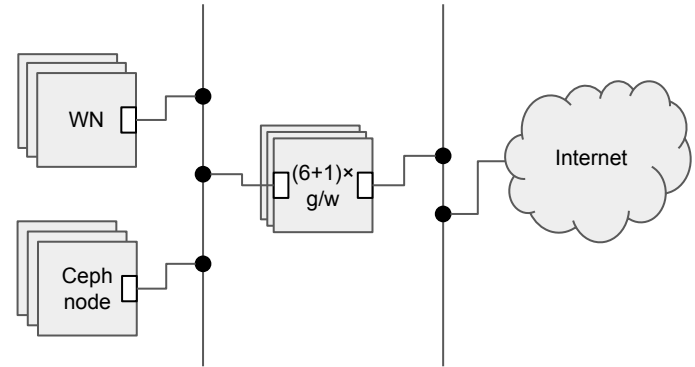


# XRootD Monitoring at Lancaster

Steven Simpson, Gerard Hand, Matt Doidge, Pete Love,  
Roger Jones

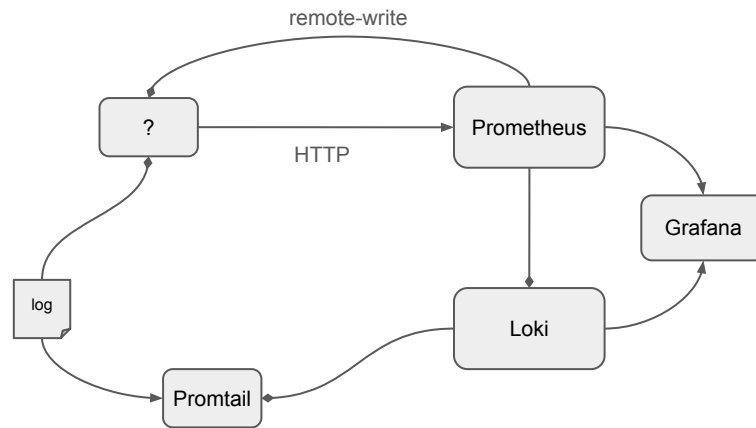
# XRootD/Storage architecture at Lancs

- Ceph storage
  - ~30 storage nodes
  - ~30×24 discs ≈ 11PiB
  - CephFS interface
    - 2 metadata servers
- XRootD
  - 1 redirector
  - 6 gateways
  - CephFS mounts on each host
  - Internal NICs
    - Ceph traffic
    - Local jobs
  - Separate external NICs
- WNs
  - Read-only CephFS mounts on each host



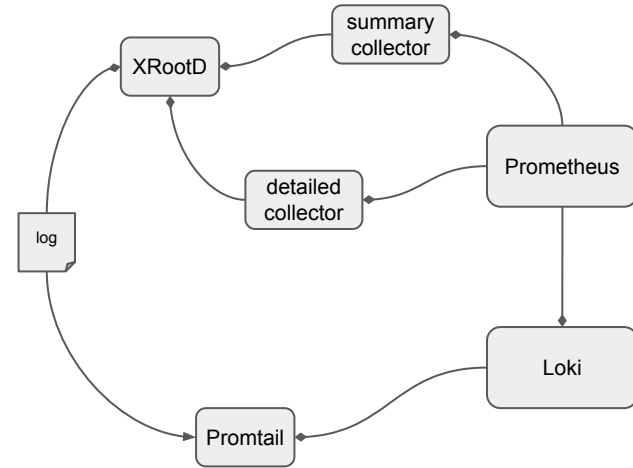
# Monitoring infrastructure at Lancs

- Prometheus
  - stores metrics
    - scrapes (pulls) periodically
    - pushable to (remote-write)
  - evaluates PromQL queries on demand
    - Results feed Grafana dashboard
- Loki
  - stores logs
    - pushed from Promtail
  - evaluates LogQL queries on demand
    - Results feed Grafana dashboard
    - Periodic evaluation generates metrics (recordings) pushed to Prometheus



