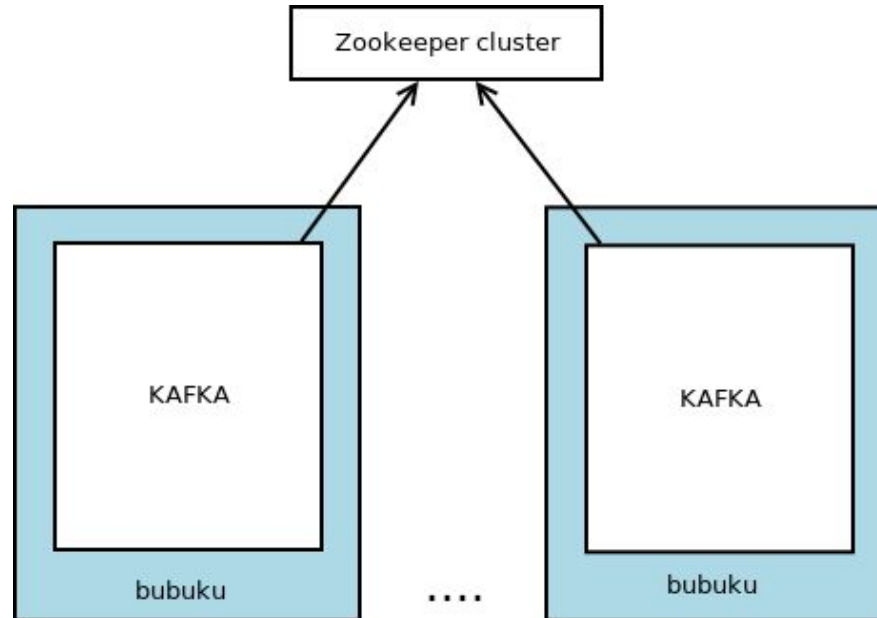


What is supervisor



From wikipedia:

blabla... controls the execution of other routines and regulates work scheduling, input/output operations, error actions, and similar functions and regulates the flow of work in a data processing system.

Do you need supervisor?

Yes:

- software not stable
- **software is running in frequently changed environment**
- **different instances have to synchronize their actions**

No:

- software is stable
- there are no frequent unpredictable restarts
- software have auto-recovery feature inside

How to create supervisor

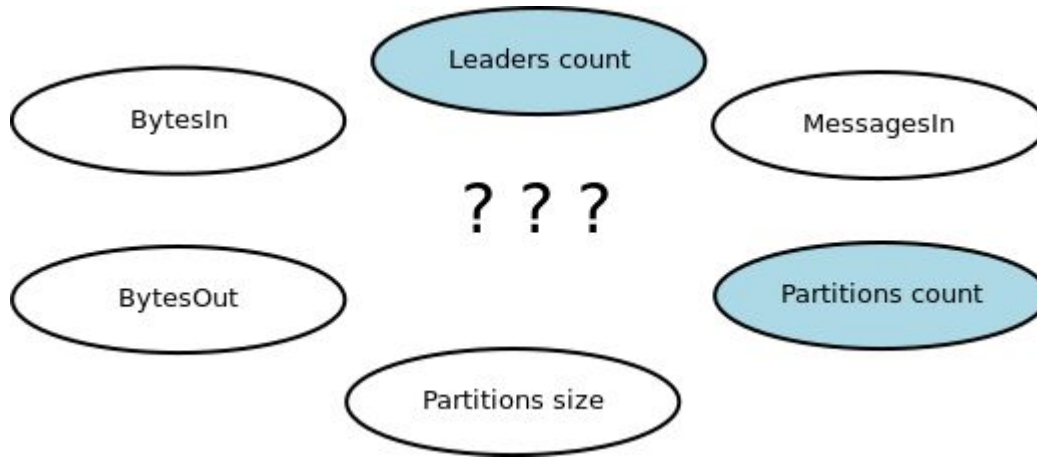
1. **Do not create supervisor, search for existing**
2. Select technology to work with
3. Think about memory
4. Think about system integration
5. Think about the way of communication with supervisor
6. If you are providing any logs - think what and how you are logging
7. If case of command line tool - **please**, think about stdout, stderr, stdin

