

Enterprise Hiccups Introducing Cloud Foundry @ SWM

Stefan Mayr, Guido Wimmel

Cloud Foundry Meetup Munich 25.01.2017

Öffentlich

M/Wasser M/Bäder M/Strom M/Fernwärme M/Erdgas M/net



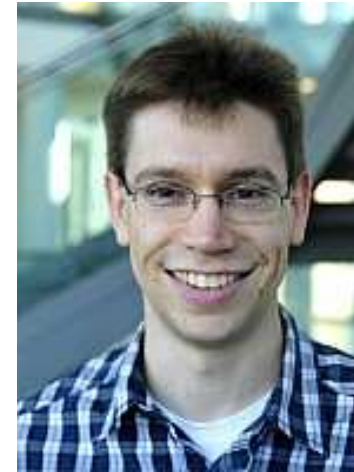
Who are we?

Dev



Guido Wimmel
IT Architect

Ops



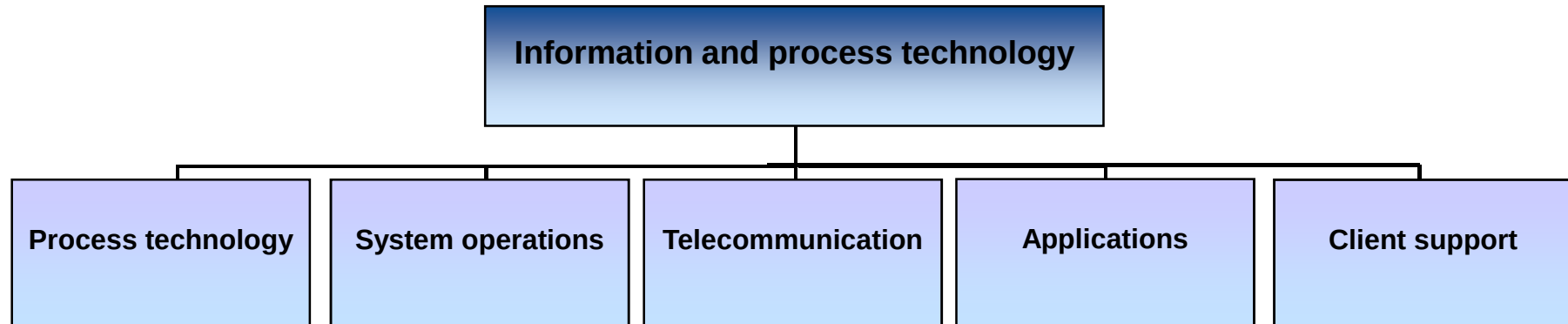
Stefan Mayr
sysadmin

Stadtwerke München



- ▶ are among the largest German utilities
- ▶ are one of the largest German municipal companies
 - ▶ total revenue 2015: 6.5 billion Euro
 - ▶ number of customers 2015: 1.2 million
- ▶ adhere to ecological principles
- ▶ are a driver of the energy revolution
- ▶ support Munich as one of the leading business centers in Europe
- ▶ provide public infrastructure for more than 130 years
 - ▶ water, electricity, gas, district heating
 - ▶ public transport
 - ▶ public baths
 - ▶ telecommunication

SWM IT department



SAP users

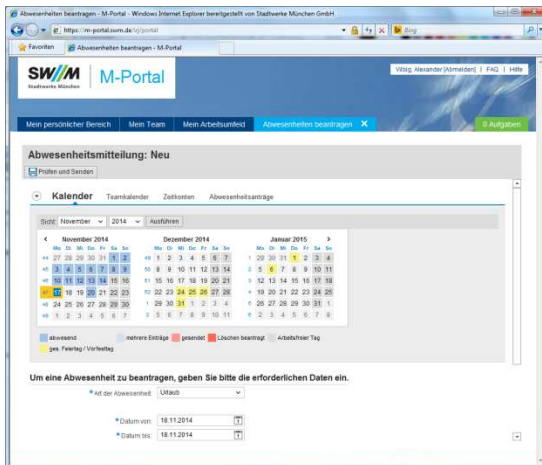
6.000

Business applications

>300

Customer satisfaction

1,3



Projects per year

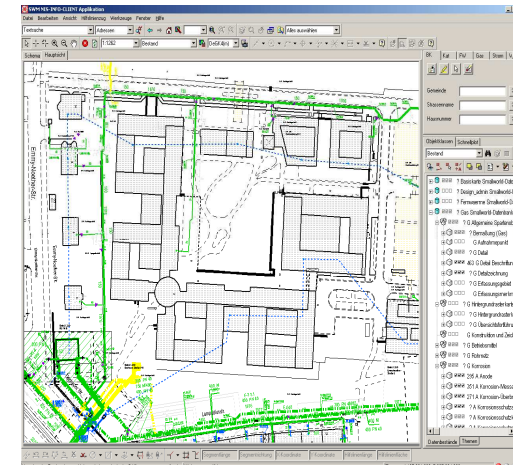
~100

Lines-of-code (Java)

