th>-A 64</th><th>-K</th><th>true</th><th>-P 25</th><th>6000</th><th>-5</th><th>54:64 -W</th></td<>	10711		20	0	35.26	20.26	227M	R 22.	2 16.0	14147:08	/usr/	11064/110	K/erts-5	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-5	54:64 -W
10/16 fto: 20 6 35.20 22/11 R 19.7 16.8 1410/2152 / Usr/11064/ridk/erts=5.9.1/bin/bedm.smp = A 64 = K true = P 250000 = S 64:64 = M 10738 fto: 20 8 35.20 22.71 S 19.7 16.8 1411651114 / usr/11064/ridk/erts=5.9.1/bin/bedm.smp = A 64 = K true = P 256000 = S 64:64 = M 10738 fto: 20 8 35.20 20.21 Z 22.71 R 19.7 16.8 14105114 / usr/11064/ridk/erts=5.9.1/bin/bedm.smp = A 64 = K true = P 256000 = S 64:64 = M 10718 fto: 20 8 35.20 20.21 Z 22.71 R 19.1 16.8 141061124 / usr/11064/ridk/erts=5.9.1/bin/bedm.smp = A 64 = K true = P 256000 = S 64:64 = M 10722 fto: 20 8 35.20 20.21 R 19.1 16.8 13164155 / usr/11064/ridk/erts=5.9.1/bin/bedm.smp = A 64 = K true = P 256000 = S 64:64 = M 10733 fto: 20 8 35.20 20.21 R 19.7 16.8 13164155 / usr/11064/ridk/erts=5.9.1/bin/bedm.smp = A 64 = K true = P 256000 = S 64:64 = M	10049		20	0	2764M	228M	2624	5 21.0	1 0.2	16033:01	/opt/	fred/erts	-5.9.3.3	3/bin/be	eam.smp –ĸ	true -	-A 5	r	00t /	opt/f	rea	-progna
10708 108 20 0 35.26 22.71 S 19.7 16.0 14451:14 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10704 10 35.26 20.26 22711 S 19.7 16.0 14451:14 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10718 108 20 0 35.26 20.26 22711 R 19.1 16.0 1440114 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 K true -P 256000 S 64:64 -W 10722 10 0 35.26 20.26 22711 R 19.1 16.0 1346:12 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 K true -P 256000 S 64:64 -W 10720 106 0 35.26 20.26 22711 R 13.6 134:6:13 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A <th>10716</th> <th></th> <th>20</th> <th>0</th> <th>35.26</th> <th>20.26</th> <th>227M</th> <th>R 19.</th> <th>/ 16.0</th> <th>19102:25</th> <th>/usr/</th> <th>11064/110</th> <th>k/erts-5</th> <th>5.9.1/b</th> <th>in/beam.smp</th> <th>-A 64</th> <th>-K</th> <th>true</th> <th>-P 25</th> <th>6000</th> <th>-5</th> <th>54:64 -W</th>	10716		20	0	35.26	20.26	227M	R 19.	/ 16.0	19102:25	/usr/	11064/110	k/erts-5	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-5	54:64 -W
18764 28 8 35.26 22.71 \$19.7 16.8 14751:14 /Usr/lib64/riak/erts=5.9.1/bin/beam.smp A 64 K true -P 256000 -S 64:64 -M 10738 10 35.26 28.26 22.71 R 19.7 16.0 14761:24 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -M 10718 10 35.26 28.26 22711 R 19.1 16.0 14161:24 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -M 10728 10 35.26 28.26 22711 R 19.1 16.0 13164:55 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -M 10720 10 35.26 28.26 22711 R 13.4 16.0 14164:33 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P	10708		20	0	35.26	20.26	227M	5 19.	/ 16.0	19h46:33	/usr/	11064/110	K/erts-5	5.9.1/b	in/beam.smp	-A 69	-K	true	-P 25	6000	-5	54:64 -W
18736 28 22711 R 19.1 16.0 14101:24 /usr/lib64/riak/erts-5.9.1/bin/beam.smp A 64 -K true -P 256000 -S 64:64 -W 10720 10k 20 35.26 28.26 22711 R 19.1 16.0 13164:55 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18720 10 35.26 28.26 22711 R 18.4 16.0 13154:85 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18721 10 35.26 28.26 22711 R <th>10704</th> <th></th> <th>20</th> <th>0</th> <th>35.26</th> <th>20.26</th> <th>227M</th> <th>5 19.</th> <th>/ 16.0</th> <th>19751:14</th> <th>/usr/</th> <th>11064/110</th> <th>K/erts-5</th> <th>5.9.1/b</th> <th>in/beam.smp</th> <th>-A 69</th> <th>-K</th> <th>true</th> <th>-P 25</th> <th>6000</th> <th>-5</th> <th>54:64 -W</th>	10704		20	0	35.26	20.26	227M	5 19.	/ 16.0	19751:14	/usr/	11064/110	K/erts-5	5.9.1/b	in/beam.smp	-A 69	-K	true	-P 25	6000	-5	54:64 -W
