News from the IT Department

PI IT Team

David Berghaus, *Oliver Freyermuth*, Frank Frommberger, Michael Hübner[1], Katrin Kohl, Helmut Kortmann[2], Peter Wienemann[3], Christian Wessel[4], Andreas Wißkirchen & more helping hands in projects and from the HISKP IT Team

it-support@physik.uni-bonn.de

27th April, 2023

[1] started April 2023
[2] retired April 2022
[3] changed institutes December 2022
[4] resigned June 2022

Outline

1. Personnel: Old and new faces
2. Device management
3. IT services in pandemic times: VPN, hybrid lectures
4. Cluster computing (BAF)
5. Power saving in IT
6. Distributing IT infrastructure across server rooms
7. Project highlights of the last three years
8. Funded IT R&D projects on Research Data Infrastructures

Personnel Changes

- since 2020: FTD IT position not filled
- April 2022: Helmut Kortmann retired
- June 2022: Christian Wessel resigned
- December 2022: Peter Wienemann changed institutes

- successor for Helmut Kortmann found, to start in June 2023
- Michael Hübner: started in April 2023, welcome 

Project-specific helpers

- Development Team for web and database projects: Philipp König, Jan Heinrichs, Luka Vomberg, starting June: Oliver But
- Research data infrastructure projects: Luka Vomberg, Simon Thiele
- IT specialist trainees: 3 months every year in cooperation with HRZ

(several personnel changes in the past years also in these projects)

Personnel Changes

- since 2020: FTD IT position not filled
- April 2022: Helmut Kortmann retired
- June 2022: Christian Wessel resigned
- December 2022: Peter Wienemann changed institutes

- successor for Helmut Kortmann found, to start in June 2023
- Michael Hübner: started in April 2023, welcome 

PI Web team

Ian Brock, Daniel Jacobi, Florian Kirfel, Barbara Valeriani-Kaminski
(coordinating also with FTD, HISKP and department web teams)

Flexible project developers

David Berghaus, Katrin Kohl

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices
- 5 managed Windows laptops

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices
- 5 managed Windows laptops
- 12 managed iPadOS devices

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices
- 5 managed Windows laptops
- 12 managed iPadOS devices
- 20 managed Windows desktops (3 for video conferencing)

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices
- 5 managed Windows laptops
- 12 managed iPadOS devices
- 20 managed Windows desktops (3 for video conferencing)
- 37 managed Linux laptops (32 in laptop pool)

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices
- 5 managed Windows laptops
- 12 managed iPadOS devices
- 20 managed Windows desktops (3 for video conferencing)
- 37 managed Linux laptops (32 in laptop pool)
- 192 managed Linux desktops (7 without login, e.g. for signage, labclasses,...)

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices
- 5 managed Windows laptops
- 12 managed iPadOS devices
- 20 managed Windows desktops (3 for video conferencing)
- 37 managed Linux laptops (32 in laptop pool)
- 192 managed Linux desktops (7 without login, e.g. for signage, labclasses,...)
- 271 managed Linux servers (112 BAF worker nodes, 75 VMs)

Managed devices

Some numbers...

- 4 managed macOS devices (2 iMacs, 2 MacBooks) + test devices
- 5 managed Windows laptops
- 12 managed iPadOS devices
- 20 managed Windows desktops (3 for video conferencing)
- 37 managed Linux laptops (32 in laptop pool)
- 192 managed Linux desktops (7 without login, e.g. for signage, labclasses,...)
- 271 managed Linux servers (112 BAF worker nodes, 75 VMs)

Only possible with automation!

Machine management: Foreman & Puppet

Foreman

Machine lifecycle management tool:

- Automated deployment of machines (bare metal and VMs), leveraging standard tools
- Configuration management via tools like Puppet or Ansible
- Foreman collects configuration status and machine metadata

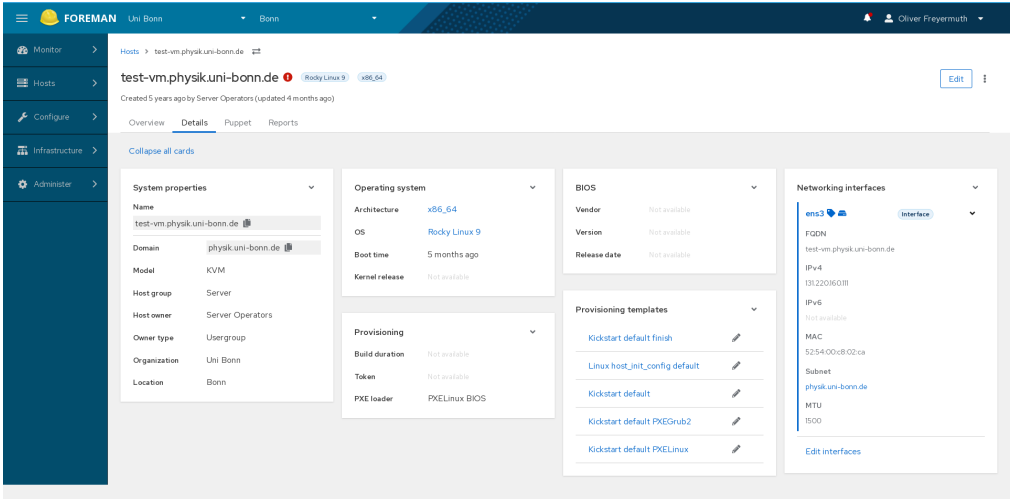
Puppet

puppet

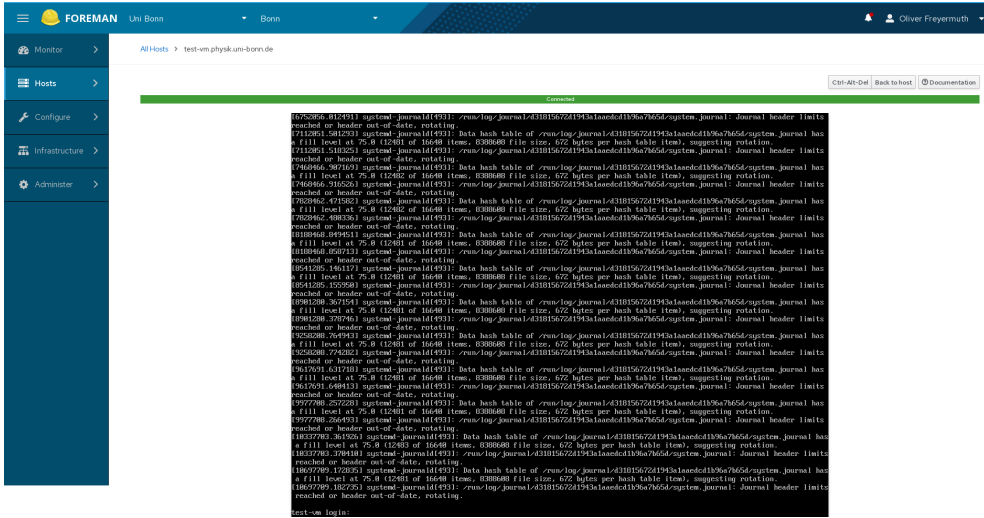
Configuration management tool:

- Runs periodically on all systems
- Ensures system is in configured state
- Reports state and hardware information

