



dCache Project Updates

Tigran Mkrtchyan for the dCache collaboration



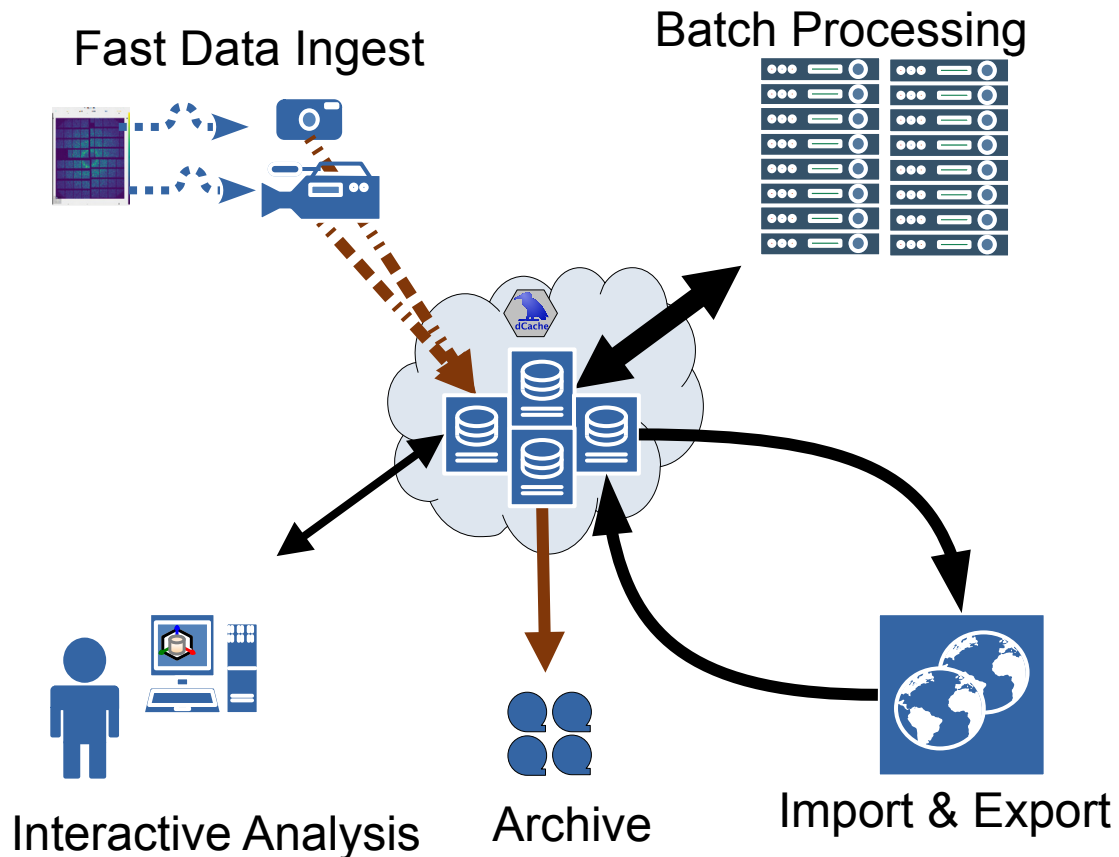
HELMHOLTZ

RESEARCH FOR
GRAND CHALLENGES

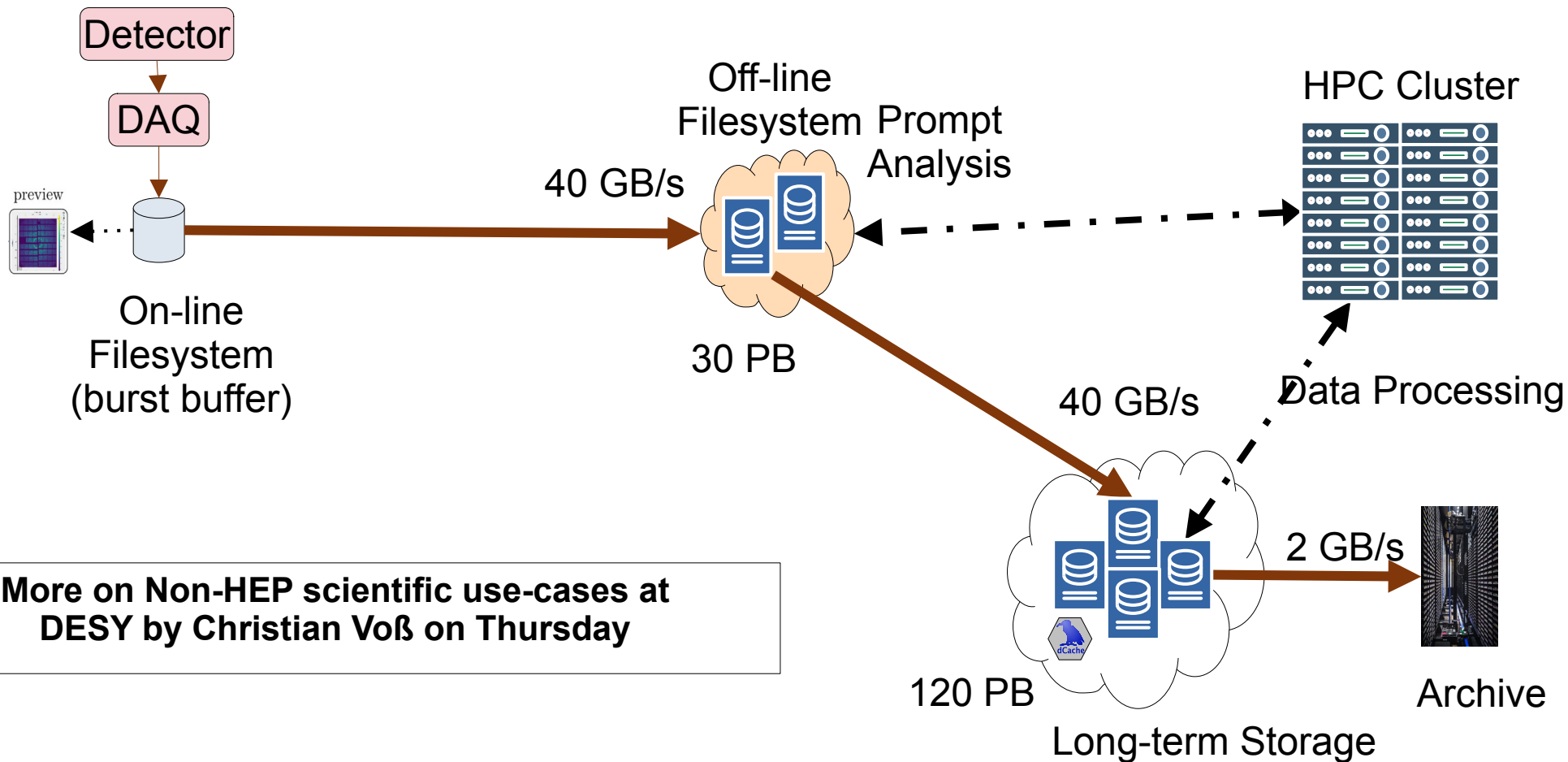
Mission Statement



- Ingest
 - Multiple parallel streams.
 - High data rate.
 - Large number of files.
- Analysis
 - High CPU efficiency.
 - Chaotic user access.
 - Standard access protocols.
- Sharing&Exchange
 - Effective WAN access.
 - In-flight data protection.
 - Federated Identity handling.
- Long-Term Archival
 - High reliability.
 - Automatic technology migration.



