



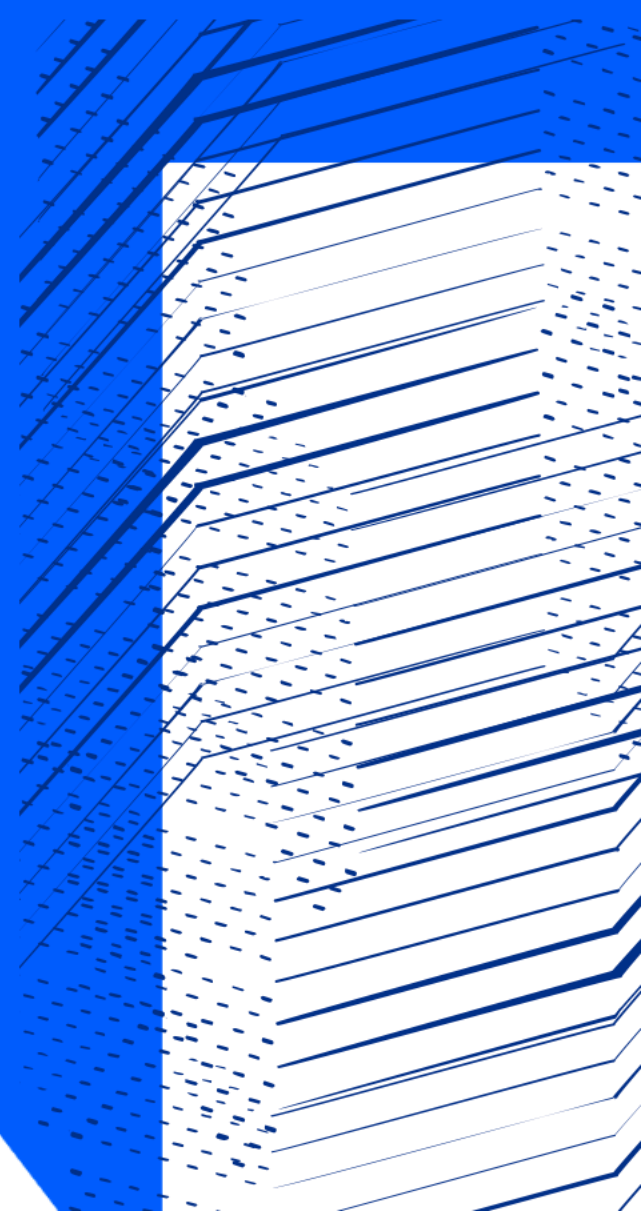
Science and  
Technology  
Facilities Council

# Evolving Monitoring and Operations services at RAL

Christos Nikitas

Will Furnell

Adam Laverack



# About me

- My name is Christos
- Studying Computer science at the University of Manchester.
  - Completed 2 Years Undergraduate work, and now doing a year in industry at STFC
- Systems Administrator in Tier 1 Production team at RAL
- Currently working remotely from Cyprus🌞.
  - Started work here in July so I have never gone to RAL to work.



# Motivation

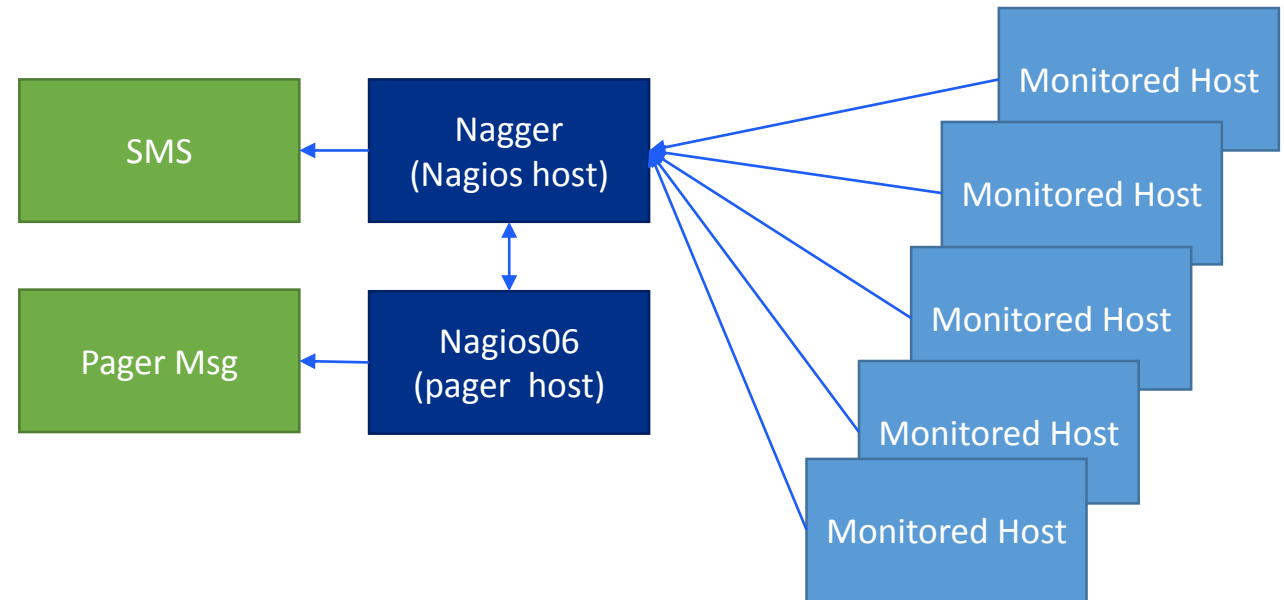
- There are an ever-increasing number of services running.
  - Increasing service level expectations from users.
- Our operations and monitoring services are becoming outdated and cannot meet up with the demands.
  - Our time series monitoring service was still running with spinning disks, which made query speeds very slow

