Large Scale Linux Management

Stefan Dietrich, on behalf of IT-Systems
Hamburg, 2025-08-14

HELMHOLTZ

DESY.

Linux@DESY

Overview

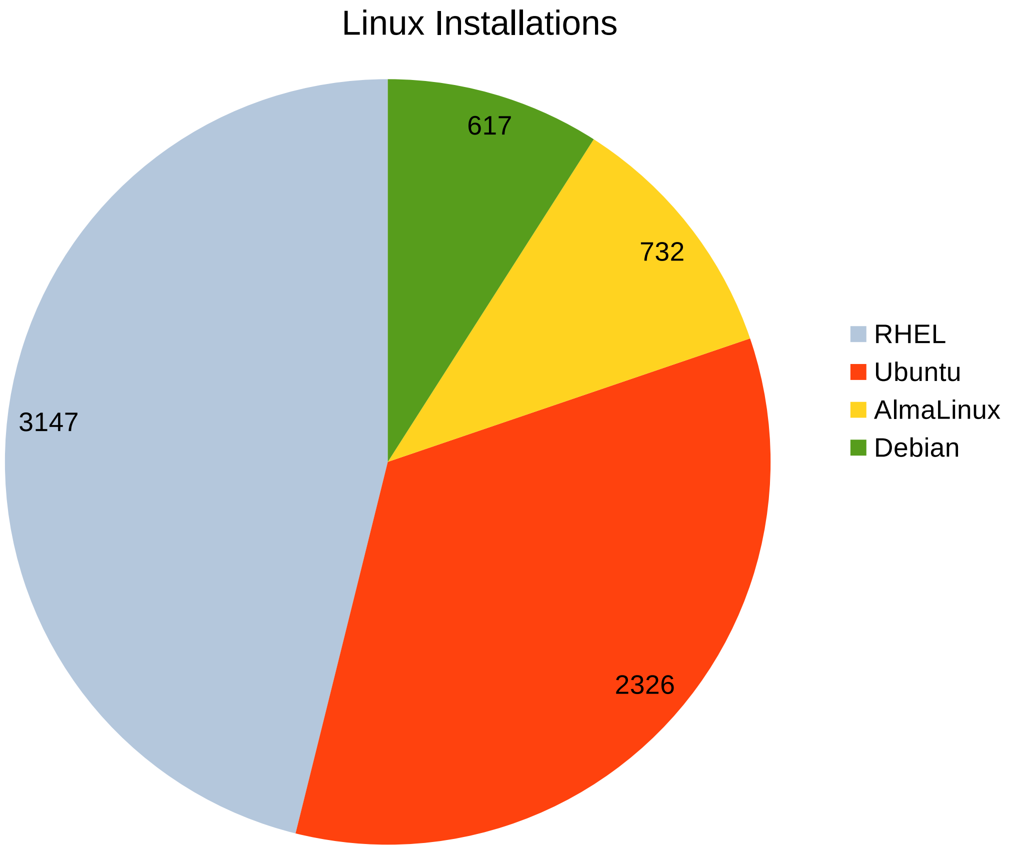
- IT provides support for several Linux distributions
  - Red Hat Enterprise Linux
  - AlmaLinux
  - Ubuntu LTS
  - Debian (limited support)
- RHEL: 5 year site-license with unlimited installations
- Installations available for all groups on the campus & Eu.XFEL
  -> lot's of users & machines

- Core Component for Compute- & Storage Systems
  - IDAF: Maxwell, GRID & NAF Compute Clusters
  - Storage: dCache
- Experiment & Machine Control
  - Debian for experiment control
  - Ubuntu for machine control
- Desktops: Ubuntu
  - Traffic Light Model:
    - Green: full support, no sudo, common tasks doable
    - Yellow: sudo, no IT-support anymore
- Many other services rely on Linux as well

Linux@DESY

Numbers

- Overall ~6800 active Linux installations
  - >= 50% RHEL & AlmaLinux
  - ~825 desktops
  - ~900 dcache nodes
  - ~1700 compute nodes
  - ~1400 virtual machines
  - ~150 ARM64
    (Raspberry Pi, 3x Ampere Altra servers)
  - ~35 POWER systems



