

OpenStack Trove

Open Source DBaaS for the Cloud

Bartosz Żurkowski

Samsung R&D Institute Poland

SOSCON 2018

SAMSUNG OPEN SOURCE CONFERENCE 2018



Outline

Database as a Service introduction

Solutions available on the market

OpenStack Trove overview

Trove in a bigger picture

Towards open source contribution

SOSCON 2018

SAMSUNG OPEN SOURCE CONFERENCE 2018



Who are we?

SOSCON 2018

- Samsung R&D Institute Poland (SRPOL)
- Building private cloud platform for 5G networks requirements based on OpenStack
- Team developing and maintaining DBaaS component - OpenStack Trove
- Active open source community members
 - Trove PTL (Project Technical Lead)
 - 3 core contributors



Dariusz Król
Trove PTL



Bartosz Żurkowski
Trove core
contrib.



Marcin Piwowarczyk
Trove core
contrib.



Przemysław Godek



Cezary Żukowski

Operating databases is hard

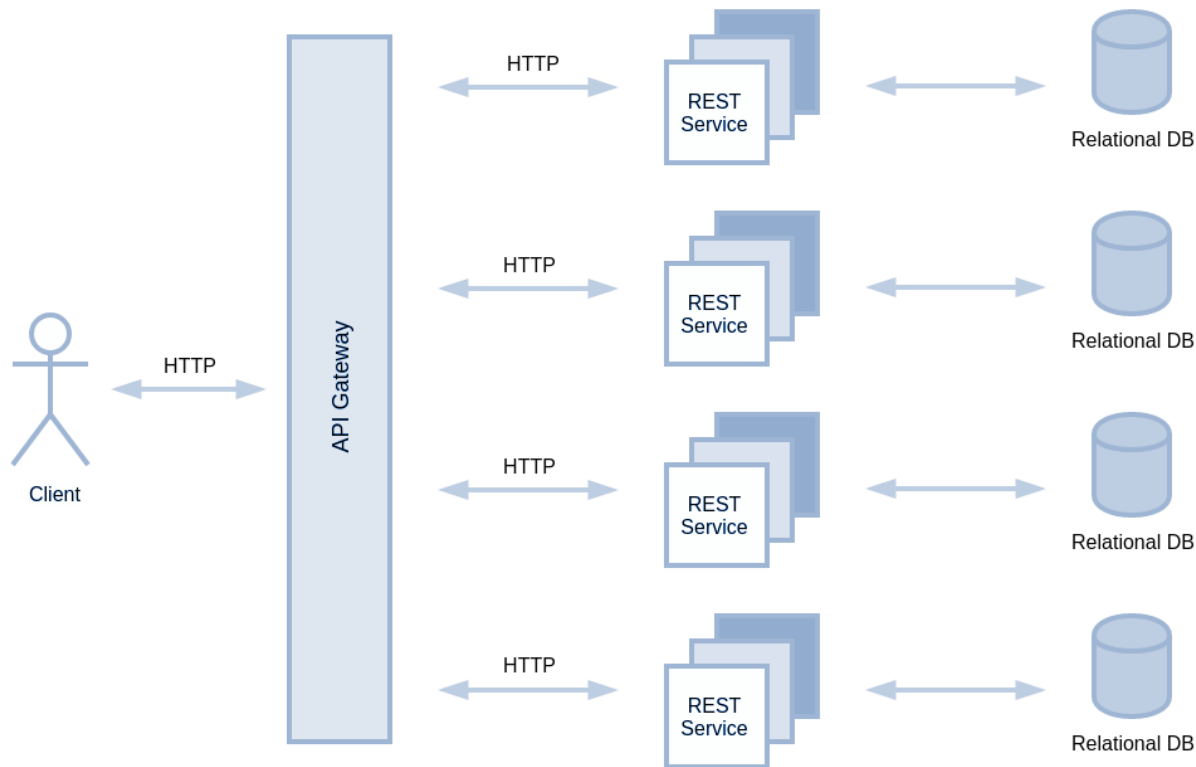
SOSCON 2018

 DBA responsibilities

Application optimization
DB performance tuning
High availability
Monitoring
Scaling
Periodic backups
DB software upgrades
Hardening
DB software setup and config
Virtual resource provisioning

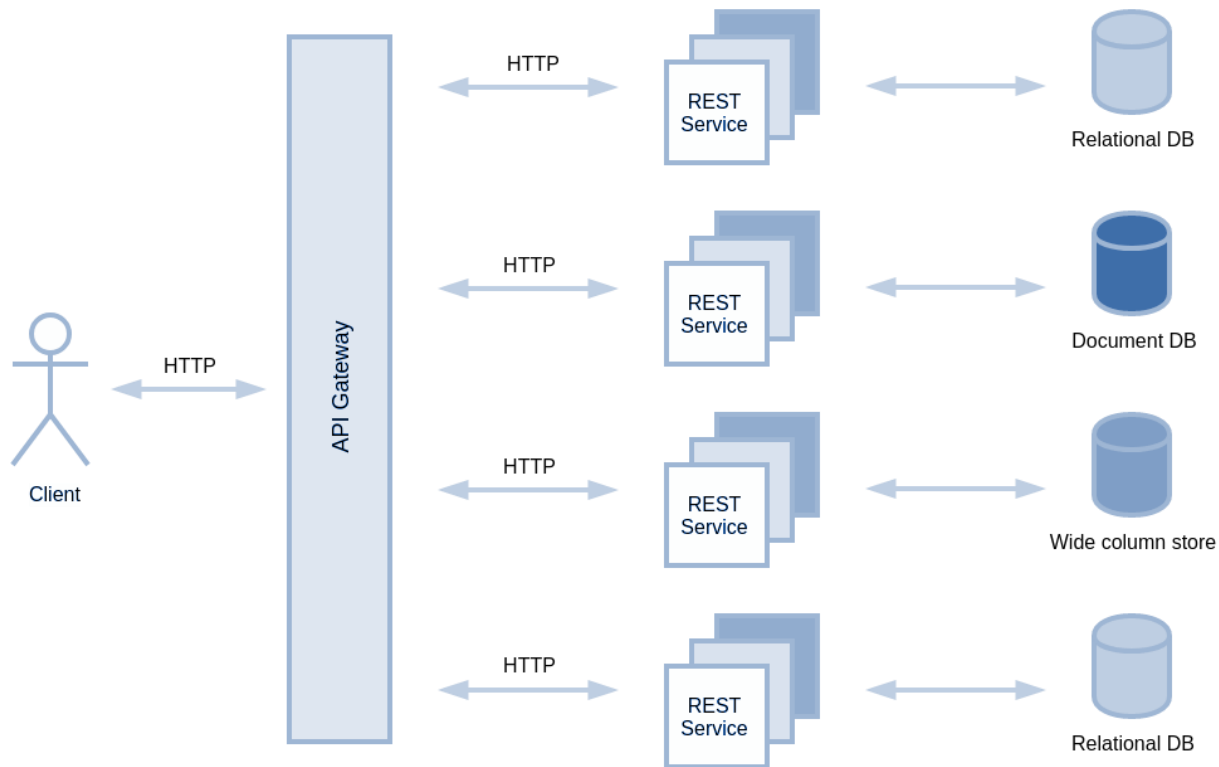
...in particular, when there are many instances