18/18 20 0 35.26 28.46 227M R 19.1 16.0 14M01124 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18/722 10 35.26 20.26 227M S 19.1 16.0 13M6:55 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18/723 10 35.26 20.26 227M R 19.1 16.0 13M6:55 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18/729 10 35.26 20.26 227M R 18.4 16.0 13M6:55 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18/721 10 35.26 20.26 227M R 17.8 16.0 13M54:05 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18/721 10 35.26 20.26 227M R 17.8 16.0 13M6:13 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W	10730		20	0	35.26	20.26	227M	R 19.	/ 16.0	13752:19	/usr/	11064/110	K/erts-5	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-5	54:64 -W
18722 18 28 8 35.26 28.26 227H R 19.1 16.0 19H08114 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256080 -S 64:4 -W 18725 19.2 20.26 227H R 18.4 16.0 14H45:3 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256080 -S 64:64 -W 18721 10.4 14H40:81 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256080 -S 64:64 -W 18727 10.8 13F	10718		20	0	35.26	20.26	227M	R 19.3	116.0	19101:24	/usr/	11064/110	K/erts-5	5.9.1/D	in/beam.smp	-A 69	-K	true	-P 25	6000	-5	54:64 -W
10733 104 20 0 35.26 20.26 22/11 R 19.1 16.0 13446:55 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10720 riak 20 0 35.26 20.26 22/11 N 18.4 16.0 13454:12 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10721 108 0 35.26 20.26 22/11 N 18.4 16.0 14145:33 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10712 108 0 35.26 20.26 22/11 N 17.2 16.0 14142:39 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10707 10 0 35.26 20.26 22/11 N 17.2 16.0 13149:27	10722		20	0	30.20	20.20	2270	5 19.	1 10.0	1900:14	/usr/	11064/110	K/erts-5	- 0 4 4	in/beam.smp	-A 04		true	-P 25	0000	-3	04:04 -W
18720 188720 28 36.26 22/11 5 16.7 16.0 16169412 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18720 rick 20 35.26 20.26 22/11 R 18.4 16.0 14h45:33 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18721 rick 20 35.26 20.26 22/11 R 17.8 16.0 14h40:01 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10707 rick 20 35.26 20.26 22/11 S 17.2 16.0 14h42:39 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10727 rick 20 35.26 20.26 22/11 R 17.2 16.0 13h58:13 /ust/lib64/riak/erts-5.9.1/bin/beam.smp	10733		20	0	35.26	20.20	2270	R 19.	1 10.0	13140:55	/usr/	11064/110	K/erts-5	5.9.1/0	in/beam.smp	-A 64	-K	true	-P 25	6000	-3	04:04 -W
10706 20 0 35.26 20.26 22/M R 10.4 10.5 14445133 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10721 10 0 35.26 20.26 22/M R 17.8 16.0 13454:05 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10712 10 0 35.26 20.26 22/M R 17.8 16.0 14440:01 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10707 10 0 35.26 20.26 22/M R 17.2 16.0 1349:27 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10705 10 35.26 20.26 22/M R 17.2 16.0 1349:58:13 /ust/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P	10720		20	0	35.26	20.20	2270	5 18.	+ 16.0	13754:12	/usr/	11064/110	K/erts-5	5.9.1/D	in/beam.smp	-A 69	-K	true	-P 25	0000	-5	04:04 -W