Installation & Configuration Management

Automation & Delegation

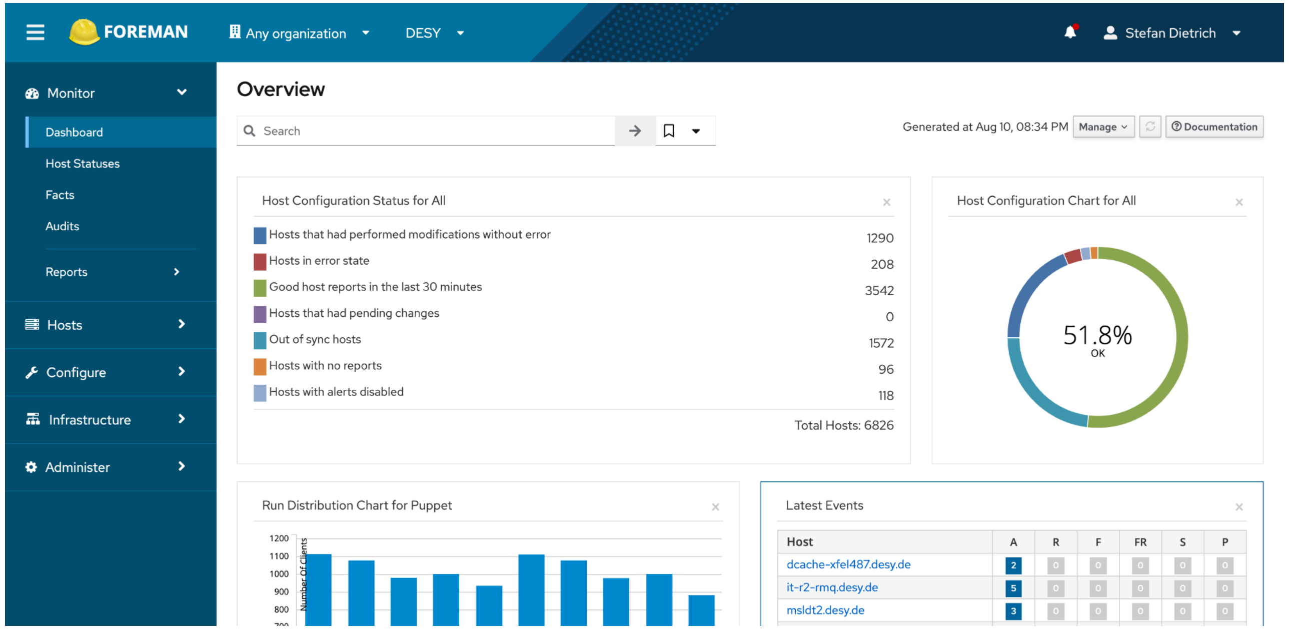
- Automating installation and configuration is mandatory at scale
  - Manual administration no longer scaling
  - Repeatable and consistent configuration
  - Avoid „balkanization“ of configurations
  - Increase security with sane, central defaults
- Delegation important factor for us
  - delegation of boring tasks to group administrators, e.g. registering new machines
- Move from pets to cattle system administration

- Current tools:
  - Foreman (https://theforeman.org/)
  - Puppet/OpenVox (https://voxpupuli.org/openvox/)
- Long-term effort
  - Started with Puppet in ~2012
  - Consolidated previous **3** config management systems to Puppet/Foreman

Foreman

OS Installation

- Node registration & hostgroup (roles) assignment
- Manages PXE configuration files for automated OS installation
  - Kickstart, Subiquity & Debian-installer
  - Deploying on bare-metal and virtual machines
- Reporting & inventory capability
- Access via web interface, CLI or REST API
- Delegation capability
  - Allow group administratos to manage their own machines



Overview

Generated at Aug 10, 08:34 PM

Host Configuration Status for All

🟦 Hosts that had performed modifications without error	1290
🟥 Hosts in error state	208
🟩 Good host reports in the last 30 minutes	3542
🟪 Hosts that had pending changes	0
🟦 Out of sync hosts	1572
🟧 Hosts with no reports	96
🟦 Hosts with alerts disabled	118