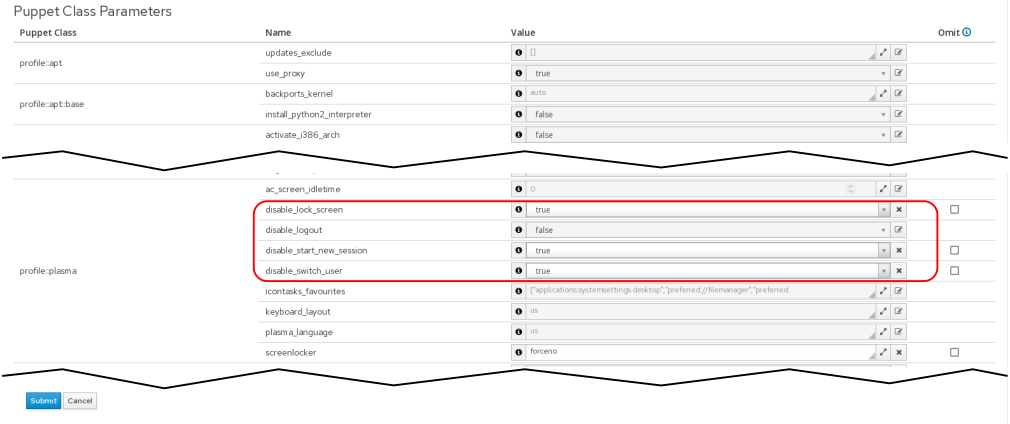
Host Details in Foreman



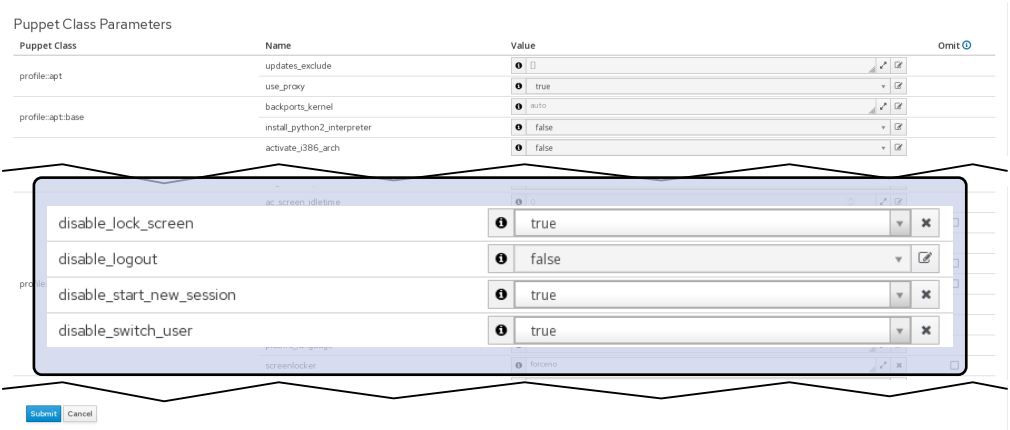
Remote access to a VM in Foreman



Configuring Hosts via Foreman



Configuring Hosts via Foreman



Interplay of Foreman and Puppet

Operational Concept

- 10 Puppet server VMs, one for each network (also take over roles of Foreman proxies for machine deployment, e.g. installing from network)
- All managed systems periodically run a Puppet agent, which reports 'facts' and asks the server for the configuration 'catalog'
- Foreman collects facts and delivers settings for individual machines or groups of hosts

Concept of 'Roles and Profiles'

- Site-specific 'profile classes' define default values and offer 'knobs' which can be turned via Foreman
- Profile classes leverage existing, Open Source Puppet modules
- Puppet modules hide the 'nitty-gritty' details of service configuration, different Linux distributions, architectures etc.

Interplay of profile class and Puppet module

Example: Configuring servers for time synchronization

```
class profile::chrony(
   Array[String] $servers = [ 'clock1.rhrz.uni-bonn.de',
                              'clock2.rhrz.uni-bonn.de' ],
) {
   class { 'chrony':
     servers => $servers,
     # Disable server mode.
     port    => 0,
   }
}
```

- Profile class assigned to hosts / hostgroups via Foreman
- `$servers` can be overwritten via Foreman
- Can query *facts* about host, use control structures, templates etc.

Puppet code development: Some numbers

- In some cases, changes to upstream Puppet modules necessary
- Changes contributed back to upstream projects (Open Source, eases maintenance)

	Lines of Code
Puppet profile code (Linux machines)	52 754
Puppet data (e.g. list of printers)	3926
Puppet profile code (Windows machines)	762
Linux-specific Puppet modules	361 403
Windows-specific Puppet modules	573 310

Significant degree of automation

Critical key component in the following timeline!

Desktop automation: Linux Desktops

- End of 2021: 'Reinstallation Campaign'
  - Migration of home directories from FSI (File System Infrastructure) to Ceph-based FDI (Research Data Infrastructure) (hoth at HRZ)
  - Migration Ubuntu 18.04 / Debian 10 $\Rightarrow$ Debian 11
  - Migration of desktop environment: Ubuntu's Unity / Gnome to KDE / Plasma
  - Migration to UEFI booting (remote reinstallation & BIOS updates possible!)
- May / June 2022: Migration of FDI mounted via SFTP to mounting via NFS
  - Until then, major lags, hangs encountered by all users
  - SFTP does not allow for local data and metadata caching nor correct file locking
  - Migration done mostly outside working hours after users had logged out
  - Now, running mostly stable, remaining hiccups still being analysed together with HRZ admins
- Likely late in 2023: Upgrade to Debian 12
  - Support for more recent hardware
  - Automated management of BIOS settings on Dell machines possible
  - First system already running with Debian 12 (OS not released yet)

Desktop automation: Windows Desktops

- Increasing number of Windows machines configured with Puppet (26 as of 2023)
- Managing both stationary and mobile machines
- Automatic installation and update of all common software without user intervention
- Manage Harddrive encryption, WireGuard VPN, WiFi, MS Office installation etc.
- Remote support possible

⇒ **Most use cases can be covered well this way!**

Desktop automation: Windows Desktops

- Increasing number of Windows machines configured with Puppet (26 as of 2023)
- Managing both stationary and mobile machines
- Automatic installation and update of all common software without user intervention
- Manage Harddrive encryption, WireGuard VPN, WiFi, MS Office installation etc.
- Remote support possible

⇒ **Most use cases can be covered well this way!**

Windows 11?

Currently, Windows 10 only — moving new systems to Windows 11 later this year, but needs special handling during installation (must prevent Microsoft Account creation).

Desktop automation: macOS

- Management crucial: Personal Apple IDs are a legal problem (and privacy issue)
- Managed (organizational) Apple IDs are likely also a GDPR issue
- Open Source management tools fail (continuous API changes in macOS, modular apps blocked by security mechanisms)
- We use Apple School Manager & JAMF School Device Management
  - Tools use managed Apple IDs, software bought by institution, assigned to devices
  - Currently, 4 production devices managed this way

However, automation is much more limited than for Linux or Windows...