# Changes Overview

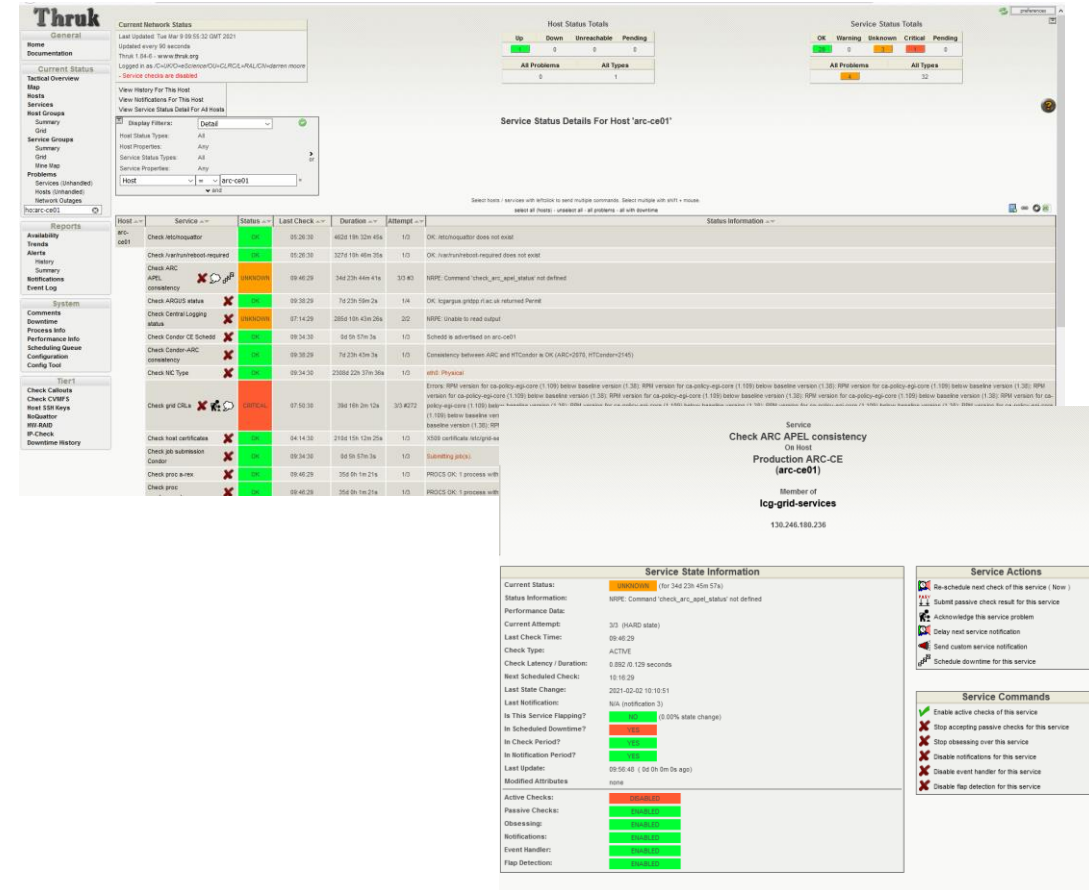
- Nagios to Icinga
  - Alert Monitoring
- Chenda -> OpsGenie
  - Callout System
- InfluxDB
  - Time Series Database
- RT -> JIRA
  - Project management and Issue Tracking

# Alert Monitoring

- Nagios – Current Alerting Tool.
  - Perform periodic checks on hosts, alerting if something is wrong
- Provides SMS and Pager notification.
- Provides alerting support for the Tier-1 On-call (Production) team.