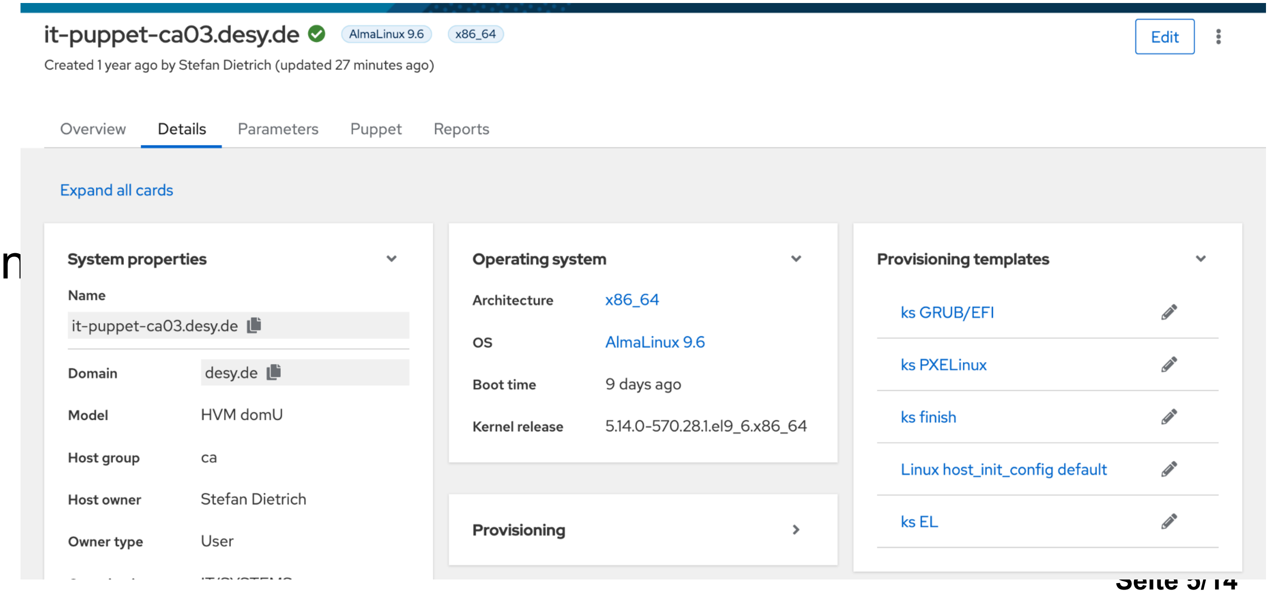
Total Hosts: 6826

Host Configuration Chart for All

Run Distribution Chart for Puppet

Latest Events

Host	A	R	F	FR	S	P
dcache-xfel487.desy.de	2	0	0	0	0	0
it-r2-rmq.desy.de	5	0	0	0	0	0
msldt2.desy.de	3	0	0	0	0	0



it-puppet-ca03.desy.de AlmaLinux 9.6 x86_64

Created 1 year ago by Stefan Dietrich (updated 27 minutes ago)

System properties

Name	it-puppet-ca03.desy.de
Domain	desy.de
Model	HVM domU
Host group	ca
Host owner	Stefan Dietrich
Owner type	User

Operating system

Architecture	x86_64
OS	AlmaLinux 9.6
Boot time	9 days ago
Kernel release	5.14.0-570.28.1.el9_6.x86_64

