

Profiling Erlang Programs using Percept2

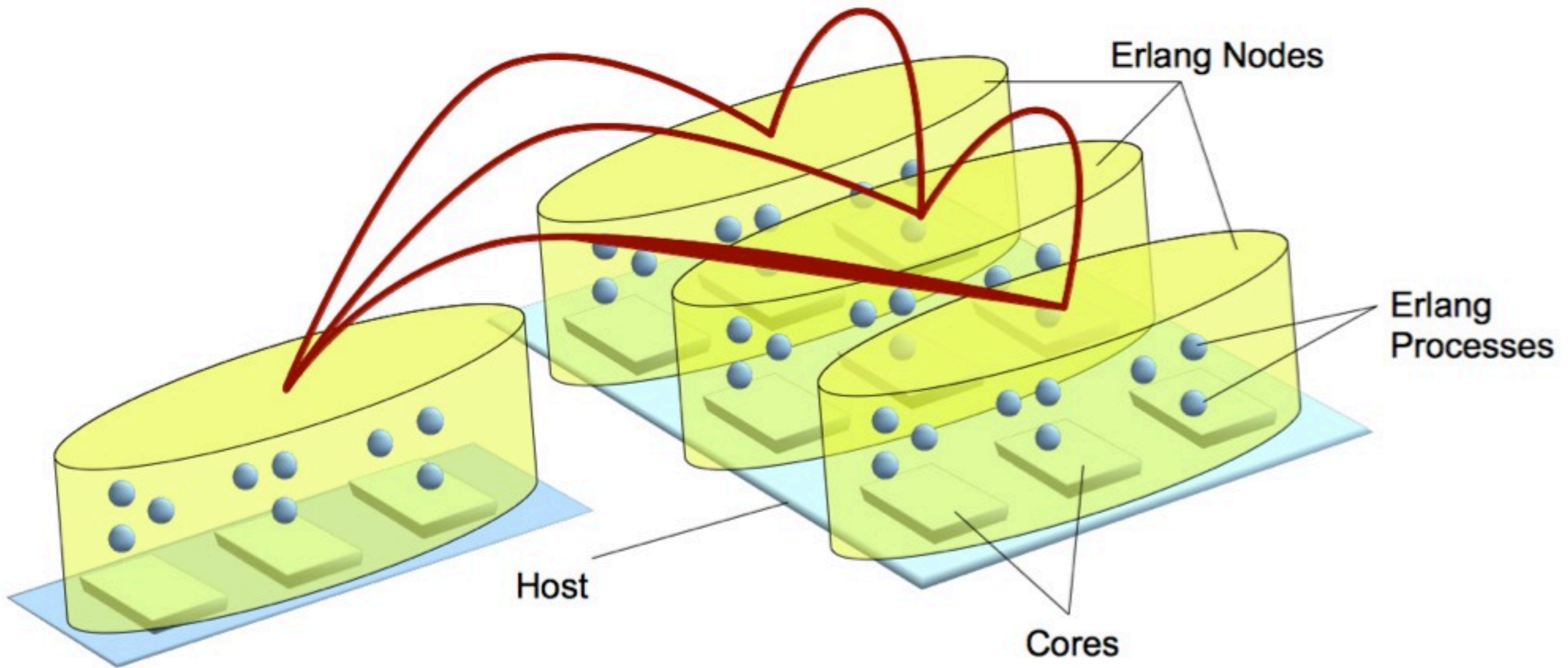
Huiqing Li and Simon Thompson
Erlang Factory Lite, London
8 November 2012



Overview

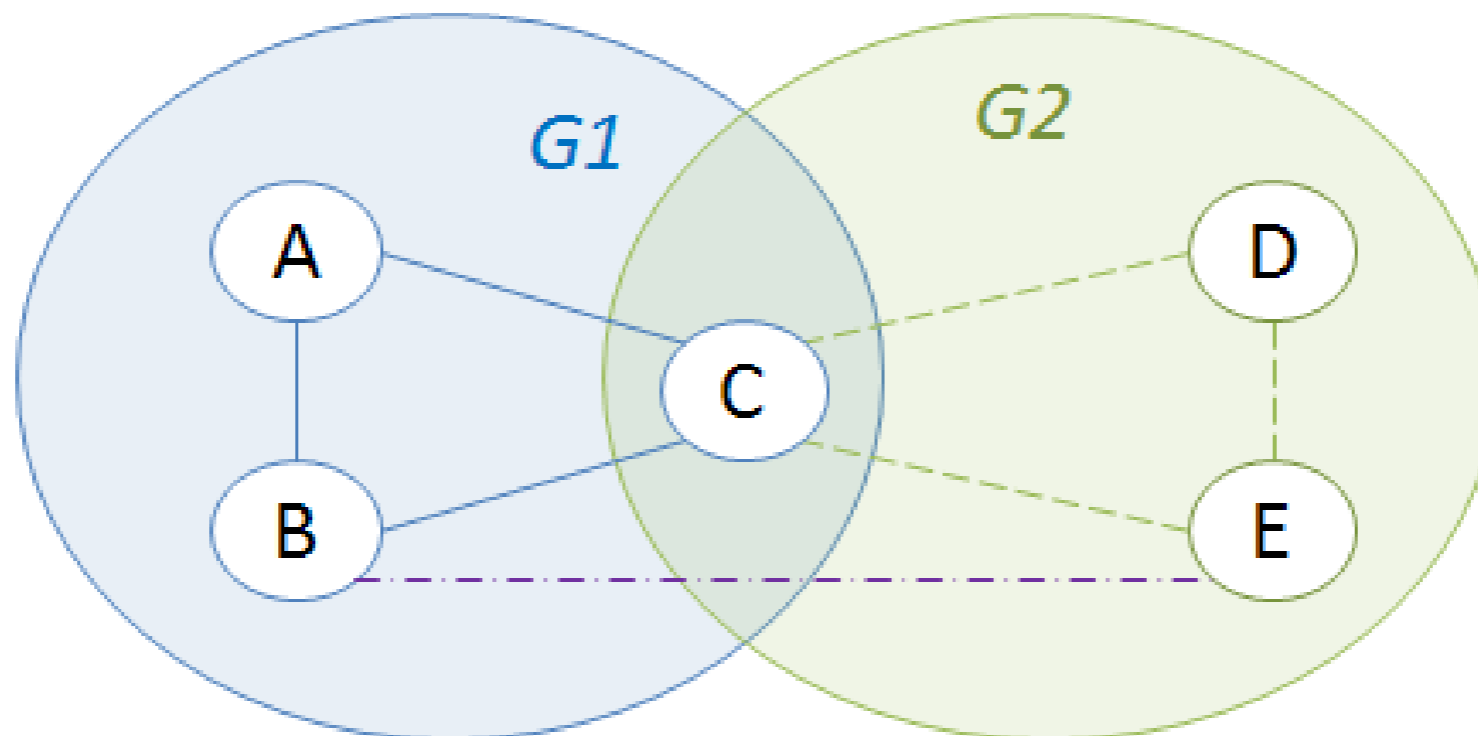
- Erlang multicore ... RELEASE
- Percept ... and Percept2
- Case study from Wrangler
- Conclusions

Erlang multicore



RELEASE

- Scalable distributed Erlang: systems and tools.
- Groups and implicit placement.



Tools

How to use
these systems
effectively?



Existing tools

- Erlang tracing
- Htop
- DTrace
- Percept

Percept

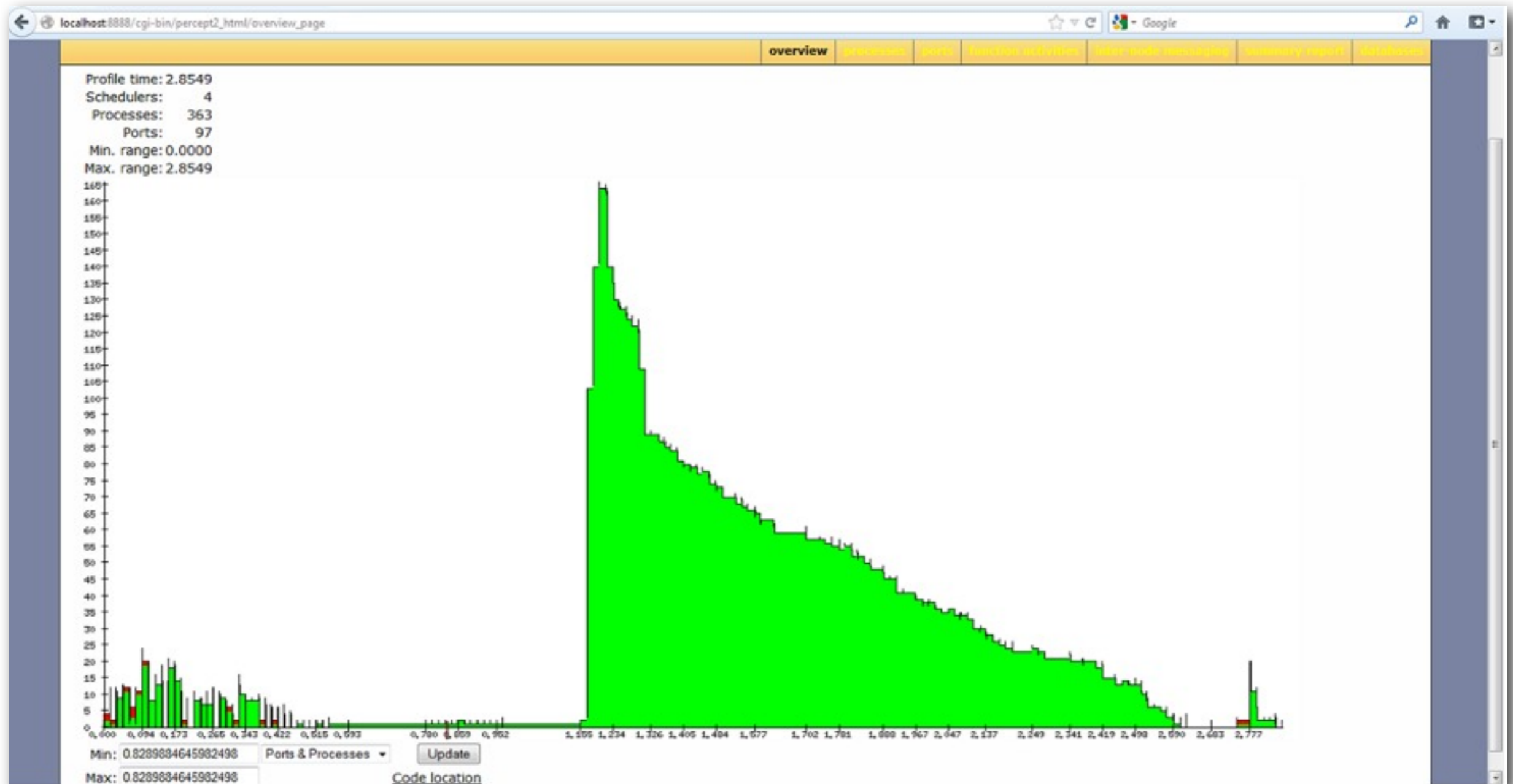
Percept

- **Percept: Erlang concurrency profiling tool.**
- In the Erlang/OTP distribution ...
- ... written by Björn-Egil Dahlberg.
- Monitors process runnability using `erlang:trace/3`, `erlang:system_profile/2`.
- Offline tool ... ‘replay’ a computation.

Percept facilities

- Histogram of active processes vs. time.
- ‘Drill down’ to process information via Pid.
- Start/end time, parent/child processes, time spent waiting for messages,
- View individual process runnability.

Percept



Percept to Percept2

- Concurrency, but multicore is 'invisible'.
 - Run queue migration,
- Analysis
 - Runnable vs running.
- Presentation of information
 - Clarify / colour.

