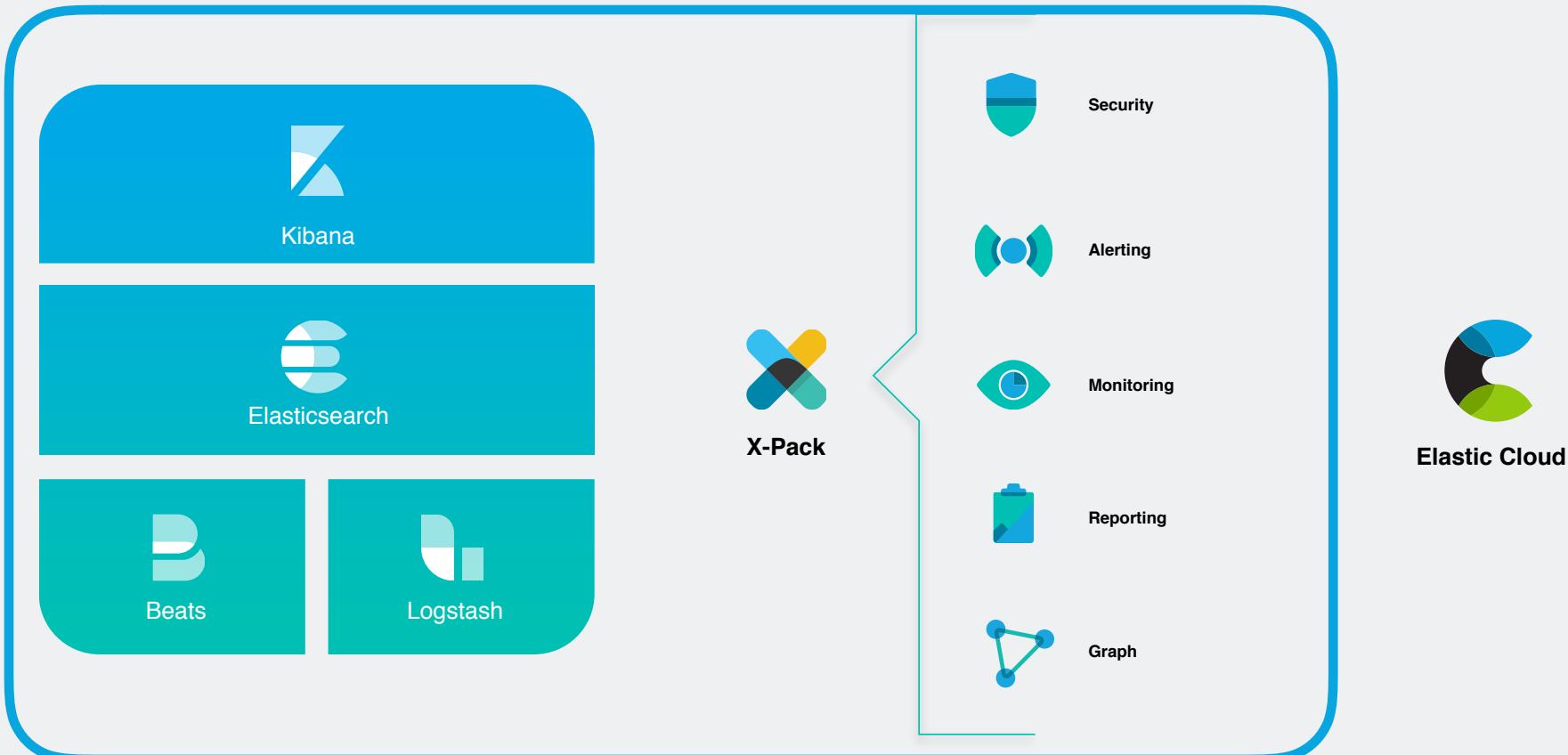




# From 50 to 330+: Scaling a distributed engineering

Patroklos Papapetrou- @softwaregarden  
Java meetup Thessaloniki, 2017-11-07

# Who are we?



# Who are we?

600

32

270

100+ million

# Nextfunders: 2012



Shay



Uri



Ron



Simon



Steven



# Distributed By Design

“

**A distributed system is a software system in which components located on network computers communicate and**

Elasticsearch  
the software,

**What does a  
distributed software  
and distributed,  
open source**

# Lots



# Distributed systems: communicati on

# Asynchronous Communication

# Email: Filters included

More ~~than~~ reached a



Mail ▾

COMPOSE

Inbox (1,412)

Sent Mail

Drafts (133)

Automated/[CI]

builds (2)

Dev (269) ■

Dev Support (99) ■

follow up ■

GH Mentions (186) ■

Github (28,175) ■

hiring (193) ■

# Email transparency

ELT Week Feb 2017 - Support Session Inbox x

 **Shay Banon**  
to all (+)

We held a session about Support, lead by Marty. A session that was the result of multiple discussions, happening what we discussed, we need to first understand what lead to this session in the first place, as it is not really what road that was taken to make said decision that matters more.

Beats update 2017-05-02 Inbox x Dev x

 **Monica Barbu**  
to dev (+)

New community Beat: **Rbeat**  
**Rbeat** is built on top of libbeat infrastructure to collect the Redis slow logs and ship them to Elasticsearch accumulated by Redis in memory, so no files are written with information about the slow command word.

What does voting mean? Inbox x

 **Steven Willmott**  
to Clusticaurus (+)

During one of the votes this question came up:  
~ Sorry if I missed something but can we clarify how and when a voting threshold needs to be reached? Something like 72h, no -1 and at least E +10? I can start a new voting thread for this question, though :)

my proposal links apply the ASF rules [1] - please read them and express your opinion

[1] <https://www.apache.org/foundation/voting.html>

# Some email become history

Tijj Ji je   Inbox x   Dev x

---

 **Boaz Leskes** <b.leskes@gmail.com>  
to Twitter, dev ▾  
UzzrzOu))6,)6.32!&)B&&&

---

 **Boaz Leskes**  
On Wed, Jan 14, 2015 at 3:00 PM, Boaz Leskes <b.leskes@gmail.com> wrote: > > ...

---

 **Boaz Leskes** <boaz.leskes@elasticsearch.com>  
to suyog, dev ▾  
Apologies. Apparently my phone the decided to defeat pocket boredom by sending mails.  
...