9 m

Employees

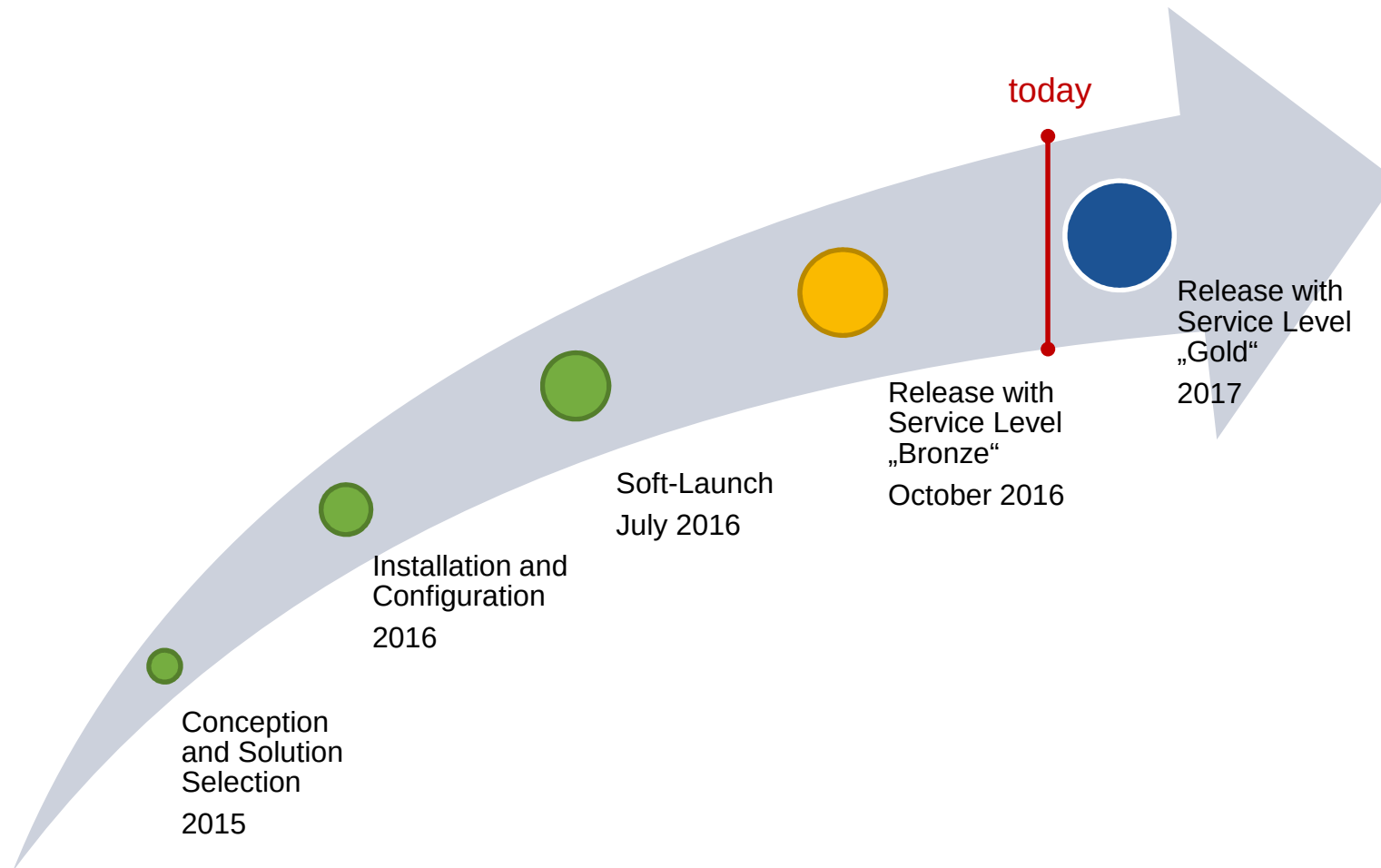
~400



SWM In-house development – main technologies

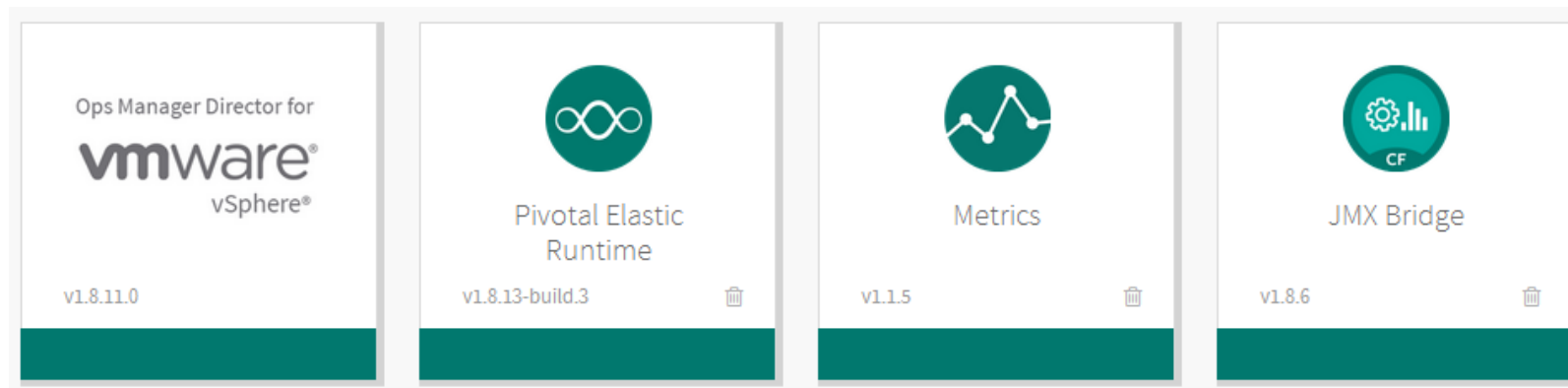


Timeline for Cloud Foundry @ SWM



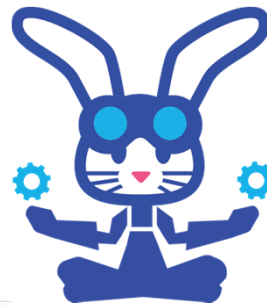
What we use

- ▶ Pivotal Cloud Foundry (1.8) on vSphere
- ▶ JMX Bridge as monitoring api
- ▶ PCF Metrics for application logs and metrics



- ▶ No data services, yet

Why Pivotal Cloud Foundry?