SOSCON 2018



...instances of different types

SOSCON 2018



... since we live in an era of polyglot persistence

SOSCON 2018

Tabular nature of data

Recommendations



Rapidly traverse links between entities

Reporting



User sessions



Rapid R/W access, no need to be durable

Products



Products make natural document aggregate

Analytics



Large-scale analytics on large cluster

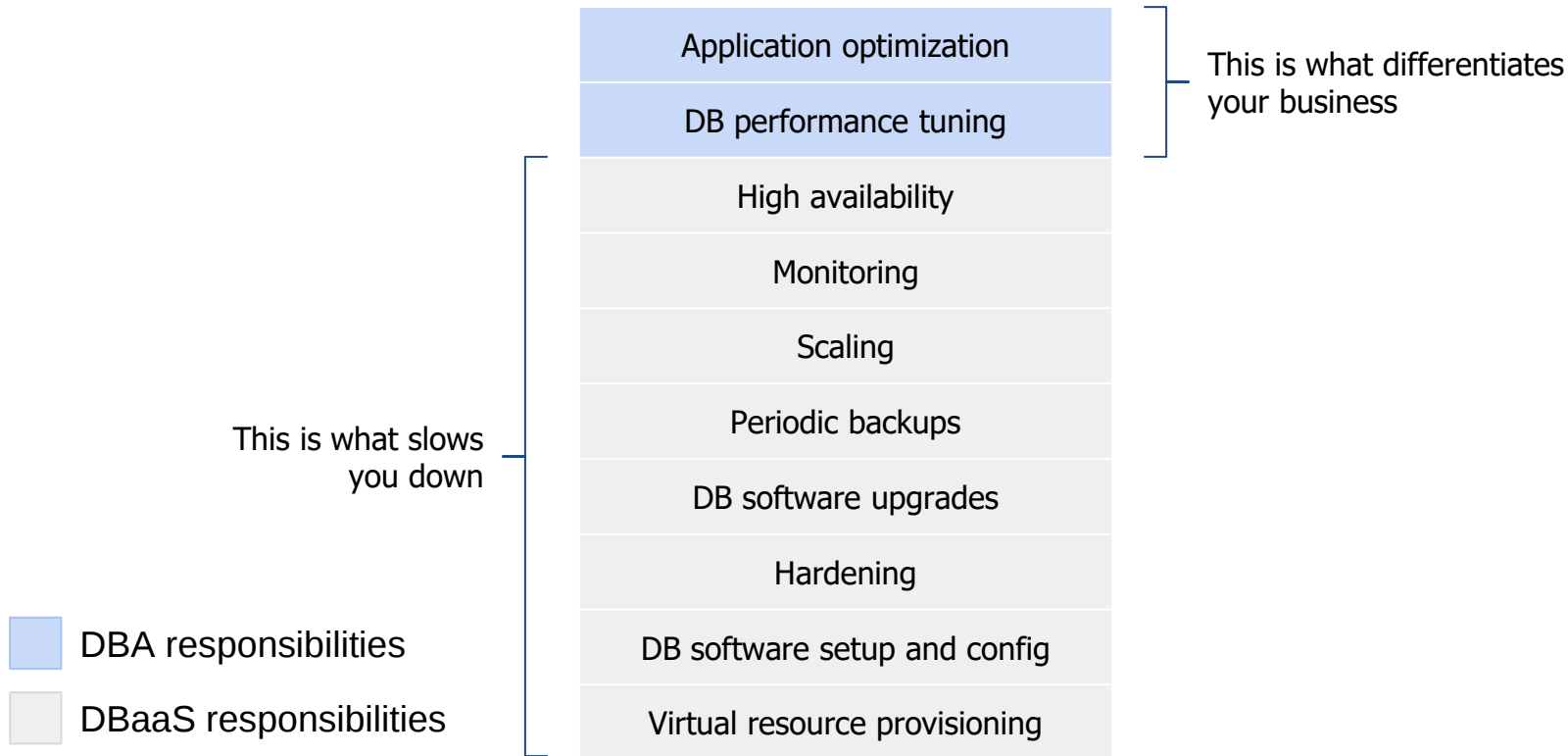
Financial data



Tabular nature of data, transactional updates

What is Database as a Service?

SOSCON 2018



DBaaS solutions available on the market

SOSCON 2018



- **AWS Aurora**
Cloud native SQL database
- **AWS RDS**
MySQL, MariaDB, PostgreSQL, Oracle, Microsoft SQL Server
- **AWS DynamoDB**
Key-value and document structs
- **AWS ElastiCache**
Redis, memcached
- **AWS Neptune**
Graph database



- **Azure SQL Database**
- **Azure Database for MySQL**
- **Azure Database for PostgreSQL**
- **Azure Cosmos DB**
Key-value, graph, column-family, and document structs
- **Redis Cache**



- **Google Cloud SQL**
Mysql, PostgreSQL
- **Google Cloud Bigtable**
Wide column store
- **Google Cloud Spanner**
- **Google Cloud Memorystore**
Redis



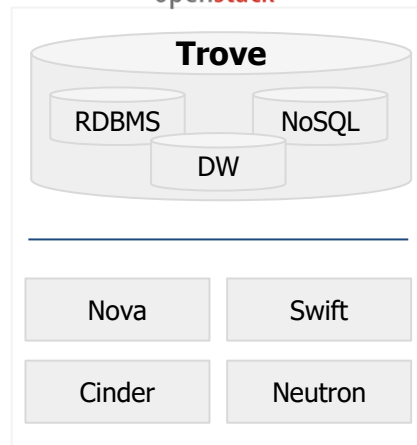
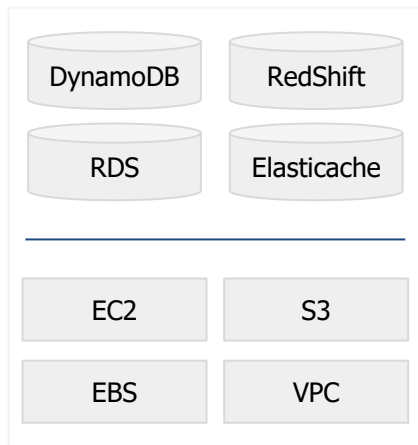
TROVE

an OpenStack Community Project

Introducing OpenStack Trove

SOSCON 2018

- Database as a Service for OpenStack
- Build entirely on OpenStack
- Started around 2012
- Entered *"The Big Tent"* in 2014
- Full database lifecycle management
- 11 datastores
- Unified user interface
- Original project sponsor: Hewlett-Packard
- Major contributors over time:
Rackspace, Tesora, HP, Red Hat, Mirantis,
EasyStack, SUSE, eBay, China Mobile, IBM
- Current PTL: Dariusz Król, Samsung SRPOL



