Dissecting the Rabbit: RabbitMQ Internal Architecture

Alvaro Videla - RabbitMQ

Alvaro Videla

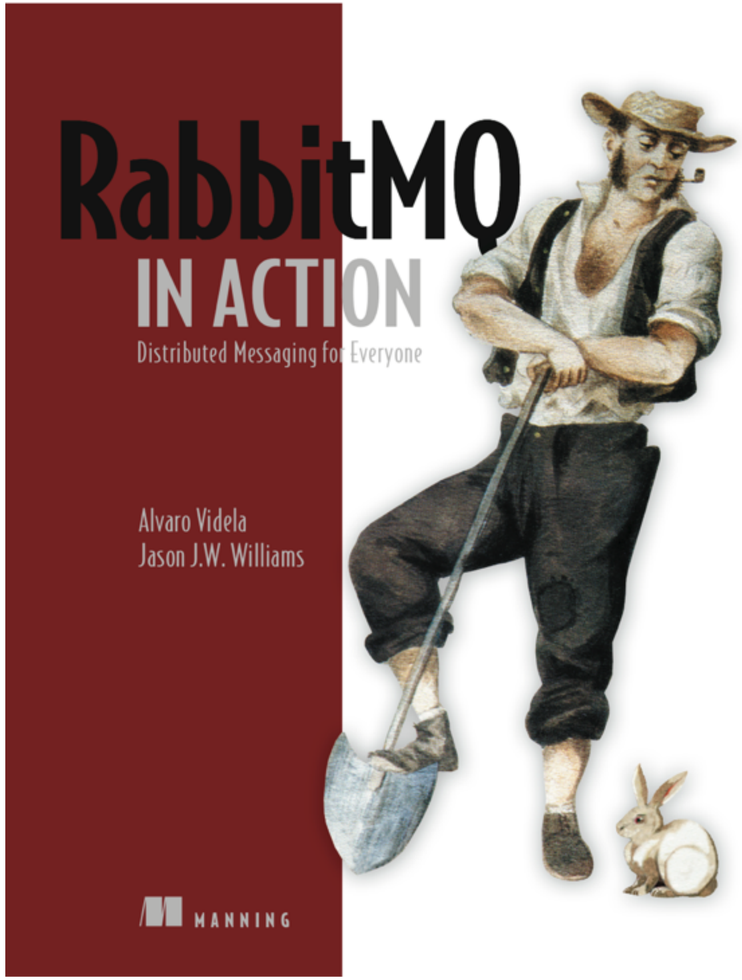
- Developer Advocate at Pivotal / RabbitMQ
- Co-Author of RabbitMQ in Action
- Creator of the RabbitMQ Simulator
- Blogs about RabbitMQ Internals: http://videlalvaro.github.io/internals.html
- @old_sound | alvaro@rabbitmq.com
  github.com/videlalvaro

About Me

Co-authored

RabbitMQ in Action

http://bit.ly/rabbitmq



Agenda

- Intro to RabbitMQ
- Dive into RabbitMQ Internals
- A day in the life of a message
- RabbitMQ message store
- RabbitMQ behaviours and extensibility

What is RabbitMQ

RabbitMQ

RabbitMQ

RabbitMQ

- Multi Protocol Messaging Server

RabbitMQ

- Multi Protocol Messaging Server
- Open Source (MPL)

RabbitMQ

- Multi Protocol Messaging Server
- Open Source (MPL)
- Polyglot

RabbitMQ

- Multi Protocol Messaging Server
- Open Source (MPL)
- Polyglot
- Written in Erlang/OTP