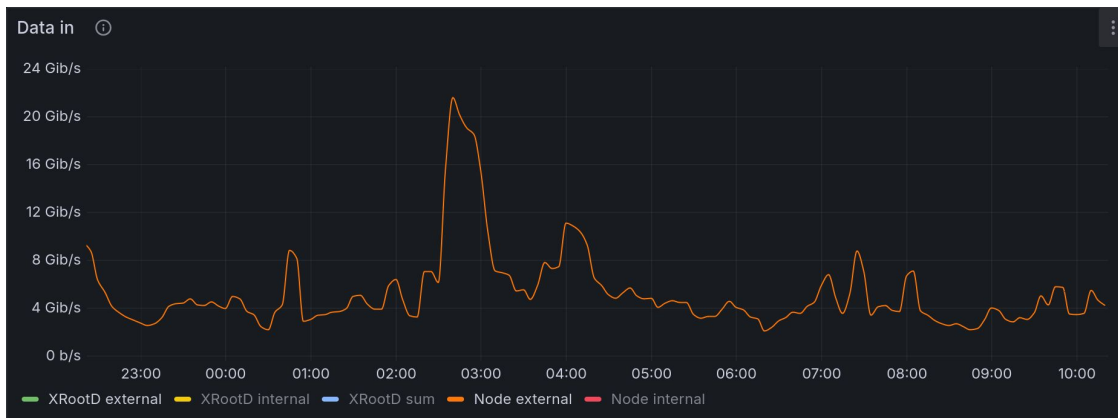
# XRootD monitoring options

- `xrd.report` (summary monitoring)
  - periodic XML/UDP summary
  - Custom Python collector converts to Prometheus remote-write message
- `xrootd.monitor` (detailed monitoring)
  - ad hoc 'binary'/UDP transfer details
    - (as used by shoveler)
  - Another custom Python collector aggregates, and turns into more detailed metrics (remote-written) and synthetic log
- Logs
  - promtail watches, and pushes to Loki
  - Loki generates metrics to write into Prometheus
- Also node exporters



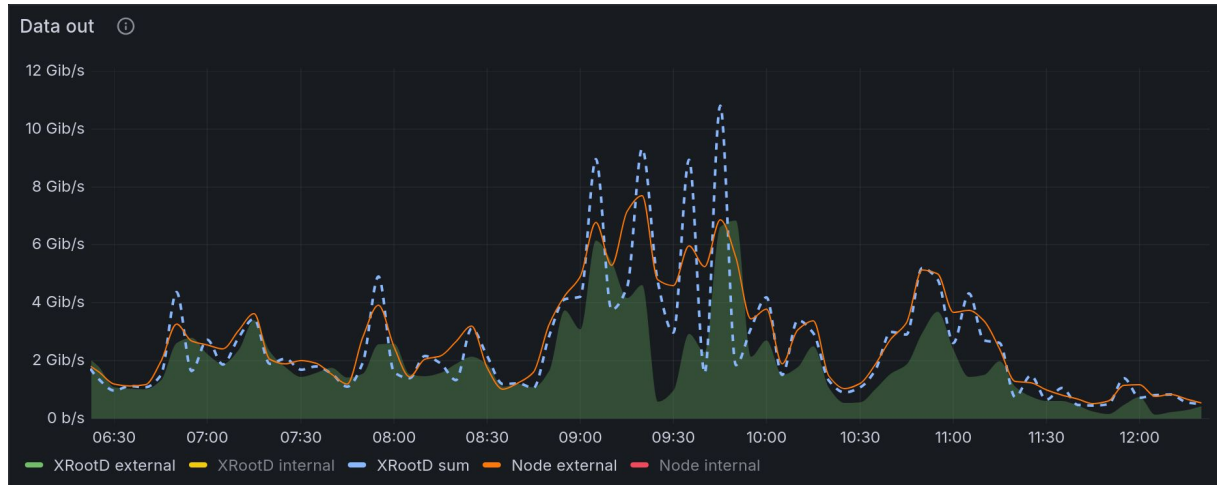
# Detailed reporting – I/O discrepancy

- External input mismatch
  - Node metrics show multi-Gib/s on external interface
  - Sum of detailed reporting for non-local clients shows almost nothing
    - Where's the green line?
- Most external data arrives by TPC pulls
  - Call to `xrdcp` contacting remote site?
  - Activity not directly seen by XRootD?
  - Stats not fed back to parent process?
    - But how does still get into the TPC g-stream?



