Parallel Programming in Erlang

John Hughes

What is Erlang?

Erlang

Haskell
- Types
- Lazyness
- Purity
+ Concurrency
+ Syntax

If you know Haskell, Erlang is easy to learn!

QuickSort again

- Haskell

```haskell
qsort [] = []
qsort (x:xs) = qsort [y | y <- xs, y<x]
               ++ [x]
               ++ qsort [y | y <- xs, y>=x]
```

- Erlang

```erlang
qsort([]) -> [];
qsort([X|Xs]) -> qsort([Y || Y <- Xs, Y<X])
                 ++ [X]
                 ++ qsort([Y || Y <- Xs, Y>=X]).
```

- Haskell

```
qsort [] = []
qsort (x:xs) = qsort [y | y <- xs, y<x]
               ++ [x]
```

qsort [] =

- Erlang

```
qsort([]) -> [];
qsort([X|Xs]) -> qsort([Y || Y <- Xs, Y<X])
                 ++ [X]
                 ++ qsort([Y || Y <- Xs, Y>=X]).
```

qsort([]) ->

QuickSort again

- Haskell

```
qsort [] = []
qsort (x:xs) = q                   <- xs
               ++ [x]
               ++ qsort [y | y <- xs
```

;

.

- Erlang

```
qsort([]) -> [];
qsort([X|Xs]) -> qsort([Y || Y <- Xs, Y<X])
               ++ [X]
               ++ qsort([Y || Y <- Xs, Y>=X]).
```

Qu gain

x:xs

- Haskell

```
qsort [] = []
qsort (x:xs) = qsort [y | y <- xs, y<x]
               ++ [x]
               ++                xs, y>=x]
```

[X|Xs]

- Erlang

```
qsort([]) -> [];
qsort([X|Xs]) -> qsort([Y || Y <- Xs, Y<X])
                 ++ [X]
                 ++ qsort([Y || Y <- Xs, Y>=X]).
```

QuickSort again

|

||

- Haskell

```
qsort [] = []
qsort (x:xs) = qsort [y | y <- xs, y<x]
               ++ [x]
               ++ qsort [y | y <- xs
```

- Erlang

```
qsort([]) -> [];
qsort([X|Xs]) -> qsort([Y || Y <- Xs, Y<X])
                 ++ [X]
                 ++ qsort([Y || Y <- Xs, Y>=X]).
```

foo.erl

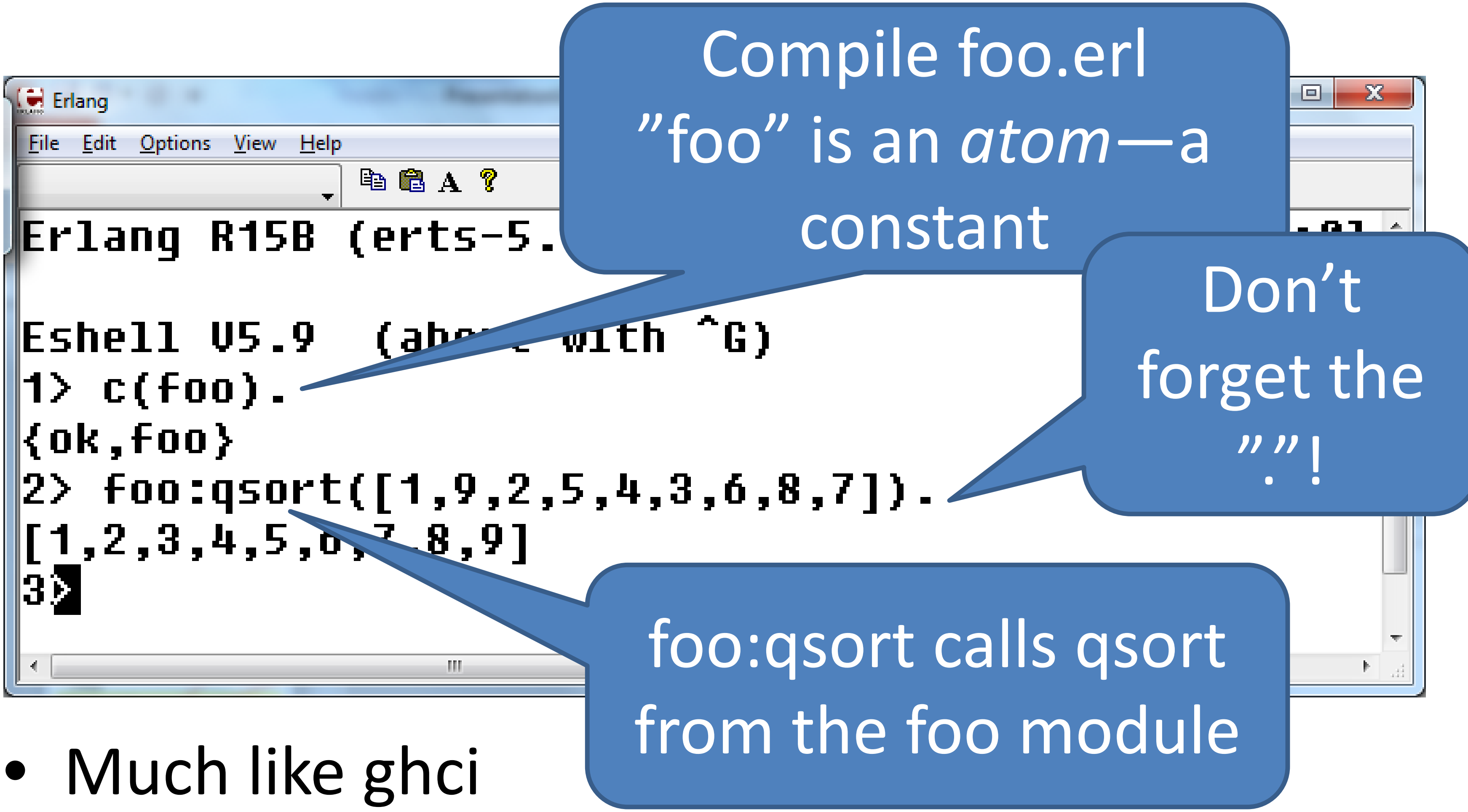
Declare the module name

```
-module(foo).
-compile(export_all).

qsort([]) ->
    [];
qsort([X|Xs]) ->
    qsort([Y || Y <- Xs, Y<X]) ++
     [X] ++
     qsort([Y || Y <- Xs, Y>=X]).
```

Simplest just to export everything

werl/erl REPL



- Much like ghci

Test Data

- Create some test data; in foo.erl:

```
random_list(N) ->
    [random:uniform(1000000) || _ <- lists:seq(1,N)].
```

Side-effects!