Desktop automation: macOS

Issues

- No clean management of apps outside of App Store (web browsers, MS Office,...)
- Lots of missing out-of-the-box functionality in JAMF (e.g. local account creation)
- Many apps expect a local administrator (e.g. scanner drivers ask users to update and expect interaction with connected device)
- macOS built-in remote desktop support not fully usable (proprietary, modified VNC)

Bottom line

- Improvement over previous situation, but solutions not satisfactory (yet).
- May investigate JAMF Pro (used by CERN, department 2,...) and ISL remote desktop

Mobile Device Management

Linux laptop pool (Debian 11)

- Covered use cases:
  - Lecture streaming / recording
  - Masterclasses and other outreach activities
  - Computer Science courses
- Reinstalled after each use (automated)
- Online 'anywhere' without login / guest accounts via WireGuard VPN

iPads with Apple Pencils

- Covered use cases:
  - Included in a central Mobile Device Management (JAMF School)
  - Explicit use with 'Guest' access only (data deleted when session is closed)
- Used in Outreach and Teaching activities
- Online 'anywhere' without login / guest accounts via WireGuard VPN

Mobile Device Management

Linux laptop pool (Debian 11)

- Covered use cases:
  - Lecture streaming / recording
  - Masterclasses and other outreach activities
  - Computer Science courses
- Reinstalled after each use (automated)
- Online ‘anywhere’ without login / guest accounts via WireGuard VPN

iPads with Apple Pencils

- Not (yet) covered use case:
  - Devices with personal user (managed Apple ID vs. GDPR)
- Devices sometimes get stuck when triggering automated updates
- Not all settings can be enforced (e.g. WiFi can be deconfigured)

Printers

- 61 printers via CUPS print server, used for PI & HISKP networks (i.e. common print server also in FTD!)
- 41 Xerox printers
- All printers allow Scan-to-Mail, most models with OCR, just search for your Uni ID
- Automated ordering of consumables
- (Mostly) automated quarterly reporting of page counters

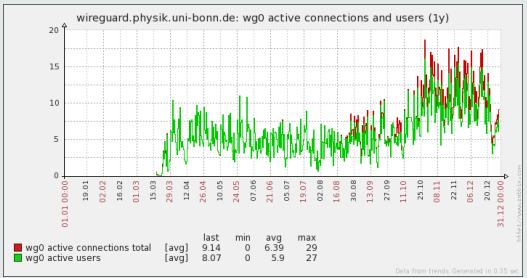
Upcoming...

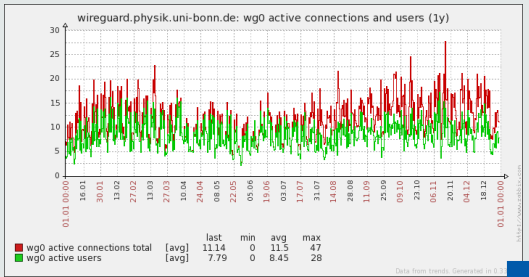
- University-wide contract with Xerox will end 2024
- New contract will be worked out, call for tenders to be prepared together with specialized consulting company

Institute VPN 'WireGuard'

- March 2020: WireGuard VPN offered
- March 2022: old Cisco ASA based VPN decomissioned (max. 10 users, slow, security issues,. . .)

WireGuard: Number of connections over time (2020 vs. 2022)

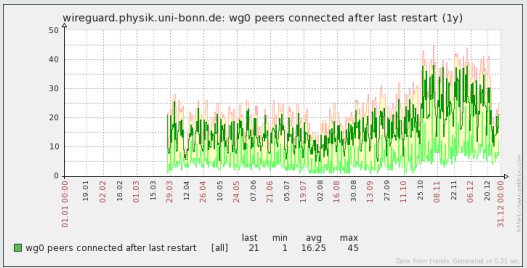


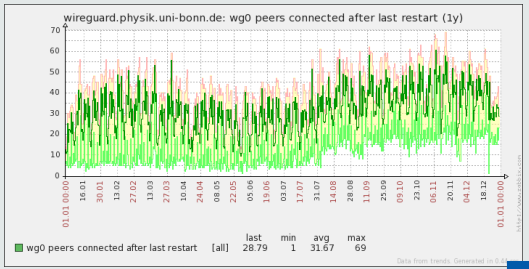


Institute VPN 'WireGuard'

- March 2020: WireGuard VPN offered
- March 2022: old Cisco ASA based VPN decomissioned (max. 10 users, slow, security issues,...)

WireGuard: Number of daily users over time (2020 vs. 2022)

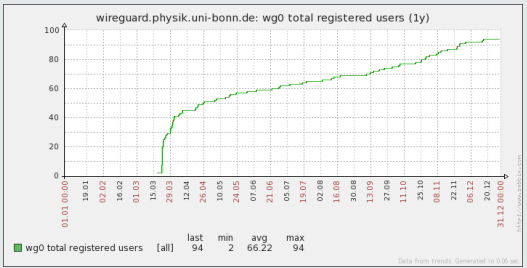


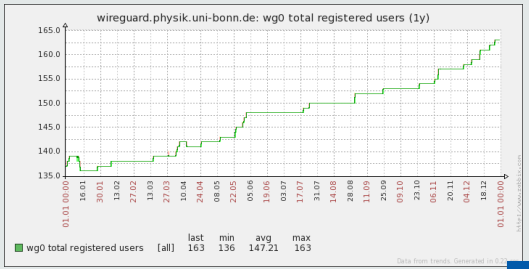


Institute VPN 'WireGuard'

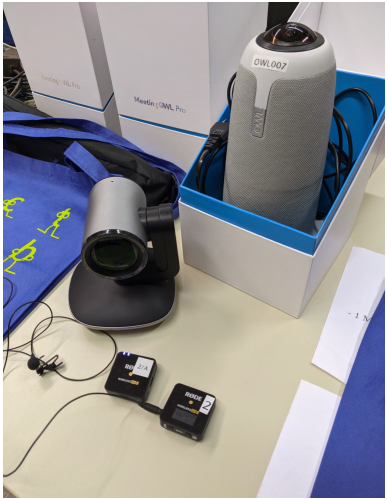
- March 2020: WireGuard VPN offered
- March 2022: old Cisco ASA based VPN decomissioned (max. 10 users, slow, security issues,. . .)

WireGuard: Number of total users (2020 vs. 2022)





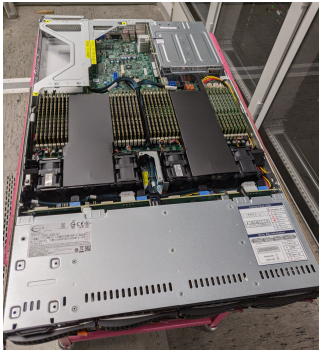
Lecture Recordings and Videoconferencing



- Started with Meeting Owls (found to be good for meetings, but insufficient for blackboards)
- Switched to mobile PTZ cameras later on, with wireless microphones
- Used pool laptops across HISKP & PI
- Technical support & student helpers trained
- Hardware re-used in conferences, meetings, workshops, colloquia,...
- More details on hardware in Confluence!