Datastore capability matrix

SOSCON 2018

SAMSUNG

SAMSUNG

Datastore	Type	Provisioning	Schema management	Backup and restore	Resizing	Replication	Clustering	Cluster backups	Cluster resizing
MySQL	Relational	✓	✓	✓	✓	✓	✓	✓	✓
MariaDB		✓	✓	✓	✓	✓	✓	✓	✓
PostgreSQL		✓	✓	✓	✓	✓	✗	✗	✗
Percona		✓	✓	✓	✓	✓	✗	✗	✗
DB2		✓	✓	✓	✓	✗	✗	✗	✗
Redis	Key-value	✓	—	✓	✓	✓	✓	✓	✓
Cassandra	Column	✓	✓	✓	✓	—	✓	✓	✓
Vertica		✓	✓	✗	✓	—	✓	✗	✗
MongoDB	Document	✓	✓	✓	✓	—	✓	✗	✗
CouchDB		✓	✓	✓	✓	—	✗	✗	✗
Couchbase	Multi-model	✓	✗	✓	✓	—	✗	✗	✗

Features overview

SOSCON 2018



Instance provisioning



Configuration



Backup and restore



Replication



Clustering

Features overview: Instance provisioning

SOSCON 2018

- Create a database instance

```
$ trove create db-inst-1 m1.xlarge
--datastore mariadb
--datastore_version 10.1.32
--size 2000
--nic 'net-id=int-net'
--databases db1 db2 db3
--users usr1:pass1 usr2:pass2
--availability-zone zone-1
--configuration mariadb-cg
```

Launch Instance

Details

Networking

Initialize Databases

Advanced

Availability Zone

nova

Specify the details for launching an instance.

Please note: The value specified in the Volume Size field should be greater than 0, however, some configurations do not support specifying volume size. If specifying the volume size results in an error stating volume support is not enabled, enter 0.

Instance Name

db-instance

Volume Size

2000

Volume Type

lvmdriver-1

Datastore

mariadb - 10.1.32

Flavor

m1.xlarge

Locality

None

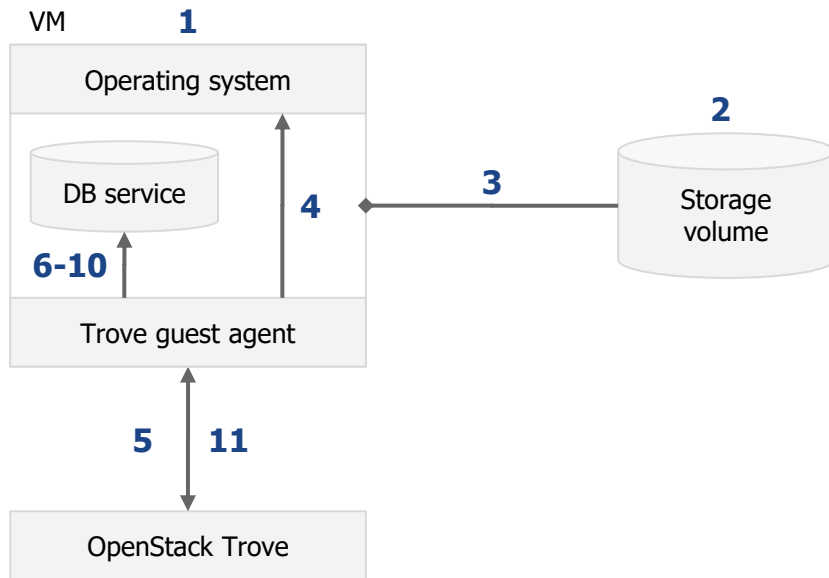
Cancel

Launch

Features overview: Instance provisioning

SOSCON 2018

- **1.** Spawn VM
- **2.** Allocate storage volume
- **3.** Attach volume to VM
- **4.** Format and mount filesystem
- **5.** Coordinate database service setup
- **6.** Apply initial configuration
- **7.** Start database service
- **8.** Secure database (e.g. remove anonymous user)
- **9.** Create initial databases and users
- **10.** Check service status
- **11.** Report database status to the control plane



- List available configuration parameters

```
$ trove configuration-parameter-list 10.1.32 --datastore mariadb
```

Name	Type	Min Size	Max Size	Restart Required
auto_increment_increment	integer	1	65535	False
auto_increment_offset	integer	1	65535	False
...	...	...	...	...
innodb_buffer_pool_size	integer	5242880	18446744073709551615	True
...	...	...	...	...
max_connections	integer	1	100000	False
...	...	...	...	...

- Create a configuration group

```
$ trove configuration-create mariadb-cg
```

```
    '{"max_connections": 1000, "wait_timeout": 50000}'
```

```
--datastore mariadb
```

Features overview: Configuration management

SOSCON 2018

- Attach configuration to an instance

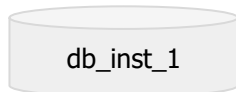
```
$ trove configuration-attach  
    db-inst-1 mariadb-cg  
$ trove configuration-attach  
    db-inst-2 mariadb-cg
```

- Update configuration

```
$ trove configuration-update mariadb-cg  
    '{"max_connections": 2000,  
    "wait_timeout": 100000}'
```