Provisioning

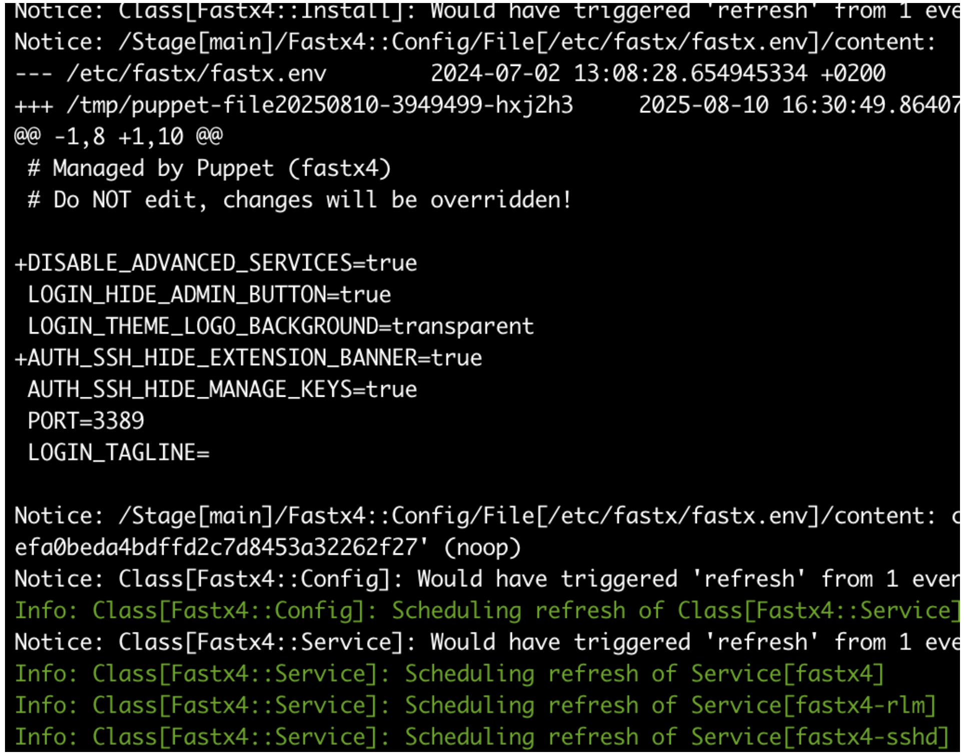
Provisioning templates

- ks GRUB/EFI
- ks PXELinux
- ks finish
- Linux host_init_config default
- ks EL

Puppet

Configuration Management

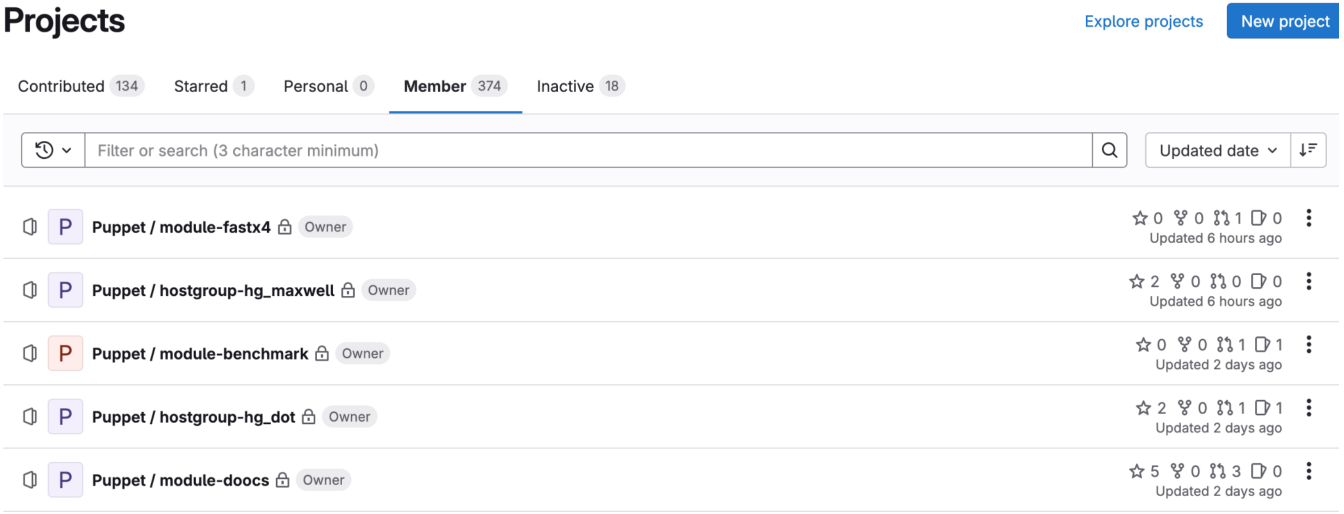
- Pull model, big community, lots of 3rdparty modules
- DSL to describe desired state, enforced by Puppet
- Setup based on Open Source Puppet 7
  - Migration to OpenVox 8 (fork) this year
- Dedicated Gitlab instance for code management
  - 374 git repositories
  - CI/CD pipeline to test for unexpected changes
- Delegation capability
  - Allow administrators to manage their own configurations
  - Core modules, like SSH:
    → merge request & review from IT



```
Notice: Class[Fastx4::Install]: Would have triggered 'refresh' from 1 eve
Notice: /Stage[main]/Fastx4::Config/File[/etc/fastx/fastx.env]/content:
--- /etc/fastx/fastx.env        2024-07-02 13:08:28.654945334 +0200
+++ /tmp/puppet-file20250810-3949499-hxj2h3     2025-08-10 16:30:49.86407
@@ -1,8 +1,10 @@
 # Managed by Puppet (fastx4)
 # Do NOT edit, changes will be overridden!

+DISABLE_ADVANCED_SERVICES=true
 LOGIN_HIDE_ADMIN_BUTTON=true
 LOGIN_THEME_LOGO_BACKGROUND=transparent
+AUTH_SSH_HIDE_EXTENSION_BANNER=true
 AUTH_SSH_HIDE_MANAGE_KEYS=true
 PORT=3389
 LOGIN_TAGLINE=

Notice: /Stage[main]/Fastx4::Config/File[/etc/fastx/fastx.env]/content: c
efa0beda4bdffd2c7d8453a32262f27' (noop)
Notice: Class[Fastx4::Config]: Would have triggered 'refresh' from 1 ever
Info: Class[Fastx4::Config]: Scheduling refresh of Class[Fastx4::Service]
Notice: Class[Fastx4::Service]: Would have triggered 'refresh' from 1 eve
Info: Class[Fastx4::Service]: Scheduling refresh of Service[fastx4]
Info: Class[Fastx4::Service]: Scheduling refresh of Service[fastx4-rlm]
Info: Class[Fastx4::Service]: Scheduling refresh of Service[fastx4-sshd]
```



