1000 YEAR-OLD DESIGN PATTERNS

Ulf Wiger
Erlang Solutions Ltd

courtesy of www.romanticism-in-art.org

What this talk is about

- Exploring a perspective on SW design
  - Human patterns for concurrency
- A walk down Memory Lane

- No easy recipes

Movie Tips

- To plant my idea...
- A few movies will be recommended
- …when they illustrate some aspect of this talk
- Try expensing them as "study of intuitive concurrency design patterns"

From Inception (2010) http://www.imdb.com/title/tt1375666/

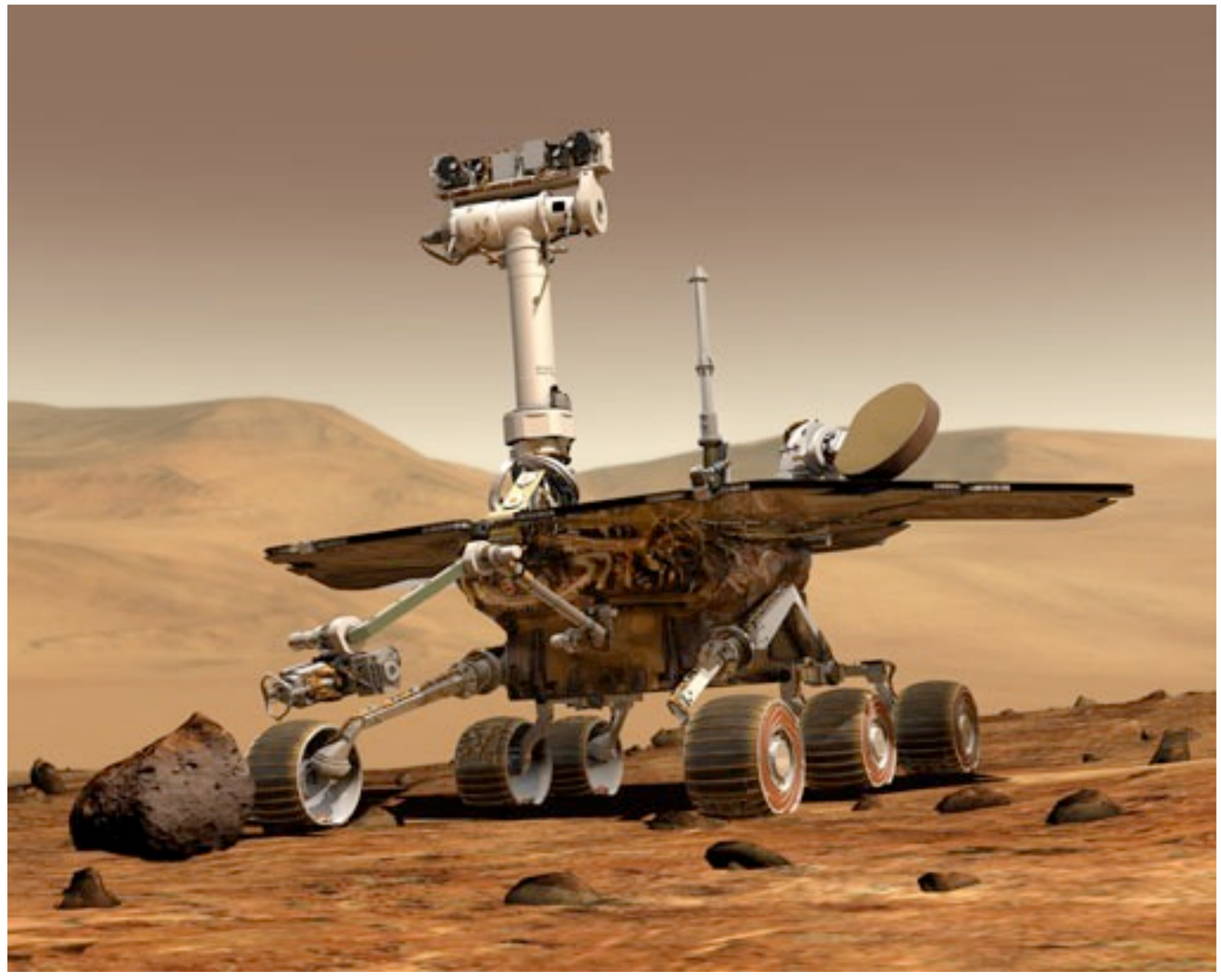
Human cooperation is naturally concurrent

- All sorts of concurrency problems are common knowledge to humans
- Mitigation strategies have been explored for millennia
- Lots of coordination and supervision design patterns

http://www.sassansanei.com/images/fullsize-trafficjam-640x480.jpg

A problem…

- Although humans document their algos,
- …they do it for human consumption (S.O.P.s)
  - Not for programmers
- Most research into ”human algorithms” is about cognitive modeling (autonomous robots)



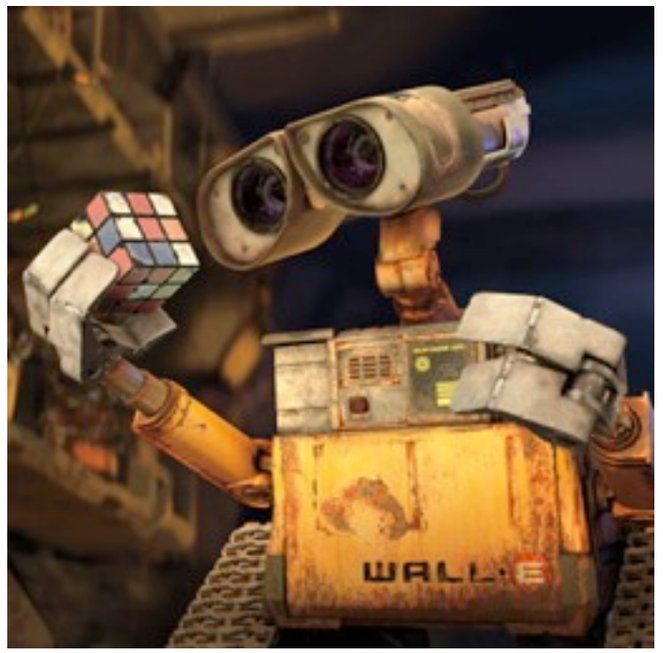
Mars Rover

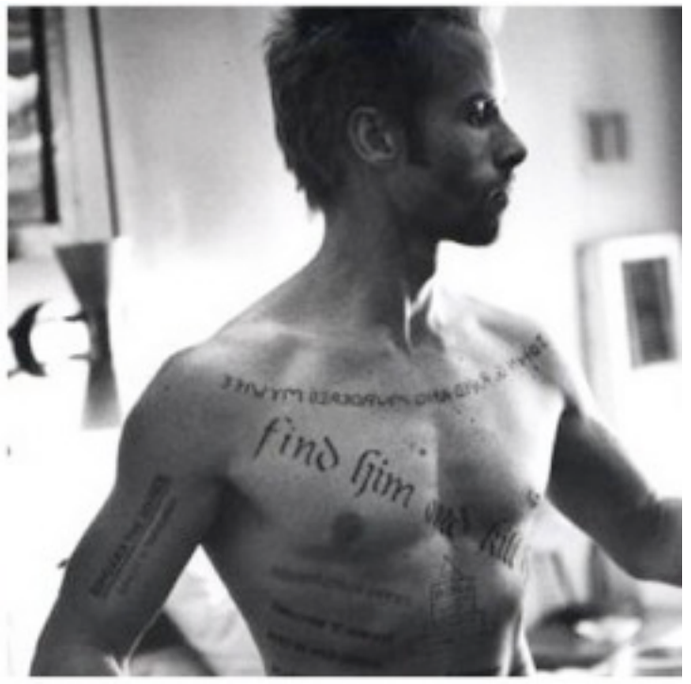
Research into human protocol

- Collect examples of how humans solve cooperation problems

- Go to the movies!

- Observe real-life patterns; consider what could transfer to software systems





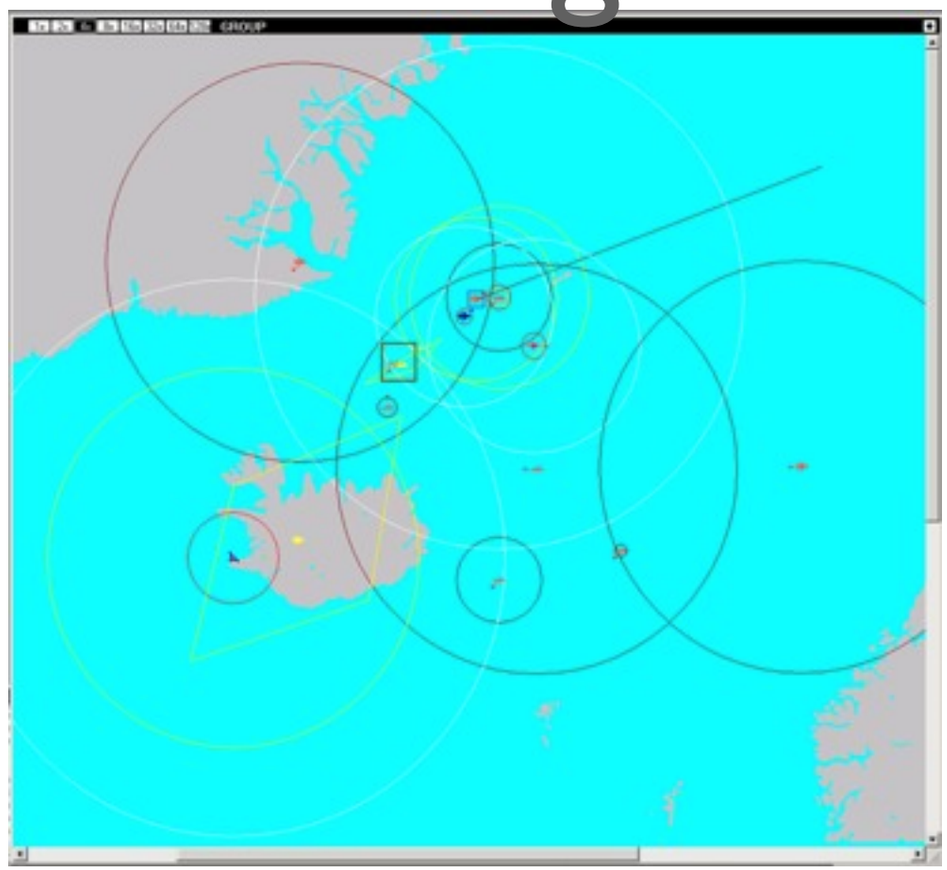
AC2SMAN - My formative years

- Alaskan Command & Control System Military Automated Network

- Built in 4 months by a fighter pilot from Memphis, and some geeks

- First ever "Overall Outstanding" rating given by NORAD 1989

The C² System Design Challenge



- Mission-critical
- Soft real-time
- Inconsistent data input
- Varying operating conditions
- Potentially global scale
- No single point of failure (40+ sites)
- Live, simulation and exercise – sometimes simultaneously

Erlang

The Competition

- One project had a $200M/year budget
- Desert Storm C2 system installation took 50K man-hours! (…!!!)
- (in our view) No alternative system came close to competing

The Secret?