Instead of [1..N]

- In the

```
L = foo:random_list(200000).
```

Timing calls

```
79> timer:tc(foo,qsort,[L]).
{390000,
 [1,2,6,8,11,21,33,37,4[illegible]
  51,59,61,69,70,75,86,[illegible]
  1[illegible],105,106,112,117,118,123|...]}
```

- Module
- Function
- Arguments
- *atoms*—i.e. constants
- Microseconds
- {A,B,C} is a tuple

Benchmarking

Binding a name… c.f. **let**

Macro: current module name

```
benchmark(Fun,L) ->
    Runs = [timer:tc(?MODULE,Fun,[L])
            || _ <- lists:seq(1,100)],
    lists:sum([T || {T,_} <- Runs]) /
        (1000*length(Runs)).
```

- 100 runs, average & convert to ms

```
80> foo:benchmark(qsort,L).
285.16
```

Parallelism

```
34> erlang:system_info(schedulers).
8
```

Eight OS threads!
Let's use them!

Parallelism in Erlang

- Processes are created *explicitly*

```
Pid = spawn_link(fun() -> …Body… end)
```

- Start a process which executes …Body…
- **`fun() -> Body end`** ~ **`\() -> Body`**
- **`Pid`** is the *process identifier*

Parallel Sorting

```
psort([]) ->
    [];
psort([X|Xs]) ->
    spawn_link(
      fun() ->
        psort([Y || Y <- Xs, Y >= X])
      end),
    psort([Y || Y <- Xs, Y < X]) ++
          [X] ++
          ???.
```

Sort second half in parallel…

But how do we get the result?

Message Passing

Pid ! Msg

- Send a message to Pid
- *Asynchronous*—do not wait for delivery

Message Receipt

```
receive
    Msg -> …
end
```

- Wait for a message, then bind it to Msg

Parallel Sorting

```
psort([]) ->
    [];
psort([X|Xs]) ->
    Parent = self(),
    spawn_link(
      fun() ->
        Parent !
            psort([Y || Y <- Xs, Y >= X])
      end),
    psort([Y || Y <- Xs, Y < X]) ++
          [X] ++
          receive Ys -> Ys end.
```

The Pid of the executing process

Send the result back to the parent

Wait for the result *after* sorting the first half

Benchmarks

```
84> foo:benchmark(qsort,L).
285.16
85> foo:benchmark(psort,L).
474.43
```

- Parallel sort is slower! *Why?*

Controlling Granularity

```
psort2(Xs) -> psort2(5,Xs).

psort2(0,Xs) -> qsort(Xs);
psort2(_,[]) -> [];
psort2(D,[X|Xs]) ->
    Parent = self(),
    spawn_link(fun() ->
      Parent !
        psort2(D-1,[Y || Y <- Xs, Y >= X])
      end),
    psort2(D-1,[Y || Y <- Xs, Y < X]) ++
    [X] ++
    receive Ys -> Ys end.
```

Benchmarks

```
84> foo:benchmark(qsort,L).
285.16
85> foo:benchmark(psort,L).
377.74
86>
foo:benchmark(psort2,L).
109.2
```

- 2.6x speedup on 4 cores (x2 hyperthreads)

Profiling Parallelism with Percept

File to store profiling information in

{Module,Function, Args}

```
87> percept:profile("test.dat",{foo,psort2,[L]},[procs]).
Starting profiling.
ok
```