Percept to Percept2

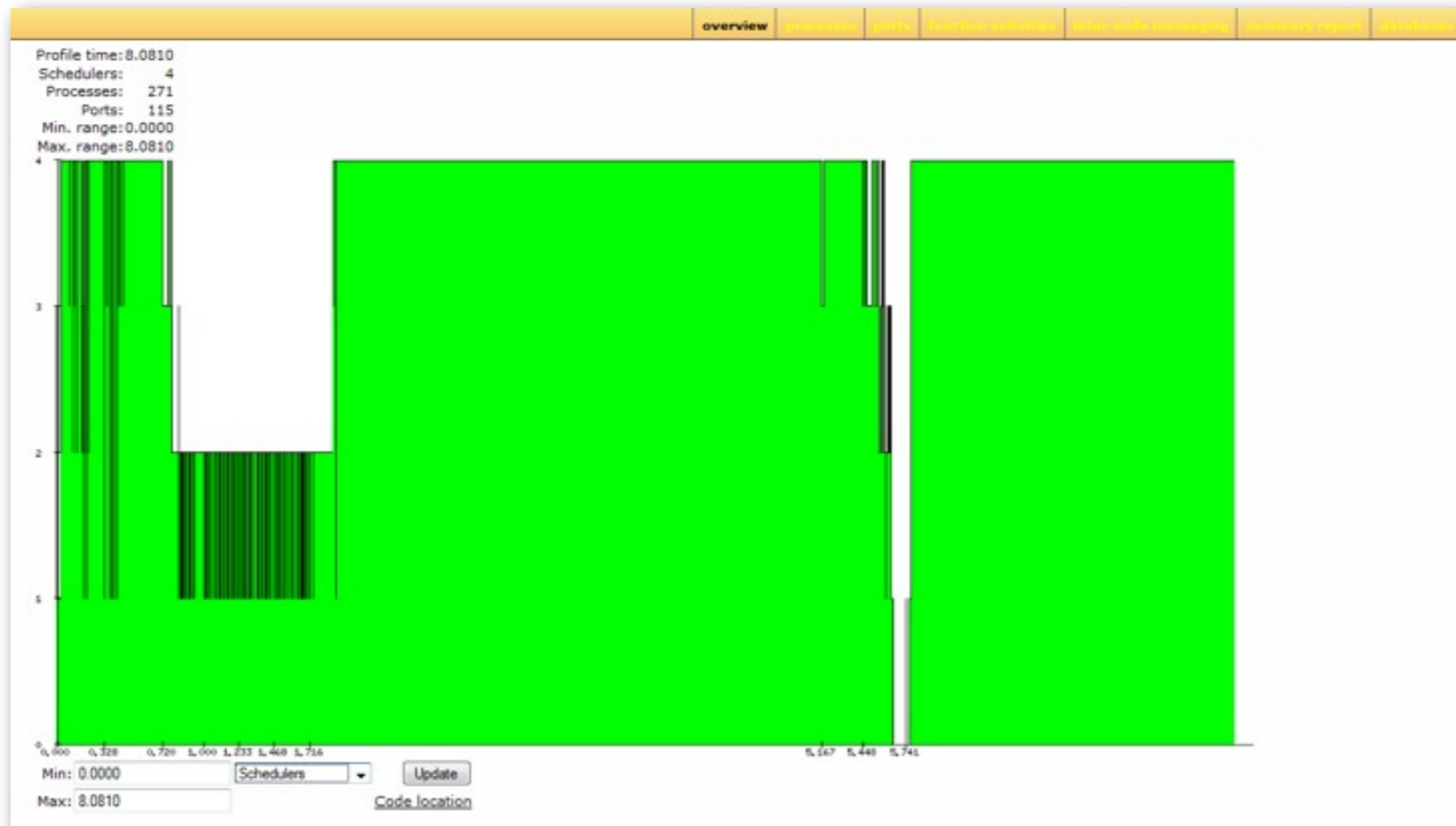
- Parallelisation
 - Dynamic callgraph information.
- Scalability of the system.
 - Usability (process tree)
 - Parallel processing

Percept2

Functionality

Scheduler activity

The number of schedulers active at any time



Process info tree



Process info presented in expandable tree format.

The screenshot displays a web interface for visualizing process information. The browser address bar shows the URL `localhost61700/cgi-bin/percept2_html/process_tree_page`. The page title is `percept2`. The main content is a table of process information with columns for `Select`, `[+/-]`, `Pid`, `Lifetime`, `Name`, `Parent`, `#RQ_chgs`, `#msgs_received`, `#msgs_sent`, `Entrypoint`, and `Callgraph`. The table is partially expanded, showing a tree structure of processes. A red box highlights a specific process and its children.

Select	[+/-]	Pid	Lifetime	Name	Parent	#RQ_chgs	#msgs_received	#msgs_sent	Entrypoint	Callgraph
<input type="checkbox"/>	[+]	<0.23.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	32	{460,1237}	{139,3920}	undefined	No callgraph/time
<input type="checkbox"/>	-	<0.24.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1909	{15795,72}	{15795,55}	undefined	No callgraph/time
<input type="checkbox"/>	-	<0.28.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	0	{6,92}	{5,85}	undefined	No callgraph/time
<input type="checkbox"/>	-	<0.30.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1	{6,150}	{12,67}	undefined	No callgraph/time
<input type="checkbox"/>	[+]	<0.37.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1	{48,1396}	{40,104}	undefined	show call graph/time
<input type="checkbox"/>	[+]	<0.340.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	10	{167,206}	{166,341}	sim_code_v3:init_hash_loop/0	show call graph/time
<input type="checkbox"/>	-	<0.341.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	21	{217,7280}	{216,167}	sim_code_v3:init_ast_loop/1	show call graph/time
<input type="checkbox"/>	[+]	<0.342.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	3	{11,24}	{1,24}	para_lib:pforeach_0/4	No callgraph/time
<input type="checkbox"/>	-	<0.543.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	8	{46,2503}	{1,36401}	sim_code_v3:init_clone_check/0	show call graph/time
<input type="checkbox"/>	[+]	<0.544.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	4	{53,24}	{1,24}	para_lib:pforeach_0/4	No callgraph/time
<input type="checkbox"/>	[+]	<0.330.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	0	{4,78}	{2,62}	undefined	No callgraph/time
<input type="checkbox"/>	[+]	<0.331.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1	{12,130}	{8,78}	undefined	No callgraph/time

At the bottom of the page, there are checkboxes for `Select all`, `Include children procs`, and `Include omitted procs`, along with a `Compare` button and a `Visualise Process Tree` button.

Process migration

		overview	processes	ports	function activities
Pid		<0.341.0>			
Name		undefined			
Entrypoint		sim_code_v3:init_ast_loop/1			
Arguments		[<0.340.0>]			
	Timestamp	Profile	Time		
Timetable	Start	{1352,219904,767820}	0.0163		
	Stop	undefined	8.0810		
Parent		<0.37.0>			
Children					
RQ_history		[1,4,1,3,1,2,1,2,1,3,1,4,3,2,1,2,1,3,2,1,2,1]			
{#msg_received, avg_msg_size}		{217,7280}			
{#msg_sent, avg_msg_size}		{216,167}			
accumulated runtime (in milliseconds)		1678			
Callgraph/time		show call graph/time			
percentage of total waiting time	total	mean	stddev	#recv	module:function/arity
	4.8195	0.6024	0.8377	8	sim_code_v3:ast_loop/1
	0.1195	0.0011	0.0041	108	sim_code_v3:insert_hash/2

Message passing stats

No. of msgs, and average msg size, sent / received.

localhost:61700/cgi-bin/percept2_html/process_tree_page

percept2

						overview	processes	ports	function activities	inter-node messaging	summary report	database
Select	[+/-]	Pid	Lifetime	Name	Parent	#RQ_chgs	#msgs_received	#msgs_sent	Entrypoint	Callgraph		
<input type="checkbox"/>	[+]	<0.23.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	32	{460,1237}	{139,3920}	undefined	callgraph/time	No	
<input type="checkbox"/>	-	<0.24.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1909	{15795,72}	{15795,55}	undefined	callgraph/time	No	
<input type="checkbox"/>	-	<0.28.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	0	{6,92}	{5,85}	undefined	callgraph/time	No	
<input type="checkbox"/>	-	<0.30.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1	{6,150}	{12,67}	undefined	callgraph/time	No	
<input type="checkbox"/>	[+]	<0.37.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1	{48,1396}	{40,104}	undefined	callgraph/time	show call graph/time	
Select	[+/-]	Pid	Lifetime	Name	Parent	#RQ_chgs	#msgs_received	#msgs_sent	Entrypoint	Callgraph		
<input type="checkbox"/>	+	<0.340.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	10	{167,206}	{166,341}	sim_code_v3:init_hash_loop/0	callgraph/time	show call graph/time	
<input type="checkbox"/>	-	<0.341.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	21	{217,7280}	{216,167}	sim_code_v3:init_ast_loop/1	callgraph/time	show call graph/time	
<input type="checkbox"/>	[+]	<0.342.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	3	{11,24}	{1,24}	para_lib:pforeach_0/4	callgraph/time	No	
<input type="checkbox"/>	-	<0.543.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	8	{46,2503}	{1,36401}	sim_code_v3:init_clone_check/0	callgraph/time	show call graph/time	
<input type="checkbox"/>	+	<0.544.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	<0.37.0>	4	{53,24}	{1,24}	para_lib:pforeach_0/4	callgraph/time	No	
<input type="checkbox"/>	+	<0.330.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	0	{4,78}	{2,62}	undefined	callgraph/time	No	
<input type="checkbox"/>	+	<0.331.0>	<div style="width: 100%; height: 10px; background-color: green;"></div>	undefined	undefined	1	{12,130}	{8,78}	undefined	callgraph/time	No	

Select all
 Include children procs
 Include omitted procs
 [Visualise Process Tree](#)