Multi Protocol

http://bit.ly/rmq-protocols

Polyglot

Polyglot

Polyglot

- Java

Polyglot

- Java
- node.js

Polyglot

- Java
- node.js
- Erlang

Polyglot

- Java
- node.js
- Erlang
- PHP

Polyglot

- Java
- node.js
- Erlang
- PHP
- Ruby

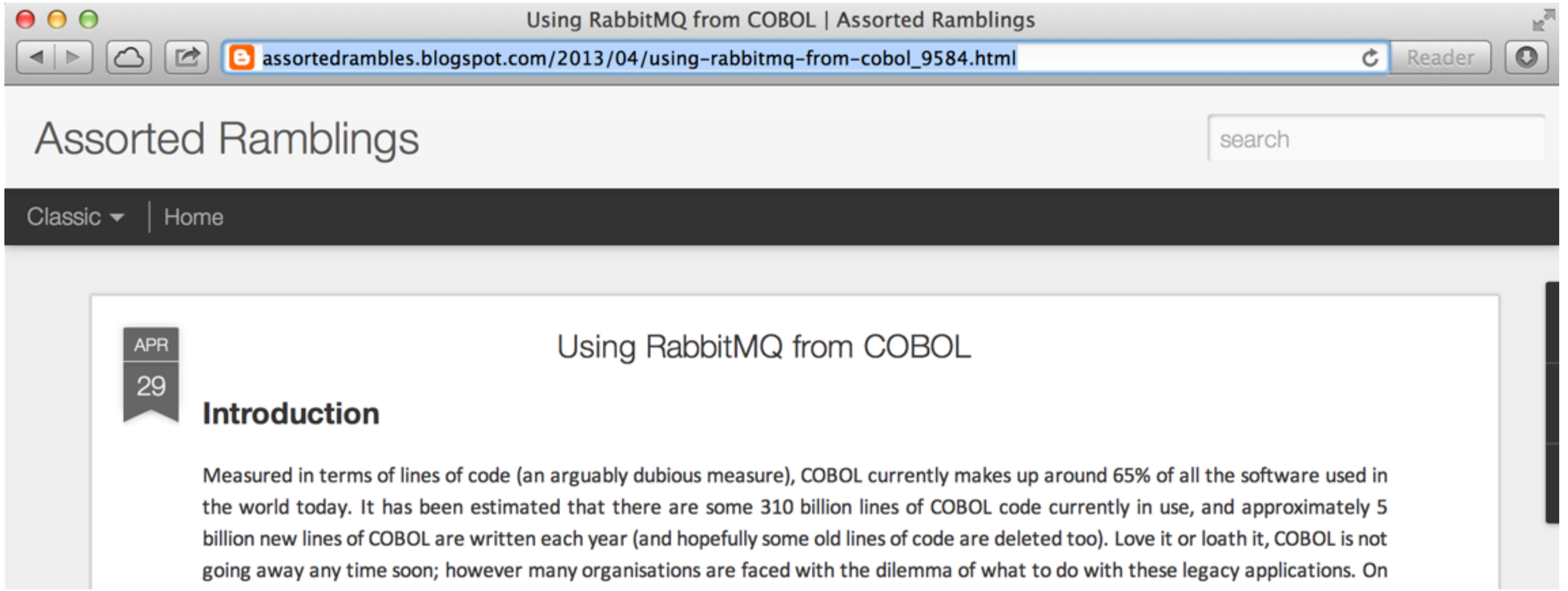
Polyglot

- Java
- node.js
- Erlang
- PHP
- Ruby
- .Net

Polyglot

- Java
- node.js
- Erlang
- PHP
- Ruby
- .Net
- Haskell

Polyglot



Even COBOL!!!11

Some users of RabbitMQ

Some users of RabbitMQ

- Instagram

Some users of RabbitMQ

- Instagram
- Indeed.com

Some users of RabbitMQ

- Instagram
- Indeed.com
- MailboxApp

Some users of RabbitMQ

- Instagram
- Indeed.com
- MailboxApp
- Mercado Libre

Some users of RabbitMQ

- Instagram
- Indeed.com
- MailboxApp
- Mercado Libre
- NHS

Some users of RabbitMQ

- Instagram
- Indeed.com
- MailboxApp
- Mercado Libre
- NHS
- Mozilla

http://www.rabbitmq.com/download.html

Unix - Mac - Windows

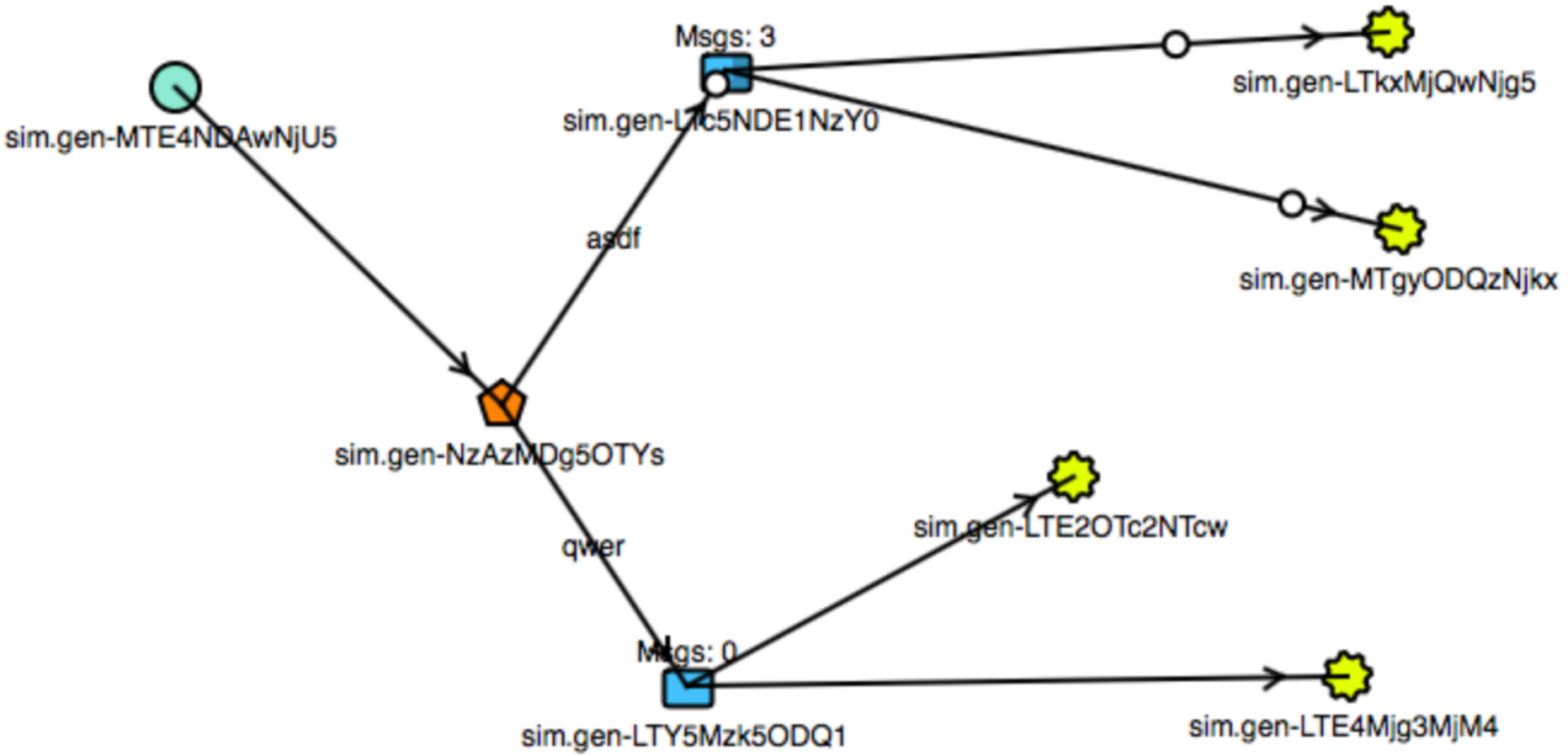
Messaging with RabbitMQ

A demo with the RabbitMQ Simulator

https://github.com/RabbitMQSimulator/RabbitMQSimulator

http://tryrabbitmq.com

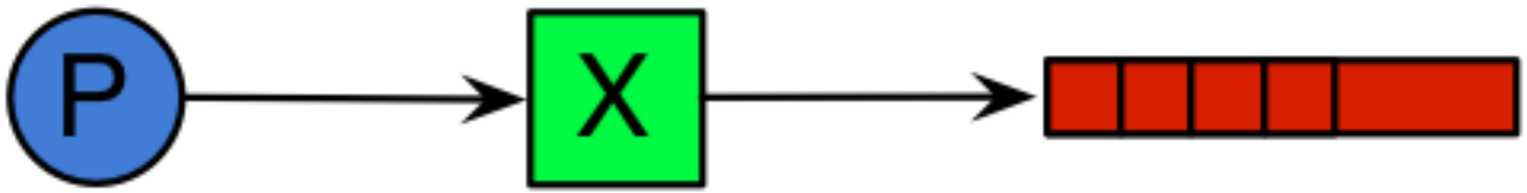
RabbitMQ Simulator



RabbitMQ Internals

A day in the life of a message

A day in the life of a message



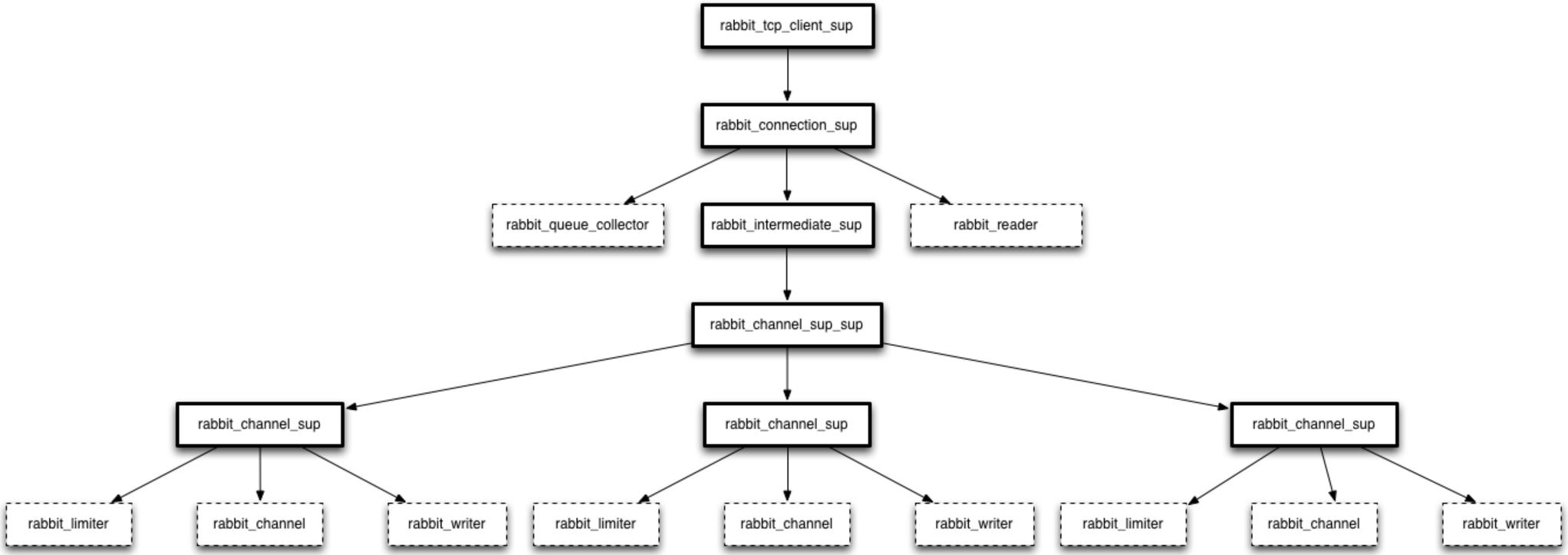
A day in the life of a message

```
$conn = new AMQPConnection(HOST, PORT, USER, PASS, VHOST);
$ch = $conn->channel();

$ch->queue_declare($queue, false, true, false, false);

$ch->exchange_declare($exchange, 'direct', false, true, false);

$ch->queue_bind($queue, $exchange);

$msg_body = implode(' ', array_slice($argv, 1));
$msg = new AMQPMessage($msg_body, array('delivery_mode' => 2));

$ch->basic_publish($msg, $exchange);
```

What happens here?

```
$conn = new AMQPConnection(HOST, PORT, USER, PASS, VHOST);
$ch = $conn->channel();
```

What happens here?



Erlang

Erlang App

- Processes (probably thousands)
- They communicate sending messages
- Each process has a message queue (don’t confuse with RabbitMQ queues)
- Virtual Machine has preemptive scheduler

Read More Here:

http://jlouisramblings.blogspot.ru/2013/01/how-erlang-does-scheduling.html

Erlang code structure

- Modules
- Functions
- Function Arity
- Arguments

```
M, F, A = Module, Function, Arguments
```

rabbit_client_sup.erl

```
-module(rabbit_client_sup).

-behaviour(supervisor2).

-export([start_link/1, start_link/2, start_link_worker/2]).

-export([init/1]).

-include("rabbit.hrl").
```

rabbit_client_sup.erl

```
start_link(Callback) ->
    supervisor2:start_link(?MODULE, Callback).

start_link(SupName, Callback) ->
    supervisor2:start_link(SupName, ?MODULE, Callback).

start_link_worker(SupName, Callback) ->
    supervisor2:start_link(SupName, ?MODULE, {Callback, worker}).

init({M,F,A}) ->
    {ok, {{simple_one_for_one, 0, 1},
          [{client, {M,F,A}, temporary, infinity, supervisor, [M]}]}};
init({{M,F,A}, worker}) ->
    {ok, {{simple_one_for_one, 0, 1},
          [{client, {M,F,A}, temporary, ?MAX_WAIT, worker, [M]}]}}.
```

Supervision Trees

Supervision tree

- Worker Processes
- Supervisor Processes
- Supervision tree as a **hierarchical** arrangement of processes

http://www.erlang.org/doc/man/supervisor.html

rabbit_client_sup

```
init({M,F,A}) ->
    {ok, {{simple_one_for_one, 0, 1},
          [{client, {M,F,A}, temporary,
            infinity, supervisor, [M]}]}};
```