- Keep the project small…
- Automate the existing workflow!
- Military C^2 goes way back...
  - Tsun Tzu, 500 B.C.
  - Julius Caesar, ca 50 B.C.
  - Genghis Khan, 1200s
  - von Clausewitz, 1800s

(Movie Tip)

- Crimson Tide (1995)
- Military command protocol
- Redundancy
- Fail-safes
- Byzantine Generals Problem
- Bully algorithm

http://www.imdb.com/title/tt0112740/

AC²SMAN Database issues

- Asynchronous, event-triggered replication
  - Across 40 sites
- No 2-phase commit – no conflicts ”possible”
- Main challenge: full replication over a 19.2Kbps modem line

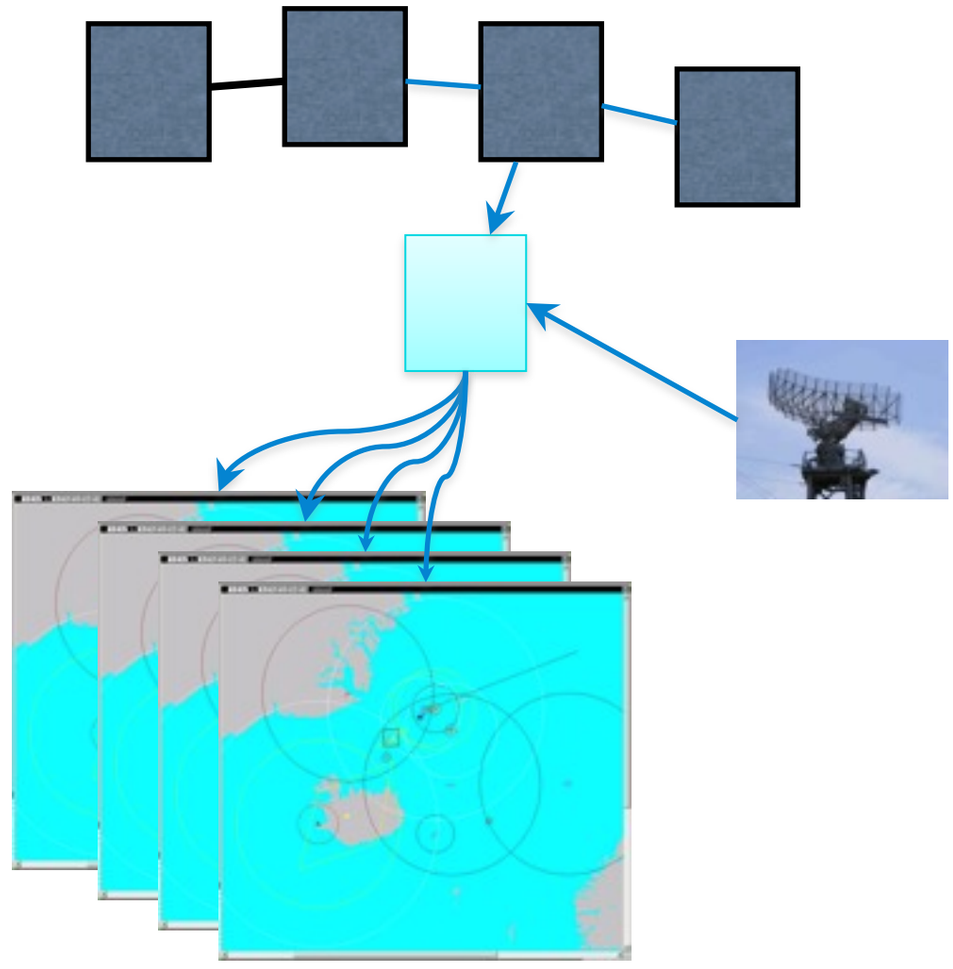
- Relational databases anno 1989 were simply non-starters

The failed alternative?

- Trying to use early-90s Distributed RDBMS technology
- This was the beginning of the hardships that led to the CAP Theorem
- The problem didn't call for an RDBMS
  - We're automating a workflow that's been around for millennia

The Feed Aggregation Problem

- Real-time subscription feed for tactical map workstations
- Messaging server was a big pile of C++ code
- Single point of failure
- Ran out of memory daily
- (Not due to programmer incompetence)

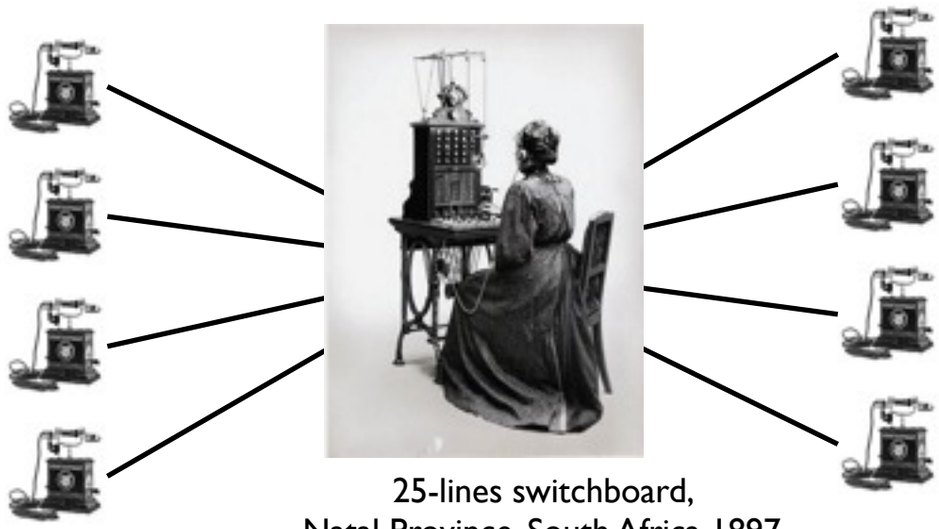


I was Searching for a Solution

- Tons of approaches evaluated
- CASE Tools, Client-Server middleware, AI middleware…
- Something suitable for a one-man development team

Eureka!

- Eventually landed in telecoms 1992
  - ”Computers in Telecommunictions” course at KTH, Stockholm
    - Teachers: **B Däcker**, **R Virding**
    - Programming language: **Erlang**
- Erlang seemed to be a perfect fit!



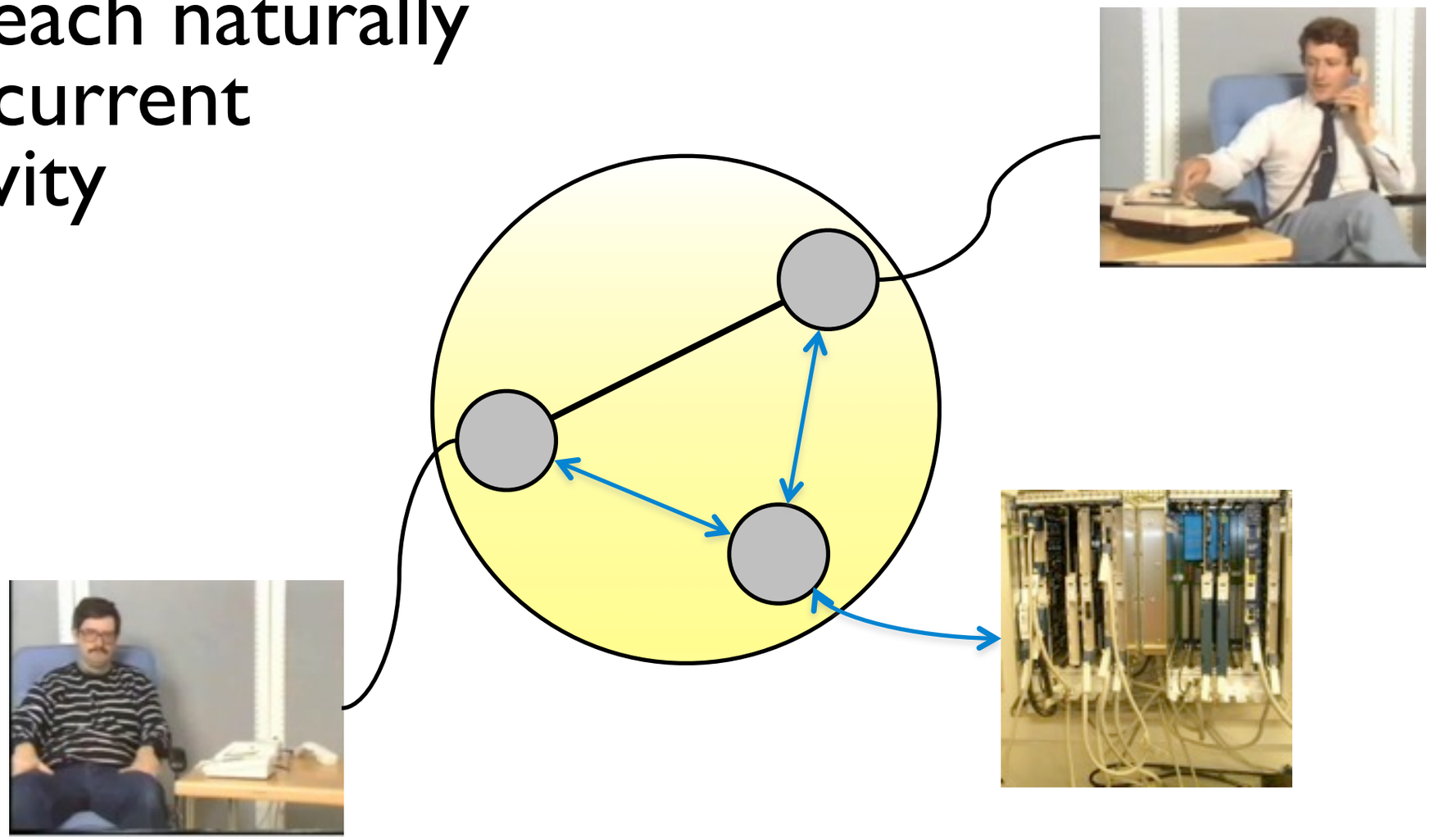
25-lines switchboard,
Natal Province, South Africa 1897
Cross-switchboard calls required
human interaction.