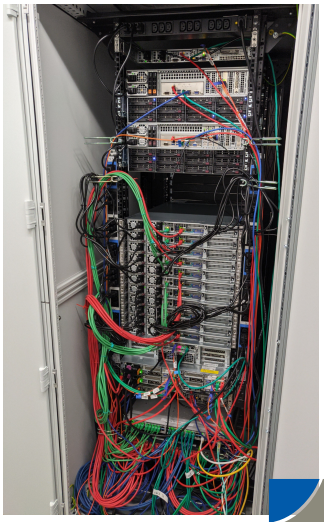
BAF Cluster

- 2017: Started with 40 worker nodes, 2240 logical cores
- 2019 and 2020: 3 waves of memory upgrades
- February 2020: 4 x NVIDIA GeForce GTX 1080 Ti, 11 GB VRAM
- July 2020: Integration of 56 worker nodes in HRZ institute machine room, 1536 logical cores total (‘CephFS_IO’)
- November 2020: Extension with 4 worker nodes, 128 logical cores each
- April 2023 (ongoing): Extension with 11 worker nodes, 256 logical cores each, 1 high-memory node: 4 TB RAM
- New nodes produce significant heat (1 kW per node)



BAF Cluster

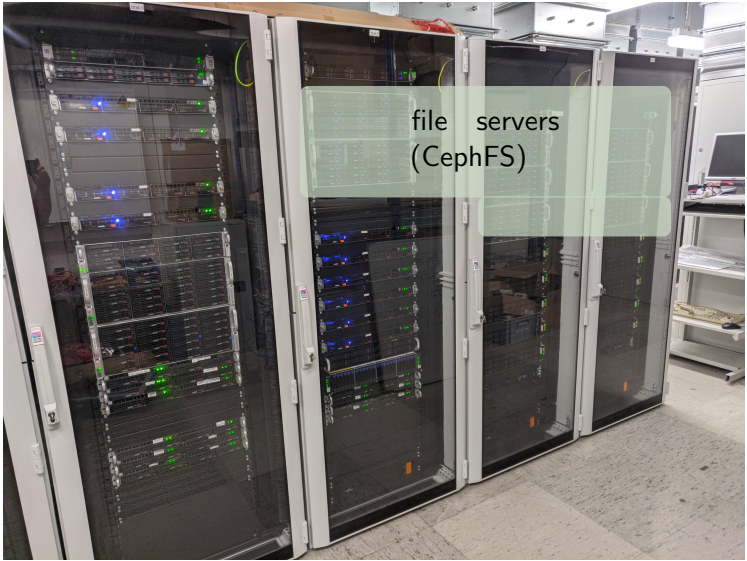
- 2017: Started with 40 worker nodes, 2240 logical cores
- 2019 and 2020: 3 waves of memory upgrades
- February 2020: 4 x NVIDIA GeForce GTX 1080 Ti, 11 GB VRAM
- July 2020: Integration of 56 worker nodes in HRZ institute machine room, 1536 logical cores total (‘`CephFS_IO`’)
- November 2020: Extension with 4 worker nodes, 128 logical cores each
- April 2023 (ongoing): Extension with 11 worker nodes, 256 logical cores each, 1 high-memory node: 4 TB RAM
- New nodes produce significant heat (1 kW per node)



BAF Cluster

BAF Cluster

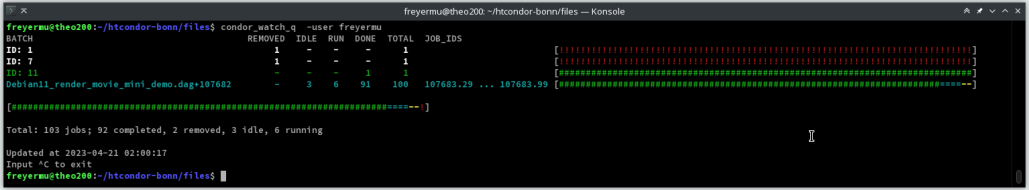
BAF Cluster



BAF Cluster

BAF Cluster: News

condor_watch_q



Operating System Containers on BAF

- Ubuntu 18.04 and 20.04, CentOS 7, Debian 10 and 11, RockyLinux 8 and 9 with site-specifics, based on official DockerHub base images
- Rebuilt at least daily with Apptainer recipe
- Deployed to our own CVMFS, kept there for at least 30 days
- Ubuntu 18.04 to be dropped, Debian 12 to be added soon

BAF Cluster: News

Organizational Developments

- Ongoing convergence to one HTC cluster for Physics Institutes
- Central HPC team: https://www.hpc.uni-bonn.de
  offering courses on Linux, Python, building your own cluster,...
- Coming soon: Large central HPC cluster 'Marvin'
- Ongoing discussions & plans to cover HTC and HPC use cases together

condor_eco_audit

```
freyermu@theo199:~$ condor_eco_audit
User                          Hours      Work        CO_2
------------------------  ----------  ----------  ---------
...
janedoe@uni-bonn.de          5394551   59.34 MWh   24.92  t
belle.cobald@uni-bonn.de     5901592   64.92 MWh   27.27  t
johndoe@uni-bonn.de          8826958   97.10 MWh   40.78  t
atlas.cobald@uni-bonn.de    21332931  234.66 MWh   98.56  t
------------------------  ----------  ----------  ---------
Cluster total              101322998    1.11 GWh  468.11  t
```

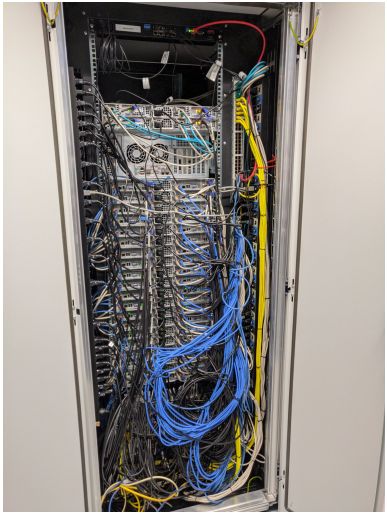
Power Saving Measures

- All desktops check whether users are logged in over night, shut down if that is not the case.
- Desktops turn on Sunday afternoon for a few hours to pull updates and configuration changes.
- Used technology: `rtcwake`, i.e. pre-programmed wake-up time.
- Note: Wake-on-LAN considered, but support inconsistent and most systems require BIOS reconfiguration (WoL disabled by vendor to save standby power)
  (will be reconsidered when BIOS configuration is automated)
- Christmas shutdown: Power consumption dropped by $\approx 30\,\text{kW}$ over two weeks in Nußallee 12:
  - 15 kW BAF cluster
  - 5 kW BAF file system
  - 10 kW service nodes, backup system, desktops
  - (plus other non-IT related energy savings)
  - used the opportunity to distribute machines across 3 locations

Server Rooms: HRZ Institute Machine Room

- 31 racks
- 1 rack filled with 56 BAF worker nodes (on the right)

Server Rooms: HRZ Institute Machine Room

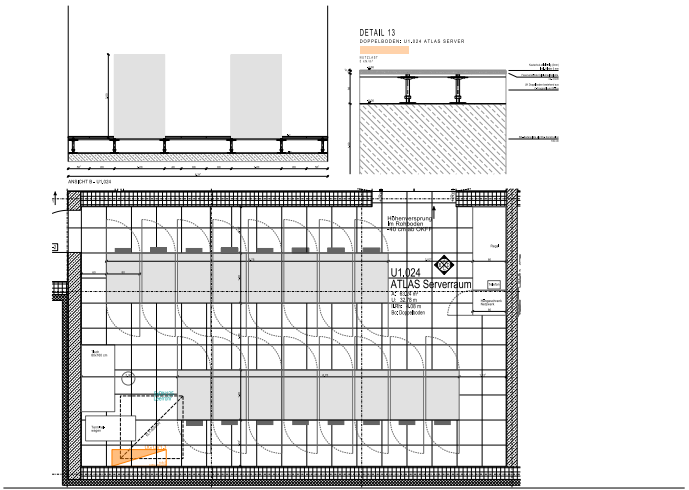


- 56 worker nodes ('rear view')
- 1 $^{\text{Gbit}}/_{\text{s}}$ ethernet, switches with 10 $^{\text{Gbit}}/_{\text{s}}$ uplink
  $\Rightarrow$ `CephFS_IO` 'medium'
- Nodes have to be drained (starting 7 days before!) if outside temperature exceeds $\approx 35\,^{\circ}\text{C}$
- Relying on DWD MOSMIX (Model Output Statistics-MIX) calculations, quite reliable (with error bands!)

