















Processes

HIPE:

ERTS as components:

The Garbage Collector I/O

The BEAM interpreter

The Scheduler

















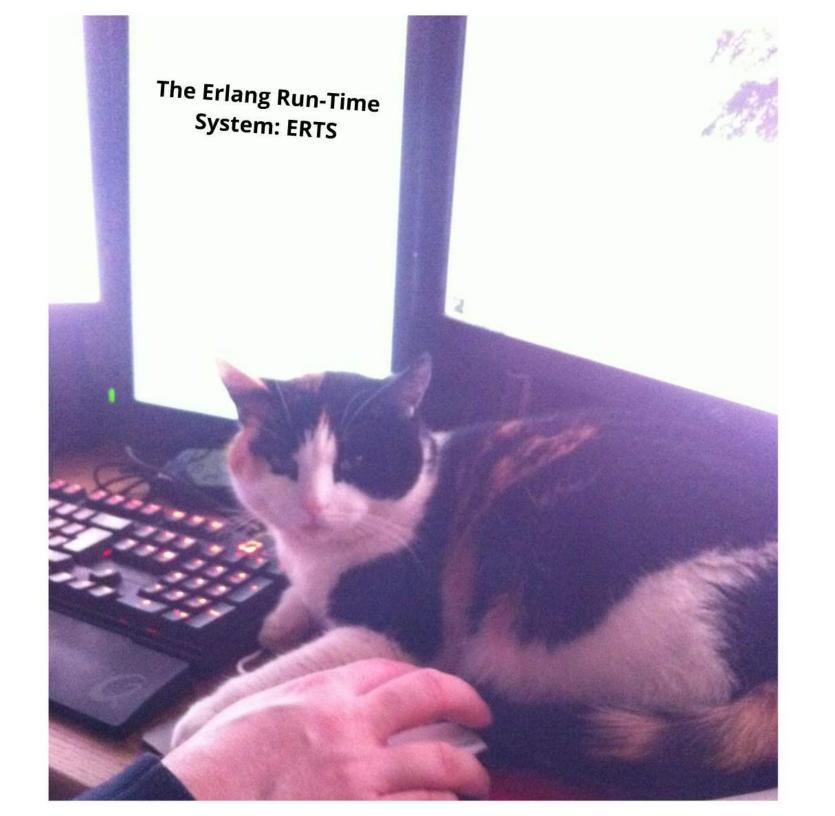














The right concurrency model.

Error & exception handling done right
Good libraries for the hard stuff



Dynamically typed.

Symbolic and transparent data structures

An interactive shell

Scalable

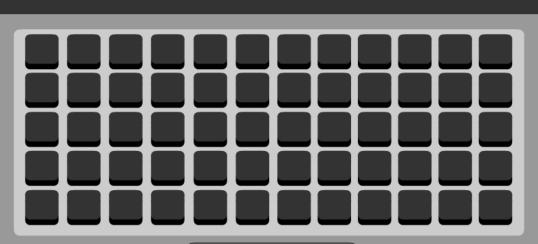
The right concurrency model.
Good libraries for the hard stuff
Weird but efficient strings for I/O

The right concurrency model

Lightweight processes Message passing Share nothing semantics Don't stop the world GC Monitors & Signals







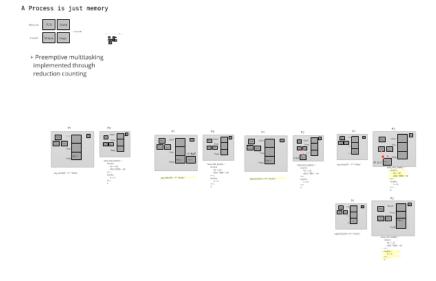
Dynamically & Strongly Typed
Typed
Symbolic and transperent data structures

Evolve:
Her. code landing
Missale hound a cacke
Transper Carbon Cacke
Transper Cacke
T

Remember I don't have opinions this is all facts.

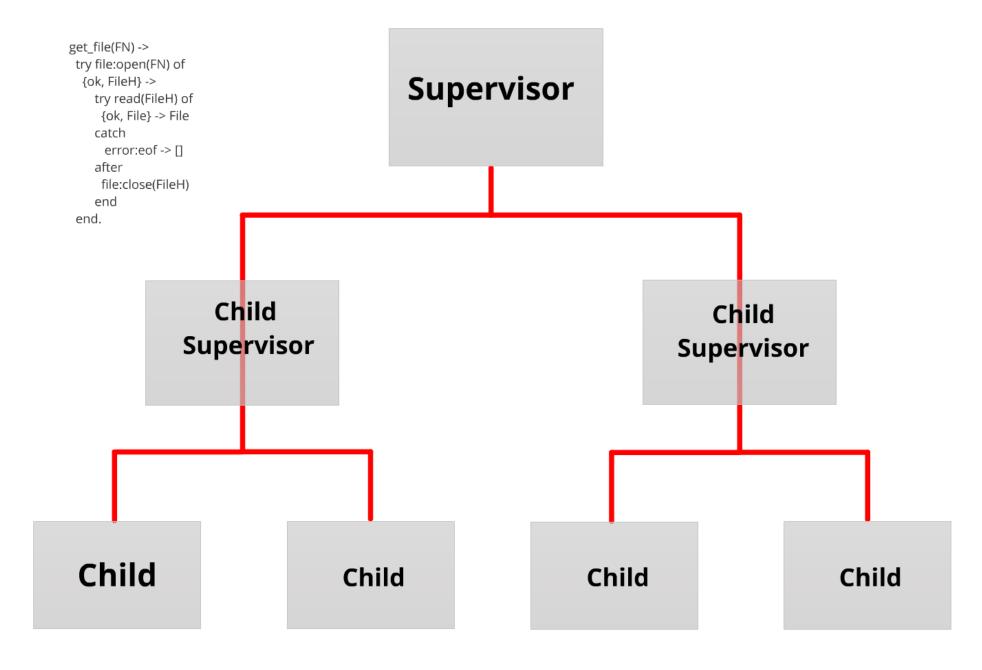
The right concurrency model

Lightweight processes
Message passing
Share nothing semantics
Don't stop the world GC
Monitors & Signals





