

Packet and Flow Marking for Global Science Domains

scitags.org

Shawn McKee / University of Michigan, Marian Babik / CERN

smckee@umich.edu | www.scitags.org

On behalf of the [Research Networking Technical Working Group](#)

At the [2nd Global Research Platform Workshop](#) Sep 20-24, 2021

Motivation

- Networks are becoming more programmable and capable with technologies such as P4, SDN, virtualisation, eBPF, etc.
- But with less and less context about the traffic they carry
 - Cloud deployments, Kubernetes, encryption, tunneling, privacy, etc.
- **Understanding scientific traffic flows in detail is critical for understanding how our complex systems are actually using the network.**
 - Current monitoring/logging tell us where data flows start and end, but is unable to understand the data in flight.
 - Dedicated L3VPNs can be created to track high throughput science domains, but with more domains requiring high throughput this will become expensive, it won't scale, won't work at big sites having to support multiple domains at the same time
- In general the monitoring we have is experiment specific and very difficult to correlate with what is happening in the network. **We suggest this is a general problem for users of the Research and Education Networks (RENs)**

Scientific Network Tags (scitags) is an initiative promoting identification of the science domains and their high-level activities at the network level.

Packet Marking History from WLCG

The Research Networking Technical Working Group(RNTWG) was formed in response to the needs expressed by the WLCG experiments at the January 2020 LHCONe/LHCOPN meeting and has three main work areas:

1. Packet Marking
2. Traffic Shaping
3. Network Orchestration

Packet marking was strongly motivated by the need to understand the source of network flows seen on our R&E networks. We have seen cases of new traffic increasing link usage by 100% for 10's of hours but we were not easily able understand the source. Is this an unintentional side product of an experiment tweaking one of their work-flow configs or distributed data management settings?

Work To Date

We formed the **RNTWG** in spring of 2020 with three sub-groups (last slide).

Packet marking was viewed as the appropriate first step and we have had regular meetings every ~2 months since summer 2020.

There are detailed results of our work in our [Packet Marking Subgroup notes](#).

- **Goal** is to make flows identifiable (Owner and Activity) anywhere along its path
- **Summary:** Packet **marking** will require IPv6 and newer kernels.
- It is possible to describe a flow via a side-channel: send special UDP packets (“Fireflies” suggested by ESnet) along the same path as the flow (works for IPv4 as well)