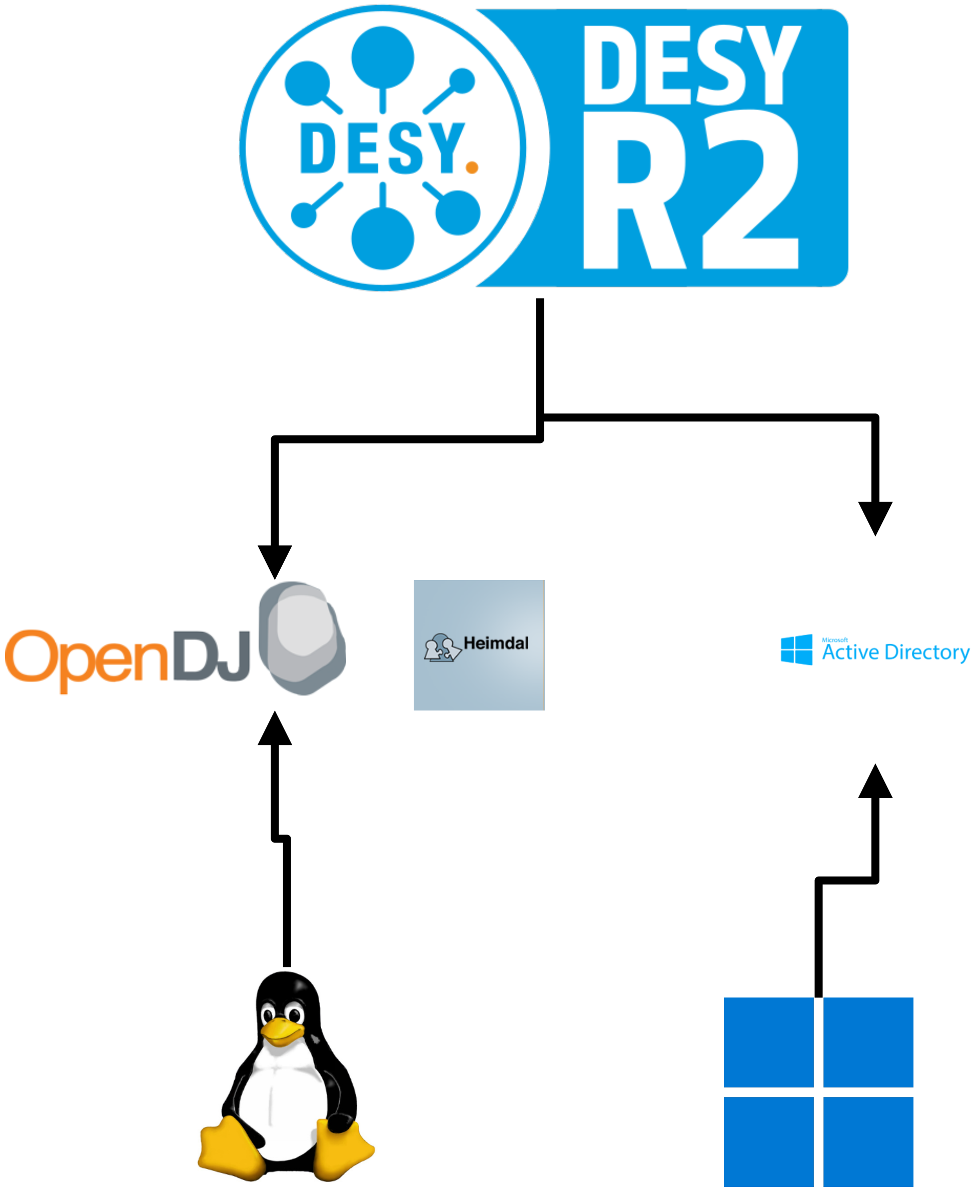
Projects

Puppet / module-fastx4	Owner	☆ 0	0	1	0	Updated 6 hours ago
Puppet / hostgroup-hg_maxwell	Owner	☆ 2	0	0	0	Updated 6 hours ago
Puppet / module-benchmark	Owner	☆ 0	0	1	1	Updated 2 days ago
Puppet / hostgroup-hg_dot	Owner	☆ 2	0	1	1	Updated 2 days ago
Puppet / module-doocs	Owner	☆ 5	0	3	0	Updated 2 days ago

Authentication & Authorization

LDAP & Kerberos

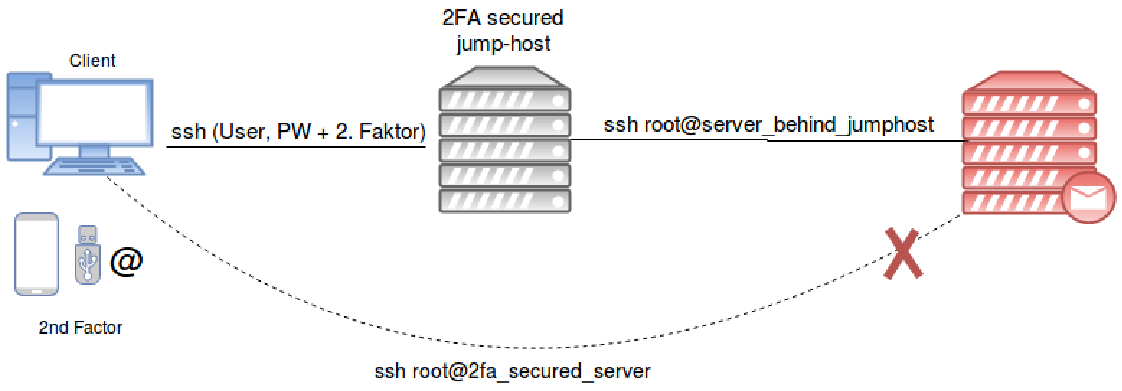
- No local user accounts, except for system users
- Central directory service
  - OpenDJ LDAP Server, holds user information
  - Heimdal Kerberos Server: User authentication
- Central user management tool:
  DESY Registry
  - Maintained by IT Information Fabrics
- DESY Registry feeds information to LDAP, Kerberos and Microsoft Active Directory
- SSSD on Linux for LDAP connections and password authentication via Kerberos
- Also: root logins via Kerberos (.k5login)

