



Troubleshooting AEM Replication

Tommaso Teofili | Software Engineer

Bē

Jon Noorlander

Agenda

- Use cases
- Common pitfalls
- Performance
- Under the covers

Use cases

WCM publish

- move pages and assets from author to publish

Flushing dispatcher cache

- invalidate paths for published pages

Move / migrate content

- Clone instances before spawning them
- Instance upgrades

The screenshot shows the 'Replication Queue' page in the AEM admin interface. At the top, there is a summary section for the 'Default Agent (publish)'. It indicates that the agent is enabled and replicating to 'http://localhost:4503/bin/receive? sling:authRequestLogin=1'. The queue status is 'active - 66 pending'. Below this are links to 'View log' and 'Test Connection'. The main area is a table titled 'Replication Queue' with columns: Time, Type, User, Size, and Path. The table lists 15 replication tasks, all of which are 'Activate' type tasks performed by user 'admin' on 25-Jul-2017 at 18:41. The paths listed are various product and category URLs. At the bottom of the table, it says 'Default Agent: Processing... (for 1 second)'.

Time	Type	User	Size	Path
1 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/coats
2 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products
3 25-Jul-2017 18:41	Activate	admin	11 kb	/content/we-retail/ca/en/products/women/shirts/rios-t-shirt
4 25-Jul-2017 18:41	Activate	admin	10 kb	/content/we-retail/ca/en/products/women/coats/sleek-insulated-coat
5 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/shirts
6 25-Jul-2017 18:41	Activate	admin	15 kb	/content/we-retail/ca/en/products/women/coats/sonja-insulated-jacket
7 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/men
8 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/pants/hiking-pants
9 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/shorts
10 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/gloves
11 25-Jul-2017 18:41	Activate	admin	11 kb	/content/we-retail/ca/en/products/women/shorts/desert-sky-shorts
12 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/gloves/gloves
13 25-Jul-2017 18:41	Activate	admin	10 kb	/content/we-retail/ca/en/products/women/shorts/bahamas-shorts
14 25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/pants/faba-running-pants
15 25-Jul-2017 18:41	Activate	admin	11 kb	/content/we-retail/ca/en/products/women/shorts/candide-trail-short

Default Agent: Processing... (for 1 second)

Common pitfalls

Stuck queues

 **Default Agent (publish)**
Agent that replicates to the default publish instance.

Agent is **enabled**. Replicating to <http://localhost:4503/bin/receive?sling:authRequestLogin=1>

Queue is **active - 85 pending**

[View log](#)

[Test Connection](#)

[Settings](#) [Edit](#)

Replication Queue

Refresh Clear Force Retry Pause

	Time	Type	User	Size	Path
1	25-Jul-2017 18:39	Activate	admin	6 kb	/content/we-retail
2	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience/climbing-in-the-massif-du-mont-blanc
3	25-Jul-2017 18:39	Activate	admin	7 kb	/content/we-retail/language-masters/en/experience/arctic-surfing-in-lofoten
4	25-Jul-2017 18:39	Activate	admin	1 kb	/content/we-retail/language-masters
5	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience
6	25-Jul-2017 18:39	Activate	admin	7 kb	/content/we-retail/language-masters/en/experience/arctic-surfing-in-lofoten
7	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience
8	25-Jul-2017 18:39	Activate	admin	7 kb	/content/we-retail/language-masters/en/experience/arctic-surfing-in-lofoten
9	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience/running-at-the-great-wall-marathon
10	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience
11	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience/climbing-on-kalymnos-island--greece
12	25-Jul-2017 18:39	Activate	admin	14 kb	/content/we-retail/language-masters/en
13	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience/running-at-the-great-wall-marathon
14	25-Jul-2017 18:39	Activate	admin	7 kb	/content/we-retail/language-masters/en/experience/arctic-surfing-in-lofoten
15	25-Jul-2017 18:39	Activate	admin	2 kb	/content/we-retail/language-masters/en/experience/climbing-on-kalymnos-island--greece

Default Agent: **Blocked** - Next Retry in 10 seconds

Stuck queues - What ?

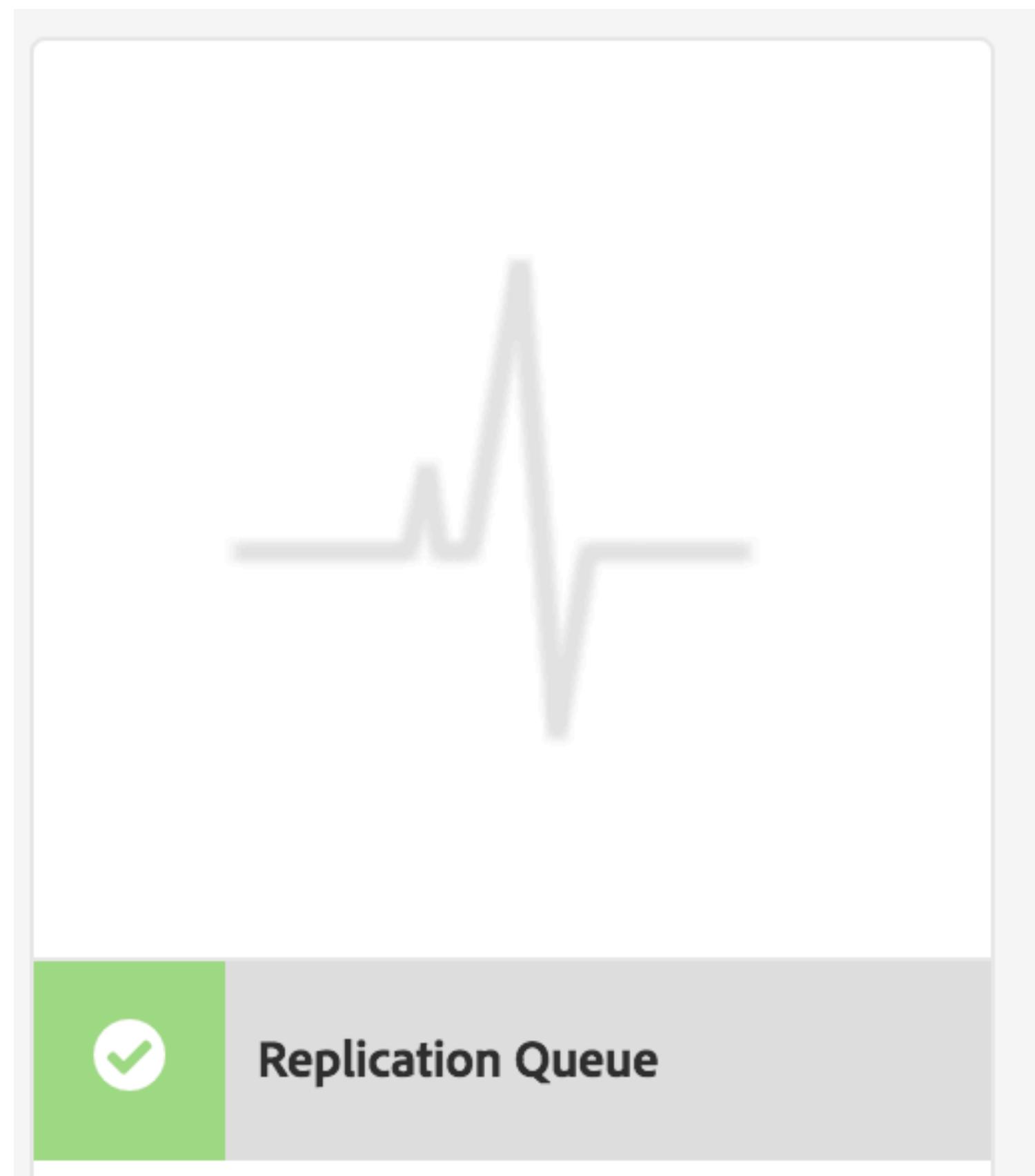
- Each replication agent has a single queue used to (asynchronously) deliver replication packages to a receiving (AEM) instance
- To preserve data integrity a replication queue is ordered (FIFO)
 - first item is processed
 - if ok -> remove it from the queue
 - otherwise -> retry
- If an item fails to be delivered or installed on the target instance, it will **NEVER** leave the queue
 - new items will start piling up and never get processed

