PERISCOPE: Standardizing and Orchestrating Looking Glass Querying

Vasileios Giotsas

UCSD/CAIDA

vgiotsas@caida.org

NANOG 68, October 17-19 2016 , Dallas, TX

Purpose of this Talk

- Inform the operational community about Periscope.
- Solicit feedback:
  - Details that we may have missed
  - Ways to make Periscope more useful
  - Technical insights, usage statistics, historical data …
- Encourage engagement and contributions

High-level Goals and Principles of Periscope

Periscope unifies the discovery and querying of Looking Glasses under a uniform API

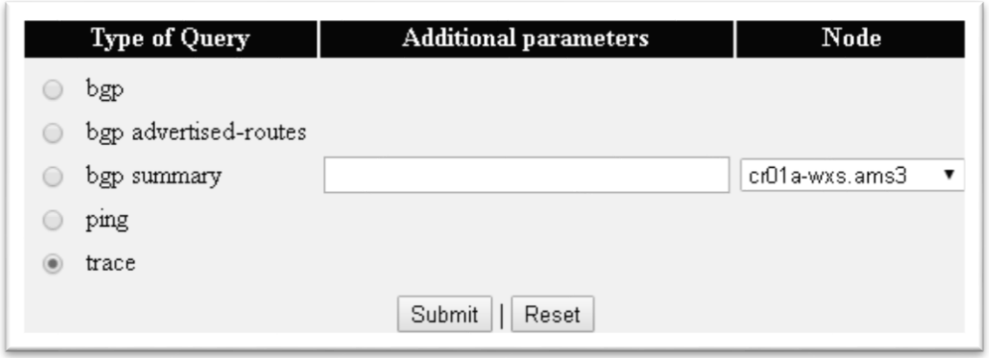
- Respect resource limitations and preserve conservative query rates.
- Provide transparency and accountability in Looking Glass querying.
- Be responsive and compliant to operators' requests.

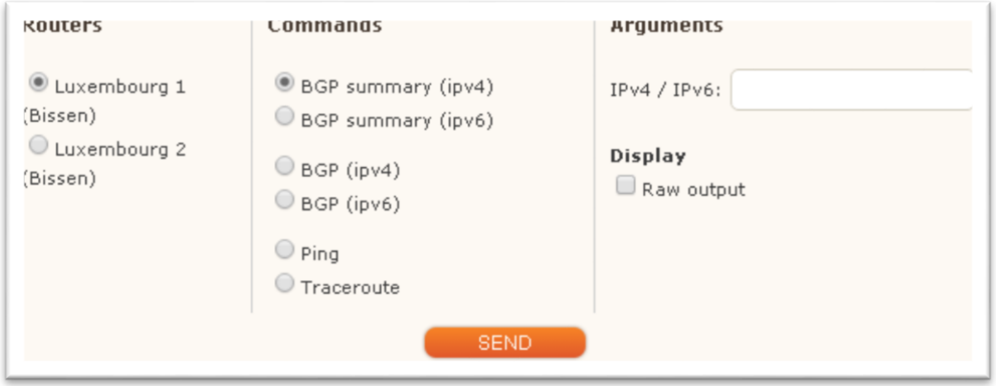
Benefits to Operators and Researchers

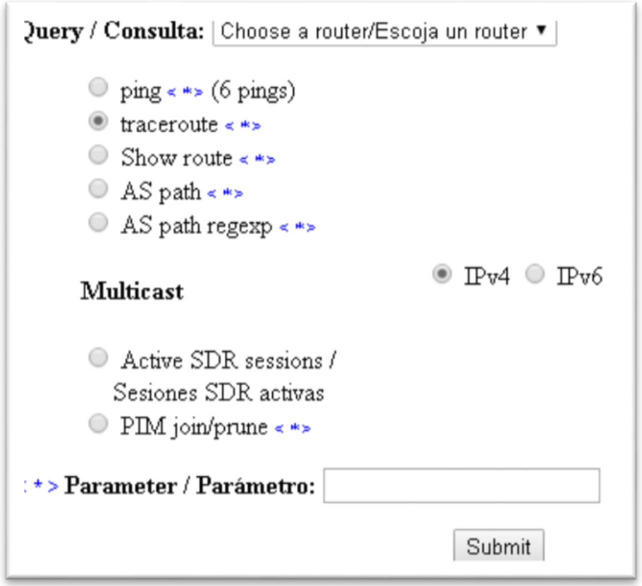
- Easier to discover and query VPs for reverse paths.
- Improved monitoring and troubleshooting capabilities.
- Easier policing of Looking Glass usage through an access-control layer.
- Improved utilization and load distribution.
- Avoid redundant measurements by capturing and making public historical measurement data.

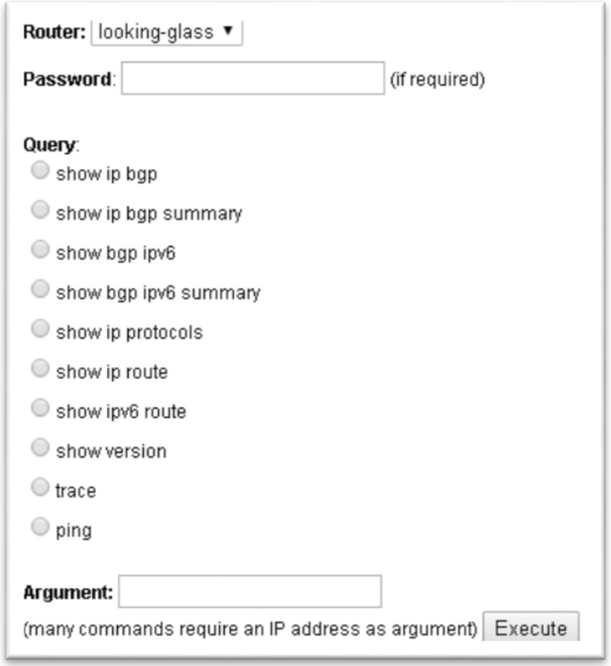
What is a Looking Glass (LG)?

- Web interfaces to routers and servers that allow the execution of non-privileged networking commands:
  - show bgp summary, show ip bgp, traceroute, ping, …









Advantages of LG measurements

- LGs combine features not available in other platforms:
  - Access to non-transitive BGP attributes (e.g. Local Preference)
  - Co-located control-plane and data-plane monitors
  - Monitors inside critical infrastructures (IXPs, Colocation Facilities, border routers)
  - Vantage points in ASes not covered by other platforms

- LGs are among the few public measurement tools that provide direct interfaces to routers.

- Motamedi R., et al. A Survey of Techniques for Internet Topology Discovery. IEEE *Communications Surveys & Tutorials*, *17*(2)
- A. Khan, T. Kwon, H.-c. Kim, & Y. Choi, “AS-level Topology Collection Through Looking Glass Servers,” in IMC ’13

LGs are widely used by researchers, operators and users

LOUP: The Principles and Practice of Intra-Domain Route Dissemination

Nikola Gvozdiev, Brad Karp, Mark Handley
University College London

IXPs: Mapped?

Brice Augustin†; Balachander Krishnamurthy‡, Walter Willinger‡

On Inferring and Characterizing Internet Routing Policies

Feng Wang
fewang@ecs.umass.edu

Lixin Gao
lgao@ecs.umass.edu

ARTEMIS: Real-Time Detection and Automatic Mitigation for BGP Prefix Hijacking

Gavriil Chaviaras, Petros Gigis, Pavlos Sermpezis, and Xenofontas Dimitropoulos
FORTH / University of Crete, Greece
{gchaviaras, gkigkis, sermpezis, fontas}@ics.forth.gr

ABSTRACT

Prefix hijacking is a common phenomenon in the Internet that often causes routing problems and economic losses. In this demo, we propose ARTEMIS, a tool that enables network administrators to *detect and mitigate* prefix hijacking incidents, against their own prefixes. ARTEMIS is based on the real-time monitoring of BGP data in the Internet, and software-defined networking (SDN) principles, and can completely mitigate a prefix hijacking within a few minutes (e.g., 5-6mins in our experiments) after it has been launched.

CCS Concepts

offer BGP prefix hijacking detection as a service to ASes. In addition, previous research focuses primarily on *accurately* detecting BGP hijacks, rather than *timely* detecting *and* mitigating them. The whole detection/mitigation cycle presently has significant delay: (i) aggregated BGP data from RouteViews [5] or RIPE RIS [6], which are commonly-used for detection, become available approximately every 2 hours (BGP full RIBs) or 15mins (BGP updates); (ii) a network administrator that receives a notification from a third-party alert system needs to *manually* process it to verify if the notification corresponds to a hijacking or is a false alarm; and (iii) for mitigation, administrators often need to manually reconfigure routers or contact administrators of other

NANOG nanog mailing list archives

Re: Frontier Internet Outage

From: Josh Reynolds <josh () kyneticwifi com>
Date: Mon, 2 May 2016 08:58:23 -0500

> From where? To what? Have you checked any carrier's looking glass?

Re: Clueful BGP from TW-Telecom/L3

From: Mel Beckman <mel () beckman org>
Date: Wed, 3 Aug 2016 16:01:05 +0000

Input Regarding Cogent and NTT

How to get TWC to stop advertising a route? (self)
58 points submitted 9 months ago

I barely know the difference between the OSI 7 layer model and a Taco Bell 7 layer burrito, but I've found myself in the middle of a routing snafu.

I work for a contractor whose client systems suddenly lost the ability to reach to our corporate mothership except via VPN. After much finger pointing and accusations hurled at the people doing firewall maintenance, we discovered that packets were reaching the destination machines, but not making it back due to Time Warner advertising a route that no longer exists.

The client owns it's own IP space, and evidently they packed up and moved to Comcast for reasons known only to themselves. I'm told our client did try talking to TWC, but got blown off. They're trying to get Comcast to do something, but don't sound hopeful.