- Detach configuration

```
$ trove configuration-detach db-inst-1
```

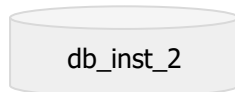


max_connections: 150
wait_timeout: 30000

max_connections: 1000
wait_timeout: 50000

max_connections: 2000
wait_timeout: 100000

max_connections: 150
wait_timeout: 30000



max_connections: 150
wait_timeout: 30000

max_connections: 1000
wait_timeout: 50000

max_connections: 2000
wait_timeout: 100000

Features overview: Backup and restore

SOSCON 2018

- Create backup of an instance

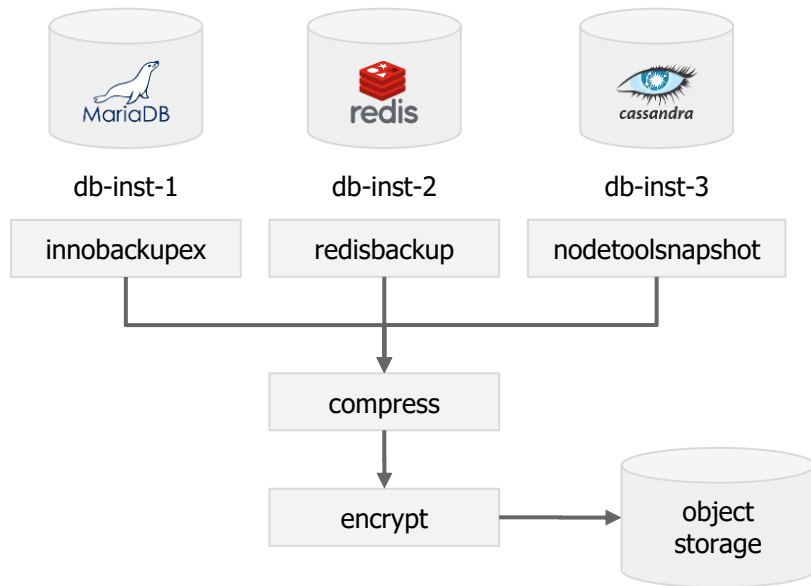
```
$ trove backup-create db-inst-1 bkp-1  
$ trove backup-create db-inst-2 bkp-2  
$ trove backup-create db-inst-3 bkp-3
```

- Restore backup into a new instance

```
$ trove create db-inst-4 m1.small  
    --datastore mariadb  
    --backup bkp-1  
    ...
```

- Schedule periodic backup

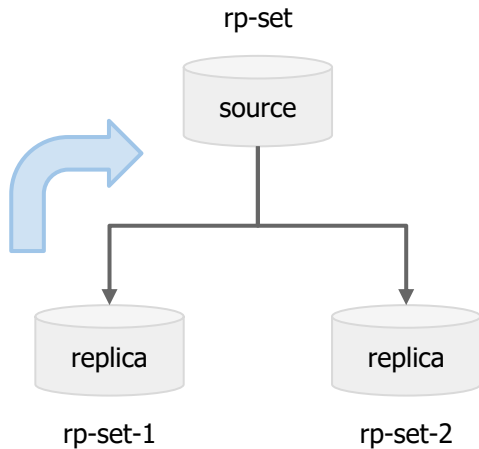
```
$ trove schedule-create  
    db-inst-1 "0 0 2 ? * *" daily-bkp
```



Features overview: Replication

SOSCON 2018

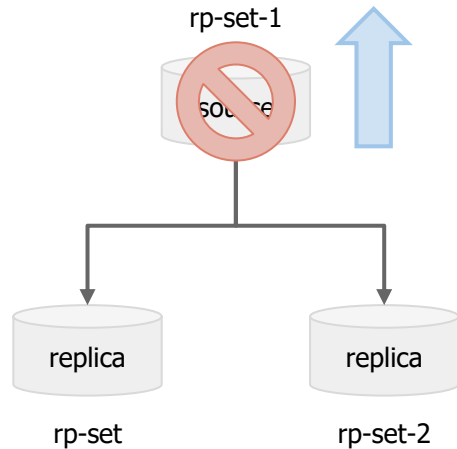
- Create a replica set
`$ trove create rp-set ...`
- Add replicas to replica set
`$ trove create rp-set
--replica_of rp-set --replica_count 2`
- Promote replica to be the new replica source of its set
`$ trove promote-to-replica-source rp-set-1`
- Eject replica source
`$ trove eject-replica-source rp-set-1`
- Detach replica from its replica source
`$ trove detach-replica rp-set-2`



Features overview: Replication

SOSCON 2018

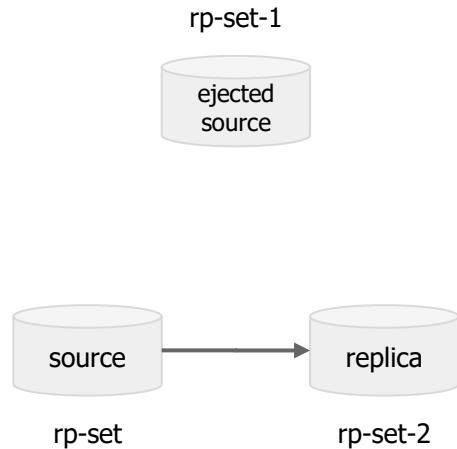
- Create a replica set
`$ trove create rp-set ...`
- Add replicas to replica set
`$ trove create rp-set
--replica_of rp-set --replica_count 2`
- Promote replica to be the new replica source of its set
`$ trove promote-to-replica-source rp-set-1`
- Eject replica source
`$ trove eject-replica-source rp-set-1`
- Detach replica from its replica source
`$ trove detach-replica rp-set-2`



Features overview: Replication

SOSCON 2018

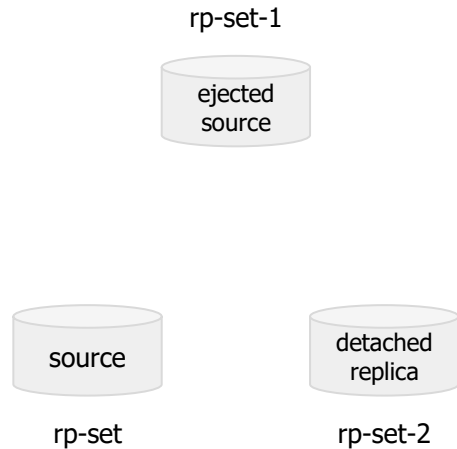
- Create a replica set
`$ trove create rp-set ...`
- Add replicas to replica set
`$ trove create rp-set
--replica_of rp-set --replica_count 2`
- Promote replica to be the new replica source of its set
`$ trove promote-to-replica-source rp-set-1`
- Eject replica source
`$ trove eject-replica-source rp-set-1`
- Detach replica from its replica source
`$ trove detach-replica rp-set-2`



Features overview: Replication

SOSCON 2018

- Create a replica set
`$ trove create rp-set ...`
- Add replicas to replica set
`$ trove create rp-set
--replica_of rp-set --replica_count 2`
- Promote replica to be the new replica source of its set
`$ trove promote-to-replica-source rp-set-1`
- Eject replica source
`$ trove eject-replica-source rp-set-1`
- Detach replica from its replica source
`$ trove detach-replica rp-set-2`