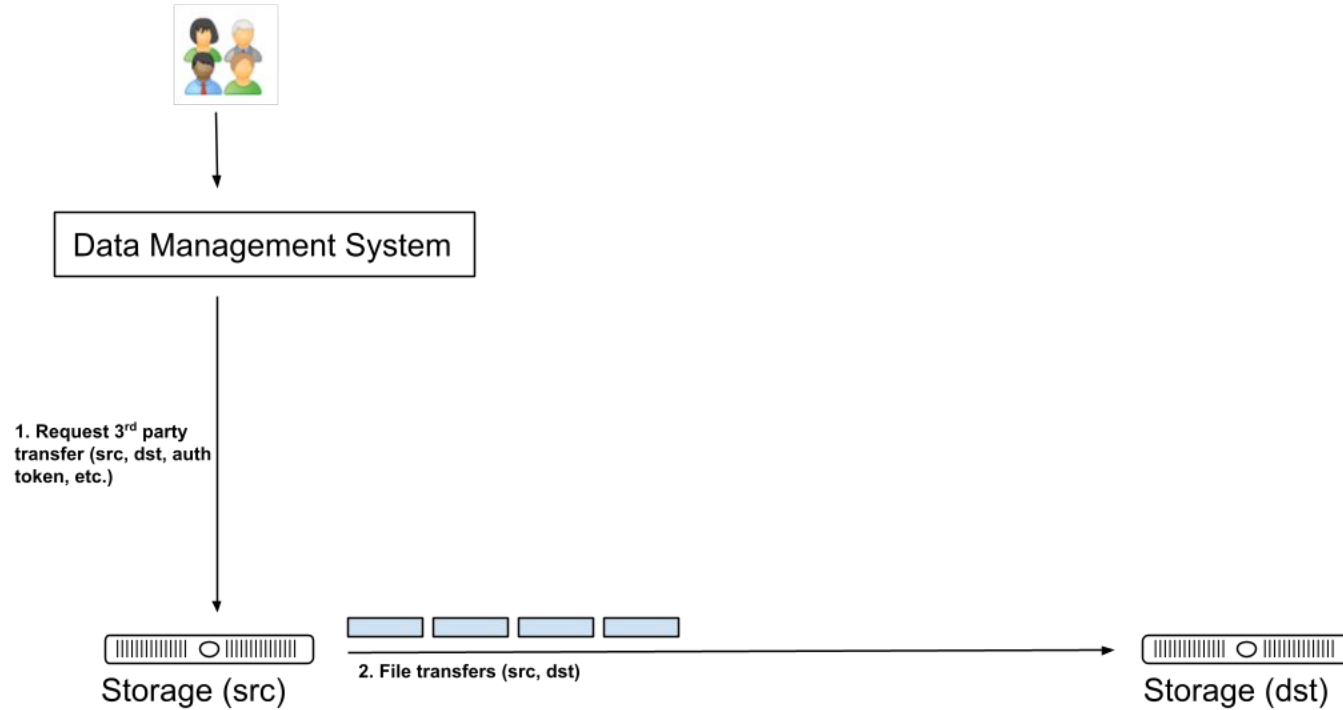
Plans and Technical Specifications

We have developed a plan to start identifying our packets and flows based on following **concepts**:

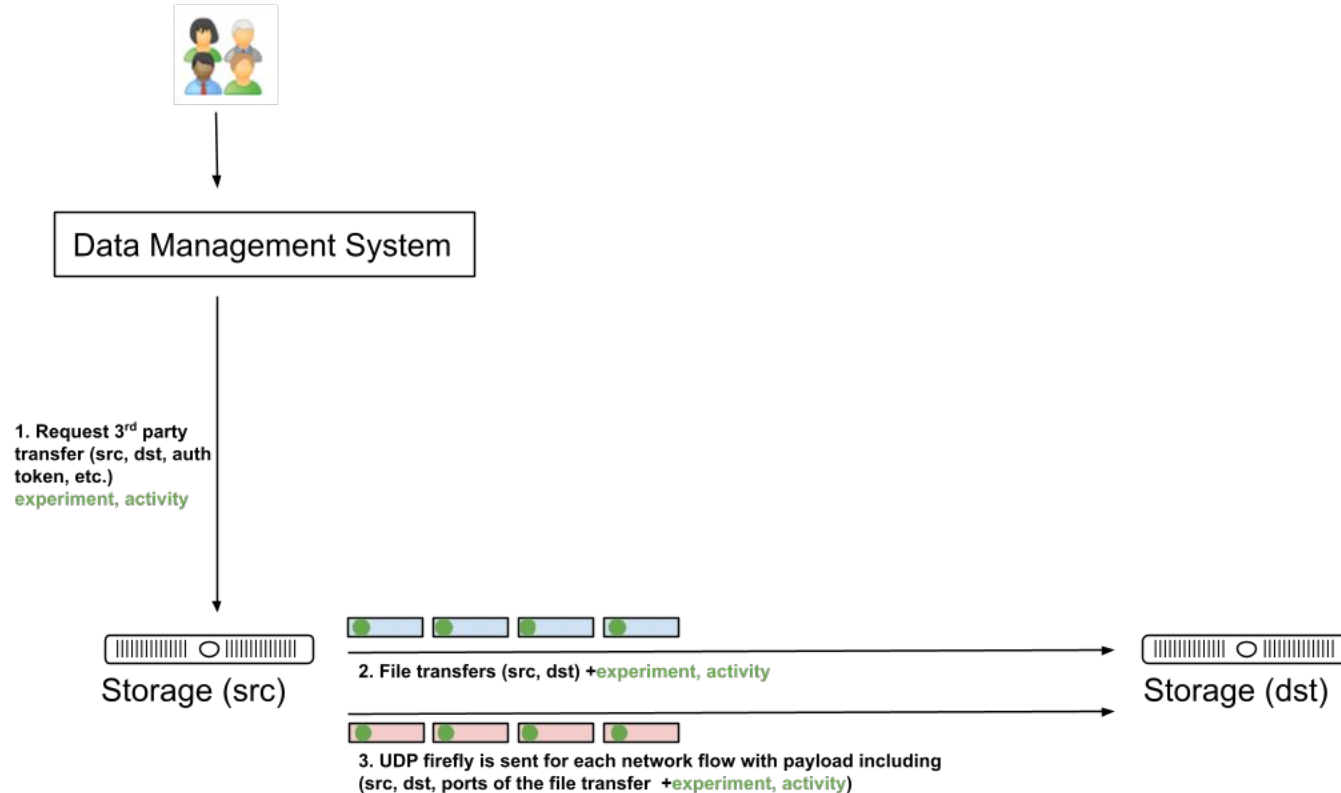
- *Flow marking* with UDP firefly and *packet marking* with IPv6 flow label
- **Flow identifier** - common identifier used by both flow and packet marking
 - Two fields: experiment and activity
- Introduces concept of **flow service** to help with the prototype
 - Dedicated service to help mark the traffic running on generators/storages/caches and exposing a well defined interface
- Introduces service to resolve flow identifier (**flow registry**)
 - Maps experiments and activities to the flow identifier and vice versa
 - Managed by all stakeholders (based on the existing Google Sheet)