Bubuku

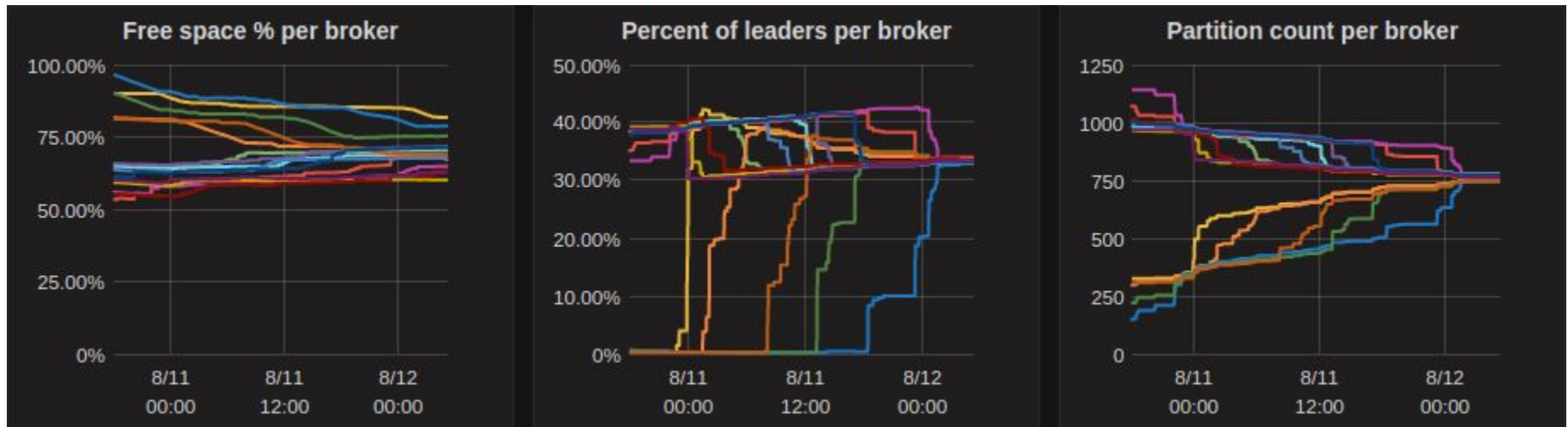
What can bubuku do?

- start/restart kafkas in a sequence
- restart kafka in case of malfunction
- provide simple to use api for managing cluster
 - balance partitions distribution
 - provide estimates on disk usage
 - optimize free space distribution across brokers

Bubuku - balancing partitions



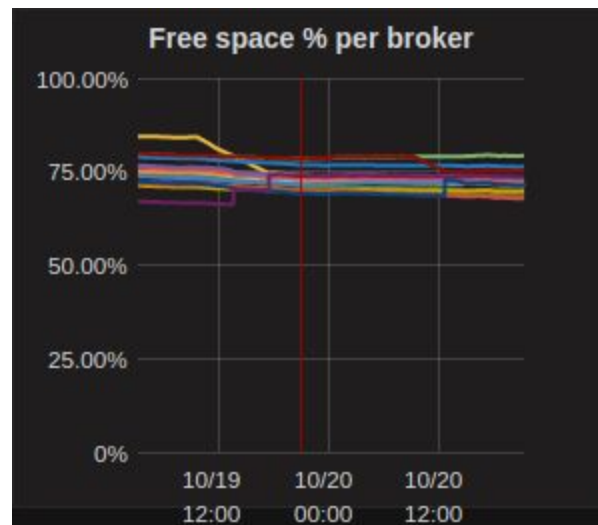
Bubuku - balancing partitions



- takes a lot of time
- free space is not evenly distributed
- overall cluster performance saturation during rebalance

Bubuku - swap fat slim

- codename: fatboy slim (Push The Tempo!)
- analyzes partitions on all brokers, taking into account rack awareness



Bubuku

<https://github.com/zalando-nakadi/bubuku>

dmitriy.sorokin@zalando.de

Questions?