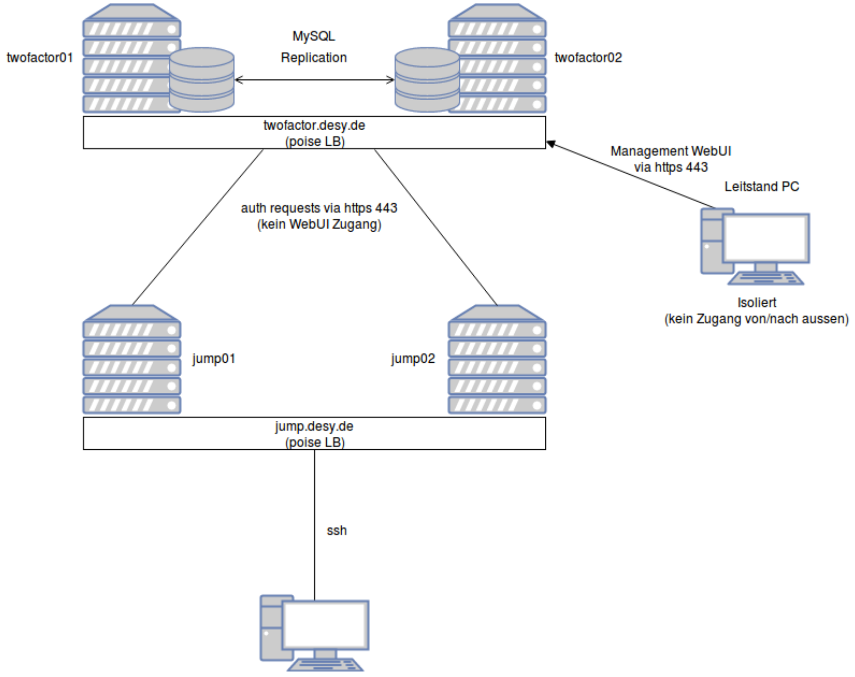


Securing root Logins

Jump Host

- Best practice: Require multi-factor authentication for administrative tasks
  - We use kerberos for root logins (.k5login)
  - Loss of password could result into root access
- Implemented “Jump Hosts”
  - Based on PrivacyIdea for MFA & plain OpenSSH
- Login flow
  - Login to Jump Host as normal user
  - Requires password **and** 2nd factor
  - All servers reconfigured to allow only root logins from Jump Hosts
- Avoids rollout of MFA directly on servers


2FA secured
jump-host
Client
ssh (User, PW + 2. Faktor)
ssh root@server_behind_jumphost
@
2nd Factor
ssh root@2fa_secured_server

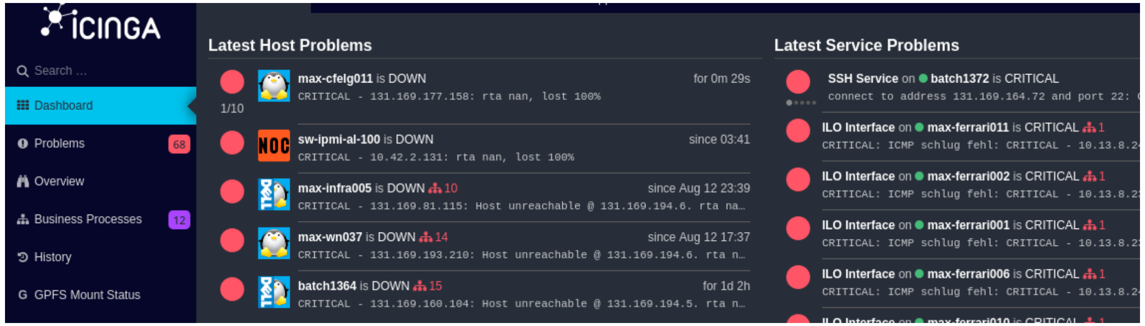


Secret Management

- 3 secret management tools in use
  - Different use-cases & historic reasons
- Host Key Distribution (HKD)
  - Kerberos keytab distribution
  - SSH known host management
  - Generic file secret management
- Keystore (GPG-based)
  - Secure storage of passwords and access from multiple users/groups
- HashiCorp Vault
  - Generic secret solution, better integration with Puppet

- Vault to supersede Keystore & (most parts of) HKD
  - e.g. rotate root passwords more often
- WiP: Distribution of X.509 certificates via Vault