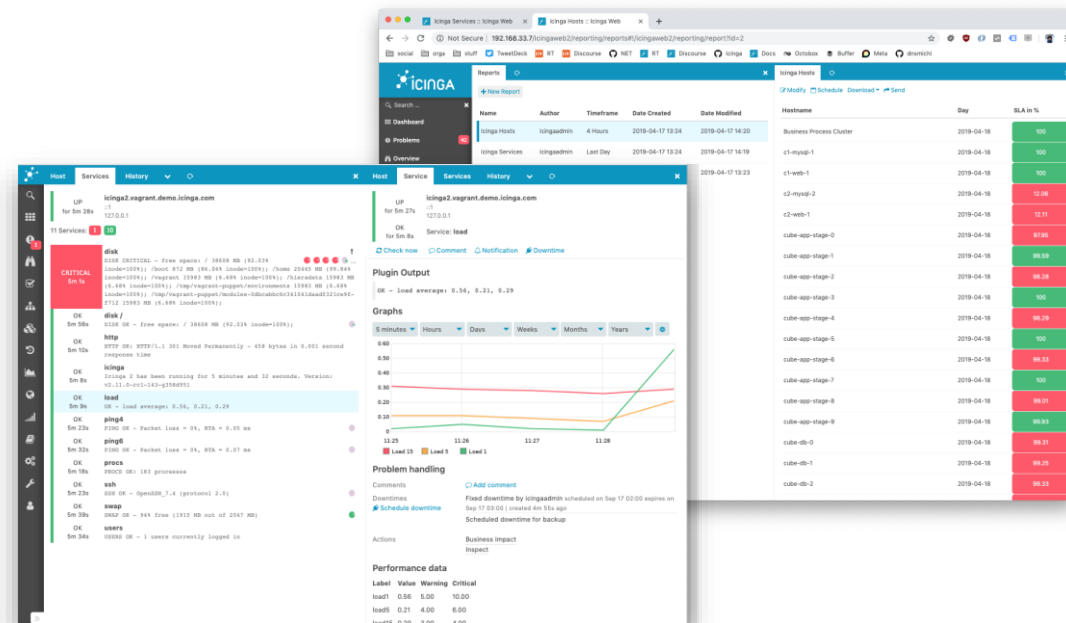
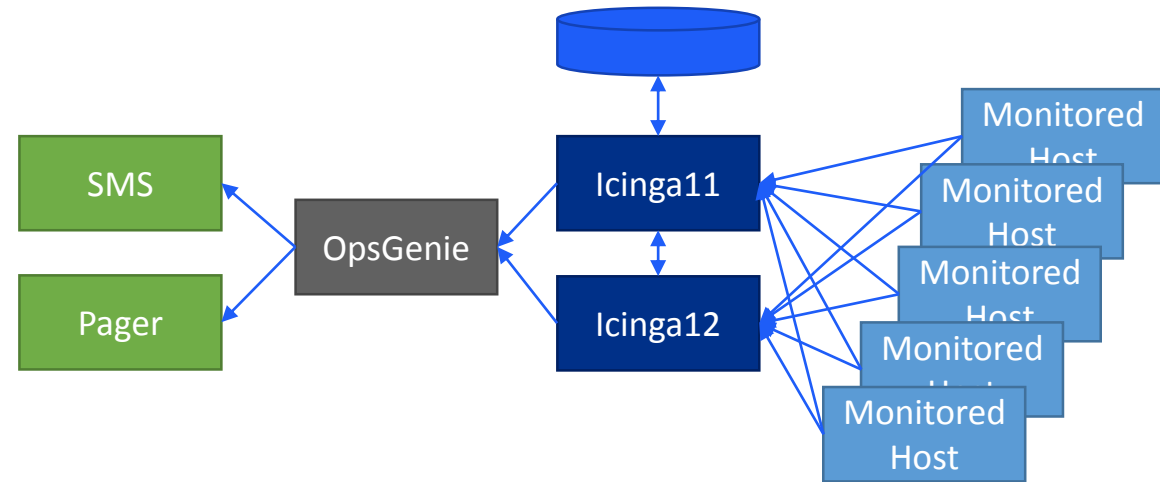
- VERY old version - Nagios(v3.5.1) released in 2013. (This is when iPhone 5s first introduced Touch ID and the galaxy s4 came out)
- Hardware of similar vintage (probably uncomfortably close to self-decommissioning)
- Single point of failure for alerting (runs on a single host)





# Icinga

- **Not antiquated!** (latest version came out in 2020)
- Integration with SCD configurations management system (Aquilon)
- Consolidation of multiple **services** (Provides a single SCD wide available alerting service)
- Consolidation of Functionality (Removal of “check creep”/duplication)
- HA cluster for reliability (No single point of failure)
- OpsGenie providing SMS/Pager alerting



# On Call System

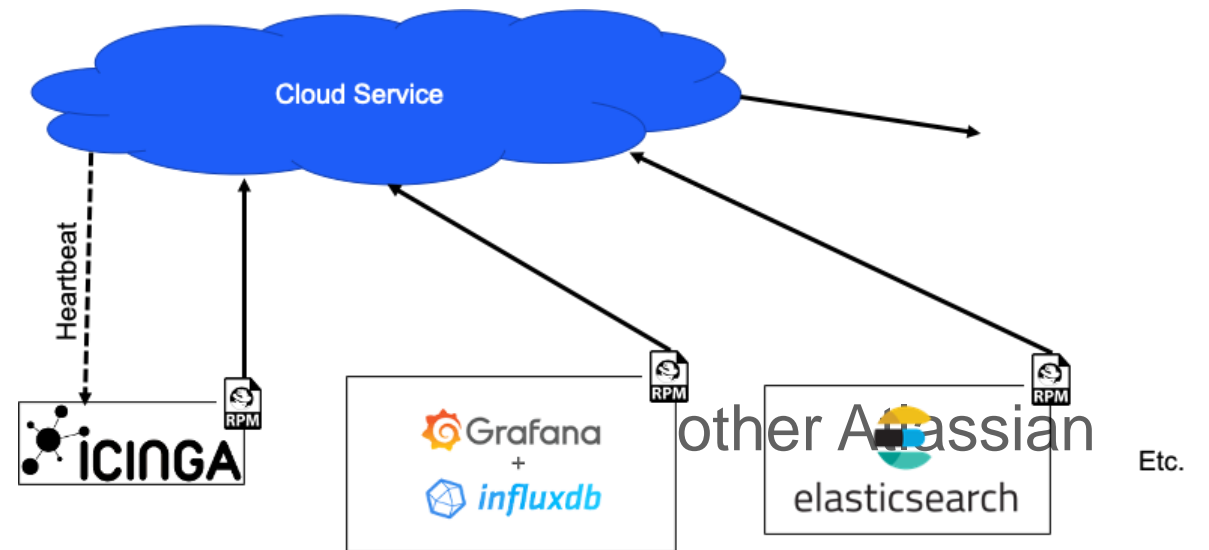
- Our Callout system needs to be integrated with Icinga
  - Sends SMS, emails, pager notifications whenever an alert pops up in icinga.
- Nagios's bespoke call-out functionality (Chenda[1]) is being replaced.
  - It was written by us and is difficult to maintain.
  - Involves a lot of scripting and going into individual files to get it working.
  - Difficult to use/change - mutated rather than evolved!
  - Seriously lacking in features

[1] The Chanda is a cylindrical percussion instrument originating in the state of Kerala and widely used in Tulu Nadu of Karnataka and Tamil Nadu in India.