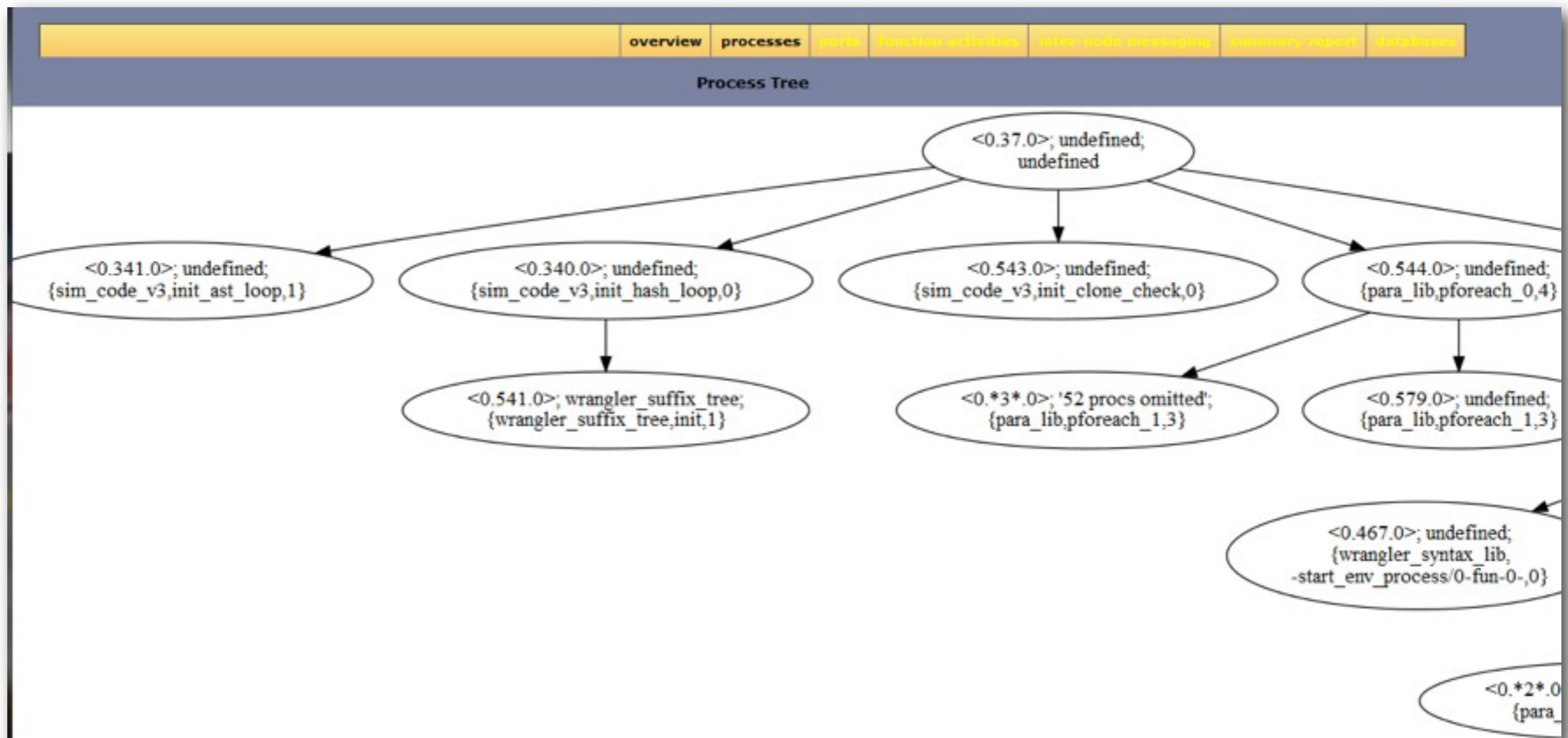
Information aggregation

											overview	processes	ports	function activities	inter-node messaging	summary report	databases
Select [+/-]	Pid	Lifetime	Name	Parent	#RQ_chgs	#msgs_received	#msgs_sent	Entrypoint	Callgraph								
<input type="checkbox"/>	<0.23.0>		undefined	undefined	32	{460,1237}	{139,3920}	undefined	No callgraph/time								
Select [+/-]	Pid	Lifetime	Name	Parent	#RQ_chgs	#msgs_received	#msgs_sent	Entrypoint	Callgraph								
<input type="checkbox"/>	<0.*4*.0>		'66 procs omitted'	<0.23.0>	0	{3375,162}	{941,994}	file_io_server:-do_start/4-fun-0-'/0	No callgraph/time								
<input type="checkbox"/>	<0.362.0>		undefined	<0.23.0>	39	{631,171}	{151,1269}	nle_io_server:-do_start/4-fun-0-'/0	No callgraph/time								
<input type="checkbox"/>	<0.24.0>		undefined	undefined	1909	{15795,72}	{15795,55}	undefined	No callgraph/time								
<input type="checkbox"/>	<0.28.0>		undefined	undefined	0	{6,92}	{5,85}	undefined	No callgraph/time								
<input type="checkbox"/>	<0.30.0>		undefined	undefined	1	{6,150}	{12,67}	undefined	No callgraph/time								
<input checked="" type="checkbox"/>	<0.37.0>		undefined	undefined	1	{48,1396}	{40,104}	undefined	show call graph/time								
<input type="checkbox"/>	<0.330.0>		undefined	undefined	0	{4,78}	{2,62}	undefined	No callgraph/time								
<input type="checkbox"/>	<0.331.0>		undefined	undefined	1	{12,130}	{8,78}	undefined	No callgraph/time								

Select all
 Include children procs
 Include omitted procs

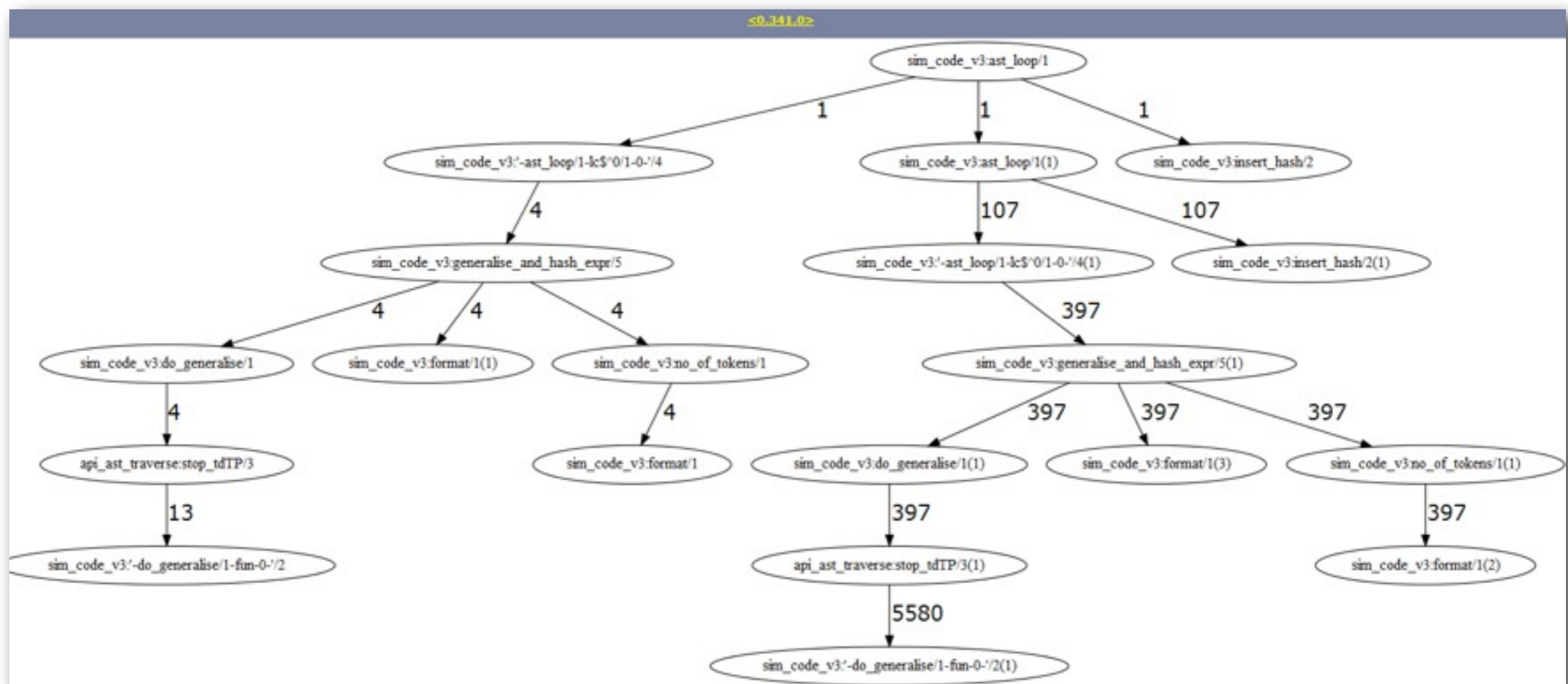
Process tree

Parent-child relationships between processes.



Dynamic function info

Call relationships b/w fns during the program run, and number of calls.

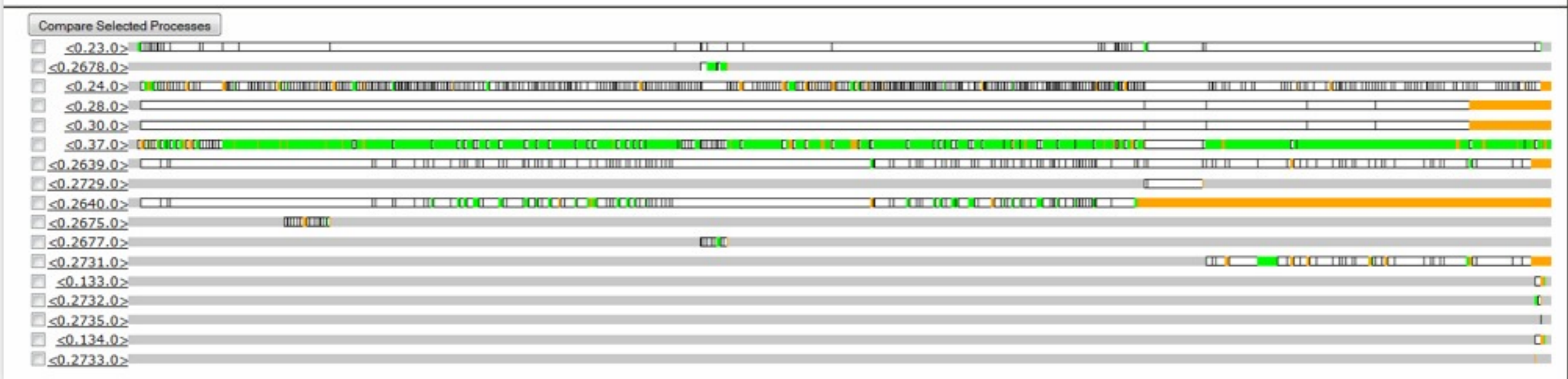
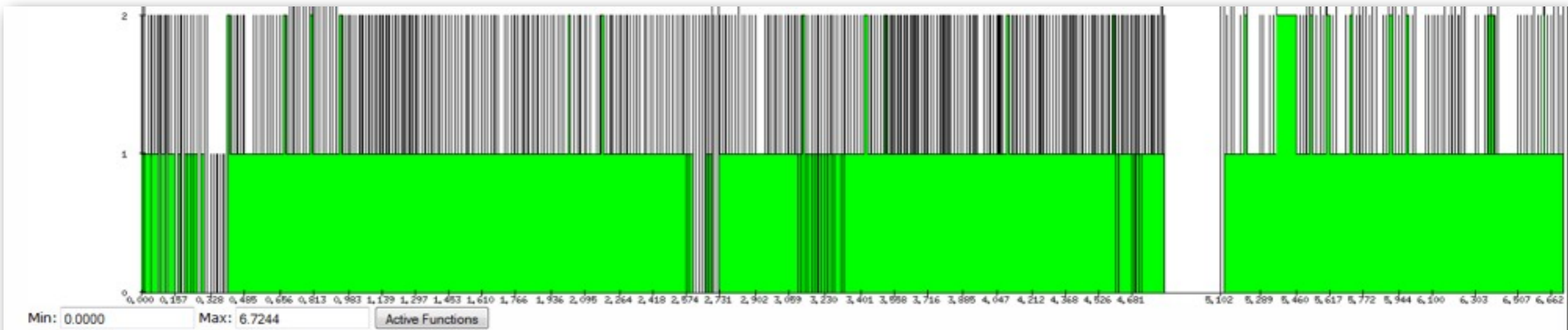


Accumulated runtime

Accumulated time during which a function within this process is in a running state.

		overview	processes
Pid		<0.37.0>	
Entrypoint		undefined	
M:F/A		<u>sim_code v3:check clone candidates/3</u>	
Call count		1	
Accumulated time (in milliseconds)		1873	
Callers	module:function/arity call count		
	<u>sim_code v3:sim code detection 1/6</u>	1	
Called	module:function/arity call count		
	<u>sim_code v3:start clone check process/0</u>	1	
	<u>sim_code v3:examine clone candidates/4</u>	1	
	<u>sim_code v3:combine clones by au/1</u>	1	
	<u>sim_code v3:stop clone check process/1</u>	1	
	<u>sim_code v3:'-check clone candidates/3-lc\$^0/1-0-'/1</u>	1	

Process activity



Scalability

Compression

overview processes ports function activities inter-node messaging summary report databases

Select [+/-]	Pid	Lifetime	Name	Parent	#RQ_chgs	#msgs_received	#msgs_sent	Entrypoint	Callgraph
<input type="checkbox"/> [-]	<0.23.0>		undefined	undefined	32	{460,1237}	{139,3920}	undefined	No callgraph/time
<input type="checkbox"/> [-]	<0.*4*.0>		'66 procs omitted'	<0.23.0>	0	{3375,162}	{941,994}	file_io_server: '- do_start/4- fun-0-' / 0	No callgraph/time
<input type="checkbox"/> [-]	<0.362.0>		undefined	<0.23.0>	39	{631,171}	{151,1269}	file_io_server: '- do_start/4- fun-0-' / 0	No callgraph/time
<input type="checkbox"/> [-]	<0.24.0>		undefined	undefined	1909	{15795,72}	{15795,55}	undefined	No callgraph/time
<input type="checkbox"/> [-]	<0.28.0>		undefined	undefined	0	{6,92}	{5,85}	undefined	No callgraph/time
<input type="checkbox"/> [-]	<0.30.0>		undefined	undefined	1	{6,150}	{12,67}	undefined	No callgraph/time
<input type="checkbox"/> [+]	<0.37.0>		undefined	undefined	1	{48,1396}	{40,104}	undefined	show call graph/time
<input type="checkbox"/> +	<0.330.0>		undefined	undefined	0	{4,78}	{2,62}	undefined	No callgraph/time
<input type="checkbox"/> +	<0.331.0>		undefined	undefined	1	{12,130}	{8,78}	undefined	No callgraph/time

Select all Include children procs Include omitted procs

Compare

Visualise Process Tree

Parallelise

- Parallelise Percept itself.
- Multiple log files ...
- ... that can be processed in parallel ...
- ... and integrated into a single result.