Stuck queues - Why ?

- Network hiccups
- Node types management
- Oak conflicts (especially on Mongo)
- Wrong permissions
- On clustered authors, look for network partitioning problems !
- ...

Stuck queues - How ?

- Use the replication healthcheck



Adobe Experience Manager

< Configure Replication Queue

HINT Click here to inspect the replication agent configurations and queues.

Status: OK	Message
[DEBUG]	Agent is disabled [dynamic_media_replication]
[DEBUG]	Agent is disabled [flush]
[DEBUG]	No items in queue for agent [publish_reverse]
[DEBUG]	Agent is disabled [s7delivery]
[DEBUG]	Agent: [scene7], first item: [2017/7/25/16/34/6e40e6a5-8c1b-4948-a343-1f14df88088b_1279], number of retries: 0
[DEBUG]	Agent is disabled [static]
[DEBUG]	Agent: [test_and_target], first item: [2017/7/25/16/34/6e40e6a5-8c1b-4948-a343-1f14df88088b_1307], number of retries: 0
[DEBUG]	Agent is disabled [youtube]
[DEBUG]	Agent: [publish], first item: [2017/7/25/16/34/6e40e6a5-8c1b-4948-a343-1f14df88088b_1277], number of retries: 0

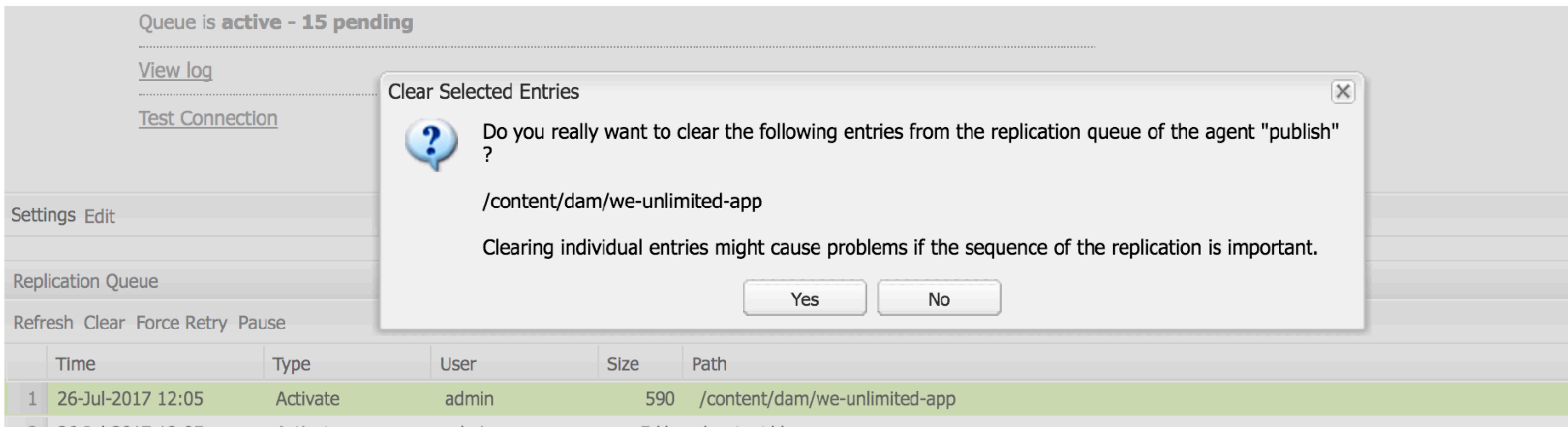
Stuck queues - How ?

- Look at (debug level) logs (sender AND receiver)
 - *com.day.cq.replication.ReplicationException: Repository error during node import: OakState0001: Unresolved conflicts in /content/dam/mac/bp6042/1*
- Look at replication log from the UI

```
: -----
: Replication (ACTIVATE) of /content/we-retail/ca/en/products/equipment/biking/compact-chain-tool successful.
: Sending POST request to http://localhost:4503/bin/receive?sling:authRequestLogin=1
: sent. Response: 200 OK
-----
: Sending message to localhost:4503
: >> POST /bin/receive HTTP/1.0
: >> Action: Activate
: >> Path: /content/we-retail/ca/en/products/equipment/biking/comfort-gel-gloves
: >> Handle: /content/we-retail/ca/en/products/equipment/biking/comfort-gel-gloves
: >> Referer: about:blank
: >> ...spooling 7682 bytes...
: --
: << HTTP/1.1 200 OK
: << Date: Tue, 25 Jul 2017 16:41:41 GMT
: << X-Content-Type-Options: nosniff
: << Content-Type: text/plain; charset=utf-8
: << Content-Length: 30
: <<
: << ReplicationAction ACTIVATE ok.
: Message sent.
-----
: Replication (ACTIVATE) of /content/we-retail/ca/en/products/equipment/biking/comfort-gel-gloves successful.
```

Stuck queues - How ?