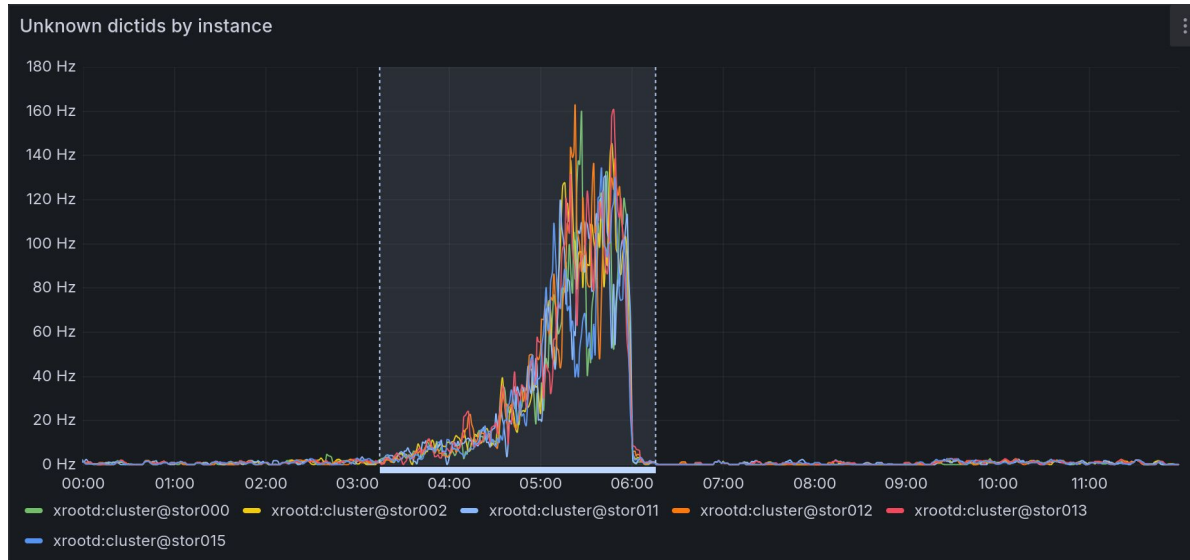
# Summary reporting – separate I/O by interface?

- Summary I/O doesn't distinguish between internal and external traffic
- Some output correlation with external (graph)
  - Needs detailed monitoring to obtain that
- Could also help to distinguish internal XRootD traffic from Ceph traffic
- (Also of interest to Glasgow)



# Detailed reporting – unresolved dictids

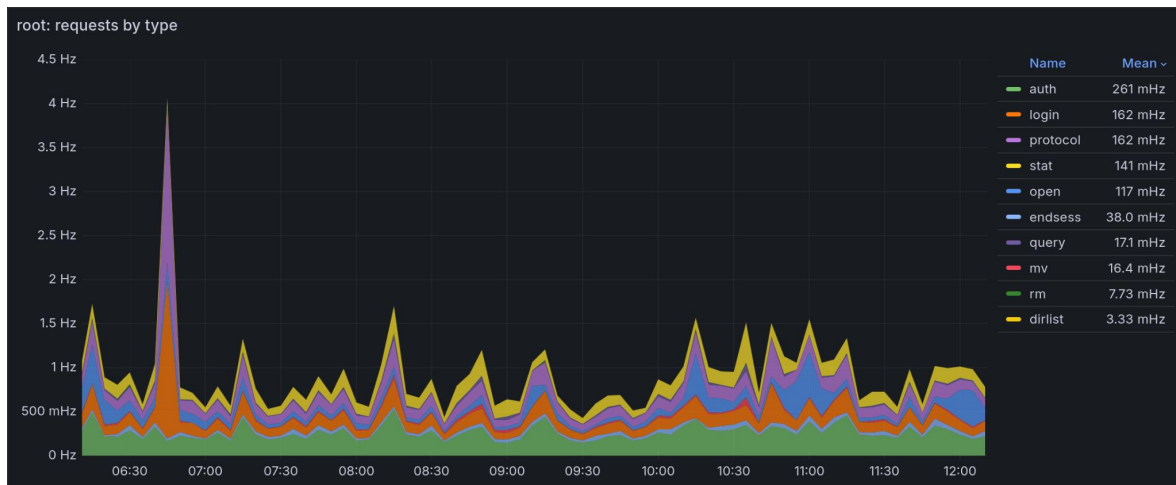
- Expect some unresolved dictids when collector starts
- Still, a low rate (<2Hz) of unresolved dictids present later
- Graph shows an increase during a rare burst of internal XRootD reads



# Log recordings

- Loki challenged by ad hoc queries
  - (We need to give it more resources)
- Recordings are periodic queries
  - Results are metrics pushed into Prometheus
  - Domain names need to be folded to prevent high cardinality
  - Derived metrics can be used in same queries as regular ones
  - Regex parsing
- Loki has several structured parsers
  - e.g., json, logfmt, ...
  - What would be most interoperable, more generally?