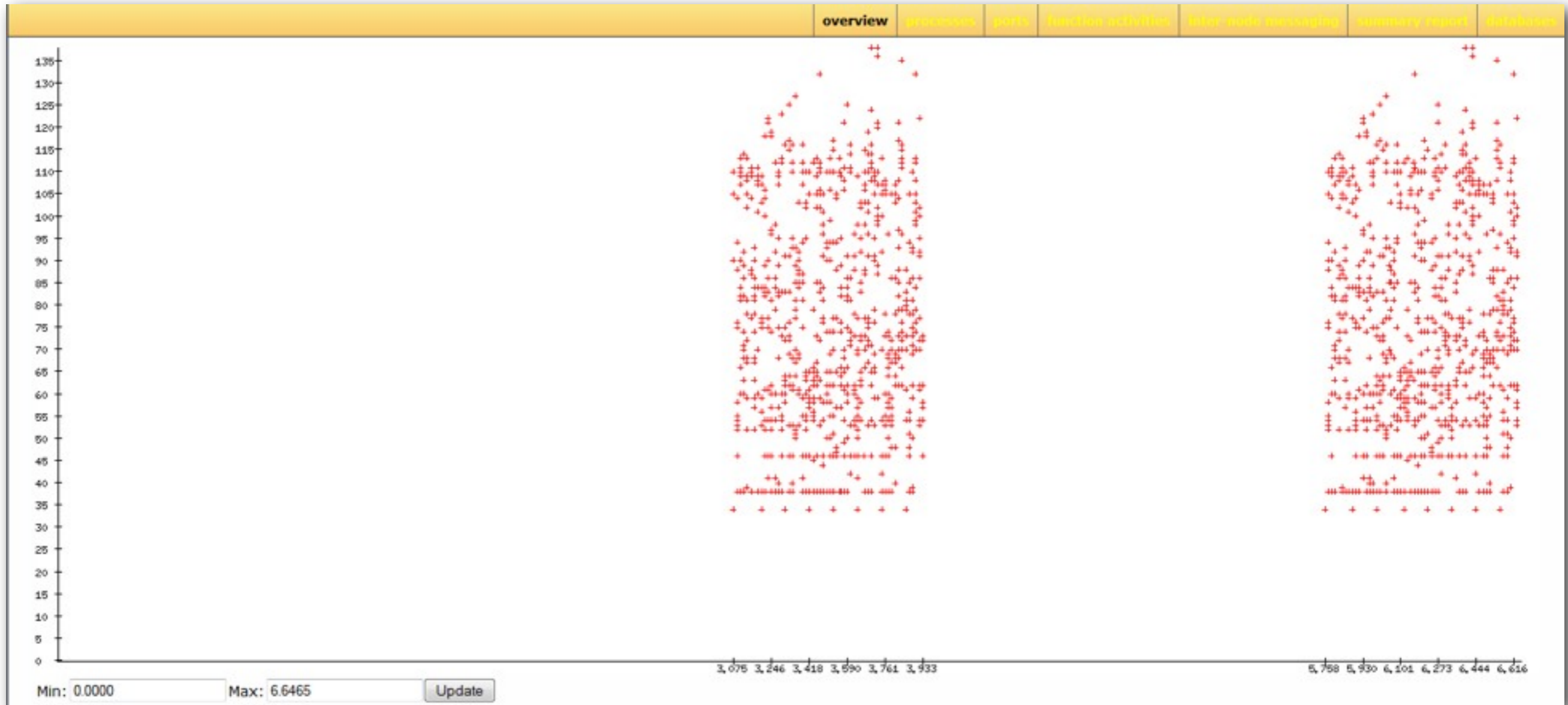
Distribution

overview processes ports function activities inter-node messaging

From node: server@hl-It
 server@hl-It
 test@hl-It
 client@hl-It

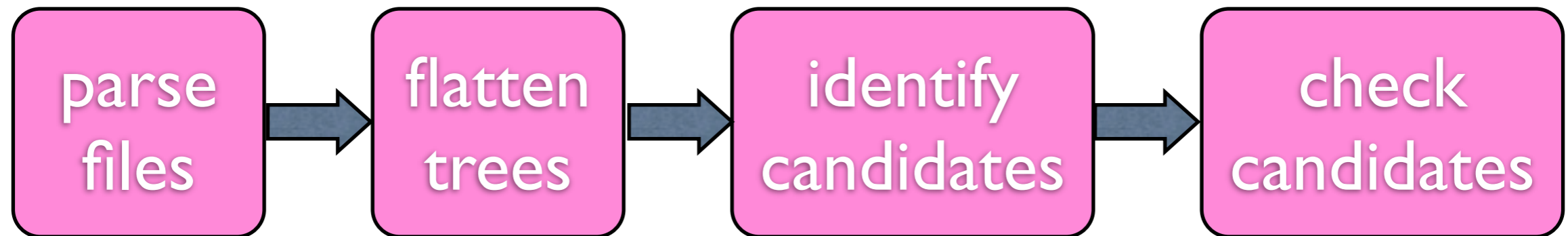
To node: server@hl-It

Generate Graph

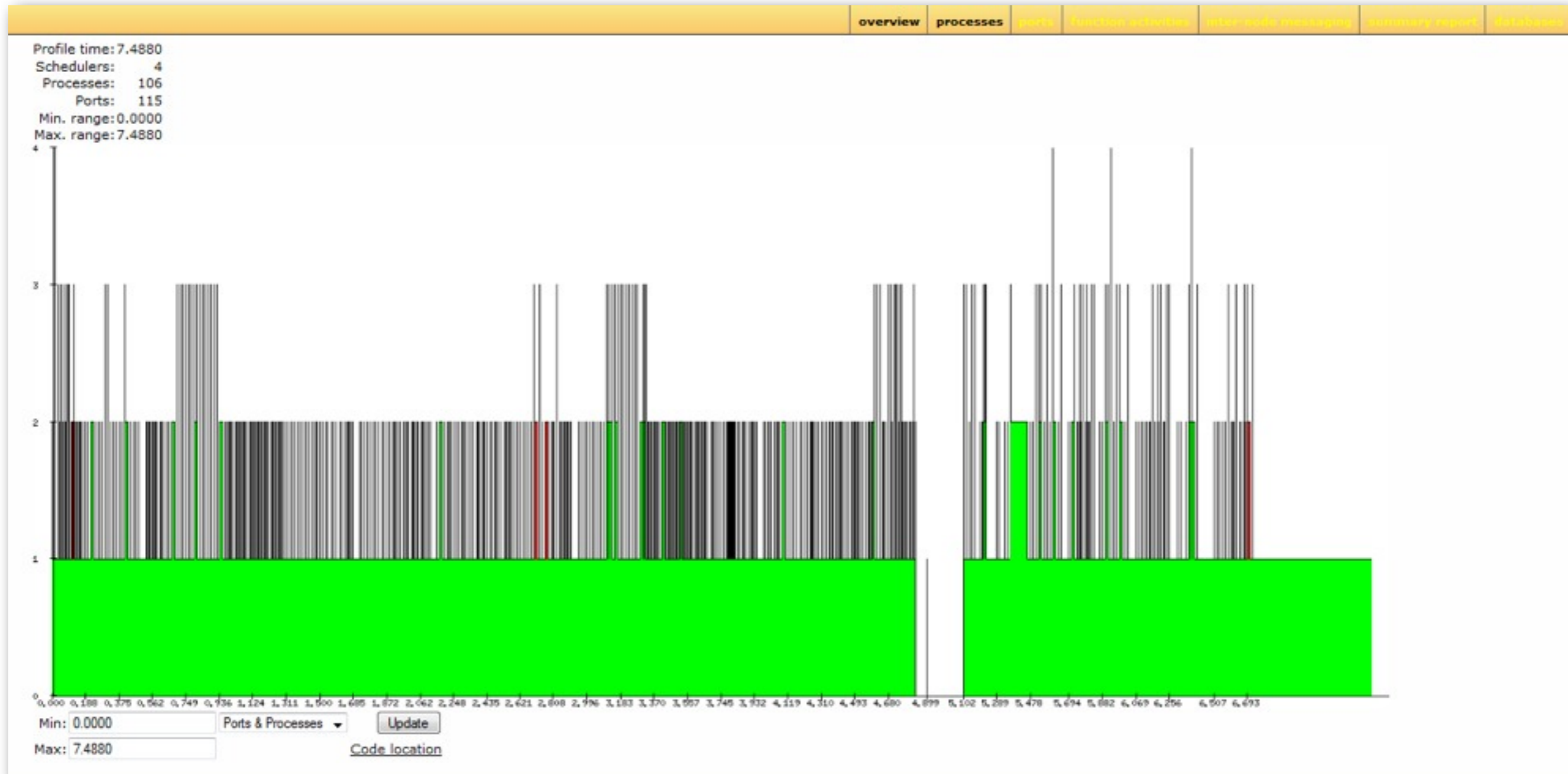


Case study

Clone detection in Wrangler



Active Processes



Processes

overview processes **ports**

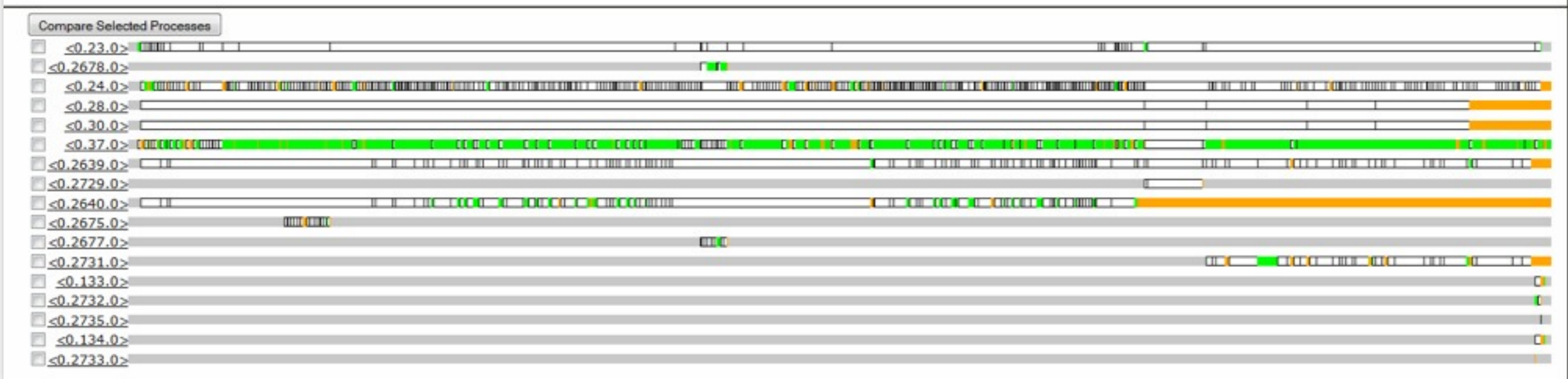
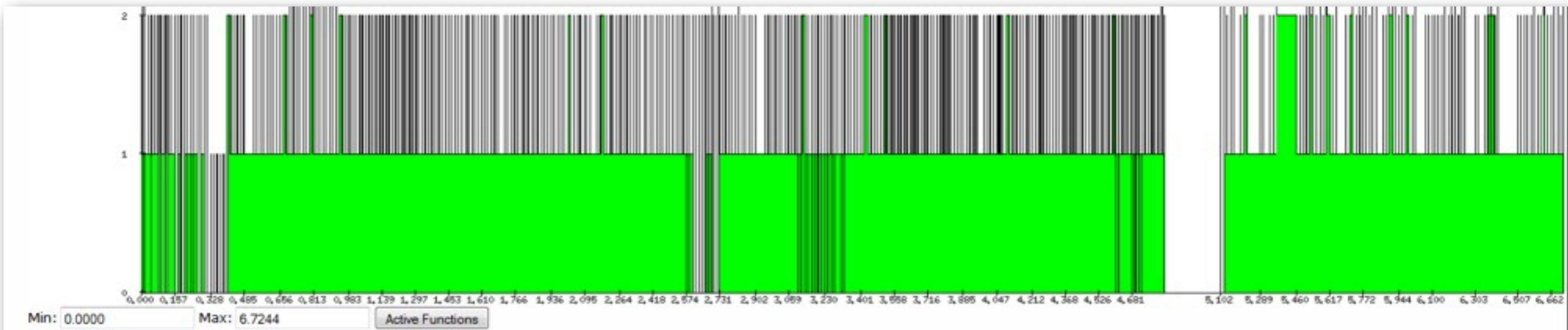
Select [+/-]	Pid	Lifetime	Name	Parent	#RQ_chgs	#msgs_received	#msgs_sent	Entrypoint	Callgraph
<input checked="" type="checkbox"/> +	<0.23.0>		undefined	undefined	2	{460,1237}	{139,3920}	undefined	No callgraph/time
<input checked="" type="checkbox"/> -	<0.24.0>		undefined	undefined	3	{15823,72}	{15823,55}	undefined	No callgraph/time
<input checked="" type="checkbox"/> -	<0.28.0>		undefined	undefined	0	{11,91}	{10,85}	undefined	No callgraph/time
<input checked="" type="checkbox"/> -	<0.30.0>		undefined	undefined	1	{11,143}	{22,66}	undefined	No callgraph/time
<input checked="" type="checkbox"/> +	<0.37.0>		undefined	undefined	6	{19938,1688}	{19971,152}	undefined	show call graph/time
<input checked="" type="checkbox"/> +	<0.133.0>		undefined	undefined	0	{4,78}	{2,62}	undefined	No callgraph/time
<input checked="" type="checkbox"/> +	<0.134.0>		undefined	undefined	0	{12,130}	{8,78}	undefined	No callgraph/time