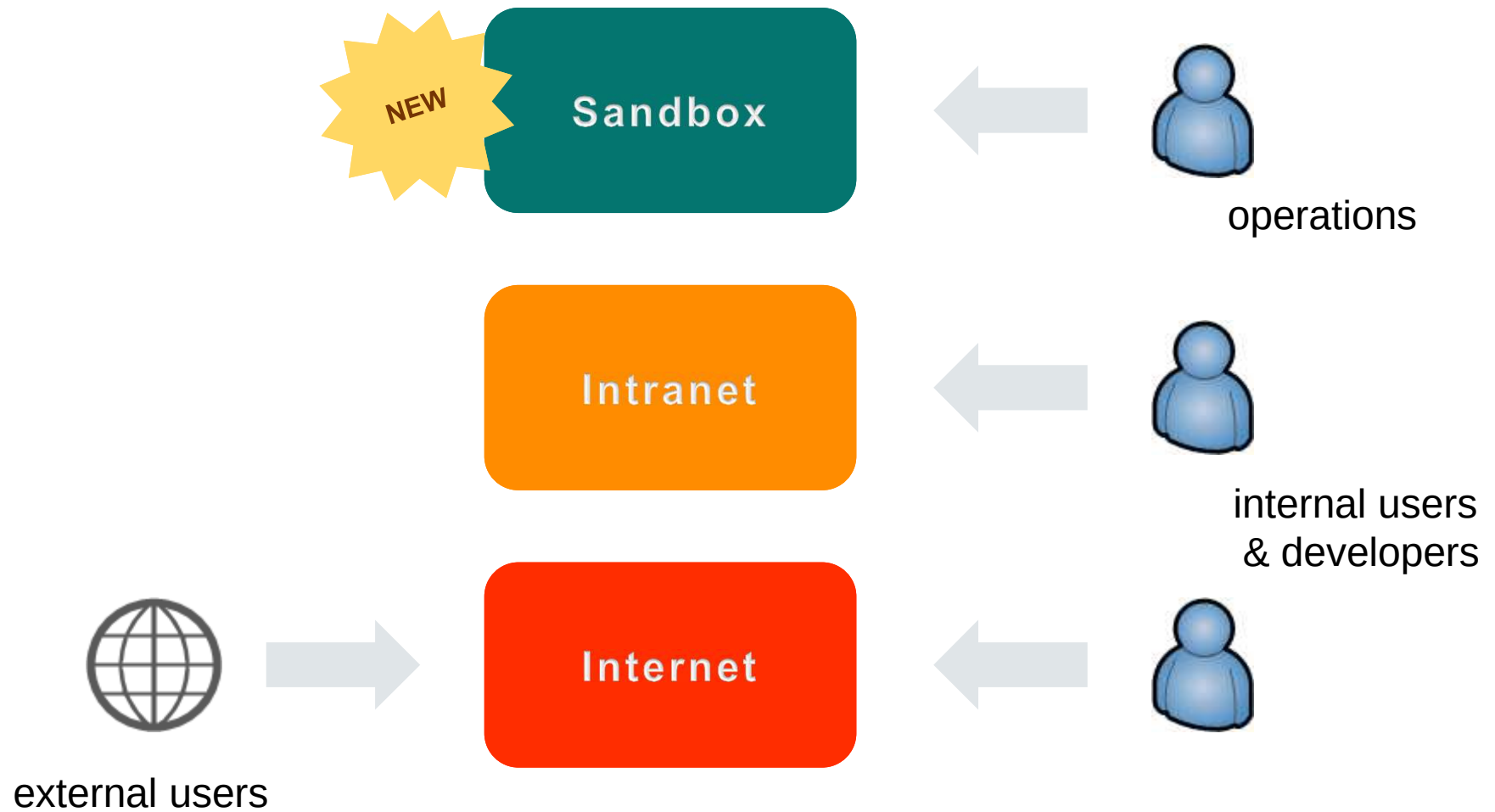


vs



- ▶ Reasons for PCF over Cloud Foundry OSS
 - ▶ Stable versions
 - ▶ GUIs
 - ▶ Ops Manager in front of BOSH (hides 1000s lines of yaml)
 - ▶ Apps Manager
 - ▶ Support

Infrastructure



Infrastructure

- ▶ Physical hardware
 - ▶ 3 Hosts with 128GB RAM
 - ▶ 4TB storage
 - ▶ Loadbalancer



<http://www.fujitsu.com/us/products/computing/servers/primergy/gallery/index.html>

- ▶ virtual machines
 - ▶ ~ 40 CF VMs
 - ▶ ~ 5 SWM „helper“ VMs
- ▶ Running ~ 50 app instances

pcf_vms					
Virtuelle Maschinen					
Aufgaben & Ereignisse Alarmer Berechtigungen Zuordnungen					
Name	Zustand	Status	Bereitgestellter S.	Verwendeter S.	
 vm-a5799635-1102-463c-bdfc...	Eingeschaltet	✓ Normal	122,16 GB	49,06 GB	
 vm-3daa618a-e5ef-42b1-8021...	Eingeschaltet	✓ Normal	150,16 GB	147,30 GB	
 vm-03b8ea92-8e7a-4cbe-98e6...	Eingeschaltet	✓ Normal	150,16 GB	147,32 GB	
 vm-fd68ff67-7c63-44a8-b115...	Eingeschaltet	✓ Normal	118,16 GB	24,57 GB	
 vm-ed15c74f-3e90-4d0d-8d85...	Eingeschaltet	✓ Normal	150,16 GB	147,34 GB	
 vm-40375c08-55d9-4e57-bf5d...	Eingeschaltet	✓ Normal	19,16 GB	15,34 GB	
 vm-4d2445e5-17cc-406d-814a...	Eingeschaltet	✓ Normal	150,16 GB	147,32 GB	