Problems with LG measurements

- Lack of standardization and consistency:
  - Disparate input interfaces, output formats, supported commands
- LGs are hard to discover
  - No centralized index of LGs and their corresponding locations
- Historical measurements are not archived:
  - Loss of reusable information, potential query redundancy
- LGs have high attrition rates:
  - Hard to maintain an up-to-date database of LGs

M. Stubbig, “Looking Glass API.” https://tools.ietf.org/html/draft-mst-lgapi-04, May 2016.

Problems with LG measurements

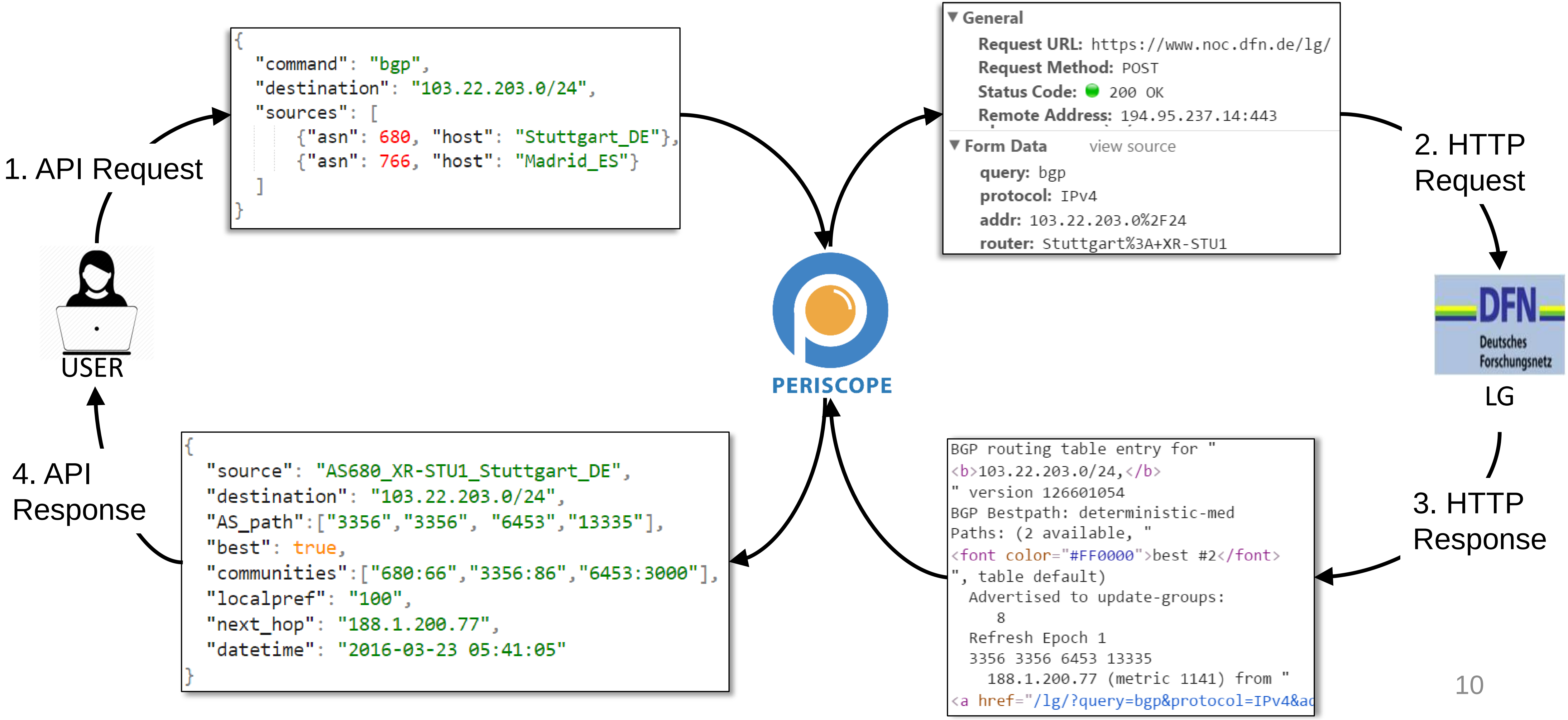
- Lack of standardization and consistency:
  - Disparate input interfaces, output formats, supported commands

Periscope implements a common querying scheme, indexing, and data persistence features.

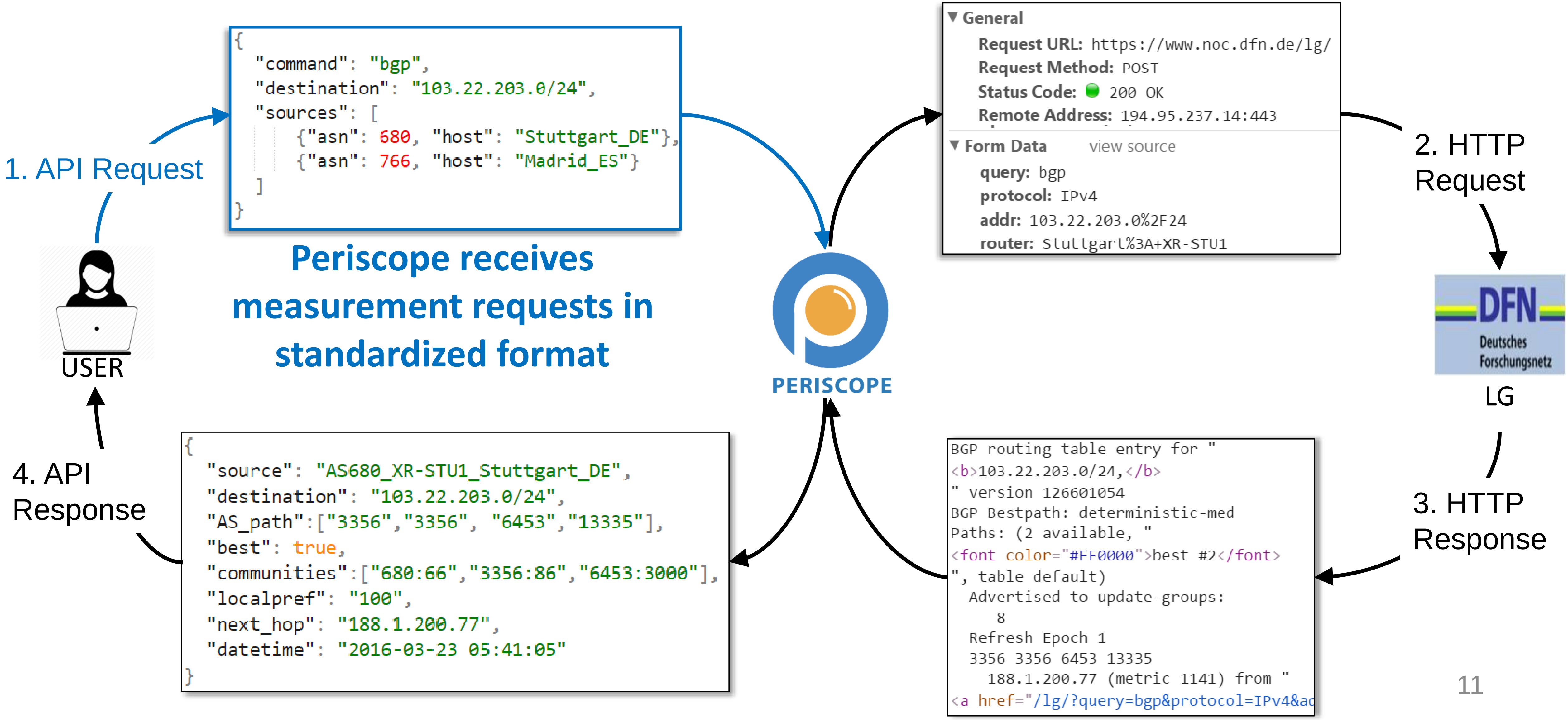
  - Loss of reusable information, potential query redundancy
- LGs have high attrition rates:
  - Hard to maintain an up-to-date database of LGs