Select all Include children procs Include omitted procs

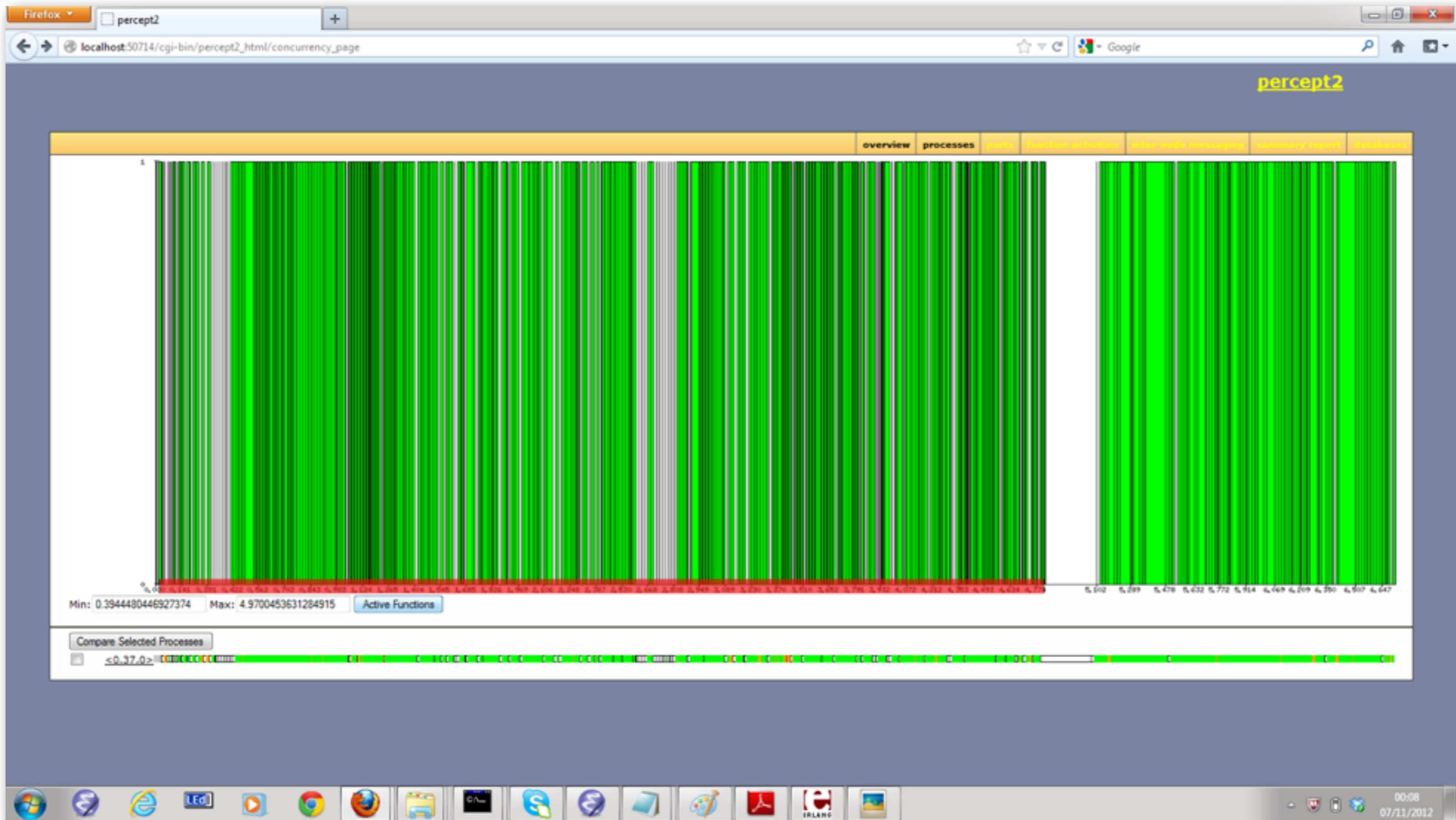
Compare

Visualise Process Tree

Process activity



Critical process



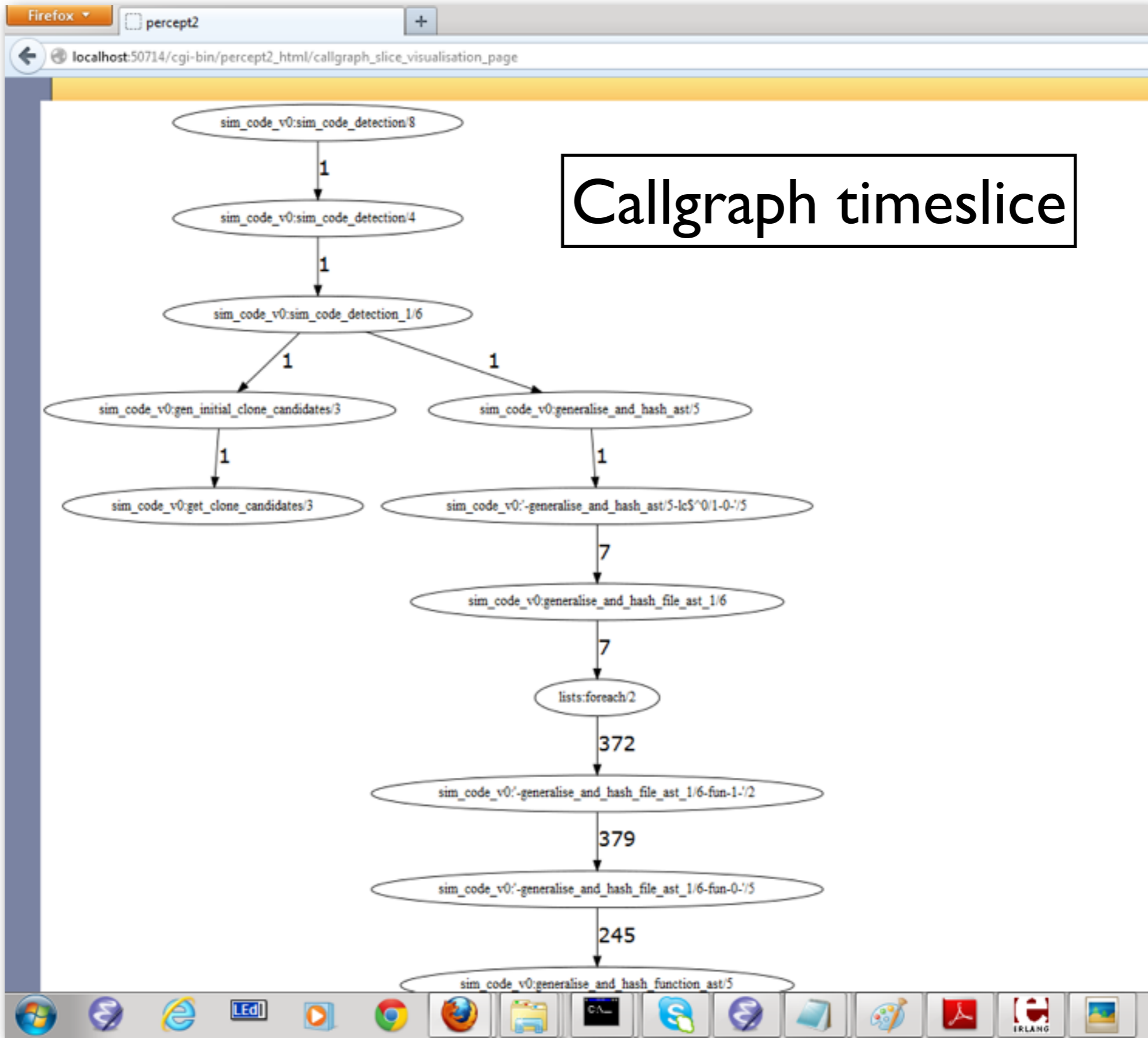
Active functions

Min. range:0.3944
Max. range:4.9700

<0.37.0> ▾

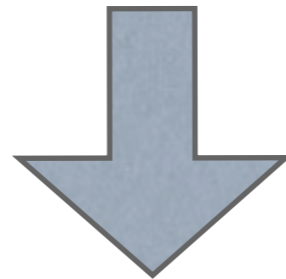
Call Graph Slice

pid	module:function/arity	active period	call count
<0.37.0>	sim code v0:sim code detection/8		1
<0.37.0>	sim code v0:sim code detection/4		1
<0.37.0>	sim code v0:sim code detection 1/6		1
<0.37.0>	sim code v0:generalise and hash ast/5		1
<0.37.0>	sim code v0:'-generalise and hash ast/5-lc\$^0/1-0-/5		1
<0.37.0>	sim code v0:generalise and hash file ast 1/6		7
<0.37.0>	lists:foreach/2		7
<0.37.0>	sim code v0:'-generalise and hash file ast 1/6-fun-0-/5		379
<0.37.0>	sim code v0:'-generalise and hash file ast 1/6-fun-1-/2		372
<0.37.0>	sim code v0:generalise and hash function ast/5		245
<0.37.0>	sim code v0:format/1		245
<0.37.0>	sim code v0:generalise and hash function ast 1/7		245
<0.37.0>	sim code v0:absolute to relative loc/2		245
<0.37.0>	api ast traverse:full tdTP/3		245
<0.37.0>	sim code v0:do abs to relative loc/2		12027
<0.37.0>	sim code v0:'-do abs to relative loc/2-lc\$^0/1-0-/2		12027
<0.37.0>	sim code v0:abs to relative loc in ann/2		80446
<0.37.0>	sim code v0:'-abs to relative loc in ann/2-lc\$^0/1-0-/2		12027
<0.37.0>	sim code v0:'-abs to relative loc in ann/2-lc\$^1/1-1-/2		12027
<0.37.0>	sim code v0:to relative/2		57209
<0.37.0>	sim code v0:'-absolute to relative loc/2-fun-0-/2		11782
<0.37.0>	sim code v0:'-abs to relative loc in ann/2-lc\$^2/1-2-/2		3418
<0.37.0>	sim code v0:'-generalise and hash function ast 1/7-fun-0-/8		12027
<0.37.0>	sim code v0:generalise and hash body/6		899
<0.37.0>	sim code v0:insert to ast tab/2		105
<0.37.0>	sim code v0:gen initial clone candidates/3		1
<0.37.0>	sim code v0:get clone candidates/3		1