 vm-bb3062cb-c4c6-4d2a-9bda...	Eingeschaltet	✓ Normal	17,16 GB	13,36 GB	
 vm-c340d3eb-6328-4bb7-98cd...	Eingeschaltet	✓ Normal	15,16 GB	12,30 GB	
 vm-da57e30d-57fa-445e-90ae...	Eingeschaltet	✓ Normal	15,16 GB	12,30 GB	

Infrastructure

- ▶ Upgrading from PCF 1.6 -> 1.7 -> 1.8

- ▶ Needs more RAM
- ▶ Needs more storage



- ▶ First load tests

- ▶ gorouter/loggregator can only do ~1000 requests/sec
 - ▶ => scale # of cpus or # job instances
- ▶ Don't trust PCF Metrics request rate

- ▶ Strange edge case for cpu shares

- ▶ # app instances == # cells => maximum performance (100%)
- ▶ # app instances == # cells +1 => 50% performance

Authentication and Authorization



Creator: Dori, unmodified work; license: CC BY-SA 3.0 US (<https://creativecommons.org/licenses/by-sa/3.0/us/deed.en>)

URL: https://commons.wikimedia.org/wiki/File:Authorized_Personnel_Only_sign_5642.jpg

Authentication and Authorization

- ▶ Our IDM system manages users and groups in a MS Active Directory
- ▶ We connected CF UAA via LDAP
- ▶ Challenges:
 1. how to enter multiple LDAP servers for redundancy?
 2. Read-Only Domain Controller (RODC): referrals + firewall = timeout
 3. you cannot persistently override the UAA config generated by Ops Manager