Erlang, Intuitively

http://video.google.com/videoplay?docid=-5830318882717959520#

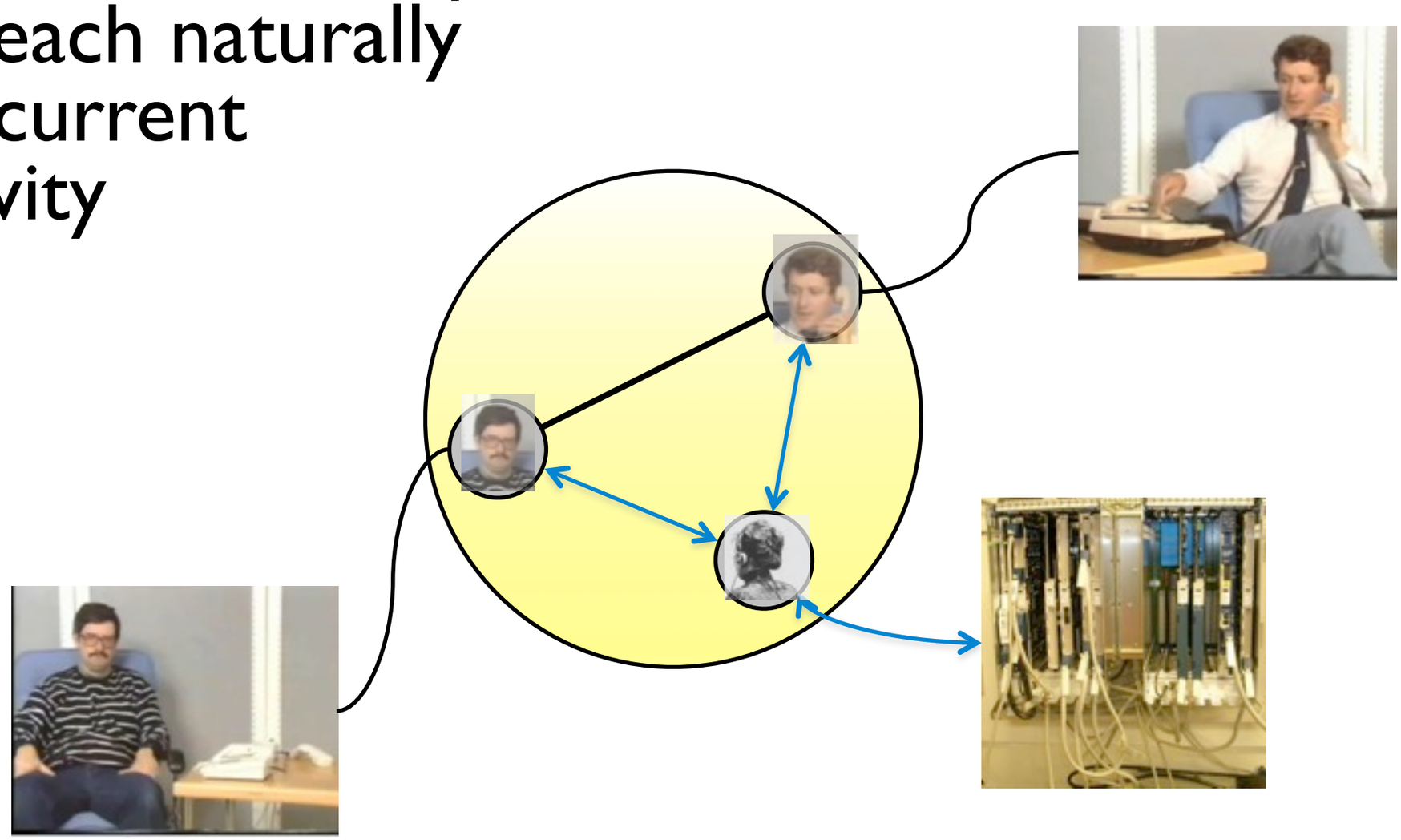
Erlang, Intuitively

- One concurrent process for each naturally concurrent activity

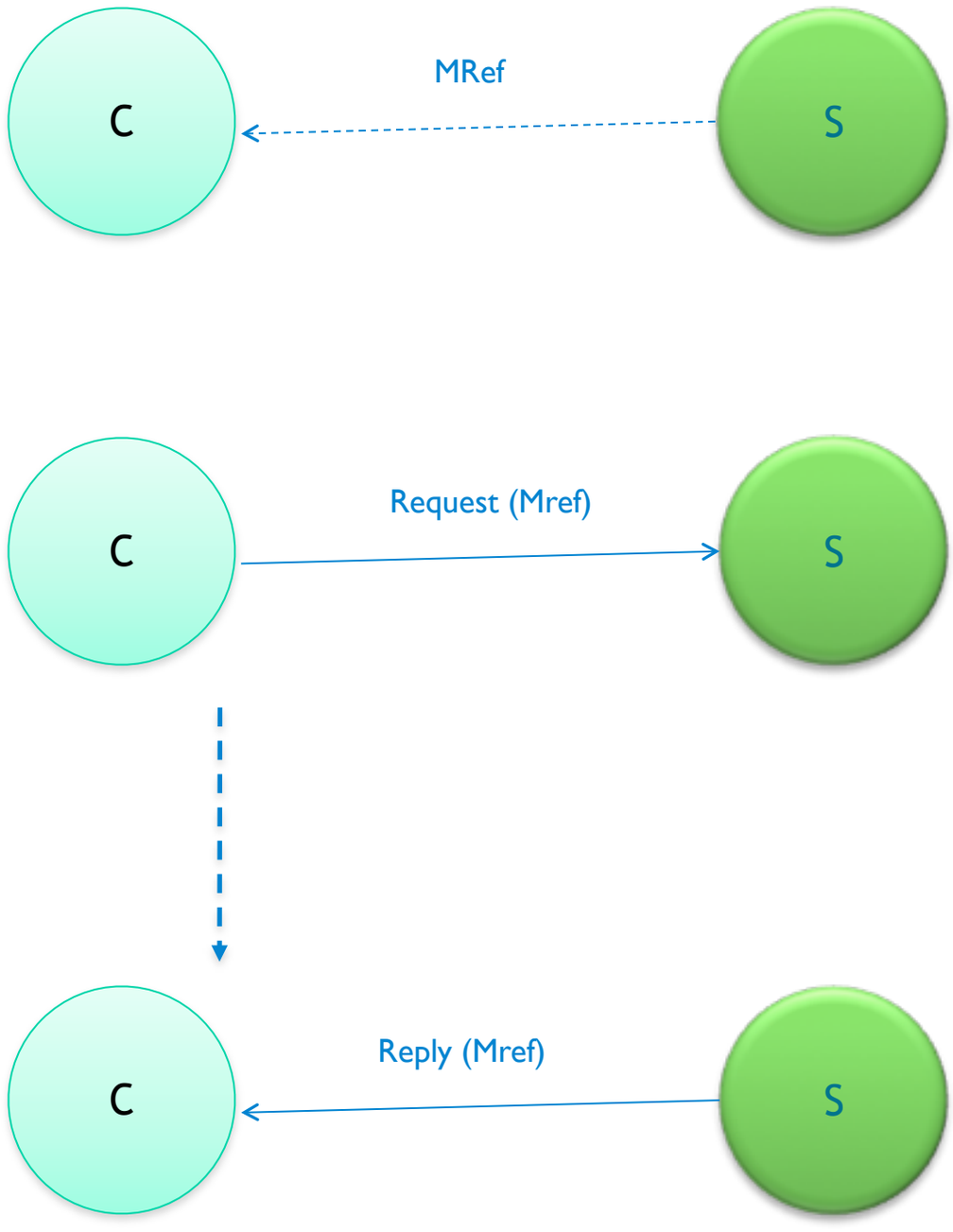


Erlang, Intuitively