M. Stubbig, “Looking Glass API.” https://tools.ietf.org/html/draft-mst-lgapi-04, May 2016.

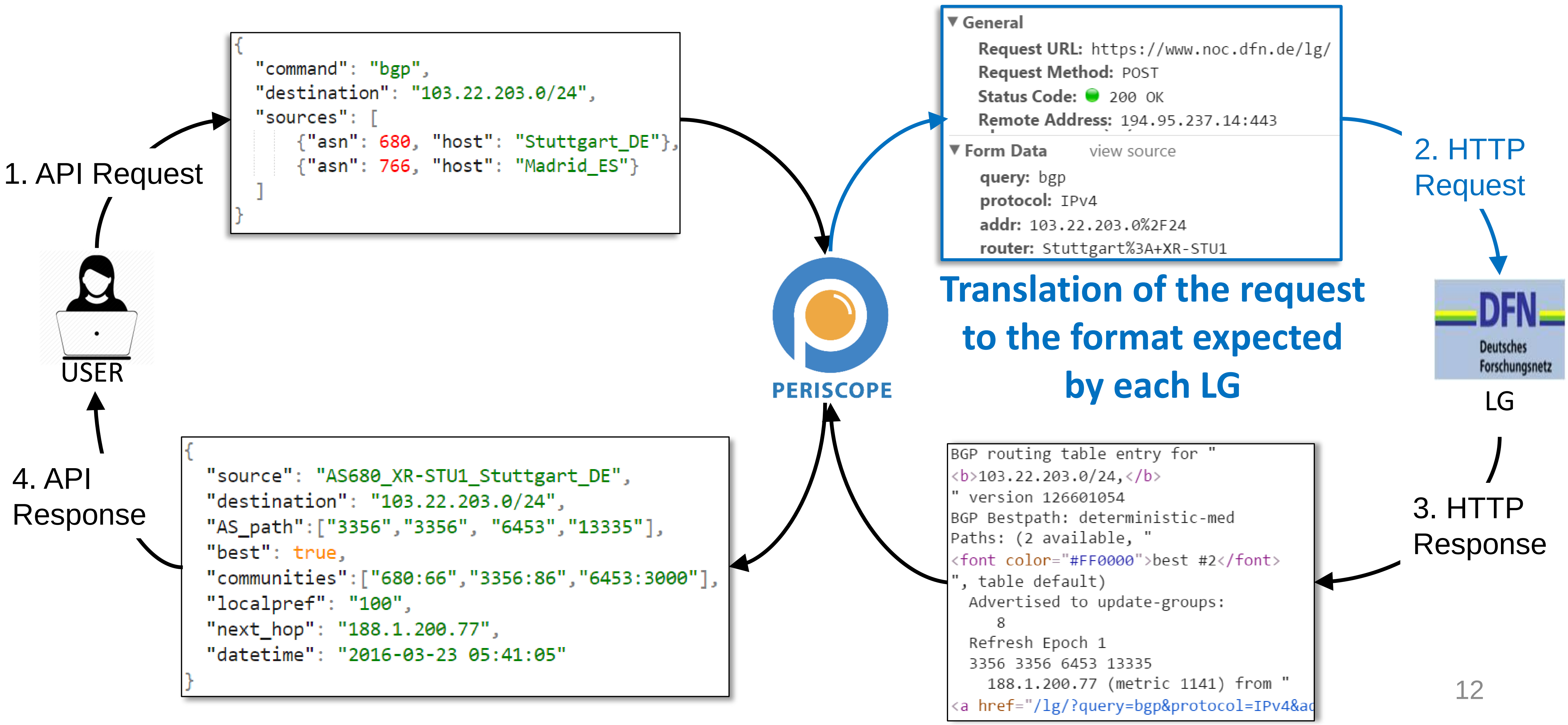
Periscope Workflow



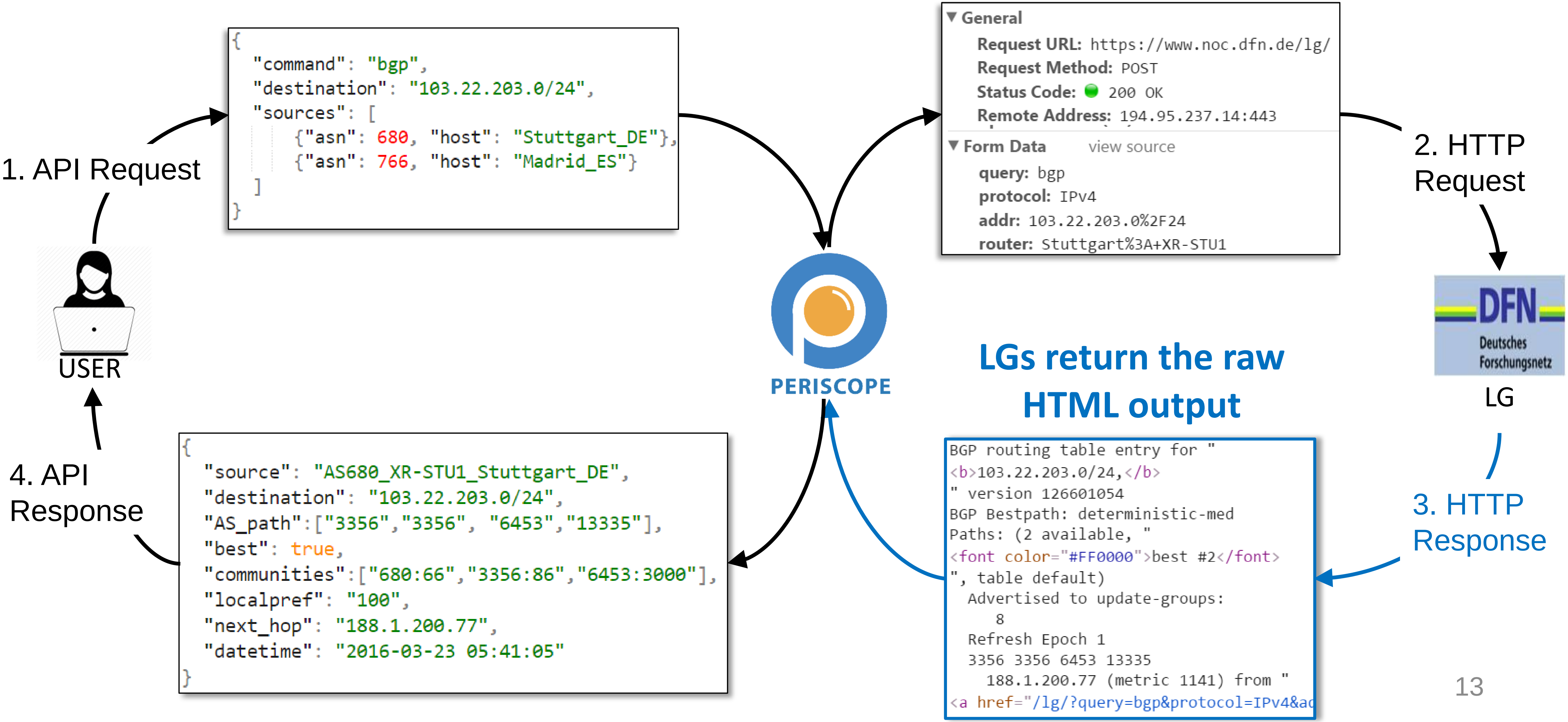
Periscope Workflow



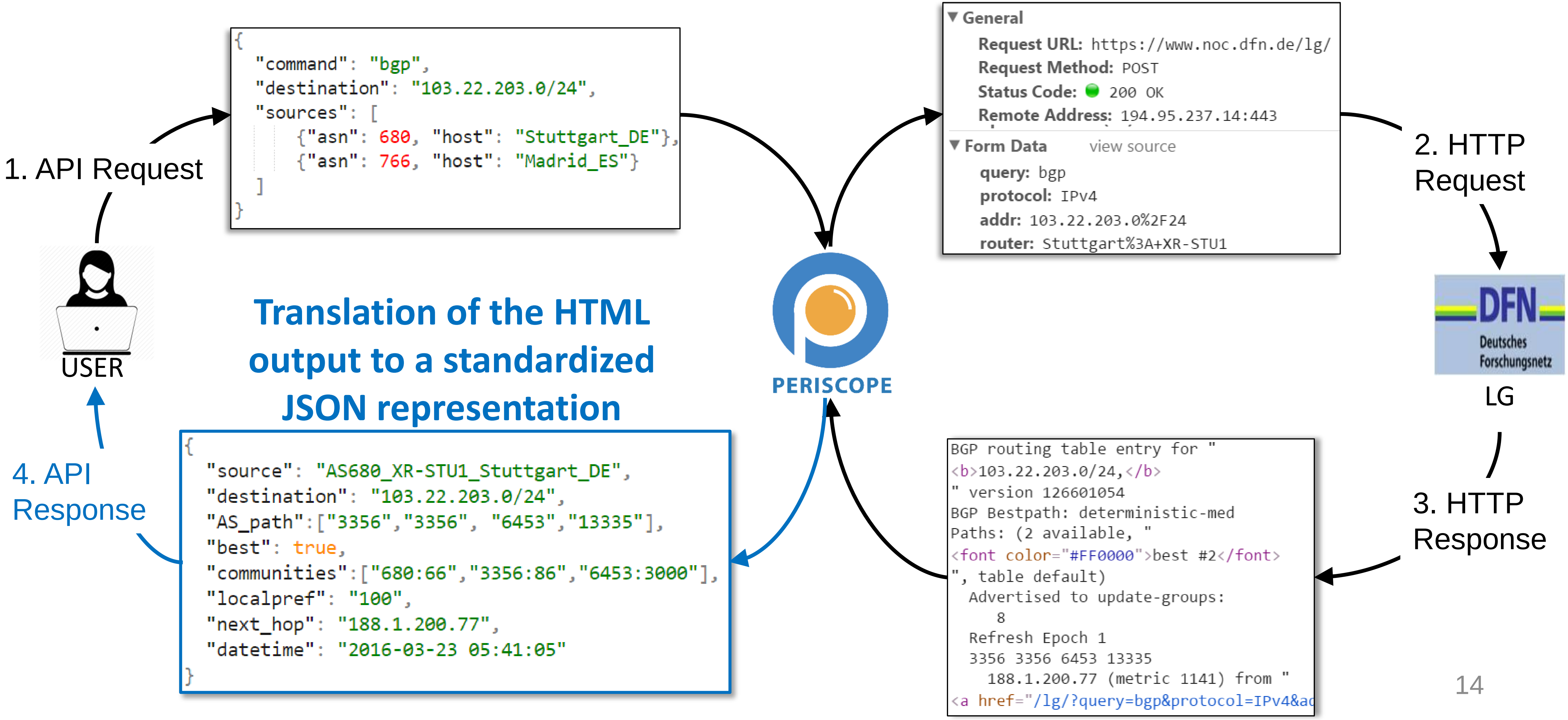
Periscope Workflow



Periscope Workflow



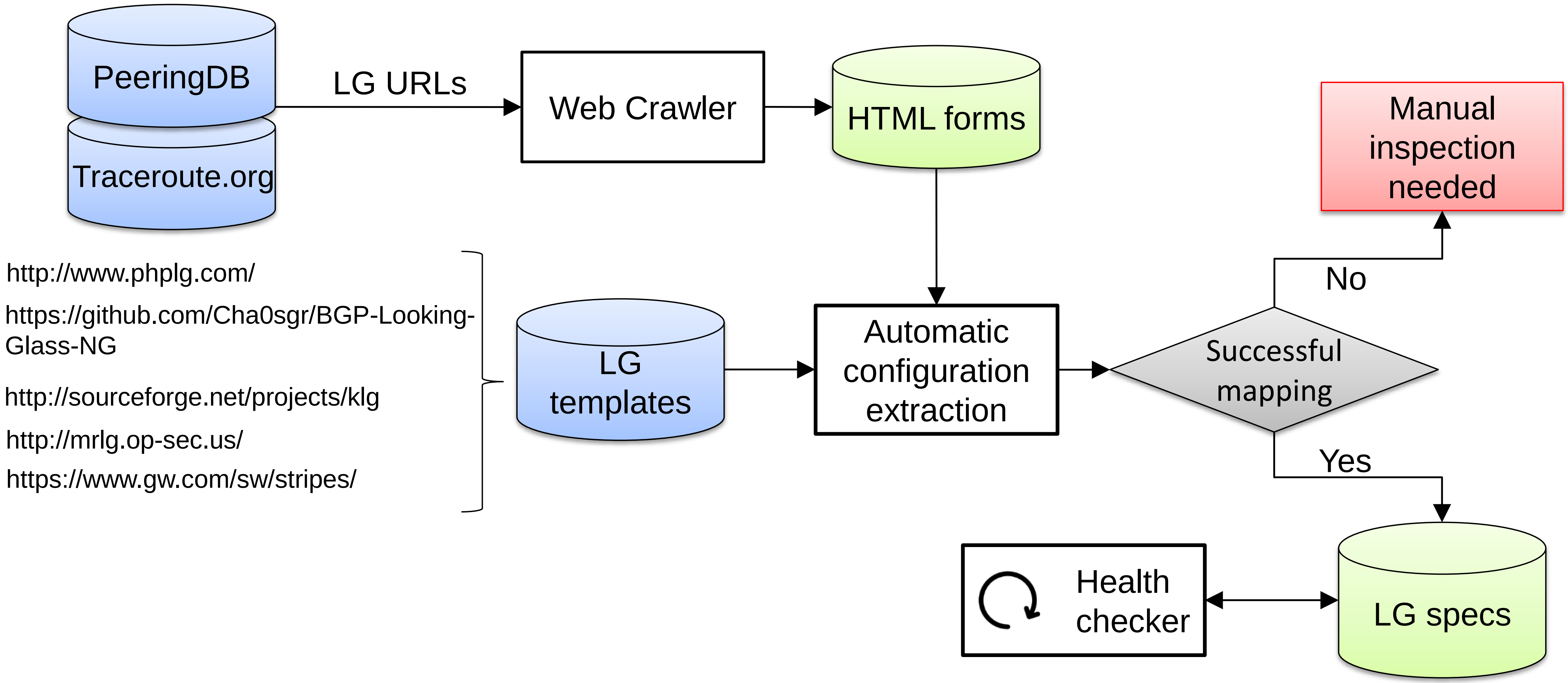
Periscope Workflow



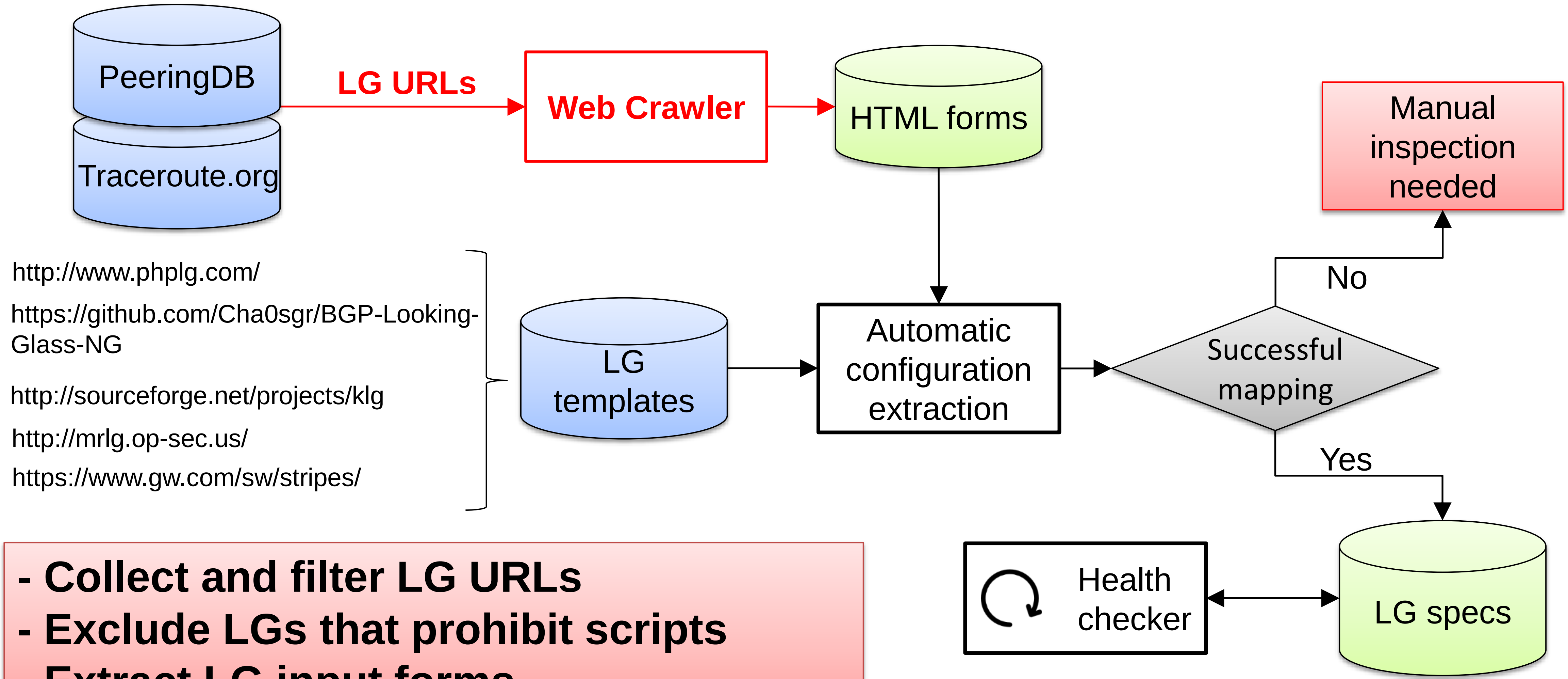
Implementation Challenges

- Automatically understand the disparate input/output formats of each LG.
- Automatically discover new LGs, detect changes in the status and capabilities of already supported LGs:
  - Manual parsing is impractical