Authentication and Authorization

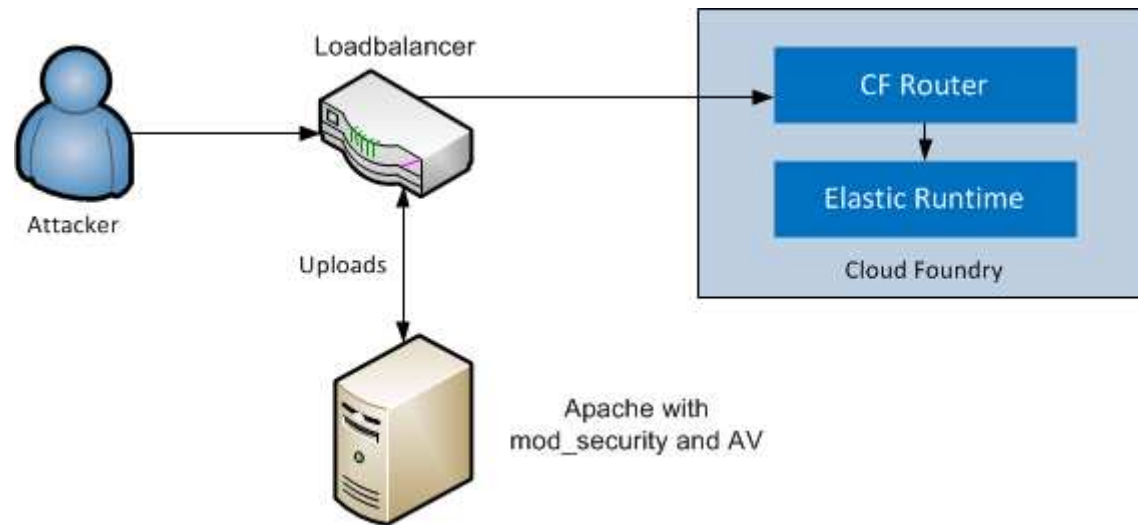
- ▶ Implemented solution:
 - ▶ We use an openLDAP instance as proxy to filter referrals
 - ▶ We have 2 of them and use keepalived with a cluster-IP
- ▶ Ops Manager now uses the cluster-IP as LDAP server
 - ▶ No more firewall problems with referrals
 - ▶ Failover of keepalived is good enough for our HA requirements
 - ▶ Generated UAA config can be used as-is

Authentication and Authorization

- ▶ Using LDAP as source for UAA users still has its quirks
 - ▶ Users have to login once before they can be assigned any role
 - ▶ Org managers have to assign new users a role in a space. They cannot assign users to an org without space roles (at least via cf CLI, possible through direct API calls)
 - ▶ Space managers can only assign space roles to users that are member of the org
 - ▶ LDAP groups can be used to assign admin privileges
 - ▶ LDAP groups cannot be used to assign org or space roles
- ▶ Possible solutions:
 - ▶ More scripting foo
 - ▶ UAA Bulk Import Tool

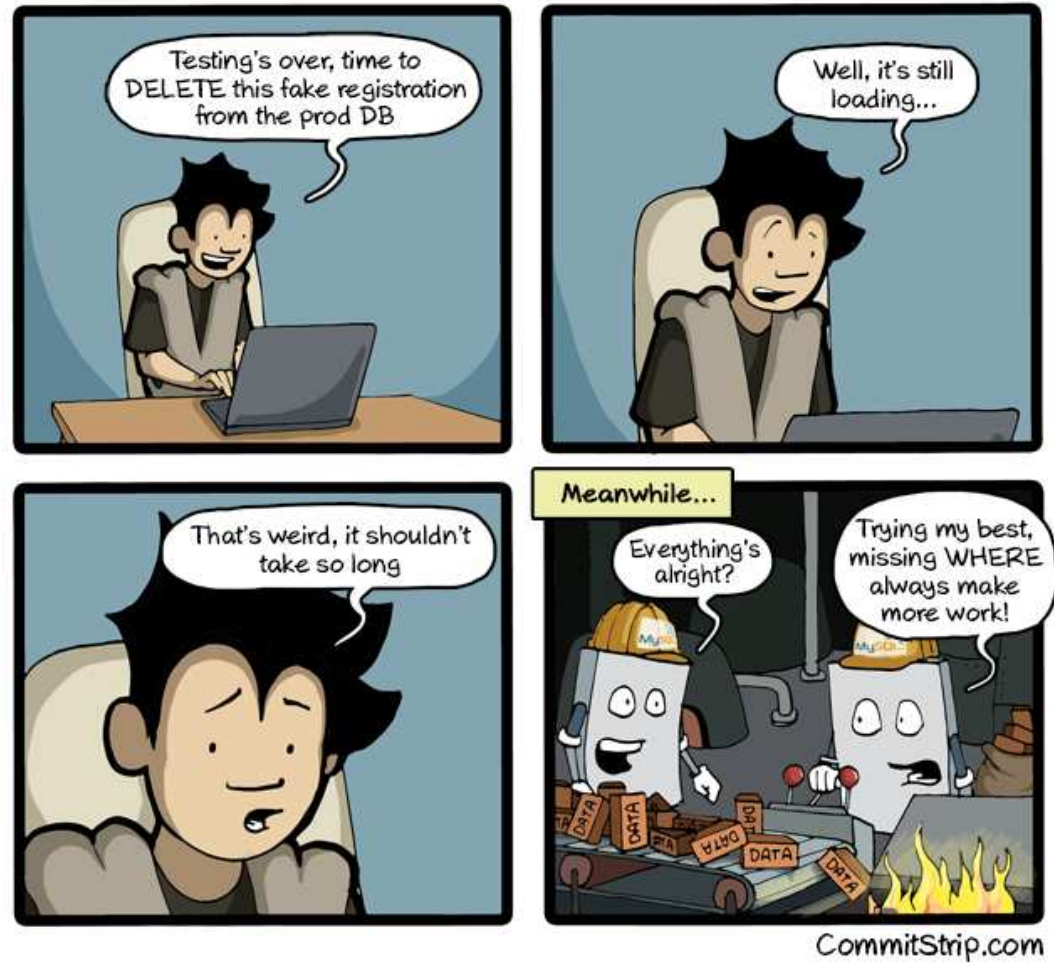
Antivirus

- ▶ Having a antivirus solution on every server is a company requirement
- ▶ We didn't find a solution that provided a bosh deployment
- ▶ Workaround: filter malicious uploads before they reach Cloud Foundry