- One concurrent process for each naturally concurrent activity

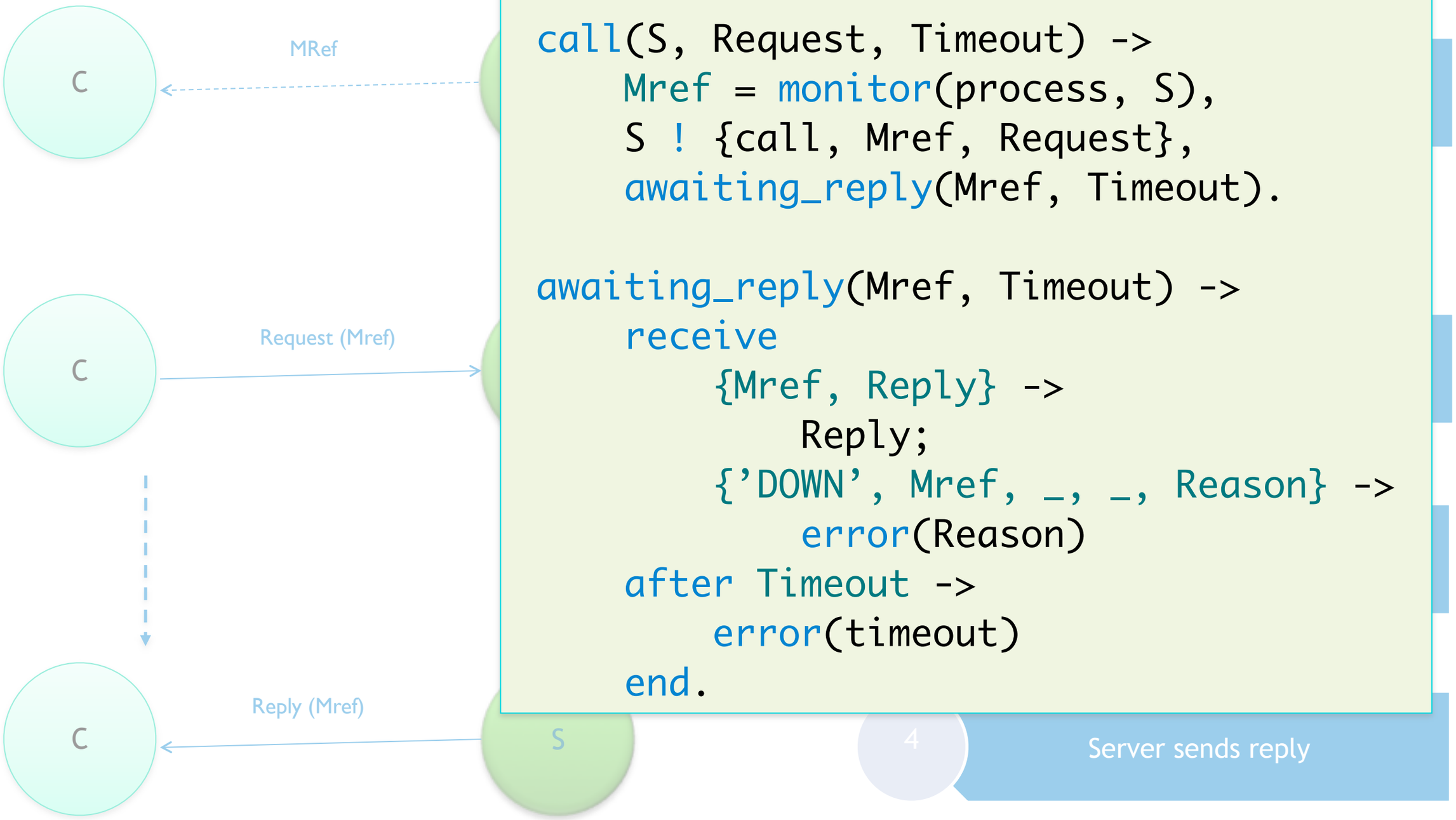


Client-server in Erlang

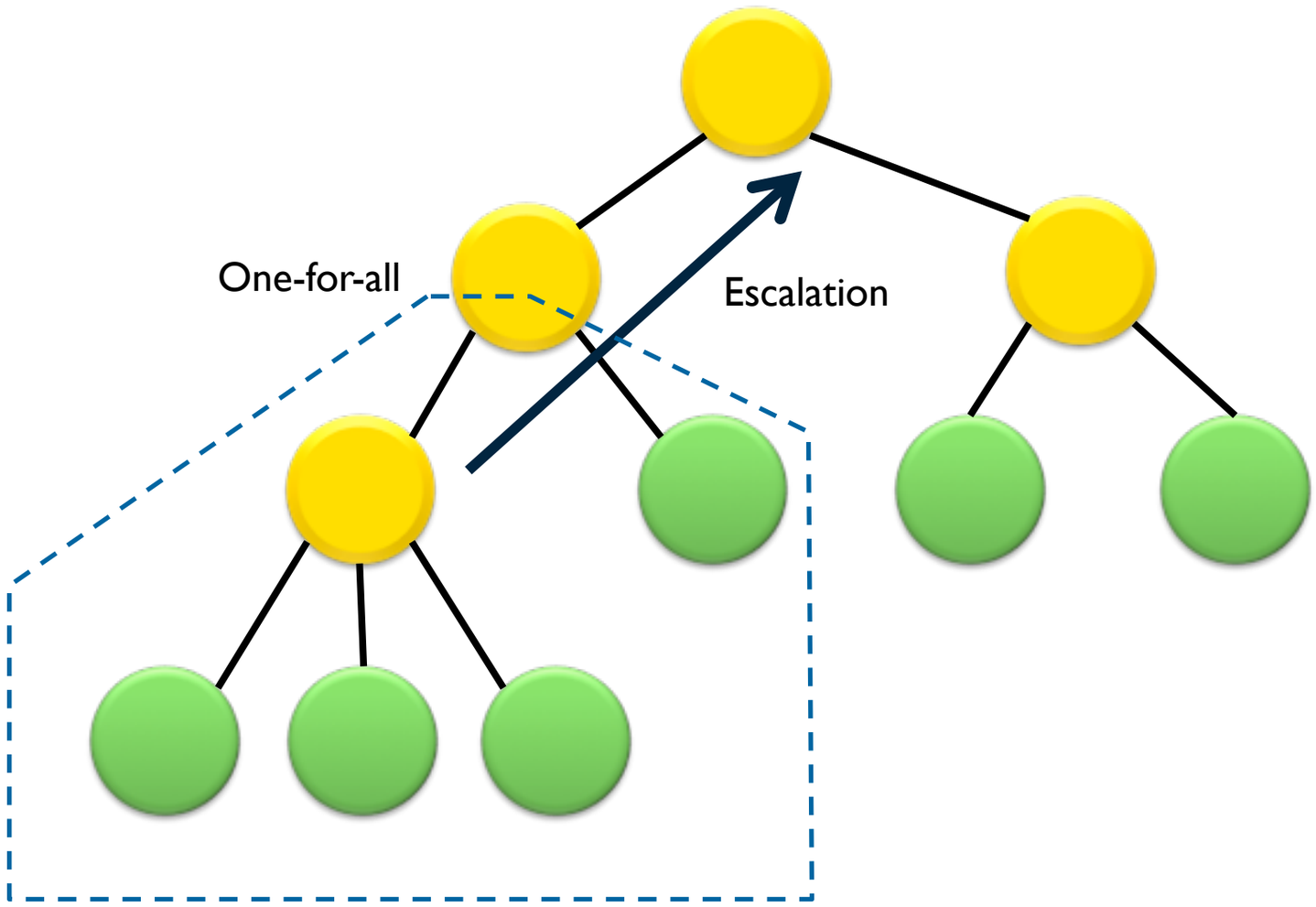


1. Client monitors server
2. Client sends a request
3. (Blocks while waiting)
4. Server sends reply

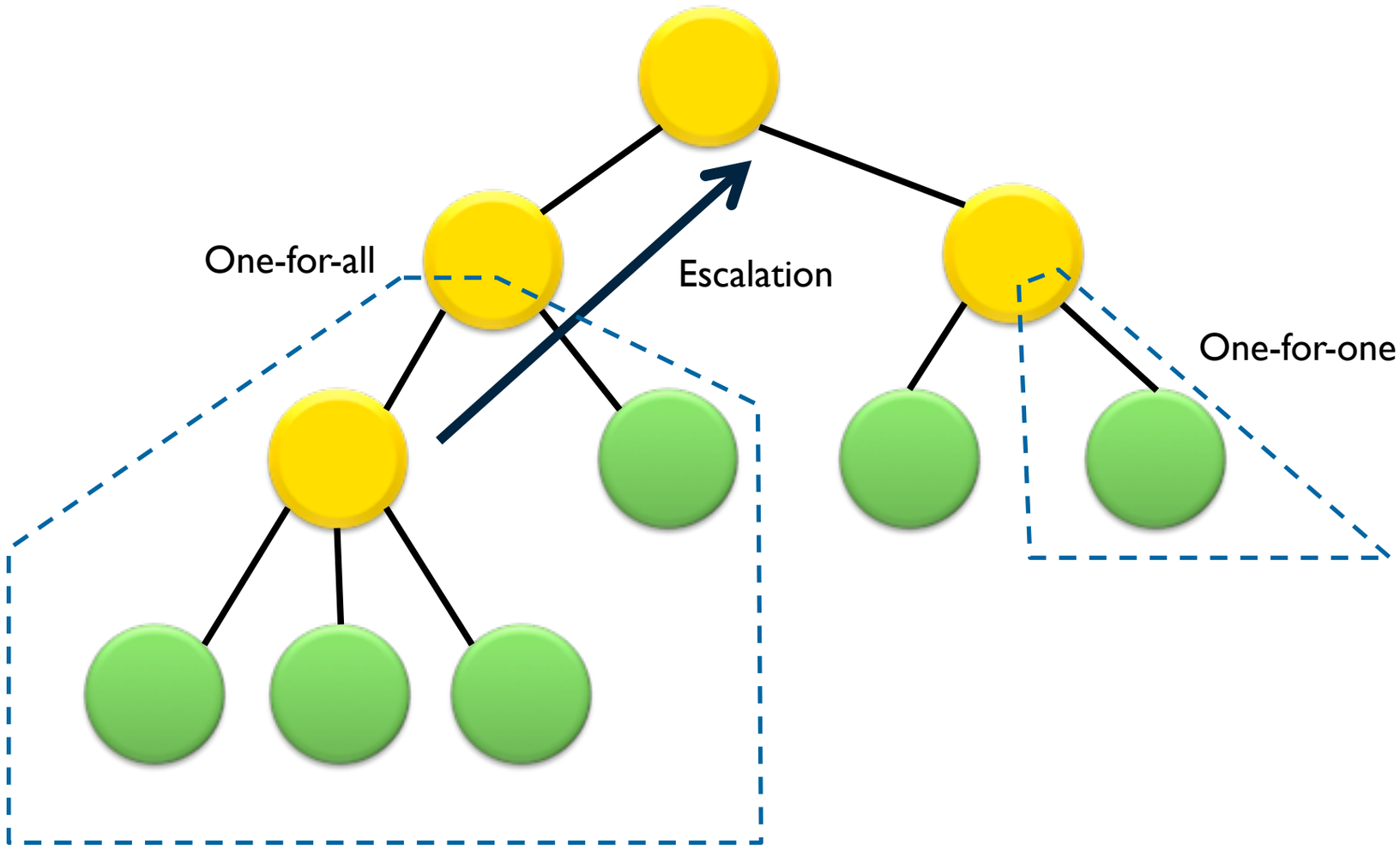
