Introduction to the Global INternet Observatory Platform

Tim Betzer

Wednesday 12th February, 2025

Chair of Network Architectures and Services
School of Computation, Information, and Technology
Technical University of Munich

Introduction

- GINO Research group at the Chair of Network Architectures and Services under Prof. Dr.-Ing. Carle
- We perform Internet-wide DNS, TLS, QUIC and port scans for common protocols [1]
- We publish an IPv6 hitlist service [2]
- Collect Internet-related data sets

[1] https://net.in.tum.de/projects/gino/
[2] https://ipv6hitlist.github.io/

Introduction

- GINO Research group at the Chair of Network Architectures and Services under Prof. Dr.-Ing. Carle
- We perform Internet-wide DNS, TLS, QUIC and port scans for common protocols [1]
- We publish an IPv6 hitlist service [2]
- Collect Internet-related data sets
- Large amount of data available for analysis but
  - as files on our storage servers
  - in lots of different file formats and complex folder structures
- Complicates Cooperation

[1] https://net.in.tum.de/projects/gino/
[2] https://ipv6hitlist.github.io/

Challenges

Time

Long setup times until one can start with research

Privacy

Data contains sensitive, private information

Restrictions

License restrictions may restrict sharing

Fine-grained access control of the data is essential for data sharing

Challenges

Time

Long setup times until one can start with research

Privacy

Data contains sensitive, private information

Restrictions

License restrictions may restrict sharing

Fine-grained access control of the data is essential for data sharing

$\Rightarrow$ GINO Platform

Goals of the GINO platform

- Allow external researchers and students to quickly start working with the available data they are allowed to access
- Create a platform for easy and quick analysis of the available data

Goals of the GINO platform

- Allow external researchers and students to quickly start working with the available data they are allowed to access
- Create a platform for easy and quick analysis of the available data

Proposed Solutions:

1. ASQ-GINO

Answering pre-defined queries

2. General purpose data access

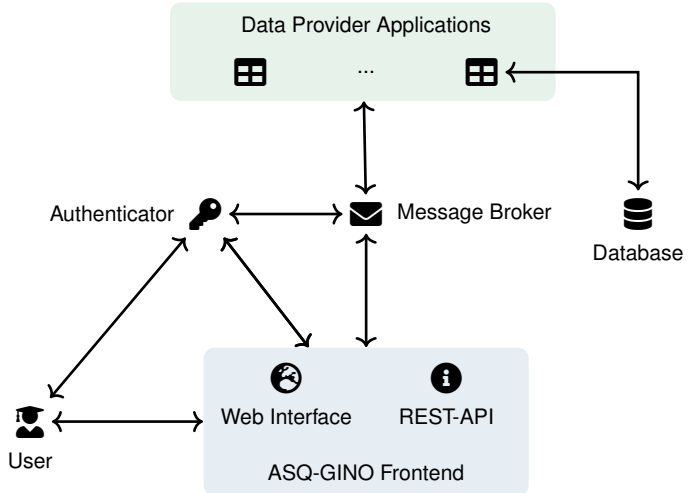
Access to data provided on general-purpose computation infrastructure

ASQ-GINO: Answering Pre-Defined Queries

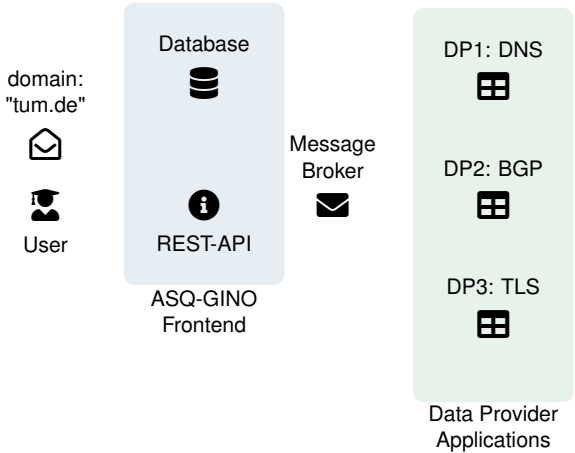
- Quick Overview of available data
- No setup required
- No Knowledge required

ASQ-GINO: Answering Pre-Defined Queries

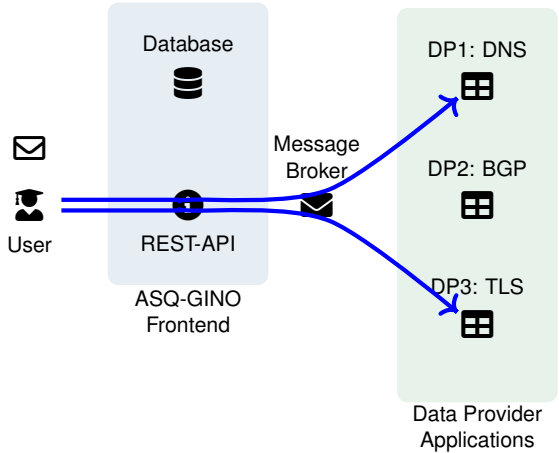
- Quick Overview of available data
- No setup required
- No Knowledge required