- ▶ Remark: it works like a route service, but it is not optional
- ▶ Alternatives: different bosh releases for ClamAV exist (by 18F, Pivotal, ...)

Backup



Backup

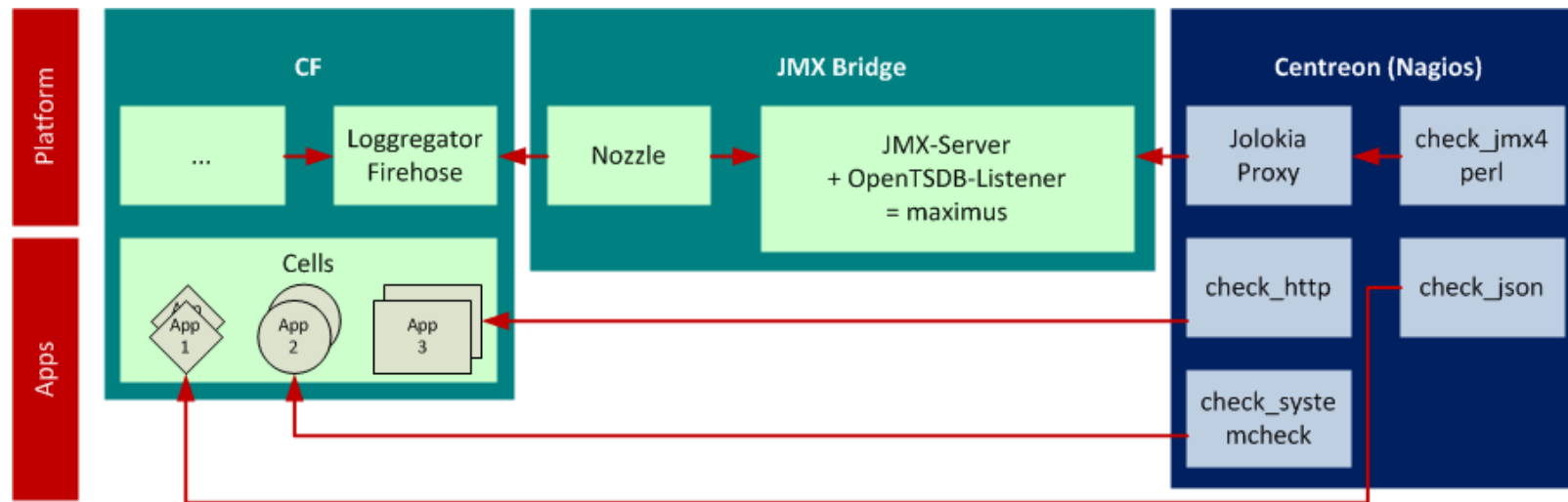
- ▶ Evaluation of PCF in 2015
 - ▶ only a manual backup was supported
 - ▶ There was no documented restore procedure – you had to call support
- ▶ Today we use the cfops tool for daily backups
 - ▶ It can only backup Pivotal Ops Manager and Elastic Runtime
 - ▶ It misses older stemcells (only the most current one is included)
- ▶ Some Pivotal AddOns
 - ▶ only have cloud backup options (AWS S3)
 - ▶ don't even have a documentation how to backup or restore data

Restore

- ▶ cfops restore works, in general
- ▶ Ops Manager restore ...
 - ▶ works only in some of our environments
 - ▶ Using curl according to Ops Manager API documentation is no valid parameter combination
 - ▶ Does not include all required stemcells

Monitoring

- ▶ We use Centreon (Nagios)



- ▶ applications running on CF
 - ▶ Website check (HTTP status code / content string)
 - ▶ JSON check (e.g. Spring Boot health endpoints)
 - ▶ SWM Systemcheck (XML endpoint with a list of all subsystems)

Monitoring

- ▶ Because the CF setup is more dynamic than our monitoring config we group multiple beans for different instances