Profiling Parallelism with Percept

Analyse the file, building a RAM database

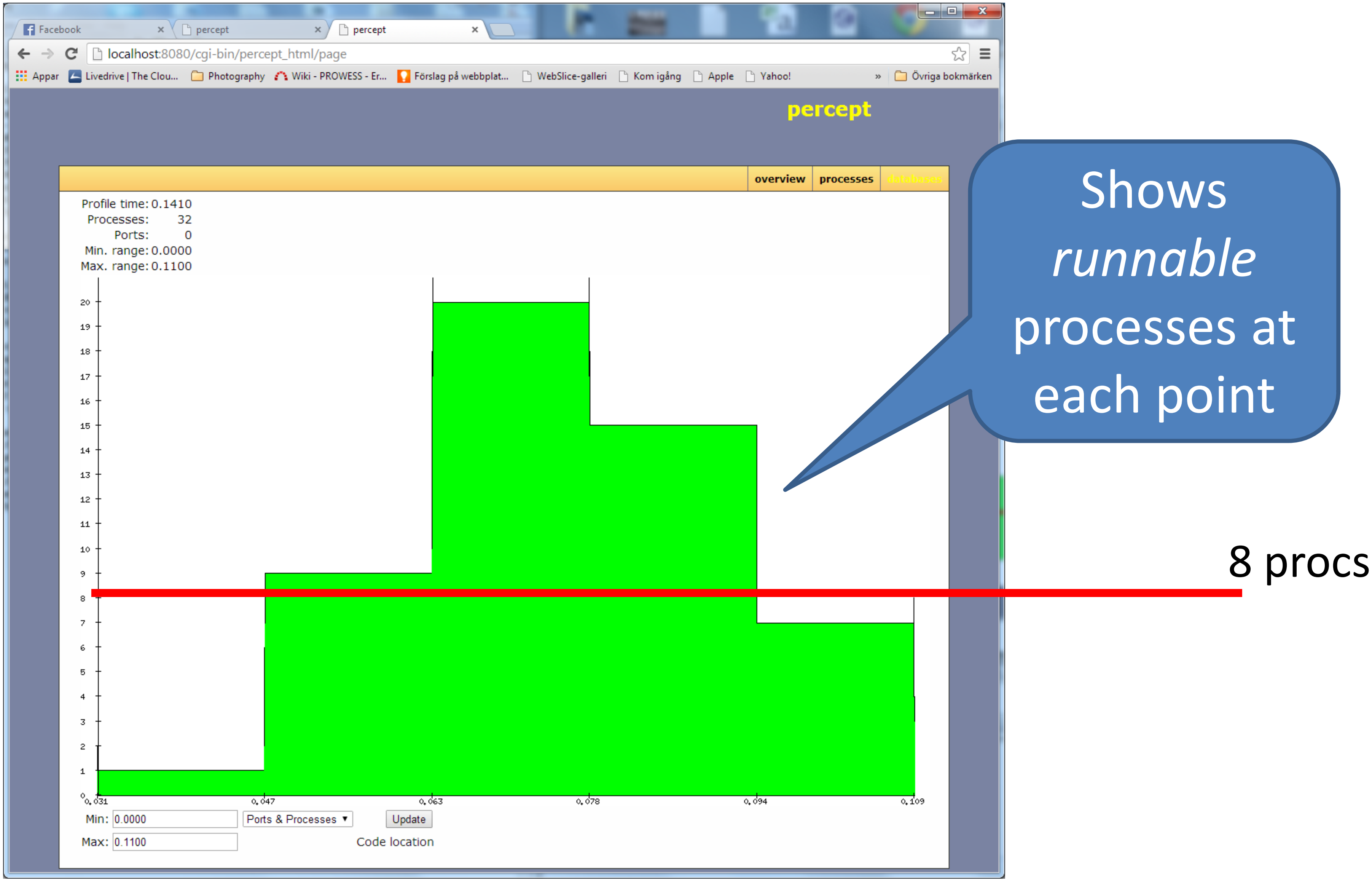
```
88> percept:analyze("test.dat").
Parsing: "test.dat"
Consolidating...
Parsed 160 entries in 0.093 s.
    32 created processes.
    0 opened ports.
ok
```

Profiling Parallelism with Percept

Start a web server to display the profile on this port

```
90> percept:start_webserver(8080).
{started,"HALL",8080}
```

Profiling Parallelism with Percept



Profiling Parallelism with Percept

percept

overview | processes | databases

Processes

Select	Pid	Lifetime	Entrypoint	Name	Parent
	<0.31.0>		undefined	undefined	undefined
	<0.21776.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.31.0>
	<0.21777.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21776.721>
	<0.21778.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.31.0>
	<0.21779.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21776.721>
	<0.21780.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.31.0>
	<0.21781.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21778.721>
	<0.21782.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21777.721>
	<0.21783.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21781.721>
	<0.21784.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21778.721>
	<0.21785.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21781.721>
	<0.21786.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21783.721>
	<0.21787.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.31.0>
	<0.21788.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21784.721>
	<0.21789.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21778.721>
	<0.21790.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21776.721>
	<0.21791.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21779.721>
	<0.21792.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21790.721>
	<0.21793.721>		foo:'-psort2/2-fun-0-'/0	undefined	<0.21777.721>

Examining a single process

percept

overview | processes | databases

Pid	<0.21776.721>
Name	undefined
Entrypoint	foo:'-psort2/2-fun-0-'/0
Arguments	
Timetable	
Parent	<0.31.0>
Children	<0.21797.721> <0.21790.721> <0.21779.721> <0.21777.721>

Timetable:

	Timestamp	Profile Time
Start	{1395,949124,702000}	0.0160
Stop	{1395,949124,827005}	0.1410

percentage	total	mean	stddev	#recv	module:function/arity
100 %	0.0470	0.0117	0.0149	4	foo:psort2/2

Correctness

```
91> foo:psort2(L) == foo:qsort(L).
false
92> foo:psort2("hello world").
" edhllloorw"
```

Oops!

What’s going on?