- Test connection (per agent) to make sure endpoint is reachable
- ***FINALLY*** when the problem is fixed -> Clear the failing item(s) from the queue

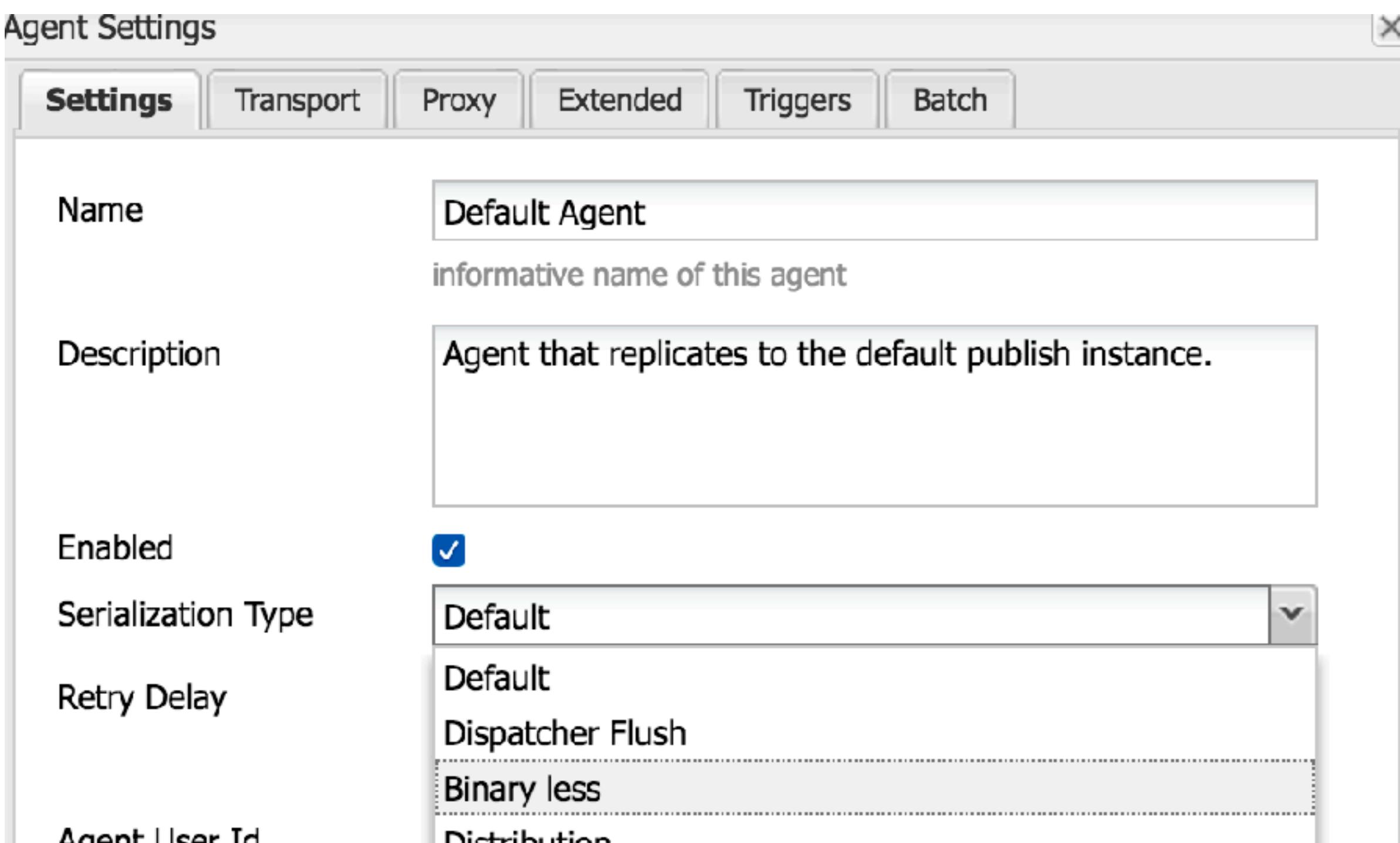
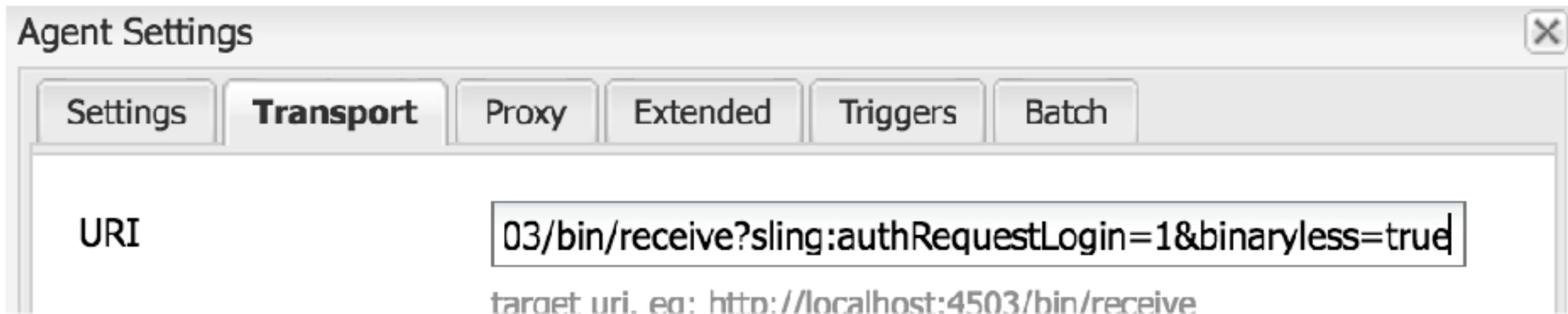


Is binaryless replication really binary less ?

- BL replication works by:
 - sending a package whose binaries are replaced with binary references (an HMAC String) so that
 - receiver will resolve the binary upon installation **by reference** from the repository
- This setup depends on setting up an Oak shared DataStore *correctly*
- When this was introduced (5.6.1) we decided to provide a fallback mechanism so that if BL replication doesn't work (e.g. wrong shared DS setup) it would fall back to normal replication (sending binaries)

Is binaryless replication really binary less ?

- Assuming shared DS is correctly configured
- Set agent serialization to “Binary less”
- Add `binaryless=true` to transport URI

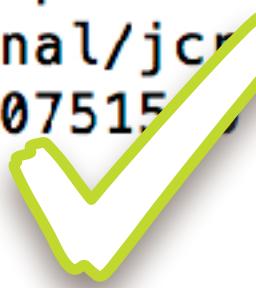


Is binaryless replication really binary less ?

```
25.07.2017 16:45:52 - INFO - publish : Sending message to localhost:4503
25.07.2017 16:45:52 - INFO - publish : >> POST /bin/receive HTTP/1.0
25.07.2017 16:45:52 - INFO - publish : >> Action: Activate [REDACTED]
25.07.2017 16:45:52 - INFO - publish : >> Path: /content/dam/we-retail/en/activities/running/marathon-shoes.jpg
25.07.2017 16:45:52 - INFO - publish : >> Handle: /content/dam/we-retail/en/activities/running/marathon-shoes.jpg
25.07.2017 16:45:52 - INFO - publish : >> Referer: about:blank [REDACTED]
25.07.2017 16:45:52 - INFO - publish : >> ...spooling 10674 bytes...
25.07.2017 16:45:52 - INFO - publish : --
25.07.2017 16:45:52 - INFO - publish : << HTTP/1.1 200 OK
25.07.2017 16:45:52 - INFO - publish : << Date: Tue, 25 Jul 2017 14:45:52 GMT
25.07.2017 16:45:52 - INFO - publish : << X-Content-Type-Options: nosniff [REDACTED]
25.07.2017 16:45:52 - INFO - publish : << Content-Type: text/plain; charset=utf-8
25.07.2017 16:45:52 - INFO - publish : << Content-Length: 591
25.07.2017 16:45:52 - INFO - publish : <<
25.07.2017 16:45:52 - INFO - publish : << FAILED PATHS START
25.07.2017 16:45:52 - INFO - publish : << /content/dam/we-retail/en/activities/running/marathon-shoes.jpg/jcr:content/renditions/cq5d
25.07.2017 16:45:52 - INFO - publish : << /content/dam/we-retail/en/activities/running/marathon-shoes.jpg/jcr:content/renditions/cq5d
25.07.2017 16:45:52 - INFO - publish : << /content/dam/we-retail/en/activities/running/marathon-shoes.jpg/jcr:content/renditions/orig
25.07.2017 16:45:52 - INFO - publish : << /content/dam/we-retail/en/activities/running/marathon-shoes.jpg/jcr:content/renditions/cq5d
25.07.2017 16:45:52 - INFO - publish : << FAILED PATHS END
25.07.2017 16:45:52 - INFO - publish : Message sent.
```