Scientific Use Case

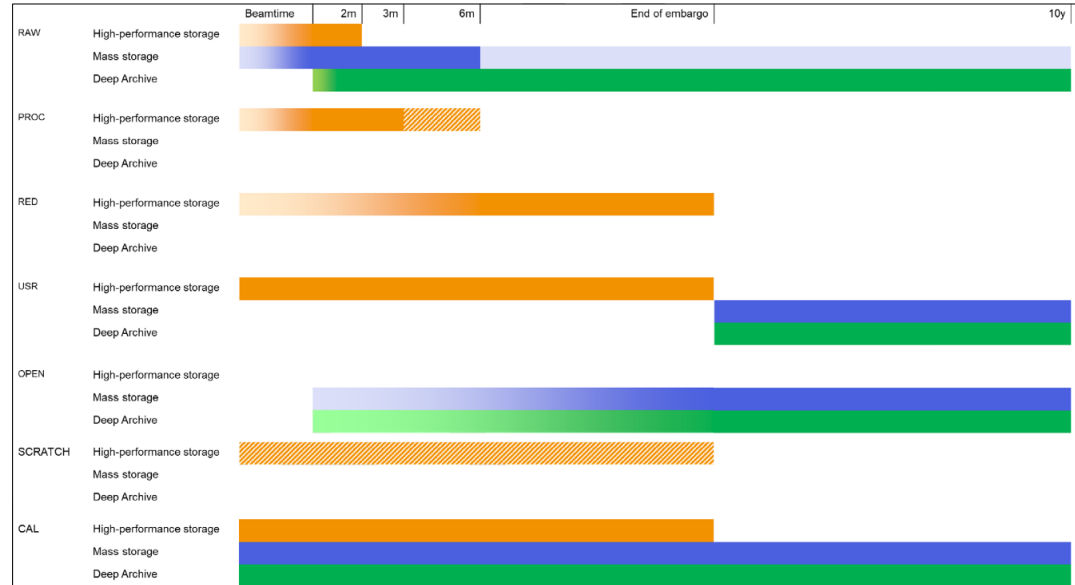


More on Non-HEP scientific use-cases at DESY by Christian Voß on Thursday

Policy driven QoS system



```
"name": "xfel-raw-policy",  
"states": [  
  {  
    "duration": "P6M",  
    "media": 1x DISK, 1x HSM  
  },  
  {  
    "media": 2x HSM  
  }  
]
```

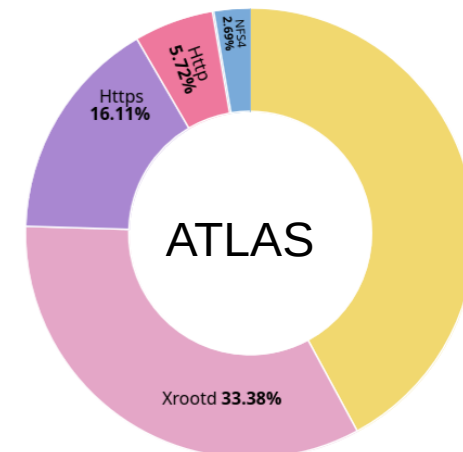
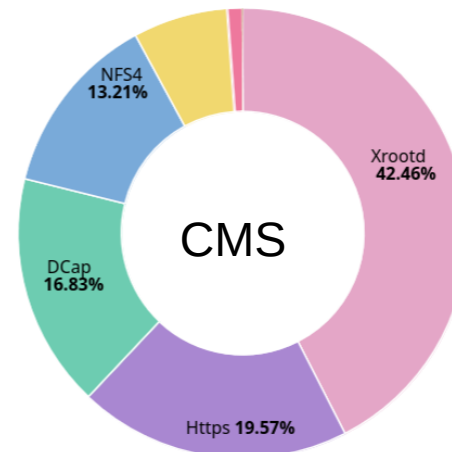
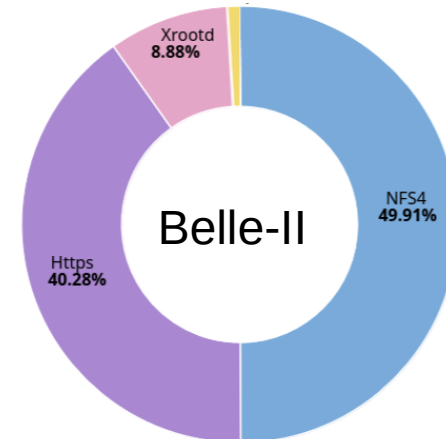
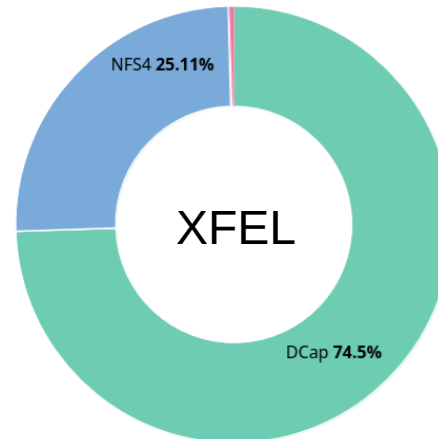


