Directory Service Defense (DSD)

Mark Kosters
CTO

Whois/Whois-RWS/RDAP

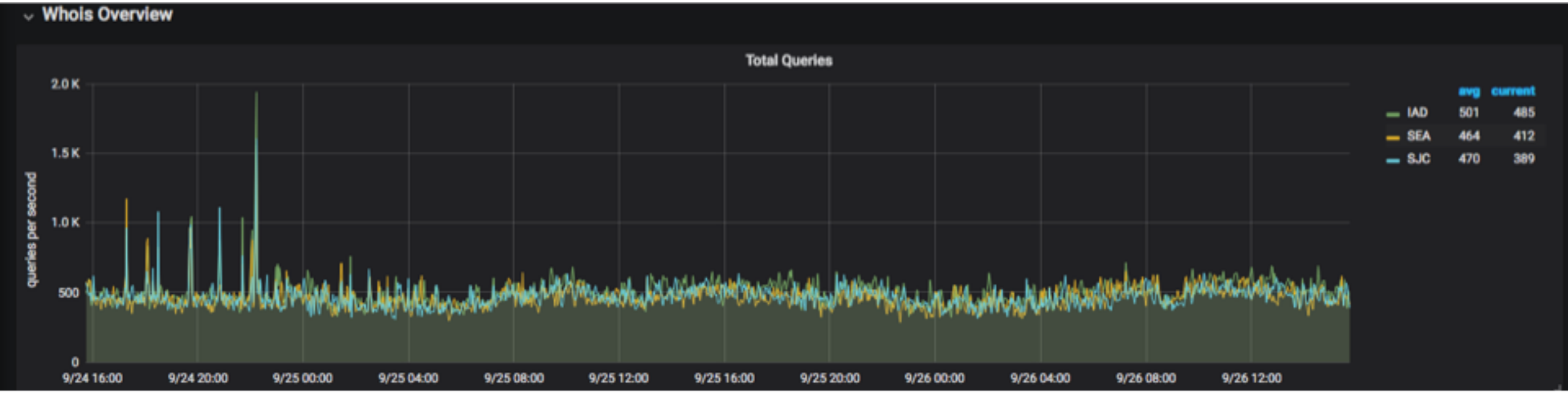
- Directory services is Whois, Whois-RWS, RDAP, and soon RPSL
- Goal for directory services availability is for people to query the service and receive results in a reasonable amount of time while abiding with the Whois Terms of Service
- Some automation is expected to be in the mix
- On automation, if the query rate is too high, that user may be tarpitted

Directory Service Abuse

- Directory service (Whois/Whois-RWS/RDAP) abuse continues
- Talked about this at ARIN 40, 41, and now 42
- Each incident requires a team response to look at the system, identify the abusers, notify the abuser, and potentially deny access to the abuser
  - Interrupts sleep or work (or both if it is a long-term event)
  - Does not scale
- Terms of Use talks about what the acceptable reasons why you can use the data (https://www.arin.net/whois_tou.html)
  - Does not talk about acceptable query rates