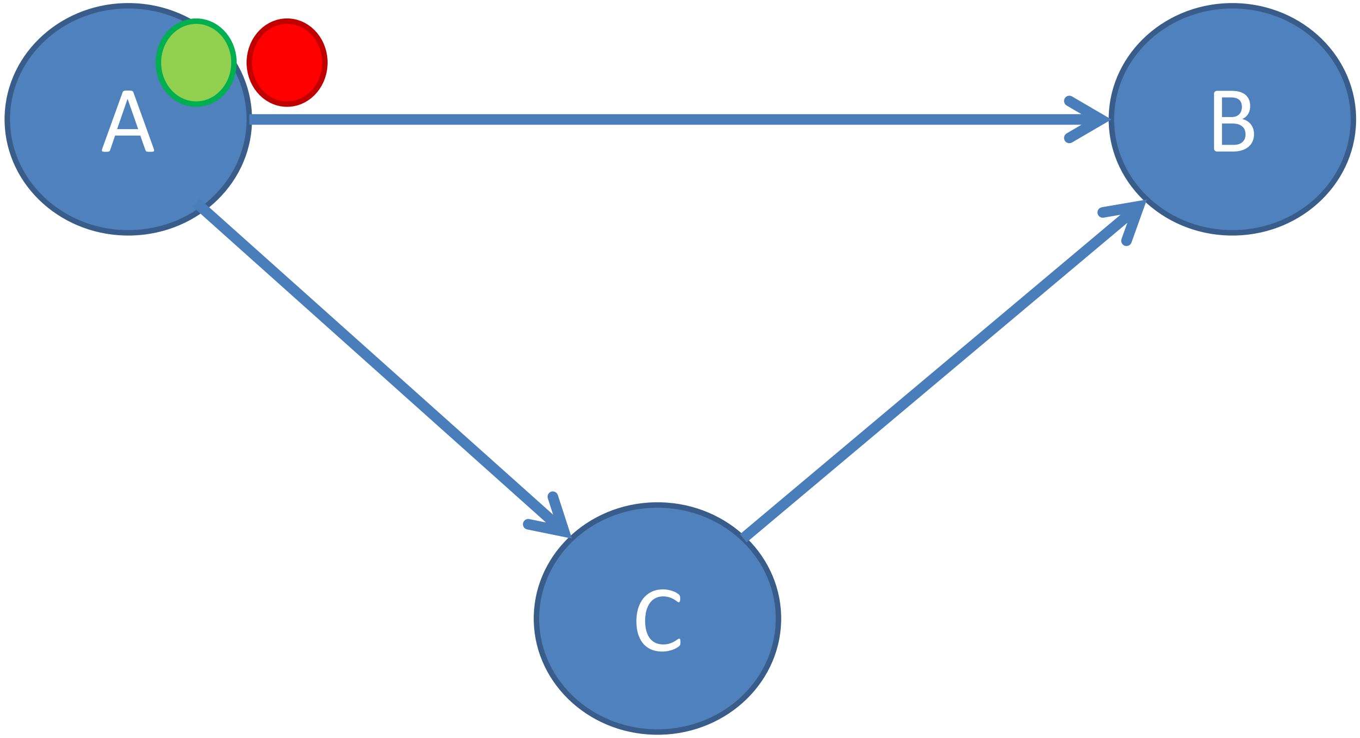
```
psort2(D,[X|Xs]) ->
    Parent = self(),
    spawn_link(fun() ->
      Parent ! …
       end),
    psort2(D-1,[Y || Y <- Xs, Y < X]) ++
    [X] ++
    receive Ys -> Ys end.
```

What's going on?

```
psort2(D,[X|Xs]) ->
    Parent = self(),
    spawn_link(fun() ->
      Parent ! …
        end),
    Parent = self(),
    spawn_link(fun() ->
      Parent ! …
        end),
    psort2(D-2,[Y || Y <- Xs, Y < X]) ++
    [X] ++
    receive Ys -> Ys end ++
    [X] ++
    receive Ys -> Ys end.
```

Message Passing Guarantees

Message Passing Guarantees



Tagging Messages Uniquely

```
Ref = make_ref()
```

- Create a globally unique reference

```
Parent ! {Ref,Msg}
```

- Send the message tagged with the reference

```
receive {Ref,Msg} -> ... end
```

- Match the reference on receipt… picks the right message from the mailbox

A correct parallel sort

```
psort3(Xs) ->
    psort3(5,Xs).

psort3(0,Xs) ->
    qsort(Xs);
psort3(_,[]) ->
    [];
psort3(D,[X|Xs]) ->
    Parent = self(),
    Ref = make_ref(),
    spawn_link(fun() ->
      Parent ! {Ref,psort3(D-1,[Y || Y <- Xs, Y >= X])}
    end),
    psort3(D-1,[Y || Y <- Xs, Y < X]) ++
      [X] ++
      receive {Ref,Greater} -> Greater end.
```

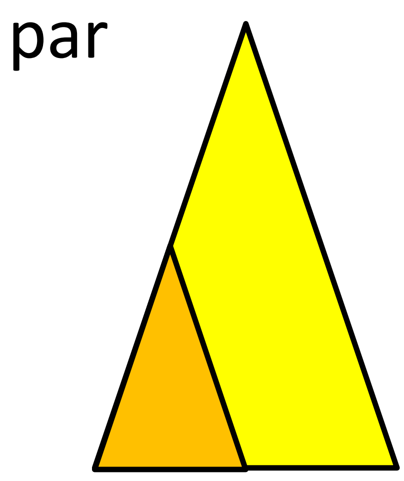
Tests

```
23> foo:benchmark(qsort,L).
285.16
24> foo:benchmark(psort3,L).
92.43
25> foo:qsort(L) == foo:psort3(L).
true
```