Source: https://www.xfel.eu/users/policies/index_eng.html

Protocols and Instances



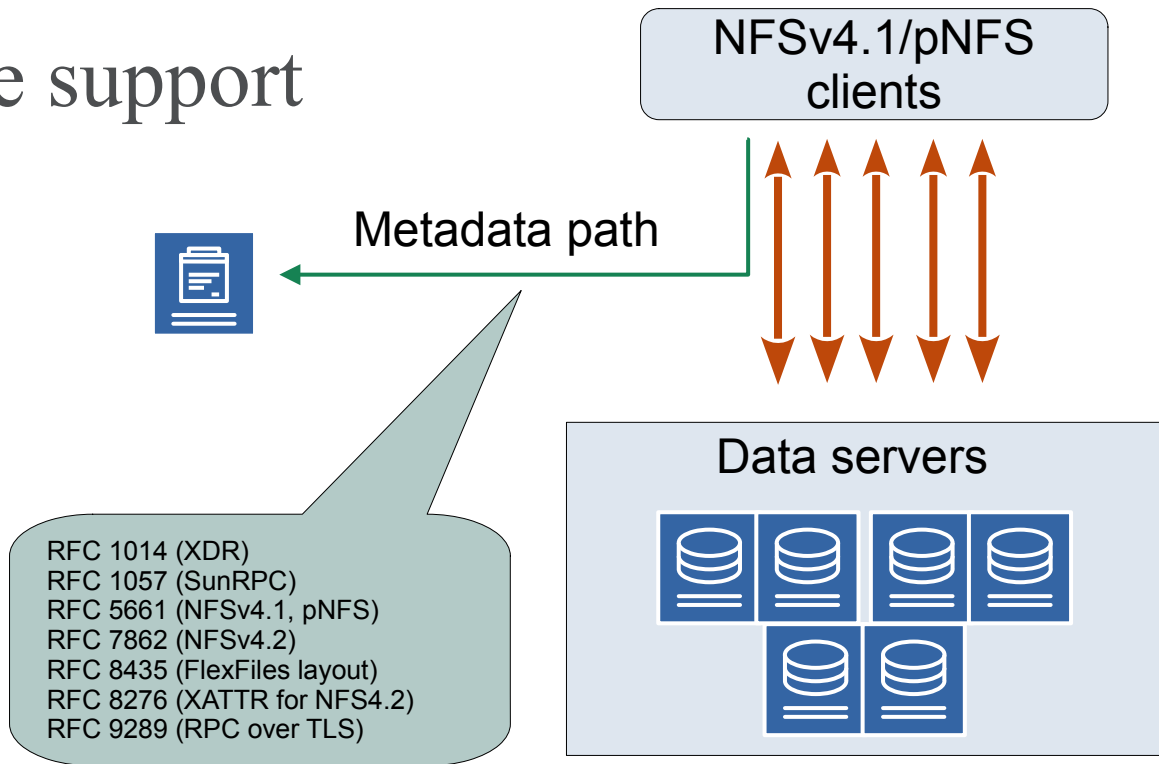
- Data is going to grow... A lot...
 - High ingest data rates.
 - More movements between sites.
- Shared Computing Resources
 - Analysis Facilities.
 - Grid Farms.
 - HPC.
 - Cloud resources (CPU&Storage).
- Standard analysis tools
 - ROOT.
 - Jupyter Notebooks.
 - Non-ROOT tools.
- Competing Tape Operations



NFS Improvements



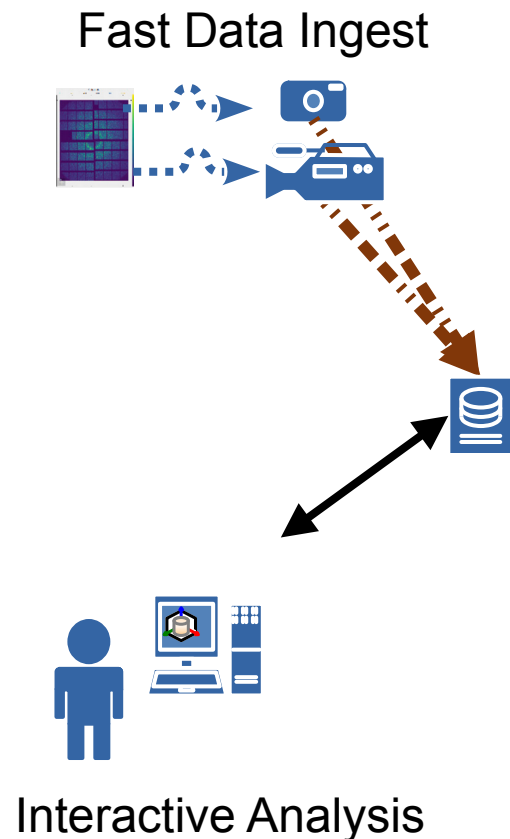
- I/O throughput
- Extended attribute support
- Quota support
 - Online files only.
- NFS-over-TLS