- Create a cluster

```
$ trove cluster-create db-cluster-1 mariadb 10.1.26
  --instance "flavor=m1.small,volume=1000,nic='net-id'=int-net"
  --instance "flavor=m1.small,volume=1000,nic='net-id'=int-net"
  --instance "flavor=m1.small,volume=1000,nic='net-id'=int-net"
  --locality anti-affinity
```

- Add nodes to a cluster

```
$ trove cluster-grow db-cluster-1
  --instance "flavor=m1.small,volume=1000,nic='net-id'=int-net"
```

- Remove nodes from a cluster

```
$ trove shrink-cluster db-cluster-1 db-cluster-1-member-4 db-cluster-1-member-4
```

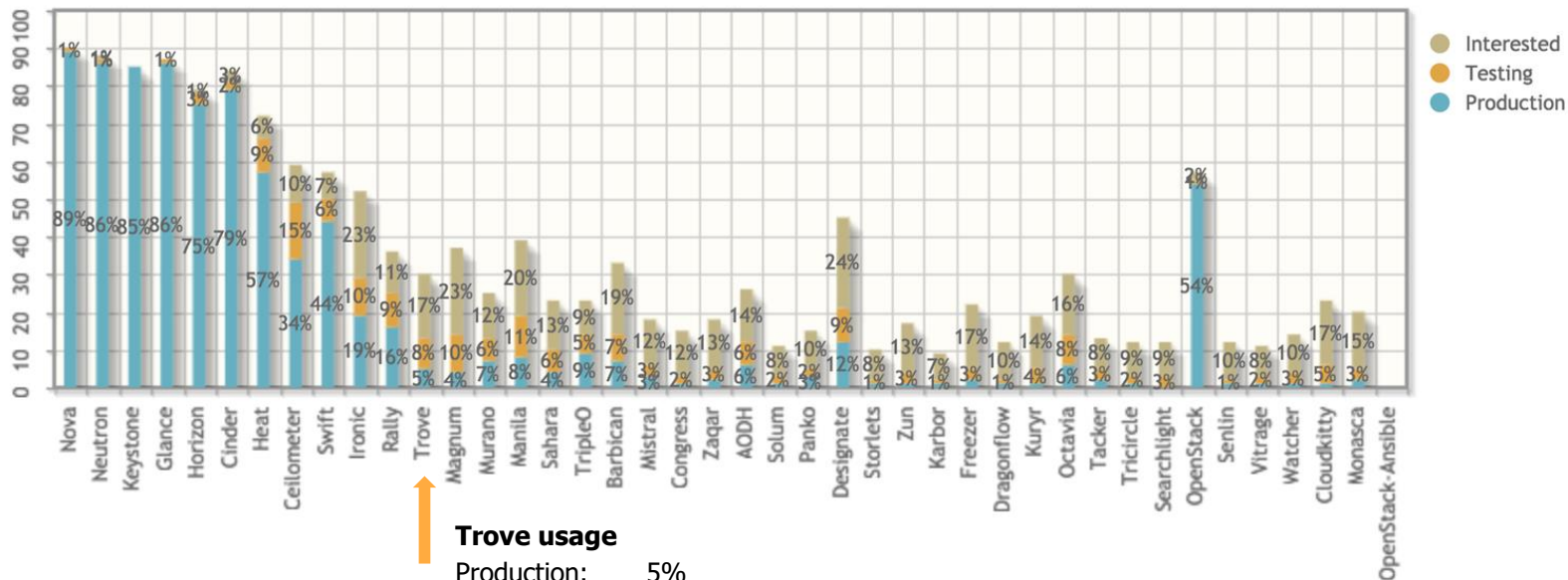
- Upgrade cluster to a new datastore version

```
$ trove cluster-upgrade db-cluster-1 mariadb 10.1.32
```

Is it production ready?

SOSCON 2018

Projects used in Production Deployments



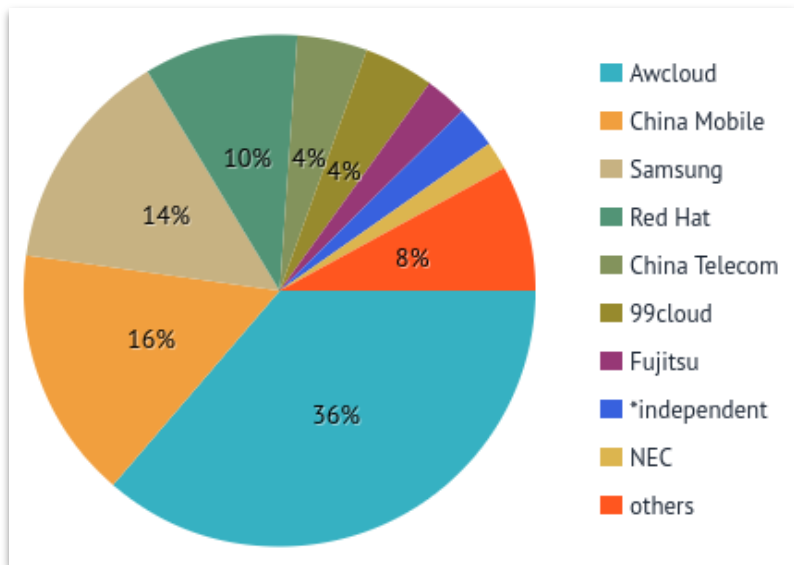
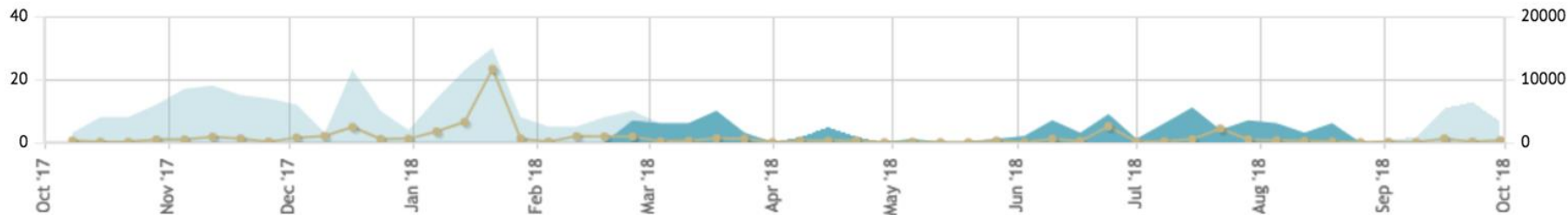
Trove usage