- Support multiple concurrent users while preserving the query rates of native LG querying.
- Optimize the number of satisfied queries within restrictive querying budgets.

LG Ingestion Process

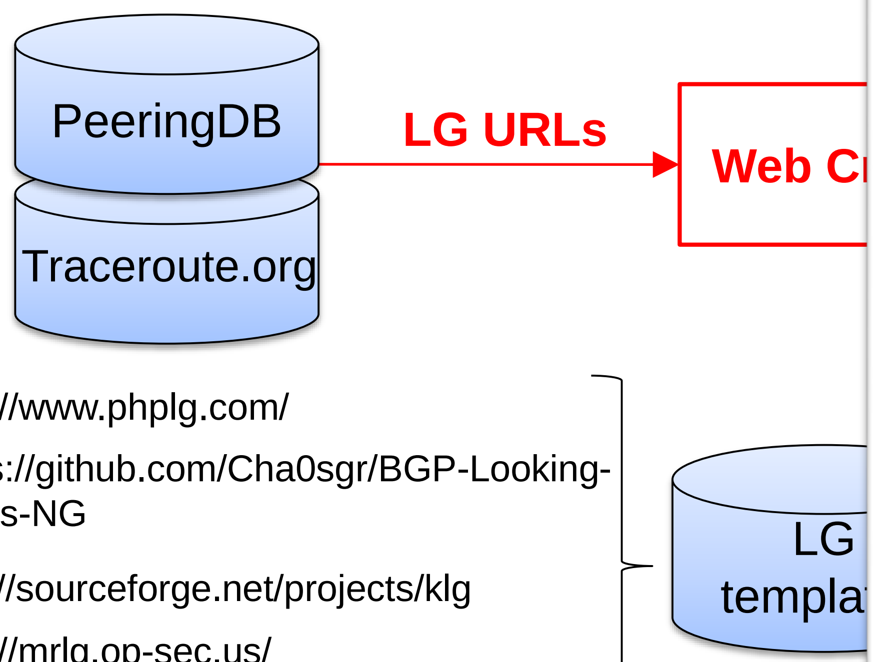


LG Ingestion Process

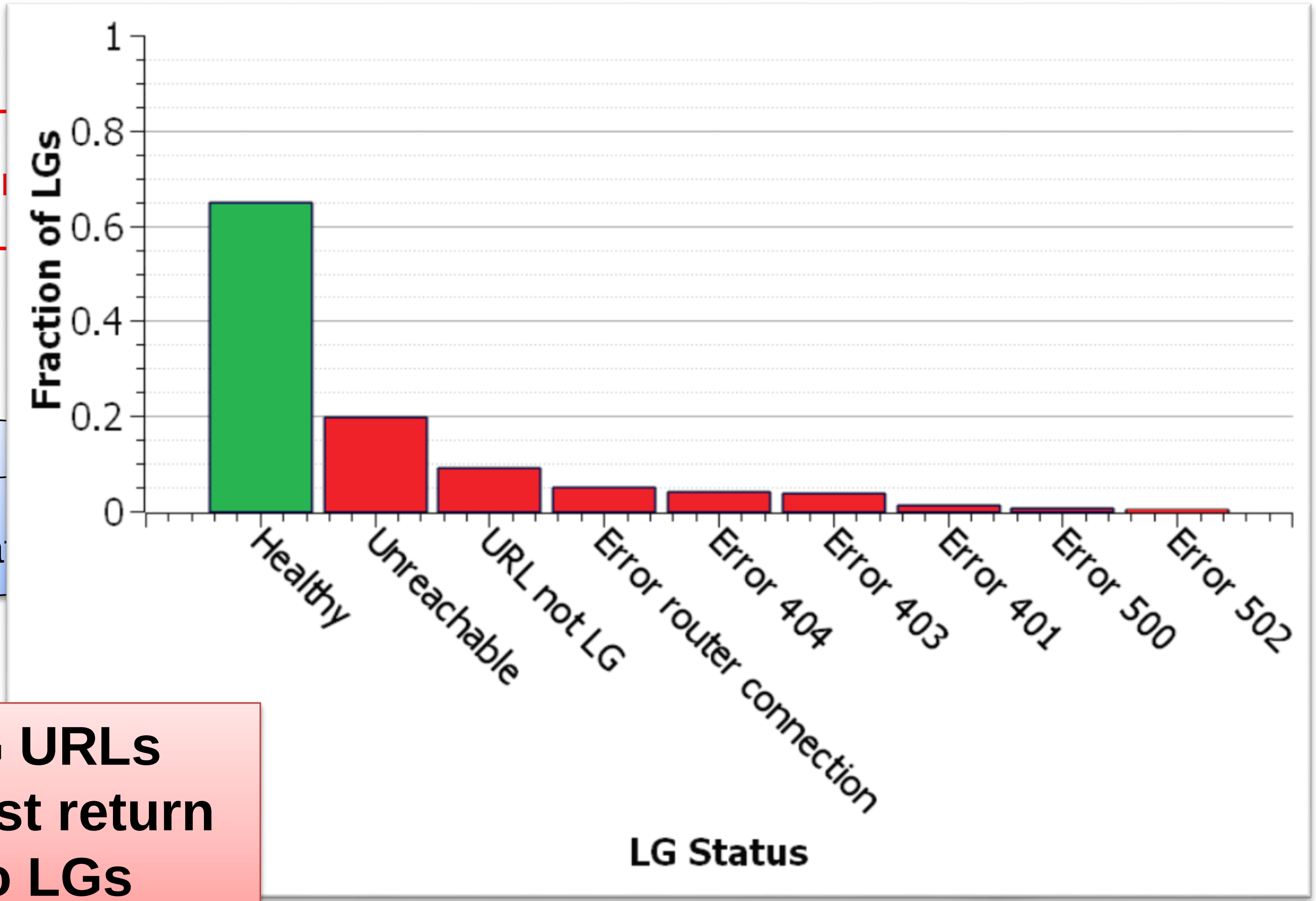


- Collect and filter LG URLs
- Exclude LGs that prohibit scripts
- Extract LG input forms

LG Ingestion Process

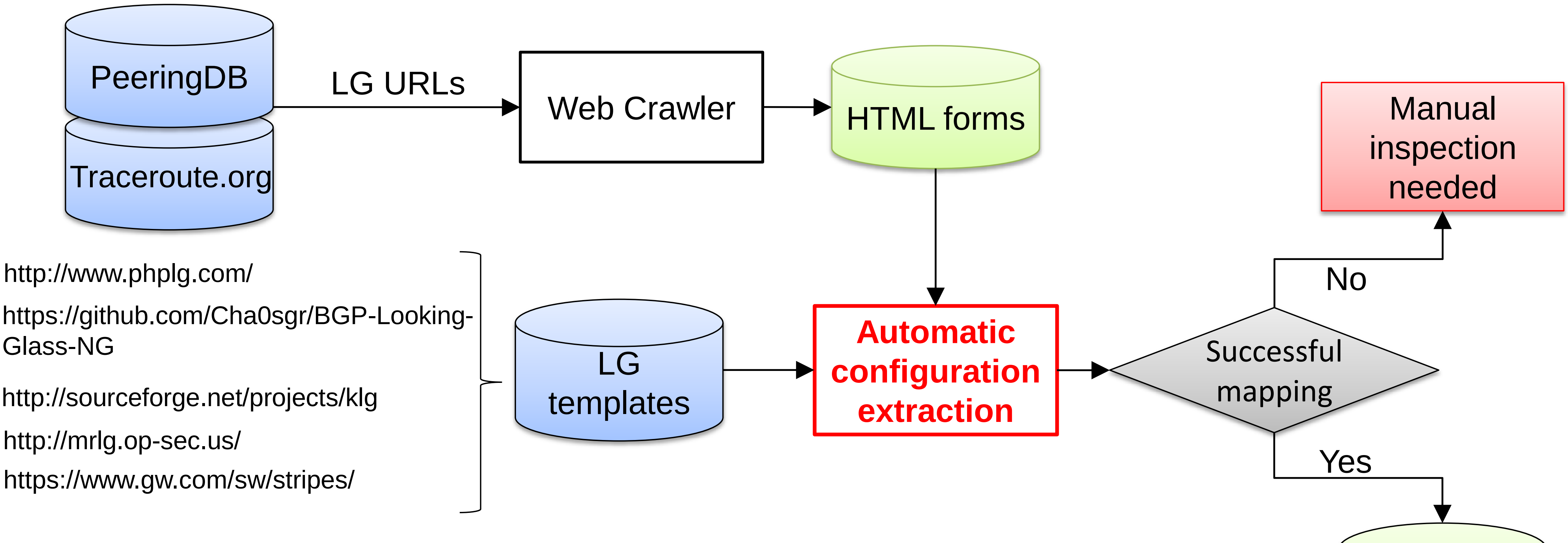


http://www.phplg.com/
https://github.com/Cha0sgr/BGP-Looking-Glass-NG
http://sourceforge.net/projects/klg
http://mrlg.op-sec.us/
https://www.gw.com/sw/stripes/