# Reply-all is fun!



Chris Cowan <chris.cowan@elasticsearch.com>

1/14/15

to Boaz, dev ↗

I was worried that you were injured in a basement and that was the only thing you could manage to type. Glad your okay :)

—  
Sent from [Mailbox](#)



Tanguy Leroux <tanguy.leroux@elasticsearch.com>

1/15/15

to Chris, Boaz, dev ↗

It looks like someone who tries to exit Vim editor for the first time



Michael McCandless <mike@elasticsearch.com>

1/15/15

to Tanguy, Chris, Boaz, dev ↗

I think Tij Ji je should be a rallying/battle cry we scream whenever we all sit down at our computers to tackle a particularly tricky test failure.

Mike McCandless

<http://blog.mikemccandless.com>

# Emails become t-shirts



# Love is in the air

all > I ❤ you all  
106 posts by 86 authors ▾

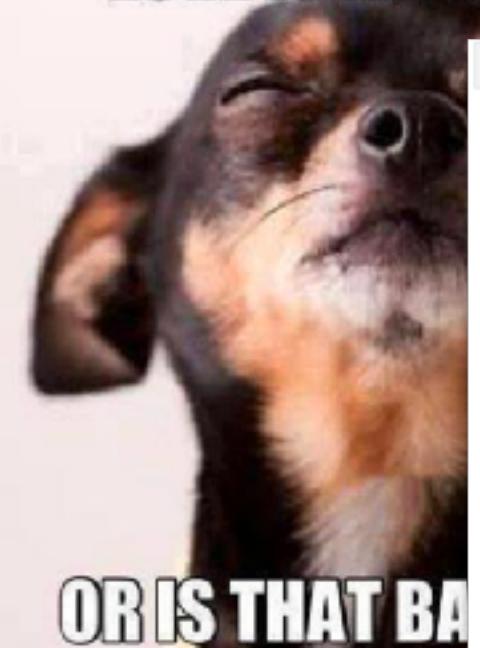
shay  
that's it, just wanted you to know.

Click here to Reply

aboyd ▾  
luisaantonio [image: Inline image]  
Jeremy.Falling On 1 August 2016  
bohyun The heart looks so pretty.  
chris.wurts ▾ 4  
sam.sharp The feeling is mutual.  
james.bowers Initiating Group Head  
gabriel.moskovicz [image: Inline image 1]  
scott\_reinke [image: Inline image 1]

Sean Gallagher <[sean@elastic.co](mailto:sean@elastic.co)>  
A Medd, Jimmie, Tyler, Jeremy, Marty, A.J., Thiago, Thom, all ▾

# LOVE IS IN THE AIR,



# OR IS THAT BA

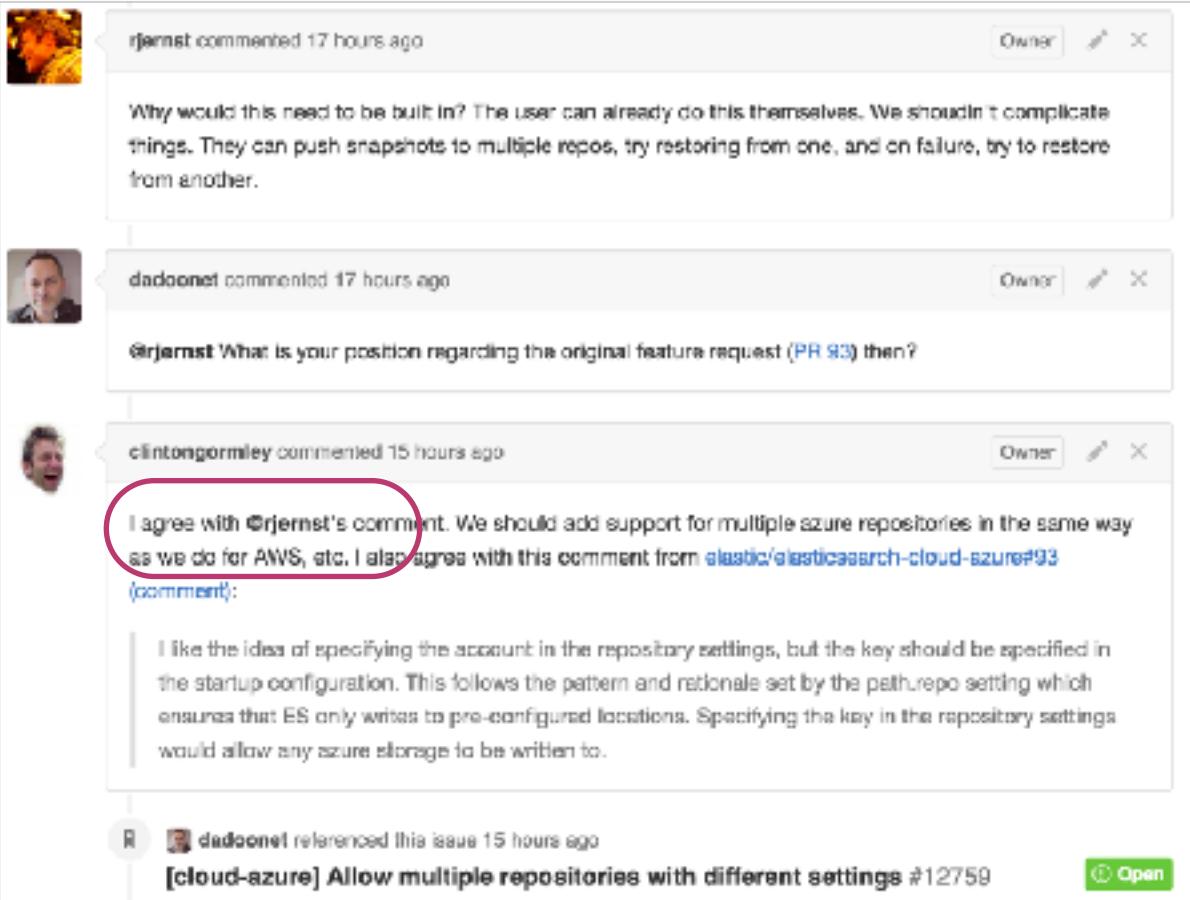
George Kobar <[kobar@elastic.co](mailto:kobar@elastic.co)>  
A Thiago, Marty, all, cpbeat\_employee ▾

So glad this thread is still a thing.



# GitHub all the things

# @mention someone directly on GH issue instead of discussing in Slack



rjernst commented 17 hours ago

Why would this need to be built in? The user can already do this themselves. We shouldn't complicate things. They can push snapshots to multiple repos, try restoring from one, and on failure, try to restore from another.

dadoonet commented 17 hours ago

@rjernst What is your position regarding the original feature request ([PR 93](#)) then?

clintongormley commented 15 hours ago

I agree with @rjernst's comment. We should add support for multiple azure repositories in the same way as we do for AWS, etc. I also agree with this comment from [elastic/elasticsearch-cloud-azure#93 \(comment\)](#):

I like the idea of specifying the account in the repository settings, but the key should be specified in the startup configuration. This follows the pattern and rationale set by the path.repo setting which ensures that ES only writes to pre-configured locations. Specifying the key in the repository settings would allow any azure storage to be written to.