Hosts ▾	Services	Status	Duration	Last Check	Tries	Status information
CF	Router Bad Gateways Delta	 	1M 18h 24m 6s	12/01/2017 15:23:55	1/3 (H)	OK - VM_e52e363b-66ed-4ab0-9ab8-886edd16e6a6: OK - 0 bad gateways, VM_ee41f117-fe9e-416c-9248-b0e4225945b7: OK - 0 bad gateways
	Router Latency	 	1d 4h 30m 31s	12/01/2017 15:26:43	1/3 (H)	OK - VM_e52e363b-66ed-4ab0-9ab8-886edd16e6a6: OK - 9 Latency, VM_ee41f117-fe9e-416c-9248-b0e4225945b7: OK - 10 Latency
	Router Latency CloudController	 	6h 5m 38s	12/01/2017 15:26:35	1/3 (H)	OK - VM_e52e363b-66ed-4ab0-9ab8-886edd16e6a6: OK - 6 Latency, VM_ee41f117-fe9e-416c-9248-b0e4225945b7: OK - 8 Latency
	Router Requests	 	1M 3d 9h 47m 51s	12/01/2017 15:23:27	1/3 (H)	OK - VM_e52e363b-66ed-4ab0-9ab8-886edd16e6a6: OK - 572997 Total requests, VM_ee41f117-fe9e-416c-9248-b0e4225945b7: OK - 572980 Total requests
	Router Requests per second	 	1M 1w 6d 17h 57m 18s	12/01/2017 15:23:11	1/3 (H)	OK - VM_e52e363b-66ed-4ab0-9ab8-886edd16e6a6: OK - 0.126246 requests per second, VM_ee41f117-fe9e-416c-9248-b0e4225945b7: OK - 0.120000 requests per second
	Router Responses	 	2M 2d 2h 19m 38s	12/01/2017 15:23:07	1/3 (H)	OK - VM_e52e363b-66ed-4ab0-9ab8-886edd16e6a6: OK - 568812 responses, VM_ee41f117-fe9e-416c-9248-b0e4225945b7: OK - 568841 responses
	VM health router	 	1M 3d 9h 47m 35s	12/01/2017 15:23:35	1/3 (H)	OK - VM_0: OK - Cellstatus=1, VM_1: OK - Cellstatus=1

OK - VM_0: OK - Cellstatus=1, VM_1: OK - Cellstatus=1

- ▶ Things we need to work on
 - ▶ How to detect problems on specific app instances
 - ▶ Automatic monitoring configuration
 - ▶ Which CF metrics are valuable?
 - ▶ Thresholds for alerts

Logging

- ▶ We use roles in CF spaces
- ▶ We also need to check permissions on logs
- ▶ Conclusion: no central syslog or elasticsearch cluster with kibana
- ▶ Requirement from devs: Multiline logs (Java stacktraces)
- ▶ Temporary solution: graylog with GELF and streams
 - ▶ One graylog stream per space
 - ▶ Apps use Logback with GELF appender (graylog extended log format)
 - ▶ Each log message contains the space_id
 - ▶ LDAP group is mapped to reader role of a stream

Logging

- ▶ Drawbacks
 - ▶ Cloud Foundry
 - ▶ Missing stdout (needs custom app logging)
 - ▶ Missing CF events
 - ▶ Graylog
 - ▶ Search is restricted to one stream
 - ▶ Only one elasticsearch index – only one retention policy
- ▶ Side note: data protection
 - ▶ In the EU/Germany logging or storing IP addresses may be forbidden
 - ▶ The router and some buildpacks log IPs.
 - ▶ Low awareness outside Germany

Customizing Java buildpack

- ▶ Inhouse development uses Oracle Java
- ▶ We have an internal public key infrastructure (PKI) – own CA certs
- ▶ Internet can only be accessed through a proxy
- ▶ Solution approach #1
 - ▶ git clone <https://github.com/cloudfoundry/java-buildpack>
 - ▶ Read docs for customizing; provide Oracle Server-JRE on an internal server
 - ▶ bundle install && bundle exec rake package OFFLINE=true
- ▶ Difficulties
 - ▶ ruby and gem versions on enterprise linux distro
 - ▶ ruby does not respect no_proxy (<https://bugs.ruby-lang.org/issues/11195>)

Customizing Java buildpack

- ▶ Solution approach #2
 - ▶ Understand the buildpack structure
 - ▶ Hack the offline buildpack manually
 - ▶ Put the customizing into a shell script

```
# ./swm-customize-java-buildpack.sh
Als Parameter muss entweder ein Offline-Buildpack oder das Verzeichnis eines Buildpacks angegeben werden

Verwendung: ./swm-customize-java-buildpack.sh [ -b SRCBUILDPACK ] [ -d SRCBUILDPACKDIR ] [ -j JAVAVERSION ]
[ -c CACERTS ] [ -v VERSION ] [ -h ]

-b SRCBUILDPACK      Angabe eines Quellbuildpacks (ZIP-Datei) das als Vorlage dient
-d SRCBUILDPACKDIR   Angabe eines Quellverzeichnis mit einem entpackten Buildback
-j JAVAVERSION       Angabe der zu verwenden Java-Version, Default: 1.8.0_111
-c CACERTS           Angabe der ZIP-Datei mit CA-Zertifikaten, Default: SWM-CA-certs-20161124.zip
-v VERSION           SWM Versionsnummer des erzeugten Buildpacks, Default: 1.0
-h                  Anzeige dieser Hilfe

Es muss ein Quellbuildpack mit -b oder -d angegeben werden. Eine Datei (-b) wird gegenüber einem Verzeichnis (-d) bevorzugt
```