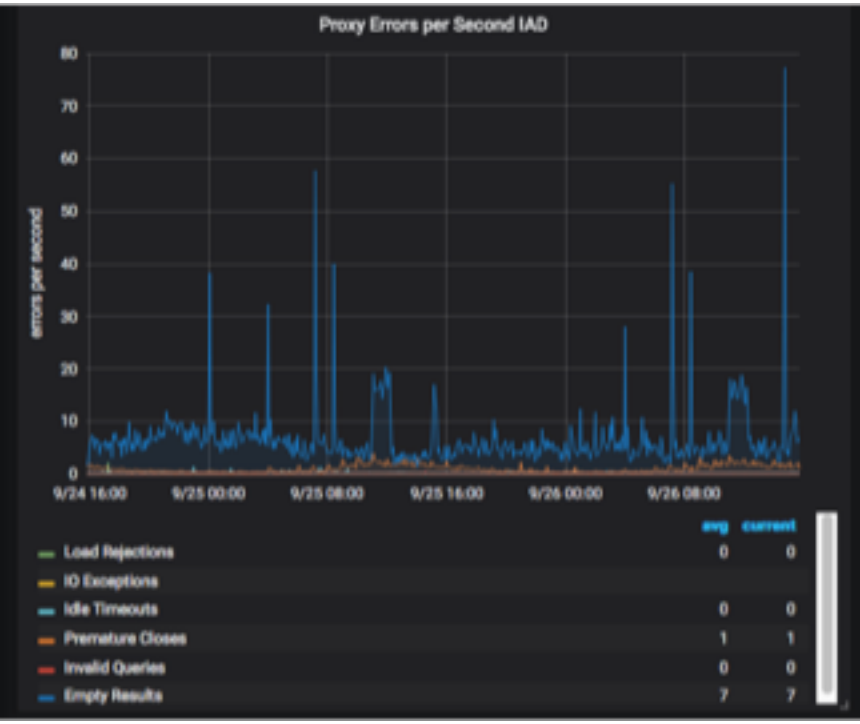
Abuse Example

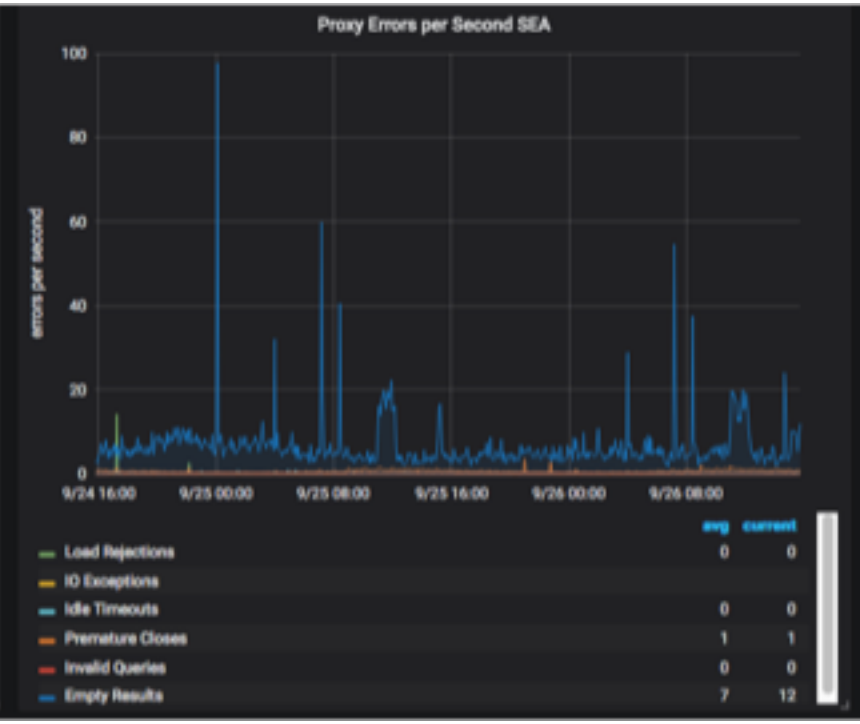
Normal Day

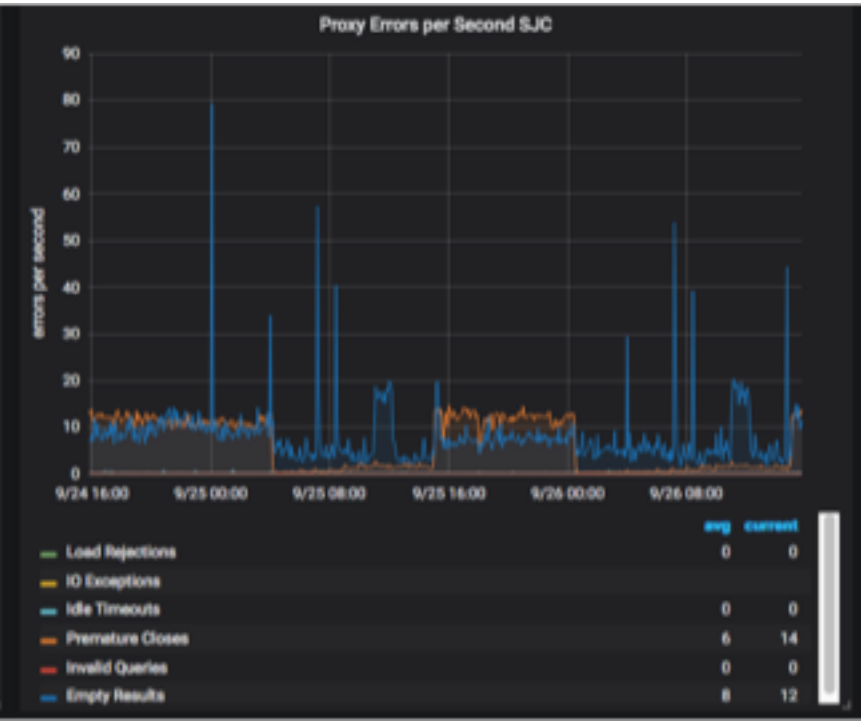