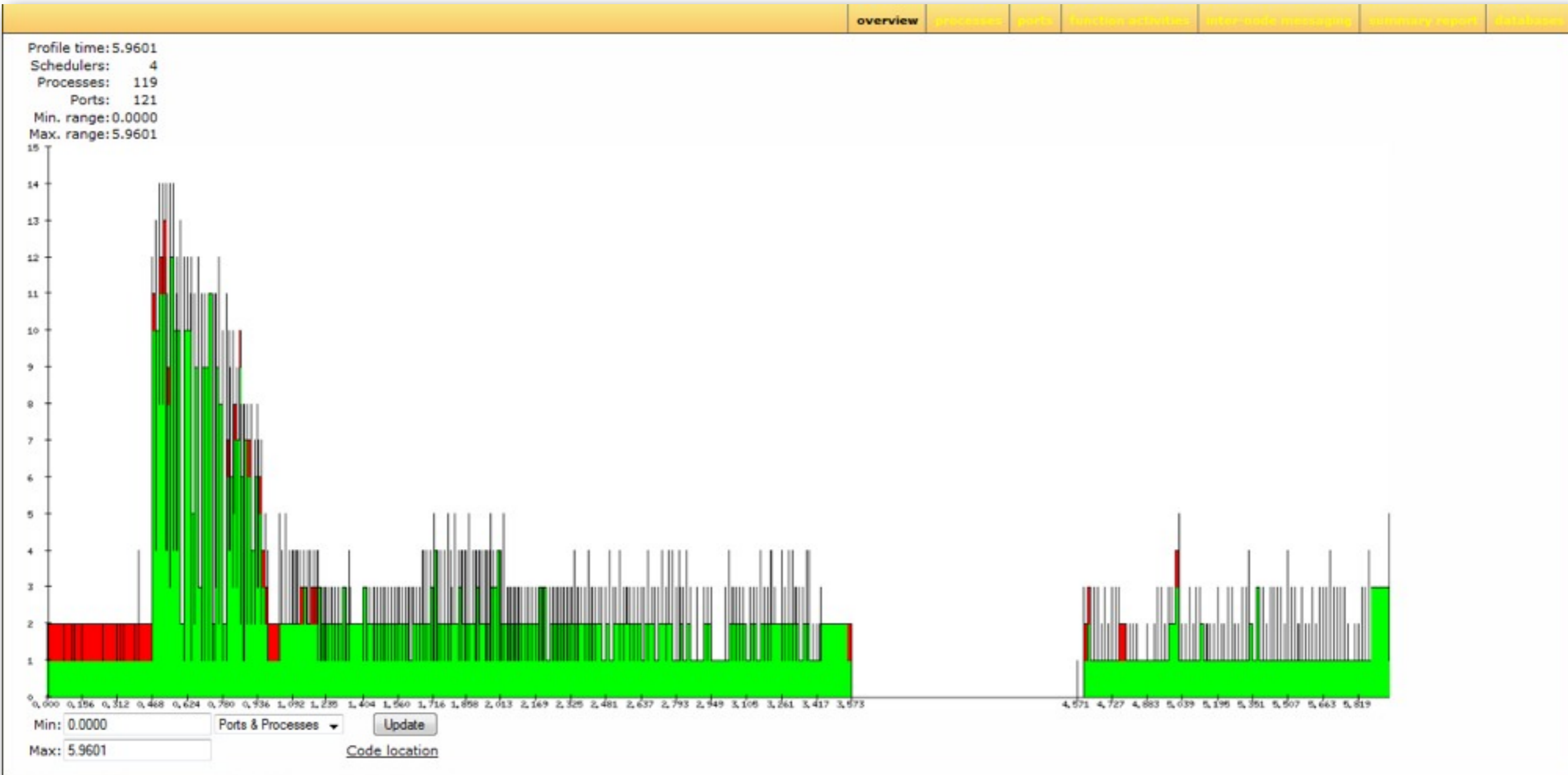
Refactoring #1

```
[generalise_and_hash_file_ast_l(  
  File, Threshold, ASTPid, true, SearchPaths, TabWidth)  
||File<-Files].
```

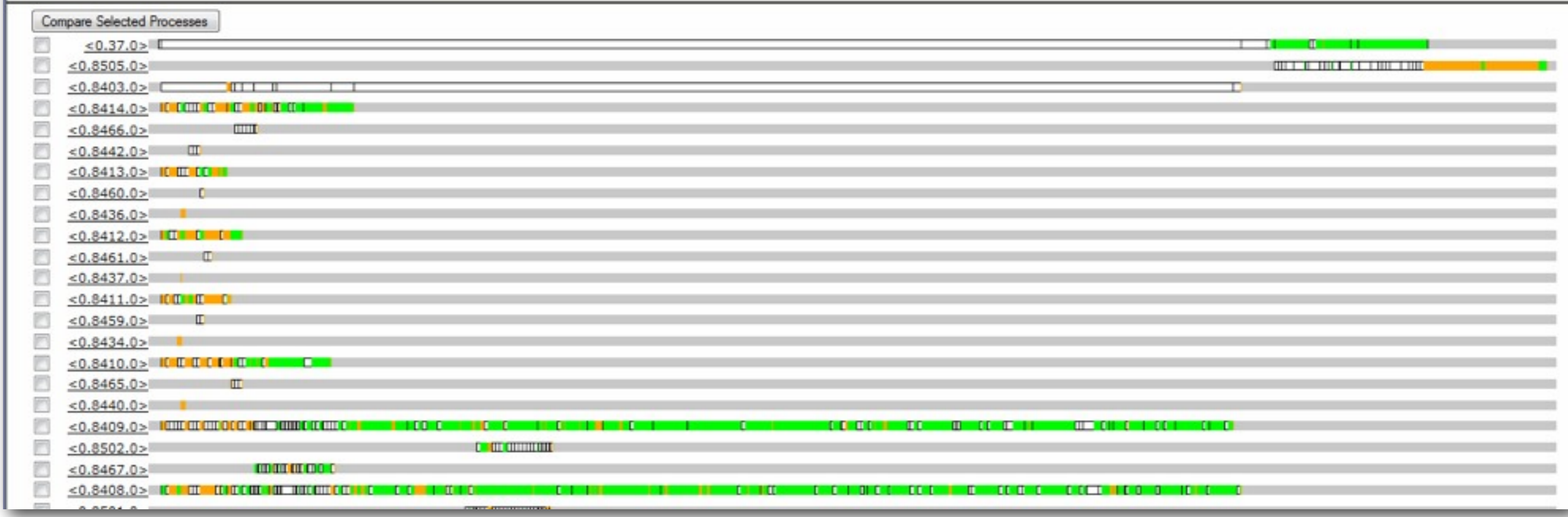
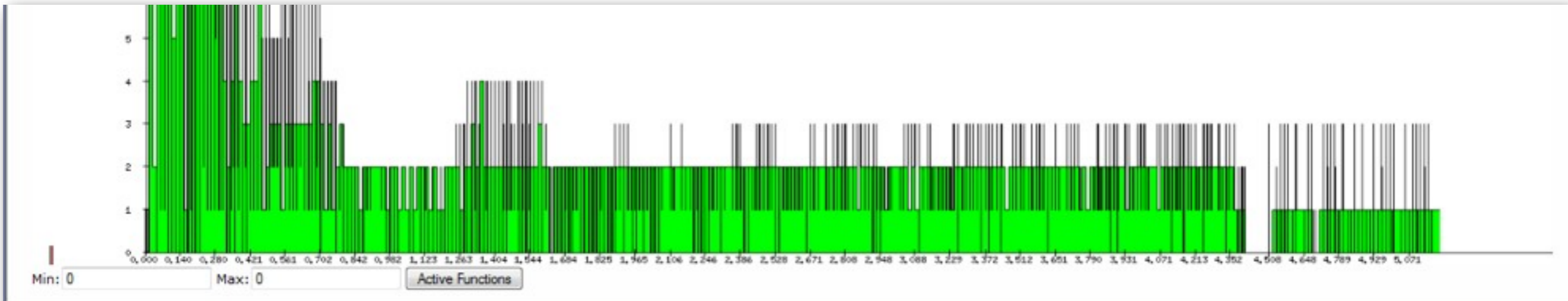


```
para_lib:pmap(  
  fun(File) ->  
    generalise_and_hash_file_ast_l(  
      File, Threshold, ASTPid, true, SearchPaths, TabWidth)  
end, Files).
```

