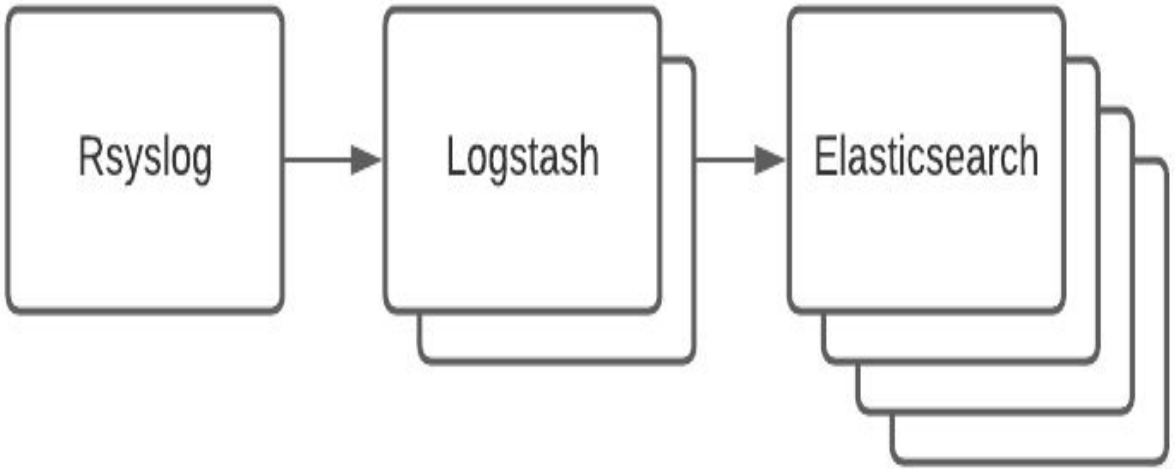
Log centralisation with rsyslog and the Elasticstack and how Sanger use it to identify issues before they are reported

Brett Hartley

wellcome sanger institute

The configuration

The layout



wellcome sanger institute

rsyslog

```
local6.info           @@logstash.example.org:10514
& ~

$WorkDirectory /var/spool/rsyslog
module(load="imfile" mode="inotify")

input(type="imfile"
 File="/var/lib/irods/iRODS/server/log/rodsLog.*"
 Tag="iRODS-rodsLog"
 Severity="info"
 Facility="local6"
)
input(type="imfile"
 File="/var/lib/irods/log/rodsLog.*"
 Tag="iRODS-rodsLog"
 Severity="info"
 Facility="local6"
)
```

wellcome sanger institute

Logstash

```
input {
  tcp {
    port => "10514"
}}
# adapted from https://www.elastic.co/guide/en/logstash/current/config-examples.html
filter {
  grok {
    match => { "message" => "%{SYSLOGTIMESTAMP:syslog_timestamp} %{SYSLOGHOST:syslog_hostname}
%{DATA:syslog_program}(?:\[%{POSINT:syslog_pid}\])?:? ?%{TIMESTAMP_ISO8601:log_timestamp}? ?%{POSINT:syslog_pid}?
?(\s%{LOGLEVEL:log_type})? %{GREEDYDATA:syslog_message}" }
    add_field => [ "received_at", "%{@timestamp}" ]
    add_field => [ "received_from", "%{host}" ]
}}
output {
  elasticsearch {
    hosts => ["elasticsearch.example.org:9200"]
    index => "logstash-%{+YYYY.MM.dd}"
}}
```