Abuse Example

Normal Day

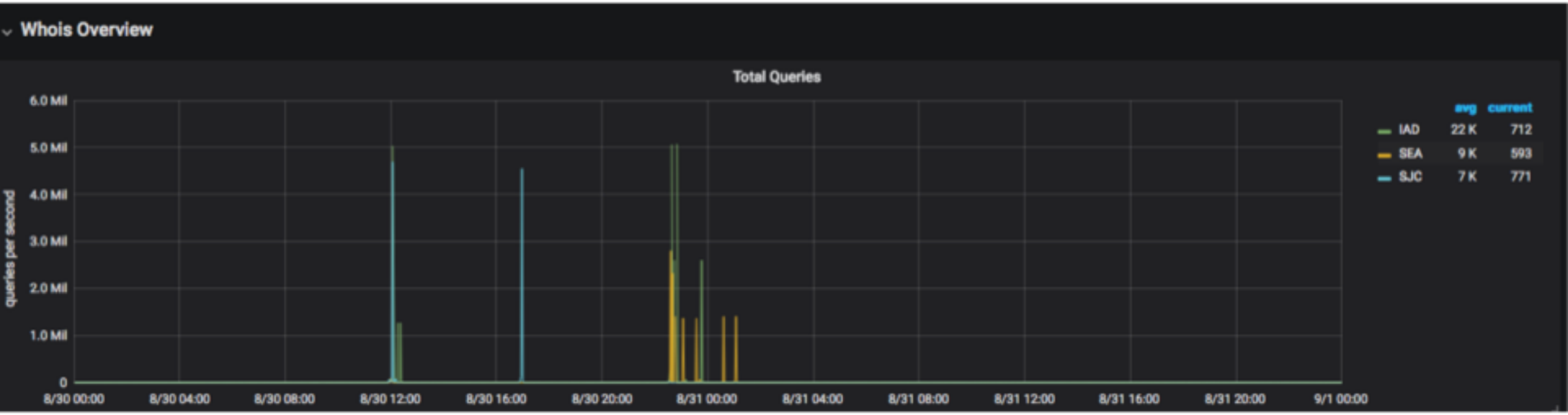






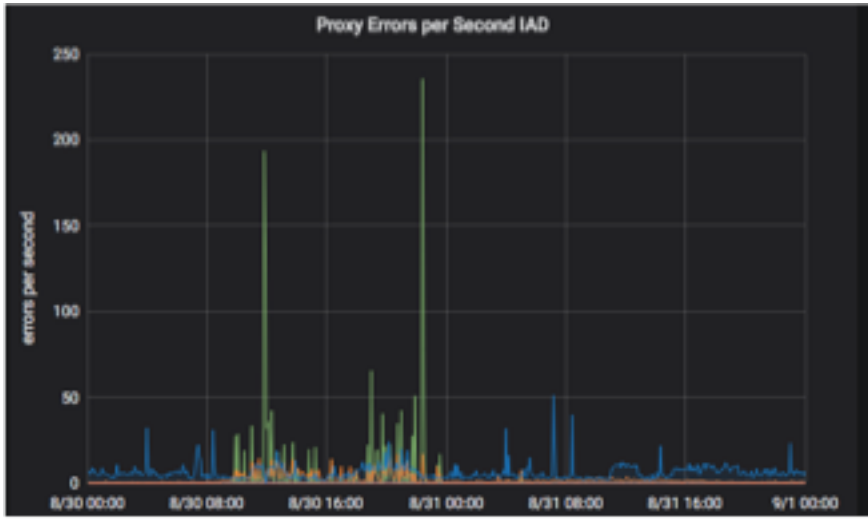
Abuse Example