Production: 5%
Testing: 8%
Interested: 17%

n=357

Project condition

SOSCON 2018



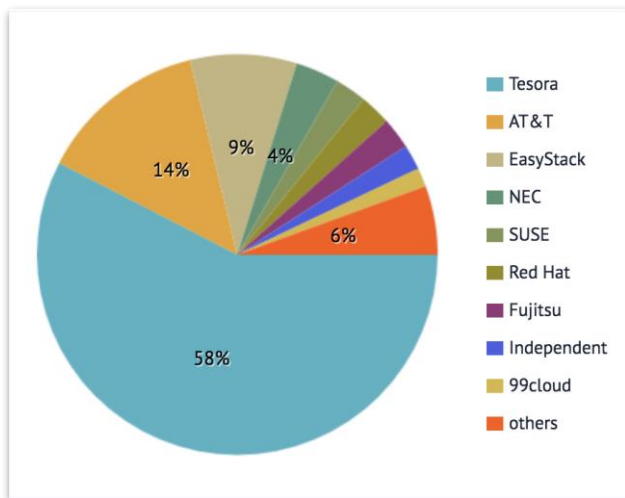
Latest release (Rocky) stats:

- Commits: 113
- Contributors: 31 (10 active)
- Reviewers: 46 (9 active)
- Core team: 10

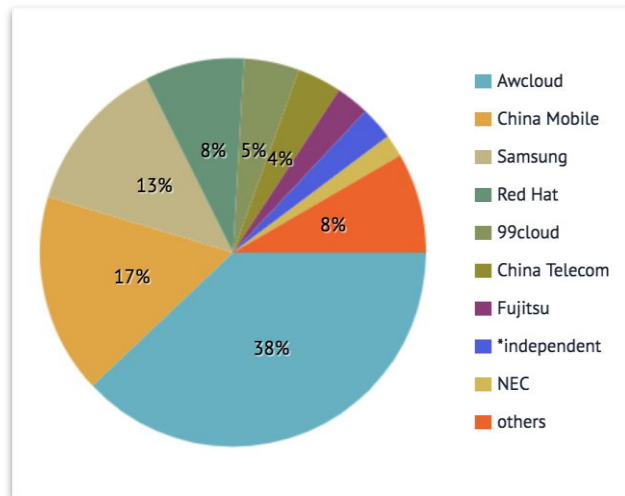
Large projects are built by companies

SOSCON 2018

Trove
Ocata



Trove
Rocky

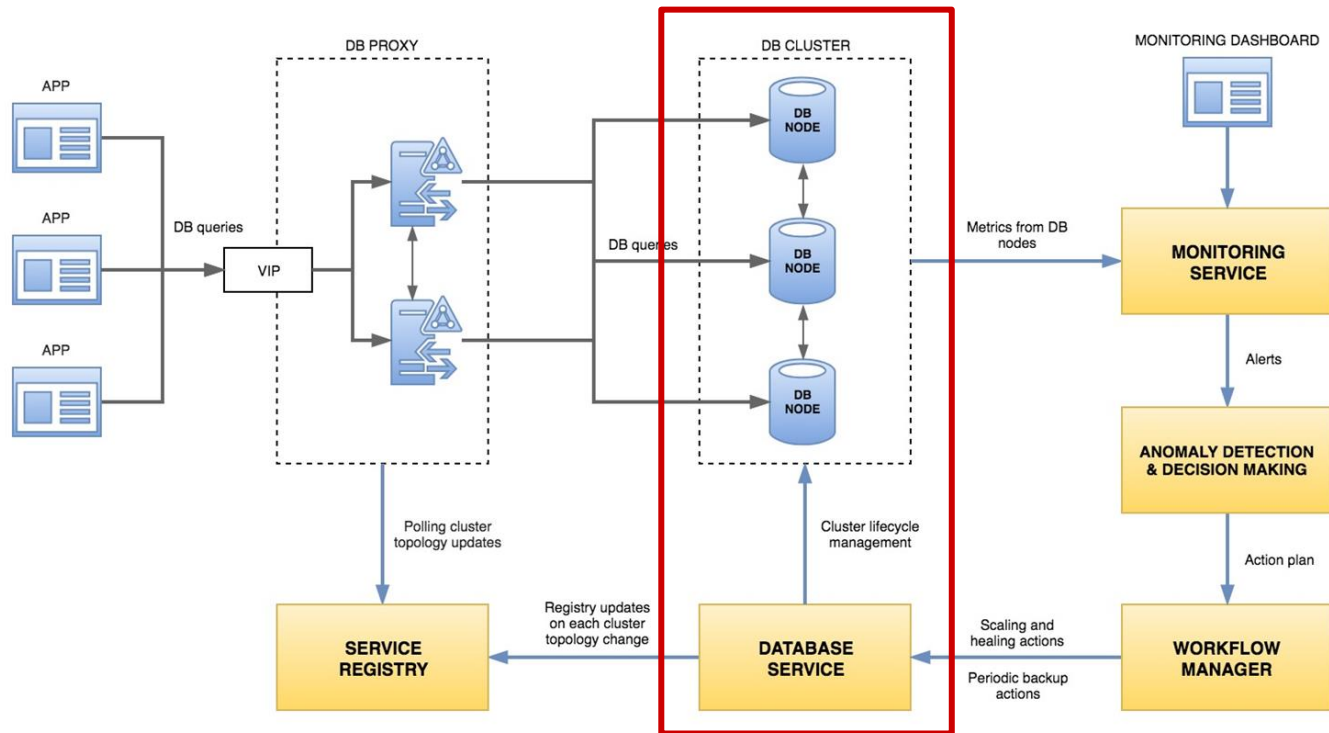


#	Company	Commits
1	Tesora	114
2	AT&T	27
3	EasyStack	17
4	NEC	7
5	SUSE	5
6	Red Hat	5
7	Fujitsu	5
8	Independent	4
9	99cloud	3
10	ZTE Corporation	2

#	Company	Commits
1	Awcloud	41
2	China Mobile	18
3	Samsung	14
4	Red Hat	9
5	99cloud	5
6	China Telecom	4
7	Fujitsu	3
	*independent	3
8	NEC	2
9	EasyStack	2