POSIX constrains



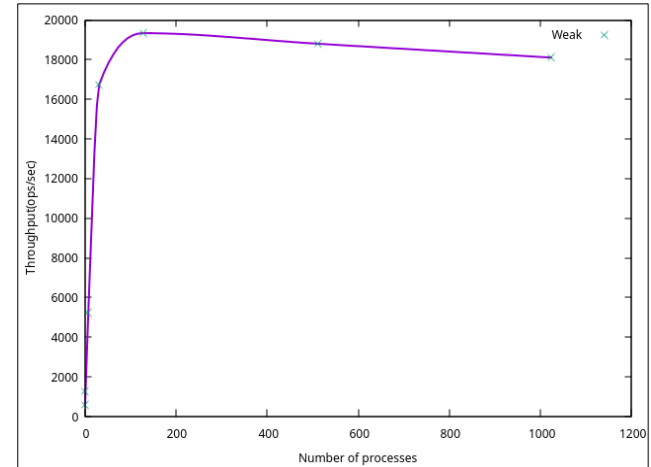
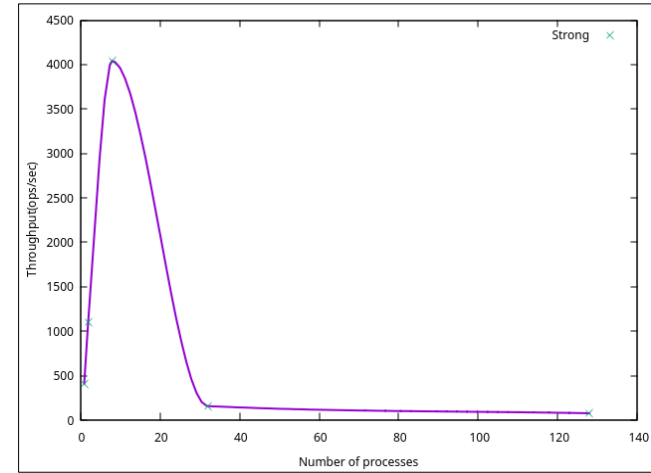
- POSIX enforces strong attribute consistency
 - On create directory attribute change must be visible to all clients
- Burst workloads are serialized, thus provide low throughput
- Not all clients need up-to-date directory attributes



Tunable Namespace Consistency



- **STRONG**
 - A creation of a filesystem object triggers simultaneous update of parent directory's mtime, ctime, nlink and generation attributes.
- **WEAK**
 - A creation of a filesystem object will eventually update (after 30 seconds) parent directory's mtime, ctime, nlink and generation attributes. Multiple concurrent modifications to a directory are aggregated into single attribute update.
- **SOFT**
 - Same as weak, but, reading of directory attributes takes into account pending attribute updates.



Labeling



- Add custom labels to existing file
- Labels dynamically create file collections
- Collections exposed as virtual-directories
- Files can be accessed by regular path
- Files can be accessed by collections