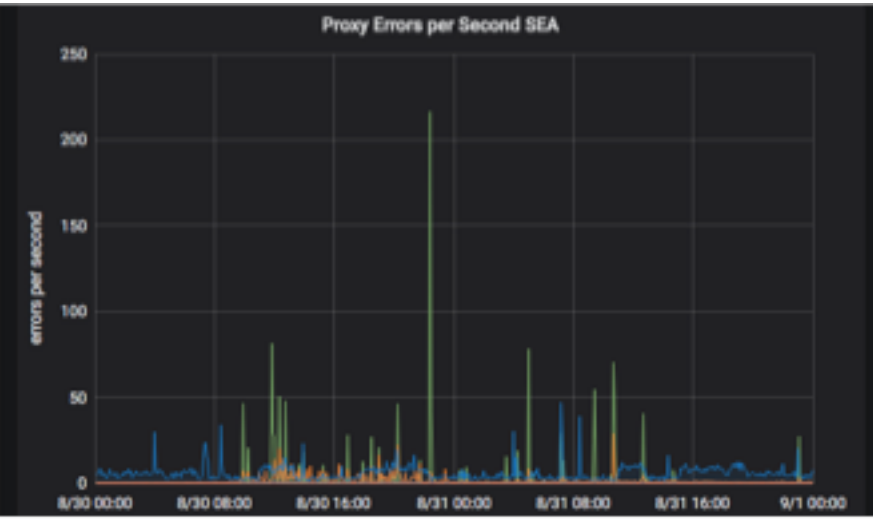
Period of abuse – scale on Y axis moves from thousands to in millions for queries received