Client-server in Erlang



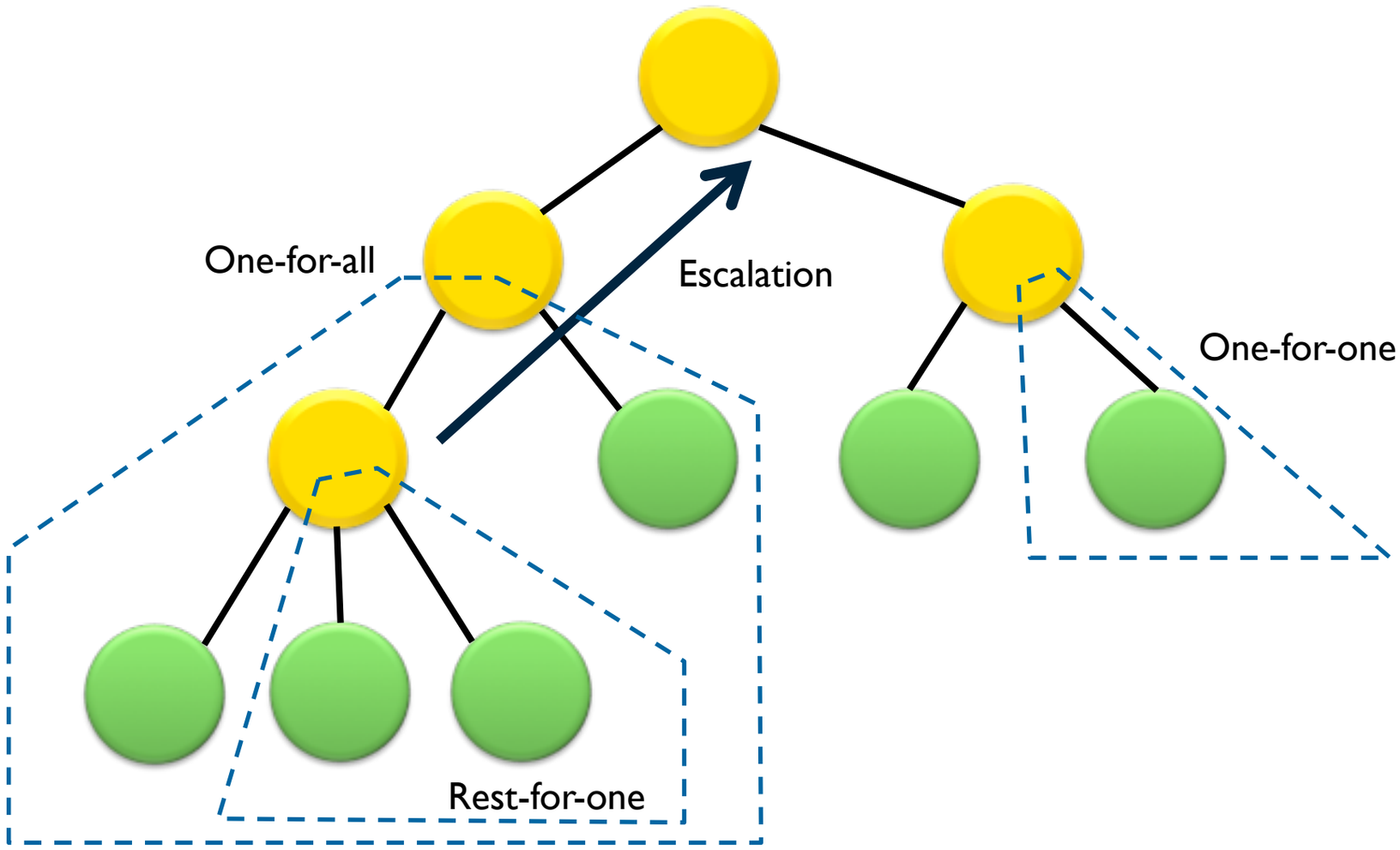
Supervisors – Out-of-Band Error Handling



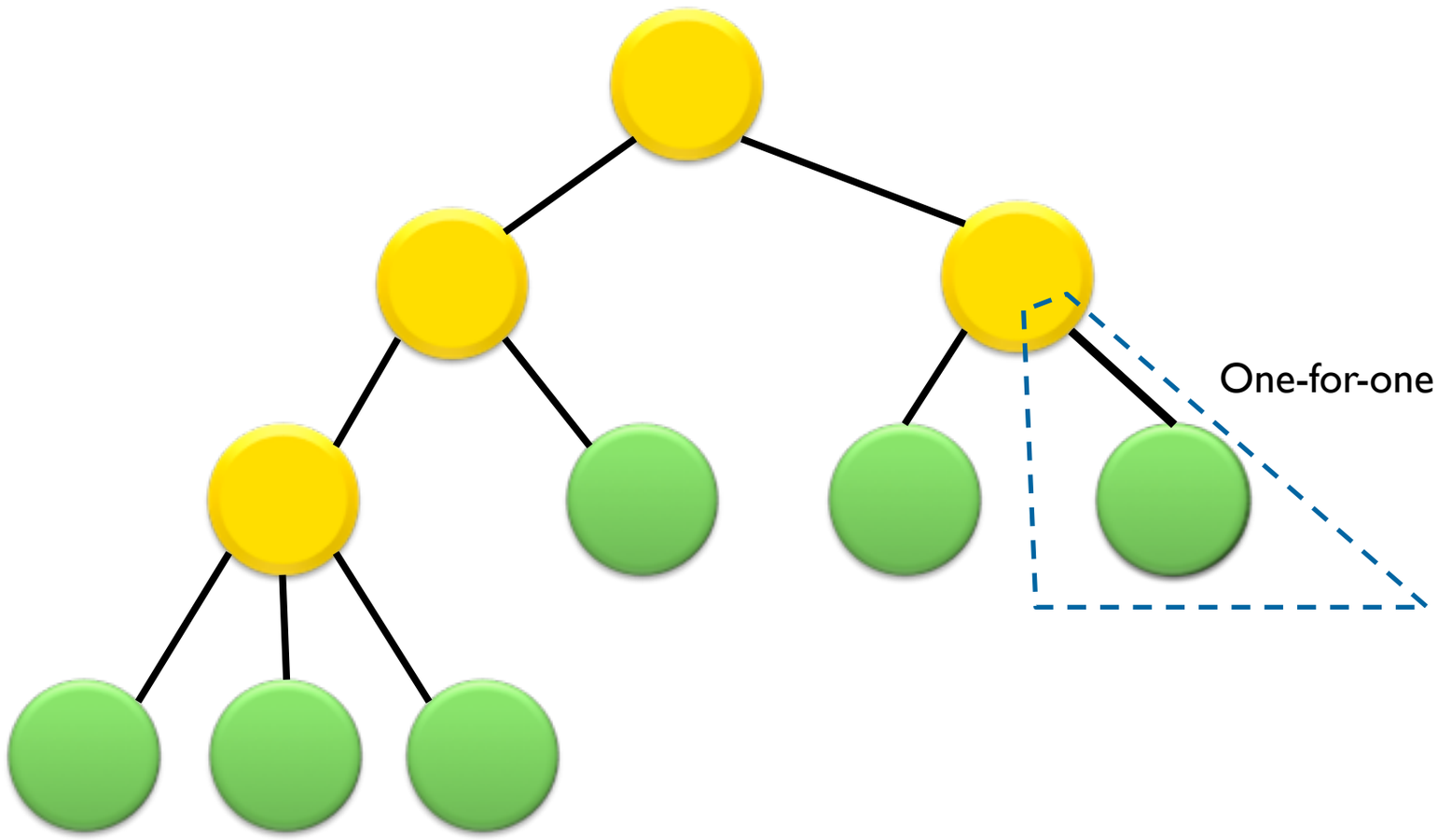
Supervisors – Out-of-Band Error Handling



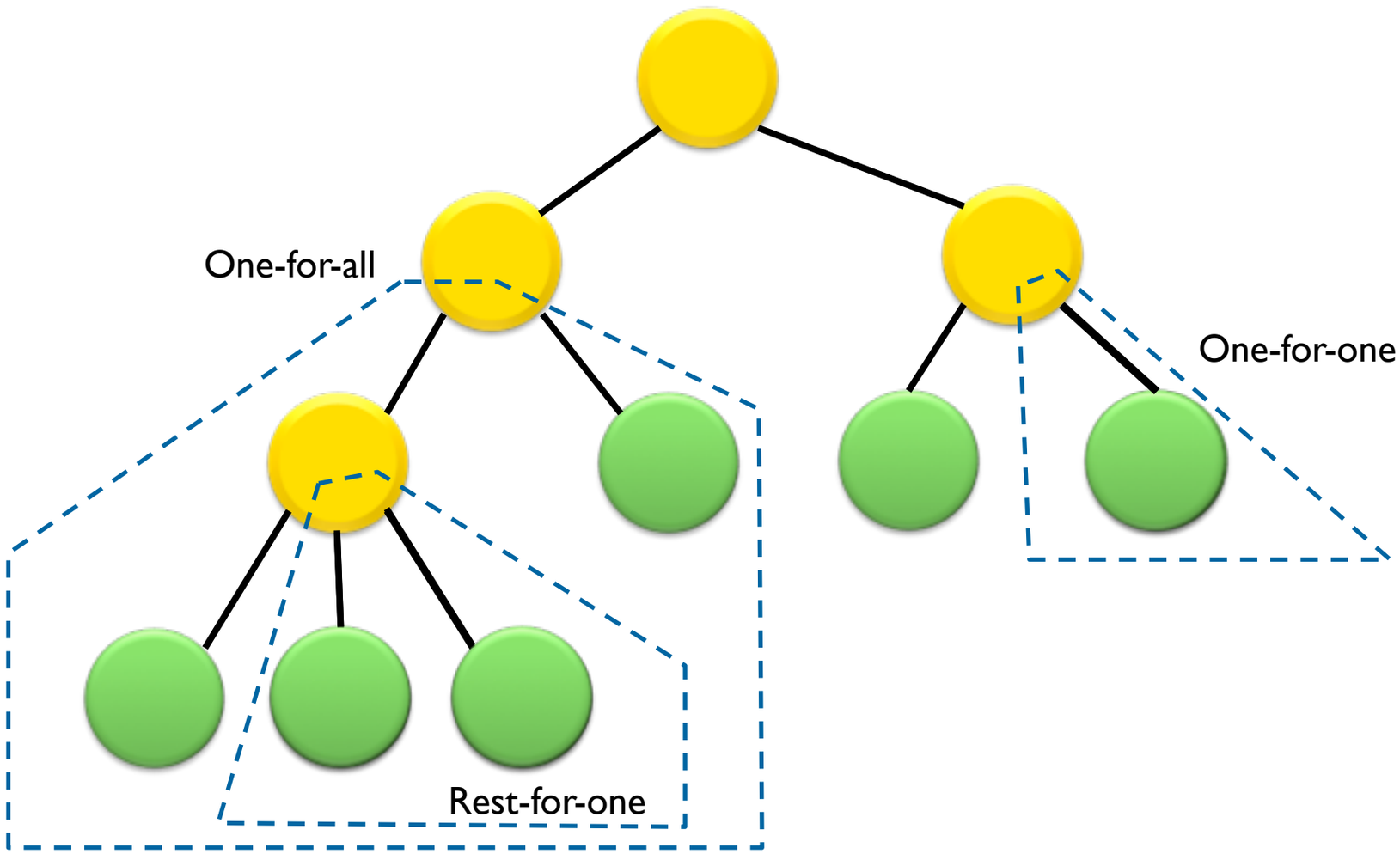
Supervisors – Out-of-Band Error Handling