- A 3x speedup, and now it works ☺

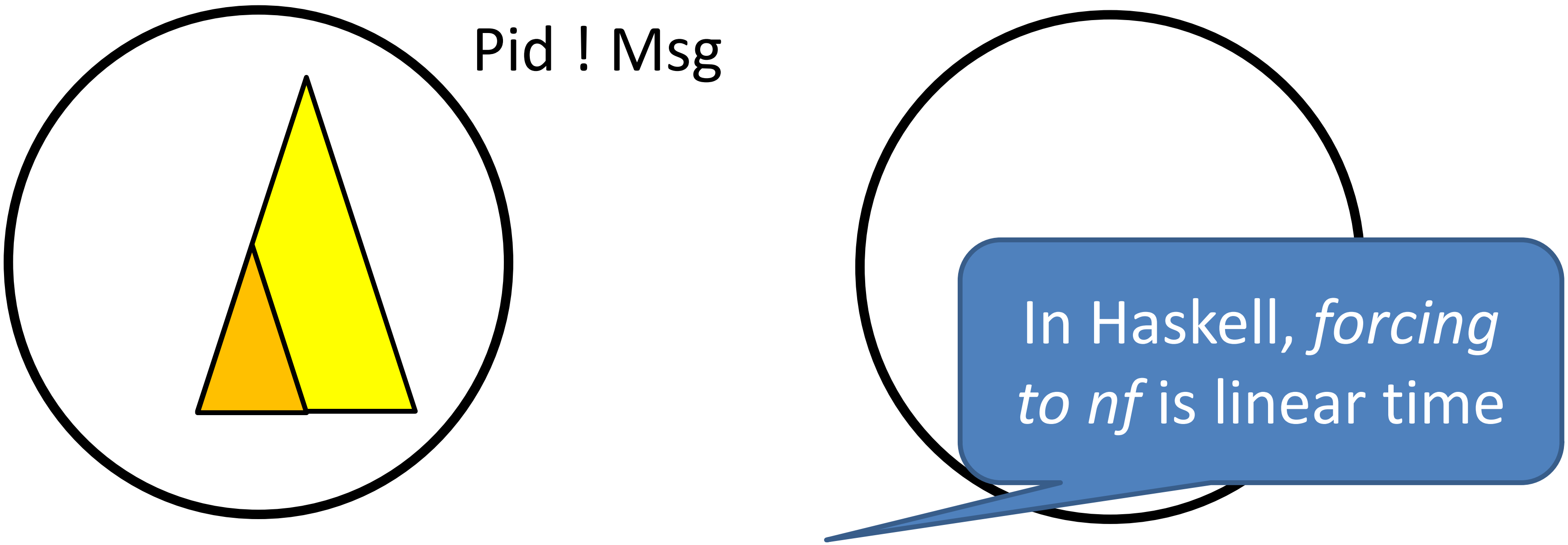
Parallelism in Erlang vs Haskell

- Haskell processes *share memory*



Parallelism in Erlang vs Haskell

- Erlang processes each have their own heap



- Messages have to be *copied*
- No global garbage collection—each process collects its own heap

What's copied here?

```
psort3(D,[X|Xs]) ->
    Parent = self(),
    Ref = make_ref(),
    spawn_link(fun() ->
      Parent !
        {Ref,
         psort3(D-1,[Y || Y <- Xs, Y >= X])}
    end),
```

- Is it sensible to copy *all of Xs* to the new process?

Better

A small improvement—but Erlang lets us *reason* about copying

```
psort4(D,[X|Xs]) ->
    Parent = self(),
    Ref = make_ref(),
    Grtr = [Y || Y <- Xs, Y >= X],
    spawn_link(fun() ->
       Parent ! {Ref,psort4(D-1,Grtr)}
    end),
```

```
31> foo:benchmark(psort3,L).
92.43
32> foo:benchmark(psort4,L).
87.23
```

3,2x speedup on 4 cores (8 threads, parallel depth increased to 8).

Haskell vs Erlang

- Sorting (different) random lists of 200K integers, on 2-core i7

	Haskell	Erlang
Sequential sort	353 ms	312 ms
Depth 5 //el sort	250 ms	153 ms

- *Despite* Erlang running on a VM!

Erlang scales much better

Erlang Distribution

- Erlang processes can run on *different machines* with the same semantics
- No shared memory between processes!
- Just a little slower to communicate…