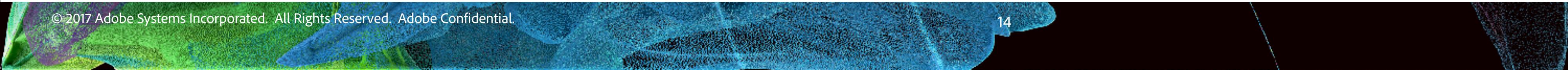


```
*DEBUG* [127.0.0.1 [1465577645518] POST /bin/receive HTTP/1.1] com.day.cq.replication.impl.content.durbo.DurboImportTransformer property jcr: data of Node[NodeDelegate{tree=/content/dam/adobetest2/iPhone_6s_Plus-4K-video.MOV/jcr:content/renditions/original/jcr:content: { jcr:primary Type = nt:resource, jcr:uuid = ae912ae7-c808-4798-b723-31af557b1ab3, jcr:lastModifiedBy = admin, jcr:data = {8407515 / bytes}, jcr:lastModified = 2016-06-10T12:54:06.667-04:00}}] set using a reference
```





Replication is **slow!!!!**

.... and NOW !??!?!


ARE YOU SURE ?

- Where (sender, receiver)
- What (package build, queueing, transport, ...)
- Why (system load, concurrency, bugs, etc.)



Sender - what is problematic performance wise ?

- 25.07.2017 19:06:57.815 *INFO* [0:0:0:0:0:0:0:1 [1501002386929] POST /libs/replication/treeactivation.html HTTP/1.1] com.day.cq.replication.impl.ReplicatorImpl Processed replication: **setup** 1ms, **checked** 5ms, **pre** 87ms, **build** 468ms, **queued** 513ms, **status** 178ms, **total** 1252ms
- get the error.log file
- cat error.log | grep Processed > repout.txt
- awk '{print \$14,\$16,\$18,\$20,\$22,\$24,\$26}' repout.txt
- put them in a CSV
- graph them !

Sender - Replication phases

- setup
 - collect agents, setup options
- checked
 - check replication permissions
- pre
 - call preprocessors (e.g. alias processor)
- build
 - build the replication packages
- queued
 - put the packages in the queues
- status
 - update status of replicated pages
- total
 - total time of sender side replication

 **Default Agent (publish)**
Agent that replicates to the default publish instance.

Agent is enabled. Replicating to <http://localhost:4503/bin/receive? sling:authRequestLogin=1>

Queue is active - 66 pending

[View log](#)

[Test Connection](#)

[Settings](#) [Edit](#)

Replication Queue

[Refresh](#) [Clear](#) [Force](#) [Retry](#) [Pause](#)

	Time	Type	User	Size	Path
1	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/coats
2	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products
3	25-Jul-2017 18:41	Activate	admin	11 kb	/content/we-retail/ca/en/products/women/shirts/rios-t-shirt
4	25-Jul-2017 18:41	Activate	admin	10 kb	/content/we-retail/ca/en/products/women/coats/sleek-insulated-coat
5	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/shirts
6	25-Jul-2017 18:41	Activate	admin	15 kb	/content/we-retail/ca/en/products/women/coats/sonja-insulated-jacket
7	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/men
8	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/pants/hiking-pants
9	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/shorts
10	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/gloves
11	25-Jul-2017 18:41	Activate	admin	11 kb	/content/we-retail/ca/en/products/women/shorts/desert-sky-shorts
12	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/gloves/gloves
13	25-Jul-2017 18:41	Activate	admin	10 kb	/content/we-retail/ca/en/products/women/shorts/bahamas-shorts
14	25-Jul-2017 18:41	Activate	admin	4 kb	/content/we-retail/ca/en/products/women/pants/faba-running-pants
15	25-Jul-2017 18:41	Activate	admin	11 kb	/content/we-retail/ca/en/products/women/shorts/candide-trail-short

Default Agent: Processing... (for 1 second)

Replication timing by phases (over time)