wellcome sanger institute

Elasticsearch

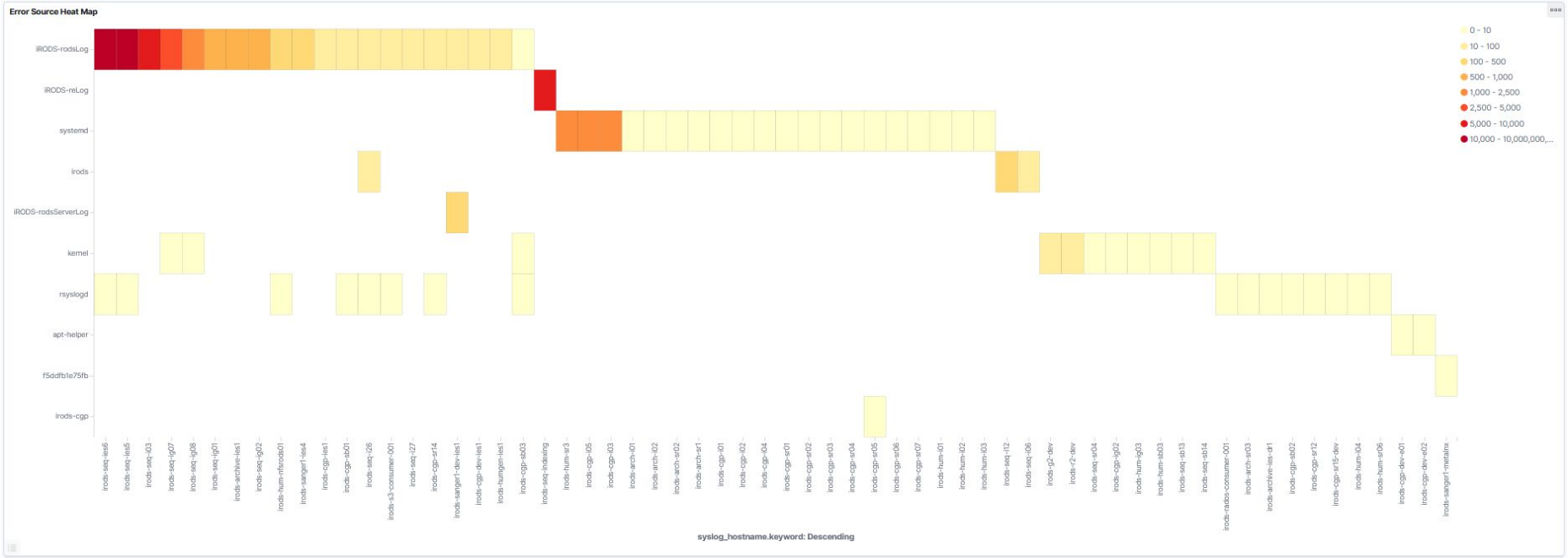
```
{
  "index_patterns" : ["logstash-*"],
  "settings" : {
    "index" : {
      "number_of_shards" : "5"
  }},
  "mappings" : {
    "dynamic_templates" : [{
      "string_fields" : {
        "mapping" : {
          "norms" : false,
          "type" : "text",
          "fields" : {
            "keyword" : {
              "ignore_above" : 256,
              "type" : "keyword"
        }}},
        "match_mapping_type" : "string",
        "match" : "*"
    }}],
    "properties" : {
      "@timestamp" : {
        "type" : "date"
  }}},
  "aliases" : { }
}
```

