

RabbitMQ Operations

About me

About me

- RabbitMQ staff engineer at Pivotal

About me

- RabbitMQ staff engineer at Pivotal
- @michaelklishin just about everywhere

About this talk

About this talk

- Brain dump from years of answering questions

About this talk

- Brain dump from years of answering questions
- Focusses on the most recent release (3.5.6)

Provisioning

Provisioning

- Be aware of mirrors: GitHub, Bintray, ...

Provisioning

- Be aware of mirrors: GitHub, Bintray, ...
- Looking into community-hosted mirrors

Provisioning

- Be aware of mirrors: GitHub, Bintray, ...
- Looking into community-hosted mirrors
- Use packages + Chef/Puppet/...

OS resources

OS resources

- Modern Linux defaults are **absolutely inadequate** for servers

ulimit -n default: 1024

1024 FILE DESCRIPTORS



**WOULD BE ENOUGH FOR
ANYBODY**

memegenerator.net

Set ulimit -n and fs.file-max to 500K
and forget about it

TCP keepalive timeout:
from 11 minutes to over 2
hours by default

```
net.ipv4.tcp_keepalive_time = 6  
net.ipv4.tcp_keepalive_intvl = 3  
net.ipv4.tcp_keepalive_probes = 3
```

enable client **heartbeats**, e.g. with an
interval of 6-12 seconds

OS resources

- Modern Linux defaults are **absolutely inadequate** for servers
- Tuning for **throughput** vs. high number of **concurrent connections**

Throughput: larger TCP buffers


```
net.core.rmem_max = 16777216  
net.core.wmem_max = 16777216
```

```
rabbit.hipe_compile = true  
(only on Erlang 17.x or 18.x)
```

Concurrent connections: smaller TCP
buffers, low tcp_fin_timeout,
tcp_tw_reuse = 1, ...

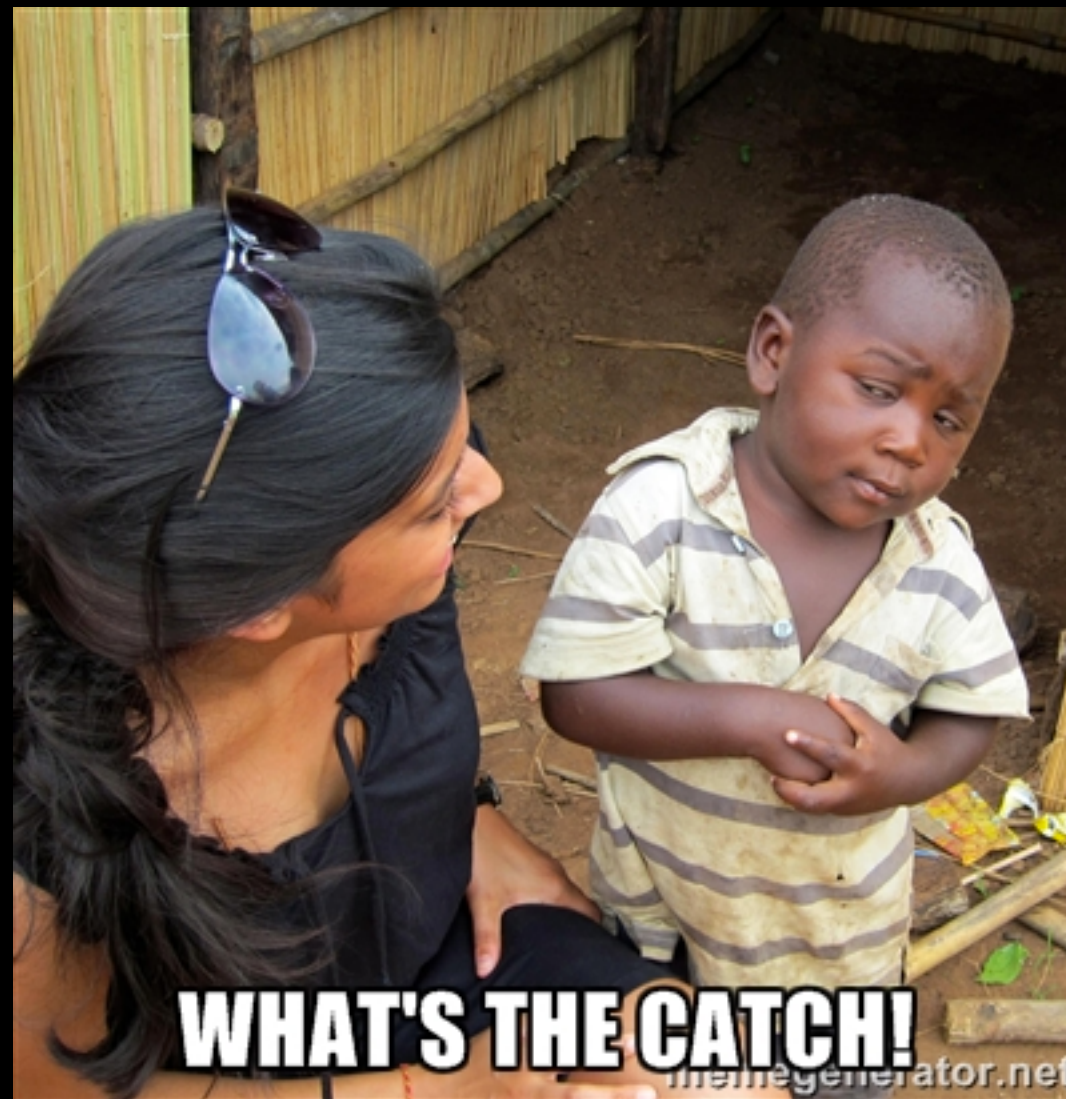
rabbit.tcp_listen_options.sndbuf

rabbit.tcp_listen_options.rcvbuf

rabbit.tcp_listen_options.backlog

Reduce per connection RAM use by 10x

```
rabbit.tcp_listen_options.sndbuf = 16384  
rabbit.tcp_listen_options.rcbuf = 16384
```