setup, checked, pre, build, queued, status, total

0, 2, 73, 163, 334, 70, 642

0, 3, 56, 414, 786, 1633, 2892

1, 3, 69, 404, 1358, 2592, 4427

0, 3, 57, 382, 3636, 900, 4978

0, 2, 78, 515, 4397, 2016, 7008

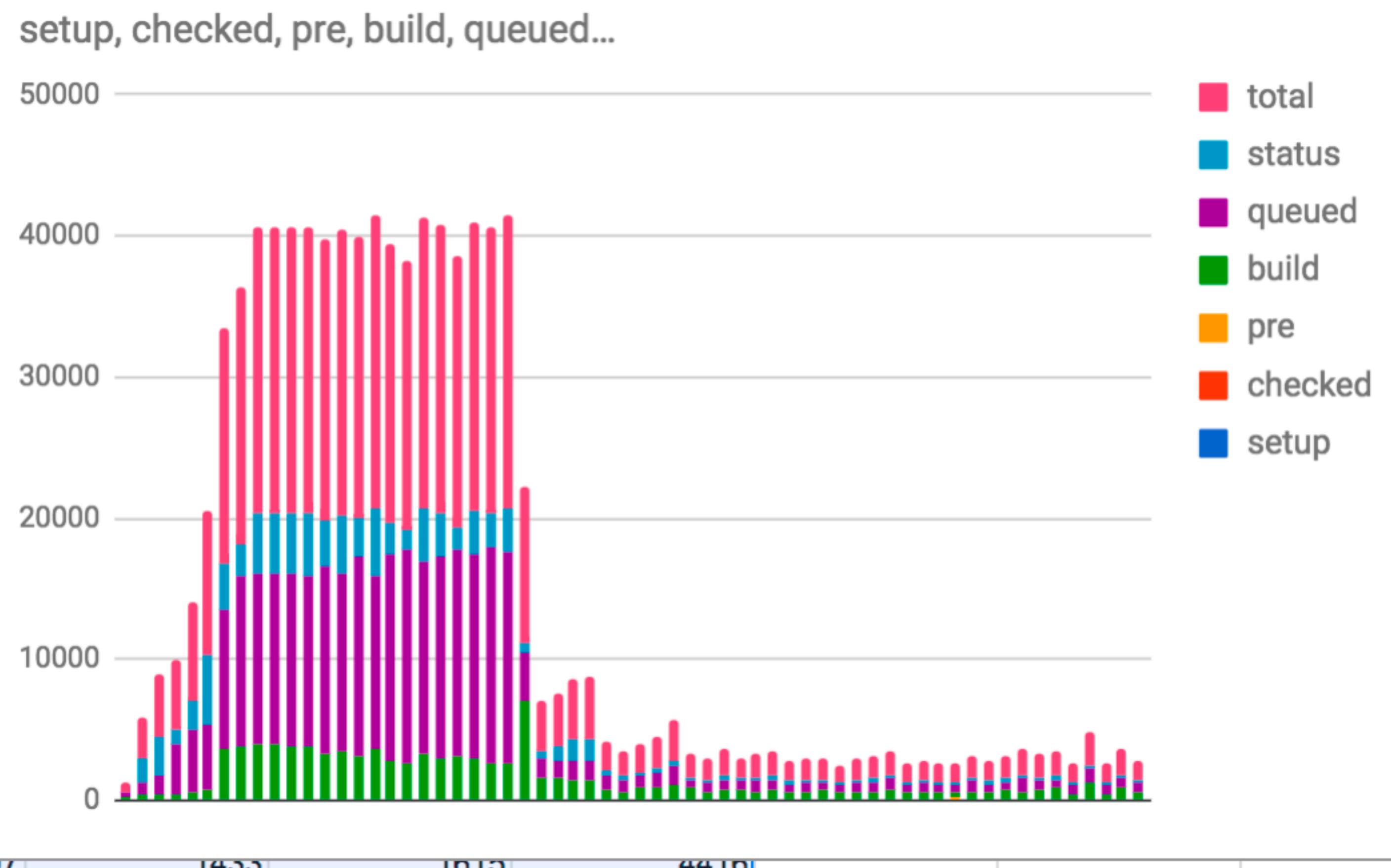
0, 2, 54, 697, 4641, 4854, 10248

2, 3, 63, 3648, 9863, 3158, 16737

0, 3, 57, 3803, 12021, 2294, 18178

0, 3, 61, 3977, 12089, 4137, 20267

0, 5, 49, 3927, 12167, 4174, 20322

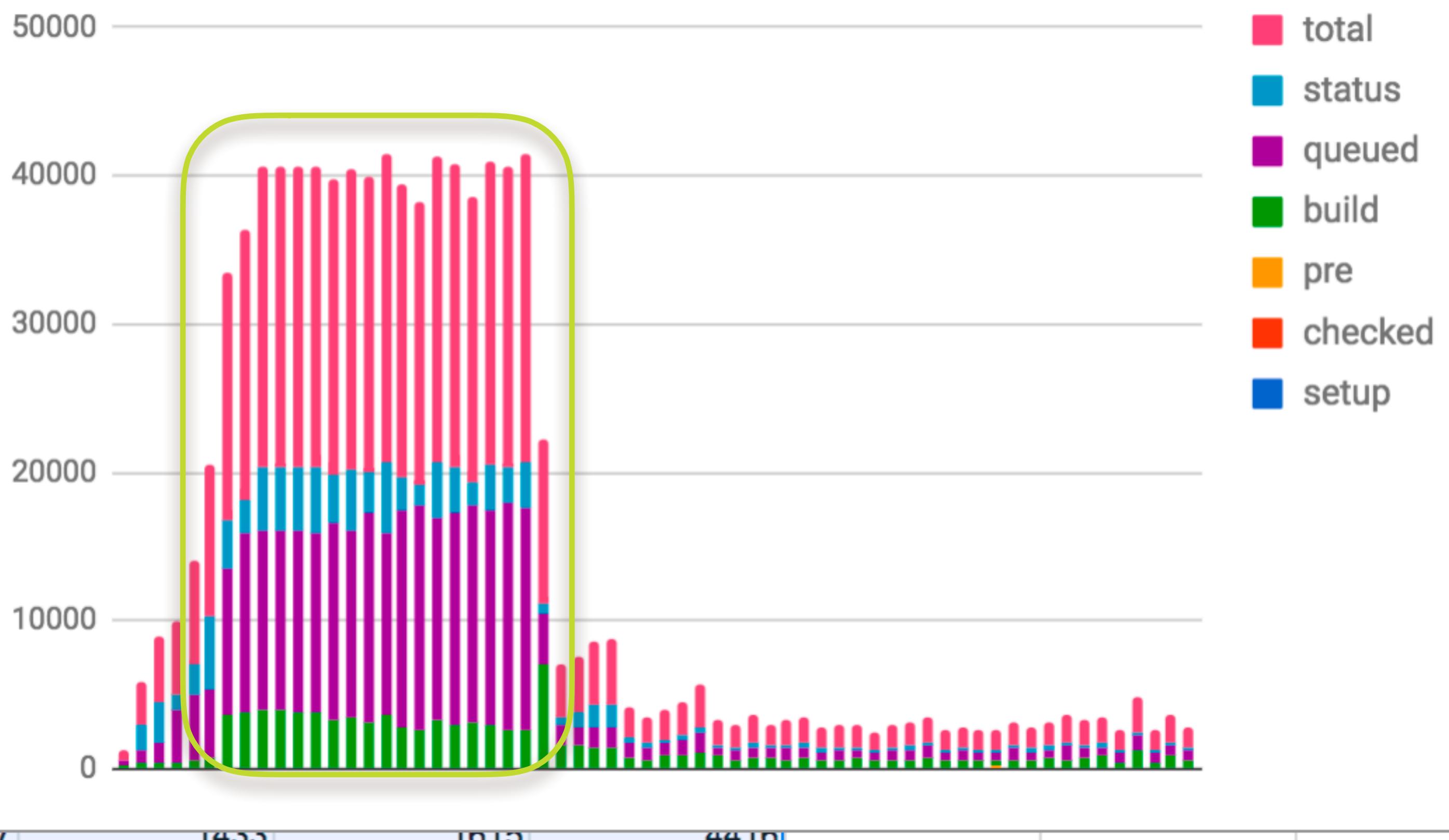


Looking at the spikes

Spikes can depend on

- concurrency
- system load
- repository slowness
- content size
- ...

setup, checked, pre, build, queued...



Timings on the receiver

- 25.07.2017 19:07:25.874 *INFO* [127.0.0.1 [1501002445869] POST /bin/receive HTTP/1.1]
com.day.cq.replication.impl.servlets.ReplicationServlet **Processed replication action in 4ms:**
ACTIVATE of /content/dam/templates/printads
- get the error.log file
- cat receiver.txt | grep Processed
- awk '{print \$14}' receiver.txt | sed s/ms:// > receiver.csv
- graph them!