Server Rooms: FTD

- 6 racks:
  - 2 network distribution and file servers
  - 2 service machines
  - 2 phone infrastructure
- central 60 kW UPS

Server Rooms: ROT



- 16 racks planned (as in current PI server room)
- increased cooling and electrical power
- cooling with cool doors (water-cooled doors at the back of the racks)
- comparable plans ongoing for move back to PI, keeping cooling and electrical power

Server Rooms

HRZ machine room	HISKP	PI	FTD

BAF Cluster: Compute Nodes

(location HISKP coming soon)

BAF Cluster: Storage

Virtualization infrastructure

almost 120 VMs, hypervisors and storage redundant, Ceph RADOS Block devices, 3 copies

Backup system

redundant, Ceph RADOS, 3 copies

Ceph – an object store

Ceph offers a '**software defined storage**':

- Distributes data across disks / machines according to configured failure domains
- Can dynamically re-balance data, include / remove nodes etc.

- Ceph is based on RADOS ('Reliable Autonomic Distributed **Object Store**')
  - 'Objects' are chunks of data (with metadata and xattrs) and a unique identifier
  - An object store allows to create, read, update or delete objects
  - Does **not** offer high-level functionality of file systems (hierarchies with folders, locking) nor structured like block storage (i.e. a 'disk' with sectors and tracks)
- Eases the implementation of replication, data life cycle, consistency checks, storage tiering,... $\Rightarrow$ Amazon S3 (Simple Storage Service)
- Backup tools (Restic, Duplicati,...) use the object store directly
- On top of the object store, Ceph implemented block devices and a file system

Virtualization Infrastructure

- Common infrastructure for PI and HISKP
- All components Open Source: Foreman, libvirt, qemu-kvm
- 120 VMs which can be 'live'-migrated between the three locations
- Backups:
  - Snapshotted virtual disks (volumes)
  - Replicated to backup Ceph system (snapshots plus live data)
  - Snapshots are backed-up daily to independent non-Ceph storage system
- Full VM life cycle (creation, backup, deletion) automated

In addition, we use the central uniVM service offered by the HRZ.

Backup System

- Common infrastructure for PI and HISKP
- All components Open Source (no extra components, 'only' Ceph)
- built from old BAF1 file servers, worker nodes and 'pre-trash' hard disks
- (almost) all components older than 10 years!
- regular scrubbing (full data at least once per week), three copies and checksums
  $\Rightarrow$ bad disks filtered out
- accessible purely as object store, i.e. dedicated backup tools needed:
  - Restic / Rustic (commandline based, Windows / Linux / macOS)
  - Duplicati (graphical, Windows / Linux / macOS)

  $\Rightarrow$ Tools offer backup plans, deduplication, compression, automatic sparsification

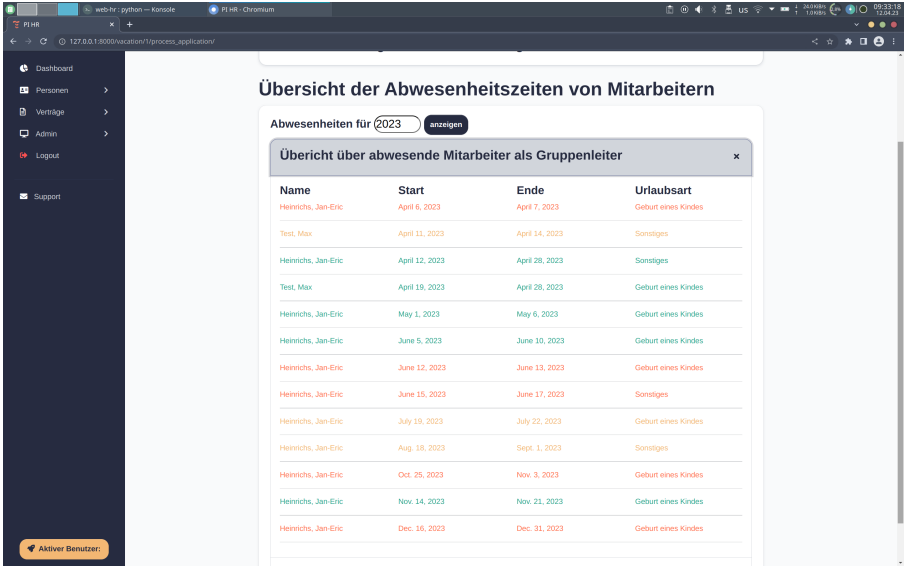
Faster, more spacy, safer and more resilient than any NAS solution!

Note: FDI team ('Forschungsdateninfrastruktur') also offers project space for experiment data storage.

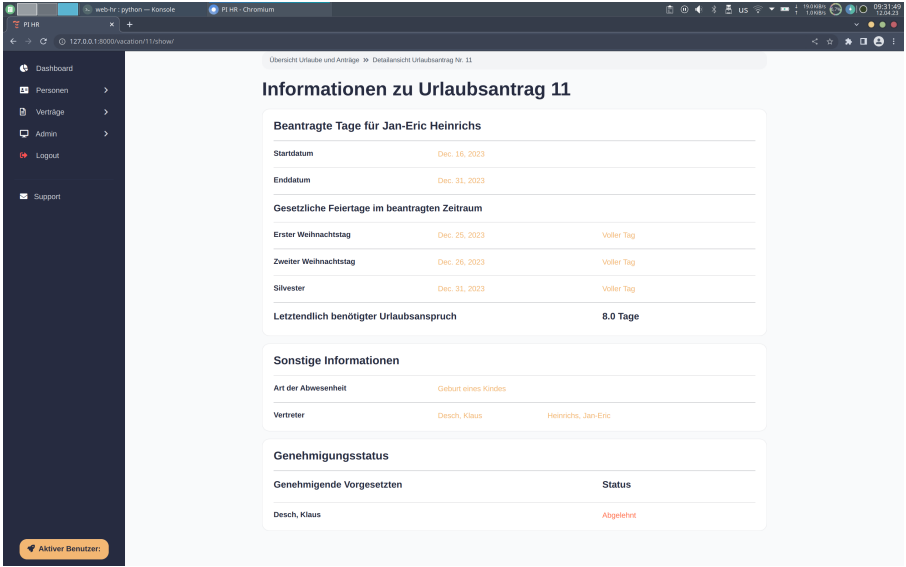
Projects (highlights of the last three years)

- 2020
  - Evaluation of MPI (jobs spanning multiple nodes) on HTC clusters (successful, but with limitations)
  - Start of PUNCH4NFDI project (next slides)
  - Backfilling the BAF cluster with jobs from ATLAS, later extended with Belle II
  - Start of JupyterHub for BAF project
- 2021
  - Publication on BAF cluster
  - 2G Semester cards
  - Start of FIDIUM project (next slides)
  - Publication 'Unleashing JupyterHub' (next slides)
- 2022 – ongoing
  - Web content team: Migration of PI web page, intranet
  - Web development team: ongoing development of HR system (next slide)
  - Joint project of PI & HISKP IT teams, secretaries, HRZ: Indico