Trove in a bigger picture

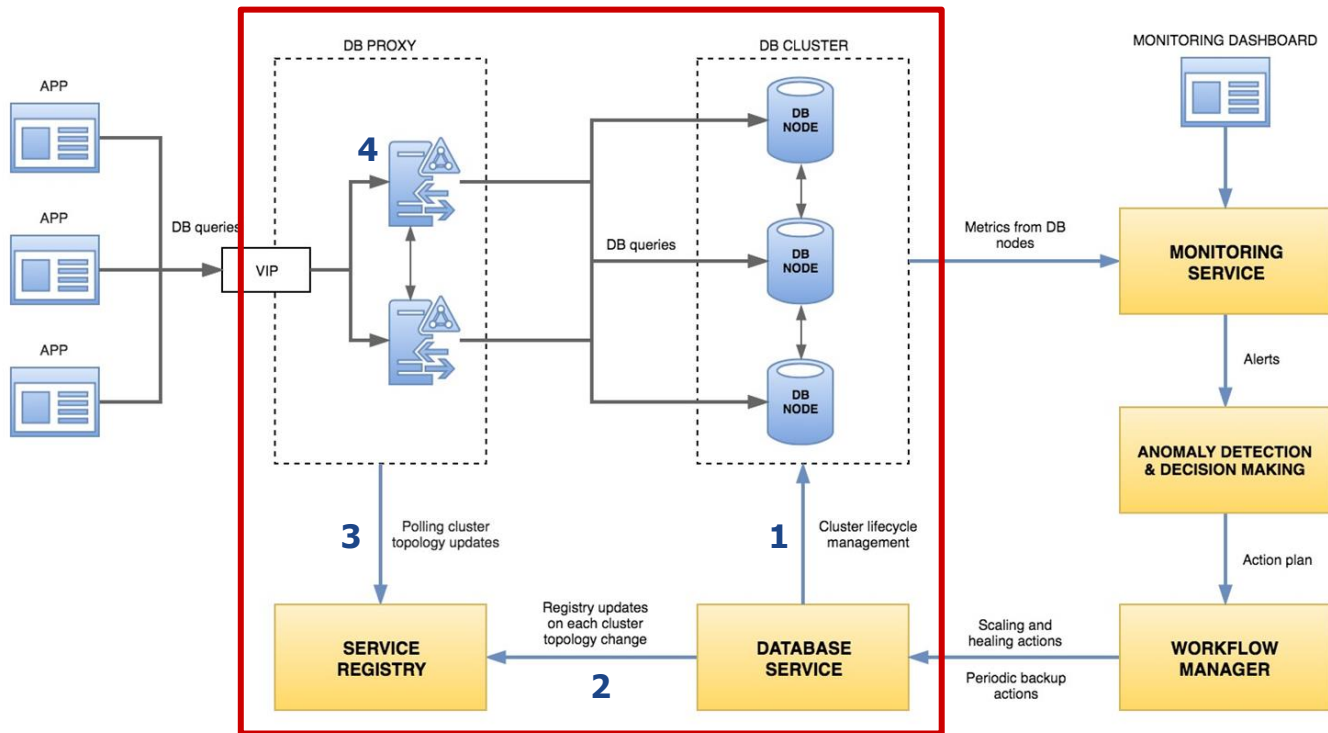
SOSCON 2018



Trove in a bigger picture: Database proxy

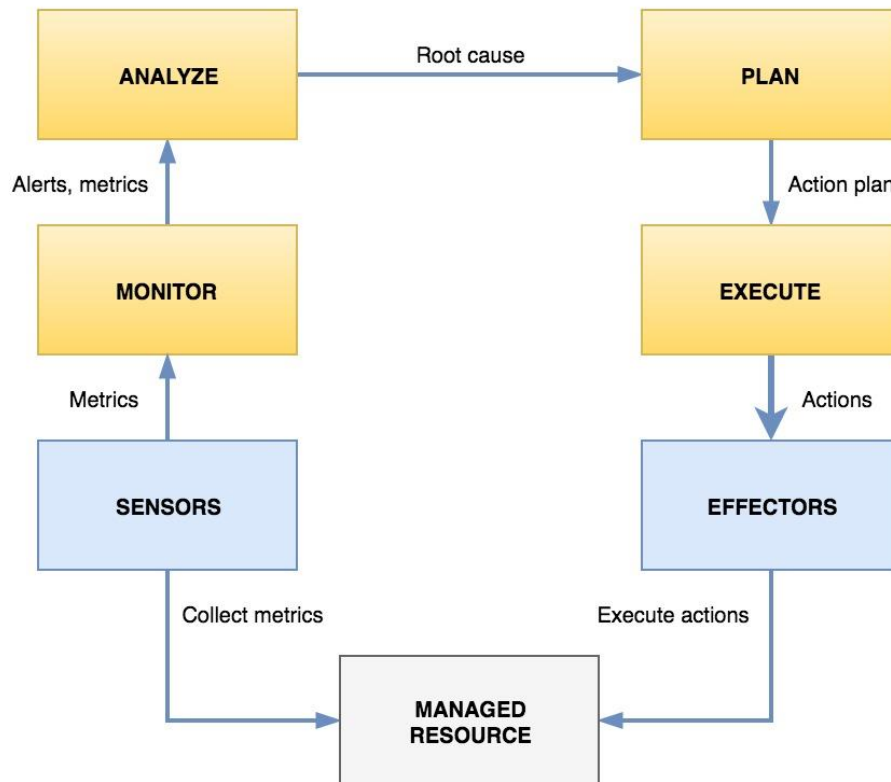
SOSCON 2018

- 1. Trove adjusts cluster topology according to user request
- 2. ... and sends updates to service registry (e.g. Consul or etcd)
- 3. Proxy instances poll registry updates from service registry
- 4. ... and reconfigure load balancing strategy