Replication import time (over time)

import time

37

32

11

10

13

7

9

37

29

18

processed (ms)

250

200

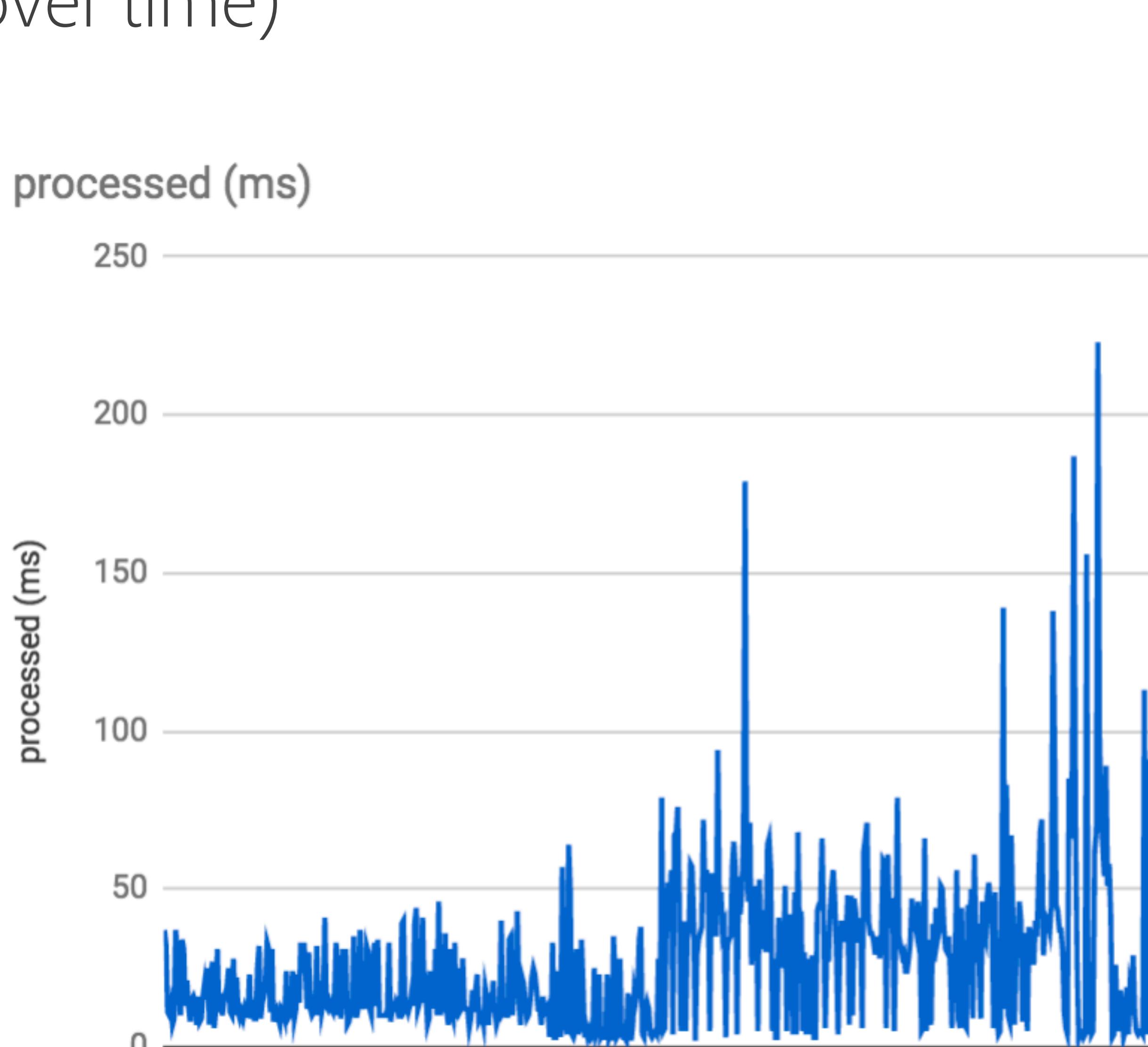
150

100

50

0

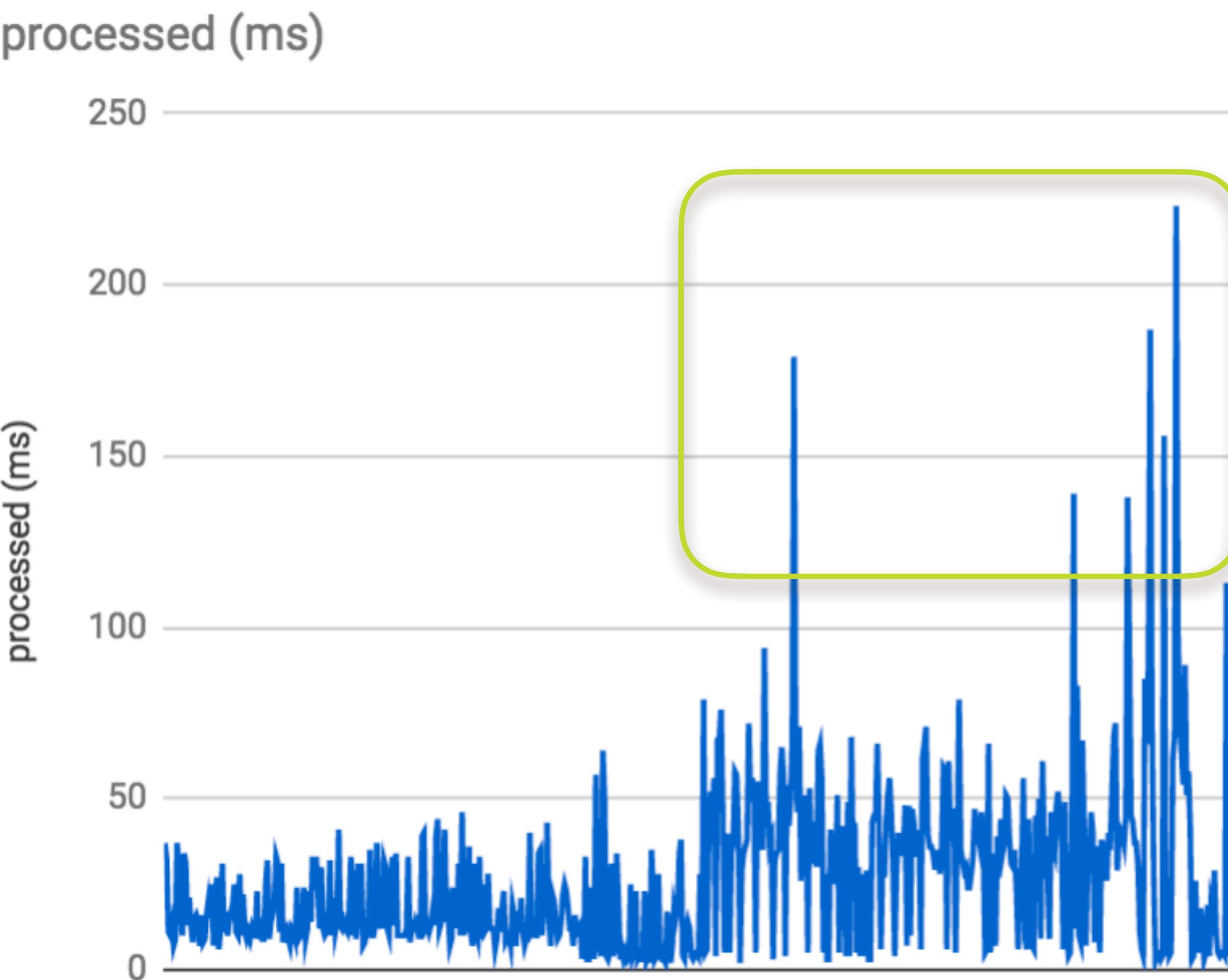
processed (ms)



Looking at the spikes

Spikes can depend on

- concurrency
- system load
- repository slowness
- content size
- ...