Named Nodes

```
werl -sname baz
```

- Start a node with a *name*

```
(baz@HALL)1> node().
baz@JohnsTablet2012
```

Node name is an atom

```
(baz@HALL)2> nodes().
[]
```

List of connected nodes

Connecting to another node

```
net_adm:ping(Node).
```

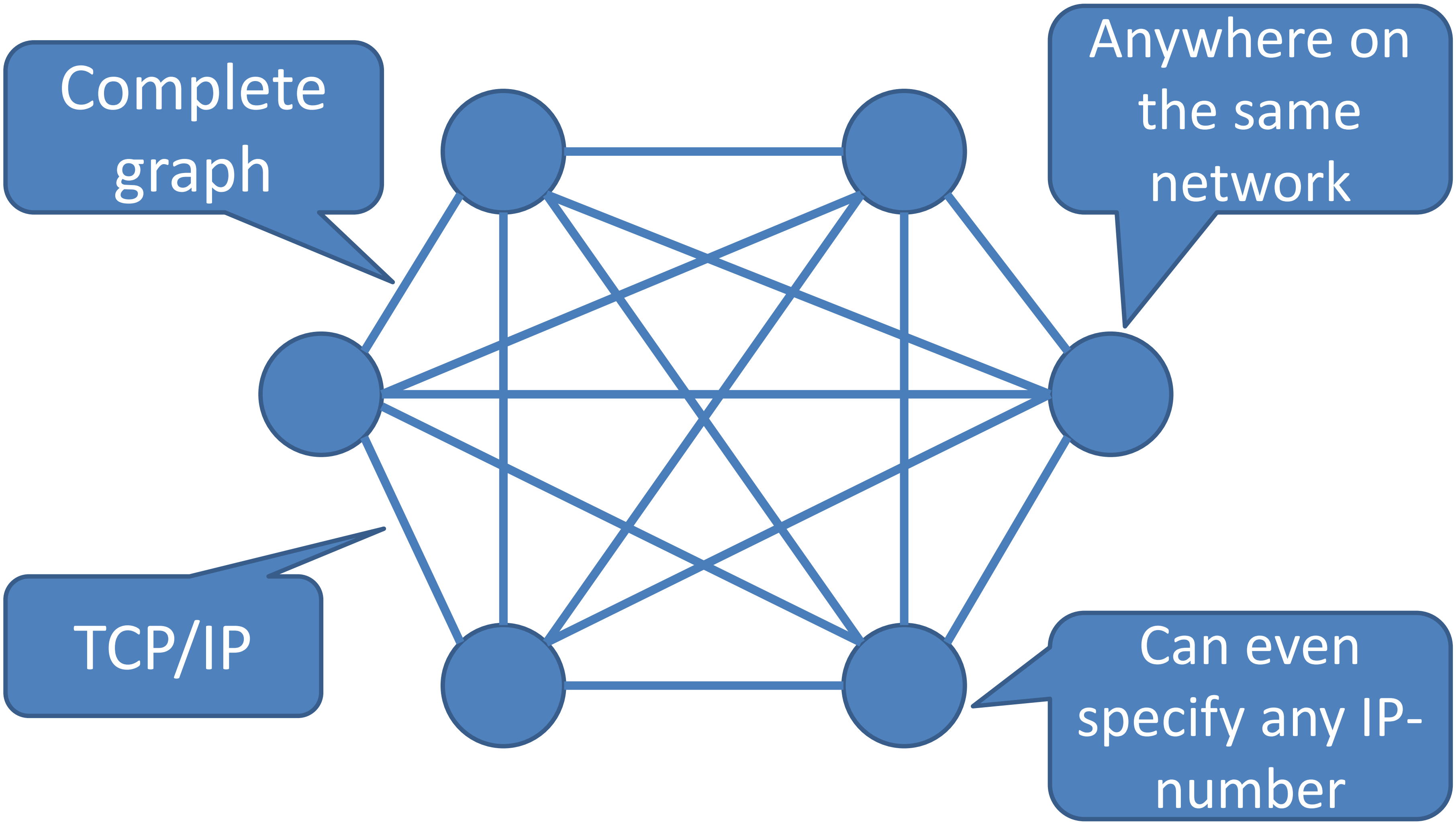
```
3> net_adm:ping(foo@HALL).
pong
```

Success—pang means connection failed

```
4> nodes().
[foo@HALL,baz@JohnsTablet2014]
```

Now connected to foo and other nodes foo knows of

Node connections



Gotcha! the Magic Cookie

- All communicating nodes must share the same *magic cookie* (an atom)

- Must be the same on all machines
  - By default, randomly generated on each machine

- Put it in $HOME/.erlang.cookie
  - E.g. cookie

A Distributed Sort

```
dsort([]) ->
    [];
dsort([X|Xs]) ->
    Parent = self(),
    Ref = make_ref(),
    Grtr = [Y || Y <- Xs, Y >= X],
    spawn_link(foo@JohnsTablet2012,
        fun() ->
             Parent ! {Ref,psort4(Grtr)}
        end),
    psort4([Y || Y <- Xs, Y < X]) ++
     [X] ++
     receive {Ref,Greater} -> Greater
end.
```

Benchmarks

```
5> foo:benchmark(psort4,L).
87.23
6> foo:benchmark(dsort,L).
109.27
```

- Distributed sort is *slower*
  - Communicating between nodes is slower
  - Nodes on the same machine are sharing the cores anyway!

OK…

A 2-core laptop… silly to send it half the work

```
dsort2([X|Xs]) ->
    …
    spawn_link(baz@JohnsTablet2014,
      fun() ->
          ….
```

```
5> foo:benchmark(psort4,L).
87.23
6> foo:benchmark(dsort,L).
109.27
7> foo:benchmark(dsort2,L).
1190.33
```

Distribution Strategy

- Divide the work into 32 chunks on the master node
- Send *one chunk at a time* to each node for sorting
  - Slow nodes will get fewer chunks
- Use the fast parallel sort on each node

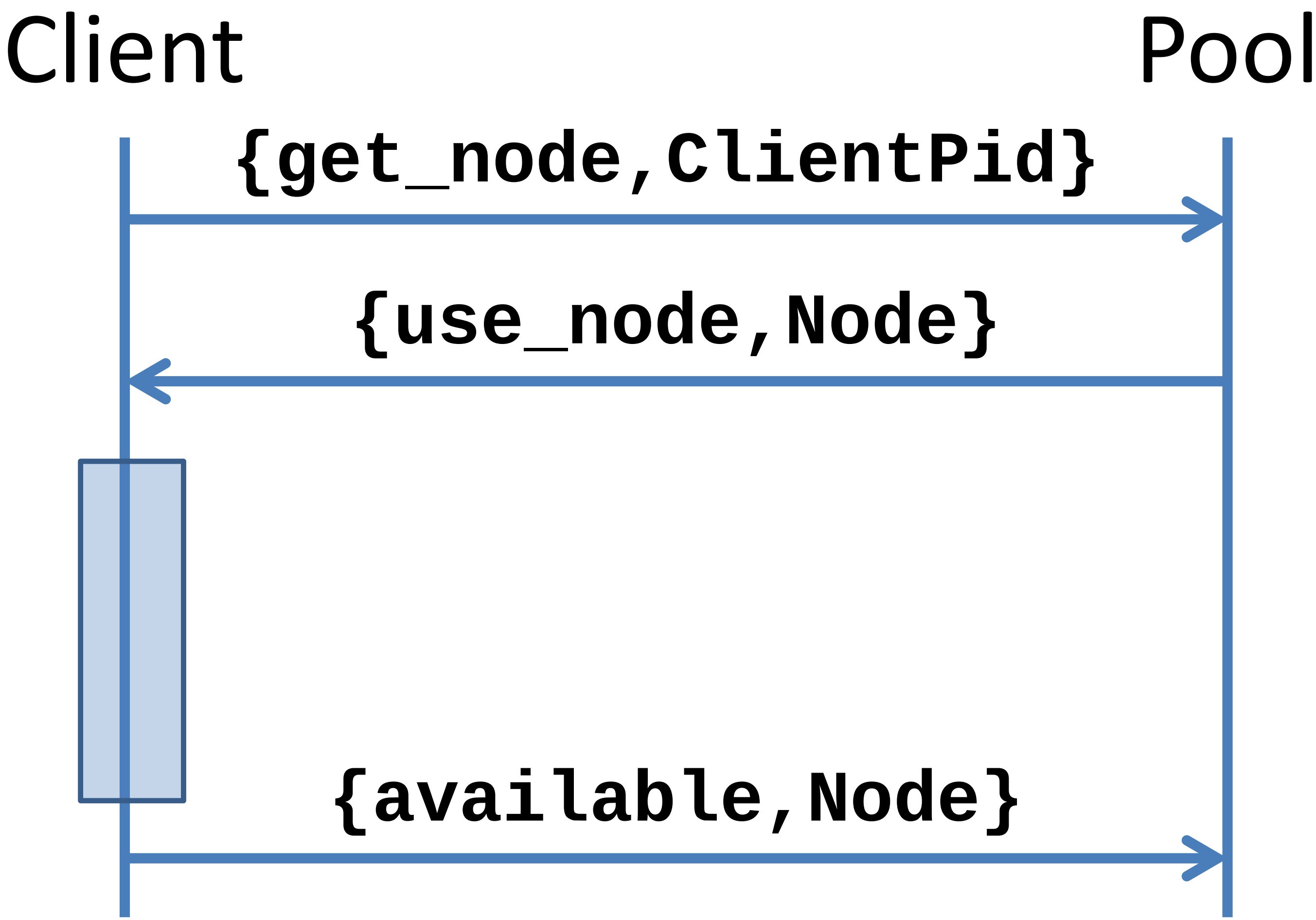
Node Pool

- We need a pool of *available nodes*

```
pool() ->
    Nodes = [node()|nodes()],
    spawn_link(fun() ->
      pool(Nodes)
    end).
```