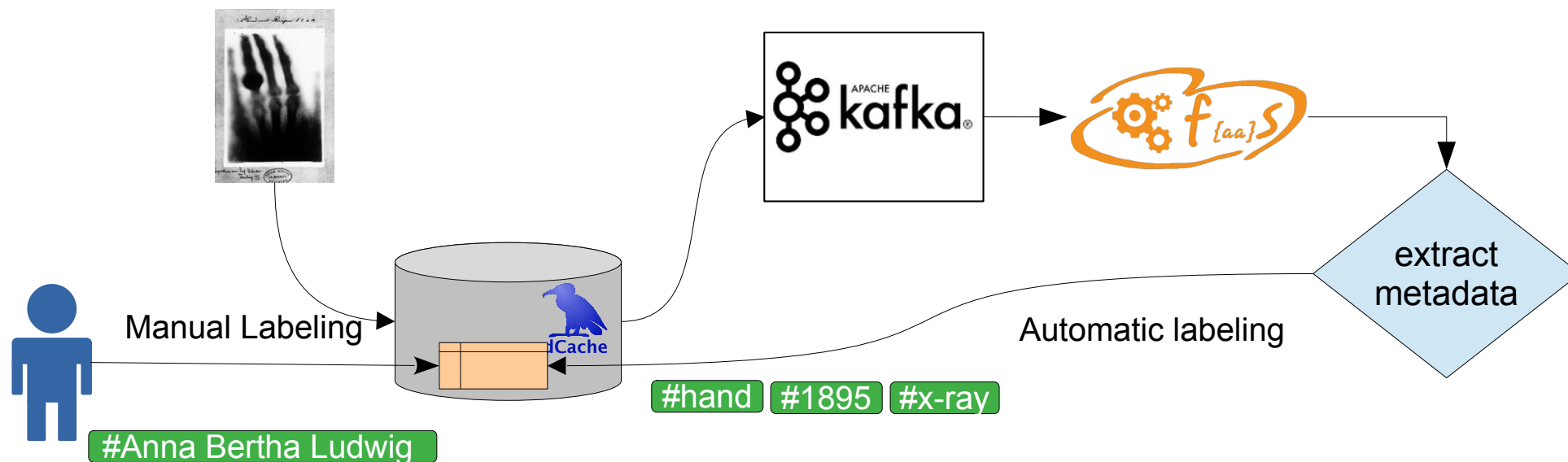
#Anna Bertha Ludwig

#hand

#1895

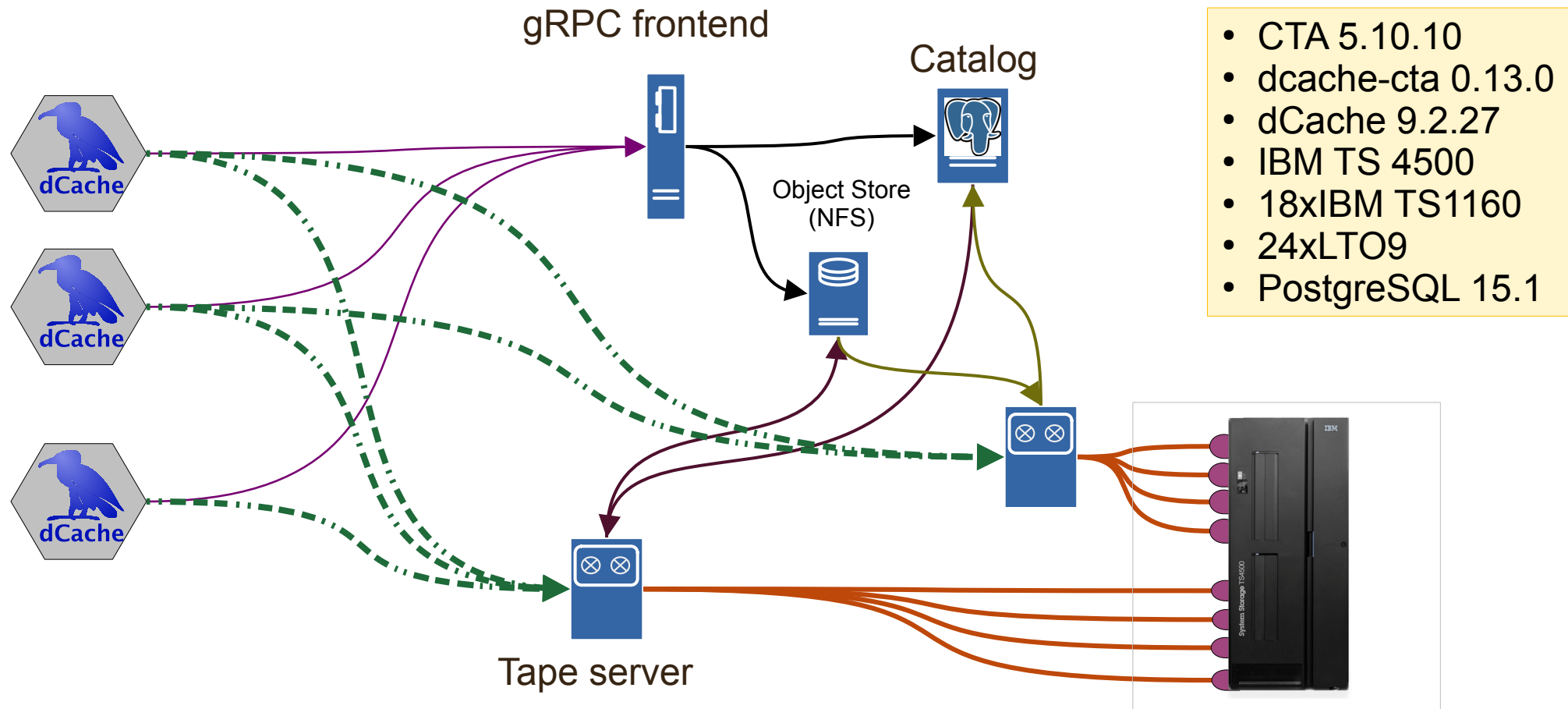
#x-ray

Metadata Population



For details, visit our Poster: label-based virtual directories in dCache, Oct 24

CTA Integration



CTA Integration

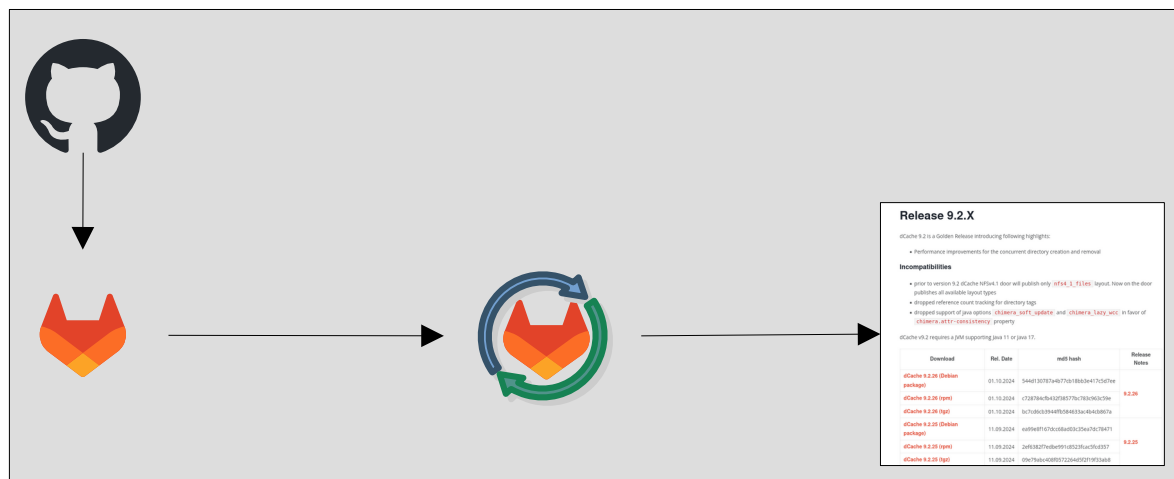
