

Lookup Service

Moving to elasticsearch and beyond...

Sowmya Balasubramanian | ESnet

2nd European perfSONAR User Workshop | Apr 15, 2021

perfSONAR is developed by a partnership of

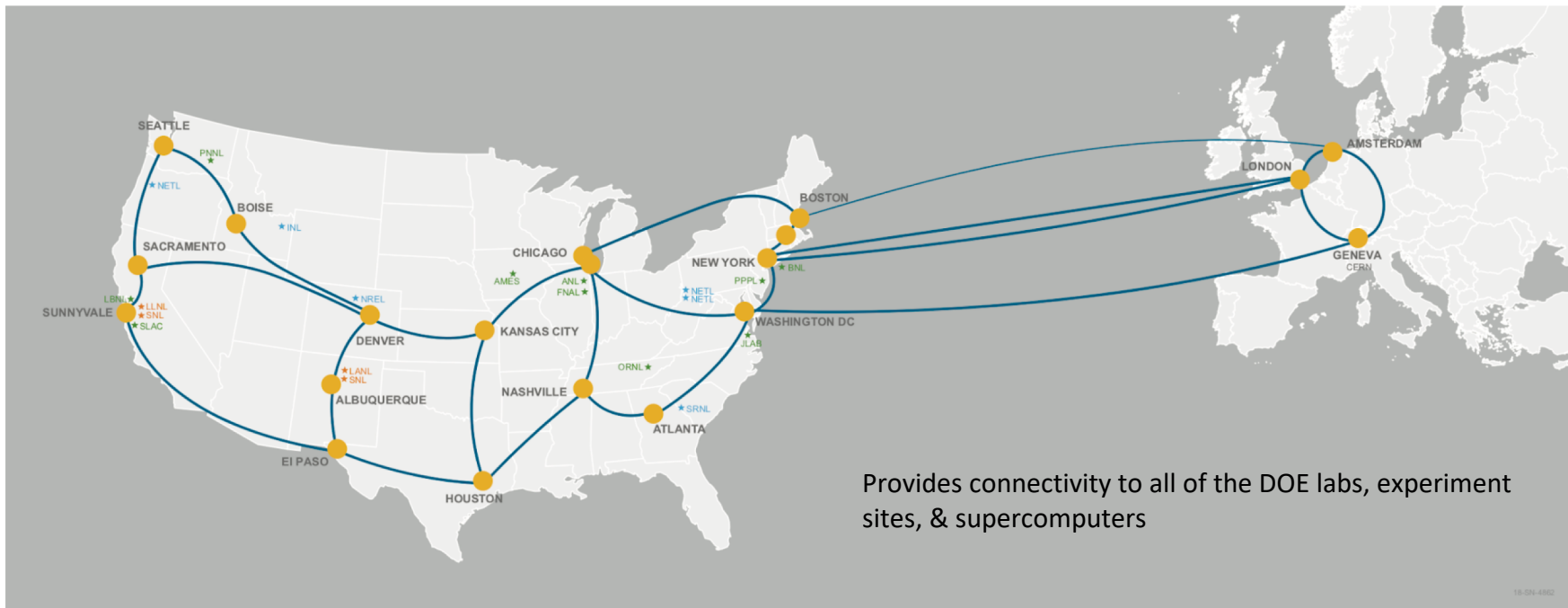


ESnet



ESnet

DOE's high-performance network (HPN) user facility optimized for enabling big-data science