10721 20 0 35.26 20.26 2271 R 17.8 16.8 13154:85 /usr/lib64/riuk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10712 20 0 35.26 20.26 2271 R 17.8 16.0 14h40:01 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10707 20 0 35.26 20.26 2271 R 17.2 16.0 14h42:39 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10707 20 0 35.26 20.26 2271 R 17.2 16.0 13h9:27 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10705 20 0 35.26 20.26 2271 R 17.2 16.0 14h46:44 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true <th>10706</th> <th></th> <th>20</th> <th>0</th> <th>30.20</th> <th>20.20</th> <th>2270</th> <th>R 10.</th> <th>t 10.0</th> <th>12654.05</th> <th>/usr/</th> <th>lib64/rlc</th> <th>k/erts-5</th> <th>5.9.1/0</th> <th>in/beum.smp</th> <th>-A 64</th> <th></th> <th>true</th> <th>-P 25</th> <th>6000</th> <th>-3</th> <th>64.64 -W</th>	10706		20	0	30.20	20.20	2270	R 10.	t 10.0	12654.05	/usr/	lib64/rlc	k/erts-5	5.9.1/0	in/beum.smp	-A 64		true	-P 25	6000	-3	64.64 -W
18712 18 20 0 35.26 20.26 227M k 17.0 16.0 14H43:01 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10707 10 0 35.26 20.26 227M k 17.2 16.0 14H42:39 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10707 10 0 35.26 20.26 227M k 17.2 16.0 13h49:27 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10705 10 0 35.26 20.26 227M k 17.2 16.0 13h58:13 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10705 10 0 35.26 20.26 227M k 16.5 16.0 14h47:45 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K <th>10721</th> <th></th> <th>20</th> <th>0</th> <th>30.20</th> <th>20.20</th> <th>2270</th> <th>R 17.0</th> <th>0 10.0 0 12 0</th> <th>14640.04</th> <th>/usi/</th> <th>11004/110</th> <th>k/erus-s k/orto E</th> <th>5.9.1/0</th> <th>in/beum.smp</th> <th>-A 04</th> <th></th> <th>true</th> <th>-P 25</th> <th>2000</th> <th>-5</th> <th>04:04 -W 64.64 -W</th>	10721		20	0	30.20	20.20	2270	R 17.0	0 10.0 0 12 0	14640.04	/usi/	11004/110	k/erus-s k/orto E	5.9.1/0	in/beum.smp	-A 04		true	-P 25	2000	-5	04:04 -W 64.64 -W
18787 20 0 35.26 20.26 227M S 17.2 16.0 1MH2:39 /usr/lib64/riuk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18727 20 0 35.26 20.26 227M R 17.2 16.0 13h49:27 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18726 20 0 35.26 20.26 227M R 17.2 16.0 13h58:13 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18705 20 0 35.26 20.26 227M S 17.2 16.0 14h46:44 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 18709 20 0 35.26 20.26 227M R 16.5 16.0 14h47:45 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P	10712		20	0 0	30.20	20.20	2270	R 17.0	0 10.0 0 46 0	14642.20	/usi/	LIDO4/ILU	K/erts-s	5.9.1/D	n/beam.smp	-A 04		true	-P 25	2000	-5	04:04 -W 64.64 -W
18727 100 20 0 35.20 20.20 22/11 17.2 10.0 10149:27 /usr/lib64/riuk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 250000 -S 64:64 -W 10726 100 35.20 20.20 22711 R 17.2 16.0 13h58:13 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10705 10k 20 35.20 20.20 22711 S 17.2 16.0 14h46:44 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10709 10k 20 35.20 20.20 22711 S 16.5 16.0 14h47:45 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10715 10k 20 35.20 20.20 22711 S 16.0 14h40:18 /usr/lib64/riak/erts-5.9.1/bin/beam.smp -A	10707		20	0	30.20	20.20	2270	D 47 4	2 10.0	42640.27	/usi/	Lib64/ric	k/erts-s	5.9.1/D	in/beum.smp	-A 04		true	-P 29	2000	-3	04:04 -W 64.64 -W