R  dadoonet referenced this issue 15 hours ago  
[cloud-azure] Allow multiple repositories with different settings #12759 

# GitHub for non-tech stuff

Screenshot of a GitHub repository interface showing issues related to non-technical events.

Repository navigation bar:

- Code
- Issues (86) (highlighted)
- Pull requests (0)
- Wiki
- Settings
- Insights

Issue search bar:

- Filters: is:issue is:open
- Labels: Labels, Milestones
- New issue button

Issue list:

checkbox	①	88 Open	✓ 517 Closed	Author	Labels	Projects	Milestones	Assignees	Sort
<input type="checkbox"/>	①	2018-03 Scatter - Austin, TX, US	2018-03	country:us	focus:none	lang:en			
		e611 opened 7 hours ago by cocoonet ↗ CFPs closing in ..							
<input type="checkbox"/>	①	2017-11 CA World Conference — Las Vegas, NV	2017-11	country:us	invited	lang:en	paid:pledging		
		e810 opened 20 hours ago by eemrich ↗ CFPs closing in ..							
<input type="checkbox"/>	①	2017-11 GDG - Vienna, AT	2017-11	country:at	focus:none	lang:en			1
		e809 opened 6 days ago by xeras ↗ CFPs closing in ..							
<input type="checkbox"/>	①	2017-11 DevOpsDays Warsaw - Warsaw, PL	2017-11	country:pl	focus:devops	lang:en			2
		e808 opened 7 days ago by miiiche ↗ CFPs closing in ..							
<input type="checkbox"/>	①	2018-01 jHawaii - Laie, Hawaii, US	2018-01	country:us	discuss	lang:en			1
		e807 opened 8 days ago by cocoonet							
<input type="checkbox"/>	①	2018-04 ML4ALL - Portland, OR, US	2018-04	country:us	focus:opensource	lang:en	paid:pledging		
		e806 opened 8 days ago by miiiche ↗ CFPs closing in ..							
<input type="checkbox"/>	①	2018-06 Monitorama - Portland, OR, US	2018-06	country:us	focus:devops	lang:en	sponsoring		1
		e805 opened 18 days ago by miiiche ↗ CFPs closing in ..							

# Synchronous Communication



No, no...  
This isn't gossip.  
It's the truth.

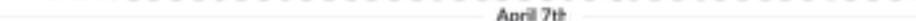


Join any room  
Technical discussions in  
public  
Public communication will  
educate everyone  
Post a summary to email or



## #craft-beer

44 members



April 7th



10:00 AM Robert Greene <http://www.nationaldaycalendar.com/days-2/national-beer-day-april-7/>



@baz: drink up!