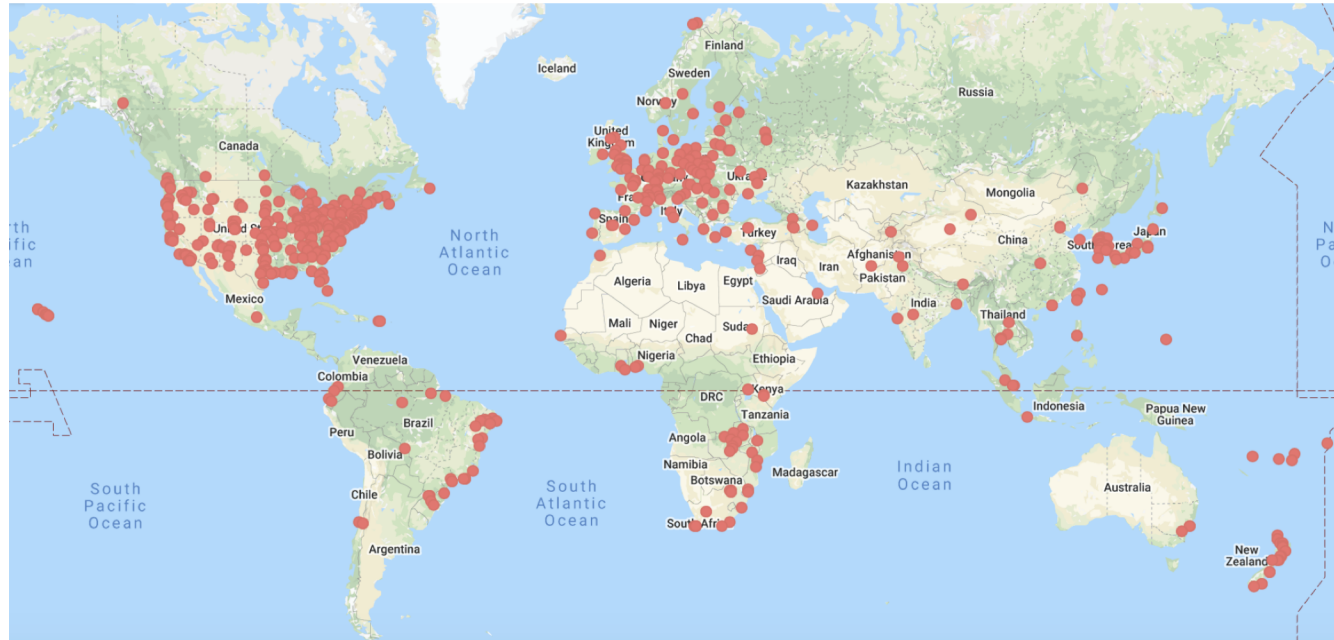


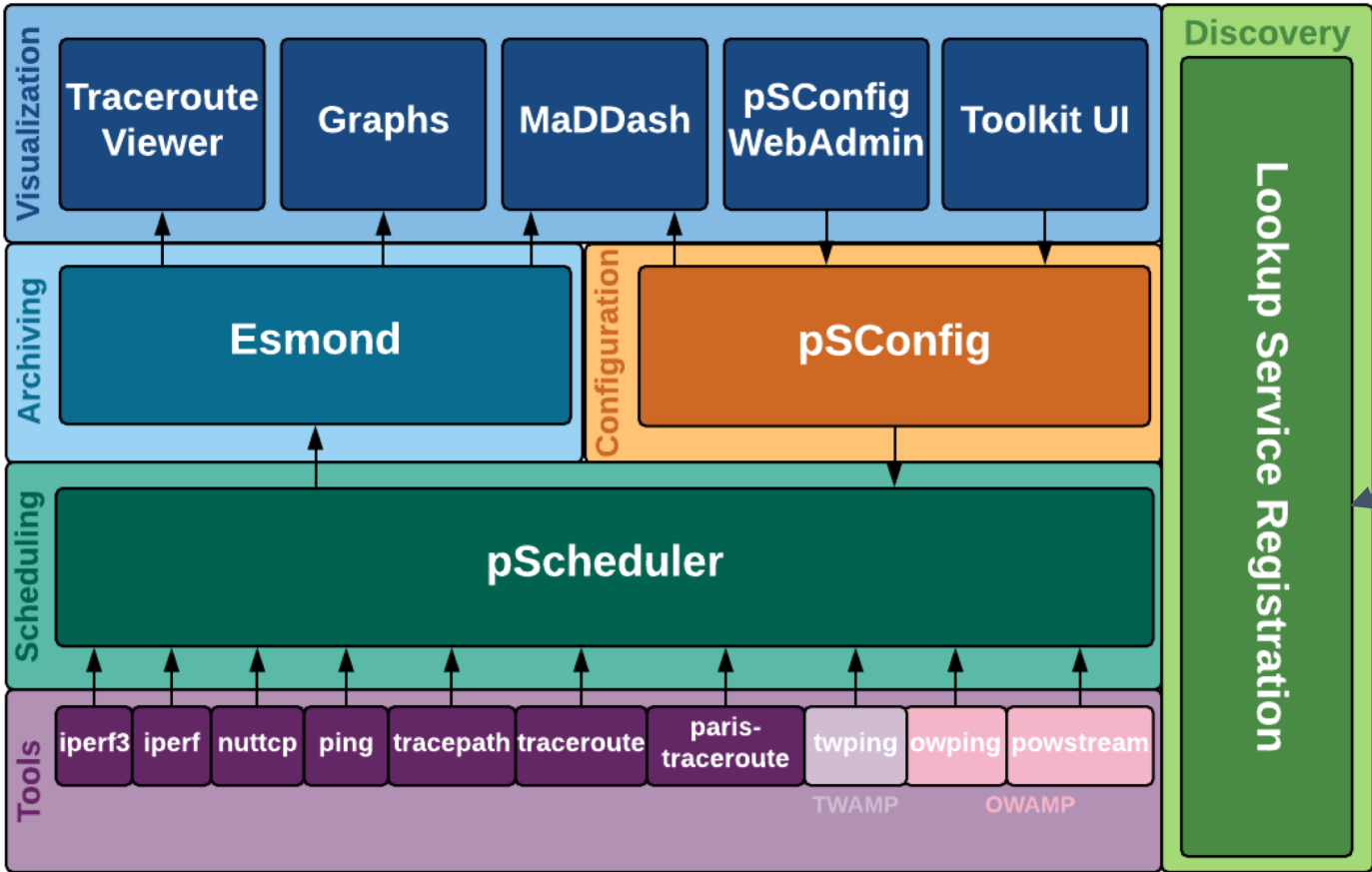
Provides connectivity to all of the DOE labs, experiment sites, & supercomputers