10720 10720 <td< th=""><th>10727</th><th></th><th>20</th><th>0 0</th><th>30.20</th><th>20.20</th><th>2270</th><th>R 17.4</th><th>2 10.0</th><th>43660.43</th><th>/usi/</th><th>lib64/ric</th><th>k/erts-s</th><th>5.9.1/D</th><th>in/beam.smp</th><th>-A 04</th><th></th><th>true</th><th>-F 29</th><th>6000</th><th>-3</th><th>64.64 W</th></td<>	10727		20	0 0	30.20	20.20	2270	R 17.4	2 10.0	43660.43	/usi/	lib64/ric	k/erts-s	5.9.1/D	in/beam.smp	-A 04		true	-F 29	6000	-3	64.64 W
10705 108 20 0 35.20 20.20 227M S 17.2 10.0 14M40.44 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10709 10715 108 20 35.20 20.20 227M R 16.5 16.0 14M47:45 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10715 108 20 35.20 20.20 227M S 16.5 16.0 14M49:32 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10732 108 20 35.20 20.20 227M R 16.5 16.0 13M48:02 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10713 108 20 35.20 20.20 227M S 15.9 16.0 13M56:23 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A	10720		20	0 0	30.20	20.20	227H	C 17	2 10.0	14646.44	/usi/	lib64/ric	k/erts-s	5 0 1 /b	in/beam.smp	-A 01		true	-F 29	6000	-3	64.64 h
10709 20 0 35.20 20.20 227M K 10.5 10.0 100.7 1	10700		20	0 Ø	35.20	20.20	227H	D 16	5 16 0	14h47.4E	Juer/	lib64/ric	n/crts-1	5 Q 1/b	in/beam.smp	-A 04		true	_P 25	6000	-9	64•64 b
10713 20 0 35.26 20.26 227M R 16.0 13h48:02 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10732 riak 20 0 35.26 20.26 227M S 16.0 13h48:02 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10713 riak 20 0 35.26 20.26 227M S 15.9 16.0 14h40:18 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10728 riak 20 0 35.26 20.26 227M S 15.9 16.0 13h56:23 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10728 riak 20 0 35.26 22.7M S 15.9 16.0 13h56:23 /usr/lib64/riak/erts=5.9.1/bin/beam.smp -A 64 -K </th <th>10715</th> <th></th> <th>20</th> <th>о р</th> <th>35.20</th> <th>20.20</th> <th>227M</th> <th>S 16</th> <th>5 16 0</th> <th>1469.32</th> <th>/usr/</th> <th>lih64/ric</th> <th>ik/erts_P</th> <th>5.9.1/b</th> <th>in/heam_smp</th> <th>-4 64</th> <th></th> <th>true</th> <th>_P 25</th> <th>6000</th> <th>_9</th> <th>64:64 _k</th>	10715		20	о р	35.20	20.20	227M	S 16	5 16 0	1469.32	/usr/	lih64/ric	ik/erts_P	5.9.1/b	in/heam_smp	-4 64		true	_P 25	6000	_9	64:64 _k
10713 rick 20 0 35.2G 20.2G 227M S 15.9 16.0 14h 40:18 /usr/lib64/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W 10728 rick 20 0 35.2G 20.2G 227M S 15.9 16.0 13h 56:23 /usr/lib64/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W	10732		20	Й	35 26	28 20	227M	R 16	5 16 0	13648:02	/usr/	lih64/ric	k/erts_F	5.9.1/b	in/heam.smp	-A 64	-K	true	_P 25	6000	-5	64:64 -4
10728 rigk 20 0 35.26 20.26 227M S 15.9 16.0 13h56:23 /usr/lib64/rigk/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W	10713		20	Й	35.26	28.26	227M	S 15	9 16.0	14h49:18	/usr/	lih64/ric	k/erts_F	5.9.1/b	in/heam.smp	-A 64	-K	true	_P 25	6999	-5	64:64 -
	10728		20	Й	35.26	28.20	227M	S 15.9	9 16.0	13156:23	/usr/	lib64/ric	k/erts_F	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-5	64:64 -
10723 rick 20 0 35.26 20.26 227M R 15.2 16.0 13h52:17 /usr/lib64/rick/erts-5.9.1/bin/beam.smp -A 64 -K true -P 256000 -S 64:64 -W	10723		20	Ø	35.2G	20.26	227M	R 15.	2 16.0	13h52:17	/usr/	lib64/ric	k/erts-5	5.9.1/b	in/beam.smp	-A 64	-K	true	-P 25	6000	-S	64:64 -W