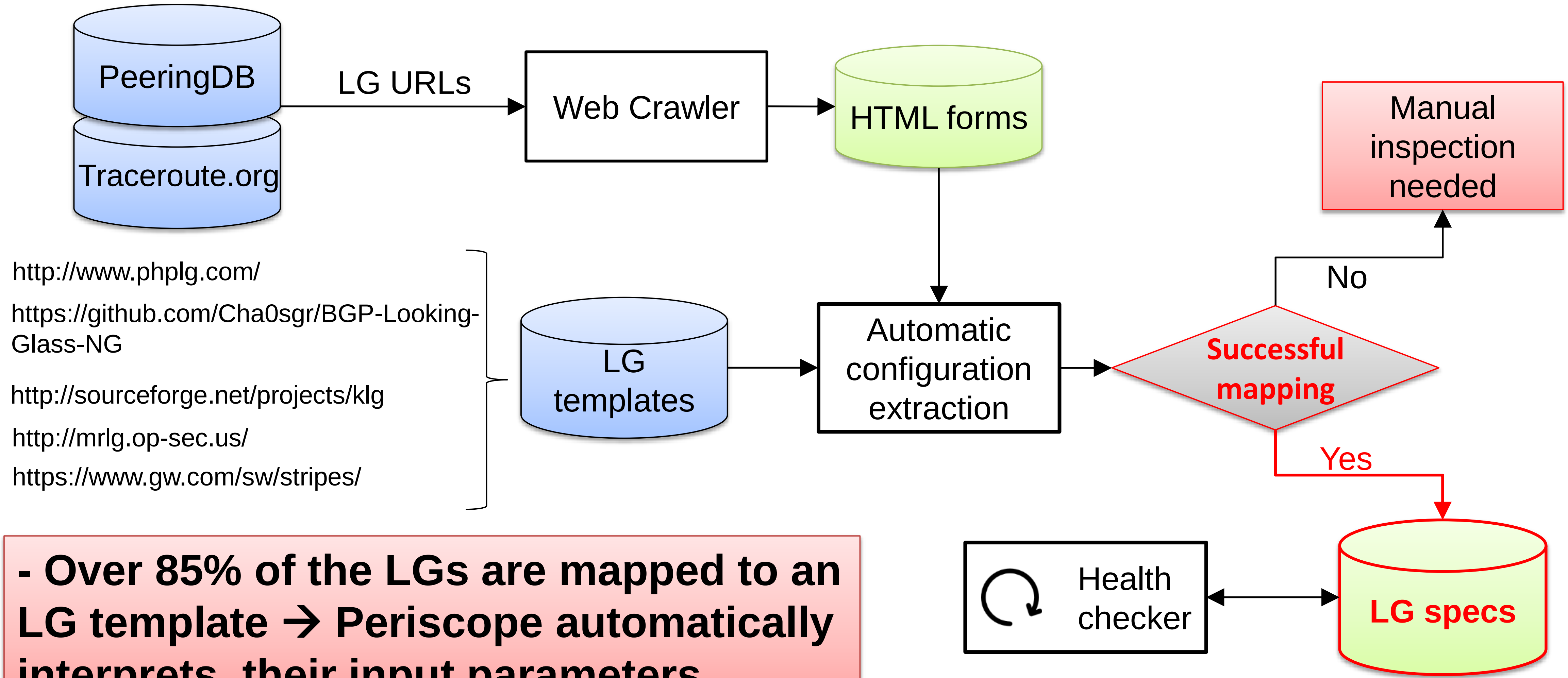
Only ~65% of the collected LG URLs point to responsive LG, the rest return errors or do not correspond to LGs

LG Ingestion Process



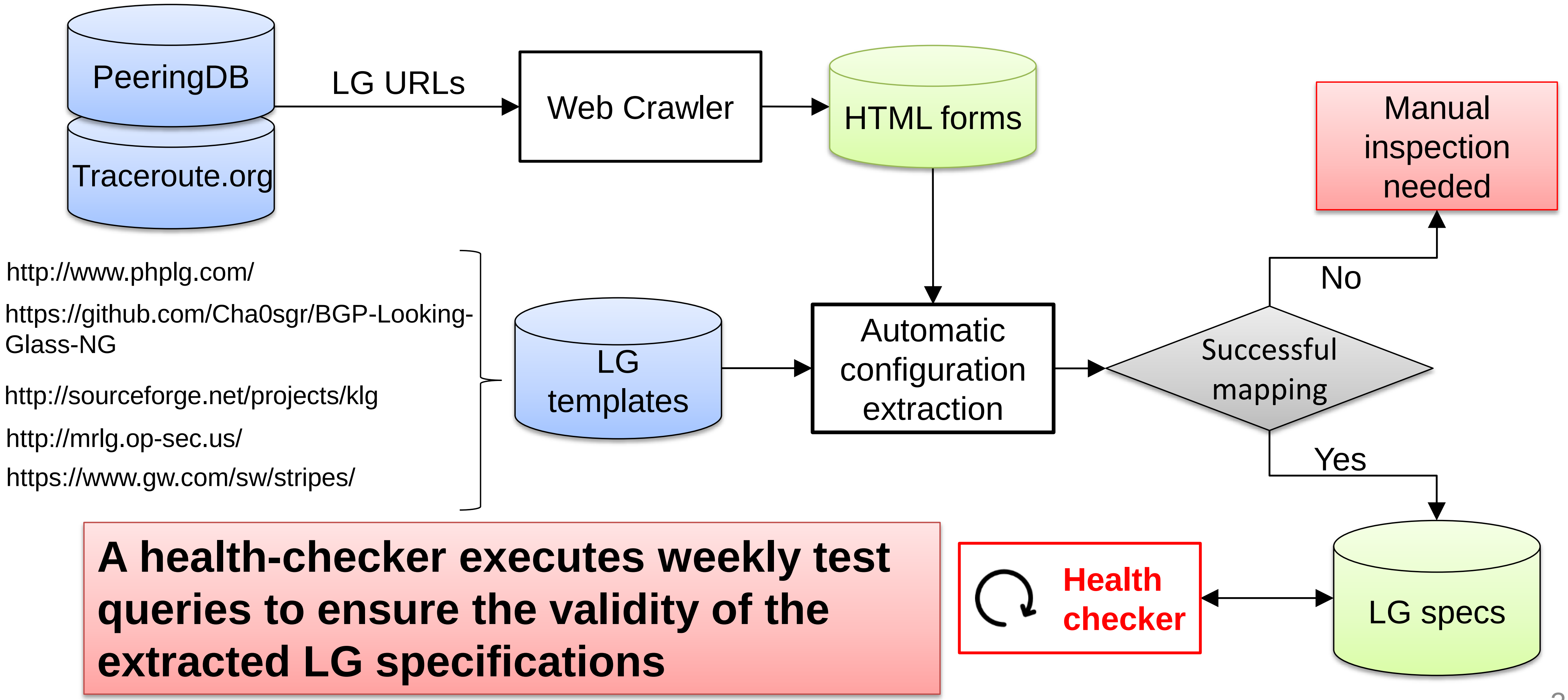
Match the extracted input forms against LG templates derived from open-source LG implementations

LG Ingestion Process

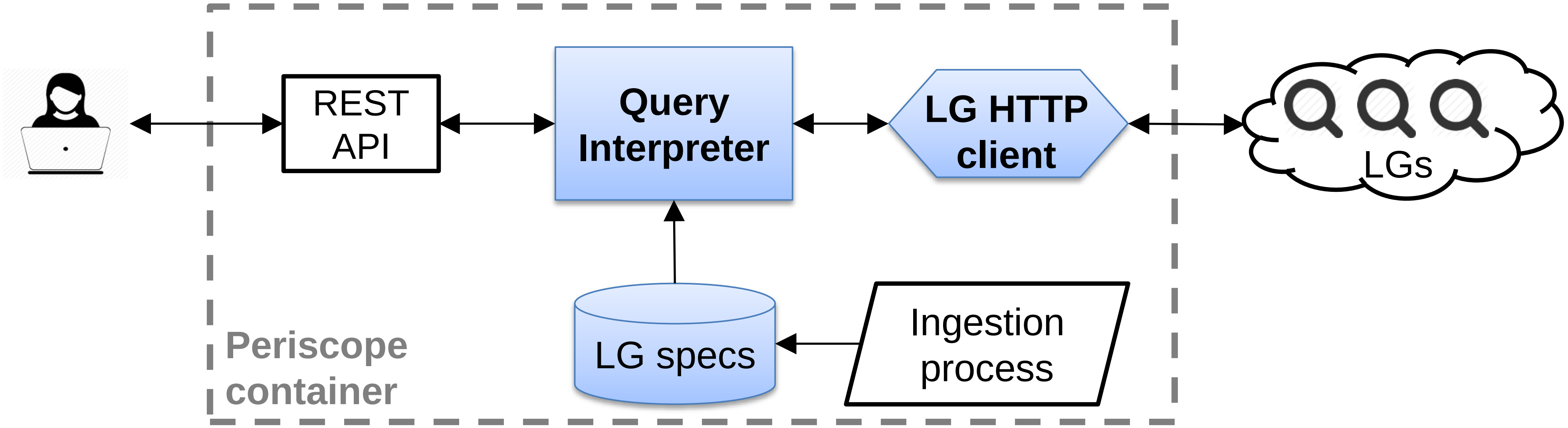


- Over 85% of the LGs are mapped to an LG template → Periscope automatically interprets their input parameters

LG Ingestion Process



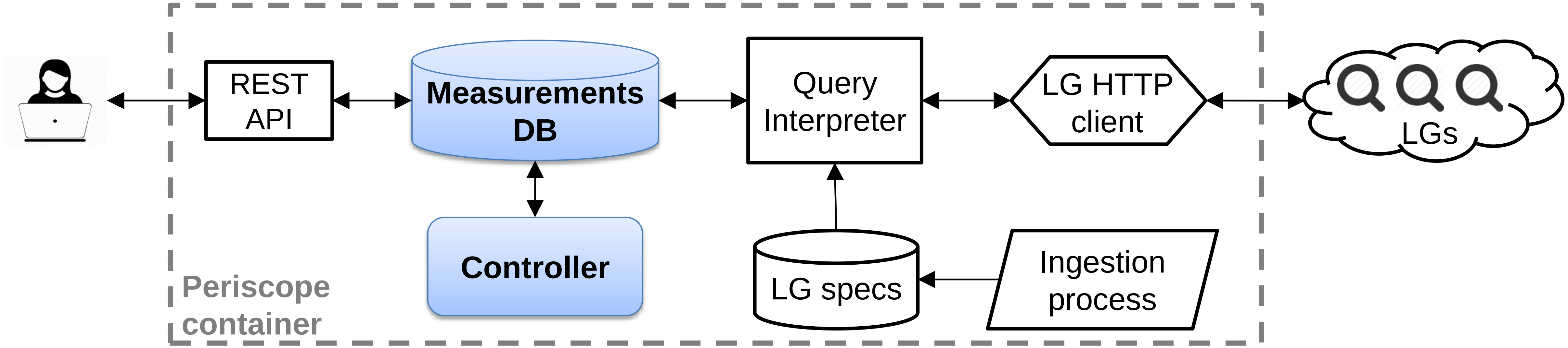
Periscope Architecture v0.1



- The **Query Interpreter** uses the outcome of the Ingestion Process to translate standardized API queries to LG-specific formats.
- The **LG Client** executes the native HTTP queries, and returns the raw HTML responses for parsing by the Query Interpreter.