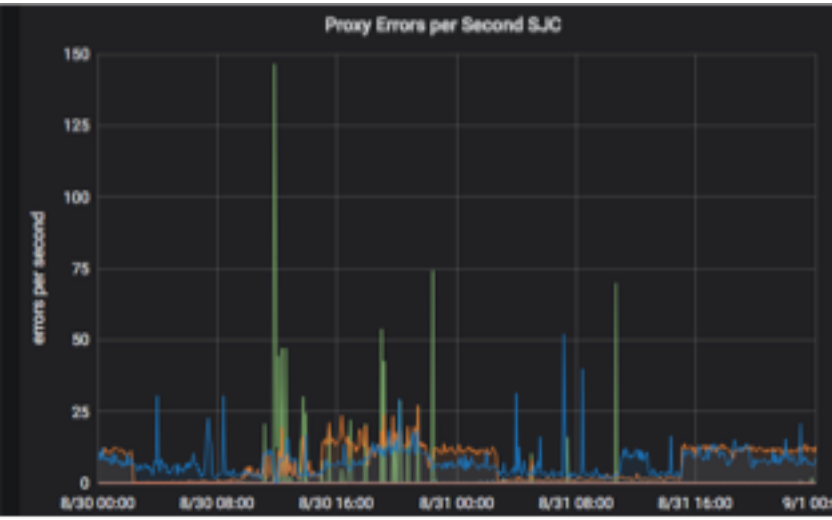


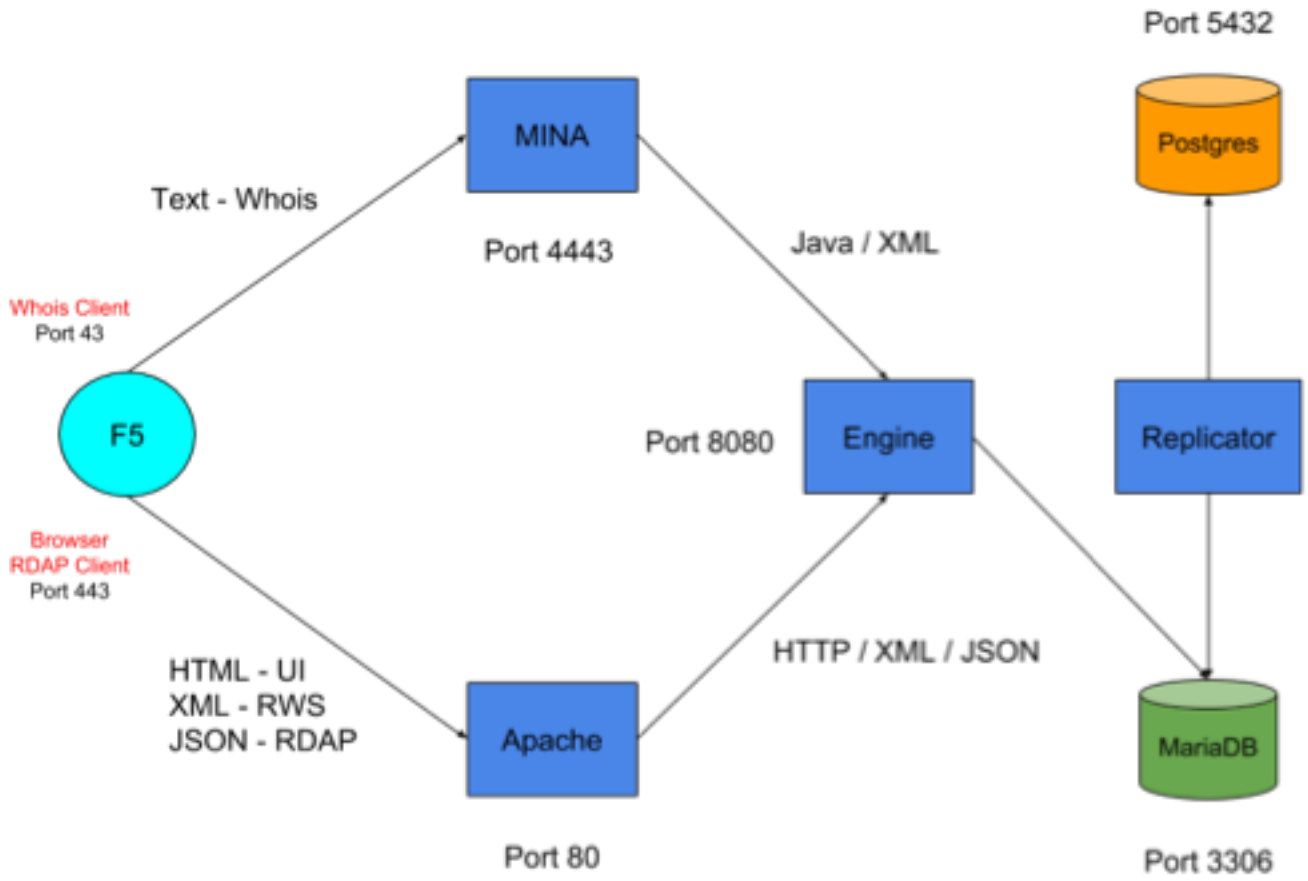
Abuse Example

Note the green lines – those are load rejections based on the high load









Whois Architecture

Traffic Trends

- 110 million requests per day
  - 90% Whois on port 43
  - 6.9% Whois-RWS
  - 3.0% RDAP
  - 0.1% Looking up whois information on ARIN's website (UI)