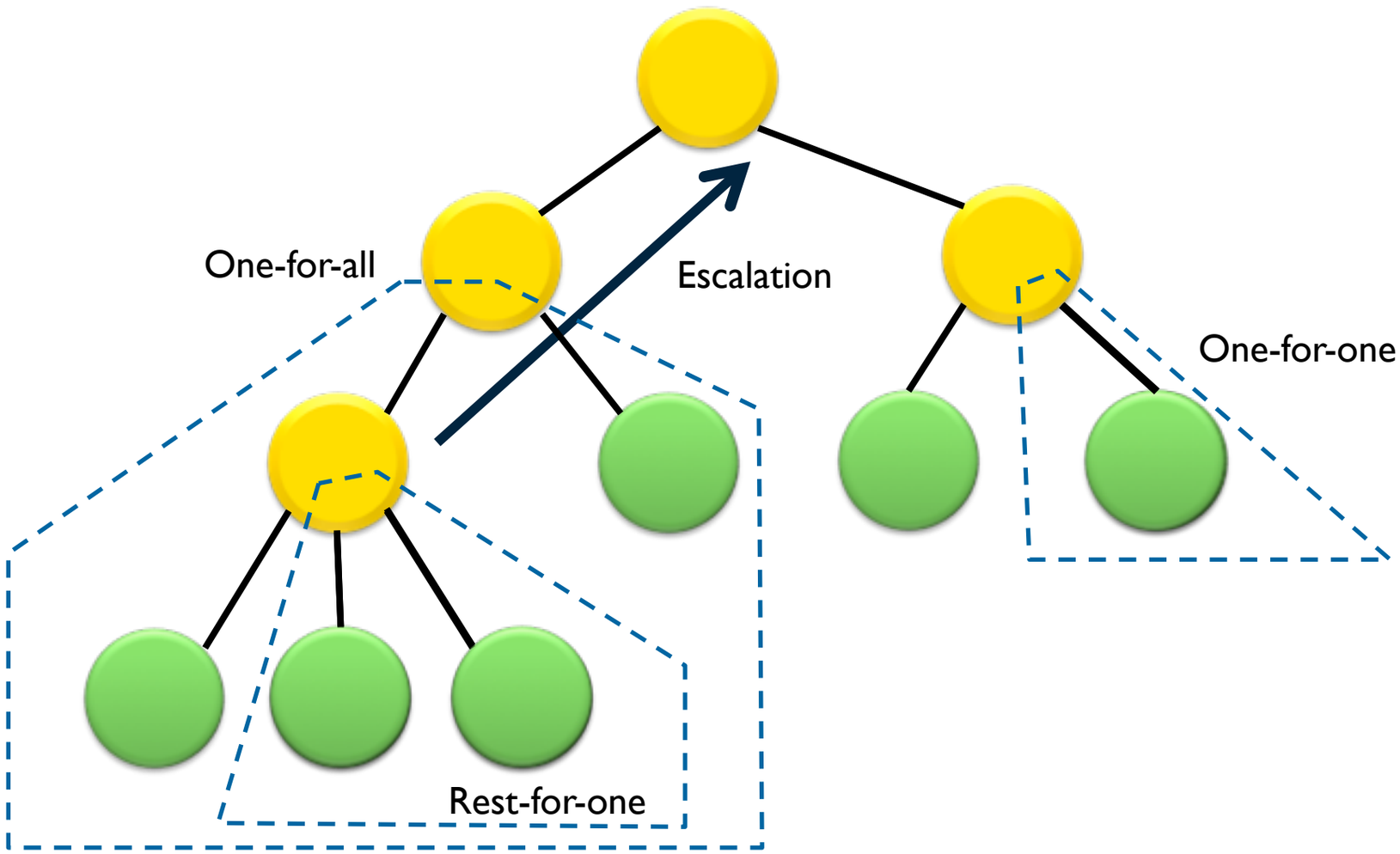
Supervisors – Out-of-Band Error Handling



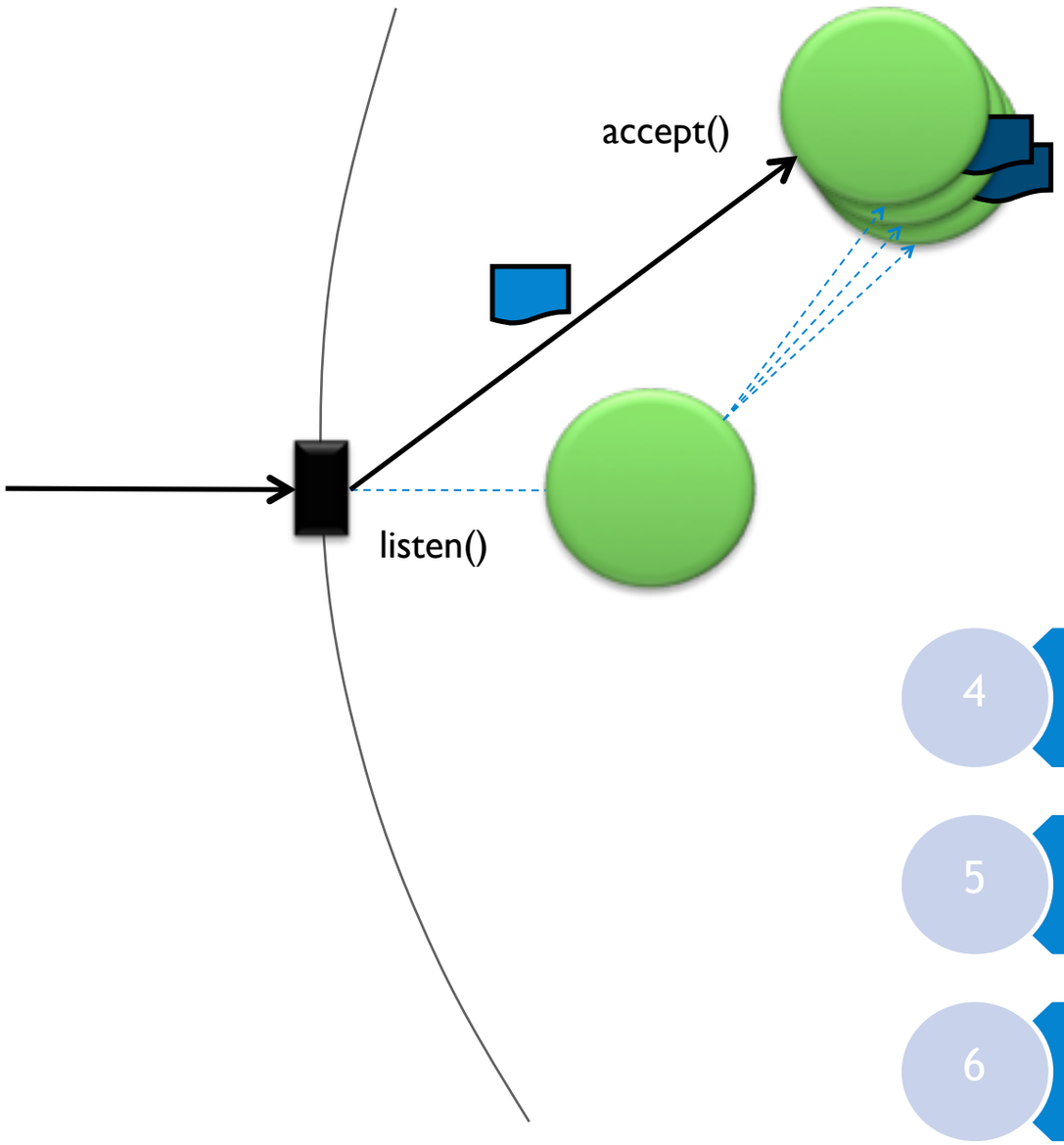
Supervisors – Out-of-Band Error Handling



Supervisors – Out-of-Band Error Handling



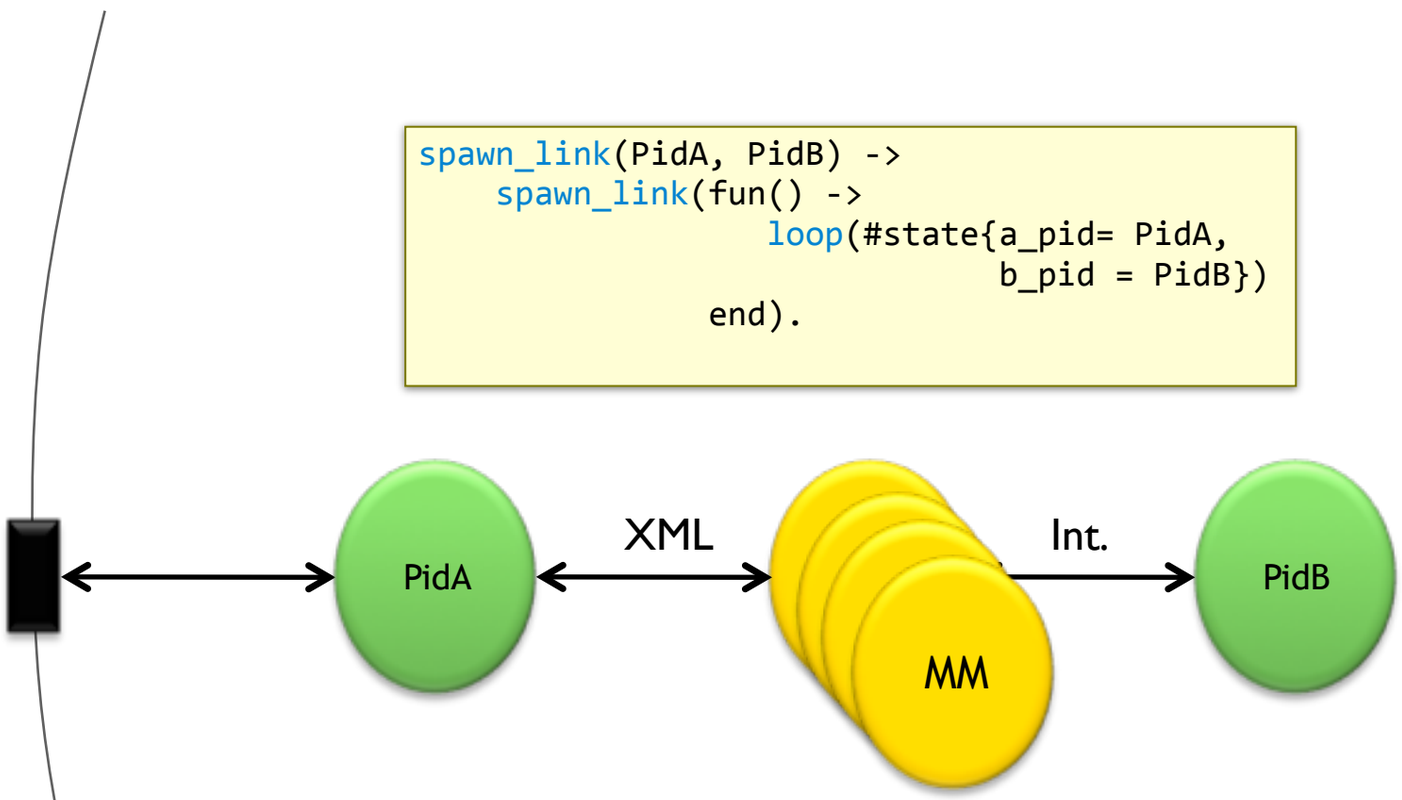
Handling sockets in Erlang



1. Static process opens listen socket
2. Spawns an acceptor process
3. Acceptor receives incoming
4. Acks back to socket owner
5. New acceptor is spawned
6. Replies sent directly to socket