Trove in a bigger picture: Autonomic control loop

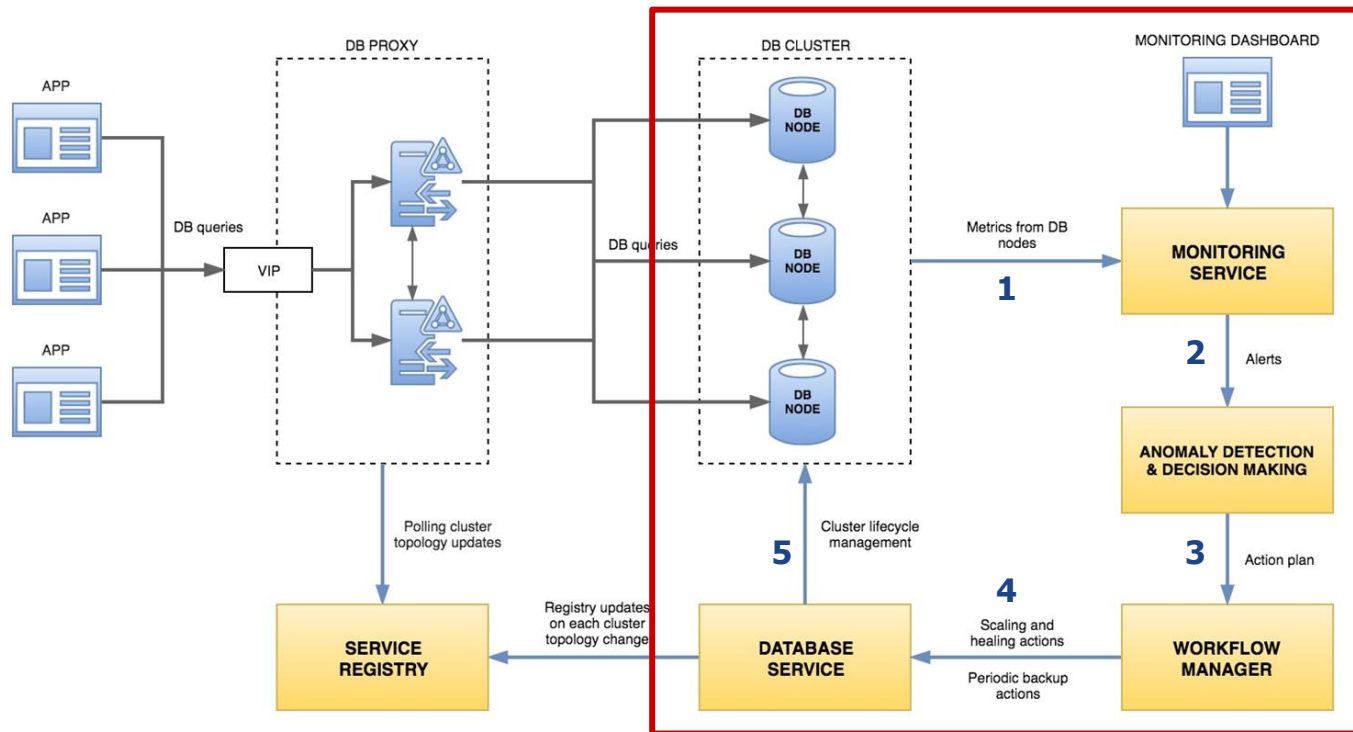
SOSCON 2018



Trove in a bigger picture: Autonomic control loop

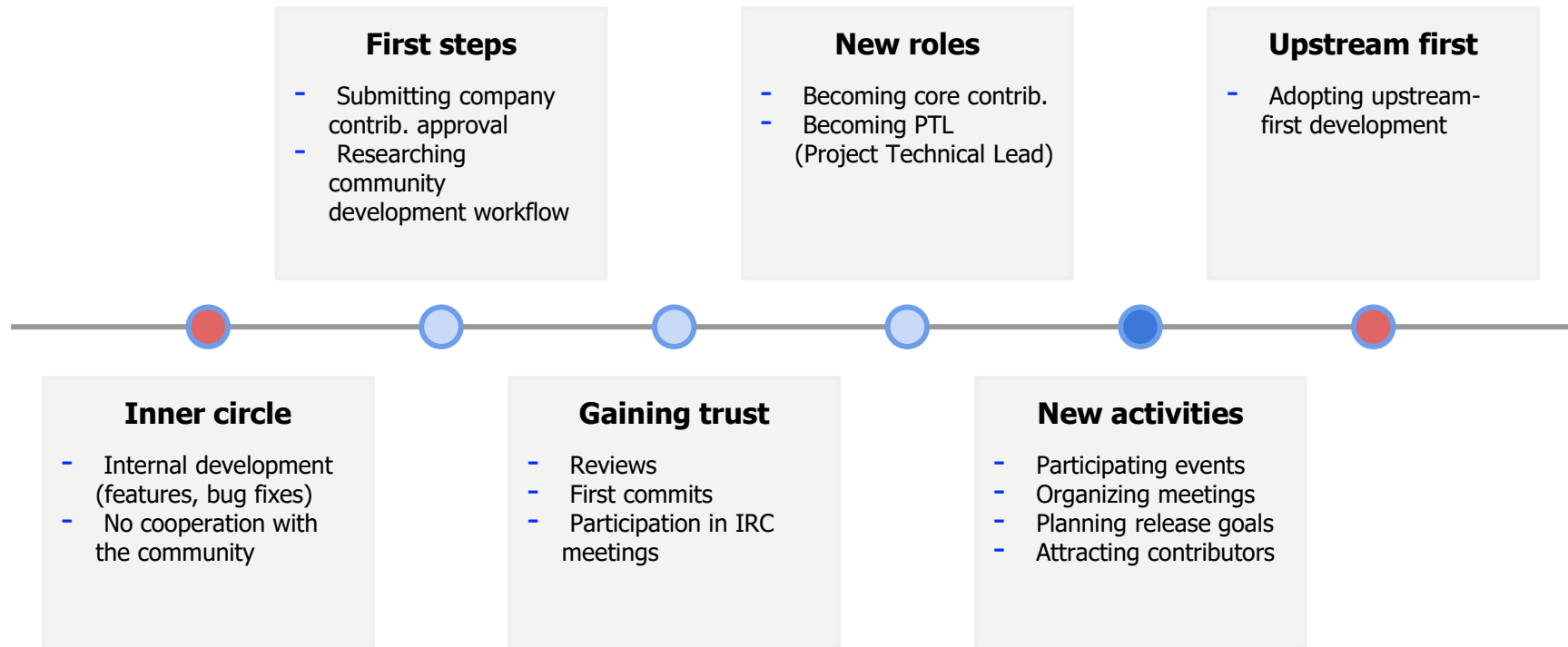
SOSCON 2018

- **1.** Database nodes report metrics to the monitoring service
- **2.** Monitoring service analyzes collected metrics and emits alerts based on static thresholds or ML predictions
- **3.** Alerts and resource entities consumed by RCA service allow to determine root cause of detected issues and produce action plan (e.g. scaling or healing)
- **4.** Workflow manager executes action plan by triggering APIs of different infrastructure services
- **5.** Infrastructure services, e.g. database service, actuate plan on database nodes



Towards open source contribution

SOSCON 2018



- Open source projects give us great opportunity to innovate
- Open source commitment brings benefits to the company
 - Reduces maintenance and update cost
 - Respect and outside reputation
 - High-quality feedback from 3rd-parties
 - Developers morale
- Open source is not a one-way street
 - It is our duty to contribute back
 - At least give feedback or report bugs
- Large projects need strong companies standing behind
 - They create the roadmap in the long run
 - They bring major contributors and users into the project

TROVE COMMUNITY

SOSCON 2018



Want to join one of the most prominent open source communities in the history?

Contact me:

Bartosz Żurkowski b.zurkowski@samsung.com

I will be glad to help You get started :)

Thank You

SOSCON 2018

SAMSUNG OPEN SOURCE CONFERENCE 2018