10:00 AM Michael Basnight oh hell yes



2:26 PM Marcelo Rodriguez /giphy homer beer

[homer beer](#) (250KB) ▾

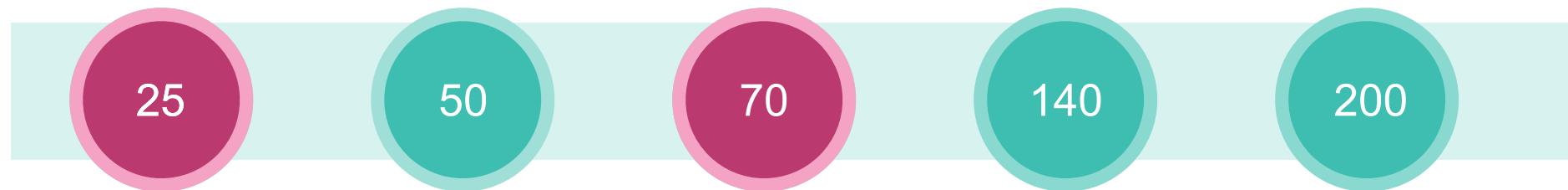


3:46 PM Michael Basnight /giphy facepalm

[facepalm](#) (809KB) ▾



# Meetings



Dev Standups  
Collapses

Demo Day  
Collapses

Table  
Groups

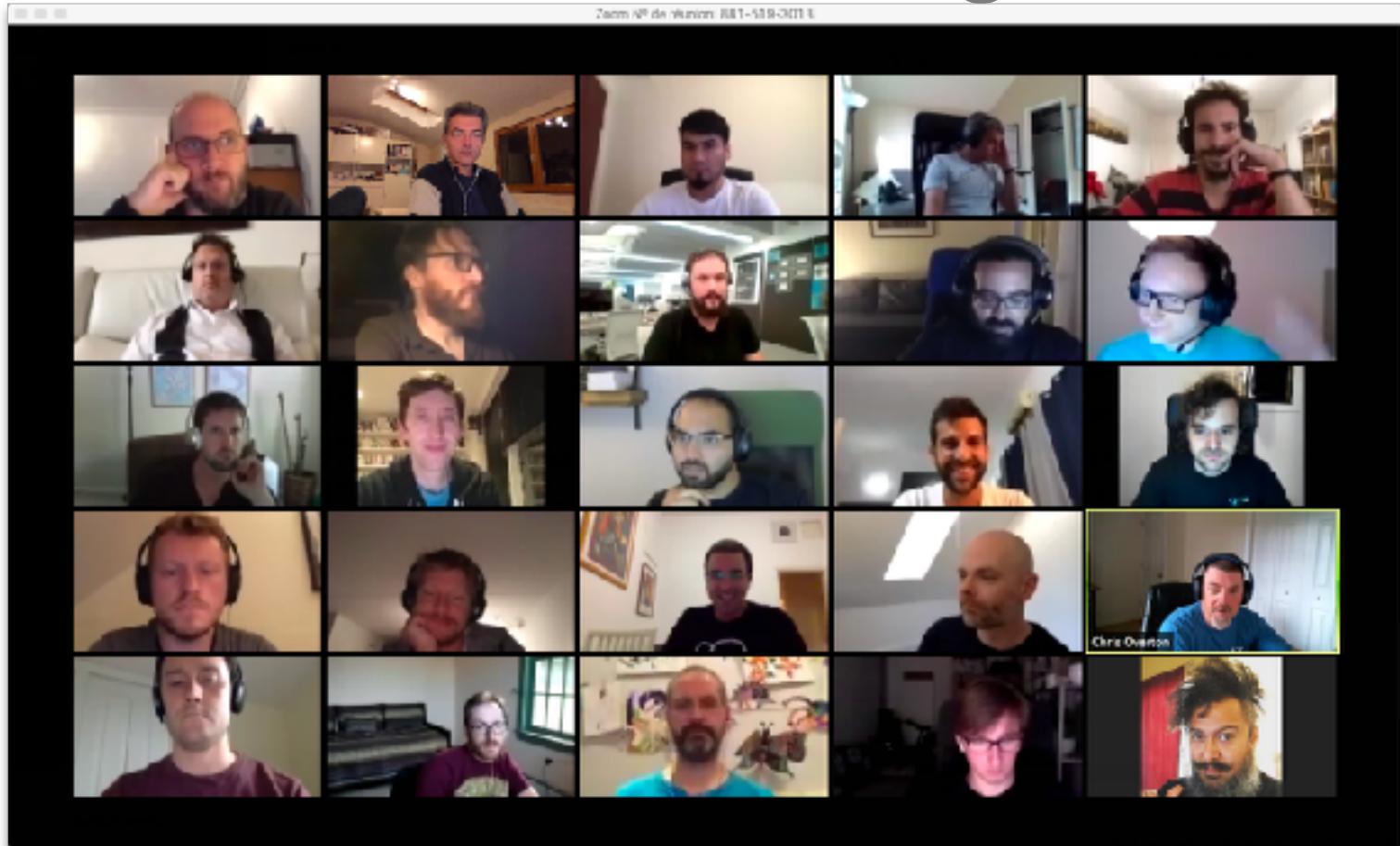
Black Belt  
Session

Cross  
Product/  
Per  
Product  
Sync

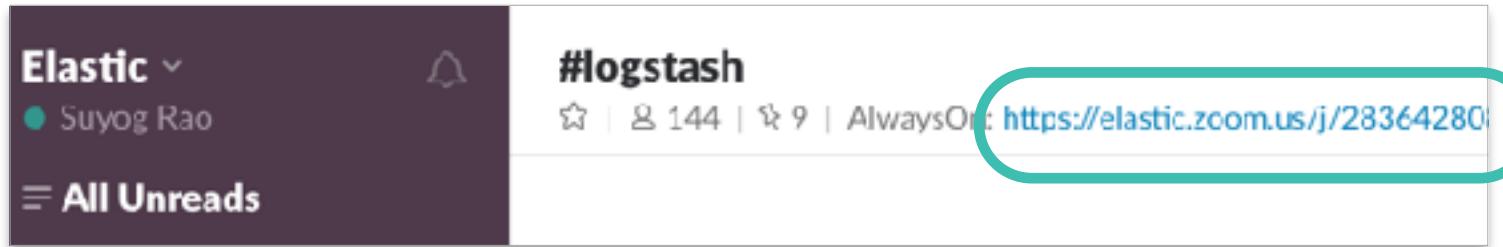
# Meetings are open: dev cal

Mon 5/8	Tue 5/9	Wed 5/10	Thu 5/11	Fri 5/12
7:30 - Clients & Integrations <a href="https://elastic.zoom.us/j/933">https://elastic.zoom.us/j/933</a>	4:30 - Bigissue ML team <a href="https://elastic.zoom.us/j/930">https://elastic.zoom.us/j/930</a>	7:30 - 8:30 ML team <a href="https://elastic.zoom.us/j/930">https://elastic.zoom.us/j/930</a>	7:30 - Dev - Lucas 8 - Core Infrastructure 8:15 - QA Daily Standup (da)	7:30 - Beats team sync <a href="https://elastic.zoom.us/j/932">https://elastic.zoom.us/j/932</a>
8 - Search & performance 8:15 - QA Daily Standup (da)	8:30 - Elastic AI X-Plat 9 - Kibana QA weekly	9 - 10 Logstash <a href="https://elastic.zoom.us/j/930">https://elastic.zoom.us/j/930</a>	8 - watcher v2 8:30 - 9:30 Demos 9:30 - 10:30 Ingest cross team <a href="https://elastic.zoom.us/j/930">https://elastic.zoom.us/j/930</a>	7 - 8 Cloud Core <a href="https://elastic.zoom.us/j/932891">https://elastic.zoom.us/j/932891</a>
10 - 11 LS Persistent Queue <a href="https://elastic.zoom.us/j/933">https://elastic.zoom.us/j/933</a>	10 - Sharing Team Sync	9 - 10 Support <a href="https://elastic.zoom.us/j/930">https://elastic.zoom.us/j/930</a>	9:30 - 10:30 Kibana All Ingest cross team <a href="https://elastic.zoom.us/j/930">https://elastic.zoom.us/j/930</a>	8 - 10 Weekly cross-beam design <a href="https://elastic.zoom.us/j/93018">https://elastic.zoom.us/j/93018</a>
			10:30 - Kibana Management	
12p - Clients	12p - Monitor	12p - Kibana Platform week		
12:30p - Kibana Visualization				
1:30p - 2:30p Beats stand-up meeting <a href="https://elastic.zoom.us/j/250289">https://elastic.zoom.us/j/250289</a>	1:30p - 2:30p Kibana Design <a href="https://elastic.zoom.us/j/930">https://elastic.zoom.us/j/930</a>	1:30p - 2:15p SRE Team <a href="https://elastic.zoom.us/j/881519">https://elastic.zoom.us/j/881519</a>		

# Zoom video conferencing



# Always On Zoom: hop on, hop off



## Persistent zoom session

A screenshot of a Slack conversation. The messages are:

- 8:33 AM | I'm game
- 8:33 AM | Aaron Bull Schaefer @dliappis @karmi just got out of my meeting...give me 5 minutes and I can jump on AoN?
- 8:33 AM | Dimitrios Liappis sure thing @abs
- 8:34 AM | Karel Minarik @abs, @dliappis, I'll try to jump in at least for a while...

The message from Aaron Bull Schaefer is highlighted with a green oval.