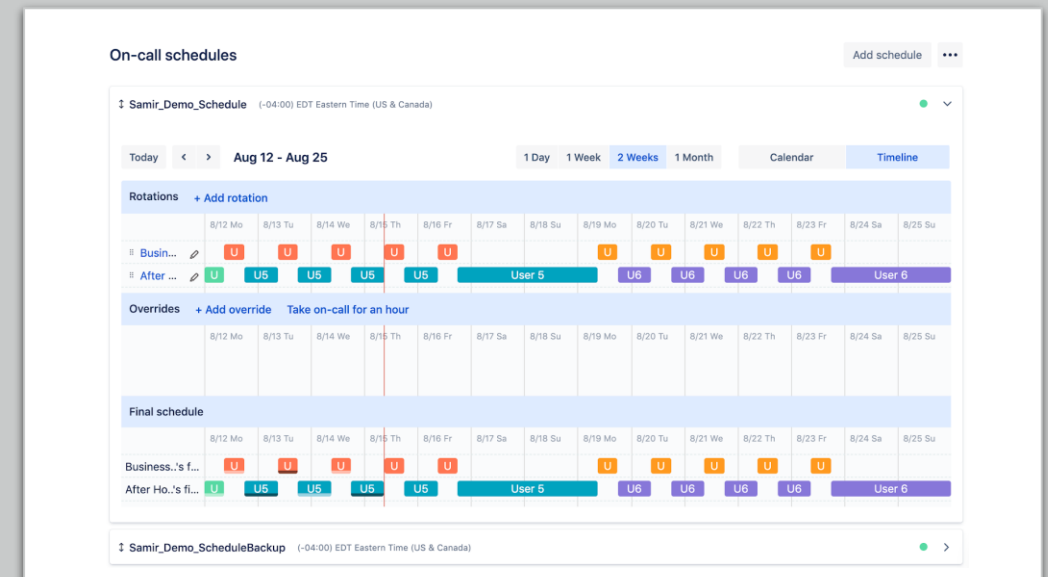
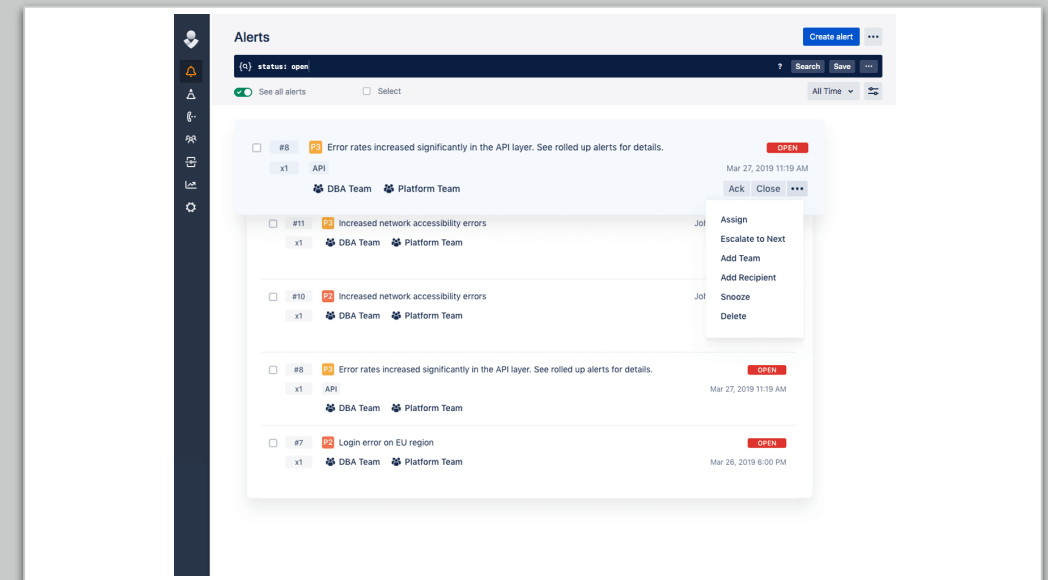
# On Call Evaluation

- In Q4 2020, we did an evaluation of cloud based on call systems.
  - OpsGenie
  - PagerDuty
- Both were very capable.
  - OpsGenie was cheaper
  - Potentially better integrated services