What about rate limits?

Periscope Architecture v0.2



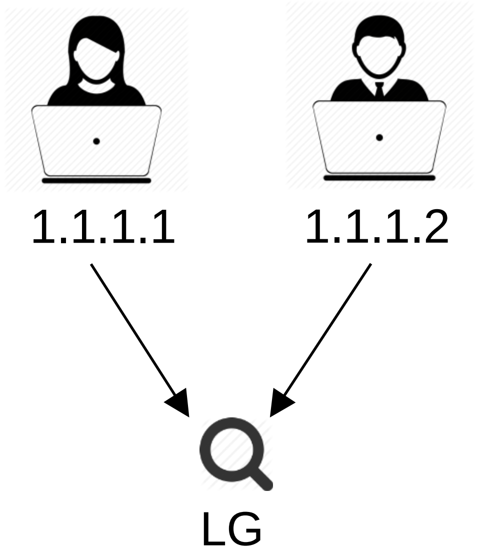
- API requests are not executed immediately, but first queued in the **Measurements DB** as pending jobs.
- The **Controller** oversees the rate limits and decides when to allocate the pending jobs to the Query Interpreter.

Periscope enforces per-user and per-LG query rate limits

- LGs may communicate their limits explicitly (through disclaimers), or implicitly (through error codes).
- Two limits control the rate of issued LG queries:
  - **User-specific**: Each user can issue only 1 query per 5 minutes to the same LG.
  - **LG-specific**: Each LG will execute up to 3 queries per minute from all the users.
- A query is allocated if neither limit is exceeded.
- Exponential back-off when LGs respond with errors.

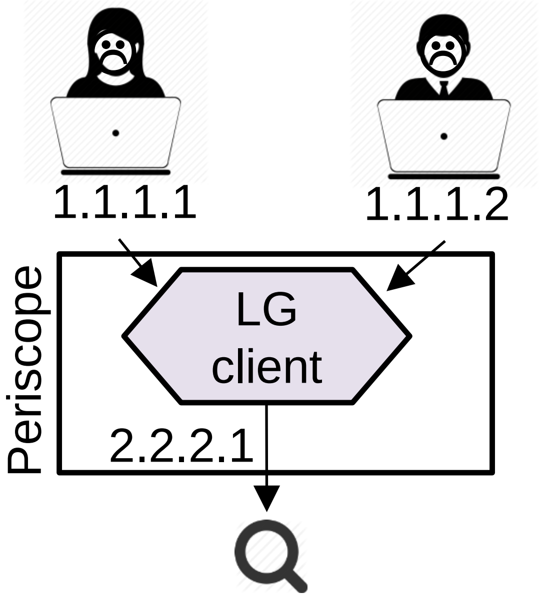
Support for multiple concurrent users requires multiple LG clients

Native LG querying



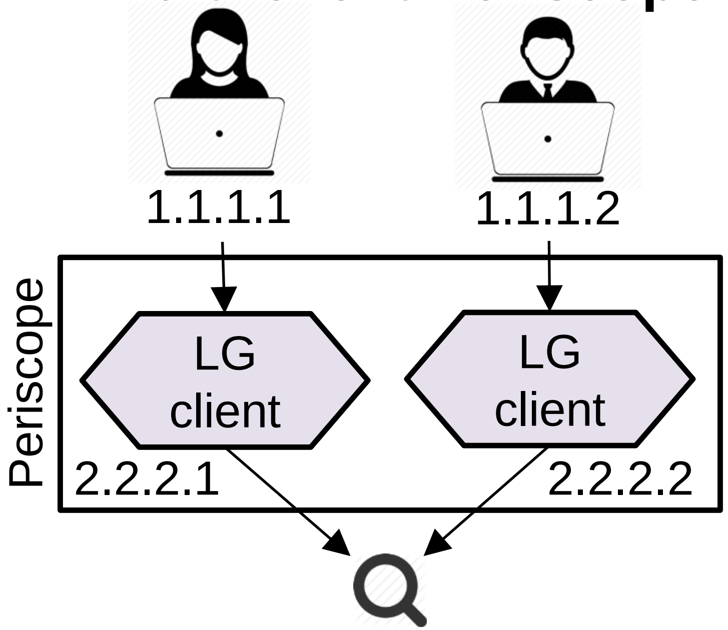
LGs use the users' IP address to impose per-user querying quotas

✘Single-client Periscope



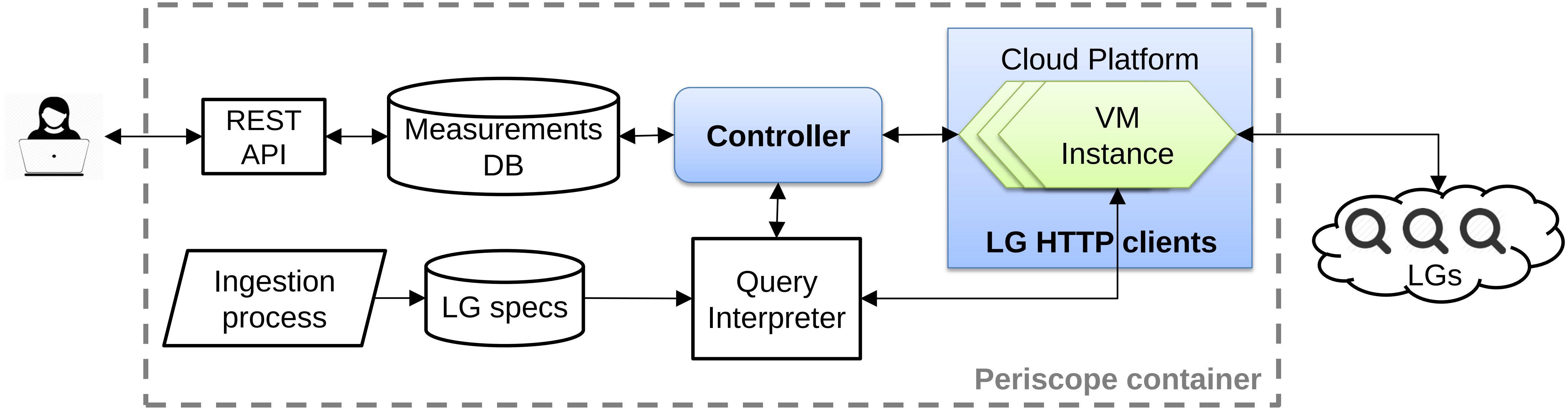
Putting multiple Periscope users behind the same IP causes all the users to share the quotas of a single user

✔Multi-client Periscope



Using different client per user allows Periscope to provide the same querying quotas as native querying

Periscope Architecture v1.0



- For each Periscope User the controller allocates a different cloud-hosted **VM instance** to execute the user queries.
- Each VM instance takes an IP address from the cloud operator's address space.
- **The same rate limits are still enforced.**

Transparency of Periscope requests

- Periscope sets three custom HTTP request headers:
  - **"X-Request-Origin: periscope"**
  - **"X-Request-For:<*user-ip*>"**
  - **"X-Request-Client:<*gcloud* OR *aws* OR *ark*>"**
- IP addresses used by Periscope LG Clients are configured with the appropriate reverse DNS record:
  - client.periscope.lg

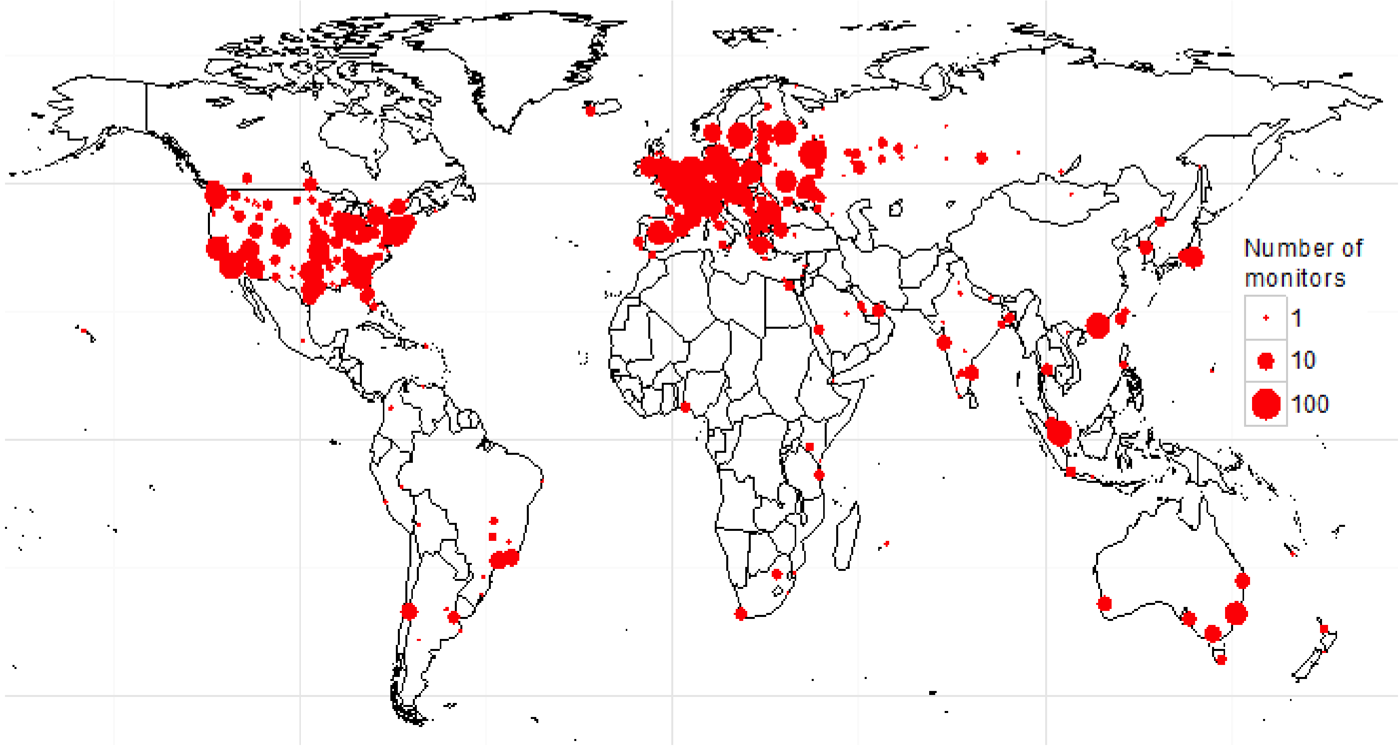
User accountability

- Periscope uses a 1-to-1 mapping between users and LG Client IP addresses:
  - A static VM Instance corresponds to each user.
  - Each VM Instance is assigned with a static IP address.
  - User-specific blocking works in the same way as native LG querying.