- We create a process to manage the pool, initially containing all the nodes

Node Pool Protocol



Node Pool Behaviour

```
pool([]) ->
  receive
    {available,Node} ->
        pool([Node])
  end;
pool([Node|Nodes]) ->
  receive
    {get_node,Pid} ->
        Pid ! {use_node,Node},
        pool(Nodes)
  end.
```

If the pool is empty, wait for a node to become

If nodes are available, wait for a request and give one out

Selective receive is really useful!

dwsort

```
dwsort(Xs) -> dwsort(pool(),5,Xs).

dwsort(_,_,[]) -> [];
dwsort(Pool,D,[X|Xs]) when D > 0 ->
    Grtr = [Y || Y <- Xs, Y >= X],
    Ref = make_ref(),
    Parent = self(),
    spawn_link(fun() ->
      Parent ! {Ref,dwsort(Pool,D-1,Grtr)}
    end),
    dwsort(Pool,D-1,[Y || Y <- Xs, Y < X]) ++
      [X] ++
      receive {Ref,Greater} -> Greater end;
```

Parallel recursion to depth 5

dwsort

```
dwsort(Pool,0,Xs) ->
    Pool ! {get_node,self()},
    receive
      {use_node,Node} ->
         Ref = make_ref(),
         Parent = self(),
         spawn_link(Node, fun() ->
            Ys = psort4(Xs),
            Pool ! {available,Node},
            Parent ! {Ref,Ys}
         end),
         receive {Ref,Ys} -> Ys end
    end.
```

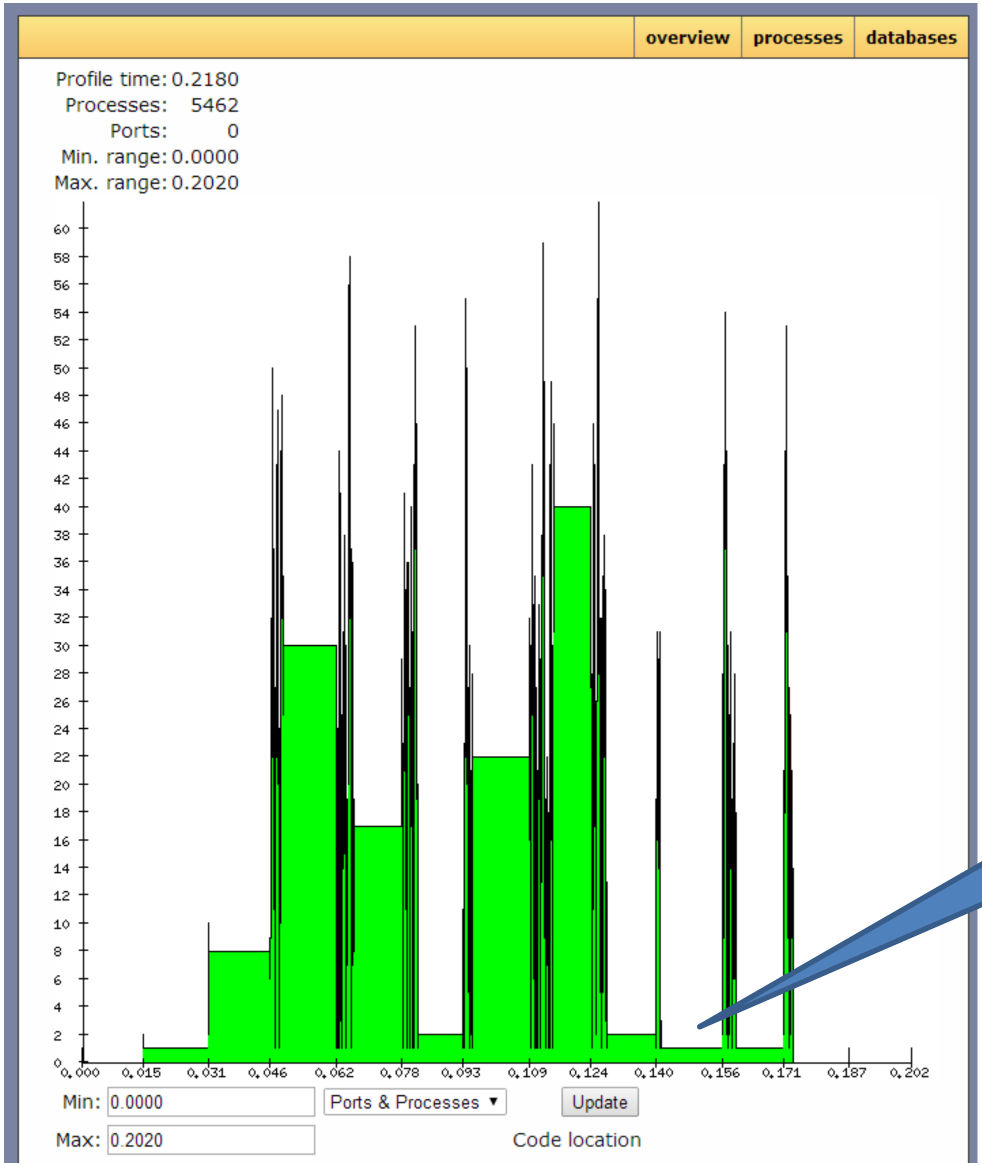
A further optimisation: if we should use the *current* node, don’t spawn a new process

