

# Baldr: Loki's Brother for HPC

Timon Vogt



## GWDG's "unified monitoring"

The GWDG has deployed Loki as part of our new "unified monitoring" last year.

- "Unified Monitoring" is a Grafana-based stack: Alloy, Prometheus, *Loki* and Grafana.
- Push-based instead of pull-based, thanks to Alloy.
- Log-monitoring was added for the first time.
- Each service gets their own OpenStack-VM, nginx proxy handles TLS and Basic Auth.
- Grafana Login handled by GWDG SSO.

# Grafana Dashboard

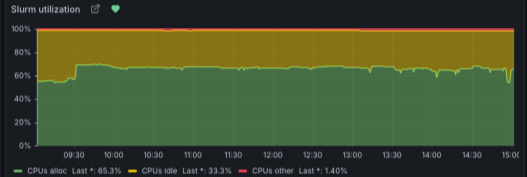
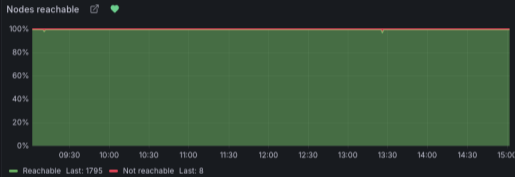
Last 6 hours 🔍 Refresh 30s ↻

Mi. 2. Juli 2025

15:02:17

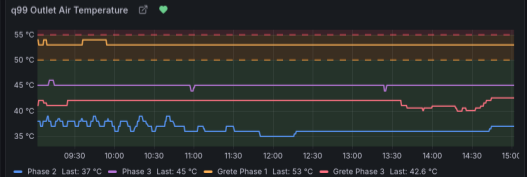
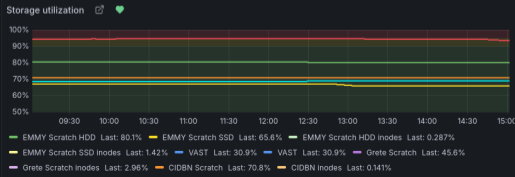
Warewolf <a href="#">🔗</a>	Logins responsive <a href="#">🔗</a>	LDAP Status
<b>OK</b>	<b>17</b>	<b>Reachable</b>

Slurmctlid <a href="#">🔗</a>	Jobs pending <a href="#">🔗</a>	Avg Waiting time <a href="#">🔗</a>
<b>OK</b>	<b>17676</b>	<b>1.11 days</b>



Storage Status <a href="#">🔗</a>	Metadata Utiliza... <a href="#">🔗</a>	Users above Limit						
<b>OK</b>	<table border="1"><tr><td>4.29%</td><td>9.83%</td><td>0.679%</td></tr><tr><td>8.28%</td><td>8.28%</td><td>0.0457%</td></tr></table>	4.29%	9.83%	0.679%	8.28%	8.28%	0.0457%	<b>12</b>
4.29%	9.83%	0.679%						
8.28%	8.28%	0.0457%						

Current energy ... <a href="#">🔗</a>	MDC Room Temperat... <a href="#">🔗</a>	AIMS Status <a href="#">🔗</a>
<b>799 kW</b>	Also Room (unreached) <b>27.0 °C</b> New Room (unreached) <b>30.5 °C</b>	<b>Failure</b>



# Loki?



## Grafana Loki

*Loki is a horizontally scalable, highly available, multi-tenant log aggregation system inspired by Prometheus*

*(<https://grafana.com/oss/loki/>)*

Open-Source (AGPLv3), Developed by Grafana Labs Inc., integrates well into the Grafana ecosystem

# Issues on the horizon

- First issue: learning that using the local file system was not officially supported by Grafana Labs.

Great for low volume applications, proof of concepts, and just playing around with Loki.

The filesystem is not supported by Grafana Labs for production environments

- ▶ You are supposed to AWS S3 or GCS

- Next up: [Issue#1502](#). Too many files.

```
root@loki:/opt/test/chunks/fake# time ls -l | wc -l
2724504
real    1m44,071s
```

# Investigating

- Loki uses **Protobuf** for transmission (within HTTPS) to serialize the **Stream** data type.
- A **Stream** consists of Labels and Entries

```
type Stream struct {  
  Labels string `protobuf:"bytes,1,opt,name=labels,proto3" json:"labels"`  
  Entries []Entry `protobuf:"bytes,2,rep,name=entries,proto3,customtype=EntryAdapter" json:"entries"`  
  Hash    uint64 `protobuf:"varint,3,opt,name=hash,proto3" json:"-"`  
}
```

- Issue found: Every time the label set changes, Alloy has to create a new **Stream** object and pushes out a new transmission.

# What is Baldr?



Logo coming soon

Baldr is my own re-write of Loki. With the parts that I needed.

- Works as a drop-in replacement for Loki (can not use its data, though)
- Saves logs in a different format (.log.gz instead of protobuf).
- Supports LogQL.
- Smaller code footprint: 2763 LoC vs 8412375 LoC

## Baldr's storage format

How does Baldr solve the storage issue?

- Baldr parses the **Stream** objects, as they come in.
- Extracts the instance label, as well as the timestamp range.
- Creates one file per instance, per day, instead of one per **Stream** object.
  - ▶ Actually appends to files.
- Writes log lines into `.log.gz` files (gzipped text).
  - ▶ Not as resilient, but far fewer files.  
2000 nodes \* 365 days = 730000 files per year
- Files can be read with simple tools (e.g. `zcat`)
- Files are understandably named, makes backups easier.

# Comparison with Loki

## Things not yet in Baldr

- LogQL is not completely supported  
(Subqueries missing, unwrapped ranges aggrs missing)
- Index not fully implemented.
- Ability to read tar files directly

## Things never in Baldr

- Clustering / Sharding, Multitenancy
- Alertmanager and WebUI: Using Grafana for that.

---

End

Thank you.

Questions?

Repository: <https://github.com/atopion/baldr>