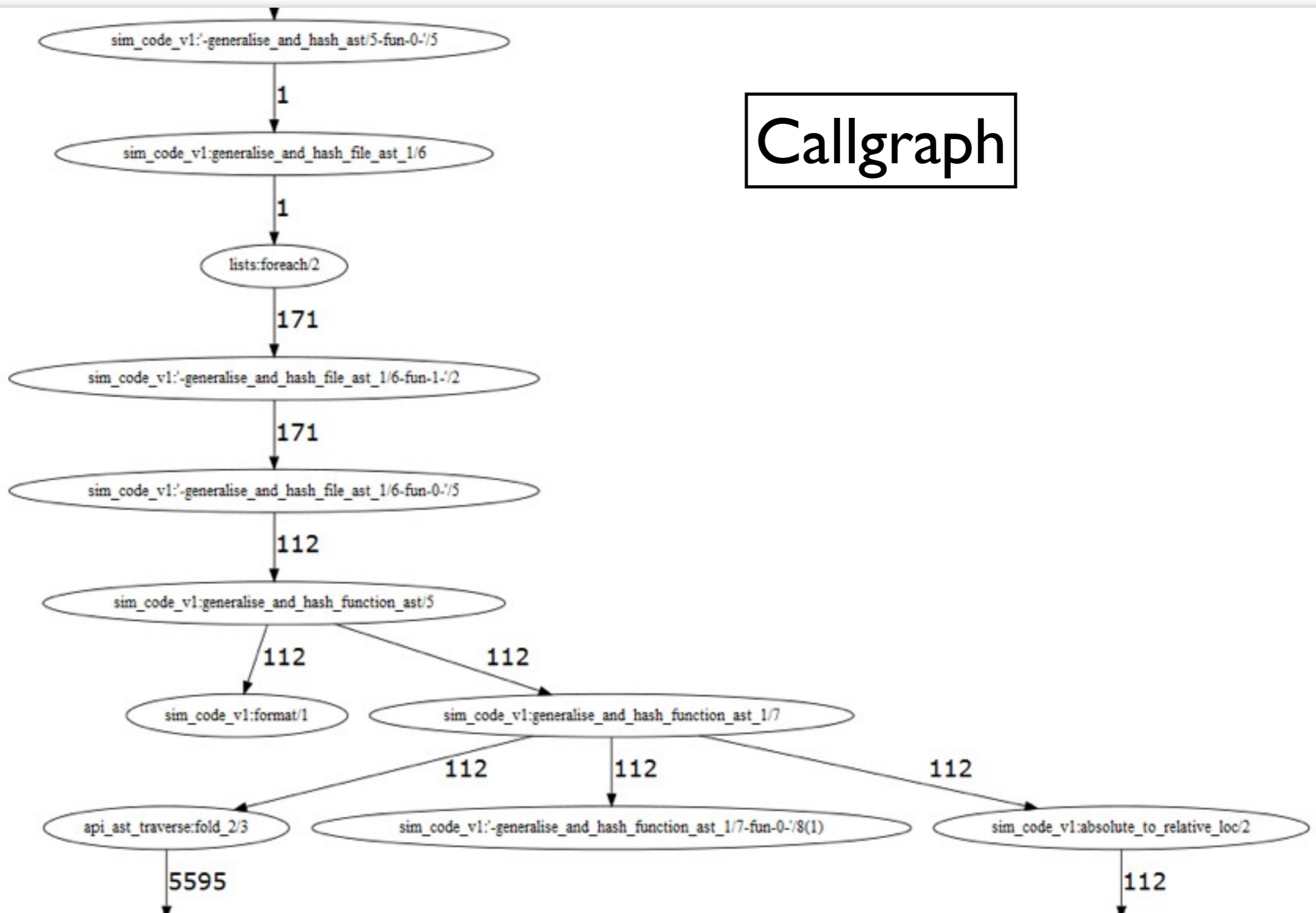
Active Processes



Processes

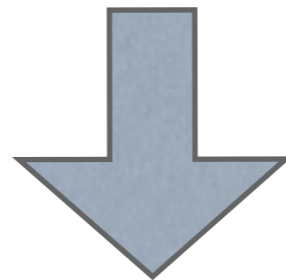


Callgraph



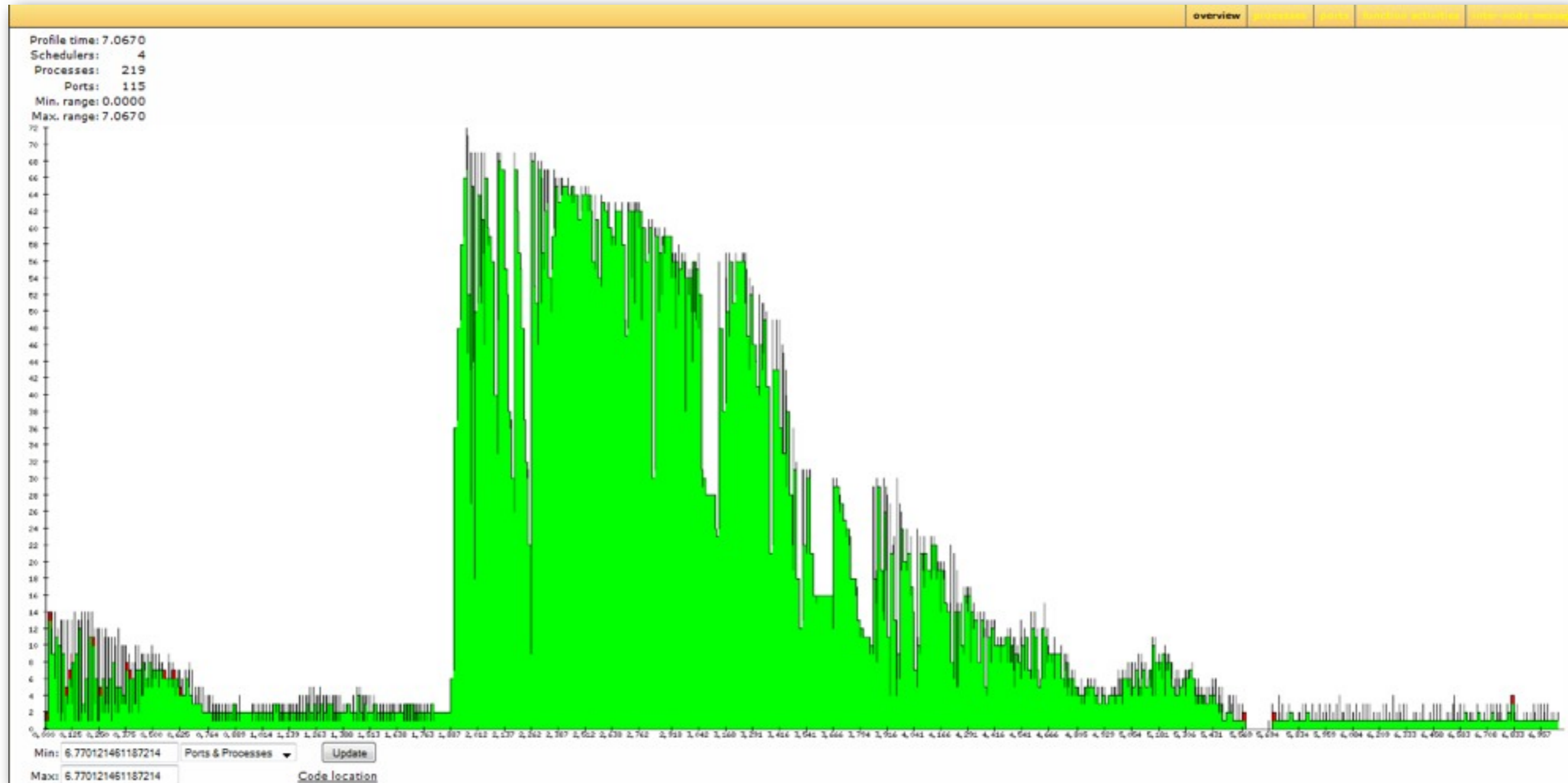
Refactoring #2

```
lists:foreach(fun (Form) -> F(Form) end, Forms).
```

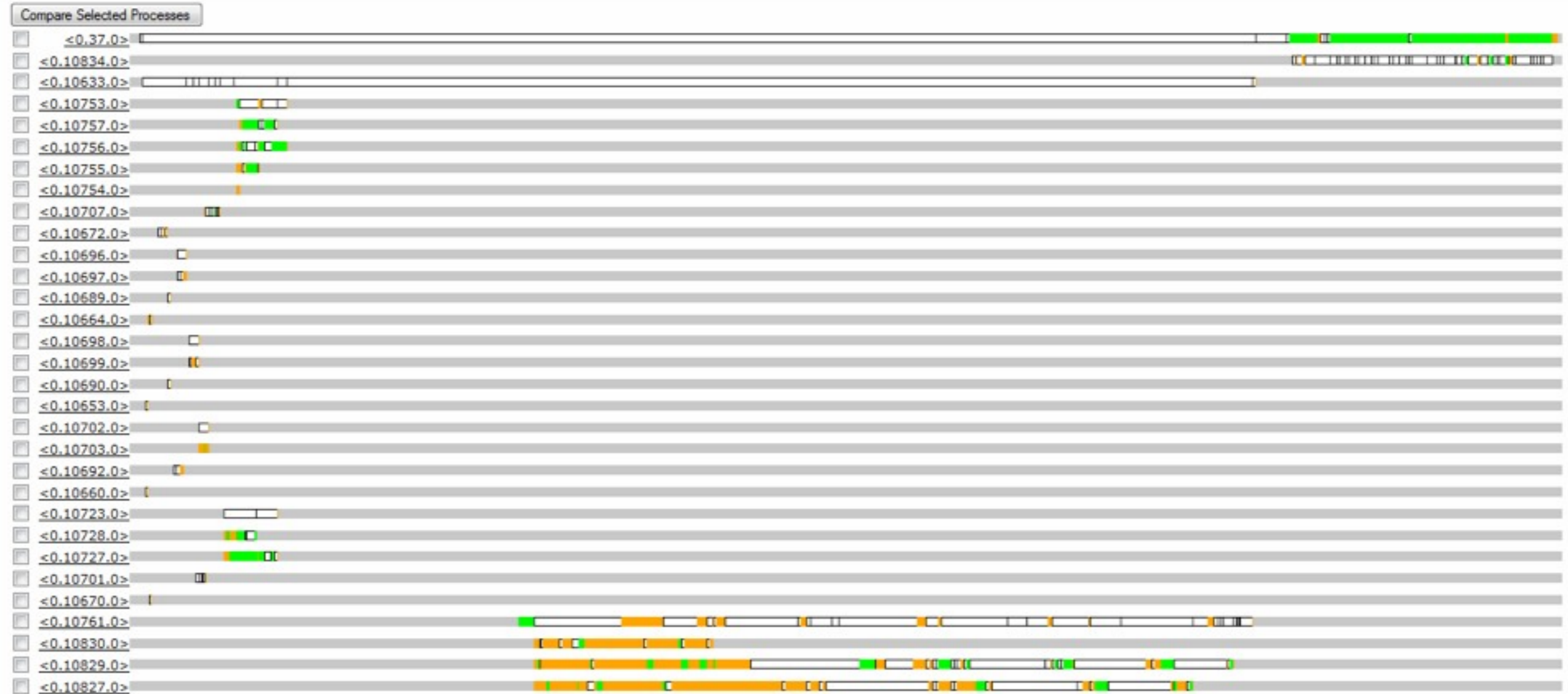
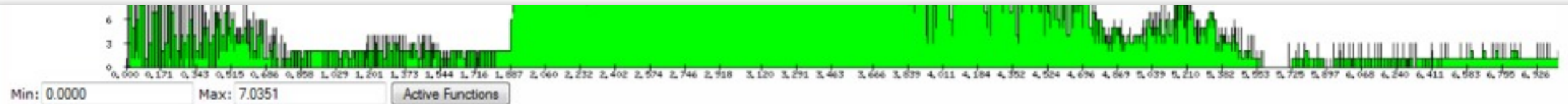


```
para_lib:pforeach(fun (Form) -> F(Form) end, Forms, 5).
```

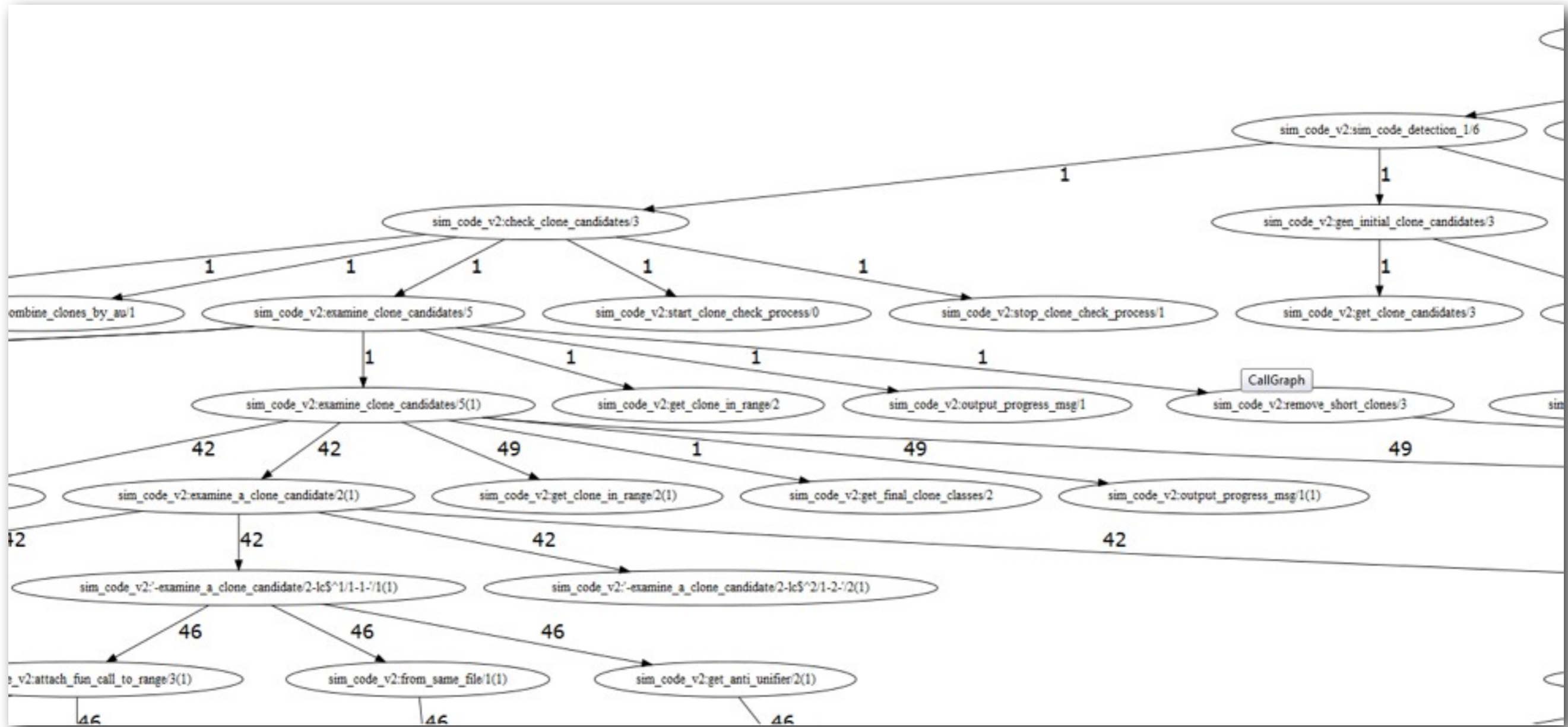
Active Processes



Processes



Callgraph



Refactoring #3

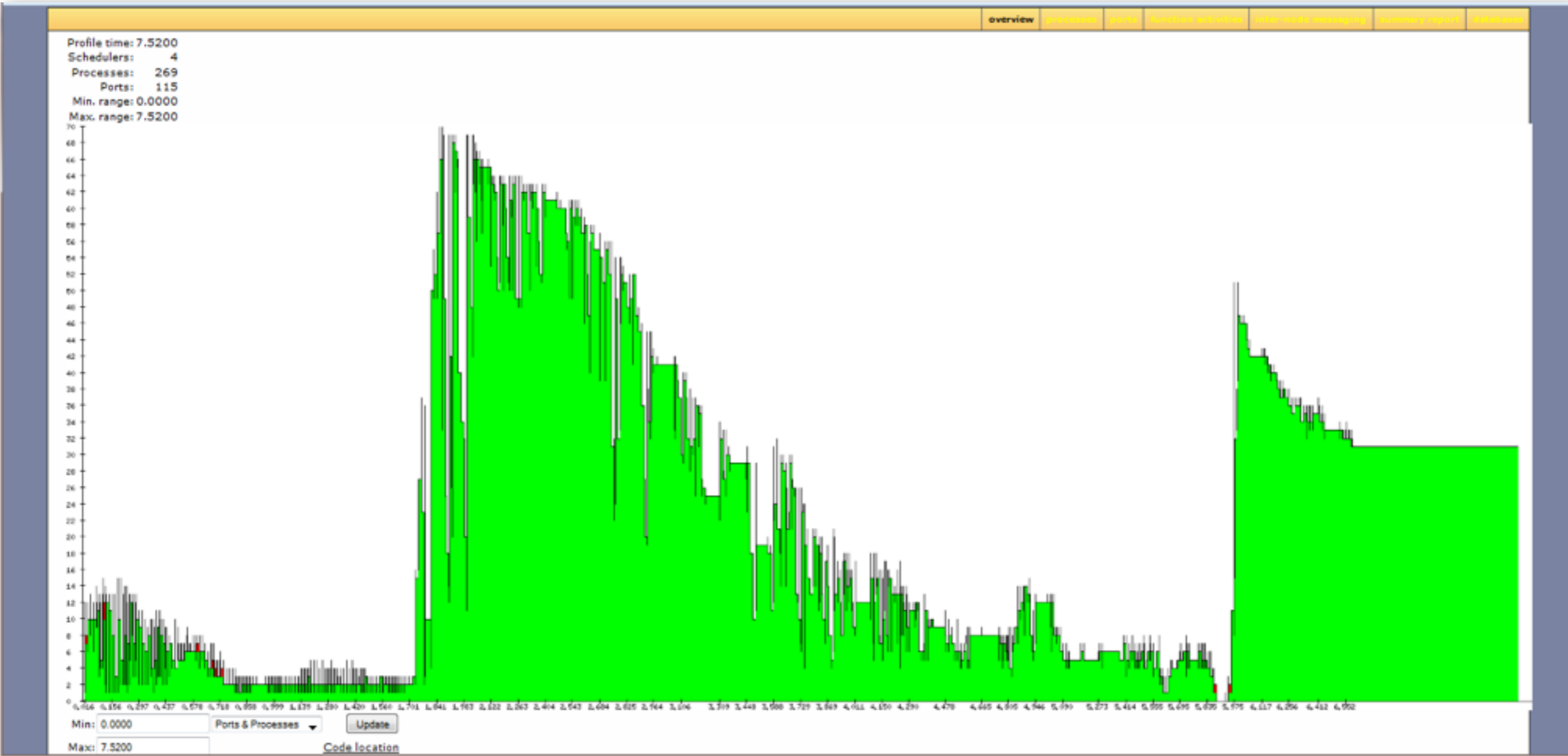
```
examine_clone_candidates([],_Thresholds,CloneCheckerPid,_Num) ->
  get_final_clone_classes(CloneCheckerPid);
examine_clone_candidates([C|Cs],Thresholds,CloneCheckerPid,Num) ->
  output_progress_msg(Num),
  CI = get_clone_in_range(C),
  NewClones=examine_a_clone_candidate(CI,Thresholds),
  add_new_clones(CloneCheckerPid,{CI, NewClones}),
  examine_clone_candidates(Cs,Thresholds,CloneCheckerPid,Num+1).
```