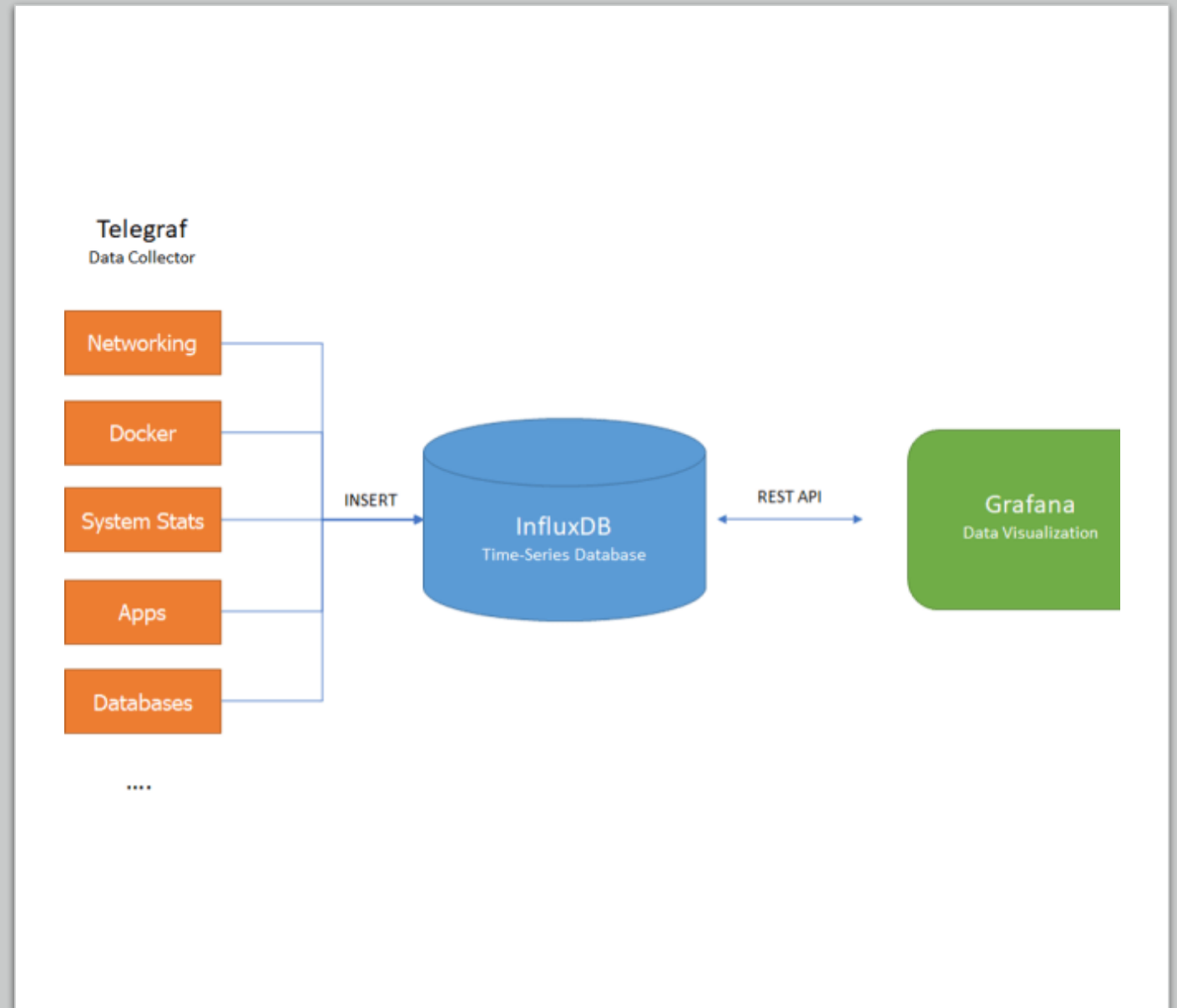
# OpsGenie

- Provides a nice UI
- Can implement escalation policies
- On-call Rota
- Can integrate with many services(Slack, Jira)
- Mobile app
- Actually cheaper than the previous pager system.



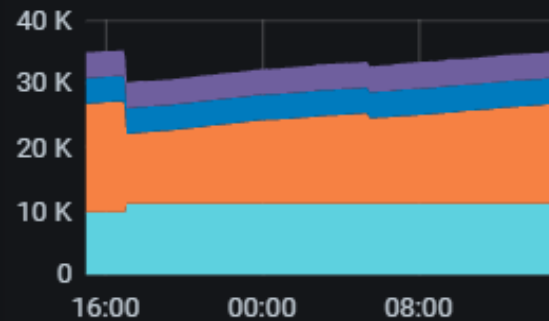
# InfluxDB

- InfluxDB is our Time Series Monitoring Database
- Used as part of the TIG stack
  - **Telegraf**: Collects metrics and sends them to influx
  - **InfluxDB**: Stores the metrics sent out by telegraf
  - **Grafana**: Displays the metric data in a nice UI. Can also make alerts.