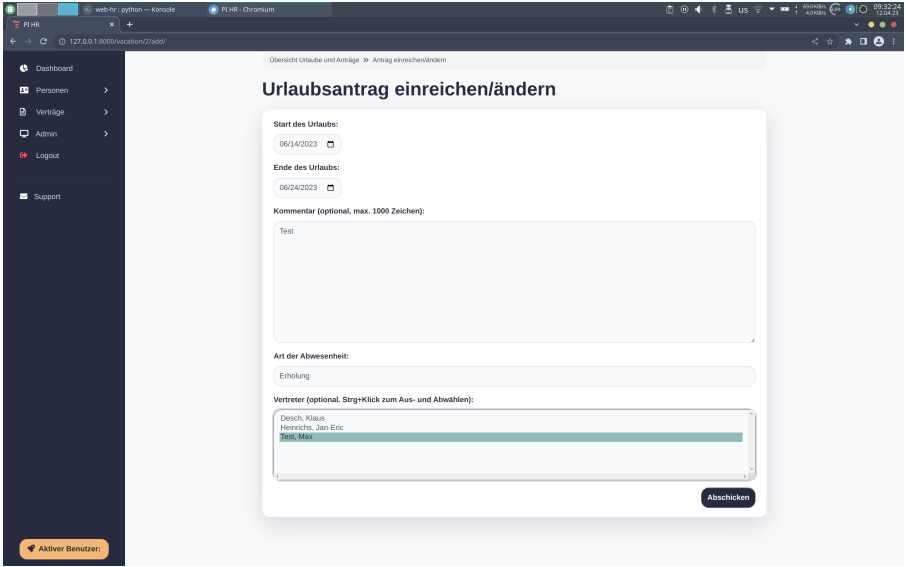
Development highlights: HR system



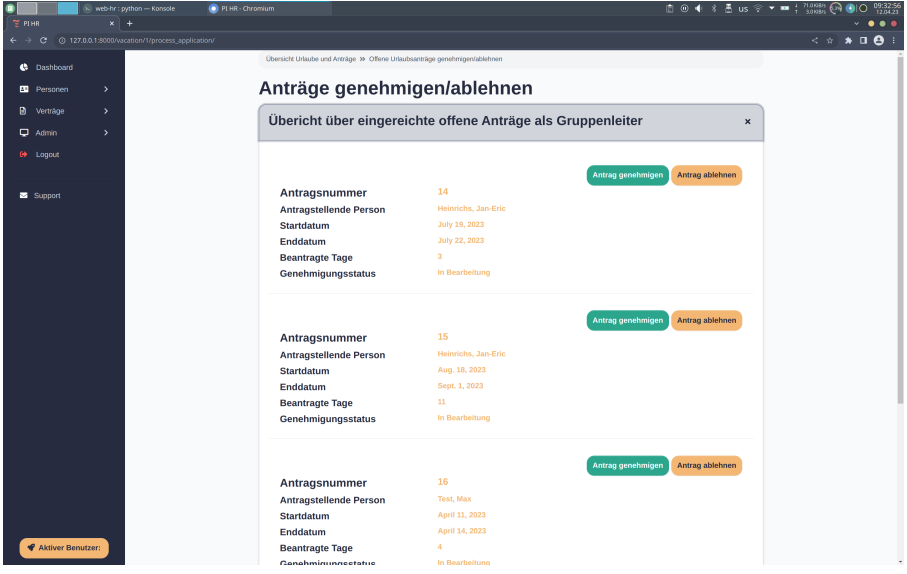
Development highlights: HR system



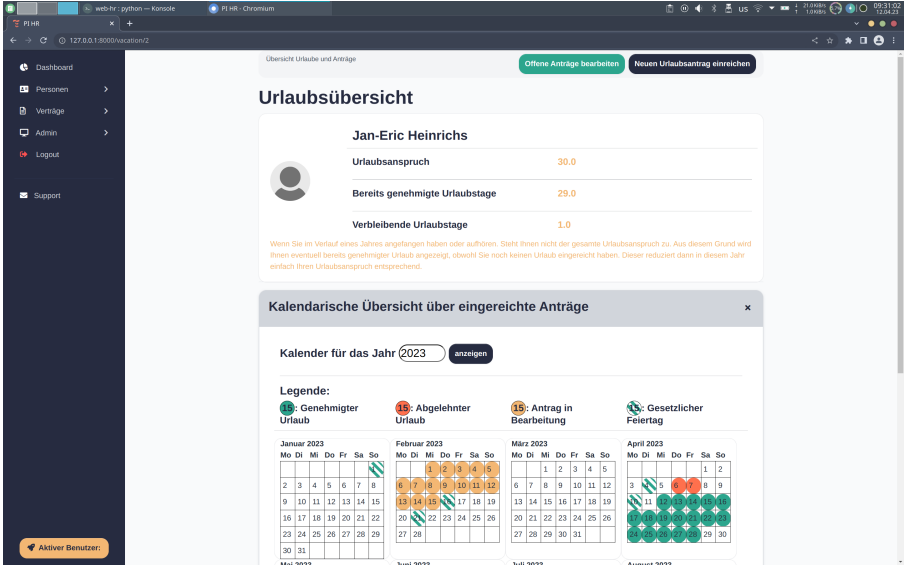
Development highlights: HR system



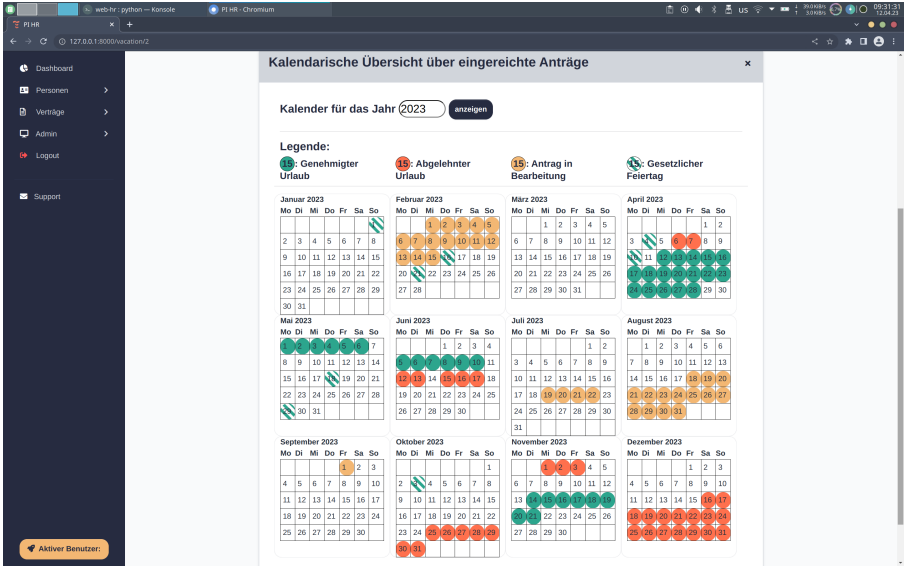
Development highlights: HR system



Development highlights: HR system



Development highlights: HR system



Project highlights: Indico event management and room booking

- Instance operated at HISKP, used for local events (from colloquia up to conferences and workshops)
- Recently: Indico in use for room booking at FTD, PI, HISKP
  Common room booking system across institutes!
- Everybody can create their own Bachelor / Master / PhD / Habilitation colloquium
  (update of abstract, upload of slides all possible in self-service)
- Interfacing tools developed, e.g. synchronizing data to Plone 5 web pages
- Indico plugin developed to grant permissions (room booking, event creation) to users with Uni ID
  (and also remove permissions if users change Universities)

Upcoming

- Login via eduGAIN (no local accounts for non-Uni-Bonn users needed anymore)
- Central Indico with similar operation model at HRZ expected later this year

Funded Projects: FIDIUM

FIDIUM project

- Federated Digital Infrastructures for Research on Universe and Matter
- Financed by the BMBF, part of ErUM-Data
- Groups: elementary particle physics, hadron physics, nuclear physics and astroparticlephysics
- 10 institutions and 4 associated partners (CERN, DESY; GridKa, GSI)
- One project in a series of projects:

  Q3/2017 → Q3/2021 ErUM-Data Pilot I: IDT-UM
  Q3/2021 → Q3/2024 ErUM-Data Pilot II: **FIDIUM**
  Q3/2024 → Q3/2025 N.N.
  Q3/2025 → Q3/2028 ErUM-Data 'föderierte Digitalinfrastrukturen' I
  Q3/2028 → Q3/2031 ErUM-Data 'föderierte Digitalinfrastrukturen' II

FIDIUM Structure & Goals

Topic Area I **Development of tools to leverage heterogeneous resources**
Coordinators: Oliver Freyermuth, Manuel Giffels

Topic Area II **Data Lakes, Distributed Data, Caching**
Coordinators: André Brinkmann, Kilian Schwarz

Topic Area III **Testing, tuning & optimization in production & analysis environments**
Coordinators: Christian Zeitnitz, Günter Duckeck

Goals

- Enable transparent use of HPCs & Cloud infrastructures
- Flexible use of resources which are only available temporarily (e.g. backfilling)
- Concentration of most production data on large centres with local caches:
  $\Rightarrow$ Concept of data lakes

FIDIUM Structure & Goals

Topic Area I **Development of tools to leverage heterogeneous resources**
Coordinators: Oliver Freyermuth, Manuel Giffels

Topic Area II **Data Lakes, Distributed Data, Caching**
Coordinators: André Brinkmann, Kilian Schwarz

Topic Area III **Testing, tuning & optimization in production & analysis environments**
Coordinators: Christian Zeitnitz, Günter Duckeck

Tasks in Bonn

- Leverage BAF (HTC) and Uni Bonn HPC for Grid jobs via an overlay batch system operated by KIT ('Single point of Entry' for ATLAS, Belle II ⇒ next slides)
- 'Compute Site in a Box': Ease deployment for other sites

FIDIUM Structure & Goals

Topic Area I **Development of tools to leverage heterogeneous resources**
Coordinators: Oliver Freyermuth, Manuel Giffels

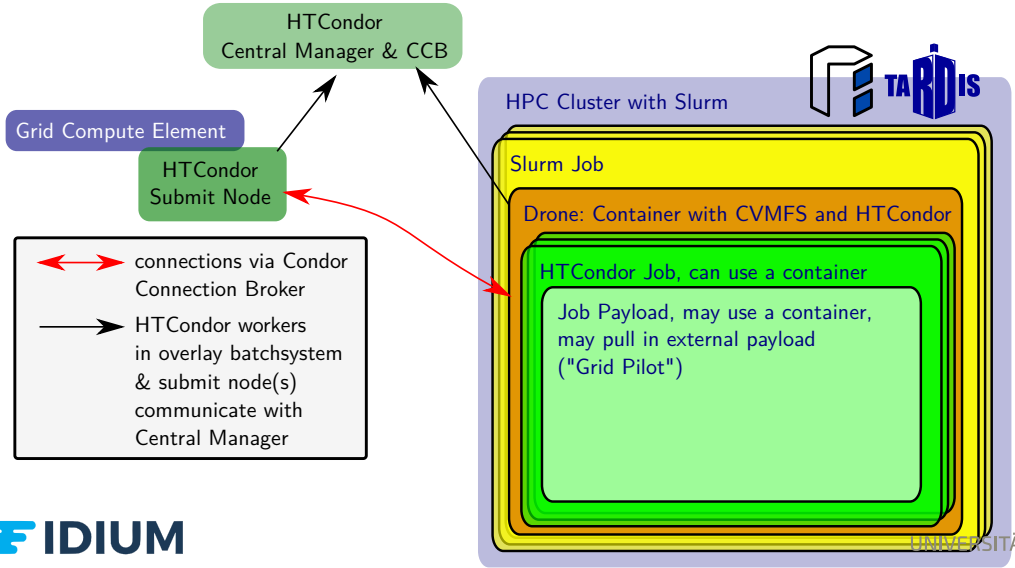
Topic Area II **Data Lakes, Distributed Data, Caching**
Coordinators: André Brinkmann, Kilian Schwarz

Topic Area III **Testing, tuning & optimization in production & analysis environments**
Coordinators: Christian Zeitnitz, Günter Duckeck

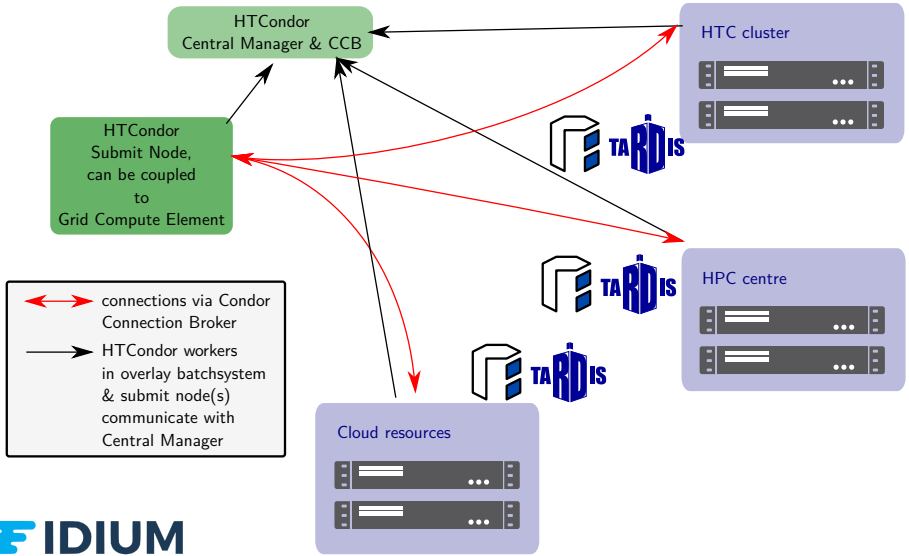
Future

Application for the successor of FIDIUM in preparation (sustainability, checkpointing compute jobs)

Federating compute resources



Federating compute resources



IDIUM

UNIVERSITÄT BONN

PUNCH: On the way to a Science Data Platform

Particles, **U**niverse, **NuC**lei and **H**adrons for the NFDI

Goal

Federated / FAIR 'Science Data Platform'
Provide infrastructure and interfaces for access to and working with data and compute resources of the communities, breaking community borders

Activities in Bonn with regard to technical infrastructure

- JupyterHub frontend for federated Compute infrastructure ('Single Point of Entry')
- Including resources in Bonn in the Compute infrastructure
- Federated storage for 'small' experiments

Other activities in Bonn

Metadata structure and definitions, Outreach activities

IT R&D for Research Data Management

- Both projects aim to federate compute and storage resources to deal with increasing resource requirements
- Connection of federated compute and storage via federated authentication
- Leverage existing compute resources via an overlay batch system
- PUNCH Goal: break community borders & offer a FAIR Science Data Platform
  - One platform / portal to allow cross-experiment analyses
  - Usage not only for Open Data
  - Allow access to data from all communities via industry standard protocols
  - Common distributed data management infrastructure including caches

FAIR: **F**indable, **A**ccessible, **I**nteroperable and **R**e-usable

FIDIUM

UNIVERSITÄT BONN

Take-home messages

- Automated device management is key to 'survival' through IT evolution
- Continuous modernization of existing services: BAF, VPN,...
- Increasingly distributed infrastructures / cross-institute collaboration
- Successful R&D projects in IT
- Debian 12 for Desktops, BAF extension on the horizon
- Check out our documentation pages on Confluence!