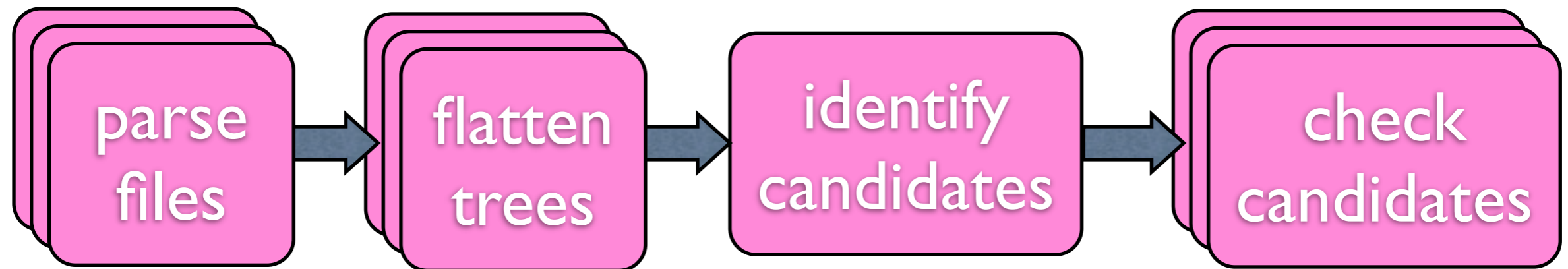
```
examine_clone_candidates(Cs,Thresholds, CloneCheckerPid) ->
  NumberedCs = lists:zip(Cs, lists:seq(1, length(Cs))),
  para_lib:pforeach(
    fun({C, Nth}) ->
      examine_a_clone_candidate(
        {C,Nth},Thresholds,CloneCheckerPid)
    end,NumberedCs),
  get_final_clone_classes(CloneCheckerPid).
```

```
examine_a_clone_candidate({C,_Nth},Thresholds,CloneCheckerPid) ->
  output_progress_msg(Nth),
  CI = get_clone_in_range(C),
  NewClones=examine_a_clone_candidate(CI,Thresholds),
  add_new_clones(CloneCheckerPid,{CI,NewClones}).
```

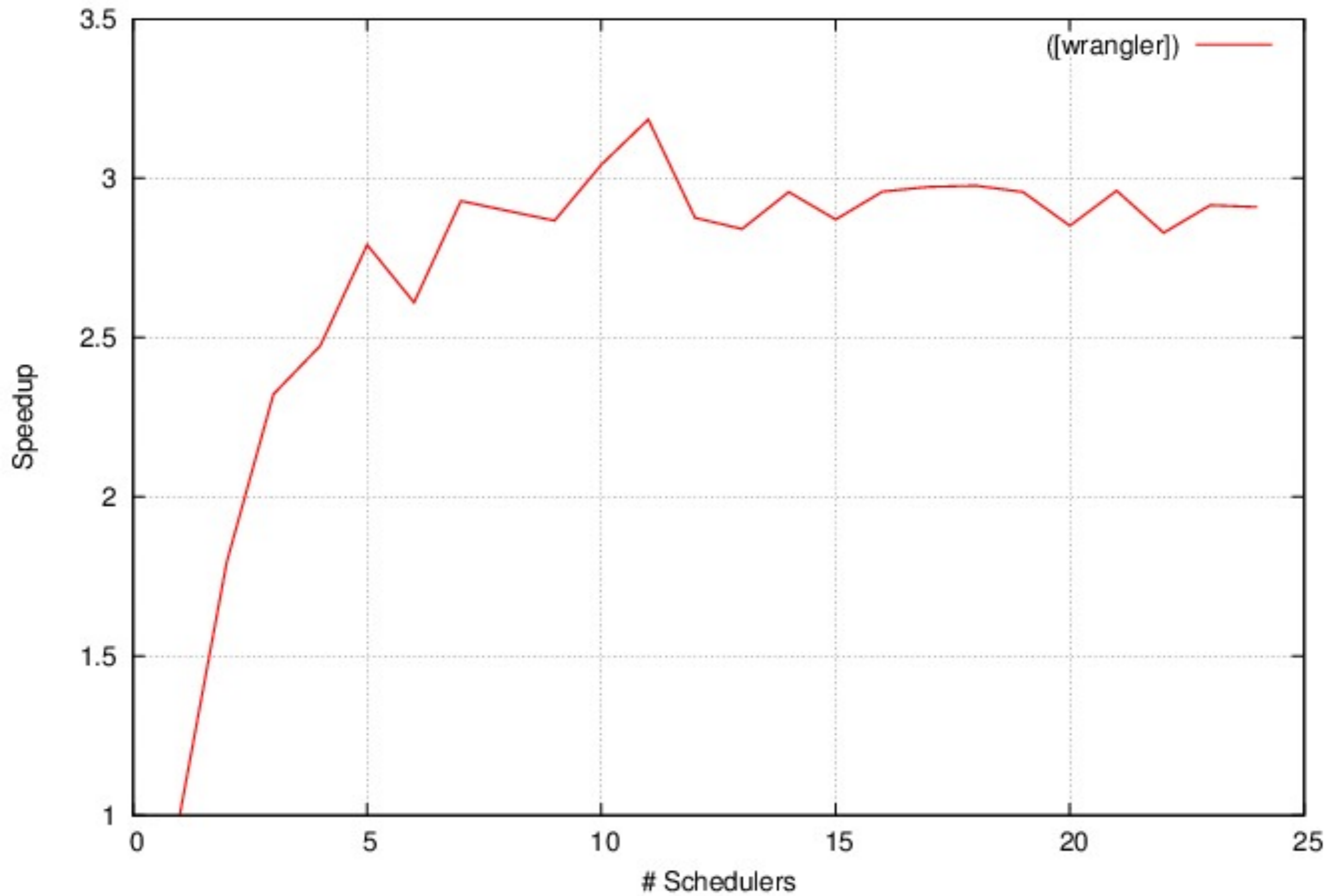
Active Processes



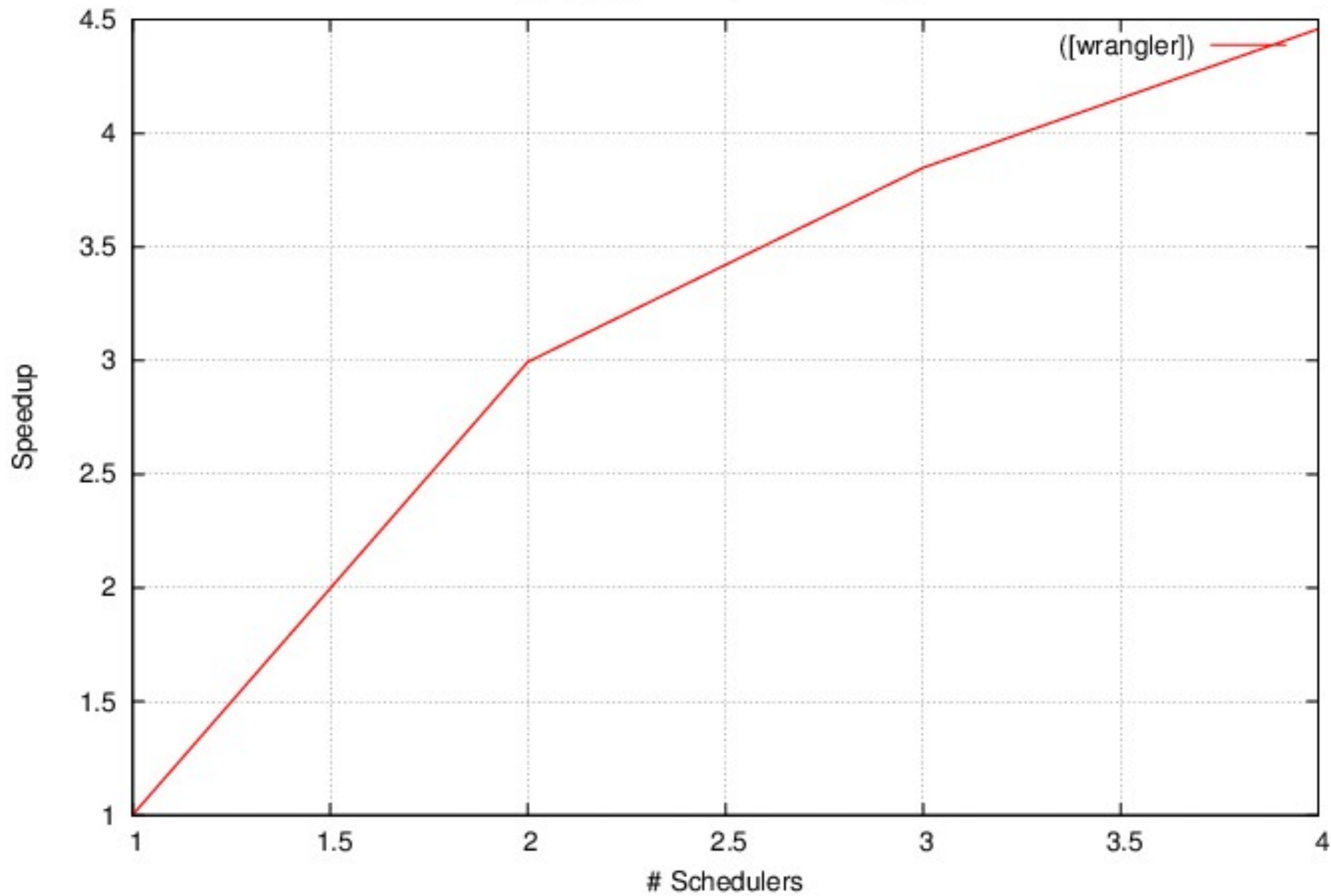
Parallel clone detection in Wrangler



sim_code - DEFOTP - DEFARGS



sim_code - DEFOTP - DEFARGS



Next steps

- Looking at the front end ... HTML5 version.
- Online SD Erlang Profiling
- Refactoring Tools



RELEASE

- Dialyzer parallelisation at Uppsala ...
- ... paper in Erlang Workshop 2012.
- BenchErl: benchmarking
- Extended tracing in Erlang and DTrace.

Conclusions

- RELEASE project.
- Tool to support scalable parallelisation.
- SD Erlang release and tools in 2013.
- Thanks to the percept team!

Get involved!

[git://github.com/huiqing/percept2.git](https://github.com/huiqing/percept2.git)