The next few slides will document how we see the process evolving.

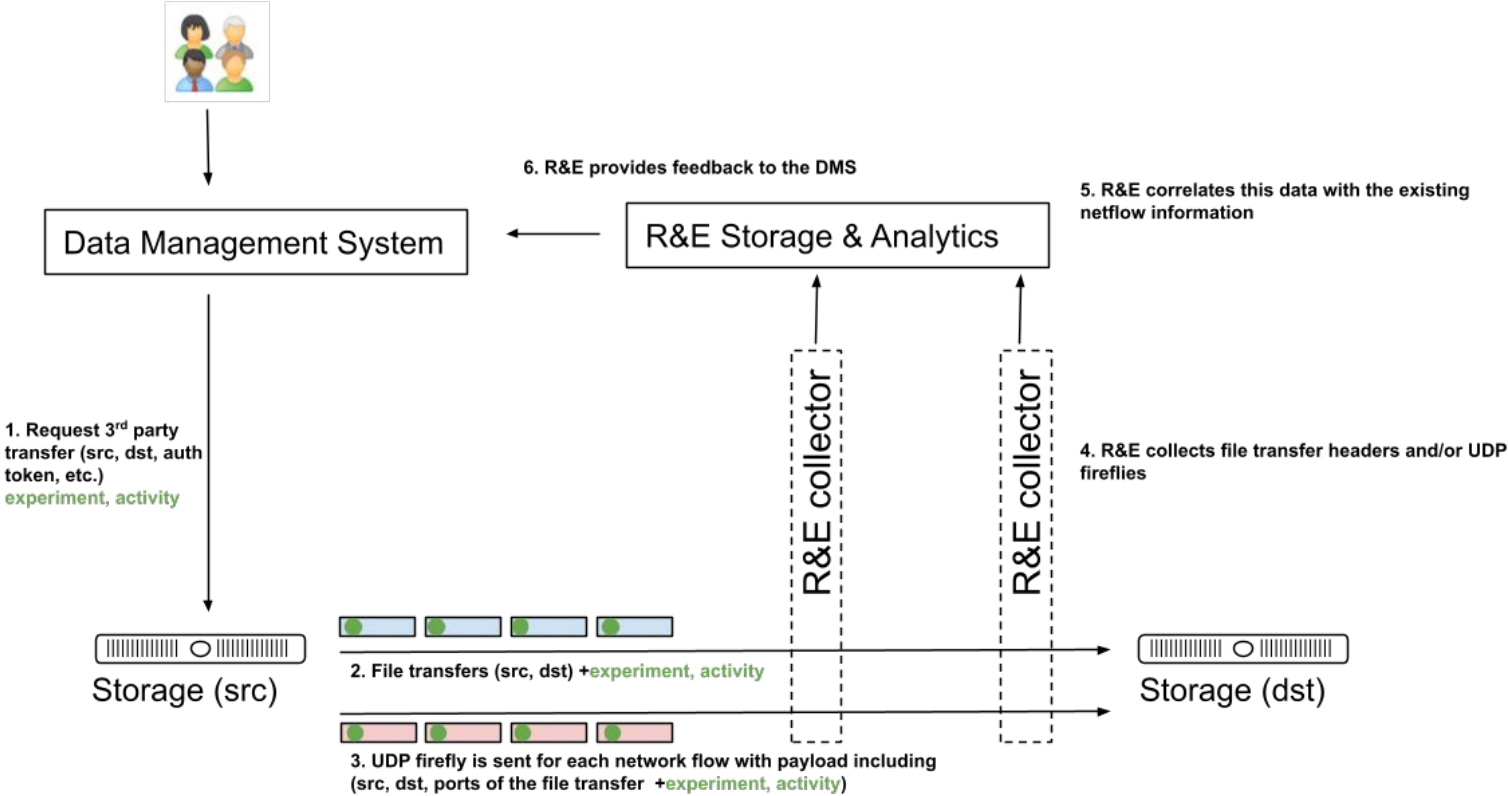
How scitags work



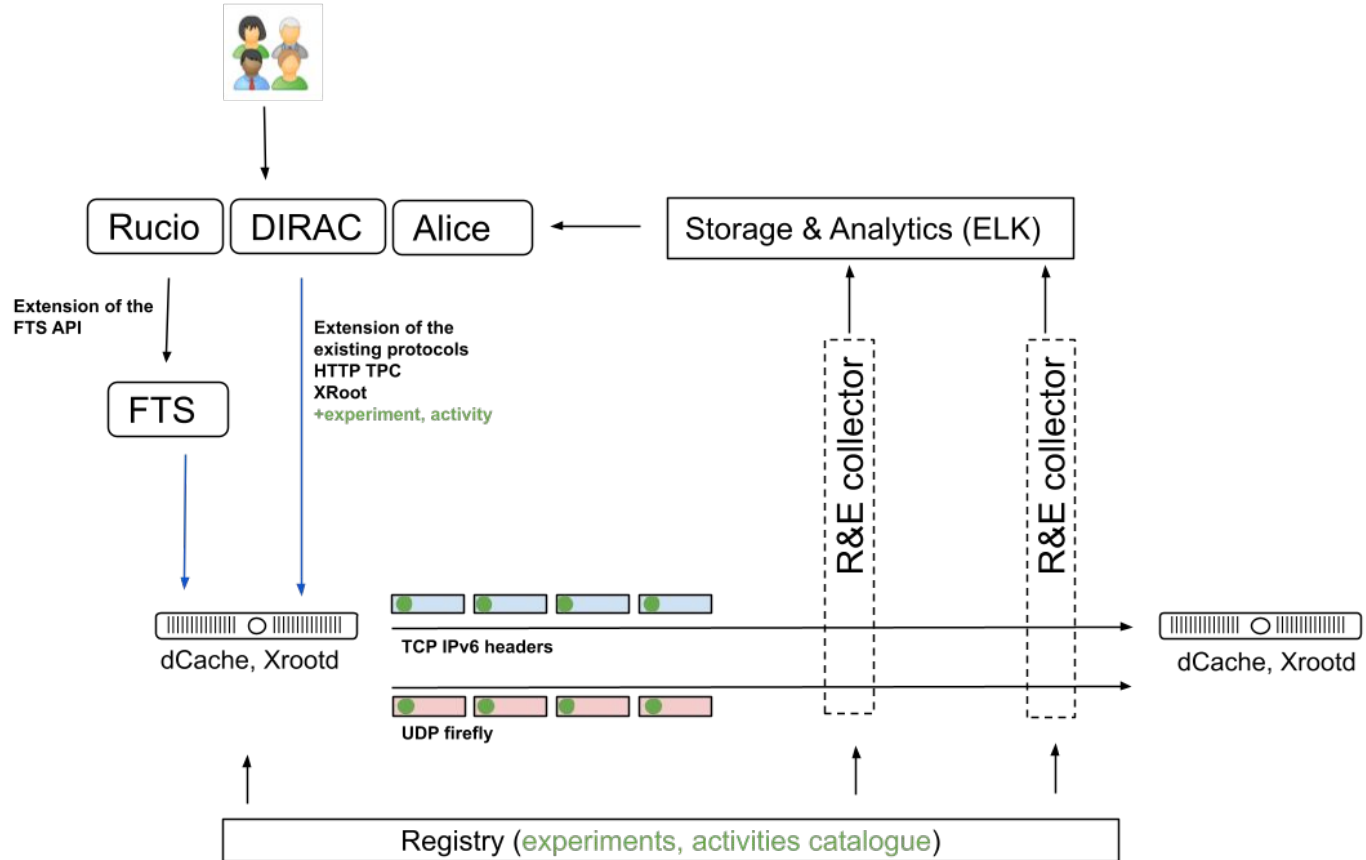
How scitags work



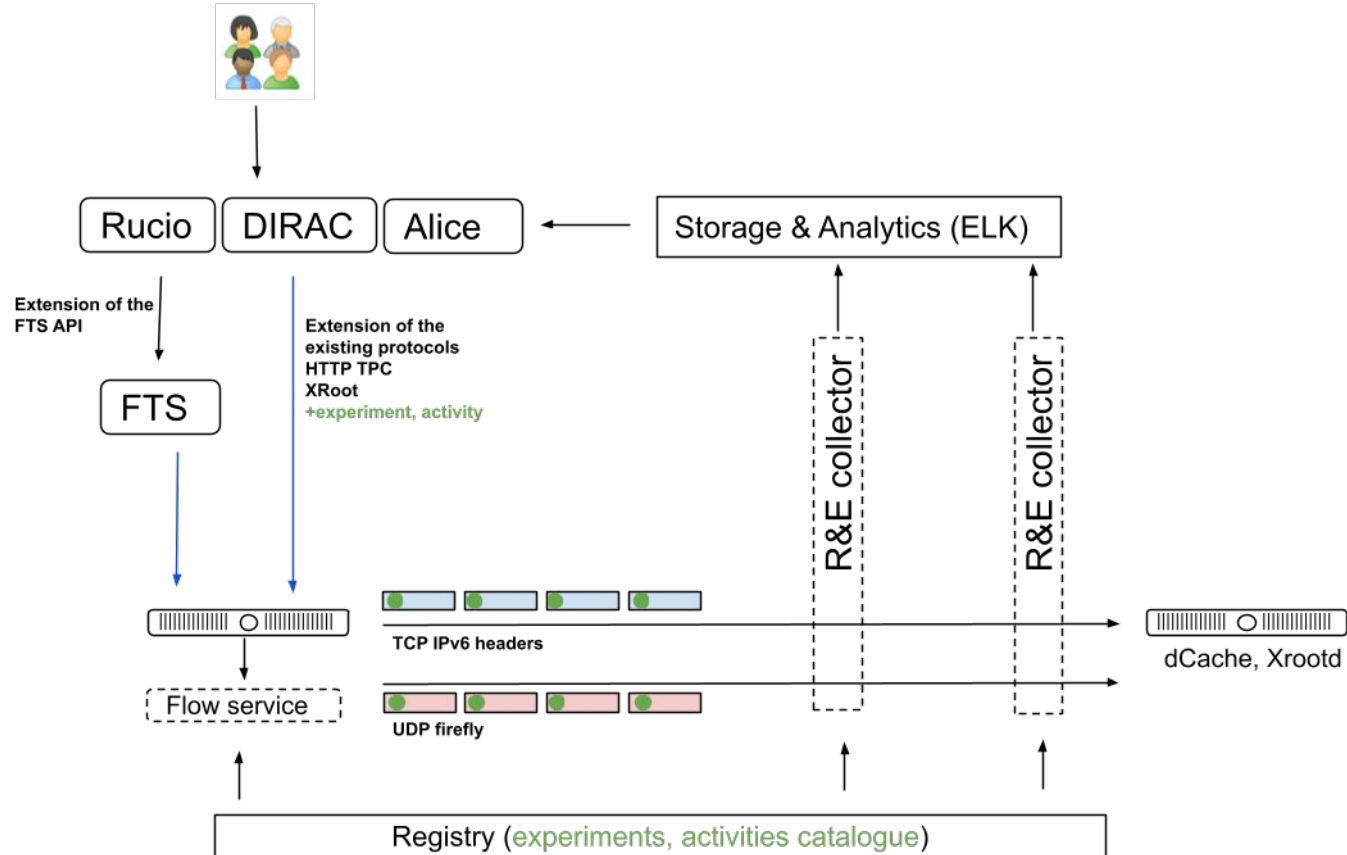
How scitags work



Scitags Architecture



Scitags Architecture



Scitags

- Marking is based on two different approaches
 - **Flow marking** using UDP fireflies (works for both IPv4 and IPv6)
 - **Packet marking** using IPv6 flow label and/or header extensions
- Both carry **flow identifier**, which at present is an encoded representation of experiment/science domain and activity
 - For UDP fireflies flow id can be extended with other fields in the future
 - For packet marking the space is restricted due to number of bits available in the headers
- Experiments and activities need to be registered prior to their usage