Error & exception handling done right

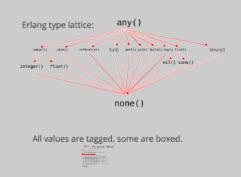


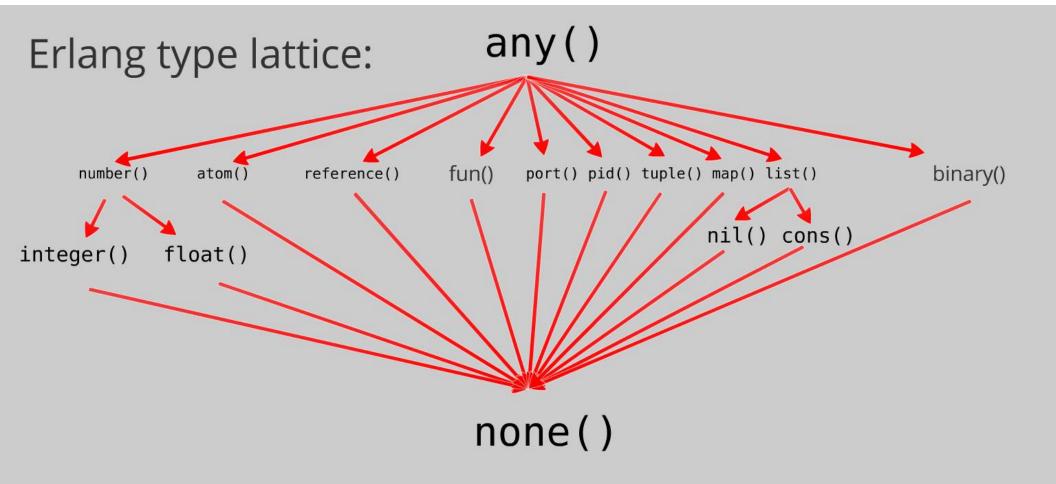
Dynamically & Strongly Typed

Symbolic and transparent data structure

Enables:

Hot code loading
Movable heaps & stacks
Transparency of data
Garbage Collection





All values are tagged, some are boxed.



What is ERTS?

ERTS is the Erlang Runtime System.



ERTS as components:

The BEAM interpreter













The Garbage Collector









Processes

A Stack A Heap

A Mailbox

A Process Control Block

A PID

HiPE 1/0

Processes

Conceptually: 4 memory areas and a pointer:

A Stack

A Heap

A Mailbox

A Process Control Block

A PID

ERTS as memory areas:

ERTS CODE

GC Scheduler **BEAM** Sockets etc...

The Tag Scheme

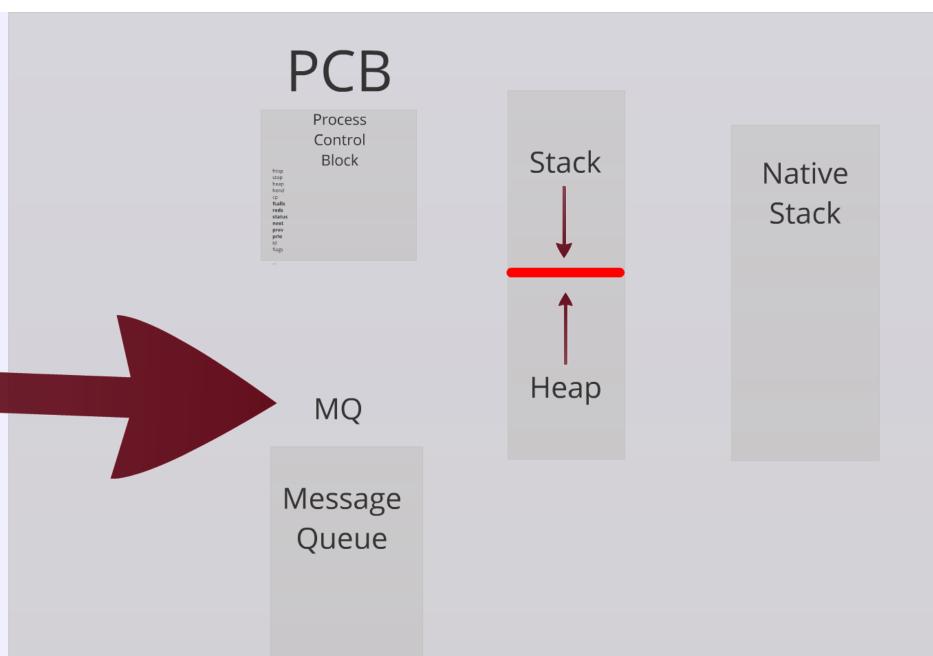
Processes C-stack P2 Queues **P3**

Binaries

P99

BEAM CODE

p2() ->L = "Hello", $T = \{L, L\},\$ P3 = mk proc(),P3 ! T.