wellcome sanger institute

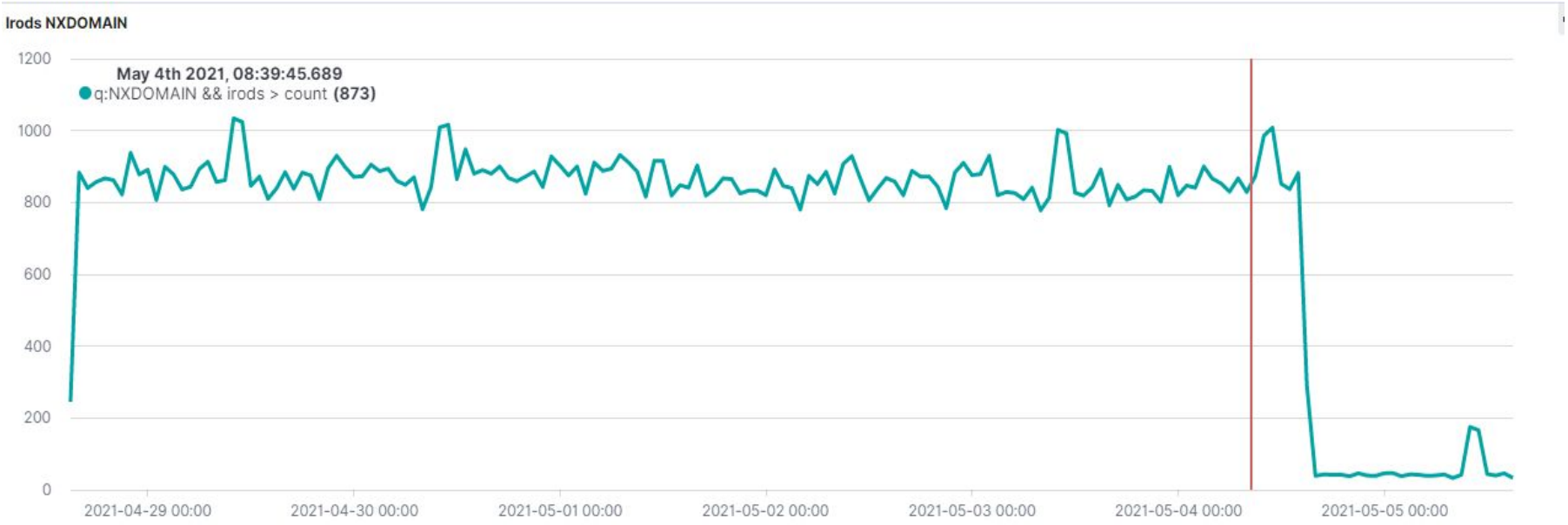
So what’s the benefit?

wellcome sanger institute

Picking out trends



An example - DNS



An example - file mismatches

object path	mismatch type	file path
/seq-dev/home/bh9#Sanger1-dev/chksum	size	/irods-seq-i16-de/home/bh9#Sanger1-dev/chksum
/seq-dev/home/bh9#Sanger1-dev/iputrecursivetest/iputrecursivetest/iput-test-12	size	/irods-seq-sr02-ddn-rd10-0-1-2/home/bh9#Sanger1-dev/iputrecursivetest/iputrecursivetest/iput-test-12
/seq-dev/home/bh9#Sanger1-dev/iputrecursivetest/iputrecursivetest/iput-test-14	size	/irods-seq-sr02-ddn-rd10-6-7-8/home/bh9#Sanger1-dev/iputrecursivetest/iputrecursivetest/iput-test-14
/seq-dev/home/bh9#Sanger1-dev/iputrecursivetest/iputrecursivetest/iput-test-5	size	/irods-seq-sr02-ddn-rd10-9-10-11/home/bh9#Sanger1-dev/iputrecursivetest/iputrecursivetest/iput-test-5
/seq-dev/home/bh9#Sanger1-dev/thisiswrong	size	/irods-seq-i15-de/home/bh9#Sanger1-dev/thisiswrong
/seq/home/bh9#Sanger1/thisiswrong	size	/irods-seq-i23-fg/home/bh9#Sanger1/thisiswrong
/seq-dev/home/bh9#Sanger1-dev/24174_5#888.cram	checksum	/irods-seq-sr02-ddn-rd10-12-13-14/home/bh9#Sanger1-dev/24174_5#888.cram

<182>Jun 4 12:30:50 irods-seq-dev-ies1 iRODS-rodsLog Jun 4 12:30:50 pid:4146 NOTICE: procChksumForClose: chksum mismatch for /seq-dev/home/bh9#Sanger1-dev/24174_5#888.cram src [82ee353d0e8ce95baa4d55df85b3b945] new [d41d8cd98f00b204e9800998ecf8427e]

wellcome sanger institute