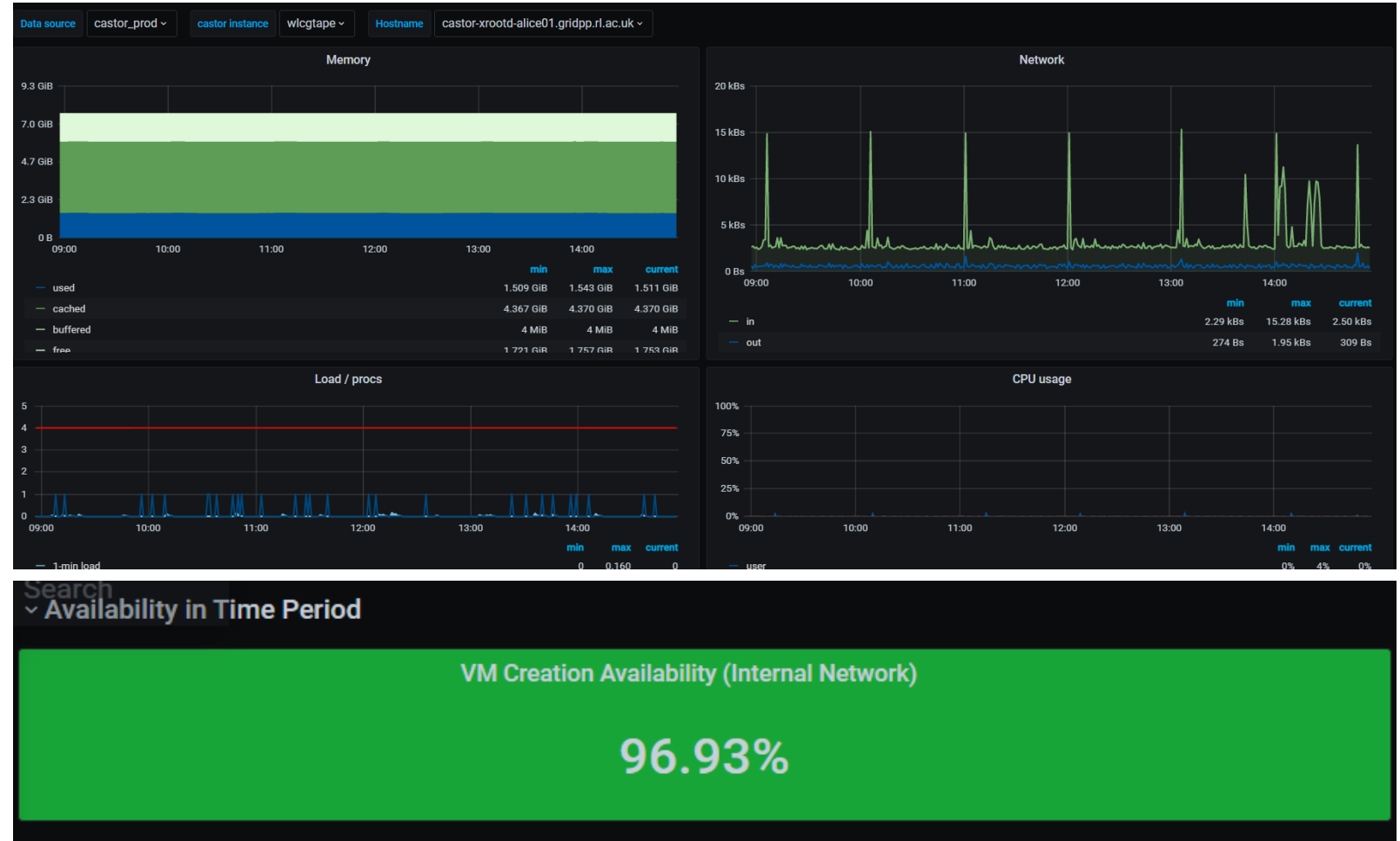


arc-ce01

Jobs by ARC status



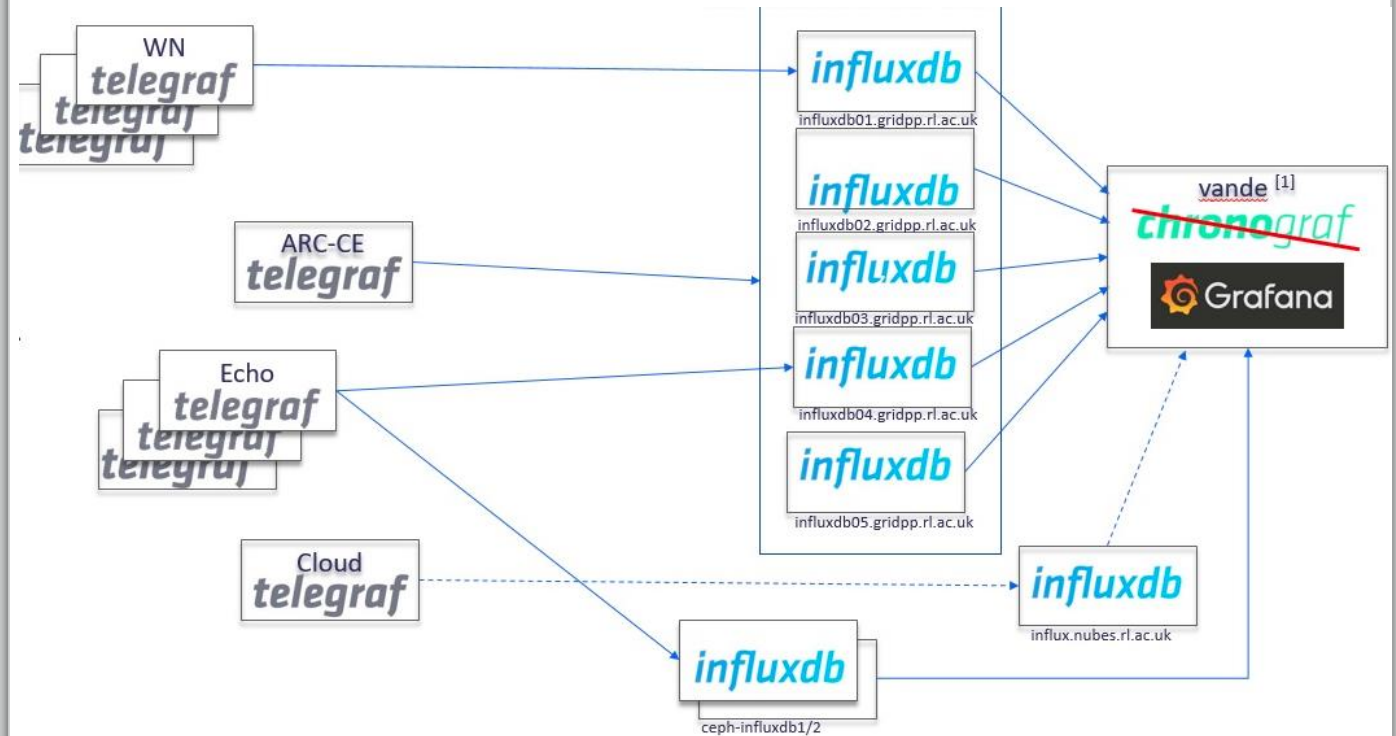
— ACCEPTED Current: 0  
— CANCELLING Current: 0  
— DELETED Current: 11.18 K