Middle-man Processes

```
spawn_link(PidA, PidB) ->
    spawn_link(fun() ->
                   loop(#state{a_pid= PidA,
                               b_pid = PidB})
               end).
```




```
loop(#state{a_pid = PidA, b_pid = PidB} = State) ->
    receive
       {PidA, MsgBin} when is_binary(MsgBin) ->
           {simple_xml, _} = Msg = vccXml:simple_xml(MsgBin),
           PidB ! {self(), Msg},
           loop(State);
       {PidB, {simple_xml, _} = Msg} ->
           Bin = vccXml:to_XML(Msg),
           PidA ! {self(), Bin},
           loop(State)
    end.
```

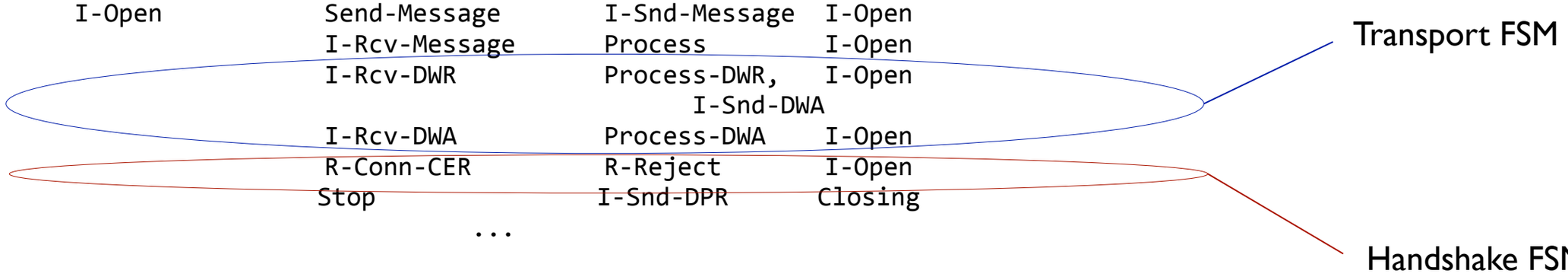
```
await_negotiation(State) ->
    receive
        {From,
         {simple_xml,
          [{"offer", Attrs, Content}]}} ->
            HisOffer =
                inspect_offer(Attrs, Content),
            Offer = calc_offer(HisOffer, State),
            From ! {self(), Offer};
        …
    end.
```

- Practical because of light-weight concurrency
- Normalizes messages
- Main process can pattern-match on messages
- Keeps the main logic clear

Language Model Affects our Thinking

Example: RFC 3588 – DIAMETER Base Protocol

```
  state             event                 action          next state
-------------------------------------------------------------------
                                  ...
    I-Open            Send-Message        I-Snd-Message   I-Open
                      I-Rcv-Message       Process         I-Open
                      I-Rcv-DWR           Process-DWR,    I-Open
                                                I-Snd-DWA
                      I-Rcv-DWA           Process-DWA     I-Open
                      R-Conn-CER          R-Reject        I-Open
                      Stop                I-Snd-DPR       Closing
                                  ...
```

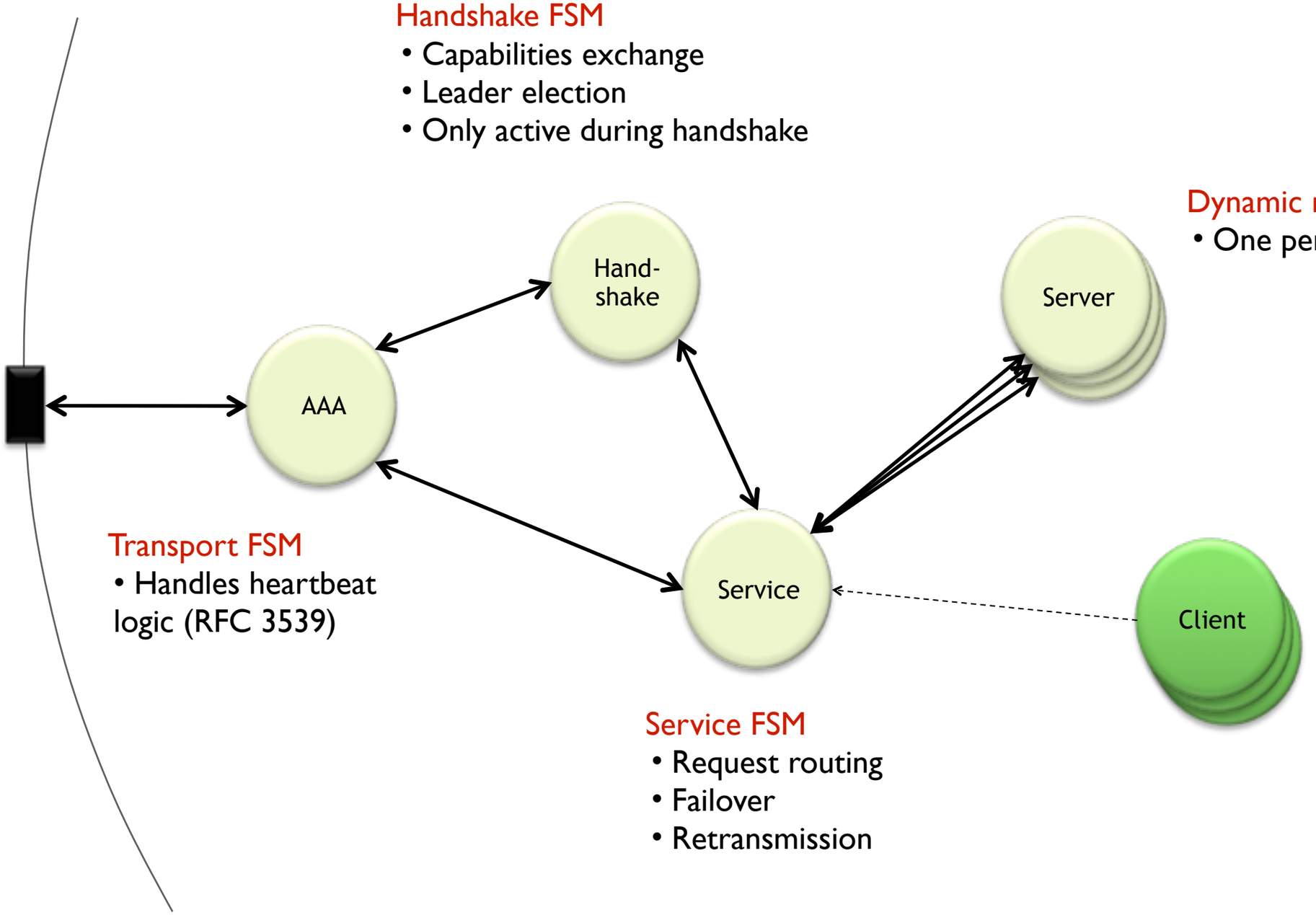


- Three state machines described as one
- Implies a single-threaded event loop
- Introduces accidental complexity

Use processes to separate concerns

Handshake FSM
- Capabilities exchange
- Leader election
- Only active during handshake

Dynamic request handler
- One per request



Transport FSM
- Handles heartbeat logic (RFC 3539)

Service FSM
- Request routing
- Failover
- Retransmission

Ericsson – The Mythical Project

- I joined Ericsson 1996 to work with Erlang
- A very large project had just been canceled
  - A well-publicized failure
- Distributed real-time, fault-tolerant complex systems in C++

Why did it crash?

- No obvious single culprit
- Obviously, the size of the project was a problem
  - But why so large?
- OO mania, featuritis, hubris?

- My thought: failure to contain the problem

AXD301 – The Pickup Project

- 200 people put into one building
- Mission: Build a product within 2 years
  - Something in the ATM domain with Telecom Characteristics
- Much leeway was given
- Erlang/OTP chosen as key implementation technology
- Result: A product was delivered in 2 years
  - Eventually returned Wireline Division to profit

Pragmatic thinking

- Shell shocked from previous project
- Fall back on what's known to work
- Straight and simple took us pretty far
  - Up to 16x16 = 256 interconnected boards
  - Up to 32 control plane processors
  - Up to 500k simultaneous phone calls
  - > 99.999% consistent uptime
    - (including maintenance & upgrades)

Erlang

Outsiders about Erlang

- Non-programmers in our projects liked Erlang
- They understood the abstractions and design patterns

Abstractions for non-determinism

- We were building complex distributed message-passing systems
- **Key challenge: contain the non-determinism!**
- Prevent explosion of the state-event matrix
- This had been identified by Ericsson already in the late 70s…