 - Registry serves this purpose and ensures RENs and DDMS have consistent view
- Designed to work with proxies, cached proxies and private networks
- Generators, collectors, storage and analytics can evolve independently

Next Steps

We are planning to test packet and flow marking during the upcoming WLCG Data Challenge(DC) the first week in October.

- Goal is to have some WLCG storage servers sending fireflies during the test
- Code (rpm) being finished and tested this week (September 20-24)
- Initial version limitations
 - The flowd service will have a firefly plugin that monitors netstat
 - New flows to external destinations will generate firefly packets (syslog encapsulated)
 - Firefly packets will be sent directly to the ESnet collector (avoiding the need to capture them)
- Direct implementations for dCache and Xrootd are planned.

We hope to learn a lot from the initial testing. Plans are to evolve the flowd service to a production service with plugins and a registry service before the second DC in 2023.

Status and Traction



Draft [Technical Specification](#) available;
[Packet Marking Overview](#)



Prototype testing as part of the WLCG Data Challenges effort in collaboration with ESnet



Prototype code of the flow service ([flowd](#)) implementing UDP fireflies

<https://www.scitags.org>

Questions ? Comments ?

Thanks!

Relevant URLs

The scitags web page: <https://scitags.github.io>

The RNTWG (sub-groups linked inside)

<https://docs.google.com/document/d/1I4U5dpH556kCnoIHzyRpBI74IPc0gpgAG3VPU98lo0/edit?usp=sharing>

Draft packet marking details:

<https://docs.google.com/spreadsheets/d/1KOkZxmCtLoU2y5DKGjvQEo-A-A3kUN2UqnWlqF-4zoQ/edit#gid=0>

Code at <https://github.com/scitags/scitags.github.io>