Child Spec - restart strategies

- one_for_one: only restart failing process
- one_for_all: restart failing process and all siblings
- simple_one_for_one: simplified version of one_for_one
- MaxR: maximum restarts allowed
- MaxT: in MaxT seconds

Child Specification

```
child_spec() = {Id,StartFunc,Restart,Shutdown,Type,Modules}
 Id = term()
 StartFunc = {M,F,A}
  M = F = atom()
  A = [term()]
 Restart = permanent | transient | temporary
 Shutdown = brutal_kill | int()>0 | infinity
 Type = worker | supervisor
 Modules = [Module] | dynamic
  Module = atom()
```

rabbit_client_sup

```
init({M,F,A}) ->
    {ok, {{simple_one_for_one, 0, 1},
          [{client, {M,F,A}, temporary,
            infinity, supervisor, [M]}]}};
```

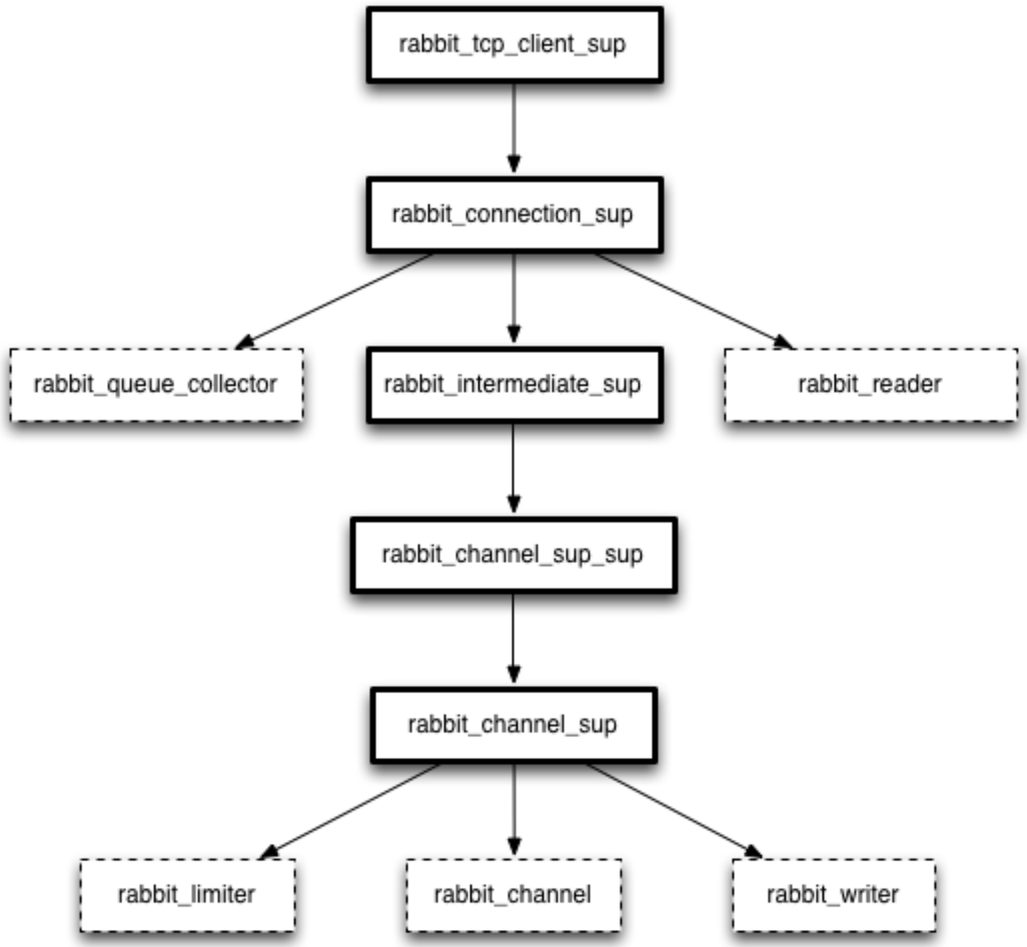
Child Spec - restart

- **permanent**: should always be restarted
- **temporary**: should never be restarted
- **transient**: should only be restarted if terminated abnormally

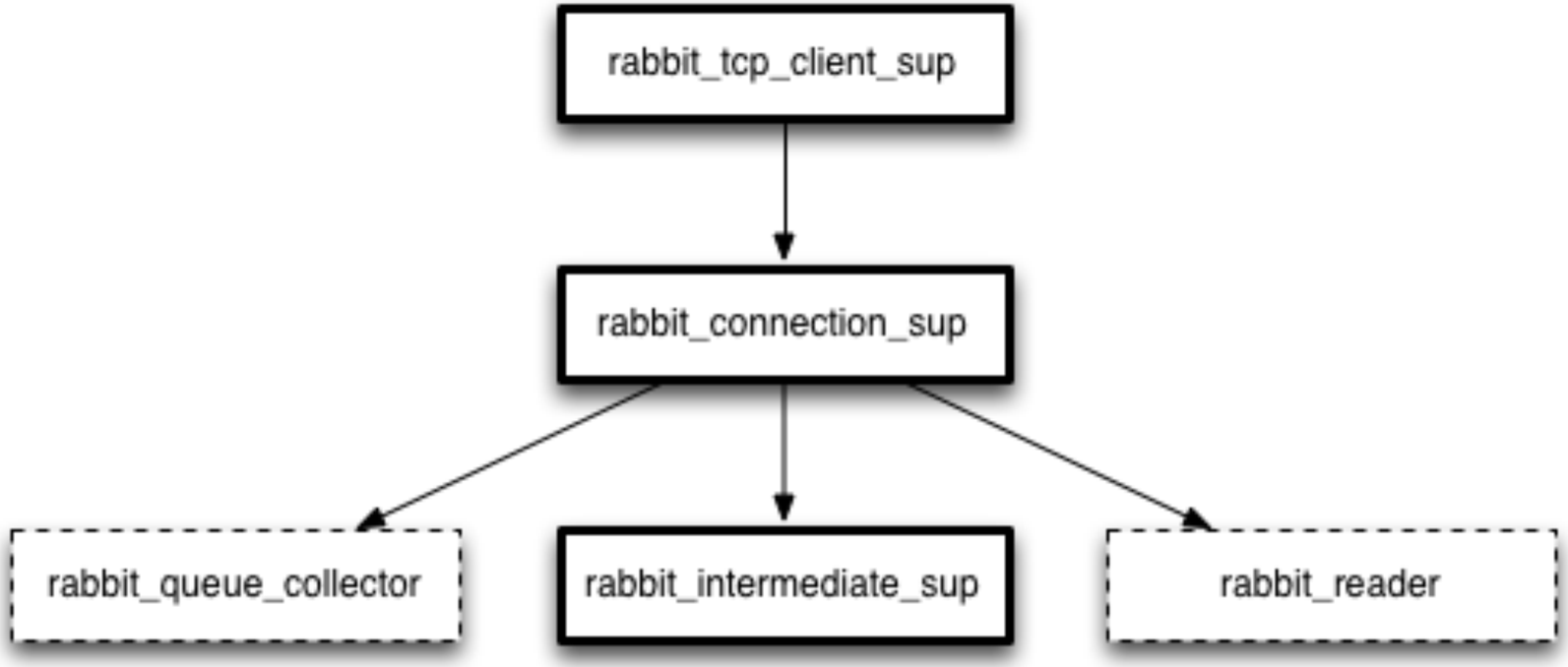
Child Spec - shutdown

- **brutal_kill**: child is terminated immediately.
- **timeout in seconds**: supervisor waits for timeout before terminating children.
- **infinity**: give enough timeout to children to shutdown its own supervision tree.