# Record meetings for posterity

Engineering [Public] > Cloud > recordings > Blackbelt Sessions

Name	Last modified	File size
Session2-CloudSaaSPostgressDataManagement.mp4	Sep 18, 2017	122 MB
Session1-CloudHighLevelArchitecture.mp4	Sep 18, 2017	122 MB
CloudSecurity-Session1_2017-07-06.mp4	Sep 18, 2017	86 MB
security_part2_2017-09-7.mp4	Sep 18, 2017	86 MB
2017-08-03 - CPU governance on Cloud.mp4	Sep 16, 2017	82 MB
security_part2_2017-09-7.mp4	Sep 16, 2017	86 MB
L1BlackbeltTraining-2017-01-11.mp4	Sep 6, 2017	42 MB
L1OnCallTrainingSREPlaybook.mp4	Sep 6, 2017	103 MB
20170617.CloudMetricsBlackbeltForSupport.mp4	Aug 25, 2017	61 MB
2017-06-15 Upgrading clusters due to EOL.mp4	May 15, 2017	25 MB
2017-06-11 Blackbelt - All things VPNs,clusters.mp4	May 11, 2017	6 KB

**Session2-CloudSaaSPostgressDataManagement.mp4**

DETAILS ACTIVITY



Type: Video  
Size: 122 MB (127,650,485 bytes)  
Storage used: 0 bytes Owned by anonymous user

Tags: #blackbelts, #cloud, #data-management, #postgress, #saas

Comments: 0

Shares: 0

Downloads: 0

Replies: 0

Likes: 0

Dislikes: 0

Views: 0

**It's not just  
virtual**

# Engineering all-hands: 2x a year



# Engineering all-hands: 2 times a

Vision, long term plans

Talks from CEO, VPs and  
investors

2 days team time

1 day cross product

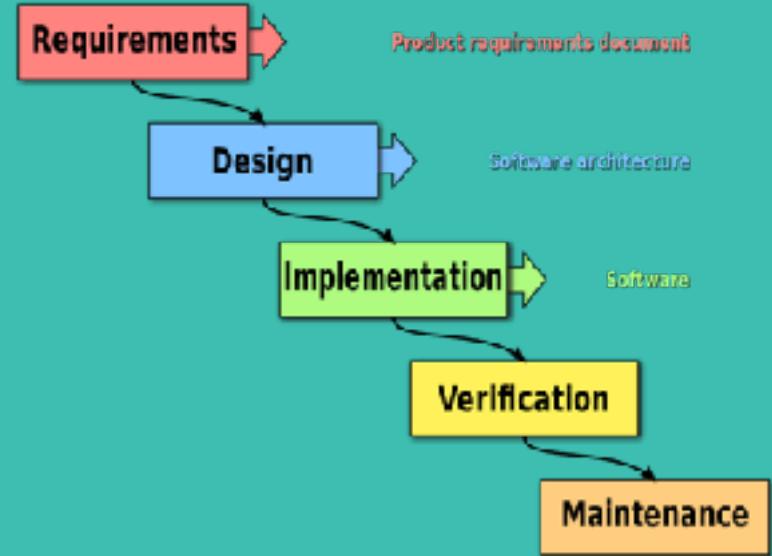
# Engineering meetings with a view



# Engineering all-hands: fun day!



~~dev~~  
~~process~~  
Agile, Waterfall



# Consensus



jayschwa commented on Oct 5, 2014



Jimmywarting commented on Oct 6, 2014



ChelseaStats commented on Oct 6, 2014

+1



dalanmiller commented on Oct 6, 2014



ryanseys commented on Oct 6, 2014



# GitHub to track things

Roadmap for 2.0 #9970

Open clintonjormley opened this issue on Mar 3 · 3 comments

clintonjormley commented on Mar 3

- #8312 Use doc values by default
- #8870 Field/type mapping refactor
- #9676 Perform computation on aggregations
- #6295 Update cluster state using diffs
- #10217 Refactor query and filter, agg, suggester parsing/serialization
- #8963 Warn when using deprecated syntax
  - + #3225 Upgrade to Natty-4
  - Lucene 5.1
  - Lucene 5.2

Index migration:

- #406 Move to one datapath per shard
- #10213 Allow upgrade API to only upgrade too-old segments
- #10214 Migration advisory \_site plugin
- #10215 Don't allow indices containing too-old segments to be opened

Edit New Issue

Owner Labels

Lgostash document improvements #10344

Open snygros opened this issue on Jun 10 · 5 comments

snygros commented on Jun 10

Meta issue to track improvements to Logstash reference docs

- Add deployment guide section for Logstash #9410 #3410
- Improve getting started/nmch/tutorials #3414
- Organize plugin documentation better #3410
- Add glossary of Logstash terms
- Add section for handling multiline logs #3059
- Consistently pass over plugin references from version 1.7 and 1.8
- Upgrading Logstash

snygros added doc meta labels on Jun 10

snygros commented on Jun 12

**Logstash document improvements #3417**

**Open** **suayograo** opened this issue on Jun 13 · 5 comments

**suayograo** commented on Jun 13

Move issue to track improvements to Logstash reference docs

- Add deployment guide section for Logstash #3410
- Improve getting started documentation #3411
- Organize plugin documentation better #3410
- Add glossary of Logstash terms
- Add section for handling multiline logs #3009
- Consistency pass over plugin reference across 1.7 and 2.0 #3413
- Upgrading Logstash

**suayograo** added **docs** **meta** labels on Jun 13

**suayograo** commented on Jun 13

Jin @thaigiai @palmerar

**Labels** **docs** **meta** **v1.6.0**

**Milestones** **Next milestone**

**Assignee** **palmerar**

**Notifications** **Unsubscribe**

*Your mailing  
subscription has been  
updated in this thread.*

**Participants**

# Fun with GitHub labels

## Unify cluster create API with UI experience #7961

Open nordbergm opened this Issue 22 days ago - 2 comments

nordbergm commented 22 days ago

There are currently a few areas where the cluster create API has options for clusters that aren't compatible with what can be done through the UI. The API experience should be the same as the UI so there aren't any "backdoors", especially for when the v1 API goes public in SaaS.

- API topologies should not have node type options, they should instead pass the node configuration id down to the plan (depends on [#7888](#)). In the absence of a node configuration id we should pass a default id into the plan which represents a "standard" cluster.
- (TBD) Nodes per zone may have to be calculated from the capacity if we end up going for a SaaS-like capacity slider in ECE. To prevent BWC breakage we could have a feature flag for this so we always have the higher level behaviour in SaaS.

/cc @uric @AlexP-Elastic

 1

 nordbergm added `area:api` `project:dedicated-node-types` `loc:medium` labels 22 days ago

 AlexP-Elastic added the `current` label 22 days ago

Edit New issue

Assignees: No one—assign yourself

Labels: `area:api` (selected), `current`, `discuss`, `loc:medium`, `project:dedicated-node-types`

Projects: None yet

Milestone: No milestone

Notifications: 

# Jira-ifying GitHub with Waffle.io

The image shows a Waffle.io board with five columns: Backlog, Current, In Progress, Needs Review, and Done. Each column lists GitHub issues with their titles, labels, and descriptions.