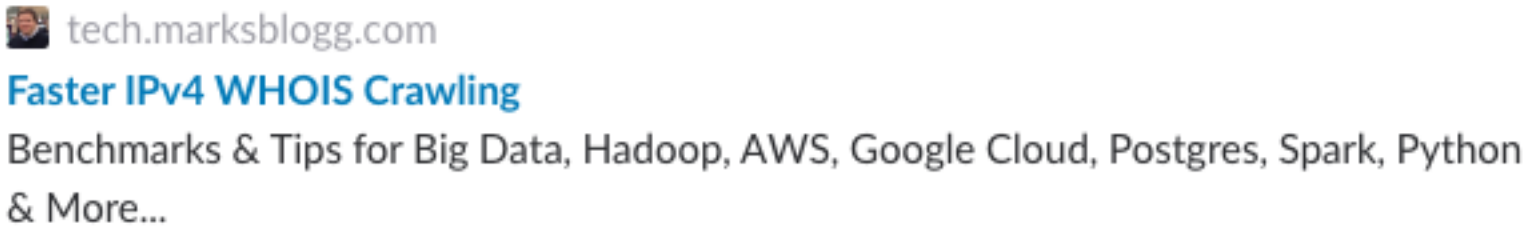
- For a Whois node (6 nodes per site, 3 sites):
  - 63 queries per second for Port 43 (Whois) requests
  - 7 queries per second for RWS/RDAP/UI HTTP requests

Traffic Trends (Continued)

- Traffic surges
  - Mostly peaks at 3 times normal traffic
  - Bots are most definitely involved (see next page)
- Many requests answered from the cache
  - Many requests answered from the cache
  - Have effective caching built in (up to 90% hit rate)
- Heavier requests not served by cache can cause:
  - Database connections peak
  - CPU usage surge

Cookbooks for Use/Abuse

- Distributed abuse of directory services
- Multiple types of abuse
  - Legitimate need but very high rates
  - Spinning up lots of VM's to target ARIN's directory services
  - Code freely available to query ARIN from AWS and others
  - http://tech.marksblogg.com/faster-ipv4-whois-crawling.html