- Periscope maintains historical logs for every measurement:
  - Violations can be traced back to the responsible users

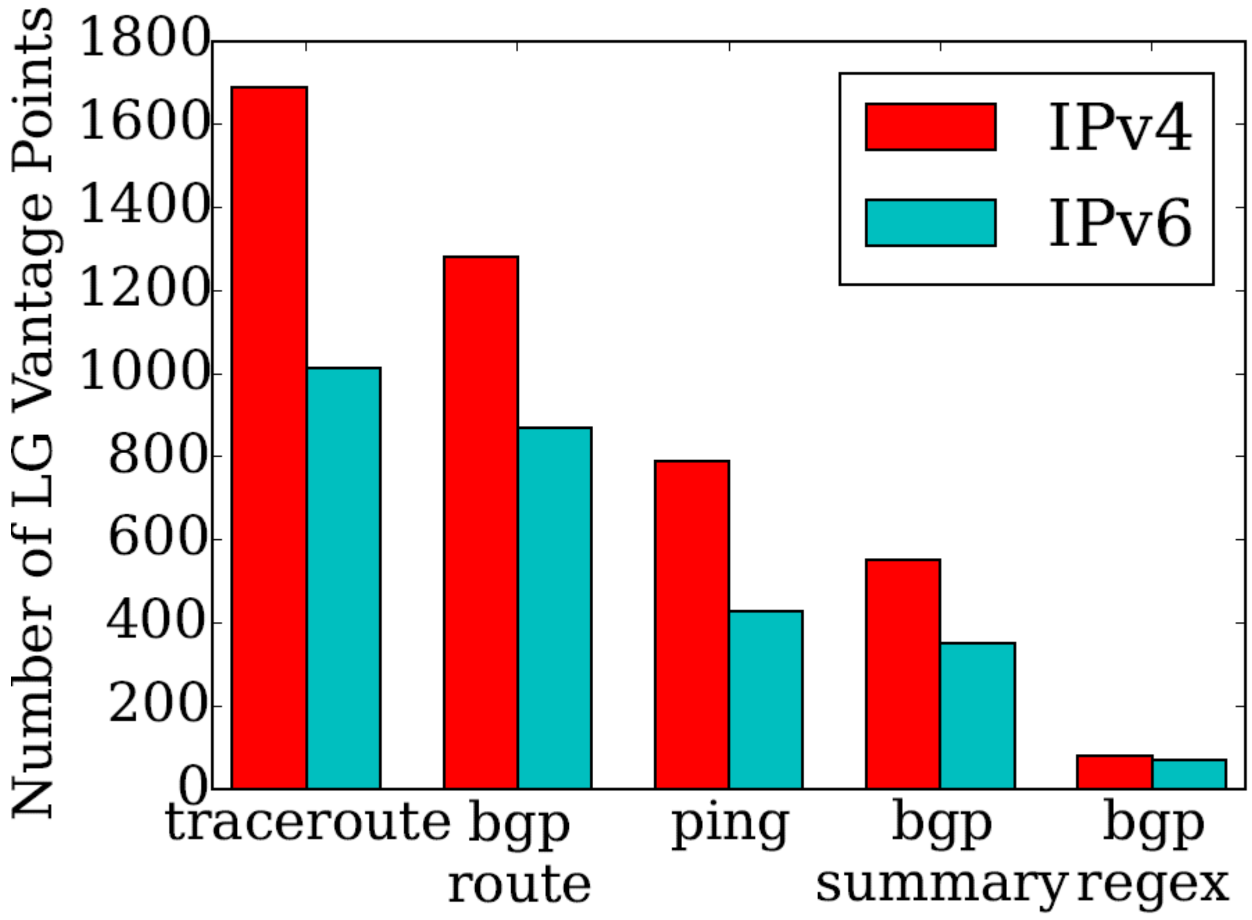
Coverage of Periscope LGs

- Monitors:
  - 572 ASNs with 2,951 VPs.
  - 77 countries, 492 cities.
- To geo-locate LGs:
  - Use locations encoded in LG interfaces.
  - Geo-locate the source IP of the LG using NetAcuity.

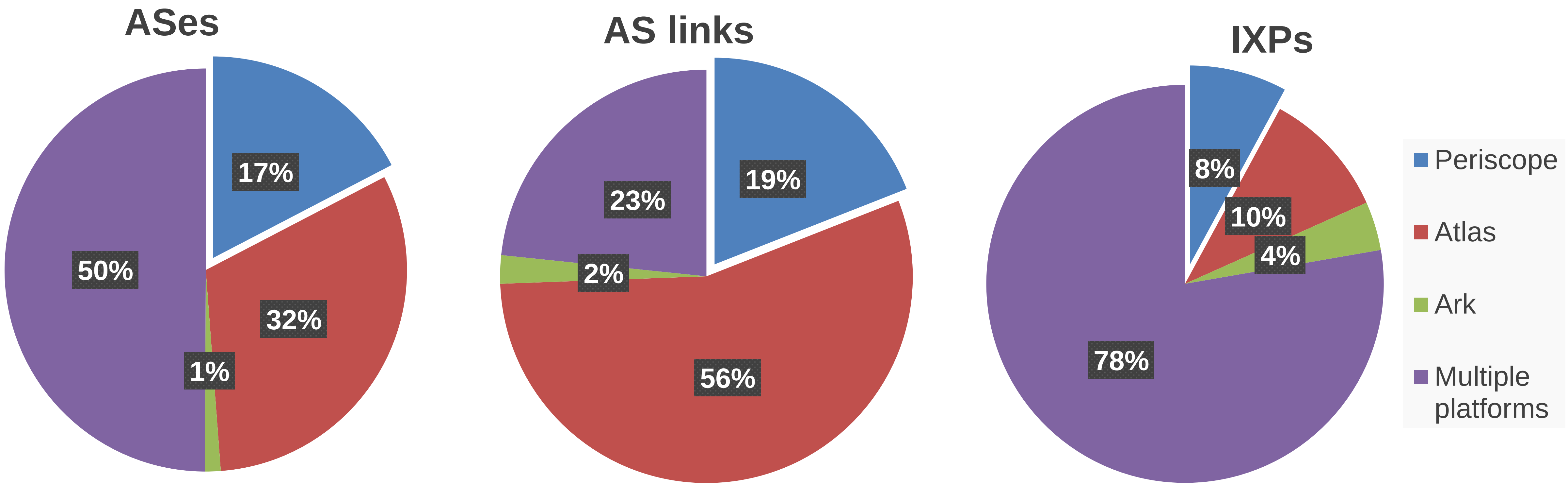


Commands supported by Periscope LGs

- Over 75% of the LG nodes provide both traceroute and BGP commands.
- Over 60% of the LGs support IPv6 queries.



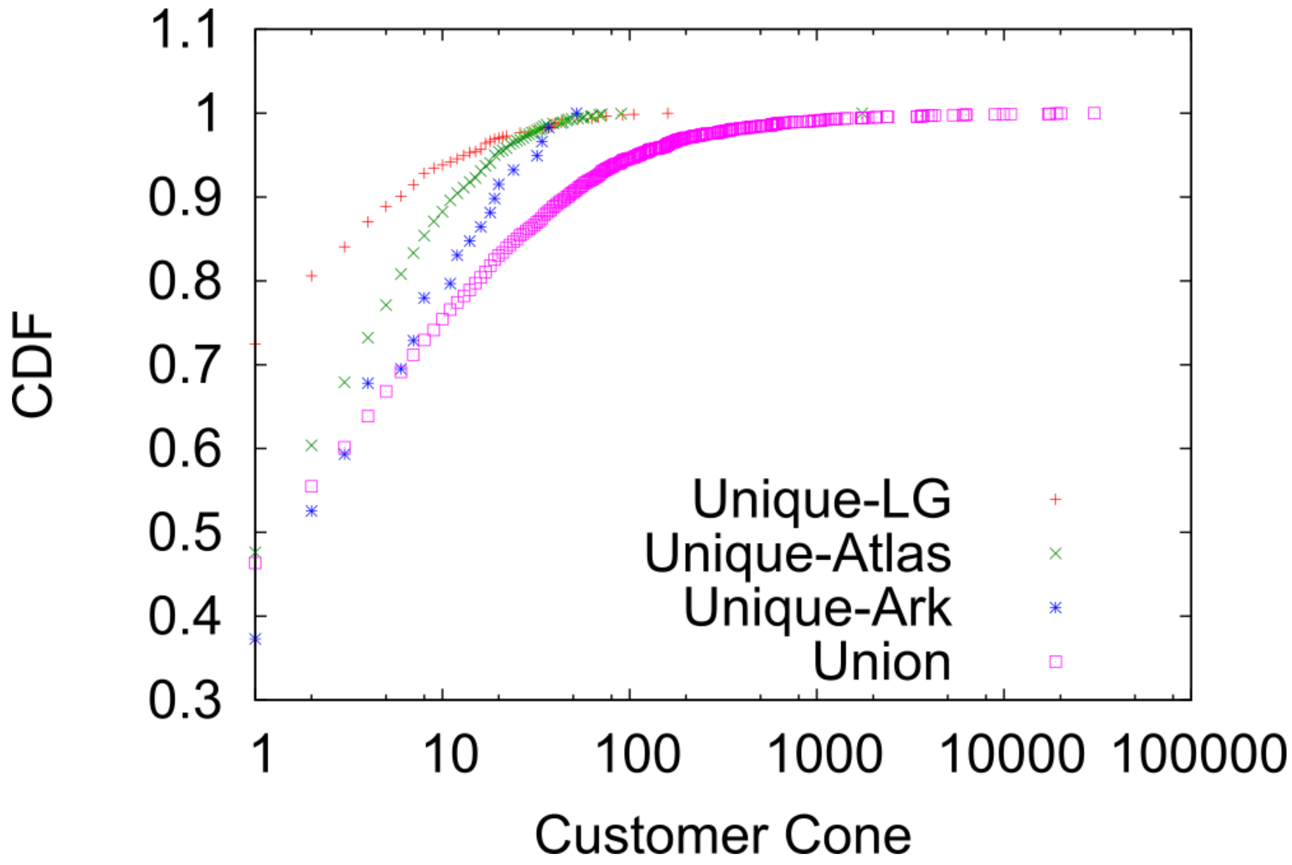
LGs capture largely complementary topology compared to other platforms



Topology obtained after querying 2,000 randomly selected IPs from each LG (2K VPs), and from every VP available in RIPE Atlas (8K probes) and CAIDA's Ark (100 VPs) during August 2015.