See: [OTP]/erts/emulator/beam/erl_process.h

Process Control Block

htop

stop

heap

hend

ср

fcalls

reds

status

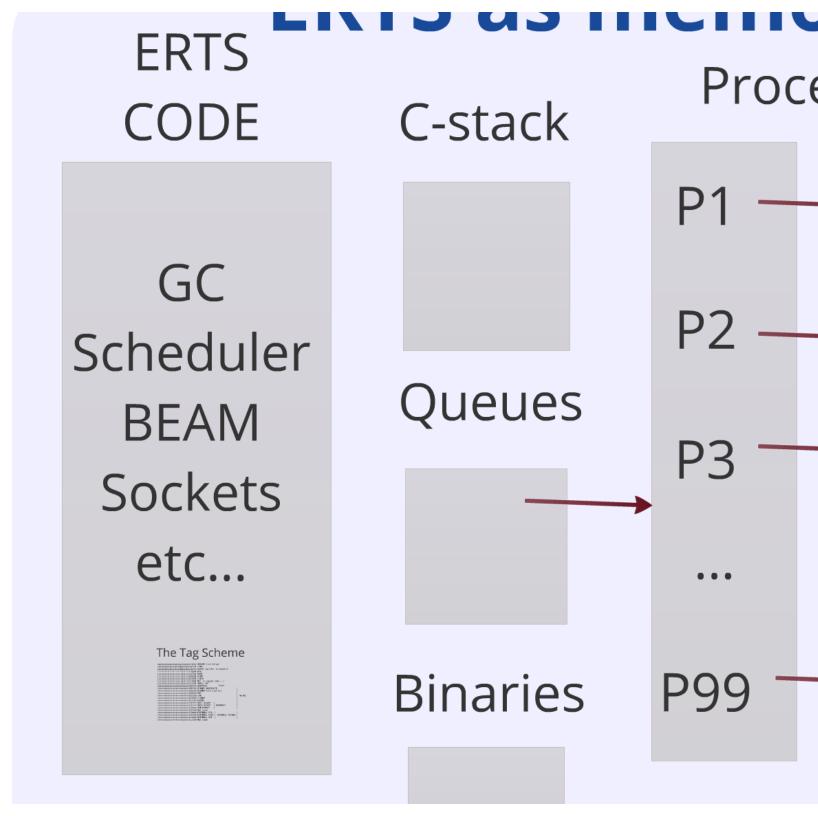
next

prev

prio

id

flags



The Tag Scheme

```
aaaaaaaaaaaaaaaaaaaaaatttt00 HEADER (see below)
ppppppppppppppppppppppppppp01 CONS
1111111111111111111111111111111110011
iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii0111 PORT
iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii
iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii
                                       Tuple
aaaaaaaaaaaaaaaaaaaaaaaa000000 ARITYVAL
VVVVVVVVVVVVVVVVVVVVVVVVVVVVO00100 BINARY AGGREGATE
vvvvvvvvvvvvvvvvvvvvvv001x00 BIGNUM with sign bit
VVVVVVVVVVVVVVVVVVVVVVVVVV010000 REF
vvvvvvvvvvvvvvvvvvvvvvv010100 FUN
                                                     THINGS
VVVVVVVVVVVVVVVVVVVVVVVVVVVVV011000 FLONUM
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVV011100 EXPOR
BINARIES
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVI01000 SUB BINARY
vvvvvvvvvvvvvvvvvvvvvvv101100 Not used
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVI10000 EXTERNAL PID
                                       EXTERNAL THINGS
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV110100 EXTERNAL PORT
VVVVVVVVVVVVVVVVVVVVVVVVVVVVIII1000 EXTERNAL REF
vvvvvvvvvvvvvvvvvvvvvvvv111100 Not used
```

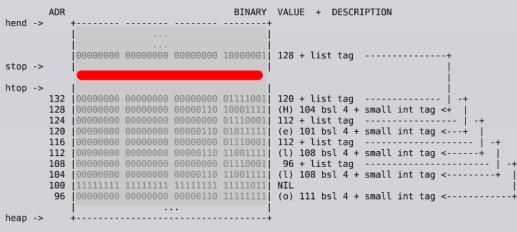
```
An example
The string "Hello", i.e. the list
[104, 101, 108, 108, 111]
```

PCB

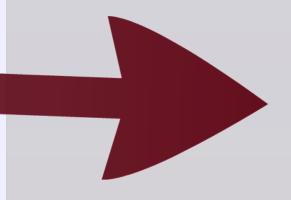
Process Control Block

stop heap hend tp fcalls reds id flags next prev

Stack

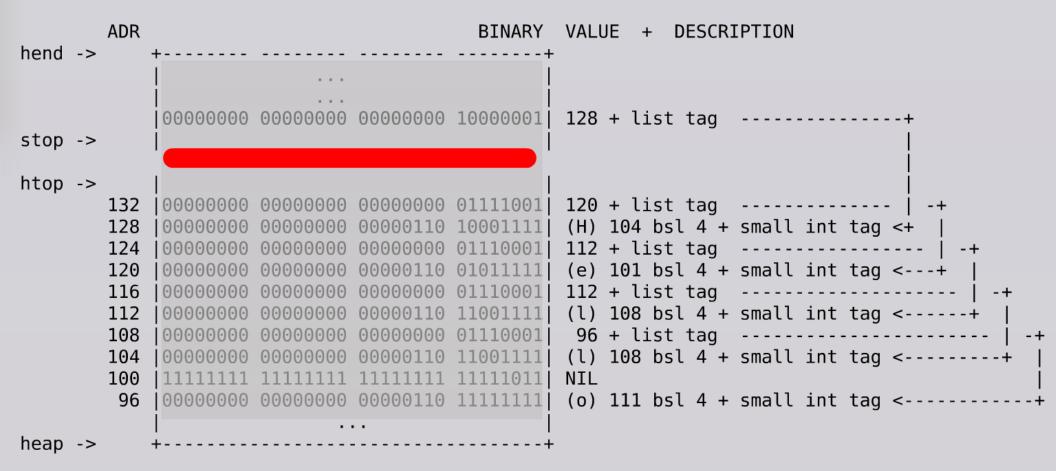


Heap



MQ

Stack



Heap

ERTS as memory areas:

ERTS CODE

GC Scheduler **BEAM** Sockets etc...

The Tag Scheme

Processes C-stack P2 Queues **P3**

Binaries

P99

BEAM CODE

p2() ->L = "Hello", $T = \{L, L\},\$ P3 = mk proc(),P3 ! T.

ERTS as components:

The BEAM interpreter













The Garbage Collector









Processes

A Stack A Heap

A Mailbox

A Process Control Block

A PID

HiPE 1/0

BEAM

- Garbage Collecting-
- Reduction Counting-
- · Non-preemptive-
- Directly Threaded-
- Register-
- Virtual-











-Machine

Memory Management:

Garbage Collection

- On the Beam level the code is responsible for:
 - checking for stack and heap overrun.
 - allocating enough space
- "test_heap" will check that there is free heap space.
- If needed the instruction will call the GC.
- The GC might call lower levels of the memory subsystem to allocate or free memory as needed.

Scheduling:

Non-preemptive, Reduction counting

- Each function call is counted as a reduction
- Beam does a test at function entry: if (reds < 0) yield
- A reduction should be a small work item
- Loops are actually recursions, burning reductions

A process can also yield in a receive.

Dispatch: Directly Threaded Code

The dispatcher finds the next instruction to execute.