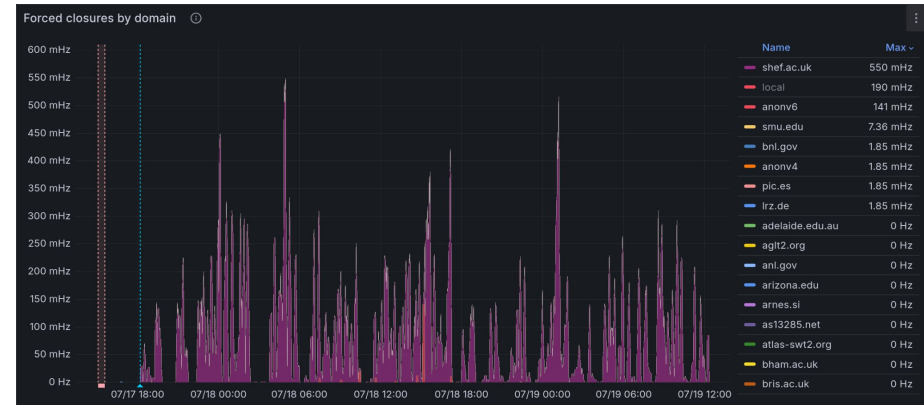
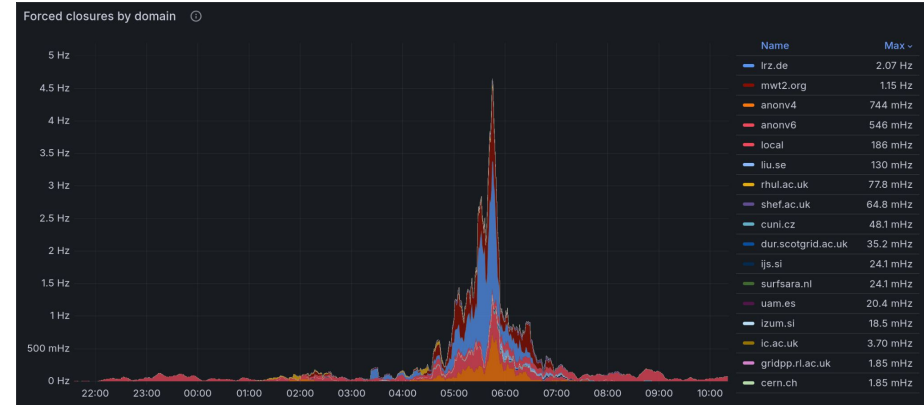
```
- record: xrootd:proto:op:domain
  expr: >-
    sum by (xrdid, pgm, user, client_domain, op)
      (label_replace(label_replace(label_replace(label_replace(label_replace(rate({job="xrootd", xrdid=~"redirector@.*"} |~ "[0-9]{6}
[0-9]{2}:[0-9]{2}:[0-9]{2} [0-9]+ (?P<user>[^\.]+\.[0-9]+:[0-9]+@(?P<client_host>[^\.]+) Xrootd_Protocol: ([0-9]+) req=' | regexp
^[0-9]{6} [0-9]{2}:[0-9]{2}:[0-9]{2} [0-9]+ (?P<user>[^\.]+\.[0-9]+:[0-9]+@(?P<client_host>[^\.]+) Xrootd_Protocol: ([0-9]+)
req=(?P<op>[^\.]+)' [5m]), "client_domain", `anonv6`, "client_host", `(.*)`, "client_domain", `anonv4`, "client_host",
`\[:ffff:([0-9]+\.[0-9]+\.[0-9]+\.[0-9]+)\]`), "client_domain", "local", "client_host", `\[:ffff:10\.*\]`), "client_domain",
"local", "client_host", `stor[^\.]+`), "client_domain", `$1`, "client_host",
`(?:[^\.]+\.[^\.]+)*?([^\.]+\.[^\.]+(?:\.edu\.au|edu\.hk|edu\.tw|gov\.pl|(?:(scotgrid\.|rl\.)?ac\.uk|ac\.cn|co\.uk|ac\.il|[a-zA-Z]+))`))
```





# Forced closures

- Detailed monitoring provides a 'forced' flag with each closure
- We can distinguish local problems from those at other sites
  - (top) local issue (slow ops on Ceph) affecting many sites
  - (bottom) issue at one remote site
- Could we get more information on why the closure was forced?
- (Also of interest to RAL)



# Wishlist

- Separate summary metrics by interface
- Increase log machine-readability
- More detail in logging
  - But not too much!
  - EWOULDBLOCK: What are you waiting for?
  - forced closures: why?
- More deletion info (volume and count)
  - Or operations in general?
- Not monitoring-related:
  - On-the-fly checksums

Thanks