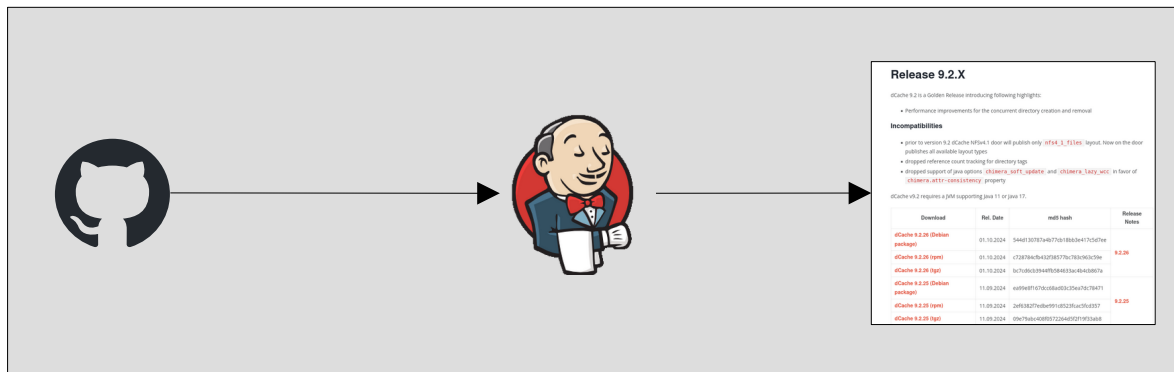


For details, visit our Poster: *dCache-CTA Integration: One Year in Production at DESY, Oct 22*

CI-Pipeline



- Migration from Jenkins to Gitlab-CI.
- The build/test process is 'documented' in repo.
- Helm charts to deploy dCache in k8s.