WHAT'S THE CATCH!

memegenerator.net

Reduce per connection
RAM use by 10x

Throughput drops by a comparable amount

```
net.ipv4.tcp_fin_timeout = 5
```



```
net.ipv4.tcp_tw_reuse = 1
```

Careful with tcp_tw_reuse behind NAT*

* <http://vincent.bernat.im/en/blog/2014-tcp-time-wait-state-linux.html>

```
net.core.somaxconn = 4096
```

<http://www.rabbitmq.com/networking.html>

Disk space

Disk space

- Pay attention to what partition `/var/lib` ends up on

Disk space

- Pay attention to what partition `/var/lib` ends up on
- Transient messages can be paged to disk

Disk space

- Pay attention to what partition `/var/lib` ends up on
- Transient messages can be paged to disk
- RabbitMQ's disk monitor isn't supported on all platforms

RAM usage

RAM usage

- `rabbit.vm_memory_high_watermark`

RAM usage

- `rabbit.vm_memory_high_watermark`
- `rabbit.vm_memory_high_watermark_paging_ratio`

rabbitmqctl status
rabbitmqctl report

RAM usage

- `rabbit.vm_memory_high_watermark`
- `rabbit.vm_memory_high_watermark_paging_ratio`
- Significant paging efficiency improvements in 3.5.5-3.5.6

RAM usage

- `rabbit.vm_memory_high_watermark`
- `rabbit.vm_memory_high_watermark_paging_ratio`
- Significant paging efficiency improvements in 3.5.5-3.5.6
- Disable `rabbit.fhc_read_buffering` (3.5.6+)


```
rabbitmqctl eval  
'file_handle_cache:clear_read_cache().'
```

recon

Ability to set VM RAM watermark as absolute
value is coming in 3.6

Stats collector falls behind

Stats collector falls behind

- Management DB stats collector can get overwhelmed

Stats collector falls behind

- Management DB stats collector can get overwhelmed
- Key symptom: disproportionately higher RAM use on the node that hosts management DB

```
rabbitmqctl eval 'P = whereis(rabbit_mgmt_db), erlang:process_info(P).'
```



```
rabbit.collect_statistics_interval = 30000
```

```
rabbitmq_management.rates_mode = none
```

```
rabbitmqctl eval 'P = whereis(rabbit_mgmt_db), erlang:exit(P, please_crash).'
```

Parallel stats collector is coming in 3.7

Cluster formation

Cluster formation

- Node restart order dependency

Cluster formation

- Node restart order dependency
- github.com/rabbitmq/rabbitmq-clusterer

Cluster formation

- Node restart order dependency
- github.com/rabbitmq/rabbitmq-clusterer
- github.com/aweber/rabbitmq-autocluster

Backups

How do I back up?

- `cp $RABBITMQ_MNESIA_DIR + tar`

How do I back up?

- `cp $RABBITMQ_MNESIA_DIR + tar`
- Replicate everything off-site with exchange federation + set message TTL via a policy

Hostname changes


```
rabbitmqctl rename_cluster_node [old name] [new name]
```


Network partition handling

Network partition handling

- When in doubt, use “autoheal”

Network partition handling

- When in doubt, use “autoheal”
- “Merge” is coming but has very real downsides, too

Misc

Misc

- Don't use default vhost and/or credentials

Misc

- Don't use default vhost and/or credentials
- Don't use 32-bit Erlang

Misc

- Don't use default vhost and/or credentials
- Don't use 32-bit Erlang
- Use reasonably up-to-date releases

Misc

- Don't use default vhost and/or credentials
- Don't use 32-bit Erlang
- Use reasonably up-to-date releases
- Participate in [rabbitmq-users](#)

Misc

- OCF resource template from Fuel (by Mirantis)

Misc

- OCF resource template from Fuel (by Mirantis)
- Use TLS

Coming in 3.6

Coming in 3.6

- In process file buffering disabled by default

Coming in 3.6

- In process file buffering disabled by default
- Queue master to node distribution strategies

Coming in 3.6

- In process file buffering disabled by default
- Queue master to node distribution strategies
- SHA-256 (or 512) for password hashing

Coming in 3.6

- In process file buffering disabled by default
- Queue master to node distribution strategies
- SHA-256 (or 512) for password hashing
- More responsive management UI with pagination

Coming in 3.6

- In process file buffering disabled by default
- Queue master to node distribution strategies
- SHA-256 (or 512) for password hashing
- More responsive management UI with pagination
- Streaming rabbitmqctl

Coming past 3.6

Coming past 3.6

- Pluggable cluster formation (à la Elasticsearch)

Coming past 3.6

- Pluggable cluster formation (à la Elasticsearch)
- On disk data recovery tools

Coming past 3.6

- Pluggable cluster formation (à la Elasticsearch)
- On disk data recovery tools
- Better CLI tools

Coming past 3.6

- Pluggable cluster formation (à la Elasticsearch)
- On disk data recovery tools
- Better CLI tools
- Easier off-site replication

Coming past 3.6

- Pluggable cluster formation (à la Elasticsearch)
- On disk data recovery tools
- Better CLI tools
- Easier off-site replication
- “Merge” partition handling strategy (no earlier than 3.8)

Thank you

Thank you

- @michaelklishin
- github.com/michaelklishin
- rabbitmq-users
- Our team is hiring!