#define Arg(N) (Eterm *) I[(N)+1] I: 0x1000 #define Goto(Rel) goto *((void *)Rel) beam_emu.c **: Loaded code*: OpCase(move_xx): { External beam format: 0x3000: x(Arg(1)) = x(Arg(0)): 0x1000: 0x3000 I += 3; $\{move,\{x,0\},\{x,1\}\}$ 0x1004: 0x0 Goto(*I); $\{move, \{y, 0\}, \{x, 0\}\}.$ 0x1008: 0x1 $\{move, \{x, 1\}, \{y, 0\}\}.$ 0x100c: 0x3200 OpCase(move_yx): { 0x3200: x(Arg(1)) = y(Arg(0)); 0x1010: 0x1 I += 3; 0x1014: 0x1 Goto(*I); 0x1018: 0x3100 *This is a lie... beam actually rewrites the 0x101c: 0x1 external format to different internal OpCase(move_xy): { instructions.... 0x3100: y(Arg(1)) = x(Arg(0)); 0x1020: 0x1 | += 3;Goto(*I);

BEAM

- Garbage Collecting-
- Reduction Counting-
- · Non-preemptive-
- Directly Threaded-
- Register-
- Virtual-







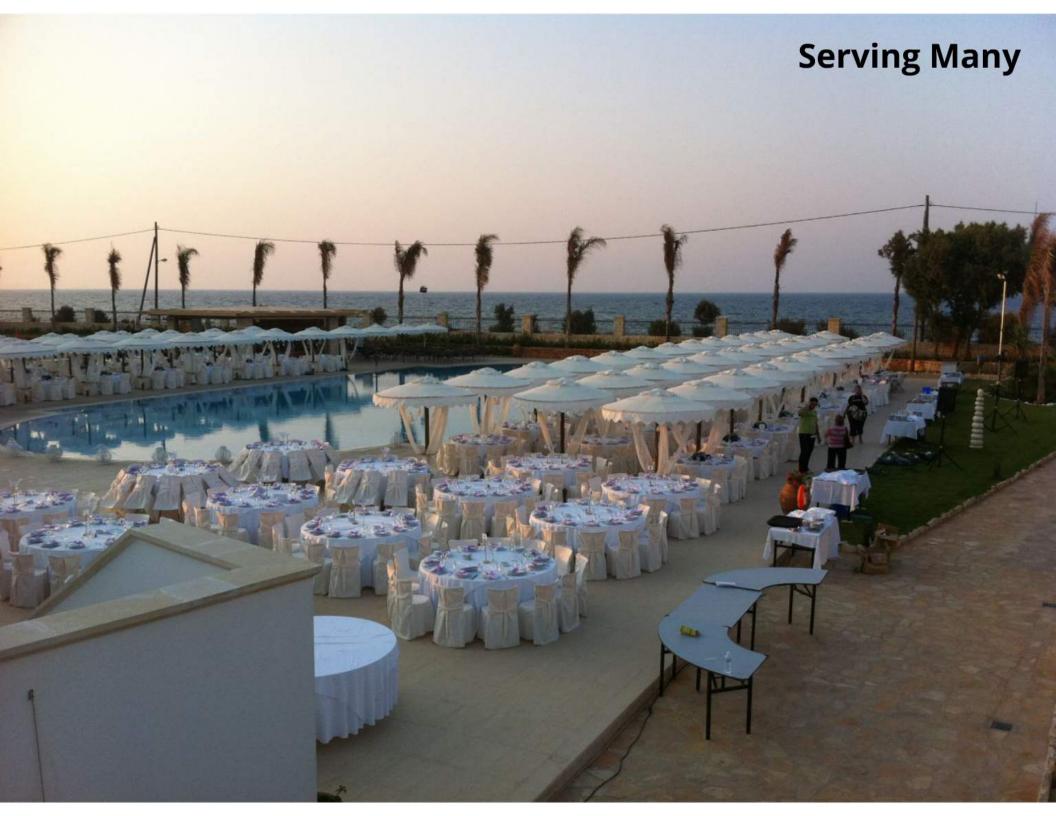




-Machine









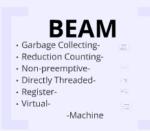














BEAM is a register machine. It has two sets of registers: x and y

- x registers are caller save and arguments.
 y registers are callee save and actually the stack.

You can look at beam code by giving the 'S' flog to the compiler:

c(test, ['S']).

See: [DTP]/erts/emulator/beam/beam_emu.c

The Scheduler



Process State Reductions Que Handling Timing wheels

Possible Problems

erl process.c schedule()

Lukas: "It is quite short and not hard to understand if you know C".

beam_emu.c process_main()

schedule() 1. Update reduction counters

- 2. Check triggered timers
- 3. If check balance reds > 4,000,000 check balance
- 4. Possibly migrate processes+ports
- 5. Execute scheduller work (load, free, trace, etc) 6. If function_calls > 4000 check IO, update time
- 7. Execute 1 to N ports for 2000 reds

Reduction count problems

BIFs uses an arbitrary amount of reductions. Seal the market Da read to show that there

A return does not use any reductions.

NIFs uses an arbitrary amount of

reductions. Should be worried? The Doll need to know about this? Yes.

Load Balancing

Load balancing operations are calculated when a scheduler has done 4.000,000 reductions.

Processes will normally migrate towards lower schedulers if there is no overload.

If a scheduler is overloaded procesesses are evicted to other schedulers.

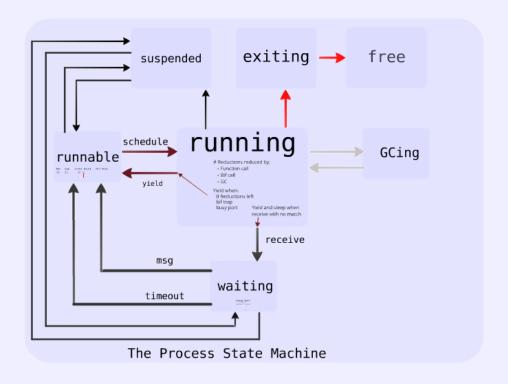
If reduction counting is messed up, starvation might occur.

Use: +sfwi to wake up sleeping schedulers

The Garbage C

One Scheduler Per Core

Schedulers Ready Q Running **Cores**



Process State Reductions Que Handling Timing wheels

Possible Problems

Priority Inversion

Should I be worried?

Do I need to know about this?

What can I do?

No No Don't mess with priorities

runing

- # Reductions reduced by:
 - Function call
 - Bif call
 - GC

Yield when:

0 Reductions left

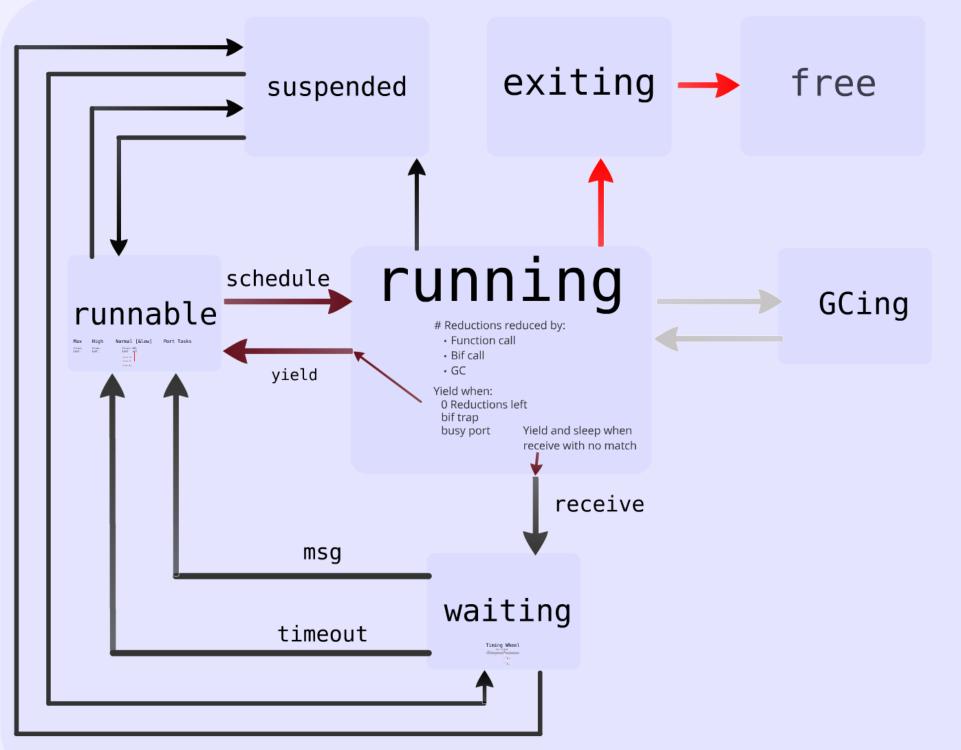
bif trap

busy port

Yield and sleep when

receive with no match





The Process State Machine

runable

Max

High

Normal [&low]

Port Tasks

First: Last:

First: Last:

First: &P1 Last: &P3 P1 next: &P2 P2 next: &P3 P3 next: NULL

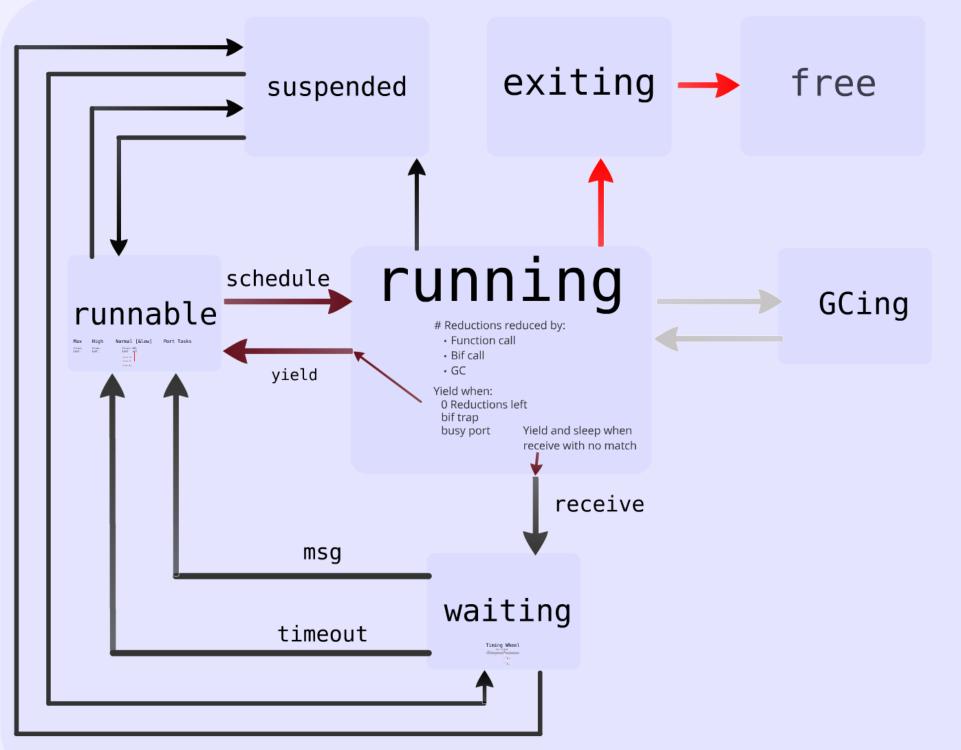
Normal [&low]

```
First: &P1
Last: &P3
```

P1 next: &P2

P2 next: &P3

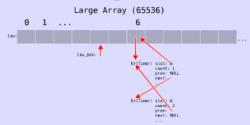
P3 next: NULL



The Process State Machine

Waiting

Timing Wheel

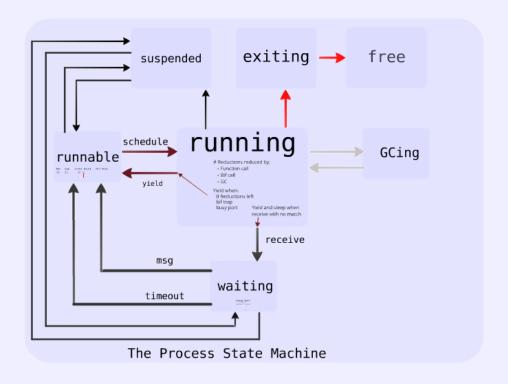


Timing Wheel

Large Array (65536)

6 0 tiw: tiw pos: ErlTimer: slot: 6 count: 1 prev: NULL next: ErlTimer: slot: 6 count 2 prev: next: NULL

. . .



Process State Reductions Que Handling Timing wheels

Possible Problems

Priority Inversion

Should I be worried?

Do I need to know about this?

What can I do?

No No Don't mess with priorities

erl_process.c schedule()

Lukas: "It is quite short and not hard to understand if you know C".

560 lines

beam_emu.c
process_main()

Execute process till yield. call schedule()

schedule()

- 1. Update reduction counters
- 2. Check triggered timers
- 3. If check_balance_reds > 4,000,000 check balance
- 4. Possibly migrate processes+ports
- 5. Execute scheduller work (load, free, trace, etc)
- 6. If function_calls > 4000 check IO, update time
- 7. Execute 1 to N ports for 2000 reds

(More or less stolen from Lukas presentation)

Load Balancing

Load balancing operations are calculated when a scheduler has done 4,000,000 reductions.

Processes will normally migrate towards lower schedulers if there is no overload.

If a scheduler is overloaded procesesses are evicted to other schedulers.

If reduction counting is messed up, starvation might occur.

Use: +sfwi to wake up sleeping schedulers.

Reduction count problems

BIFs uses an arbitrary amount of reductions. Should De Worried?