- A low overhead ($\sim 2\%$) profiler built into JVM
- Simple admin commands to start/stop recording
 - More sophisticated configs can be enabled with java tools.
- The collected recordings can be shared with developers for later analysis
- Detailed information in the *Debugging* section of “The dCache book”

The Mapping and Access Handling



```
"wlcg.ver": "1.0",  
"sub": "2ea41db8-ecd7-47a1-8f86-7ac97ebcc466",  
"aud": "https://wlcg.cern.ch/jwt/v1/any",  
"nbf": 1662472763,  
"scope": "openid profile storage.read:/ storage.create:/",  
"iss": "https://cms-auth.web.cern.ch/",  
"exp": 1662476363,  
"iat": 1662472763,  
"jti": "88a1f764-8311-4210-a8cd-d5604b942d8c",  
"client_id": "106ceb94-e89c-421a-9c7d-1424e2fa4668",  
"wlcg.groups": [  
  "/cms",  
  "/cms/dcms",  
  "/cms/integration"  
]
```

Tokens carrying **storage.*** OR **compute.*** scopes are explicit AuthZ tokens and exempt from any further permission checks!

Python-gPlazma module



- Admin friendly.
- Quick way to prototype site-specific mapping/authentication.
- Dynamically loaded.

```
def py_map(principals):  
    if "user:tigran" in principals:  
        principals.add("uid:1234")  
        return True  
    else:  
        return False
```

Other Changes

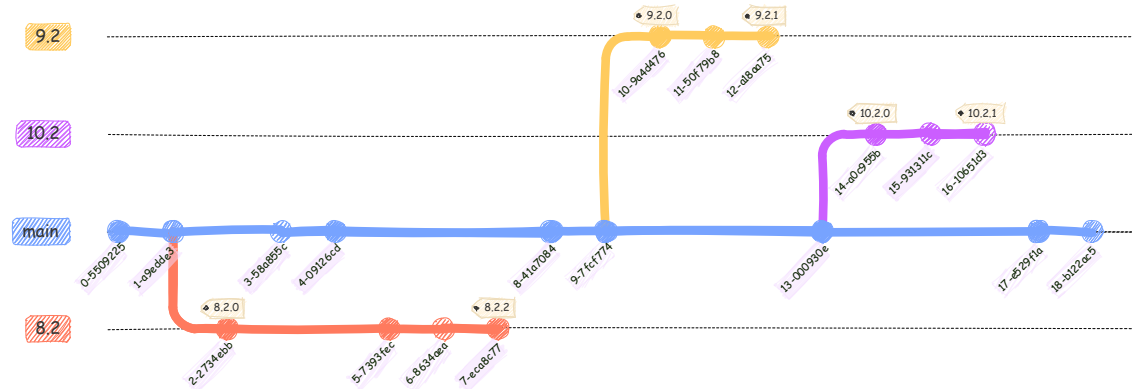


- Persistent port number per protocol on pools
- XROOT proxy-IO through the door
- Java Versions (17 recommended, 21 under evaluation)
- Direct I/O for scrub jobs
- `storage.stage` scope support
- Environment variables passed as dCache properties.
- Network fireflies (PoC)
- Json format for billing files (no need for kafka).
- New look-and-feel

dCache versions



- 8.2 out of support
 - Security fixes only.
- 9.2 Current
 - Support till fall 2025.
- 10.2 the next **Golden Release (LTS)**
 - Release last week.
 - Support till fall 2026.



Technical Directions



- Scaleout
 - Namespace.
 - Number of pools (SW/HW).
- BULK operations
- QoS
- Token-based Authentication
- Better *Analysis Facility* support
 - POSIX access and compliance.
 - HPC workload support (DDoS protection).
- Tape integration



Thank you!



Questions?