Unique ASes in each dataset differ in terms of customer-cone sizes

- LGs tend to capture more peripheral and stub ASes.
- Ark and Atlas capture ASes with larger customer cone.



Intelligent Load Distribution

- Some LGs receive higher query loads than other LGs:
    - LGs in large providers receive more queries than their customers.
- Some queries can be satisfied by multiple LGs:
    - `“What is the path between Level3 and Cloudflare?”`
    - The path can be returned either by the Level3 LG, or by an LG that reaches Cloudflare through Level3
- Some queries can be satisfied by multiple platforms:
    - Overlap of vantage points among Atlas, Ark and Periscope.

Intelligent Load Distribution

```
BGP Routing table entry for 103.22.203.0/24
Paths: (2 available, best #1)

 6453 13335
  AS-path translation: 103.22.203.0/24
   edge4.Frankfurt1 (metric 0)
    Community: Europe Backbone_2 Lclprf_86 Germany Level3_Peer Frankfurt
    6453:3000 6453:3100 6453:3103
    Origin: IGP, metric 0, localpref 86, Used  Valid  Best  IGP  Group-Best
    Originator: edge4.Frankfurt1

 6453 13335
 AS-path translation: 103.22.203.0/24
  edge4.Frankfurt1 (metric 0)
   Community: Europe Backbone_2 Lclprf_86 Germany Level3_Peer Frankfurt
    6453:3000 6453:3100 6453:3103
    Origin: IGP, metric 0, localpref 86, Valid  IGP
    Originator: edge4.Frankfurt1
```

AS3356 Looking Glass
(high query load)

```
BGP routing table entry for 103.22.203.0/24, version 126601054
BGP Bestpath: deterministic-med
Paths: (2 available, best #2, table default)
  Advertised to update-groups:
     8
  Refresh Epoch 1
  3356 3356 6453 13335
    188.1.200.77 (metric 1141) from 188.1.200.81 (188.1.200.81)
      Origin IGP, metric 0, localpref 80, valid, internal
      Community: 680:66 3356:2 (Europe) 3356:22 3356:86 3356:501 (Germany)
                 3356:666 (Peer route) 3356:2065 (FRF - Frankfurt) 6453:3000
                 6453:3100 6453:3103
      Originator: 188.1.200.77, Cluster list: 188.1.200.81
      rx pathid: 0, tx pathid: 0
  Refresh Epoch 1
  3356 3356 6453 13335
    188.1.200.77 (metric 1141) from 188.1.200.77 (188.1.200.77)
      Origin IGP, metric 0, localpref 80, valid, internal, best
      Community: 680:66 3356:2 (Europe) 3356:22 3356:86 3356:501 (Germany)
      3356:666 (Peer route) 3356:2065 (FRF - Frankfurt) 6453:3000 6453:3100
      6453:3103
```

AS680 Looking Glass
(low query load)

Utility Optimization

- If the same path can be satisfied by multiple LGs, allocate the query to the LG with the lowest load.
- Learn which LGs use the same paths through historical measurements and CAIDA's AS relationships dataset.
- Use historical measurements from multiple platforms to improve the selection of alternative LGs.

4x increase in satisfied queries for the same querying budget

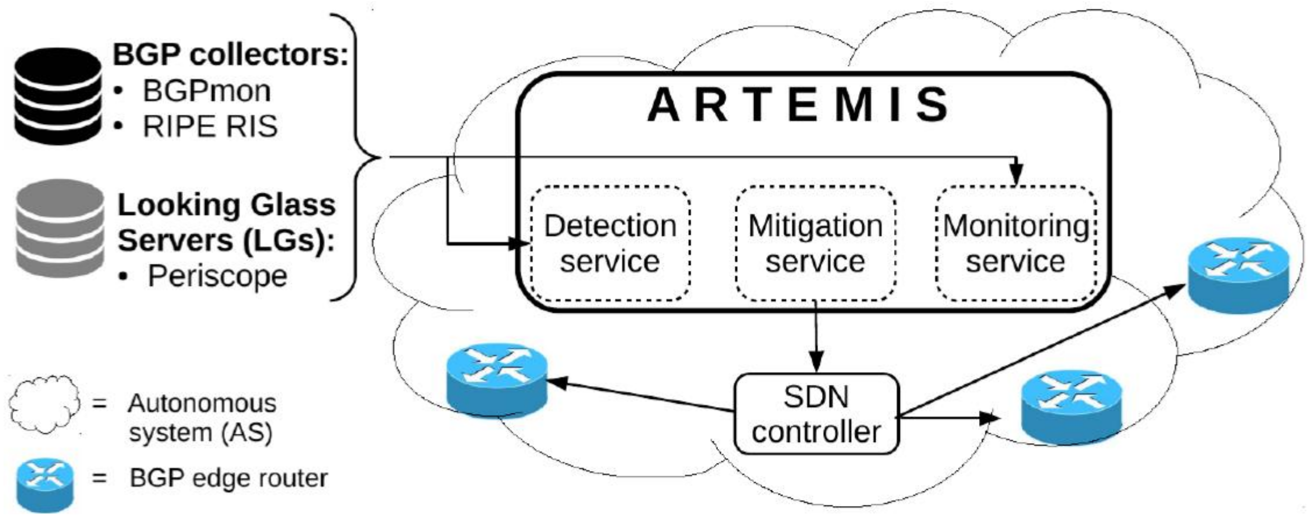
Cunha Í, Marchetta P, et al. Sibyl: a practical Internet route oracle. In13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16) 2016 (pp. 325-344).

Reduction of query redundancy

- In native LG querying users are not aware of archived and concurrent queries:
  - Redundant queries to popular destinations (e.g. 8.8.8.8)*.
- Periscope has a bird's-eye view of measurements across users and across LGs:
  - Redundant queries are bundled together to reduce query load.
- Measurement results are made publicly available:
  - Users can consume completed queries instead of issuing new.

* https://www.reddit.com/r/sysadmin/comments/1f9kv4/what_are_some_public_ips_that_are_ok_to/
http://superuser.com/questions/769005/what-is-a-external-reliable-ip-address-to-ping-to-check-if-internet-is-available
http://serverfault.com/questions/132805/why-do-we-ping-the-ip-4-2-2-2-to-test-connectivity

Case study 1: Prefix Hijack Detection



- Combining Periscope with passive BGP collectors enables the detection of prefix hijacks in less than 1 minute.
- Faster detection in 60% of the hijack cases compared to using only BGPMon.

G. Chaviaras et al. ARTEMIS: Real-Time Detection and Automatic Mitigation for BGP Prefix Hijacking. *ACM SIGCOMM 2016*

Case study 2: Troubleshooting Network Disruptions

- NetDiagnoser (ND): Identify the location of network failures through distributed traceroutes:
  - Unresponsive hops (*) or private IP address can cause errors
- Combine BGP and traceroute data from LGs to infer the ASes of the unidentified hops:
  - Up to 60% improvement in correct diagnoses in simulations

Periscope enables the practical implementation of the simulation-based methodology

Dhamdhere, Amogh, et al. "NetDiagnoser: Troubleshooting network unreachabilities using end-to-end probes and routing data." *ACM CoNEXT* 2007.

Current status

- Periscope is accessible after email request: periscope-info@caida.org
- API documentation: http://www.caida.org/tools/utilities/looking-glass-api/
- Ongoing and future work:
  - Development of a graphical user interface.
  - Hosting of Periscope clients inside Ark monitors.
  - Improve optimization of LG utilization through cross-platform interoperability.

Request for Contributions

- Please contribute feedback regarding:
  1. Per-user query limits
  2. Global query limits
  3. Opt-in or opt-in requests
- Utilization statistics and archived queries.
- Infrastructure support (e.g. VM instances, cloud-computing credit).

Contact us at:
periscope-info@caida.org

Conclusion

- Periscope goals:
  - Unify LGs under a standardized overlay API.
  - Enforce per-user and per-LG rate limits.
  - Provide transparency and accountability.
- Benefits of Periscope:
  - Extends topology coverage.
  - Optimizes LG utilization.
  - Improves troubleshooting capabilities.

Questions?

BACKUP SLIDES

Periscope uses path prediction to optimize query distribution

- Periscope is based on SIBYL to predict which LGs will return the same path, and selects the LG with the lower query load:
  - Use **previously issued measurements** to predict **unmeasured** paths.
  - Use **path splicing** when there are no archived measurements for the all the possible (source,destination) pairs.
  - Use **RuleFit** to assess the confidence in the predicted paths.

Cunha Í, Marchetta P, et al. Sibyl: a practical Internet route oracle. In13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16) 2016 (pp. 325-344).

Optimization of Load Distribution across measurement platforms

Traceroute to caida.org (192.172.226.78), 48 byte p

1	193.174.247.1	kr-sgs1-vlan100.sgs.dfn.de	AS680	1.82ms
2	188.1.230.45	cr-fra2-pwether1.x-win.dfn.de	AS680	7.656ms
3	62.40.124.217	dfn.mx1.fra.de.geant.net	AS21320	7.173ms
4	62.40.125.18	internet2-gw.mx1.fra.de.geant.net	AS21320	104
5	198.71.45.6	et-7-3-0.4072.rtsw.atla.net.internet2.edu	AS11537	
6	198.71.45.13	et-10-2-0.105.rtr.hous.net.internet2.edu	AS11537	
7	198.71.45.21	et-7-1-0.4070.rtsw.losa.net.internet2.edu	AS11537	
8	137.164.26.200	hpr-lax-hpr2--i2-r&e.cenic.net	AS2153	174.1
9	137.164.26.34	hpr-sdsc-10ge--lax-hpr.cenic.net	AS2153	176.3
10	192.12.207.10	medusa-mx960.sdsc.edu	AS195	176.49ms
11	192.172.226.78	rommie.caida.org	AS1909	176.606ms

AS680 RIPE Atlas probe

```
Type escape sequence to abort.
Tracing the route to ns1.caida.org (192.172.226.78)
VRF info: (vrf in name/id, vrf out name/id)
  1 xr-fzk1-pc2.x-win.dfn.de (188.1.145.81) [MPLS: Label 1274
  2 cr-fra2-be9.x-win.dfn.de (188.1.144.121) 4 msec 8 msec 4 m
  3 dfn.mx1.fra.de.geant.net (62.40.124.217) [AS 20965] 4 msec
  4 internet2-gw.mx1.fra.de.geant.net (62.40.125.18) [AS 2096
  5 et-7-3-0.4072.rtsw.atla.net.internet2.edu (198.71.45.6) [
  6 et-10-2-0.105.rtr.hous.net.internet2.edu (198.71.45.13) [
  7 et-7-1-0.4070.rtsw.losa.net.internet2.edu (198.71.45.21)
  8 hpr-lax-hpr2--i2-r&e.cenic.net (137.164.26.200) [AS 2153]
  9 hpr-sdsc-10ge--lax-hpr.cenic.net (137.164.26.34) [AS 2153]
 10 medusa-mx960.sdsc.edu (192.12.207.10) [AS 195] 180 msec 1
 11 ns1.caida.org (192.172.226.78) [AS 1909] 188 msec 176 msec
```

AS680 Looking Glass