OS Tools - iostat

000)	企	martin.k	cjellin —	martin.	kjellin@	fred-test	-app1:~	— ssh —	111×1	8		R M
Every 5.0)s: iost	at -x								Thu J	un 51	3:25:45	2014 🔳
Linux 2.6	5.32-431	.3.1.el6	6.x86_64	(fred-te	st-app1)	6	06/05/2014	xi	36_64_	(32	CPU)		
avg-cpu:	%user 10.20	%nice 0.02	%system 47.70	%iowait 0.01	%steal 0.00	%idle 42.07							
Device: sda sdb dm-0 dm-1 dm-2 dm-3 dm-4 dm-5		rrqm/s 0.12 0.02 0.00 0.00 0.00 0.00 0.00	wrqm/s 160.55 807.69 0.00 0.00 0.00 0.00 0.00	r/s 0.26 10.16 0.29 0.00 10.17 0.00 0.07 0.01	w/s 99.39 49.09 0.89 0.00 856.78 0.44 258.44 0.17	rsec/s 17.02 2028.57 9.45 0.01 2028.56 0.02 7.34 0.10	wsec/s 2079.49 6854.24 7.14 0.00 6854.24 3.51 2067.48 1.37	avgrq-sz 21.04 149.92 14.02 8.00 10.25 8.00 8.03 8.00	avgqu-sz 0.01 0.04 0.00 0.00 0.66 0.00 0.01 0.00	await 0.06 0.64 0.96 0.64 0.75 0.69 0.06 0.81	svctm 0.06 0.15 0.32 0.53 0.01 0.52 0.02 0.03	%util 0.57 0.90 0.04 0.00 0.90 0.02 0.50 0.00	

Average values since last invocation (or system start) On linux, use -x and keep an eye on await (i/o op latency in ms) Use watch: watch -n 5 iostat -x