Service Monitoring & Log Analysis



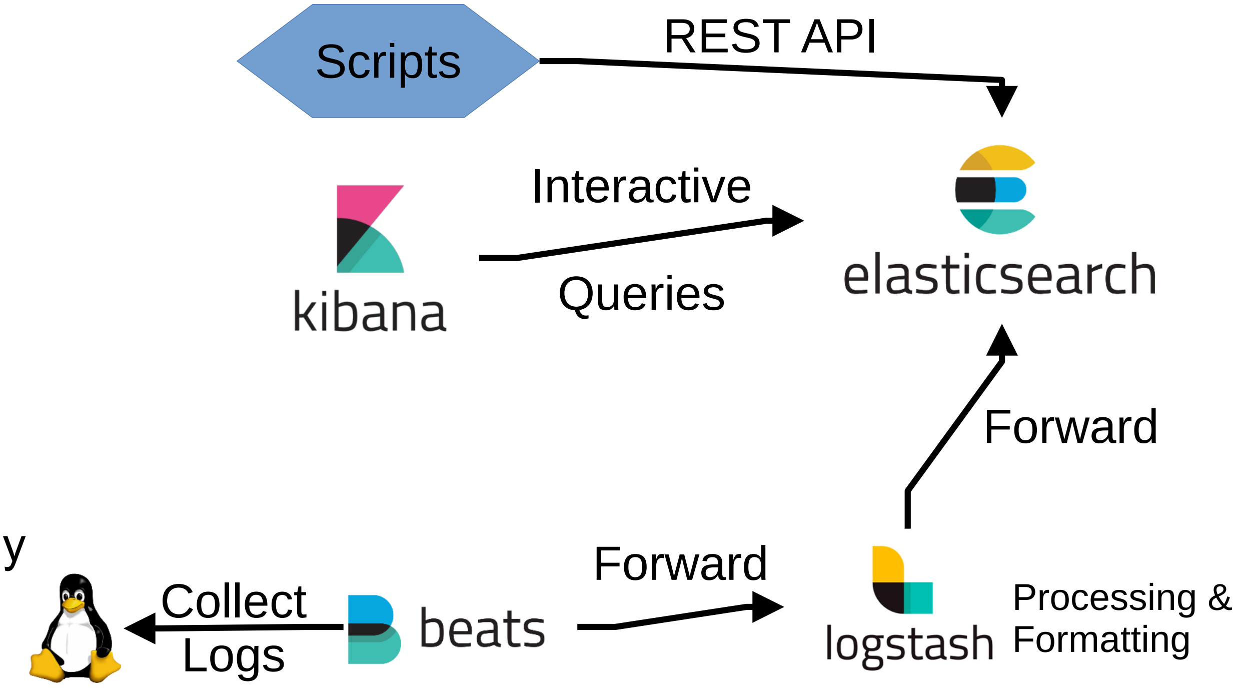
Latest Host Problems

```
max-cfelg011 is DOWN                                                   for 0m 29s
CRITICAL - 131.169.177.158: rta nan, lost 100%
1/10
sw-ipmi-al-100 is DOWN                                                 since 03:41
CRITICAL - 10.42.2.131: rta nan, lost 100%
max-infra005 is DOWN 10                                    since Aug 12 23:39
CRITICAL - 131.169.81.115: Host unreachable @ 131.169.194.6. rta na…
max-wn037 is DOWN 14                                       since Aug 12 17:37
CRITICAL - 131.169.193.210: Host unreachable @ 131.169.194.6. rta n…
batch1364 is DOWN 15                                                   for 1d 2h
CRITICAL - 131.169.160.104: Host unreachable @ 131.169.194.5. rta n…
```

Latest Service Problems

```
SSH Service on batch1372 is CRITICAL
connect to address 131.169.164.72 and port 22: (
ILO Interface on max-ferrari011 is CRITICAL 1
CRITICAL: ICMP schlug fehl: CRITICAL - 10.13.8.24
ILO Interface on max-ferrari002 is CRITICAL 1
CRITICAL: ICMP schlug fehl: CRITICAL - 10.13.8.2
ILO Interface on max-ferrari001 is CRITICAL 1
CRITICAL: ICMP schlug fehl: CRITICAL - 10.13.8.2
ILO Interface on max-ferrari006 is CRITICAL 1
CRITICAL: ICMP schlug fehl: CRITICAL - 10.13.8.24
ILO Interface on max-ferrari010 is CRITICAL 1
```

- Central Icinga 2 instance for hardware/service monitoring & alerting
  - includes on-call duty notification for business critical services
  - Central view for all machines in the data center
- Central Linux log management with **E**lasticsearch, **L**ogstash & **K**ibana (ELK)
  - Central view for all Linux related logs, allows automated analysis or centralized queries
  - Logstash: convert unstructured logs to structured logs with key/value pairs
  - ~700 GiB/day for Linux syslogs; overall ~3 TiB/day
- Stack operated by IT Network & Operations group



Scripts

REST API

kibana

Interactive Queries

elasticsearch

Forward

Collect Logs

beats

Forward

logstash

Processing & Formatting

Open Source & Monetization

- We rely on lot’s of open source software
- Observing increased monetization from companies behind popular open source tools
  - Idea of open source is declining
- Examples
  - CentOS Linux → CentOS Stream
  - Perforce buying Puppet: discontinued open source edition; alienating community, while participating from OSS
  - HashiCorp: License change to target cloud vendors
  - Anaconda Python: very popular solution, mixing free and paid channels, using paid parts by accident unavoidable
    -> blocking access to Anaconda in central firewall