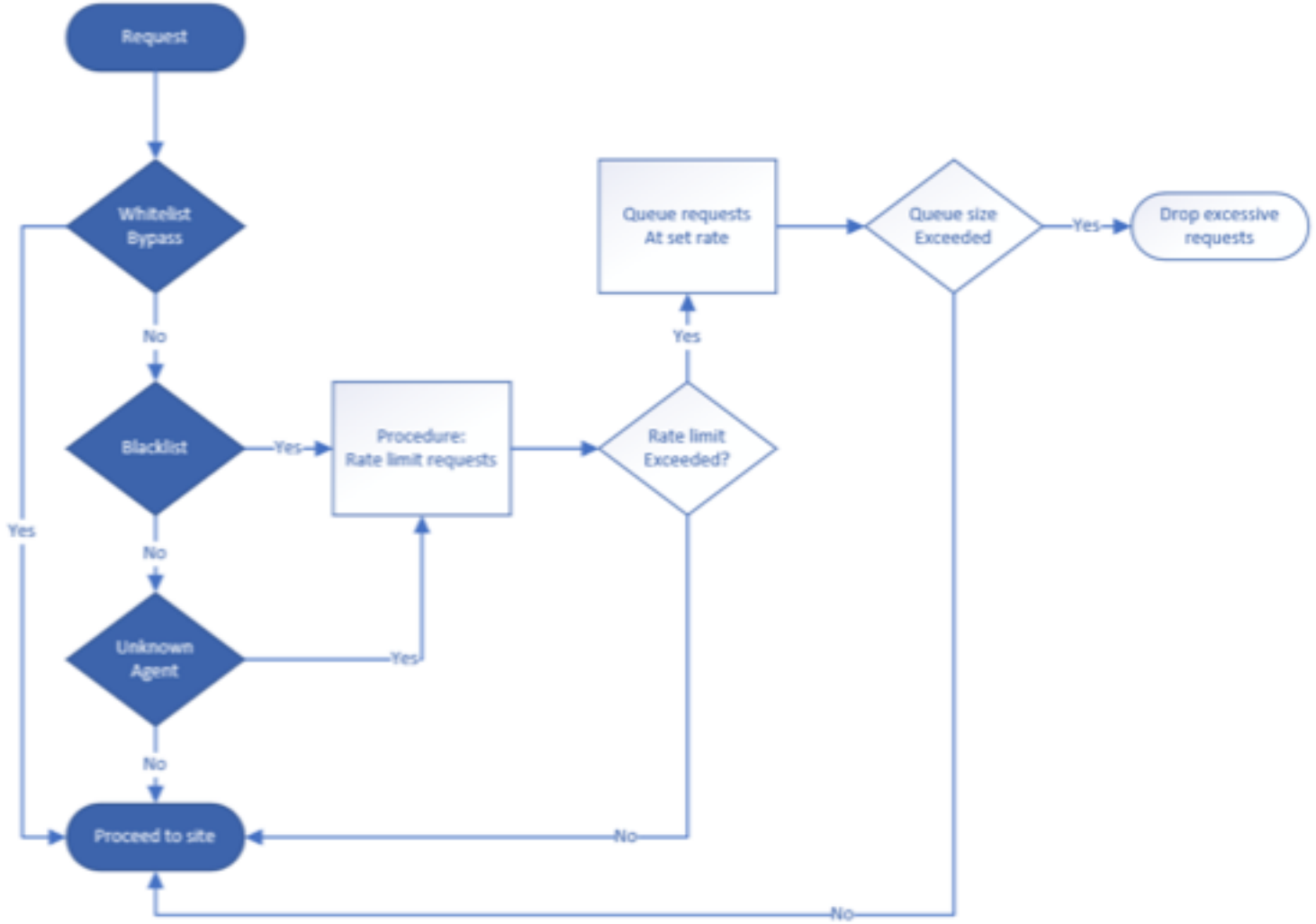
Improvements so far

- Better cache efficiency for Port 43, RWS, and RDAP traffic (have up to 90% cache hit rates)
- Direct Java calls from Proxy to Engine instead of HTTP calls that we previously used
- Optimized various queries that where expensive calls to the database (for example: quickly reject whois queries for domain names e.g. example.com)
- Tarpitting installed on our F5 load balancers

Tarpitting

- Addressed the automated abuse of our directory services with a concept called tarpitting.
- How tarpitting works:
  - If the rate limit is exceeded, any queries over that rate limit are put on a queue.
  - This queue is looked at every 2 seconds and queries are then allowed to be processed as long as the current queries do not exceed the limit.
  - If the rate is sustained and the queue limit has been met, then the queries on the queue are popped off in a FIFO fashion with a TCP reset back to the source.

Tarpitting diagram



Tarpitting example

- Here are some examples with 100 queries per second as the threshold:
- Example A: 105 queries per second done in a burst - 100 are immediately processed and 5 are put on the queue. In 2 seconds, the 5 on the queue are then processed.
- Example B: 105 queries done in a second, followed by a sustained rate of 99 queries per second. At every 2 second interval, 1 query is taken off the queue to be processed.
- Example C: 105 queries per second at a sustained rate. At every second interval, 5 of those queries are put on the queue with the remaining 100 being processed. At every 2 second interval, 0 queries are processed. Since the queue continues to grow at 5 added to the queue every second, it will hit the queue limit with 10 queries every two seconds being issued TCP resets back to the sender.

???

Thank you.

Any Questions?