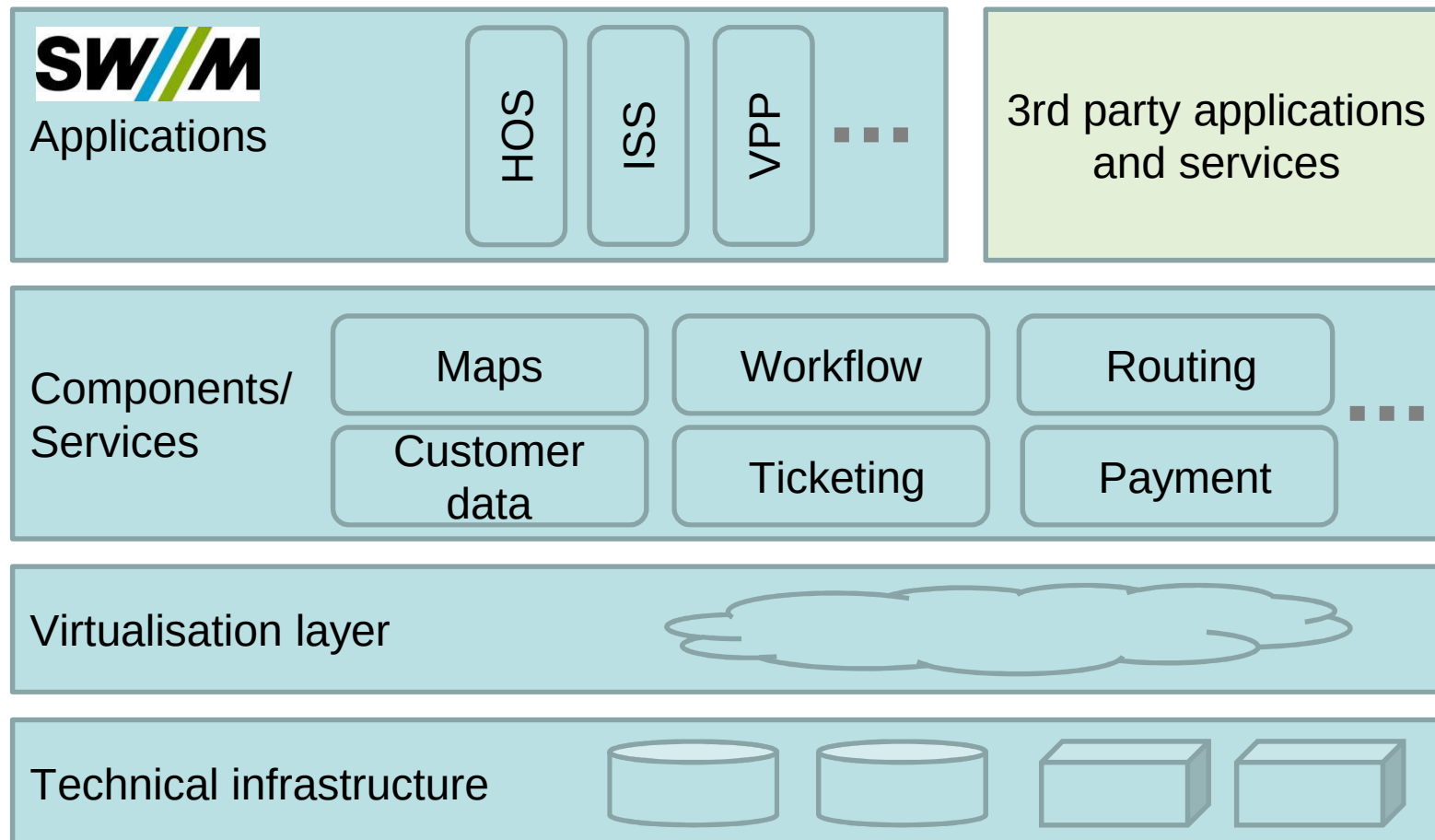
Customizing Java buildpack

- ▶ Details
 - ▶ Toggle required flags using sed against yml files
 - ▶ Add a fake download URL to config/oracle_jre.yml
 - ▶ URL-encode fake URL and create resources/cache/\$fakeurl/index.yml.cached
 - ▶ Copy Oracle Server-JRE to resources/cache/\$fakeurl/\$serverjre.cached
 - ▶ Extract Oracle cacerts to resources/oracle_jre/jre/lib/security
 - ▶ Import SWM CA certs into cacerts file

Component-Based Development



Conceptual Architecture



HOS: Handy Online Ticket System, ISS: Internet Self Services, VPP: Virtual Power Plant

Thank you