Back to htop

1 2 4 5 6 7 8 Mem Swp		 		26. 38. 27. 28. 22. 28. 21. 26. 11111		9 [10 [11 [12 [13 [14 [15 [16 [25	28 24 29 17 28 22 24 254/12901 187/281	.5%] .2%] .2%] .3%] .5%] .3%] .3%] .3%] .3%] .3%] .3%]	17 [18 [19 [20 [21 [22 [23 [24 [Tasks Load Uptim	 : 59 , 3 average e: 50 a	1111 111111 159 thu 11 17 10ys, 2	2 2 2 1 2 1 2 1 2 1 2 1 5 5 2 0 :08 :00	3 4] 6 9] 7 1] 7 3] 6 1] 6 1] 14.68	25 26 27 28 29 30 31 32			 	11		17 . 68 25 . 58 24 . 78 24 . 58 15 . 88 15 . 88 15 . 88 17 . 68 18 . 48	
PID	USER	PRI	NI	VIRT	RES	SH	R S	CPU%	MEM%	TIME+	Comma	nd												
10505	riak	20	0	35.20	20.20	5 227	MS	933.	16.0	857h	/usr/	lib64/ri	ak/erts	\$-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	-\
10710		20	0	35.26	20.20	227	MS	46.4	16.0	14h45:52	/usr/	lib64/ri	ak/erts	\$-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	-W
10717		20	0	35.26	20.20	227	M S	37.5	16.0	1905:09	/usr/	11064/11	ak/erts	5-5.9.3	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-5	64:64	-W
10719		20	0	30.20	20.2	227	MR	22.9	16.0	1902:33	/usr/	11064/11	ak/erts	8-5.9.3	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-5	64:64	-W
10711		20	0	35.26	20.20	221	MR	22.2	16.0	1914/:08	/usr/	11064/ri	ak/erts	5-5.9.3	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-3	64:64	
10049		20	0	2764M	2280	1 262	4 5	21.0	0.2	10033:01	/opt/	fred/ert	8-5.9.3	5.3/01	n/beam.s	:mp -к	true -	AD]	root	/opt/	rrea	1 -pro	gna
10716		20	0	35.20	20.20	227	MR	19.7	16.0	19102:25	/usr/	11064/11	ak/erts	5-5.9.1	1/01n/be	am.smp	-A 64	-K	true	-P	255000	-5	64:64	-W
10708		20	0	35.26	20.20	227	m 5	19.7	10.0	14140:33	/usr/	11064/11	ak/erts	5-5.9.1	1/D1N/De	am.smp	-A 64	-K	true	-P	256000	-5	64:64	-W
10704		20	0	35.26	20.20	227		19.7	10.0	19751:14	/usr/	11004/11	ak/erts	5-5.9.1	1/Din/De	am.smp	-A 04		true	-P	256888	-0	64:04	-w
10730		20	0 0	30.20	20.20	227		19.7	10.0	13102:19	/usr/	11004/11	uk/erts	5-5.9.3	1/DIN/De	am.smp	-A 04		true	-P	250000	-0	24.24	-w
10710		20	0 0	30.20	20.20	227	Me	10.1	16.0	44600.14	/usi/	11004/11	ak/erts	5-9.9.3	1/bin/be	am emp	-4 04		true	-P	250000	-3	64.64	-w
10722		20	0 G	35.20	20.20	227	мр	10 1	16.0	13646.55	/usi/	lib64/ri	ak/erts	5-9.9.J	1/bin/be	am smp	- 4 01		true	-F	256000	-5	64.64	-w
10733		20	o a	35.20	20.20	227	MS	18 4	16.0	1364.12	/usi/	lib64/ri	ak/orts	5 0 4	1/bin/be	am smp	- 4 64		true	-r	256000	-5	64.64	-w
10720		20	a a	35 20	20.20	227	MD	18 4	16 0	14645.33	/usi/	lib64/ri	ak/orts	-5.9.9.1	1/hin/he	am smp	64		true	_P	256000	-0	64.64	-n _lu
10721		20	a a	35 20	20.20	227	MD	17.8	16 0	1364.05	/usi/	lih64/ri	ak/erts	-5.9.7.	1/hin/he	am smp	_4 64		true	_P	256000	-5	64.64	-n lu
10712		20	ñ	35 26	20.20	227	MR	17.8	16.0	14h40:01	/usr/	lih64/ri	ak/erts	-5.9.1	1/hin/he	am.smp	_4 64		true	_P	256000	-5	64:64	
10707		20	й	35.26	29.20	227	MS	17.2	16.0	14h42:39	/usr/	lih64/ri	ak/erts	-5.9.1	1/hin/he	am.smp	-4 64		true	_P	256000	-5	64:64	W
10727		20	й	35.26	28.20	227	MR	17.2	16.0	13649:27	/usr/	lih64/ri	ak/erts	-5.9.1	1/hin/he	am.smp	-4 64	-K	true	_P	256000	-5	64:64	W
10726		20	ø	35.2G	20.20	227	MR	17.2	16.0	13158:13	/usr/	lib64/ri	ak/erts	-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	W
10705		20	0	35.2G	20.20	227	MS	17.2	16.0	14h46:44	/usr/	lib64/ri	ak/erts	-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	- W
10709		20	0	35.26	20.20	227	MR	16.5	16.0	14h47:45	/usr/	lib64/ri	ak/erts	-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	-W
10715		20	0	35.2G	20.20	227	MS	16.5	16.0	14h09:32	/usr/	lib64/ri	ak/erts	-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	-W
10732		20	0	35.2G	28.20	227	MR	16.5	16.0	13h48:02	/usr/	lib64/ri	ak/erts	-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	
10713		20	0	35.26	20.20	227	MS	15.9	16.0	14h40:18	/usr/	lib64/ri	ak/erts	-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	-W
10728		20	0	35.2G	20.20	227	MS	15.9	16.0	13h56:23	/usr/	lib64/ri	ak/erts	\$-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	-W
10723		20	0	35.2G	20.20	227	M R	15.2	16.0	13h52:17	/usr/	lib64/ri	ak/erts	-5.9.1	1/bin/be	am.smp	-A 64	-K	true	-P	256000	-S	64:64	-W