18-0N-4802



What is perfSONAR Lookup Service?





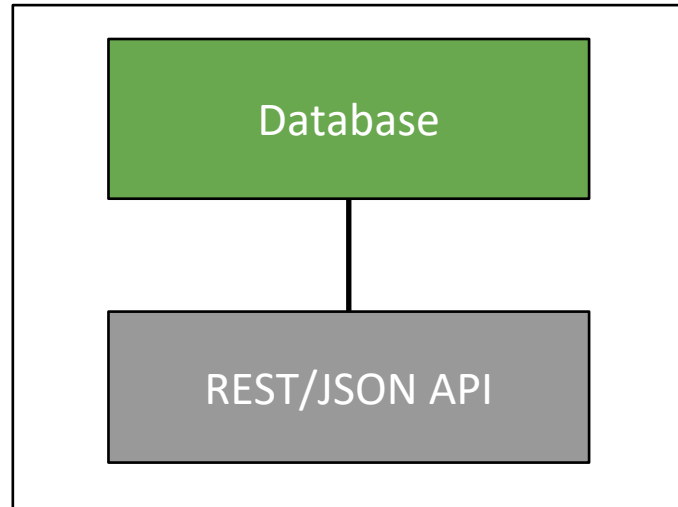
Data visualized in LS Service Directory

<https://stats.es.net/ServiceDirectory/>

Lookup Service Overview

Elasticsearch

Java



Runs in a VM in Google Cloud

Over the years....



2013

Simple Lookup Service

REST/JSON based lookup service was released. It had MongoDB as backend.

Few Instances deployed across the world. Deployment model was similar a gLS that existed in previous version



2018

Moved to Google cloud

Consolidated into a single instance



2020

Elasticsearch backend

Replaced MongoDB with elasticsearch backend.



2021

...

Data Validation, Improved APIs and more...

Data Validation using schemas,

Improved APIs - better query performance, improve analysis capabilities..