- Backlog:** 126 / 1701 0 / 18
- Current:** 76 / 177
- In Progress:** 22 / 77
- Needs Review:** 5 / 56
- Done:** 36 / 615

Issues in the Done column (partial list):

- #442 elastic/cloud: Bring back RAM/Disk field in the ECE UI. Labels: area:api, area:ui, blocker. Version: 1.1.0ea.
- #444 elastic/cloud: Pr/8215. Labels: area:api, area:core, project:dedicated-node-types.
- #420 elastic/cloud: Set repositories API should validate repository type. Labels: area:api.
- #351 elastic/cloud: Purging "AWS Resource" buckets fails to purge buckets of clusters who have been scaled down. Labels: affect: production, area:cloud, priority:1, blocker.
- #845 elastic/elasticsearch: Add disk usage on allocators endpoint. Labels: area:api, area:core, project:api-v1-migration, priority:1.
- #190 elastic/cloud: Plan attempt logs - clean up doesn't parse correctly? Labels: area:api, bug, loemail, priority:2, project:api-v1-migration, releases-as-release-september-2011, version:1.1.0ga.
- #3855 elastic/cloud: Make license safety period > the warning time. Labels: area:api, loemail, priority:3, releases-as-release-september-2011, version:1.1.0ga.
- #7998 elastic/cloud: Restrict IP address of adminconsole users with tryuser\_ip. Labels: area:api, area:elasticsearch, loemail, low-medium, priority:3, releases-as-release-september-2011, security, triage.
- #961 elastic/cloud: Unify cluster create API with III experience. Labels: area:api, discuss, low-medium, project:dedicated-node-types.
- #4779 elastic/cloud: Meter Usage for Billing Purposes (aka usage API) Labels: area:api, area:core, loemail, project:cloud-business-ops, project:new-cloud-billing-system, releases-as-release-september-2011, version:1.1.0ga.
- #1082 elastic/cloud: Possible issue with restoring 2x snapshots into 5x clusters. Labels: affects:production, area:api, bug, releases-as-release-september-2011, triage.
- #1165 elastic/elasticsearch: VI system settings could be politer in terms of letting the constructor know what's changed. Labels: area:api, loemail, priority:2, project:api-v1-migration.
- #247 elastic/cloud: Finalize 1.1 fine-grained authz model and fill in from DB. Labels: area:api, storage, percent-complete:80, project:dedicated-node-types.
- #643 elastic/cloud: Change metadata raw body field to object. Labels: area:api, bug, loemail, project:api-v1-migration.
- #1105 elastic/cloud: Fix reserved word mappings for coverage tests client. Labels: area:api.
- #7811 elastic/cloud: Add missing 412 response in updateElasticsearch Cluster Plan. Labels: area:api, loemail.
- #7999 elastic/elasticsearch: Fix api-coverage-tests build after #7864. Labels: area:api.
- #8933 elastic/cloud: Add admin id to user service and change source. Labels: area:api.
- #6194 elastic/cloud: Improve maintainability of swagger via macros. Labels: area:api, loemail, priority:3, project:api-v1-migration.
- #7284 elastic/cloud: VI API to support changes to rolling reboots. Labels: area:api.
- #7922 elastic/cloud: VI API to support changes to rolling reboots. Labels: area:api.

# Pull Requests: Peer Review

Add logs on runner & allocator connection status change

Edit

#7014

Merged

swallez merged 2 commits into `elastic:master` from `swallez:add-logs-on-disconnect-7006` on Jul 26

Conversation 4

Commits 2

Files changed 2

+33 -6



swallez commented on Jul 24



Fixes #7006

Logs a warning whenever the connection status of a runner or an allocator changes (the `/connected` ephemeral ZK node).

Also adds a bit of whitespace to improve readability :-)