Some similar projects

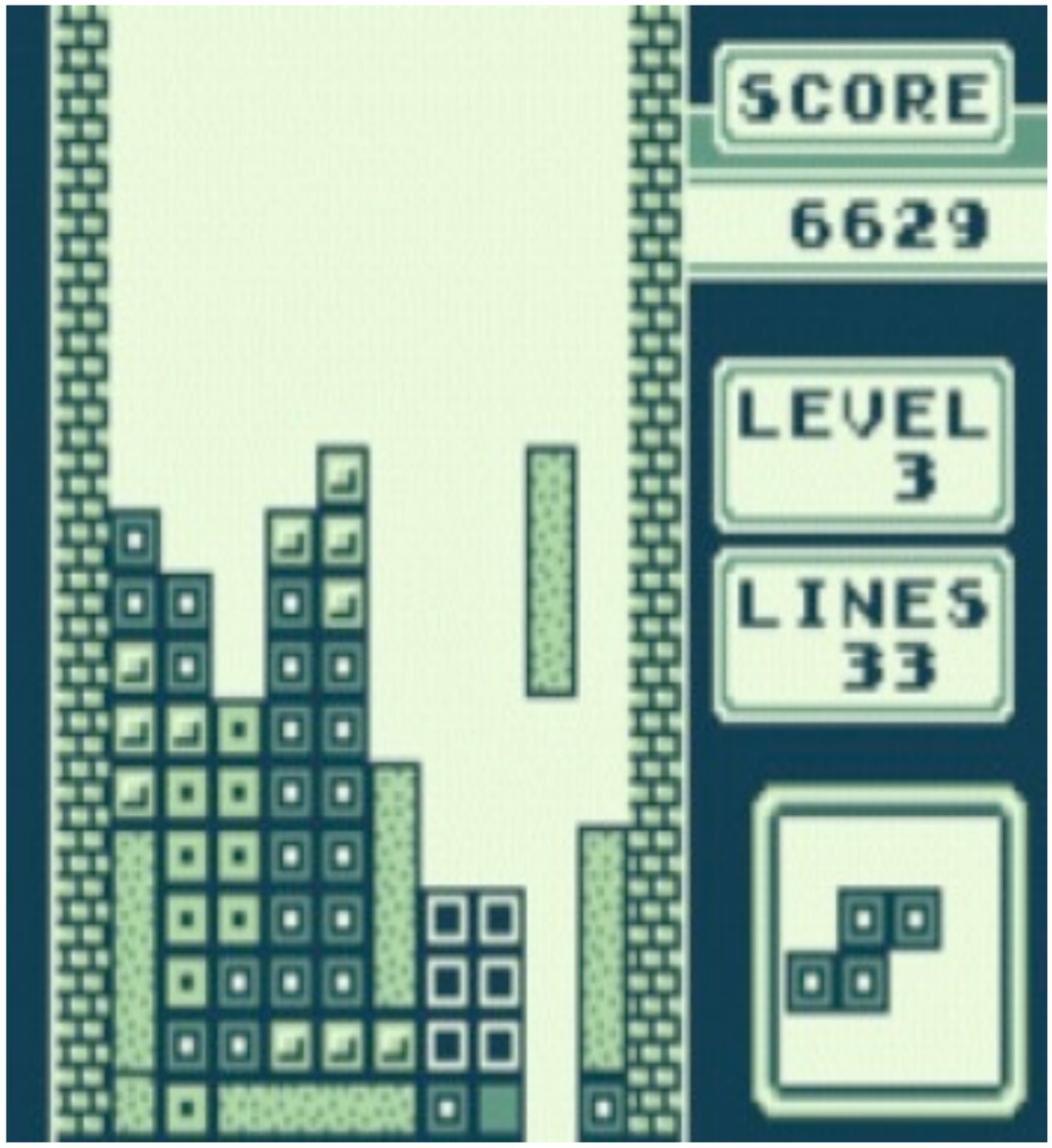
- In one (mature) UML/C++ project, 10% of all bugs were related to unexpected order of events
- Inadequate methods for abstracting away accidental ordering

Programs modeling ”human protocols”

- Must have their own thread of control
- Communicate with messages
- A sense of time
- Adapt to changes/problems
- Control order of input processing

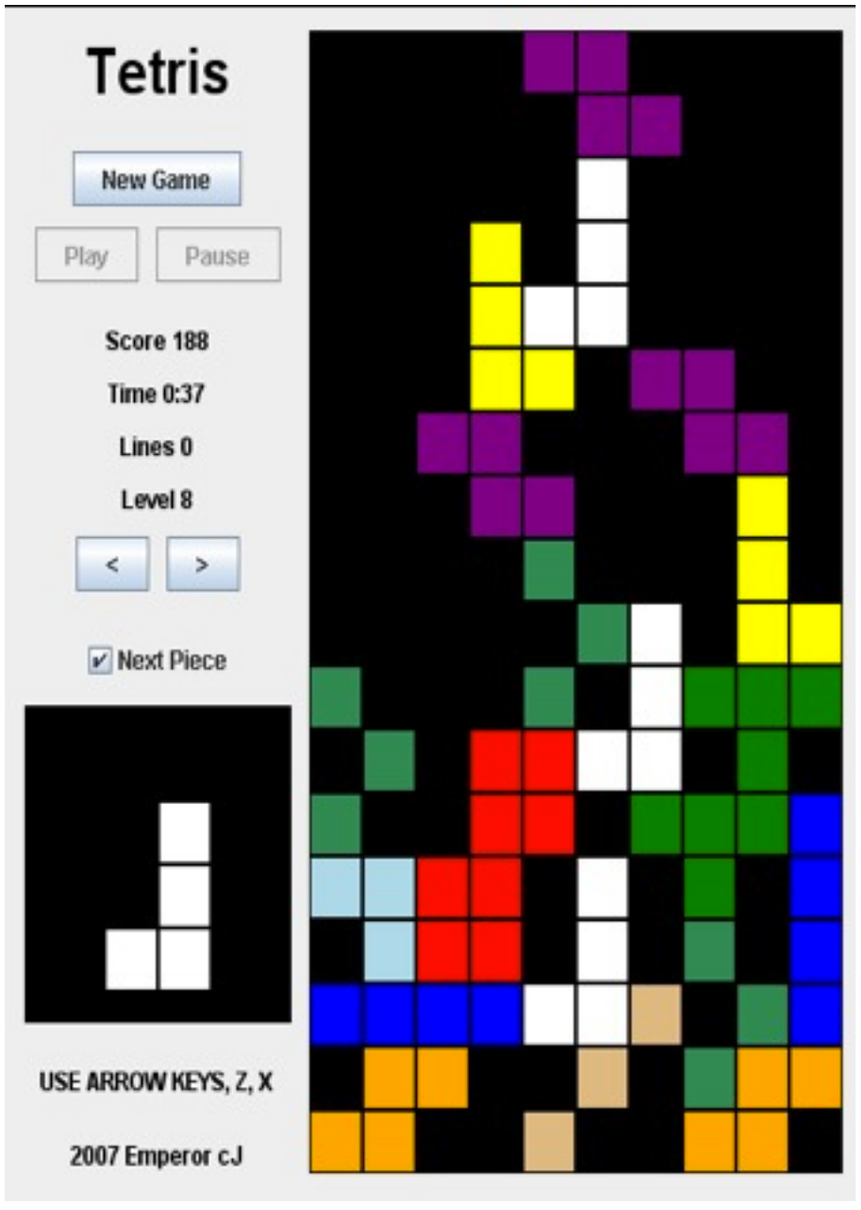
Tetris Management

- The age-old classic has coined a new time management method
- The idea: learn how to keep the pile small



Tetris Management

- Used in a derogatory sense at a major software development project
- As in ”reactive management without a plan”
- Basically, don’t let your project become a tetris game



A different kind of puzzle

- What if your problem more resembles this?
- Would you attack this problem with a tetris approach?

http://www.worldslargestpuzzle.com/hof-008.html

Event Handling Strategies

- Twist and place the next piece — before it lands
- In cheat mode, you get to peek at the next one
- Otherwise, hope for the best

- Search for a specific piece
- Put away pieces that don’t fit
- Keep at it until fitting piece found

Event Handling in Software

- FIFO, run-to-completion event handling
- Not allowed to block
- Fine, as long as the pieces fit

- Blocking, selective receive
- Wait, until the next *desired* piece arrives
- Ignore unknown pieces

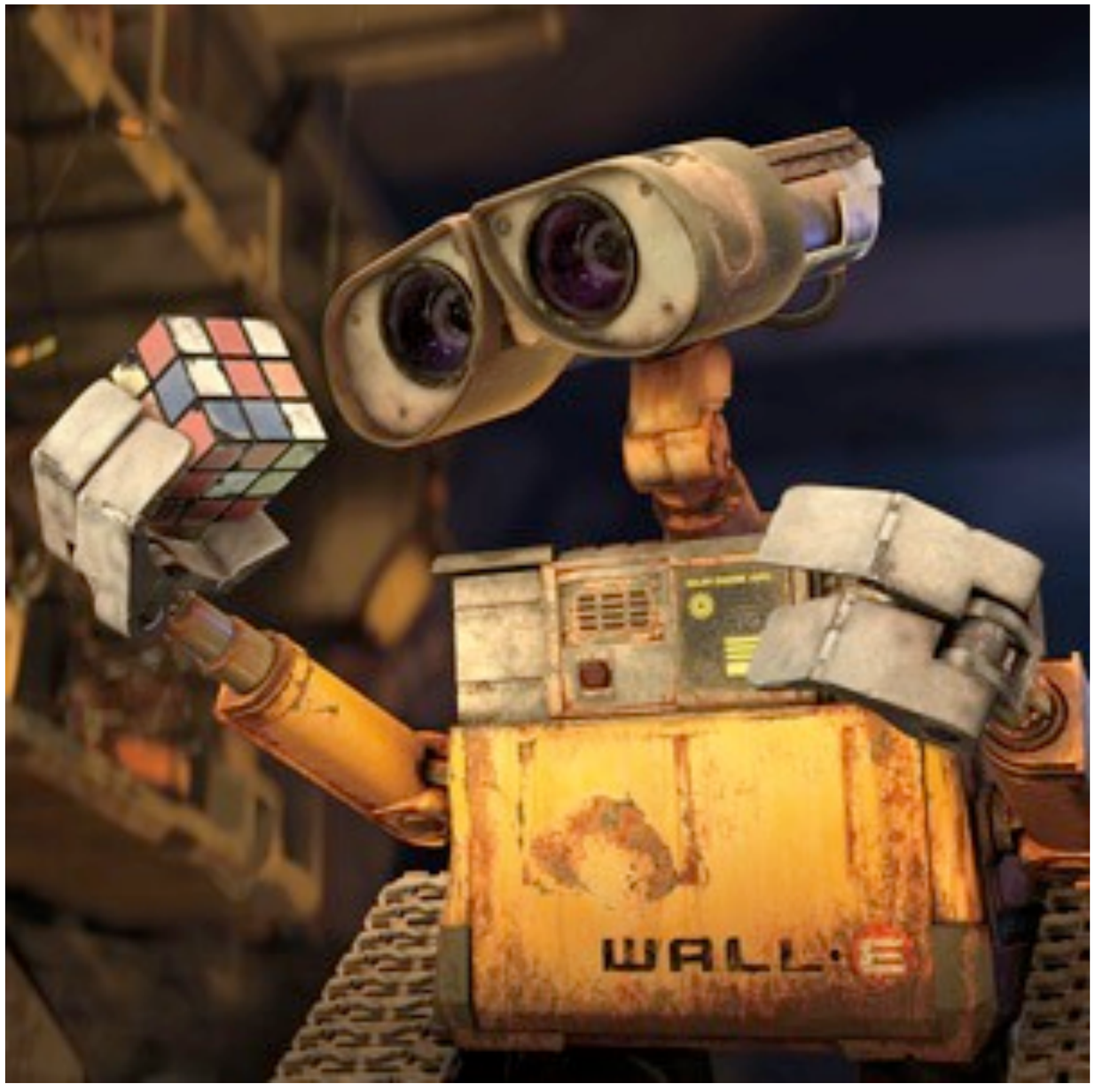
(Movie Tip)

- Memento (2000)
- Human FIFO, run-to-completion event handling
- Storing context for future reference

Memento (2000) http://www.imdb.com/title/tt0209144/

In conclusion

- Our mental models greatly influence how we attack software problems
- Our real-life experience is full of useful patterns for concurrency
- Actor-style programming is a pretty good fit for modeling such patterns



Wall-E (2008) http://www.imdb.com/title/tt0910970/

Questions?