Swap in use but large amount of free memory?

On redhat-like systems, /proc/buddyinfo holds memory fragmentation info:

Node 0, z	zone	DMA	2	1	1	1	1	0	1	0	1	1	3
Node 0, z	zone	DMA32	397	281	227	485	417	263	144	52	83	11	6
Node 0, z	zone	Normal	1043	830	485	194	69	375	265	72	18	35	21
Node 1, z	zone	Normal	2981	435	411	251	0	44	0	20	0	6	1

Use watch to monitor:

```
watch -n5 cat /proc/buddyinfo
```

Dropping page cache restored write performance to initial numbers.

Not all bottlenecks directly visible!

A few last words

Trace BIFs Powerful but complex, use redbug when debugging under load. What to profile? Profile the entire system including disk, network & third party software Which tools? It depends - but use observer Measure? Yes! and visualise your measurements Debug Read other people's code Cry Drink Coffee Laugh Profile Learn from mistakes Observer Redbug Is that a timer:sleep? eprof prof cprof flags dtop percept Scream Priority Ask that guy with the beard Bugs everywhere! Nobody Knows Why Works on my machine Pain Optimize I should have studied law instead Strace Reductions

WAR STORY

Cry Throughput My eyes hurt Everyone else is at the party 95th Percentile load Tracing htop percept2 Graphite UI sucks Stress Test microseconds kcachegrind Queues Expectations Metrics Latency WTF Measure It's dark outside erlgrind kcachegrind suspend Garbage Collection scheduler make_ref I cannot believe it



WHAT BREAKS FIRST?

ystem Load	Charts Applications Processe	s Table Viewe	r Trace Ove	erview	
Pid	Name or Initial Func	Reds	Memory	MsgQ	Current Function
<0.148.0>	lager_event	69116427	3722407424	46139	gen_event:server_notify/4
<0.3.0>	erl prim loader	0	263600	0	erl prim loader:loop/3
<0.5.0>	error logger	0	24840	0	gen event:fetch msg/5
<0.6.0>	application controller	0	17032	0	gen_server:loop/6
<0.8.0>	proc_lib:init_p/5	0	7032	0	application_master:main_loop/2
<0.9.0>	application_master:start_i	0	2760	0	application_master:loop_it/4
<0.10.0>	kernel_sup	0	1802000	0	gen_server:loop/6
<0.11.0>	гех	0	2824	0	gen_server:loop/6
<0.12.0>	global_name_server	0	2904	0	gen_server:loop/6
<0.13.0>	erlang:apply/2	0	2720	0	global:loop_the_locker/1
<0.14.0>	erlang:apply/2	0	2720	0	global:loop_the_registrar/0
<0.15.0>	inet_db	0	16752	0	gen_server:loop/6
<0.17.0>	net_sup	0	13856	0	gen_server:loop/6
<0.18.0>	erl_epmd	0	2864	0	gen_server:loop/6
<0.19.0>	auth	0	2824	0	gen_server:loop/6
<0.20.0>	net_kernel	7	9016	0	gen_server:loop/6
<0.21.0>	inet_tcp_dist:accept_loop/2	0	2792	0	prim_inet:accept0/2
<0.22.0>	net_kernel:ticker/2	2	2720	0	net_kernel:ticker_loop/2
<0.23.0>	global_group	0	2824	0	gen_server:loop/6
<0.24.0>	file_server_2	3525	26632	0	gen_server:loop/6
<0.25.0>	code_server	0	5686136	0	code_server:loop/1
<0.26.0>	standard_error_sup	0	2864	0	gen_server:loop/6
<0.27.0>	standard_error	0	2864	0	standard_error:server_loop/1