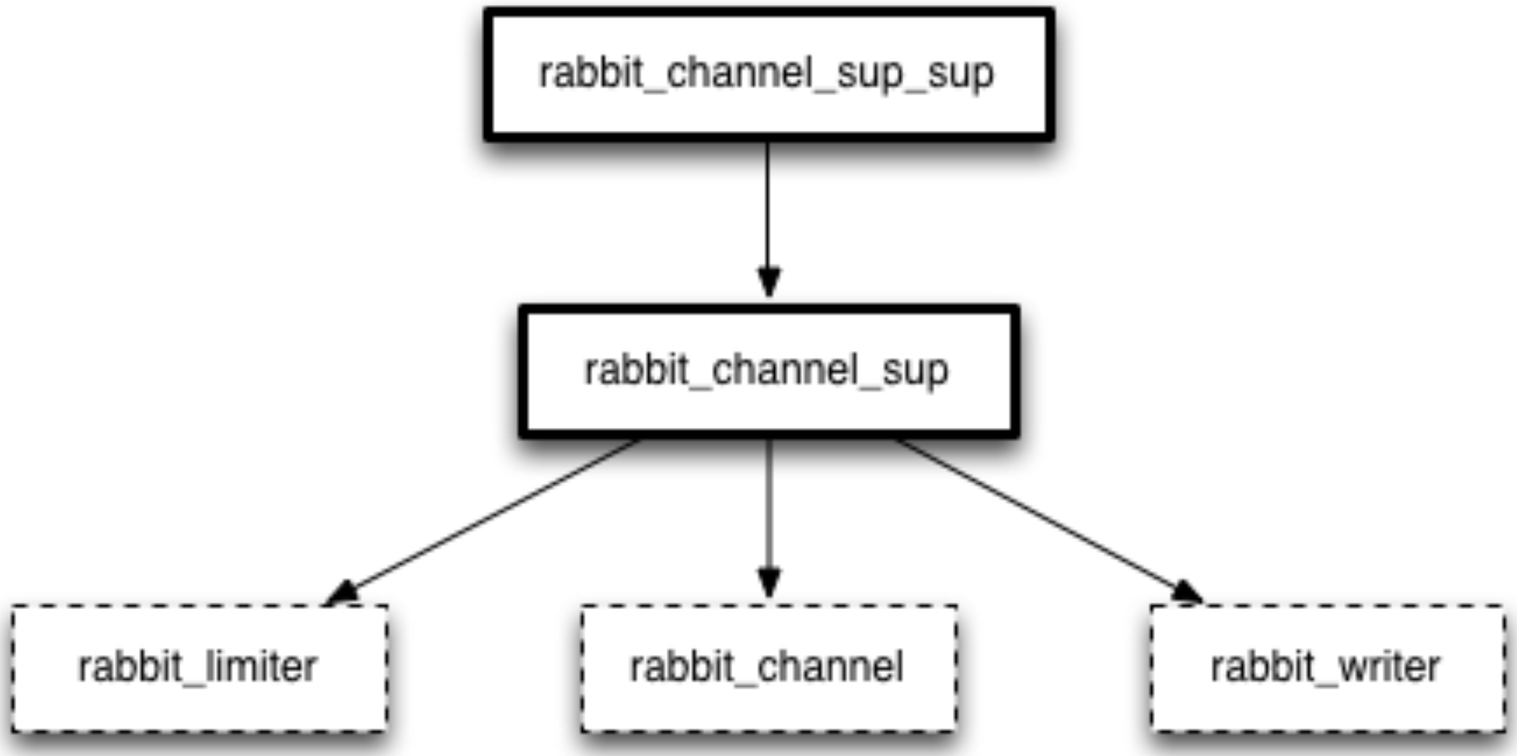
Connection Supervision Tree



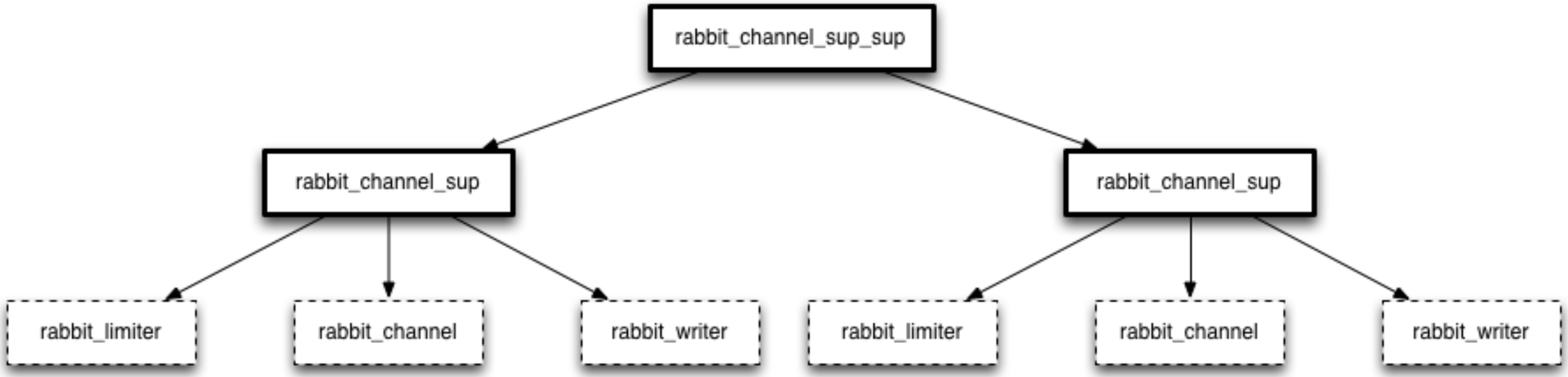
Connection Supervision Tree



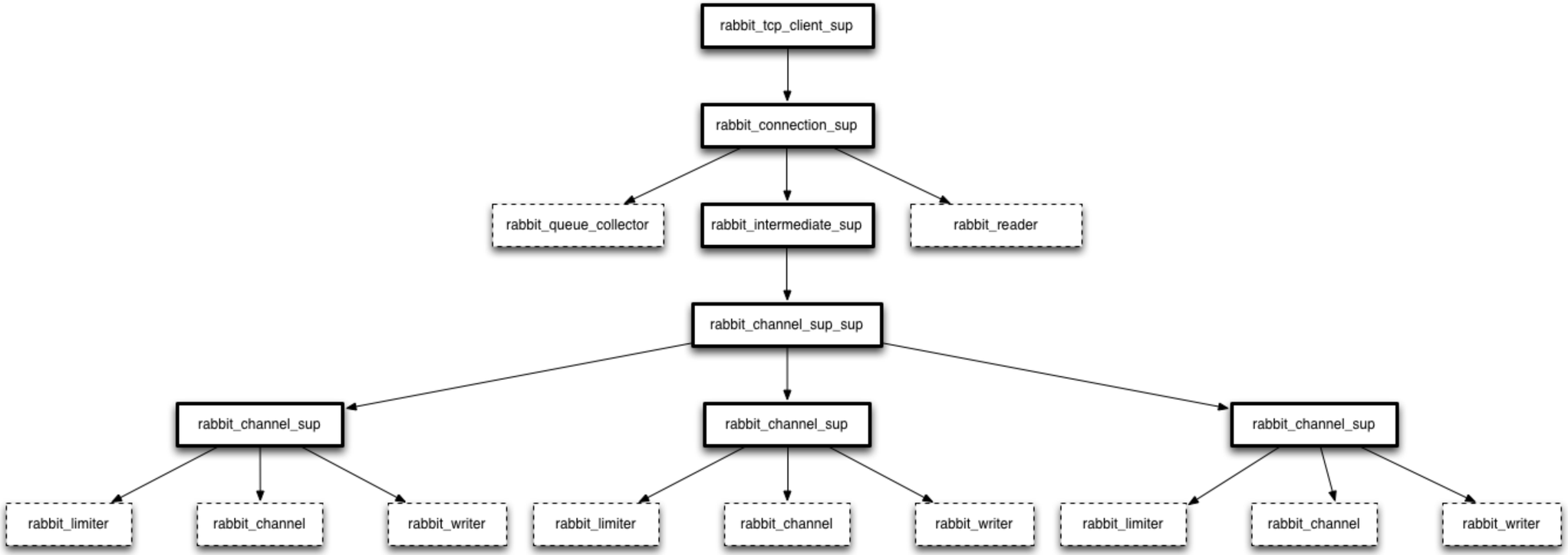
Connection Supervision Tree



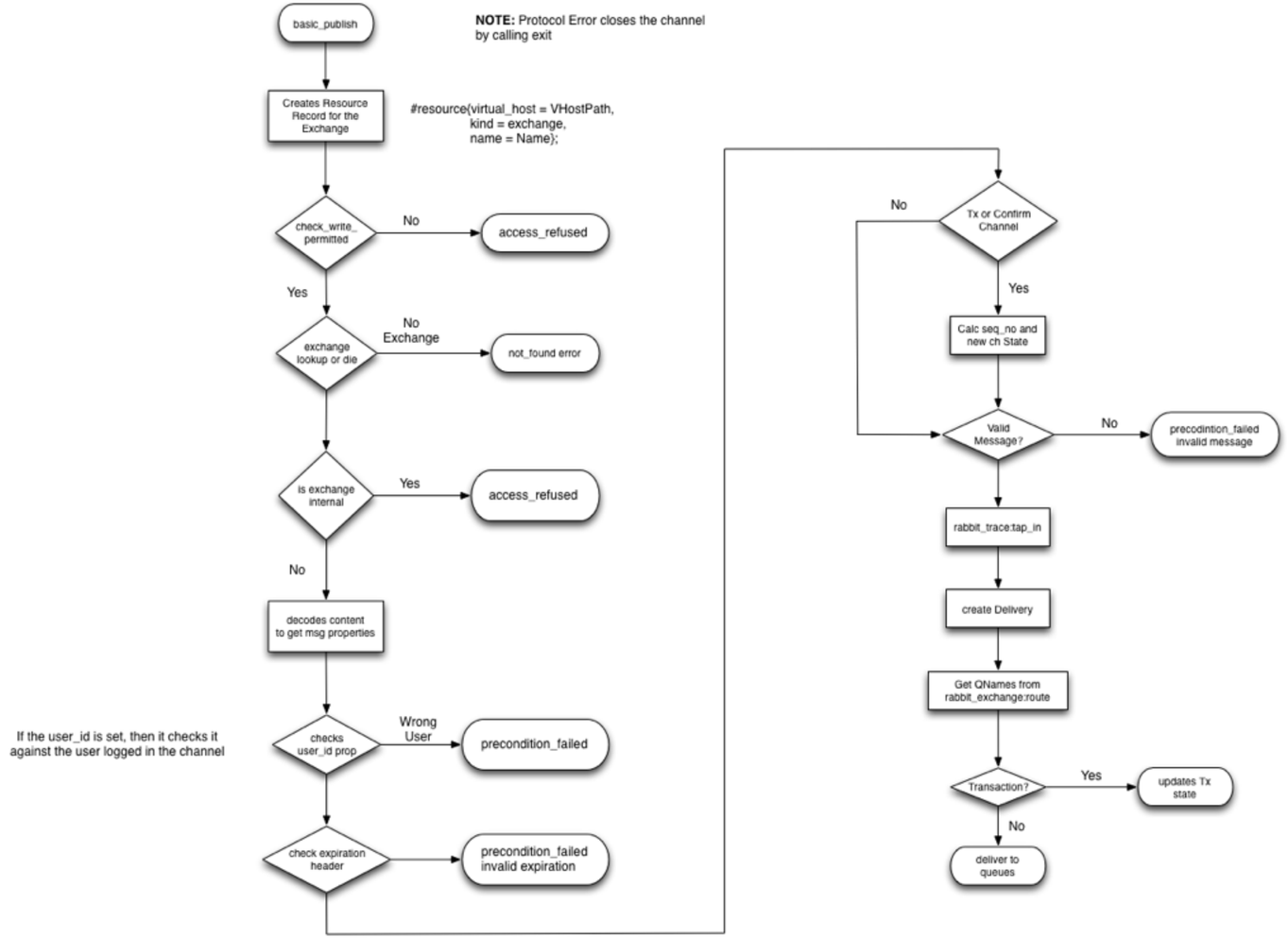
Connection Supervision Tree



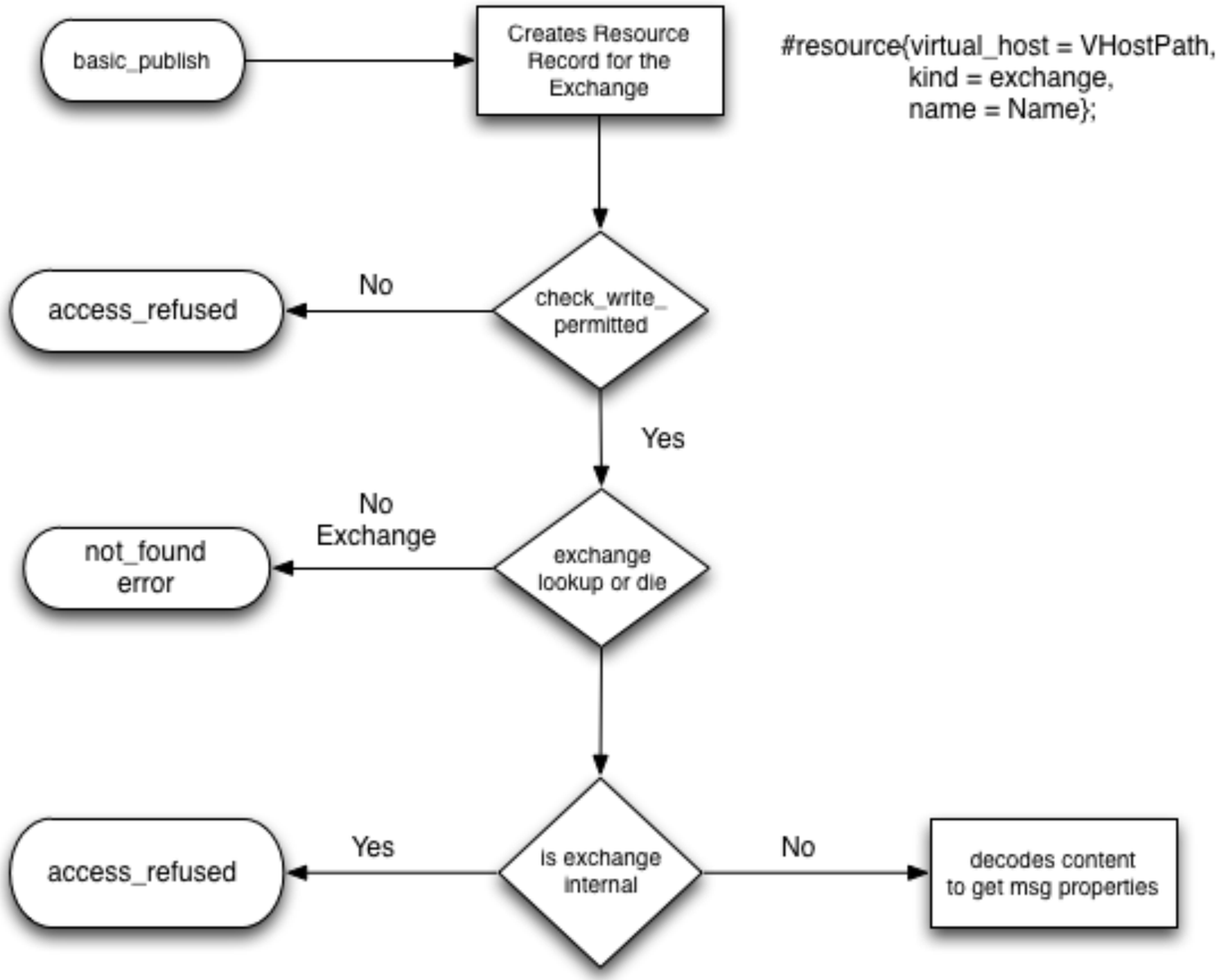
Connection Supervision Tree



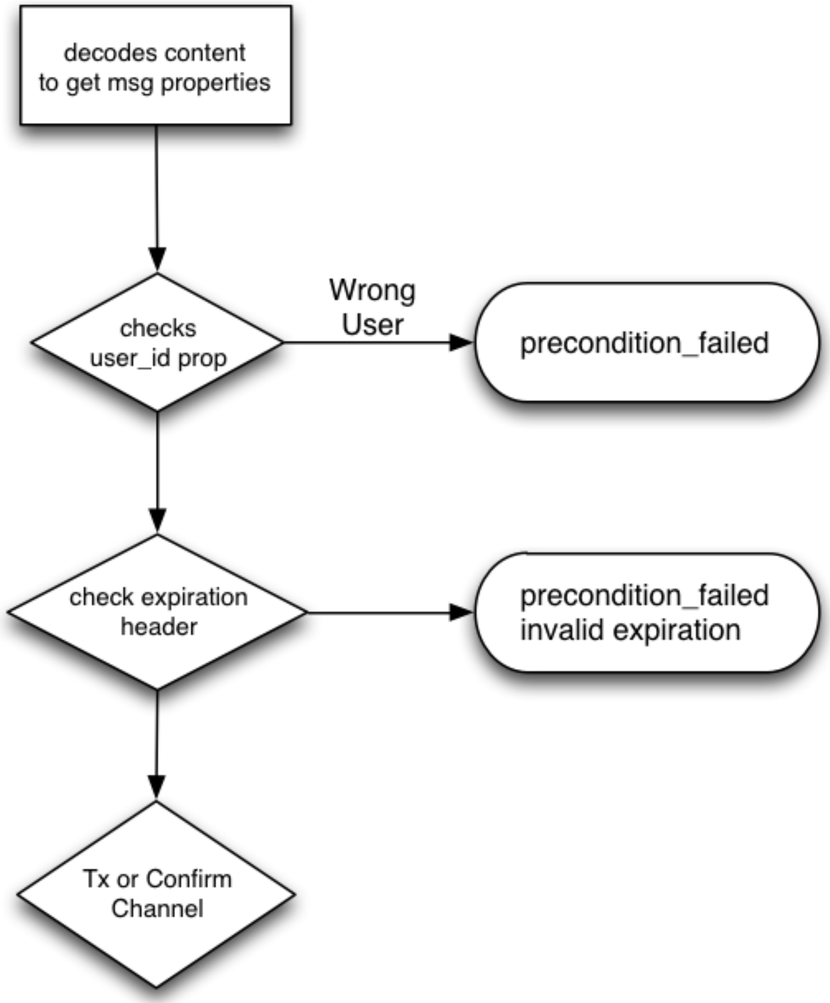
A day in the life of a message



A day in the life of a message