Ongoing future project

Common institute firewall for Physics institutes:
More bandwidth, IPv6, less poking of firewall holes

In case of any questions, don't hestitate to contact us!

it-support@physik.uni-bonn.de

Thank you

for your attention!

FIDIUM Structure & Goals

Topic Area I **Development of tools to leverage heterogeneous resources**
Coordinators: Oliver Freyermuth, Manuel Giffels

Topic Area II **Data Lakes, Distributed Data, Caching**
Coordinators: André Brinkmann, Kilian Schwarz

Topic Area III **Testing, tuning & optimization in production & analysis environments**
Coordinators: Christian Zeitnitz, Günter Duckeck

Goals

- Enable transparent use of HPCs & Cloud infrastructures
- Flexible use of resources which are only available temporarily (e.g. backfilling)
- Concentration of most production data on large centres with local caches:
  $\Rightarrow$ Concept of data lakes

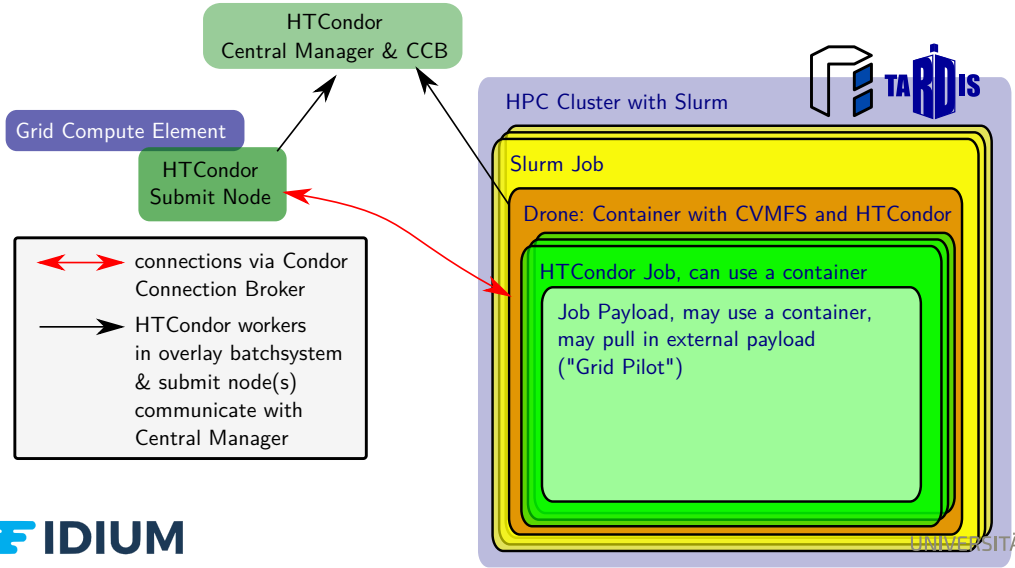
FIDIUM Activities in Bonn

- Project ‘Compute Site in a Box’
  - Usage of COBaID/TARDIS (tools developed at KIT for opportunistic computing)
  - Opening and extending existing Puppet configuration, test deployment with partners
  - Generalize unprivileged container workflows for COBaID/TARDIS resources
  - Leverage publicly accessible container infrastructure
  - Partners Mainz and GSI Darmstadt as ‘testbeds’
  - Documentation and workshops for inside and outside particle physics community
- Funding: 1 FTE for three years, equal own contribution
- Close integration with developments in other topic areas (Accounting, Caching, Monitoring)

Future

Application for the successor of FIDIUM in preparation (sustainability, checkpointing compute jobs)

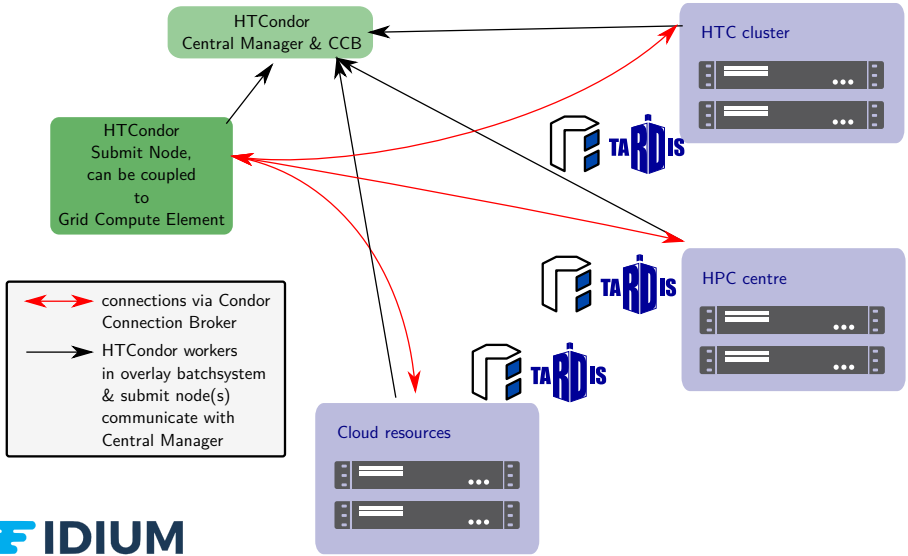
Federating compute resources



IDIUM

UNIVERSITÄT BONN

Federating compute resources



IDIUM

UNIVERSITÄT BONN

Transparent integration of Compute resources

Main steps

- Jobs submitted locally ⇒ Execution on **Overlay Batch System**
- Software stack as container (shipped via Cern-VM FS)
- Unprivileged containers started as jobs on site: Leverage User Namespaces
- Inside, an HTCondor `startd` (execute node) is started ⇒ 'Drone'
- Drone registers with Central Manager of the Overlay Batch System
- Jobs inside Overlay Batch system can use containers themselves

Scaling of number of 'Drones' based on resource usage

- COBaID — the Opportunistic Balancing Daemon
- TARDIS — The Transparent Adaptive Resource Dynamic Integration System

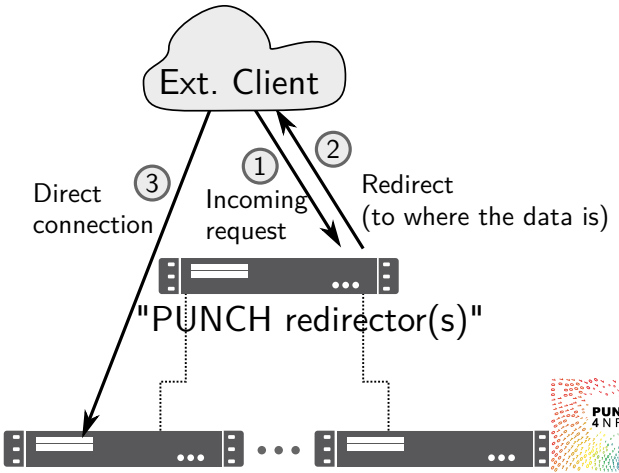
Successfully used with HPC BONNA

IDIUM

UNIVERSITÄT BONN

XRootD: Federating data access across sites

- XRootD storage provides HTTPS / WebDAV and XRootD protocols
- Global namespace: Use **g**lobal **L**ogical **F**ile **N**ames
- Redirection from central redirector to site with physical files ('merge' sites)
- Compute jobs can be routed to sites with data, or close to them
- Concept also works with caches



HTCondor Networking: JupyterHub