2



2



2

swallez added some commits on Jul 24

Log runner & allocator connection state changes

0b80bfaf

Add whitespace, remove unused variable

x 2aF267c

Reviewers

matschaffer



joegallo



Assignees

swallez



Labels

None yet



Projects

None yet



astic

LGTMs ==  
ACKs



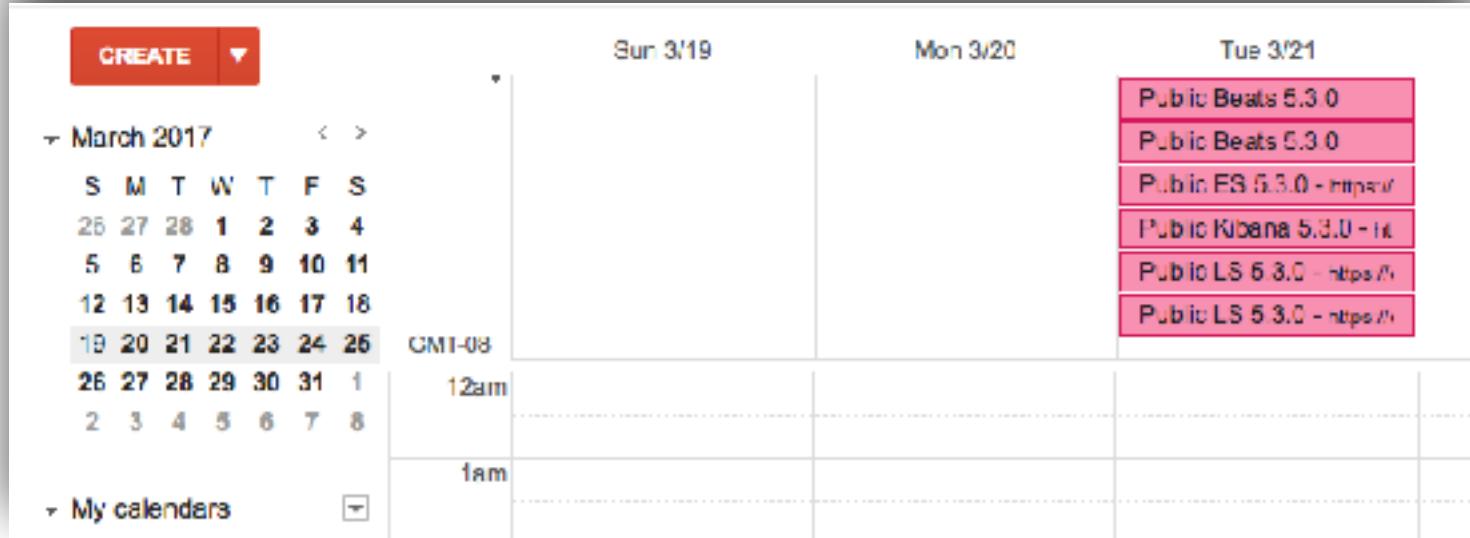
# Releases

# Before: Independent releases



# Time based releases

# Release Process: Time based



x.y.z →

major minor patches

# Release: Train analogy



# Release Process: Time based

Merge when ready and  
scalable

Otherwise catch the next train

No pressure on devs for

# Team dynamic

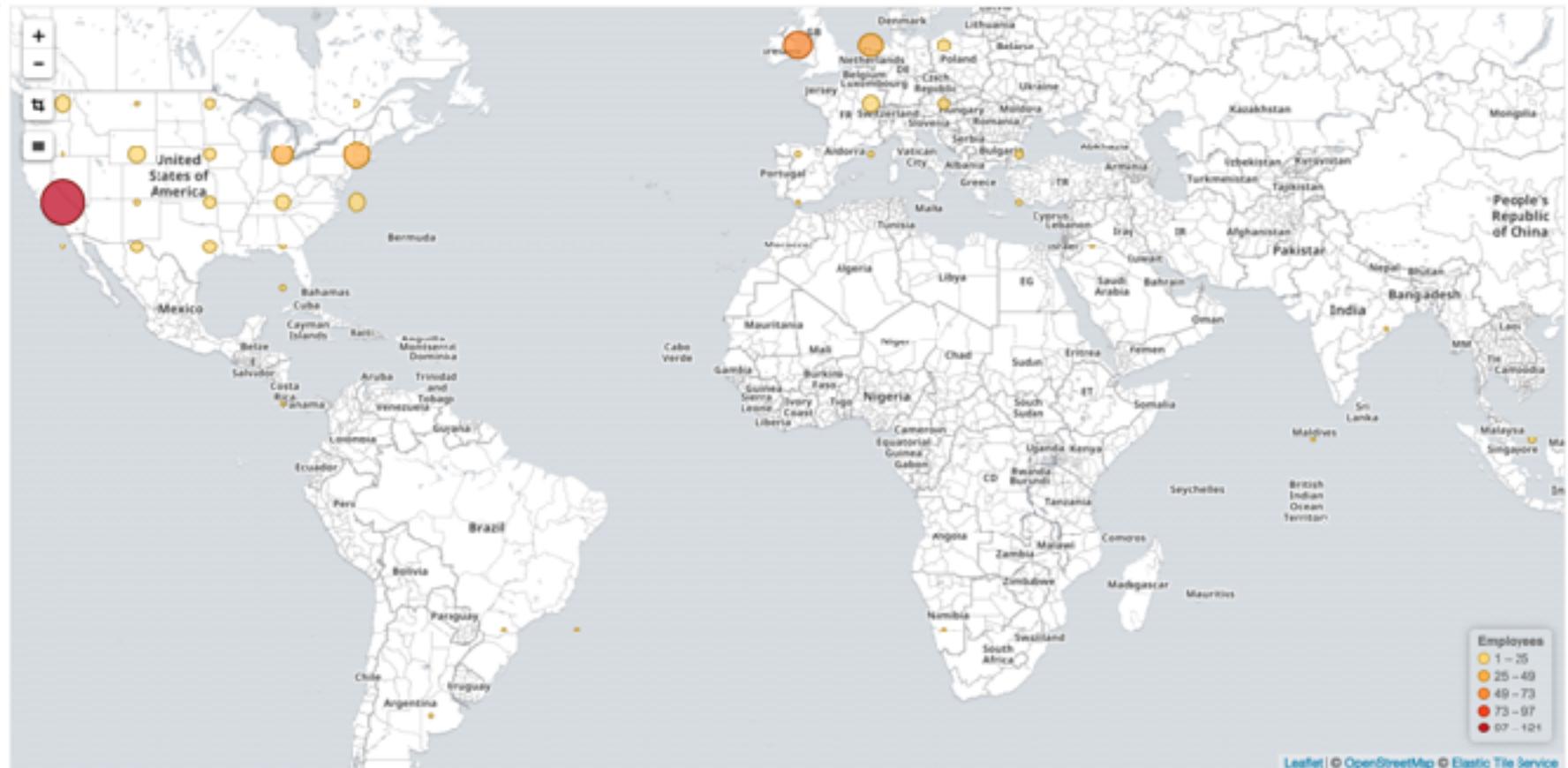


# How's the org structure

Flat org

Self organising teams

Highly available



Employees

- 1 - 25
- 25 - 49
- 49 - 73
- 73 - 97
- 97 - 121

## Golang Developer - Beats Team

APPLY NOW

Location: Anywhere

Department: Engineering

# Globally distributed

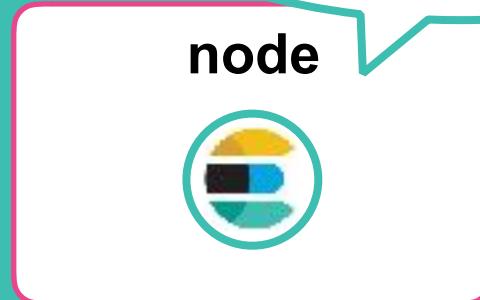
# Hire the best talent

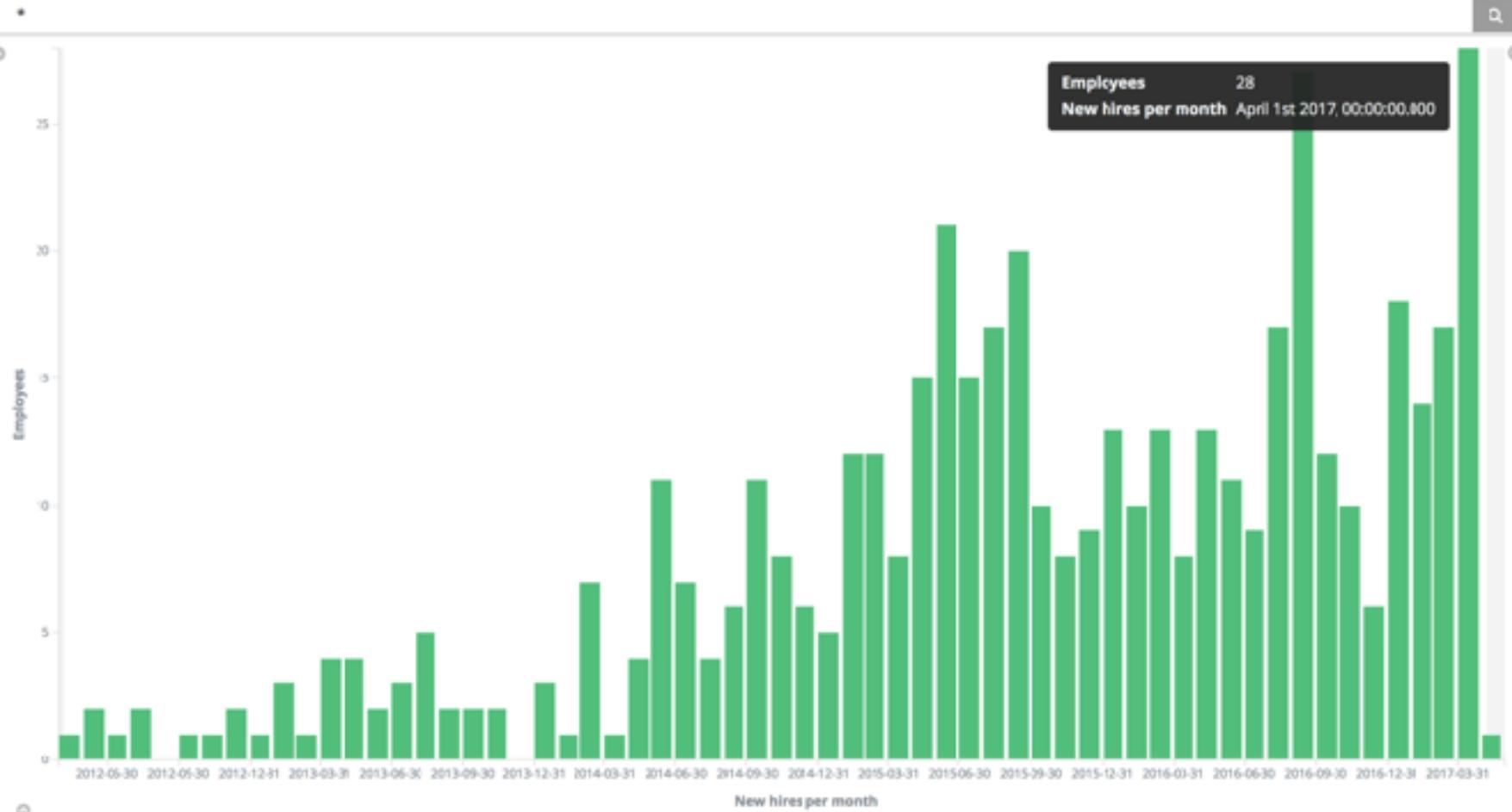
A node has many  
roles

Engin ~~blète~~odes

# node discover y

*looking for other  
nodes in elastic*