A day in the life of a message



Intermezzo: pattern matching

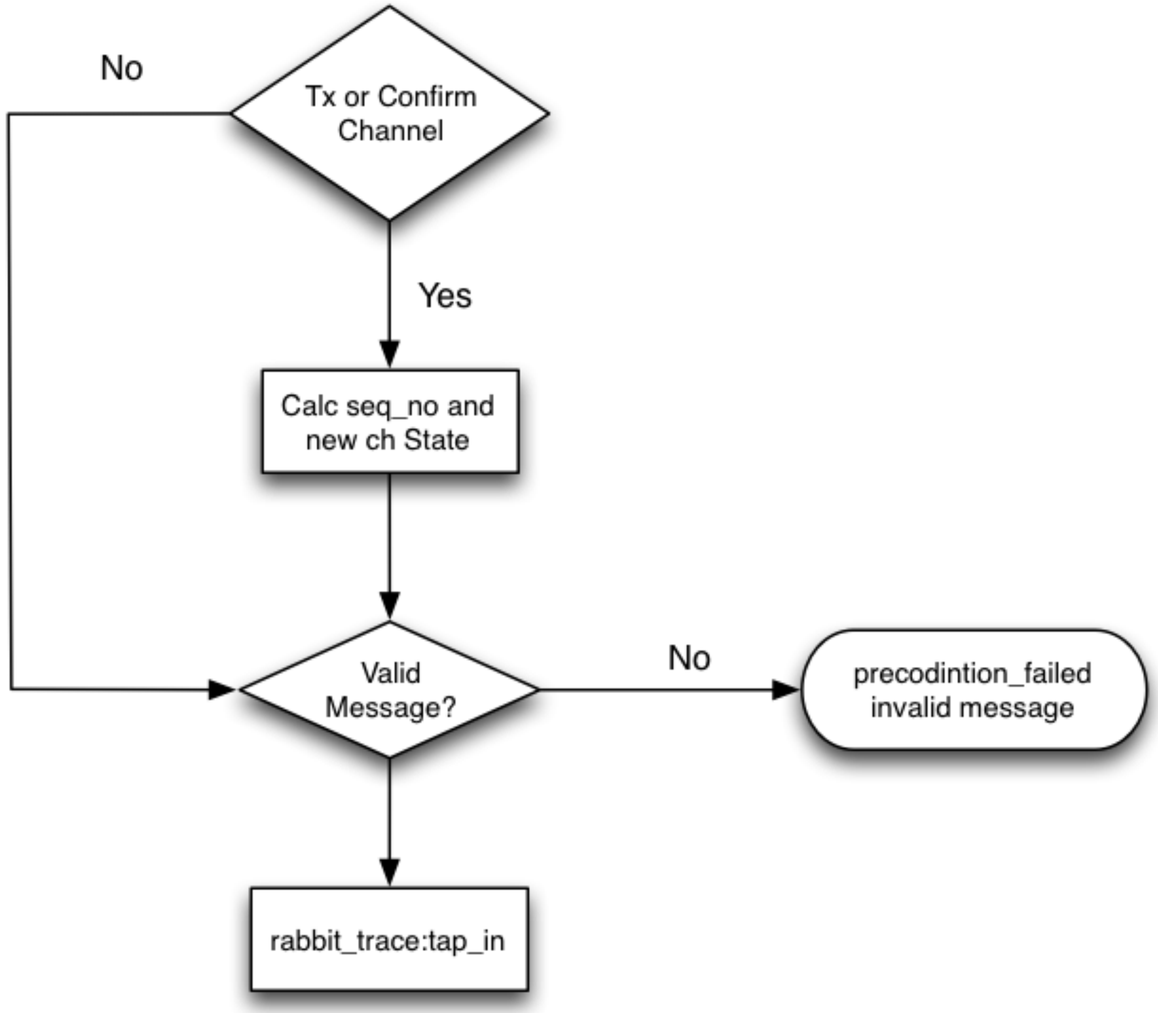
```
check_user_id_header(#'P_basic'{user_id = undefined}, _) ->
    ok;
check_user_id_header(#'P_basic'{user_id = Username},
                     #ch{user = #user{username = Username}}) ->
    ok;
check_user_id_header(#'P_basic'{user_id = Claimed},
                     #ch{user = #user{username = Actual,
                                      tags     = Tags}}) ->