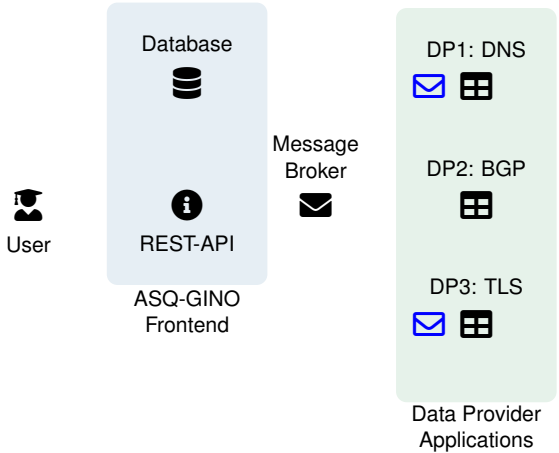
ASQ-GINO: Answering Pre-Defined Queries



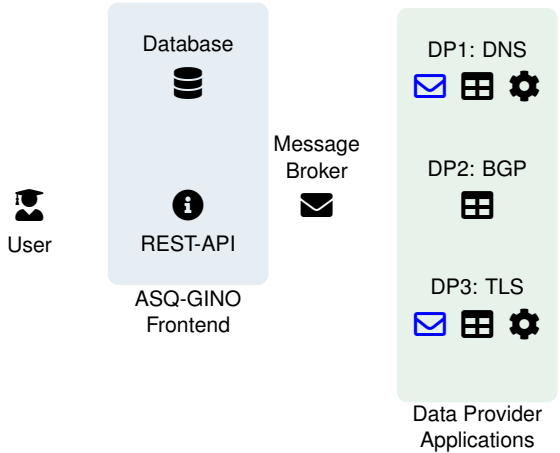
ASQ-GINO: Answering Pre-Defined Queries



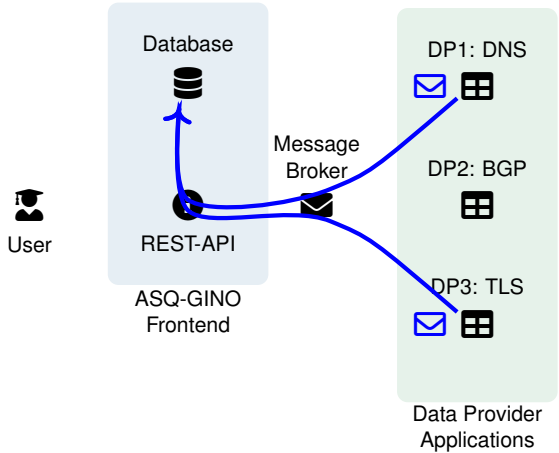
ASQ-GINO: Answering Pre-Defined Queries



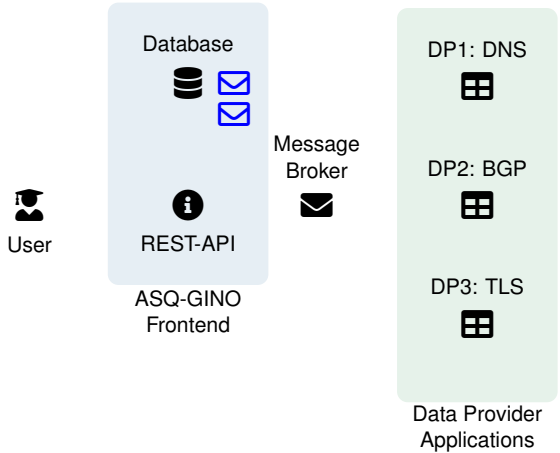
ASQ-GINO: Answering Pre-Defined Queries



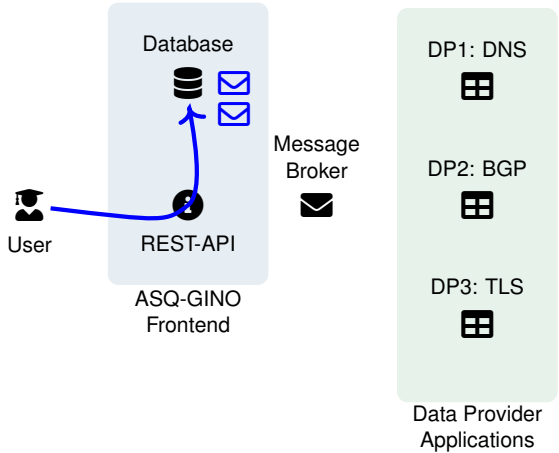
ASQ-GINO: Answering Pre-Defined Queries



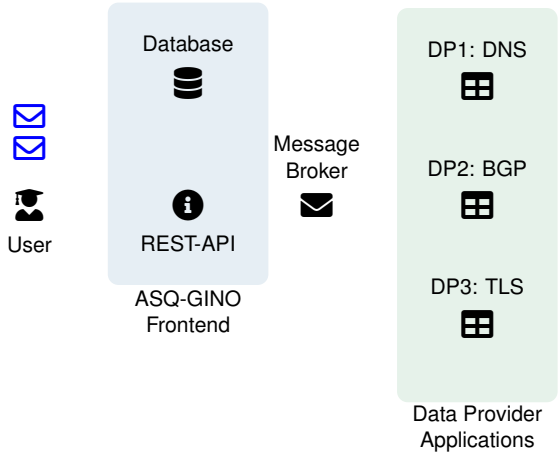
ASQ-GINO: Answering Pre-Defined Queries



ASQ-GINO: Answering Pre-Defined Queries

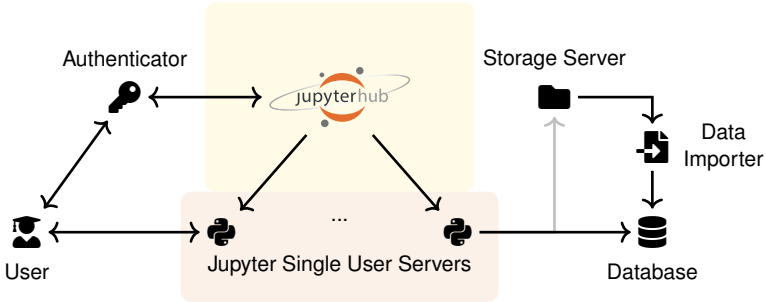


ASQ-GINO: Answering Pre-Defined Queries



General purpose data access

- Allow broader and more general access to raw data