# Hiring

~~“how many golf balls fit in~~

# Hiring

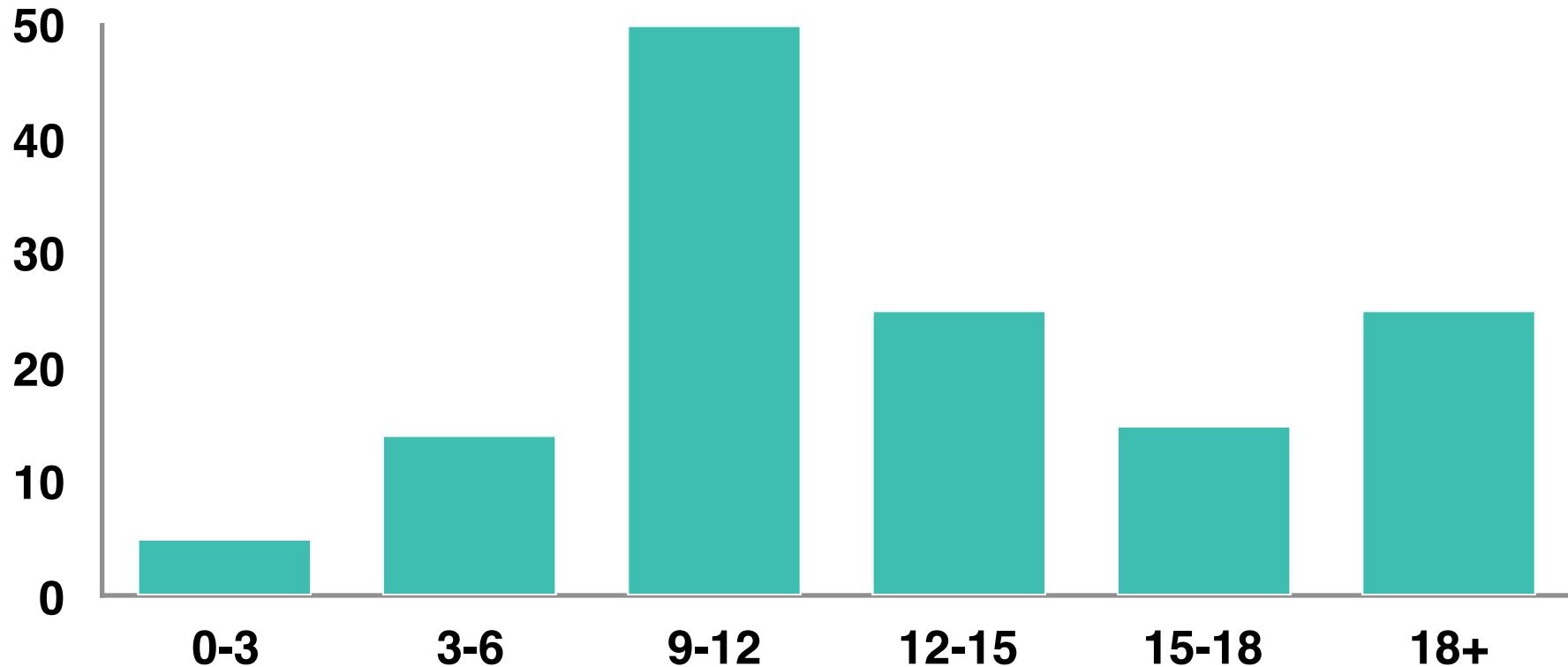
No puzzle questions

Each team has different interview styles

Technical and team fit for distributed environment

**Nodes are  
independent**

# Average experience



# On Boarding: Mentoring

Day 1

High touch  
mentoring

Mentor !=



# On Boarding: X-School

Week-long orientation in  
Mountain View, CA

Meet team leads, product team  
Big Picture: Understand



# Takeaways

# Key takeaways

Transparency & openness

# Key takeaways

Trust by default

# Key takeaways

Don't build processes

# Key takeaways

Progress over perfection

# In closing...

They say it won't work at 50,

# Questions?

# More resources

Initially presented as OSCon, Austin: <https://conferences.oreilly.com/oscon/oscon-tx/public/schedule/detail/56992>

<https://www.elastic.co/elasticon/conf/2016/sf/inside-elastic-engineering>

<https://www.elastic.co/elasticon/conf/2017/sf/on-distributed-systems-and-distributed-teams>

<https://zachholman.com/posts/how-github-works/>

Fonts and slide design were inspired by Zach Holman: <https://zachholman.com/posts/slides-design-for-developers/>