    case lists:member(impersonator, Tags) of
        true  -> ok;
        false -> precondition_failed(
                   "user_id property set to '~s' but authenticated user was "
                   "'~s'", [Claimed, Actual])
    end.
```

http://videlalvaro.github.io/2013/09/rabbitmq-validating-user-ids-with-erlang-pattern-matching.html

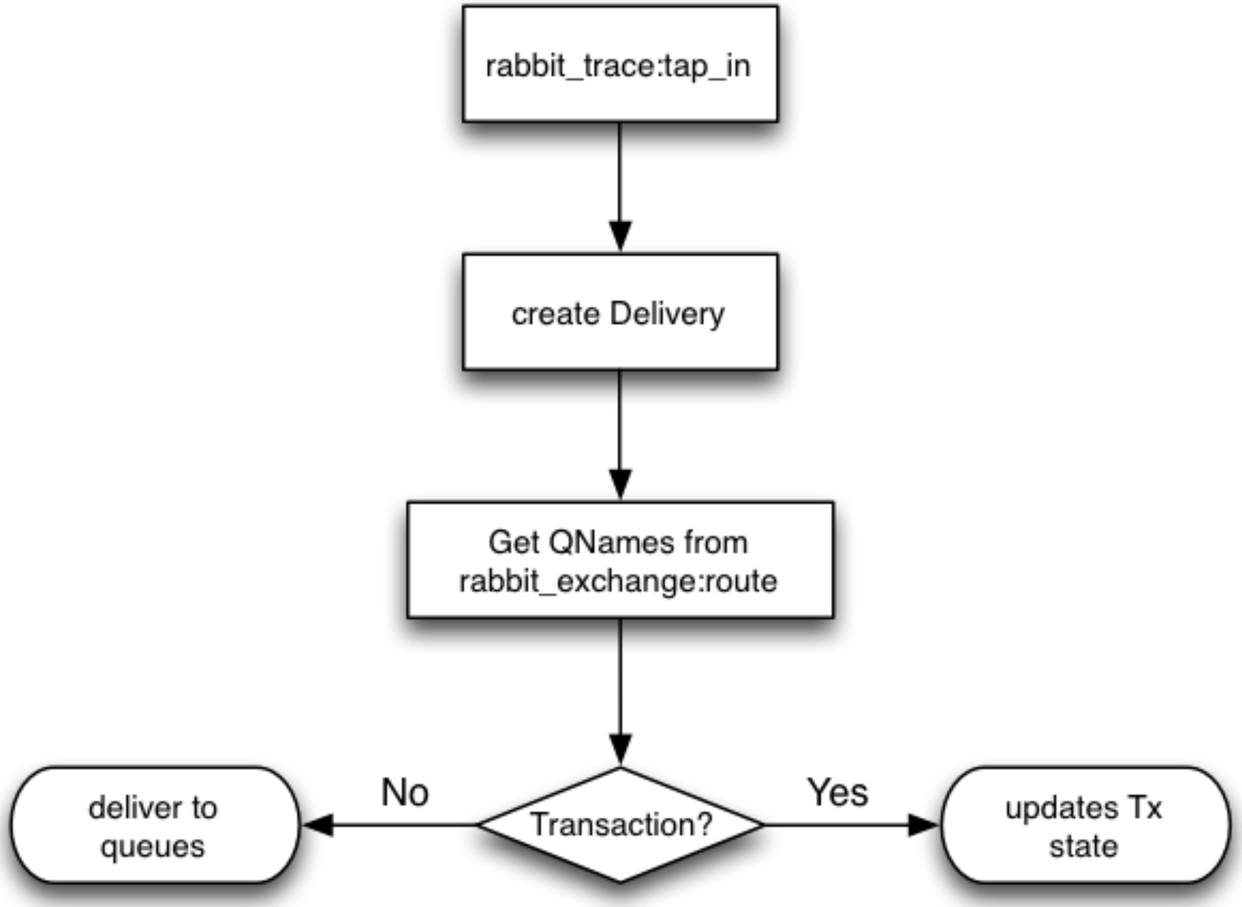
Read More Here:

http://videlalvaro.github.io/2013/09/rabbitmq-validating-user-ids-with-erlang-pattern-matching.html