Benchmarks

```
(baz@HALL)17> foo:benchmark(qsort,L).
271.97
(baz@HALL)18> foo:benchmark(psort4,L).
88.65
(baz@HALL)19> foo:benchmark(dsort2,L).
1190.33
(baz@HALL)20> nodes().
[baz@JohnsTablet2014]
(baz@HALL)21> foo:benchmark(dwsort,L).
295.59
(baz@HALL)22> foo:benchmark(dwsort2,L).
195.05
```

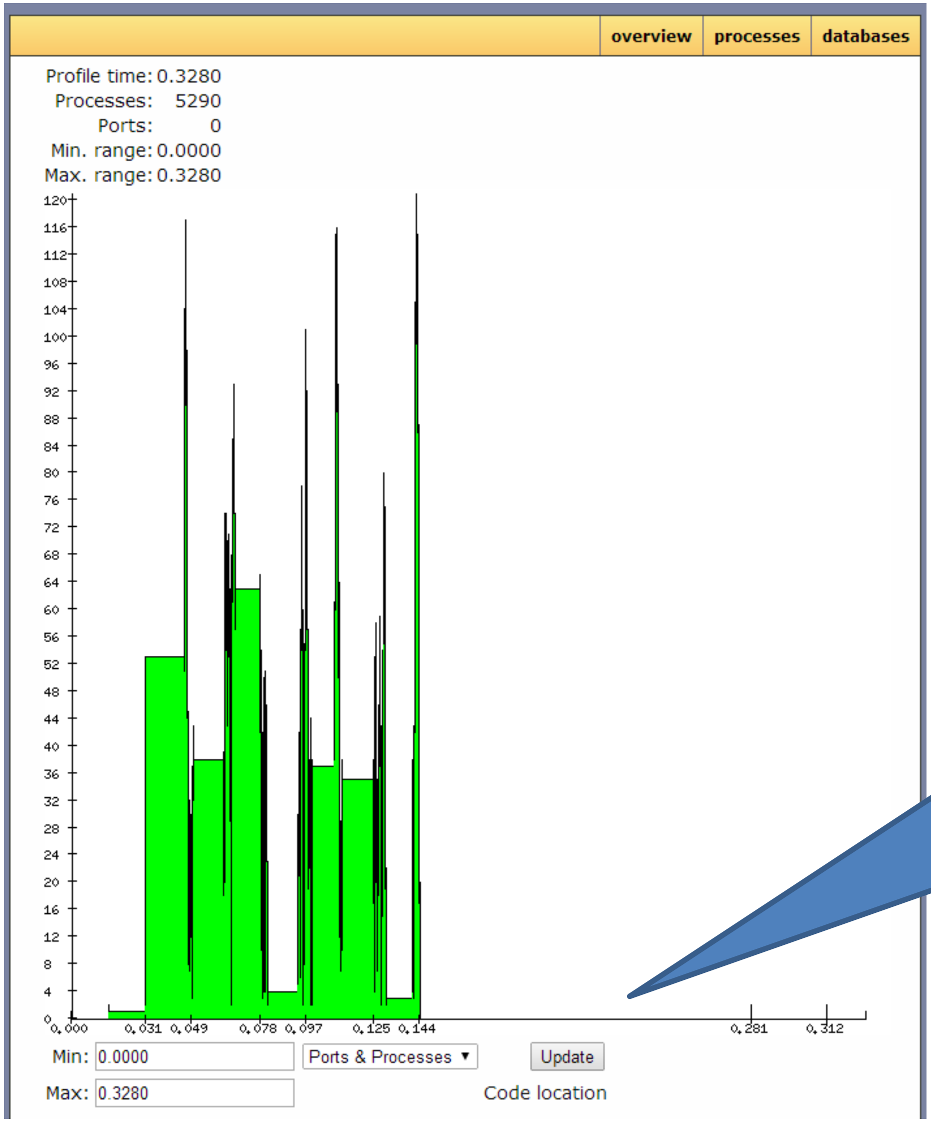
With each node in the pool *twice*, to overlap communication and computatíon

dwsort



Lots of time with only one or two runnable processes

dwsort2



Better parallelism on the local node, followed by a long wait for remote results to come back!

Oh well!

- It’s quicker to *sort* a list, than to send it to another node and back!

Another Gotcha!

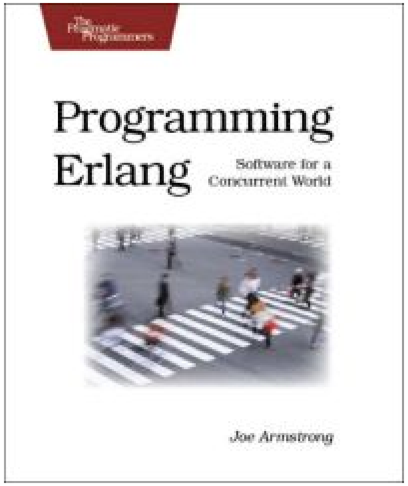
- All the nodes must be running *the same code*
  - Otherwise sending functions to other nodes cannot work
- **`nl(Mod)`** loads the module on *all* connected nodes.

Summary

- Erlang parallelism is more explicit than in Haskell
- Processes do not share memory
- All communication is explicit by message passing
- Performance and scalability are strong points
- Distribution is easy
  - (But sorting is cheaper to do than to distribute☹)

References

- *Programming Erlang: Software for a Concurrent World*, Joe Armstrong, Pragmatic Bookshelf, 2007.



- *Learn you some Erlang for Great Good*, Frederic Trottier-Hebert , http://learnyousomeerlang.com/