Import time histogram

import time

37

32

11

10

13

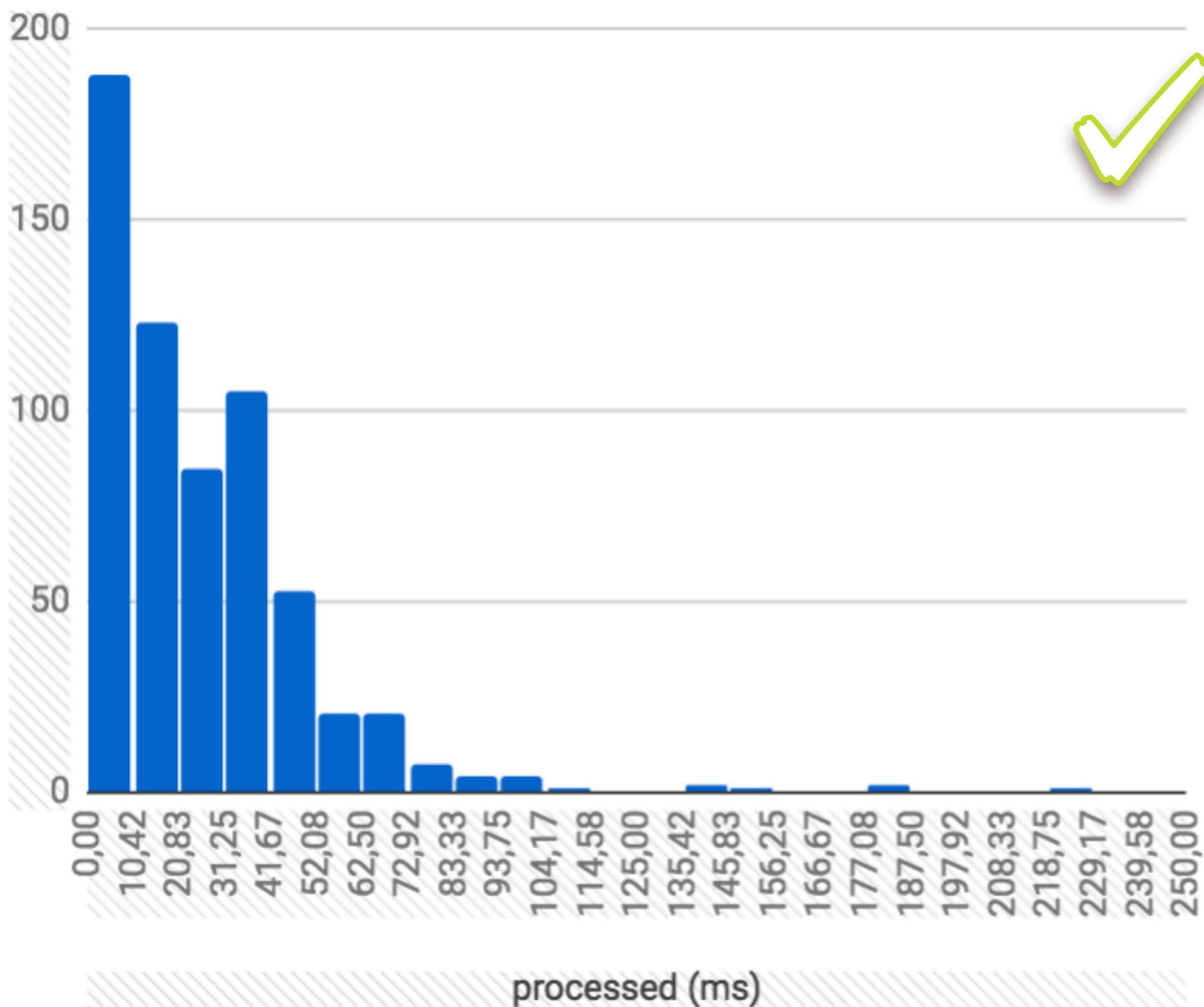
7

9

37

29

18



JMX

Look at agent JMX Mbeans

- **over time!**
- this can tell you :
 - avg no. of items
 - avg. processing time
 - blocked time percentage
 - etc.

com.adobe.granite.replication: "publish" (agent)

Provides information about a replication agent and its queue.

Attributes

Attribute Name	Attribute Value
Id	publish
Valid	true
Enabled	true
QueuePaused	false
QueueNumEntries	9
QueueBlocked	false
QueueStatusTime	2017-07-25T16:39:23+0200
QueueNextRetryTime	
QueueProcessingSince	2017-07-25T16:39:23+0200
QueueLastProcessTime	2017-07-25T16:39:23+0200

Operations

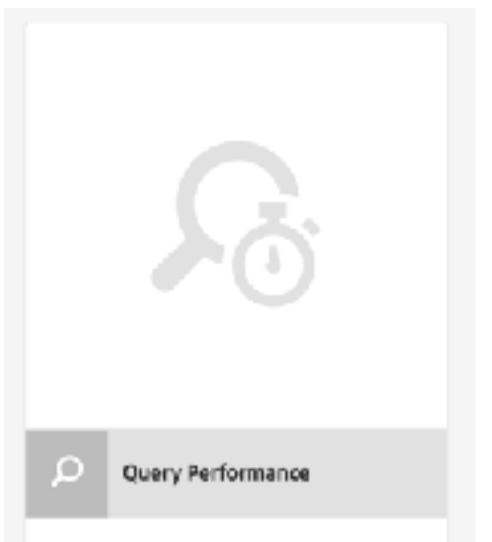
Return Type	Name	
void	queueClear()	Clears the queue and removes all queue entries
void	queueForceRetry()	Forces a retry attempt on a blocked entry. this has only an effect, if the queue is blocked.