A day in the life of a message

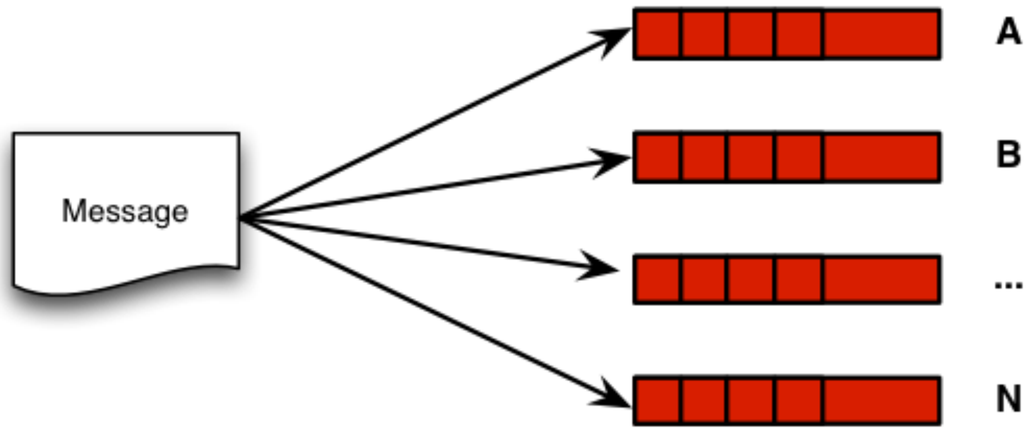


A day in the life of a message

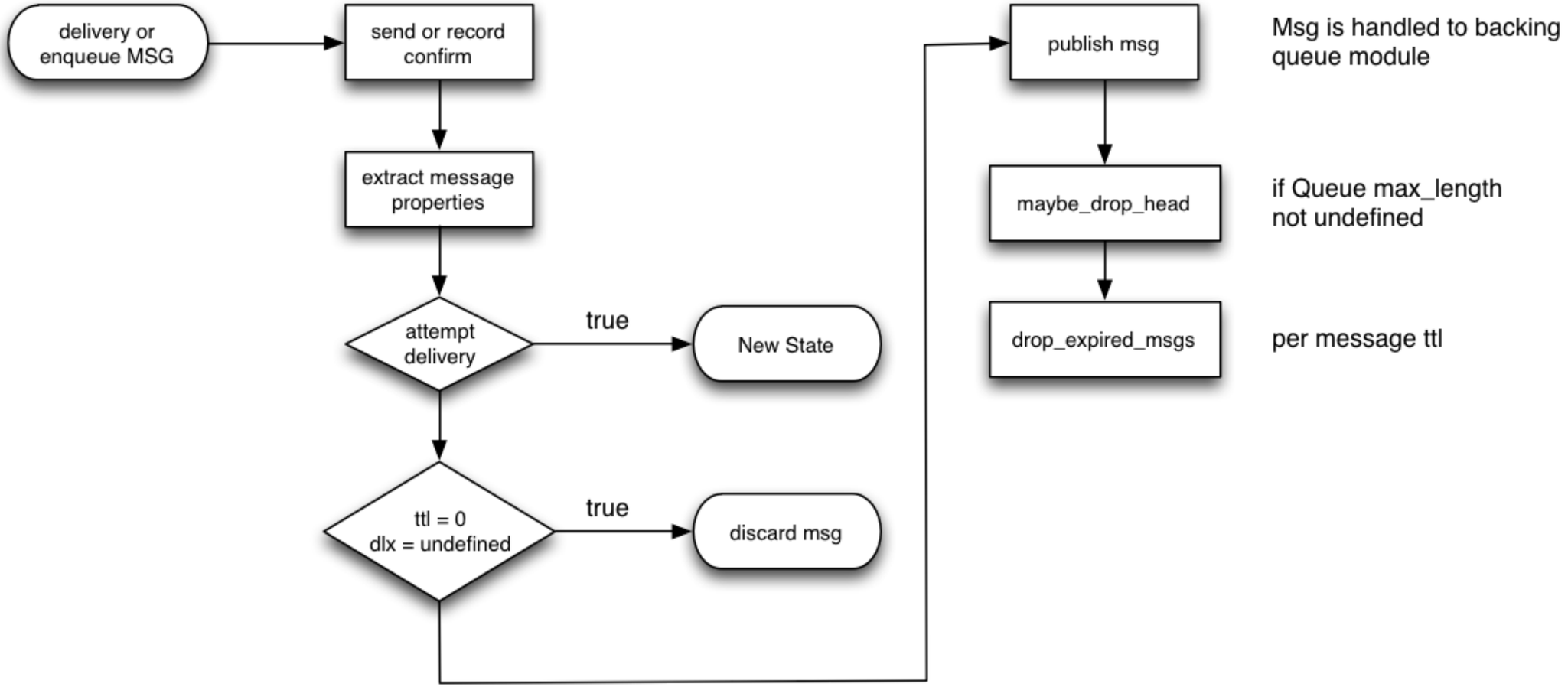


A day in the life of a message

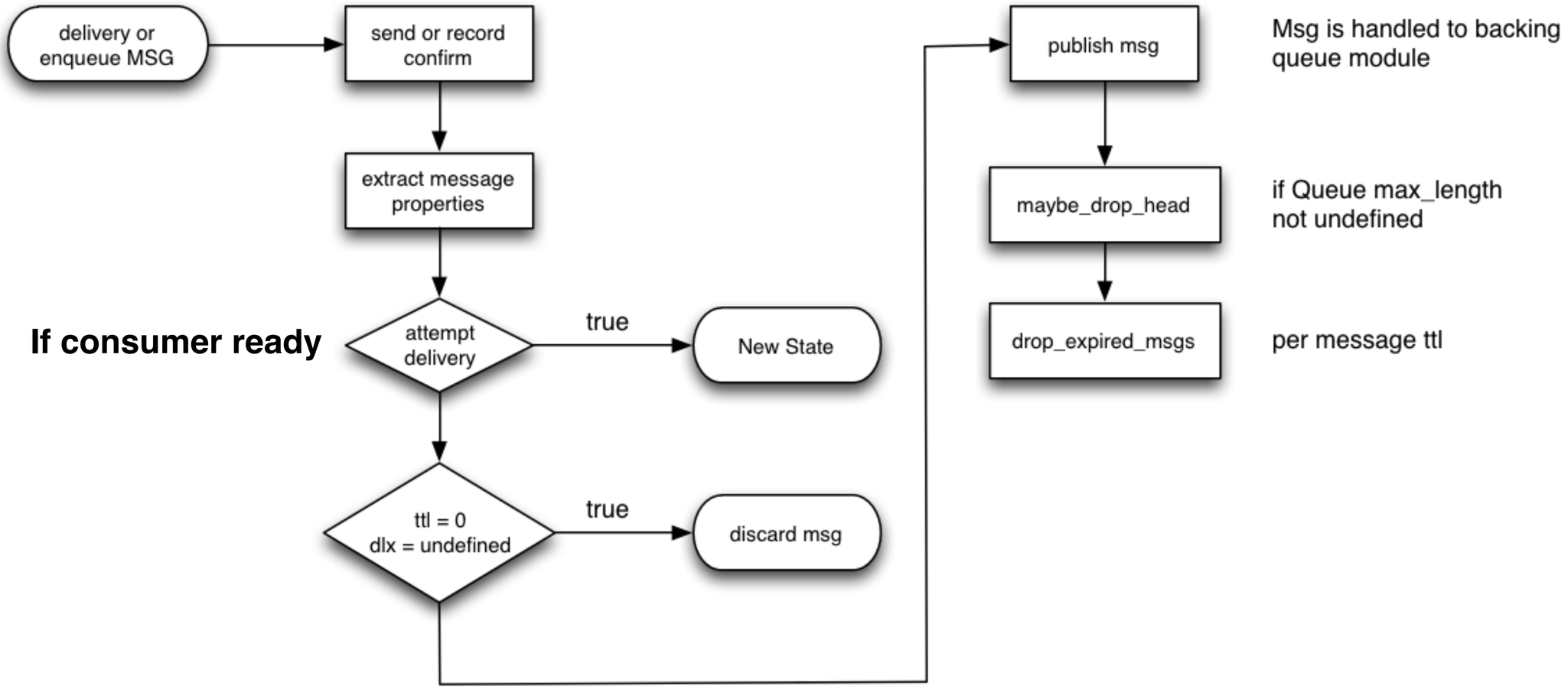
We have a list of queues where the channel will deliver the messages



A day in the life of a message



A day in the life of a message



RabbitMQ Message Store

- Made specifically for messaging
- Keeps messages in memory and (sometimes) on disk
- Bounded by disk size
- per node message store (transient, persistent)
- per queue “queue index”

RabbitMQ Message Store

- Message written to disk when:
  - Message published as persistent (delivery_mode = 2)
  - Memory pressure

Read More Here:

Read More Here:

Why a custom message store?

http://www.rabbitmq.com/blog/2011/01/20/rabbitmq-backing-stores-databases-and-disks/

Read More Here:

Why a custom message store?

http://www.rabbitmq.com/blog/2011/01/20/rabbitmq-backing-stores-databases-and-disks/

How the message store compacts files

http://hg.rabbitmq.com/rabbitmq-server/file/56d190fd4ea3/src/rabbit_msg_store.erl#l181

Read More Here:

Why a custom message store?

http://www.rabbitmq.com/blog/2011/01/20/rabbitmq-backing-stores-databases-and-disks/