Do I need to know about this?

What can I do?

Fix the BIF::
Use small data sets, add a call to erlang:bump_reductions;

A return does not use any reductions.

Should I be worried?

Do I need to know about this?

What can I do?

One tail recursion, don't have insanely long callchains

NIFs uses an arbitrary amount of reductions. Should I be worried? Yes.

Do I need to know about this? Yes.

What can I do?

Don't use NIFs ;)
Make sure your NIFs are yielding and using reductions.
Wait for "dirty schedulers".

n arbitrary amo

Should I be worried?

Do I need to know about this?

What can I do?

Yes.

Yes.

Fix the BIF;)
Use small data sets, add a call to erlang:bump_reductions()

es not use any

Reduction count problems

BIFs uses an arbitrary amount of reductions. Should I be worried?

Do I need to know about this?

What can I do?

Fix the BIF.)
Use small data sets, add a call to erlang bump_reductions()

A return does not use any reductions.

Should I be worried?

Do I need to know about this?

What can I do?

So No
Probably not
Onthave Insanely long callchains

NIFs uses an arbitrary amount of reductions. Should I be worried? Yes.

Do I need to know about this? Yes.

What can I do?

Don't use NIFs ;)

Make sure your NIFs are yielding and using reductions. Wait for "dirty schedulers".

es not use

Should I be worried?

Do I need to know about this?

What can I do?

No

Probably not

Use tail recursion, don't have insanely long callchains.

Do I need to know about this? What can I do?

Use small data sets, add a call to erlang:bump_reductions/

A return does not use any reductions

Should I be worried? Do I need to know about this? Probably not What can I do? don't have insanely long callchains.

NIFs uses an arbitrary amount of reductions. Should I be worried? Yes.

Do I need to know about this? Yes.

What can I do?

Don't use NIFs:)

Make sure your NIFs are yielding and using reductions.

Wait for "dirty schedulers".

ERTS as components:

The BEAM interpreter













The Garbage Collector









Processes

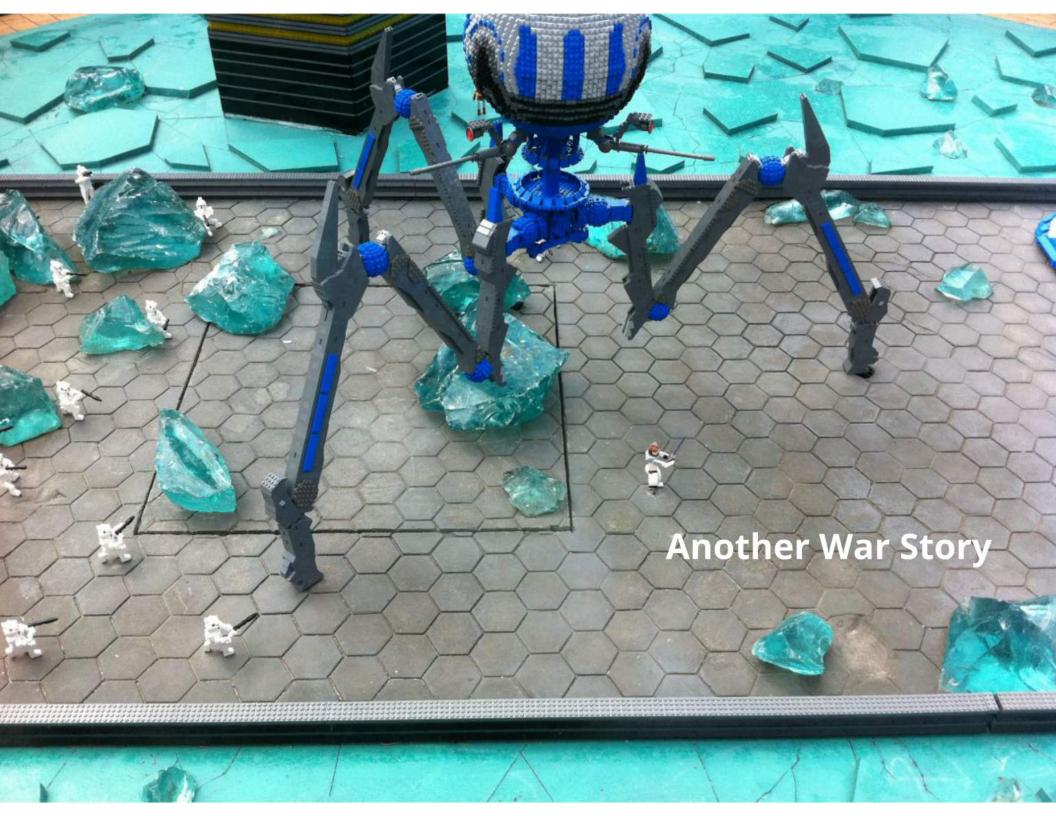
A Stack A Heap

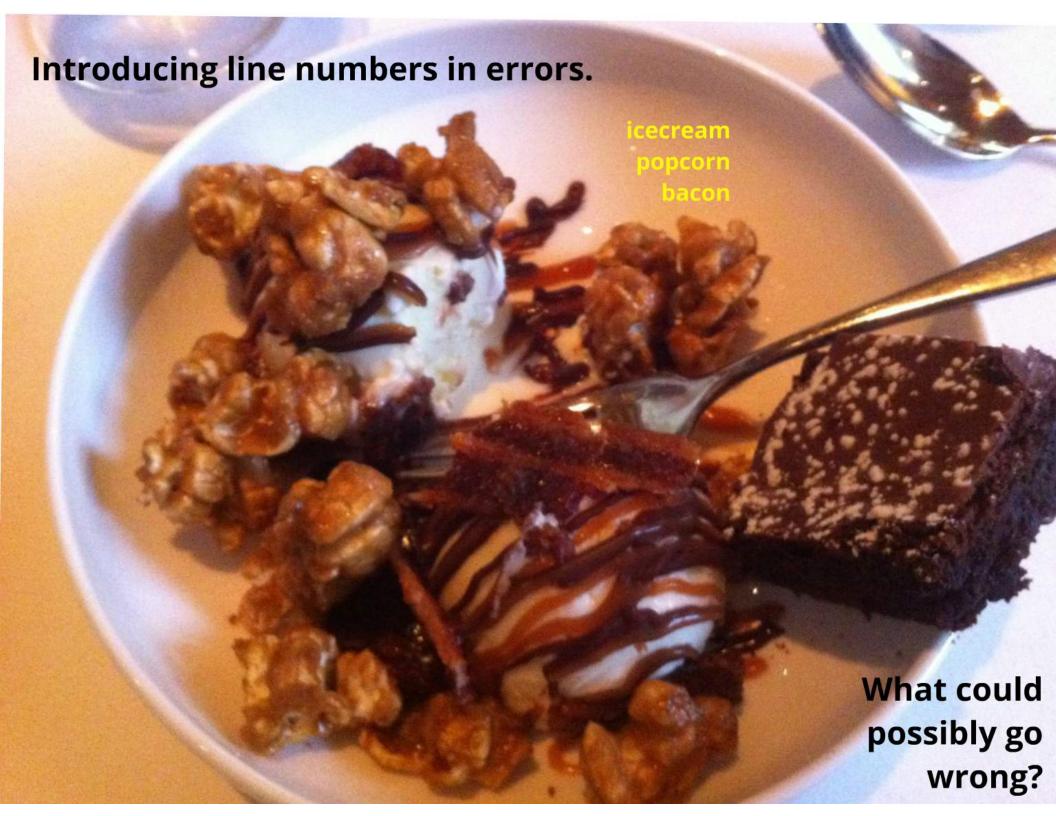
A Mailbox

A Process Control Block

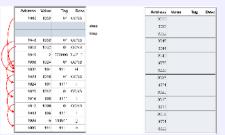
A PID

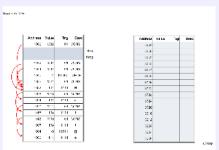
HiPE 1/0





Copying Generational Garbage Collector





	4ddmm	Vri.n	Tag	Desc	Address	Value	Tag	Case	
- 1	1000	2300	21	00.46	18 (4)				
					20.00				
- 1					28.97				
_	12.70	150		11,106	80,46				
<u> </u>	10.11	100		11100	64.6-				
/ E	10.40		10 20 1	19.5	50.40				
ΑЕ	17.96	2360		C5.,98	3036				
V3	17.52	- 5		WWCD	3022				
	1036	1212		CCAR	3010				
$I \perp$	123-	01	111		3632				
V	1230	1010	21	00.46	20.43				
P*	1216	100	111	1	18/99				
и.	15.99	19%	5.	90.99	1877				
m	11.00	***	.11.	1	18 06				r) ka
VI.	0.0		111111	El.	8.5	17.84	- 0	2796	
- 4	17.34	111	1111		5000	94	111	н	

Add	****	Victor	Tr	Case	Address	Value	Tag	Dat
	181	2000	***	1000	9.60			
					300			
					2060			
	OIE	1032		CC35	300			
	044	10.52		cose	3014			
	OIE.		60000		30.61			
	ues.	2020			3002			
	0.7	- 4		M-Y-E	30.0			
	OC.5	2020			900			
	1024	- 4		NETZ-E	80.4			
	020	10.2		CCNC	960			
	1111	244			3015			
	012	1000		CUMC	3012	1016	64	00.96
	+21	7.00			94.5	121	211	
	004	-5		0	3001	2200	64	00395
	280	111	1223		3000	134	111	- 11

17 37 11 12 13 14 15 15 15 15 15 15 15	1										
Dec Col. Dec Col.	100 100		Stres	24.0	Top Devi		Address	Velue	Tay	Direct	1
30 CG 31 CGC 31 CGC 31 CGC 31 CGC 31 CGC 32 31 CGC 33 31 CGC 34 31 CGC 35 31 CGC 35 31 CGC 35 35 CGC 35	10	- 0	1,67	423	01.0065		3000				
0-2 C2 01 C02	9-2 CS 9 CO 2 9-1 CS 9 CO 2 9-2 CS 9 CS 9 9-2 CS 9 CS 9 9-2 CS					ansp.	3236				
Sel. GO St. Chr.	Col.					lakes .	32.22				
1	1	-	1945	C22	OI COR.S	\ \	3610				
1	1	60	1911	1022	01 6262	\ \	14.77				
1	1	ш	5942	- 2	DECK BOSTS	1	20.00				1 170p
15	7: 4 15 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	W	261	100			27.90		Trer	п	
C C K S D C C C C C C C C C	C C C C C C C C C C	13	202	- 4	FIRESE	1	85.53	211	211		
OCT X S OT	007 66 67 70 70 70 70 70		200	414		\rightarrow	37.86	2352		0036	
010 0 CHCCCC	000 XD 01 00 000000 00 000 000 000 000 000 00		1004	- 9	0.00500	- \(\frac{1}{2} \rightarrow \frac{1}{2} \rightarrow \f	228	05	- 11	- 1	
100 CM 100 F 100 CM 100	915 X2 31 11 11 11 11 11 11 11 11 11 11 11 11			X 16		ブプラント	3230	2321	5.	00.95	
	900 0 CHOND / DE 90 90 000 c					リートレントラー	32.6	100	111	- 1	
	/ //		1912	XV	01	7 16	32.9	2316		35.90	
901 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	28 8 8 2		1000	- 0	0 MONTO	/ \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	11.50	20	1111		

Fablures.	College:	143	0.4	Adding	Votes	Tree	Ces
000	2230	04	6246	139			CONT
				226		-11	
				117			
1986				224			_
011				239			
				224			_
	_			227		1113.1	0
				276		1111	
	+			170			CORP
				274			
G:				127			0690
				227			_
:				22:1			0690
				228		1111	
				220			0690
786				220			

Copying Generational Garbage Collector

Advantages with 1 heap/process:

- + Free reclamation when a process dies
- + Small root set
- + Sort of "Incremental"
- + Improved cache locallity
- + Cheap stack/heap test
- + Smal extra footprint during GC

Disadvantages with 1 heap/process:

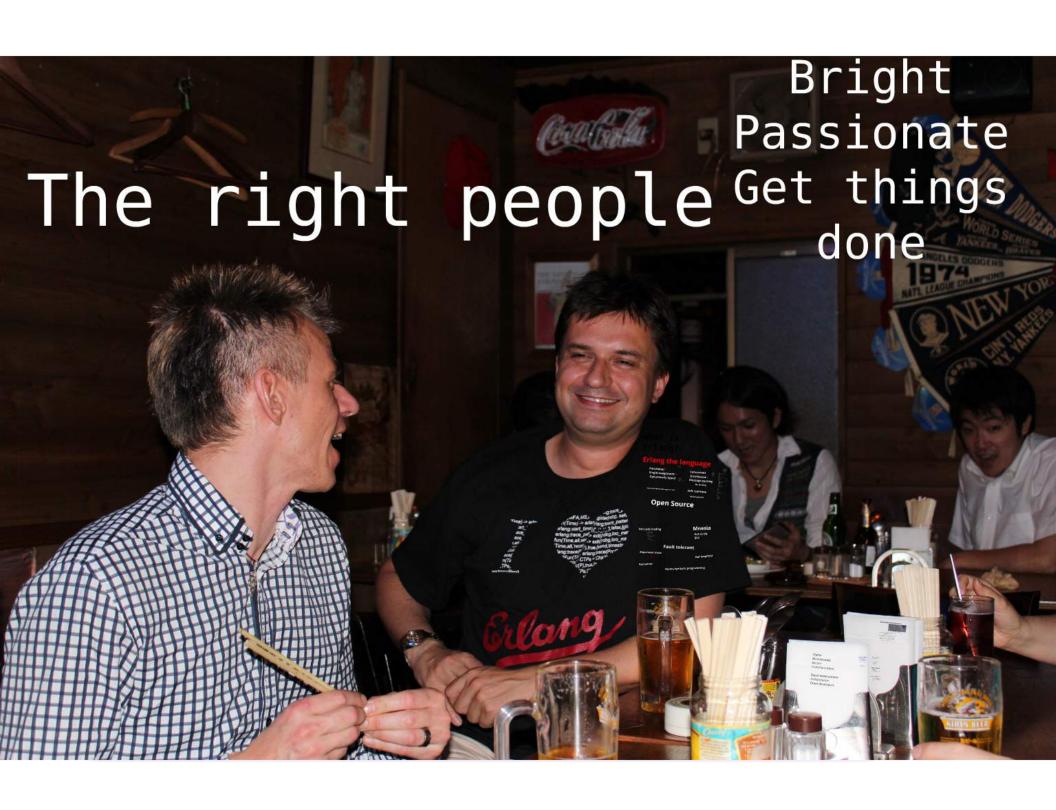
- Message passing is expensive
- Uses more space (fragmentation)

Lessons learned:

- **ERTS the Erlang RunTime System is** the defacto standard implementation of Erlang
- Each process has its own stack and heap
- The Erlang VM, BEAM, executes the **Erlang code**
- Process scheduling is controled by reduction count
- GC is local to a process
- GC is generational and copying







-module(helio), -export([run/0]), run() -> io:format("Helio, World"in")

Functional Single-assignment Dynamically typed

Concurrent

Distributed

Message passing

No sharing

Automatic Memory Management (GC)

Soft real-time

Fault tolerant

%% File: hello.erl

- -module(hello).
- -export([run/0]).

run() -> io:format("Hello, World!\n").

-module(helio), -export([run/0]), run() -> io:format("Helio, World"in")

Functional Single-assignment Dynamically typed

Concurrent

Distributed

Message passing

No sharing

Automatic Memory Management (GC)

Soft real-time

Fault tolerant

-module(helio), -export([run/0]), run() -> io:format("Helio, World"in")

Functional Single-assignment Dynamically typed

Concurrent

Distributed

Message passing

No sharing

Automatic Memory Management (GC)

Soft real-time

Fault tolerant

```
(foo@frodo)1> X=42.
42
(foo@frodo)2> X.
42
(foo@frodo)3> X=43.
** exception error: no match of right hand
side value 43
(foo@frodo)4>
```

-module(helio), -export([run/0]), run() -> io:format("Helio, World"in")

Functional Single-assignment Dynamically typed

Concurrent

Distributed

Message passing

No sharing

Automatic Memory Management (GC)

Soft real-time

Fault tolerant

```
(foo@frodo)4> F= fun(X,Y) -> X + Y end.
#Fun<erl_eval.12.113037538>
(foo@frodo)5> F(1,2).
3
(foo@frodo)6> F(1.0, 2).
3.0
(foo@frodo)7> F("1", 2).
** exception error: bad argument in an arithmetic expression
  in operator +/2
    called as "1" + 2
```

-module(helio), -export([run/0]), run() -> io:format("Helio, World"in")

Functional Single-assignment Dynamically typed

Concurrent

Distributed

Message passing

No sharing

Automatic Memory Management (GC)

Soft real-time

Fault tolerant

Demo 1

```
%%% DEMO 1
%%% Show code
cd("C:/Users/happi").
S=demo:start server().
W = demo:start_worker(S).
W2 = demo:start_worker(S).
S! {broadcast, hi}.
W!terminate.
exit(W2, kill).
S!{broadcast, hi}.
exit(S, kill).
S! {broadcast, hi}.
```

Hot code loading

Im() > (c:)(M) | M < mm()).

mm() > modified_modules().

**Modified_modules().

**Modified_

Mnesia

Hof on DB QLC

Fault tolerant

Supervisor trees

Demo3

Hof map/fold

Genserver

Atoms/symbolic programming

Easy to make fault-tolerant systems.

- Erlang was designed from the ground up with the purpose of making it easy to develop faulttolerant systems.
- Erlang was developed by Ericsson with the telecom market in mind.
- Erlang supports processes, distributed systems, advanced exception handling, and signals.
- Erlang comes with OTP-libraries (Open Telecom Platform), e.g. supervisors and generic servers.

Low maintenace easy upgrade

Hot code loading.

Distribution.

Interactive shell.

Simple module system.

No shared state.

Virtual machine.

Network programming is easy

Distributed Erlang solves many network programming needs.

Setting up a simple socket protocol is a breeze.

The binary- (and now bit-) syntax makes parsing binary protocols easy.

There are simple but powerful libraries for HTTP, XML, XML-RPC and SOAP.

Ability to leverage multi-core

The concept of processes is an integral part of Erlang.

No shared memory -- easier to program. The Erlang Virtual machine (BEAM) has support for symmetric multiprocessing.

"Each year your sequential programs will go slower. Each year your concurrent programs will go faster."

-- Joe Armstrong

Rapid development

- Automatic memory management.
- Symbolic constants (atoms).
- An interactive shell.
- Dynamic typing.
- Simple but powerful data types.
- Higher order functions and list comprehensions.
- Built in (distributed) database.

God way to get great programmers.



Lennart

Phil

SPJ

John

Nice paradox:

The lack of Erlang programmers makes it easier for us to find great programmers.

There are many great C and Java programmers, I'm sure, but they are hidden by hordes of mediocre programmers.

Programmers who know a functional programming language are often passionate about programming.

TM

Passionate programmers makes Great Programmers