LS Requirements Revisited - [Link](#)

Data & APIs

Improve Query APIs, Updated registration , query response times, data volume to support, Data validation using schemas

Monitoring and Alerting

Performance monitoring to improve stability, alerting, dashboard, etc

Security

Provide support for authentication of records

Architecture & Deployment

Ease of deployment in cloud and on-prem, support private LS. Redundancy and availability



Reviewed in 2018-2019



ESnet GÉANT



INDIANA UNIVERSITY



INTERNET2



RNP



UNIVERSITY OF MICHIGAN

Security

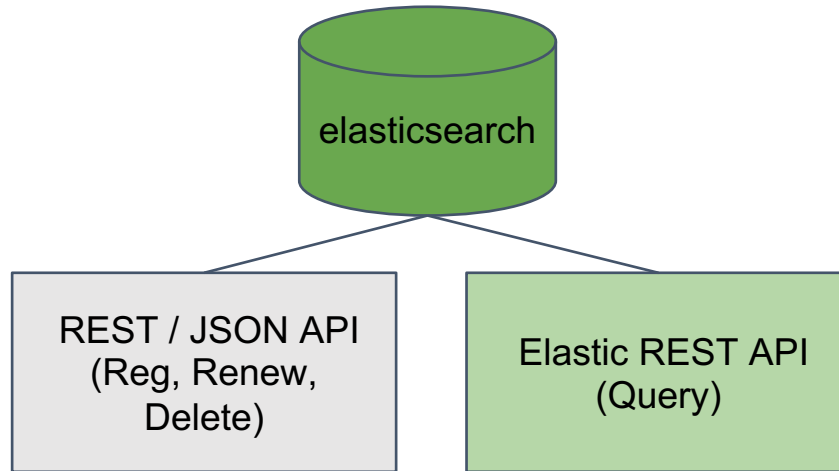
- Supports Digital signing of records
 - Each record can be digitally signed using PKI
- Certificate can be registered with the LS

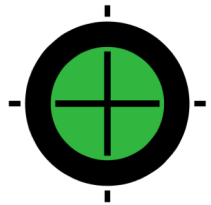


Improved APIs

Bulk renewals

Replaced backend with elasticsearch to leverage REST/JSON APIs to query





What's next?

Data normalization and Analysis



Using Logstash and Elasticsearch

Build logstash pipelines that take the data from the Lookup Service, normalize and write to a different index. Potential to perform roll-ups, transforms, etc

Why Logstash?

Start working on data normalization without affecting the production LS



Initial Goals

Create a pipeline to read the data from LS periodically, use logstash to normalize and write normalized data to a different elastic index

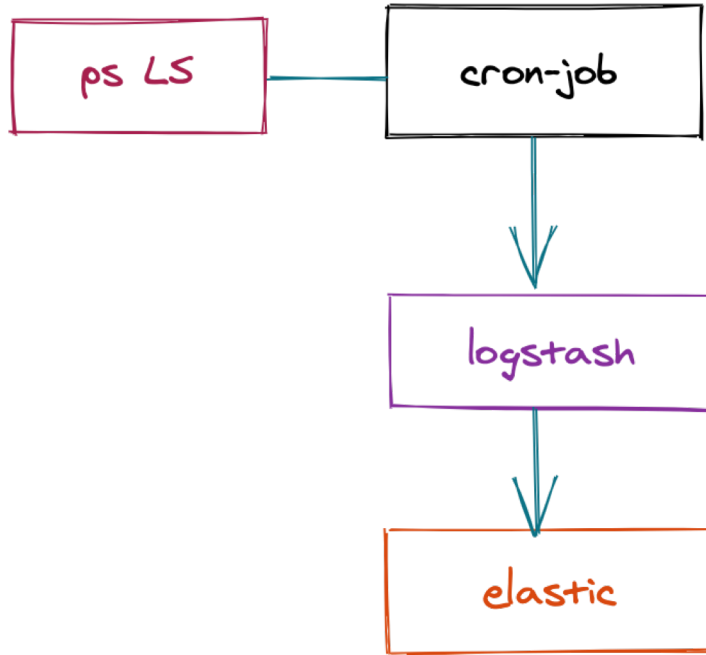
Make the elasticsearch index available to users to experiment.

Maintain periodic snapshots of LS data

Experiment with summary queries using transforms and roll-ups on the snapshots



Architecture



Can share ps LS backend instance

Prior Work

GEANT did some to normalize and analyze the Lookup Service data using splunk. The findings are factored into the Logstash pipelines.



Current Status

- Proof of concept stage
- A basic logstash pipeline that reads the elasticsearch data and normalizes few fields and writes to a different index.



Next steps

- Complete the logstash pipeline
- Make the normalized data available to users.





Thank You

Questions?