How the message store compacts files

http://hg.rabbitmq.com/rabbitmq-server/file/56d190fd4ea3/src/rabbit_msg_store.erl#l181

How the message store reacts to memory pressure

http://hg.rabbitmq.com/rabbitmq-server/file/56d190fd4ea3/src/rabbit_variable_queue.erl#l35

Credit Flow

Prevent processes from overflowing each other's mailboxes

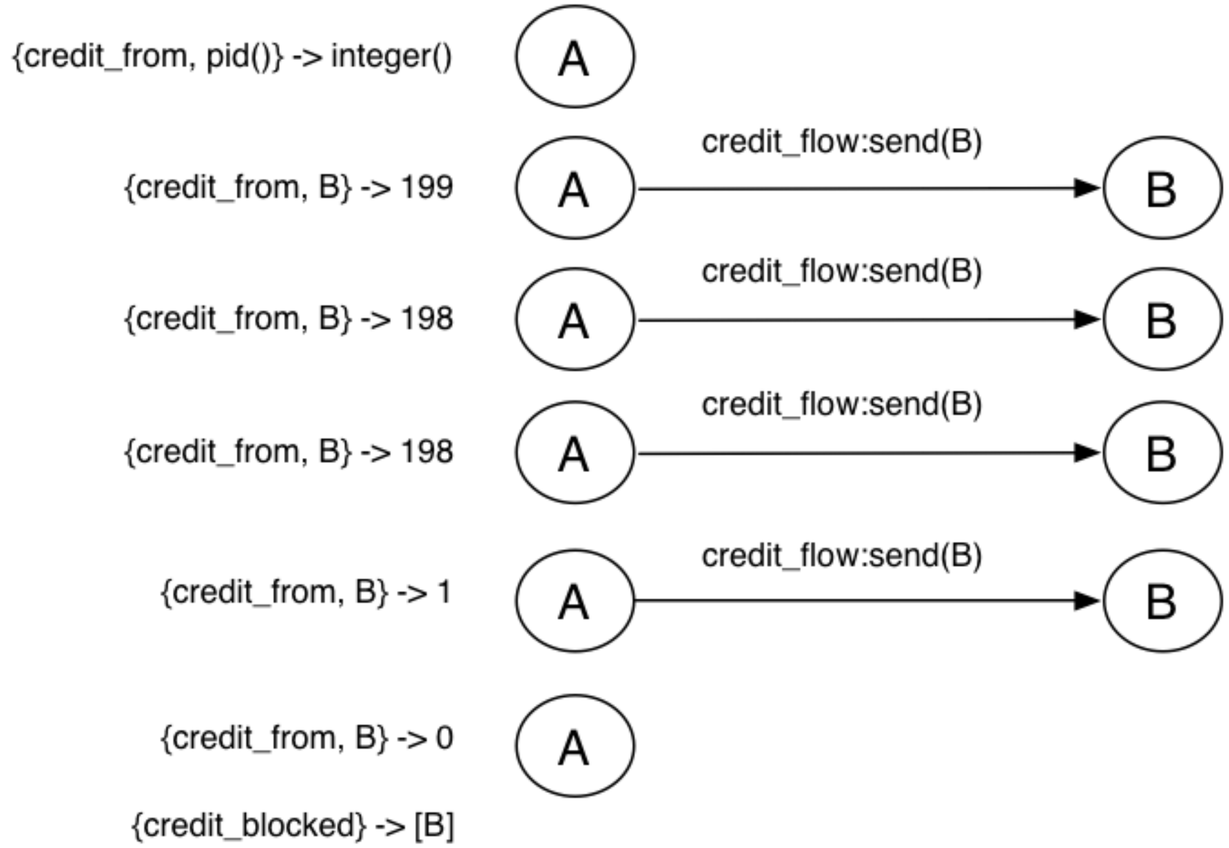
Credit Specification

```
{InitialCredit, MoreCreditAfter}
```

Credit Specification

```
{200, 50}
```

Process **A** sends messages to **B**



Process will

- Grant more credit
- Block other processes
- Delay granting credits

Read More Here:

http://videlalvaro.github.io/2013/09/rabbitmq-internals-credit-flow-for-erlang-processes.html

RabbitMQ Behaviours

RabbitMQ Behaviours

- Add a common interface for different components

RabbitMQ Behaviours

- Add a common interface for different components
  - Exchanges

RabbitMQ Behaviours

- Add a common interface for different components
  - Exchanges
  - Queues

RabbitMQ Behaviours

- Add a common interface for different components
  - Exchanges
  - Queues
  - Decorators

RabbitMQ Behaviours

- Add a common interface for different components
  - Exchanges
  - Queues
  - Decorators
  - Authentication methods

RabbitMQ Behaviours

- Add a common interface for different components
  - Exchanges
  - Queues
  - Decorators
  - Authentication methods
- Add extensibility

rabbit_exchange_type

```
-module(rabbit_exchange_type).

-callback description() -> [proplists:property()].

-callback serialise_events() -> boolean().

-callback route(rabbit_types:exchange(), rabbit_types:delivery()) ->
    rabbit_router:match_result().

-callback validate(rabbit_types:exchange()) -> 'ok'.

-callback validate_binding(rabbit_types:exchange(), rabbit_types:binding()) ->
    rabbit_types:ok_or_error({'binding_invalid', string(), [any()]}).
```

rabbit_exchange_type

- You can add your own exchange type via plugins
  - consistent hash exchange
  - random exchange
  - recent history exchange
  - riak exchange

RabbitMQ Behaviours

- rabbit_auth_backend (config)
- rabbit_msg_store_index (config)
- rabbit_backing_queue (config)

RabbitMQ Behaviours

- rabbit_auth_backend (config)
- rabbit_msg_store_index (config)
- rabbit_backing_queue (config)

```
[{rabbit,
  [{auth_backends, [rabbit_auth_backend_http,
                    rabbit_auth_backend_internal]}]
 }].
```

RabbitMQ Behaviours

- rabbit_auth_mechanism (registry)
- rabbit_exchange_decorator (registry)
- rabbit_exchange_type (registry)
- rabbit_mirror_queue_mode (registry)
- rabbit_policy_validator (registry)
- rabbit_queue_decorator (registry)
- rabbit_runtime_parameter (registry)

RabbitMQ Behaviours

- rabbit_auth_mechanism (registry)
- rabbit_exchange_decorator (registry)
- rabbit_exchange_type (registry)
- rabbit_mirror_queue_mode (registry)
- rabbit_policy_validator (registry)
- rabbit_queue_decorator (registry)
- rabbit_runtime_parameter (registry)

```
-rabbit_boot_step({?MODULE,
                   [{description, "exchange type direct"},
                    {mfa,         {rabbit_registry, register,
                                   [exchange, <<"direct">>, ?MODULE]}},
                    {requires,    rabbit_registry},
                    {enables,     kernel_ready}]}).
```

RabbitMQ Federation Plugin

- Message replication across WANs
- Uses an exchange decorator
- Uses a queue decorator
- Uses parameters
- Uses policies

RabbitMQ Federation Plugin

```
rabbitmqctl set_parameter federation-upstream my-upstream \
'{"uri":"amqp://server-name","expires":3600000}'
```

RabbitMQ Federation Plugin

```
rabbitmqctl set_parameter federation-upstream my-upstream \
'{"uri":"amqp://server-name","expires":3600000}'
```

```
rabbitmqctl set_policy federate-me "^amq\." '{"federation-upstream-
set":"all"}'
```

RabbitMQ Federation Plugin

```
-module(rabbit_federation_queue).

-rabbit_boot_step({?MODULE,
                   [{description, "federation queue decorator"},
                    {mfa, {rabbit_registry, register,
                           [queue_decorator, <<"federation">>, ?MODULE]}},
                    {requires, rabbit_registry},
                    {enables, recovery}]}).

-behaviour(rabbit_queue_decorator).
```

RabbitMQ Federation Plugin

```
-module(rabbit_federation_exchange).

-rabbit_boot_step({?MODULE,
                   [{description, "federation exchange decorator"},
                    {mfa, {rabbit_registry, register,
                           [exchange_decorator, <<"federation">>, ?MODULE]}},
                    {requires, rabbit_registry},
                    {enables, recovery}]}).

-behaviour(rabbit_exchange_decorator).
```

Read More Here:

http://www.rabbitmq.com/federation.html

Read More Here:

Making sure the user provided the right behaviour:

http://videlalvaro.github.io/2013/09/rabbitmq-internals-validating-erlang-behaviours.html

How RabbitMQ prevents arbitrary code execution:

http://videlalvaro.github.io/2013/09/rabbitmq-sanitzing-user-input-in-erlang.html

RabbitMQ Boot Step System:

https://github.com/videlalvaro/rabbit-internals/blob/master/rabbit_boot_process.md

RabbitMQ uses Erlang's features to be robust, fault tolerant and very extensible

Go grab the source code!

http://hg.rabbitmq.com/rabbitmq-server/

Questions?

Thanks

Alvaro Videla - @old_sound