# Some Examples

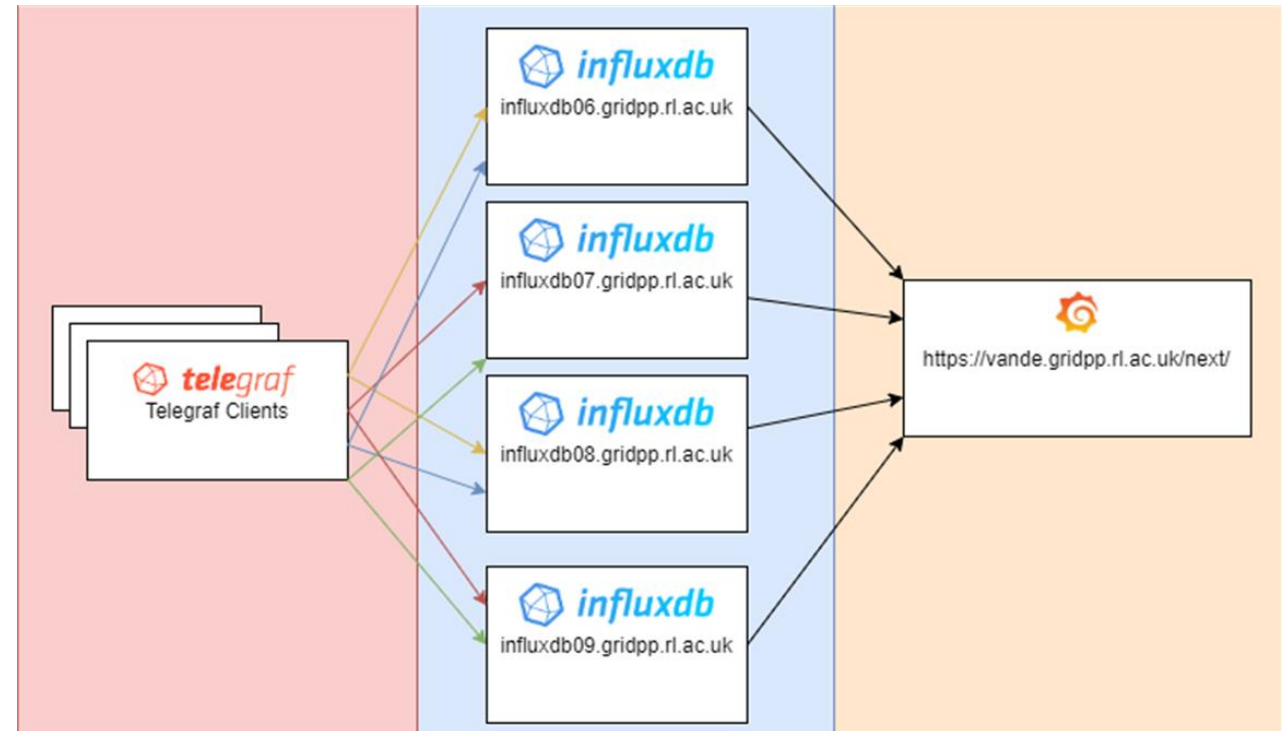
# InfluxDB

- Up until recently we had 5 influxdb servers running on “influxdb{01-05}.gridpp.rl.ac.uk”.
  - Every database was only stored once in the hosts.
  - Some hosts had 80GB worth of data while others had 3GB
  - Some hosts had 200GB storage capacity while others had 1TB
  - Some hosts use spinning disk drives



# InfluxDB

- We are replacing the old servers with 5 new ones. “influxdb{06-10}.gridpp.rl.ac.uk”.
  - Store every database twice to increase high availability
  - Each has 900GB of storage
  - All hosts use SSDs
  - Data is spread out more evenly



# InfluxDB

- Another major benefit from this migration is the change of retention policies and data granularity. (How long data is kept and its resolution)
  - We used to store data at 1-minute intervals for 31 days.
  - Now we keep data like this
    - raw data for 24h
    - 1m summary for 31d
    - 10m summary for 6months
    - 60m summary for 5 years
  - More accurate metrics and stored for a much greater amount of time.



# RT -> JIRA

- STFC Cloud team did an evaluation of ticket systems.
- RT wasn't able to cater for their growing use case:
  - Serviceable, but becoming outdated
    - Limited metrics production
    - No feedback functionality
  - No Integration with tools like Slack, JIRA, Confluence
  - Meh UI
- JIRA Service Desk was chosen
  - On premise license currently for 25 users (with academic discount)
  - Other services (e.g. Tier-1) will be migrating in 2021.

RT at a glance

New ticket in

3d

Search...

Edit

Bookmarked Tickets

Edit

My reminders

10 highest priority tickets I own

Edit

10 newest unowned tickets

Edit

#	Subject	Queue	Status	Created	
398677	UPS: Passed a self-test.	Fabric	new	11 minutes ago	Take
398676	UPS: Started a self-test.	Fabric	new	12 minutes ago	Take
398675	[ral-tier1][PROBLEM] ha-icinga on xrootd-cms-uk.gridpp.rl.ac.uk is WARNING!				
398674	[ral-tier1][PROBLEM] ha-load db server on dbadb11.fds.rl.ac.uk is CRITICAL!				
398673	[ral-tier1][PROBLEM] ha-load db server on dbadb11.fds.rl.ac.uk is WARNING!				
398672	[ral-tier1][PROBLEM] check_service ncm-cdispd on xrootd-cms-uk.gridpp.rl.ac.uk is CRITICAL!				
398671	Nagios issued service alarm Tier1_service_OPN_CERN_Link_Cluster_check_on_host_naggar [NoCallout]				
398669	icinga issued and cleared ha-check_castor_ns_showq_down on fac-central-services01.fds.rl.ac.uk				
398667	icinga issued and cleared ha-check_castor_ns_showq_down on fac-central-services01.fds.rl.ac.uk				
398666	[ral-tier1][PROBLEM] ha-load db server on dbadb11.fds.rl.ac.uk is CRITICAL!				