https://github.com/basho/lager

"Lager (as in the beer) is a logging framework for Erlang. Its purpose is to provide a more traditional way to perform logging in an Erlang application that plays nicely with traditional UNIX logging tools like logrotate and syslog."

- DISCLAIMER -

The problems we found were not in lager itself, but in the way lager was used





LAGER IS BASED ON GEN_EVENT

Did you know?

In the Erlang gen_event behaviour event manager and event handlers run within the same context



gen_event yes





USE Avoid synchronous calls in the handlers Use the manager as a "dispatcher" **<u>gen</u>** event Spawn new processes whenever meaningful cautiously Avoid too many handlers

OR YOU MIGHT END UP WITH...



OK, THERE IS A QUEUE IN THE EVENT MANAGER

WHY IS A BIG MESSAGE QUEUE BAD?

(Aside from the obvious reasons, such as memory consumption and having an overloaded process)

In Erlang if you send a message to a process which has a big mailbox you get punished

Did you know?

https://github.com/erlang/otp/blob/master/erts/emulator/beam/bif.c



lager_event

Lager Overload Protection

Lager 1.x

log(Event) ->
gen_event:sync_notify(lager_event, Event).

Lager Overload Protection

Lager 2.x

```
log(Event) ->
  case lager_config:get(async, false) of
    true -> gen_event:notify(lager_event, Event);
    false -> gen_event:sync_notify(lager_event, Event)
end;
```

async flag automatically toggled based on the mailbox size async messaging used until the message exceeds **async_threshold** sync messaging used after the threshold is passed async messaging reverted when size is back below **async_threshold** - **async_threshold_window** "The problem with the current behaviour is it just pushes the problem onto the rest of the app. Sure, you don't have queuing in lager's mailbox, but now the rest of the app is slowed down instead. This isn't the right tradeoff for me at all. Logging should have minimal impact on the performance of the rest of the system, but instead we're effectively blocking at every log statement."





"So, assuming you're logging at a rate faster than you can actually write those logs to disk/syslog/whatever, your choice is either: slow down the rest of the app to compensate or let the mailbox balloon and have logging slow to a crawl anyway and possibly OOM the node."

https://github.com/basho/lager/pull/113

SO WE DID SOME PROFILING.

An Interesting Behaviour

- After stopping the load generator lot of activity is visible in htop
- Activity is restricted to one single core
- dtop shows that lager_event is busy
- no i/o wait is visible in the system

- Most of the time is spent in file:write/2 -

A single file write operation takes 12 ms And we have 400.000 messages to log That's more than 1 hour to catch up



THE COST OF FILE:WRITE/2

is directly proportional to the length of the message queue of the writing process

(at least in R15B03-1)

SELECTIVE RECEIVE



A WORKAROUND (the make_ref trick)

OTP-8623 == compiler erts hipe stdlib ==

Receive statements that can only read out a newly created reference are now specially optimized so that it will execute in constant time regardless of the number of messages in the receive queue for the process.

See gen:do_call/4 for an example of a receive statement that will be optimized.

Erlang R14A (June 2010) Release Notes

THE ALTERNATIVE: A MIDDLEMAN PROCESS



- We went from 12 ms to <1 ms per write -

https://github.com/klarna/lager_middleman_backend



PROVE IT WORKS!

STRESS TEST SCENARIO

3 File handlers (debug, warning, error)
6 Concurrent workers
100 Log messages / sec
1Kb Message size
10 Minutes (interrupted)

- 1800 file write operation / second -

MESSAGE QUEUE GROWTH



Seconds

FILE WRITE TIME



Seconds







