

**adaptTo()**

APACHE SLING & FRIENDS TECH MEETUP  
BERLIN, 26-28 SEPTEMBER 2016

**AEM in Docker**  
**Andrei Darashenka, freelancer**

# Andrei Darashenka

DevOp und Software Architect from 1998

Docker from 2014

AEM from 2015

Freelancer from 2016

# Docker

# Docker benefits

- Software delivery
- Predictable environments
- Fast rollout
- Cheap “Virtualisation”
- Overlay filesystem

# Docker vocabulary

- Image like a filesystem snapshot
- Container like a running program
- Docker hub like an app store for docker images
- Registry like a private docker hub

# AEM in Docker

**BOOM**

# AEM is an antipattern for Docker

- Docker is designed for microservices
- AEM is a monolith
  
- Docker runs the best with stateless containers
- AEM has a big filesystem snapshot
  - 50 GB is a common



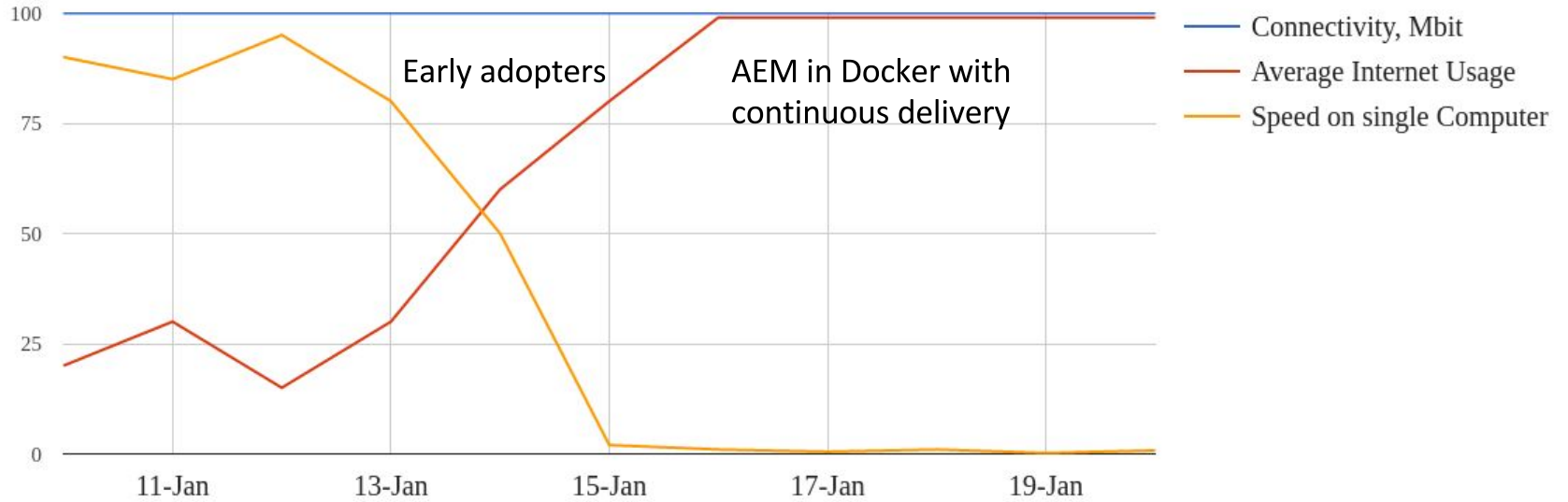
# You cannot drive fast



# Alternative without docker



## Internet usage in the office with more than 5 persons



# Good news

- Tarmk is append-only storage
- Mongomk is a remote
- Solrcloud
- Data Store on external filesystem or on S3

# Test vs. Live instances

- **Live instances: Many publishers, same dataset**
  - Data should be external to docker
  
- **Test instances: Many publishers, many datasets**
  - Every container should have his own dataset
  - We can build a docker image with initial dataset

# Building AEM images for test

- Dockerfile to create base, author, publisher
- Start author+publisher, install project, import products, rollout pages, replicate
- Stop author+publisher, compact tar, commit
- Image: 3-50 GB

# Building AEM images for test

- Feature-branch image
  - Build at PR creation time
  - Push to registry
- Build time 15-20 min, Image 50-200 MB
- 2-5 min to start

# Building AEM images for test

- Volumes
  - /tmp
  - /crx-quickstart/logs
  - /var/log/apache
  - /var/www/dispatcher-cache



# Building AEM images for live with tarmk

- **Use data container: backup etc...**
  - /crx-quickstart/repository
  - /crx-quickstart/launchpad
- **Use log container: send logs to ELK**
  - /crx-quickstart/logs
  - /var/log
- **/tmp**

# Starting new publisher with tarmk

1. Add this publisher to author with fake IP:port
2. Pause other publisher + commit data-container
3. Start new publisher from committed data-image
4. On author set correct IP:port of the publisher
5. Enable publisher in load-balancer

# Use AEM in cluster

- Use S3 with binaryless replication
- Put publisher/dispatcher in one pod
- Commit your containers just before/after rollout
- On problems simply recreate a publisher
- Hot-standby is important ;)
- Varnish 5.0 is released :)

# Appendix

# Andrei Darashenka

- `contact@darashenka.com`
- [www.xing.com/profile/Andrei\\_Darashenka](http://www.xing.com/profile/Andrei_Darashenka)
- [github.com/darashenka/aem-cmd](https://github.com/darashenka/aem-cmd)
- [github.com/darashenka/docker-registry-cache](https://github.com/darashenka/docker-registry-cache)

# Questions