Queues uncovered

- Replication queues are implemented on top of Sling Jobs
 - by default queue states are persisted in the repository
 - jobs history can be enabled (all jobs remain archived in the repository, disabled by default)
- Each replication queue is associated with a Sling Jobs queue / topic
- Retrieving the queue items results in:
 - finding all the jobs with a certain topic
 - issuing a (JCR) query
 - **this is costly !**
- With 20 agents and 50ms per query, it means 1s just to retrieve the entries from the repo
- Not counting returning all those elements (e.g. 1000 entries per queue)

JCR queries for replication queues

- Look at the query performance tool



Diagnosis Tools

[Slow Queries](#) [Popular Queries](#) [Explain Query](#)

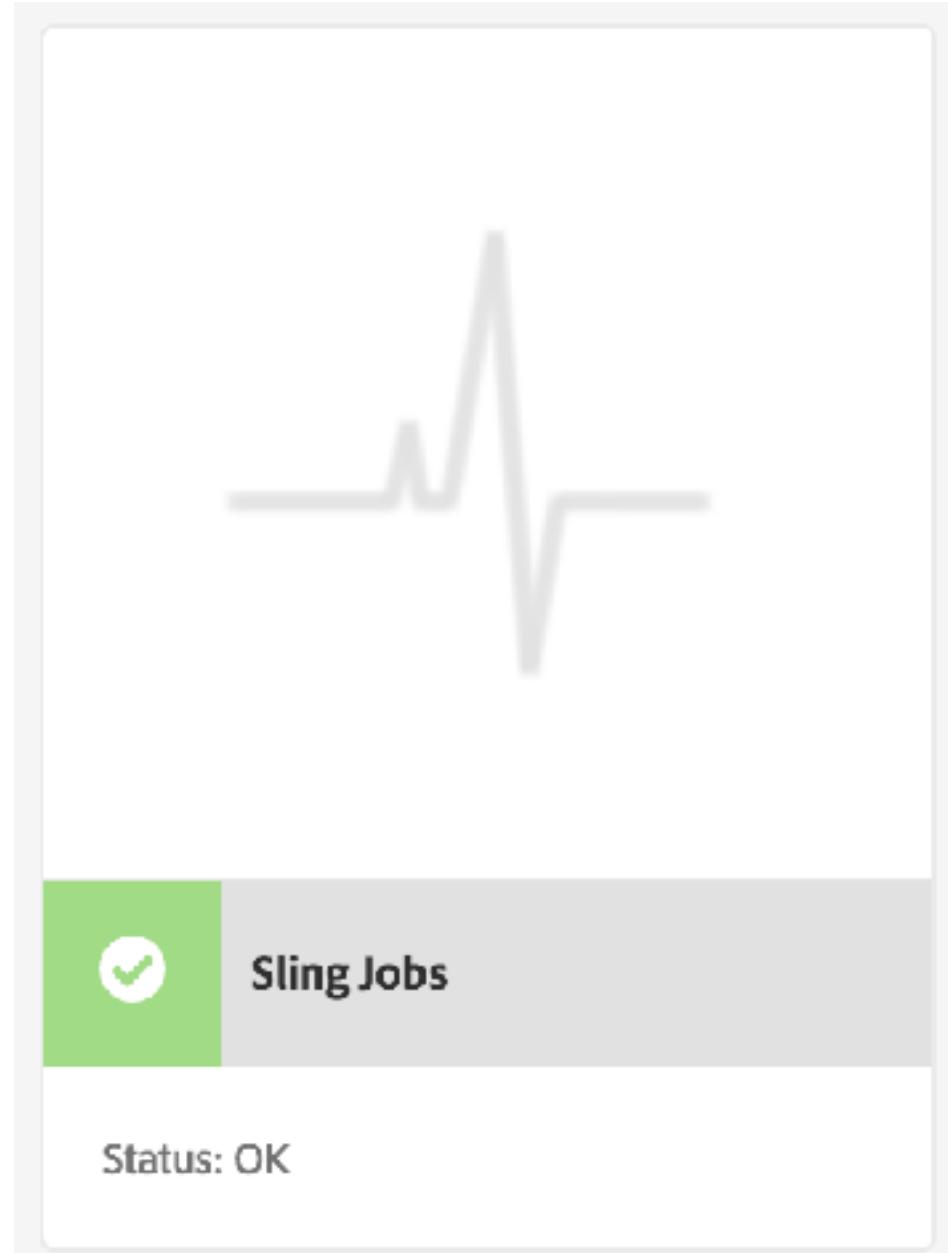
<input type="checkbox"/>	Creation Date	Occurrences / Language	Statement	Duration
<input type="checkbox"/>	Wed, Jul 26 2017 05:12:41 CEST	13 XPATH	/jcr:root/home/users//element(*,rep:Token)	<div style="width: 15667 ms"></div> 15667 ms
<input type="checkbox"/>	Wed, Jul 26 2017 05:12:41 CEST	16 SQL	SELECT * FROM nt:base WHERE sling:resourceType = 'cq/reporting/components/reportpage' AND jcr:path like '/etc/reports/%'	<div style="width: 15654 ms"></div> 15654 ms
<input type="checkbox"/>	Wed, Jul 26 2017 05:12:41 CEST	1 XPATH	/jcr:root/var/eventing/jobs//element(*,slingevent:Job)[@slingevent:eventId = '2017/7/26/2/12/6e40e6a5-8c1b-4948-a343-1f14df88088b_9254']	<div style="width: 15639 ms"></div> 15639 ms
<input type="checkbox"/>	Tue, Jul 25 2017 12:36:50 CEST	1 SQL	SELECT sling:vanityPath, sling:redirect, sling:redirectStatus FROM nt:base WHERE sling:vanityPath IS NOT NULL	<div style="width: 6064 ms"></div> 6064 ms

JCR queries for replication queues

- What can we do ?
 - find out why they're slow!
 - look at the system status and load, it could be a transient state
 - look at the replication timings
 - transport is slow
 - receiver is slow
 - in that case most of the items stay for a long time in the queue waiting the processed item to be delivered and installed
 - report !

Sling Jobs

- Look at jobs healthcheck



A screenshot of the Adobe Experience Manager Sling Jobs configuration page. The header features the AEM logo and the title "Sling Jobs". On the left, there's a "Configure" button with a gear icon. The main content area has two columns: "Status: OK" and "Message". The "Status" column contains a green "INFO" button. The "Message" column displays the text "[INFO] Found 29 jobs queued.".

Sling Jobs

- Look at jobs queue JMX Mbean
 - over time!
 - avg no. of jobs per state
 - processed
 - active
 - queued
 - failed
 - ...

org.apache.sling: com_day_cq_replication_job_publish (queues)

Information on the management interface of the MBean

Attributes

Attribute Name	Attribute Value
Name	com_day_cq_replication_job_publish
StartDate	2017-07-25T16:33:17+0200
LastActivatedJobDate	2017-07-25T16:40:15+0200
LastFinishedJobDate	2017-07-25T16:40:15+0200
StartTime	1500993197237
NumberOfJobs	0
NumberOfProcessedJobs	866
NumberOfActiveJobs	0
NumberOfQueuedJobs	0
AverageWaitingTime	1193
AverageProcessingTime	153
NumberOfFinishedJobs	866
NumberOfCancelledJobs	0
NumberOfFailedJobs	0
LastActivatedJobTime	1500993615479
LastFinishedJobTime	1500993615709

Operations

Return Type	Name
void	reset() Operation exposed for management

Q & A



Adobe

MAKE IT AN EXPERIENCE