- Keeping the setup easy to use and maintain

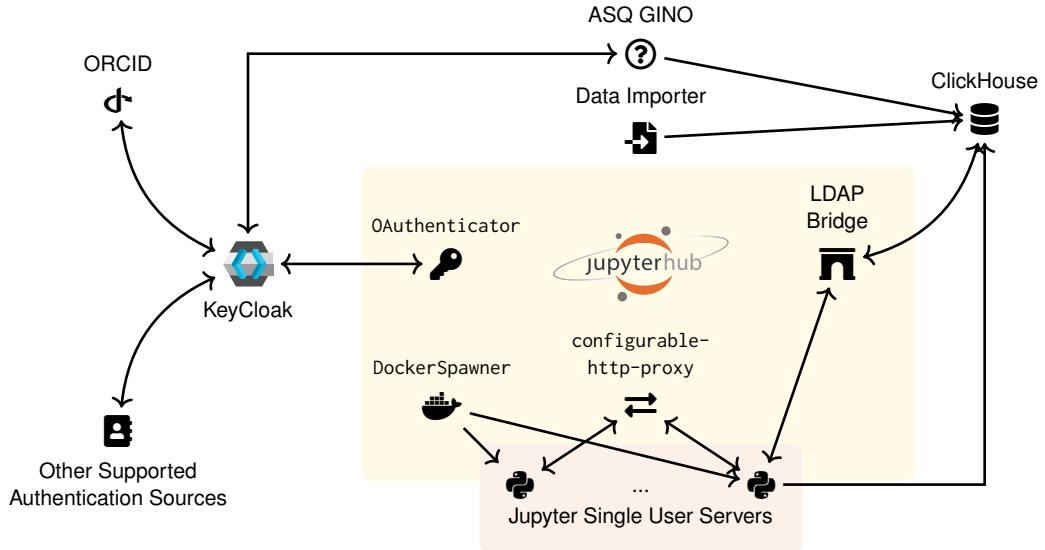


Next Steps

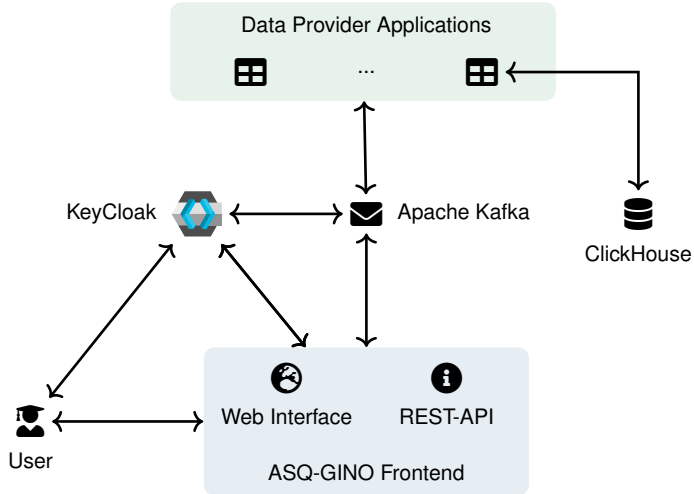
- Currently in the setup phase.
- Most components are available.
- Now moving to the hardware step.

More information on project GINO can be found under https://net.in.tum.de/projects/gino/

Appendix


ASQ GINO
ORCID
Data Importer
ClickHouse
LDAP
Bridge
OAuthenticator
jupyterhub
KeyCloak
configurable-
http-proxy
DockerSpawner
Other Supported
Authentication Sources
...
Jupyter Single User Servers

Appendix


Data Provider Applications
...
KeyCloak
Apache Kafka
ClickHouse
Web Interface
REST-API
User
ASQ-GINO Frontend