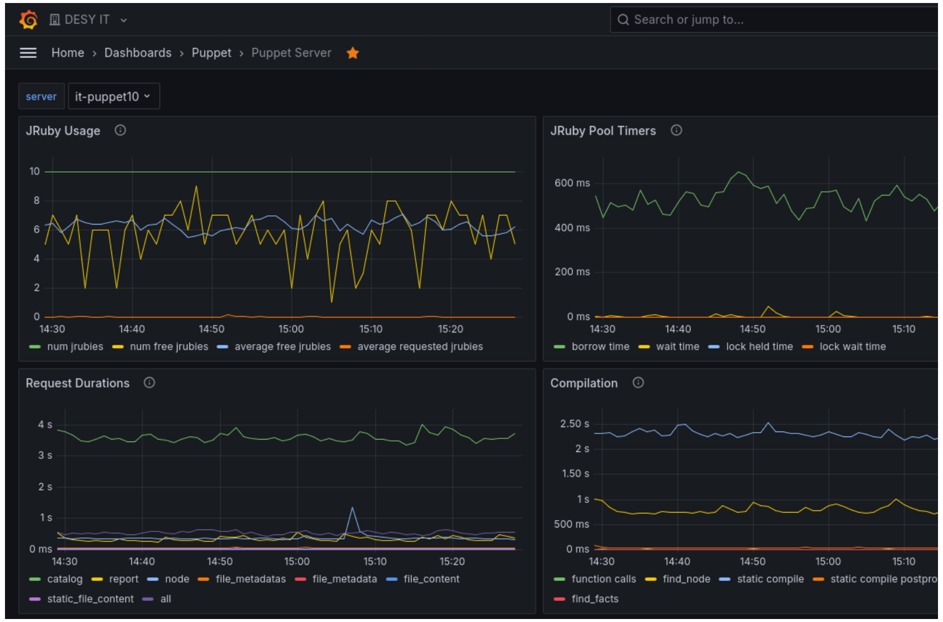
Summary

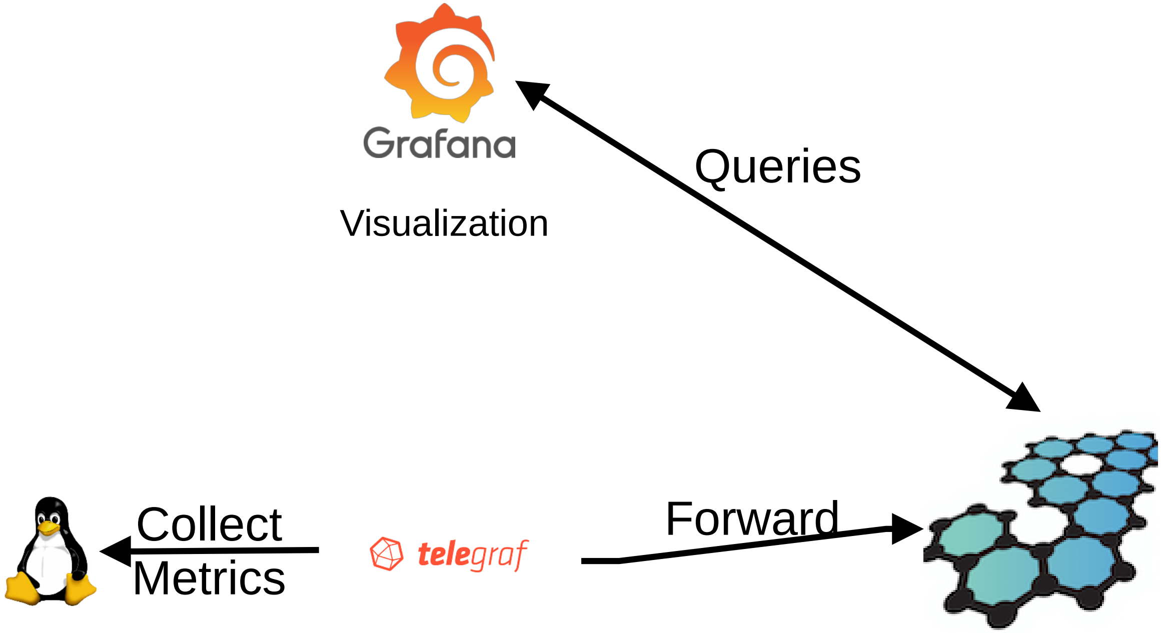
- Automation at scale is an important task
- Just providing Linux installations is not enough
  - Additional tools are necessary to keep up with the increasing number of machines
- Central views for services, logs, performance etc. are crucial
- What about Containers?
  - Starting containers for simple services can be achieved with Puppet & Quadlets
  - More complex orchestration necessary? K8s – see next talk

Vielen Dank!
Fragen?

Performance Monitoring (2)

- Performance metrics of Linux machines
- Telegraf collecting several metrics by default
  - Utilization of CPU, memory, disk…
  - Can be extended with additional inputs
- Time series database: Graphite
- Grafana for visualization





Visualization

Queries

Collect Metrics

Forward