Quick ticket creation

Subject:

Queue: 3d Owner: Me

Requestors: christos.nikitas@stfc.ac.uk

Content:

Create

Queue list

Edit

Queue	new	open	stalled
3d	-	-	-
Castor	21	6	-
Ceph	20	31	-
Ceph-Disk	99	81	3
Ceph-Tasks	1	4	6
Certificate Requests	-	-	-

Create a new ticket in Production

Requestors: christos.nikitas@stfc.ac.uk

Cc:

(Sends a carbon-copy of this update to a comma-delimited list of email addresses. These people will receive future updates.)

Admin Cc:

(Sends a carbon-copy of this update to a comma-delimited list of administrative email addresses. These people will receive future updates.)

Subject:

Sign ☐ using Queue's key Encrypt ☐

Describe the issue below:

Attach:

Drop files here or click to attach



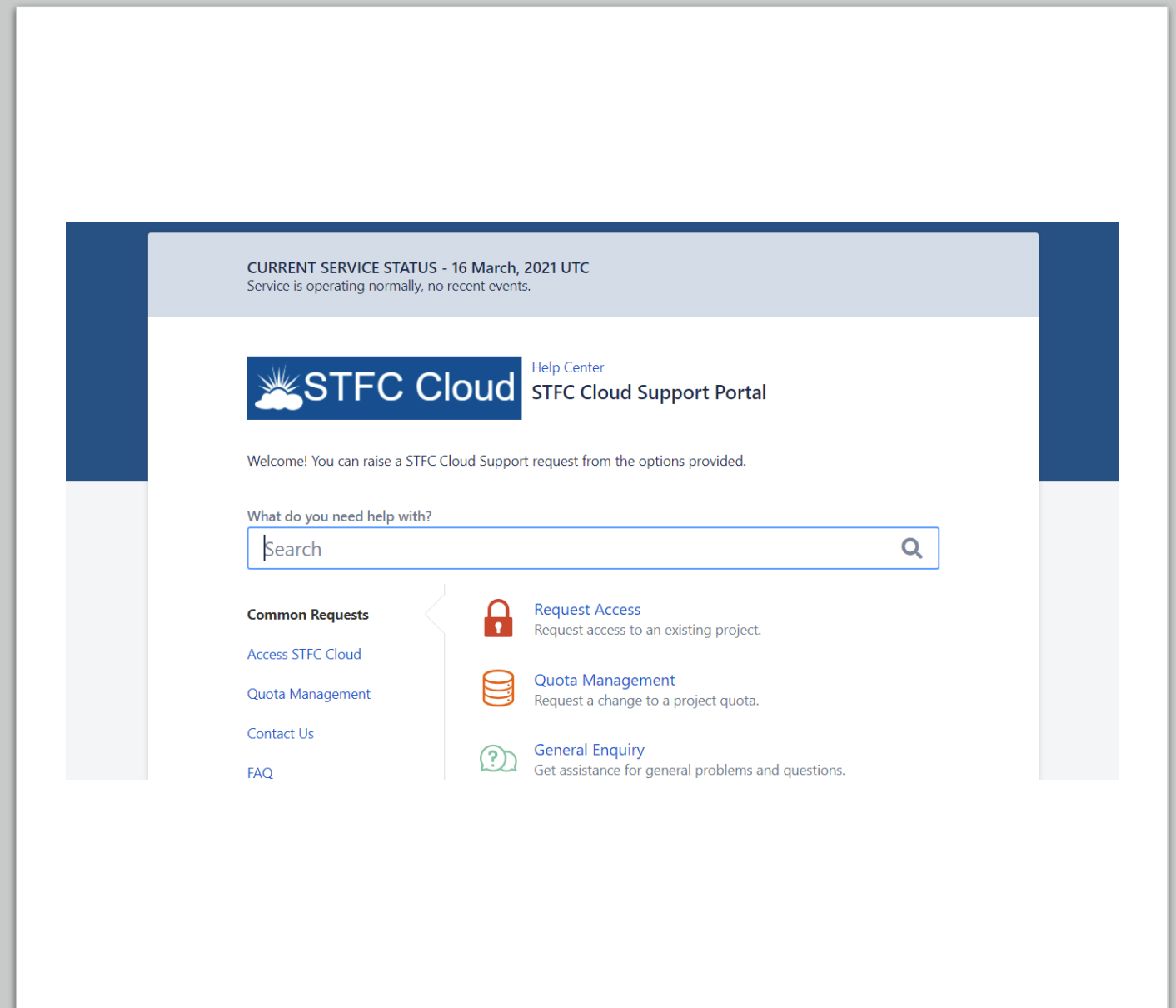
Science and  
Technology  
Facilities Council

Create



# JIRA Service Desk

- Nicer / More modern UI (Very easy to use)
- Templated tickets
- Integrated Knowledgebase
- Integration with Slack, JIRA and Confluence





Science and  
Technology  
Facilities Council

# Thank you



Science and Technology Facilities Council



@STFC\_matters



Science and Technology Facilities Council