Running a highly available, ad-blocking, private DNS setup in Kubernetes

Cool Overengineered DNS setup

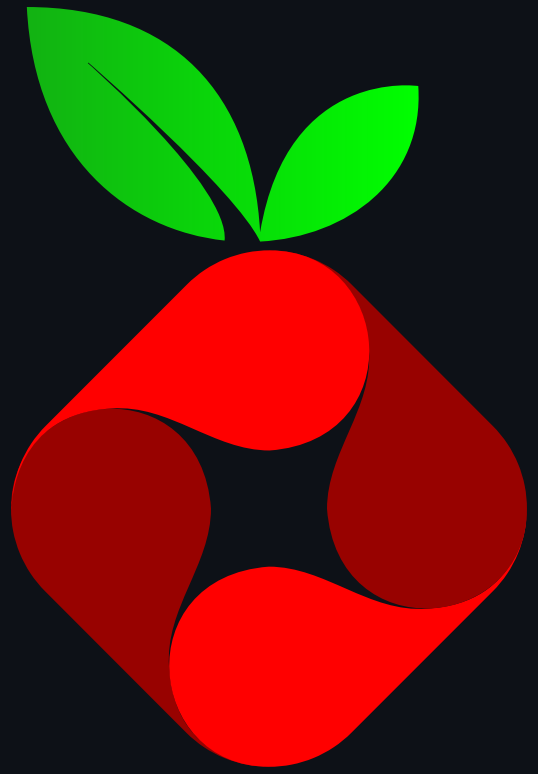
Who am I

- Nadia Santalla (she/her 🏳️‍⚧️)
  - 🌐 https://nadia.moe
  - ✉️ `mailto:nadia@nadia.moe`
- Senior software engineer @ Grafana Labs
- I read books and collect machine tools
- ❤️ All things infrastructure

Agenda

1. in
2. Wishlist
3. Why ?
4. How ? (metallb)
5. DHCP (dnsmasq)
6. Upstream (dnscrypt-proxy)
7. Kubernetes (stateless-dns)
8. Caching (dnsmasq)

PiHole?

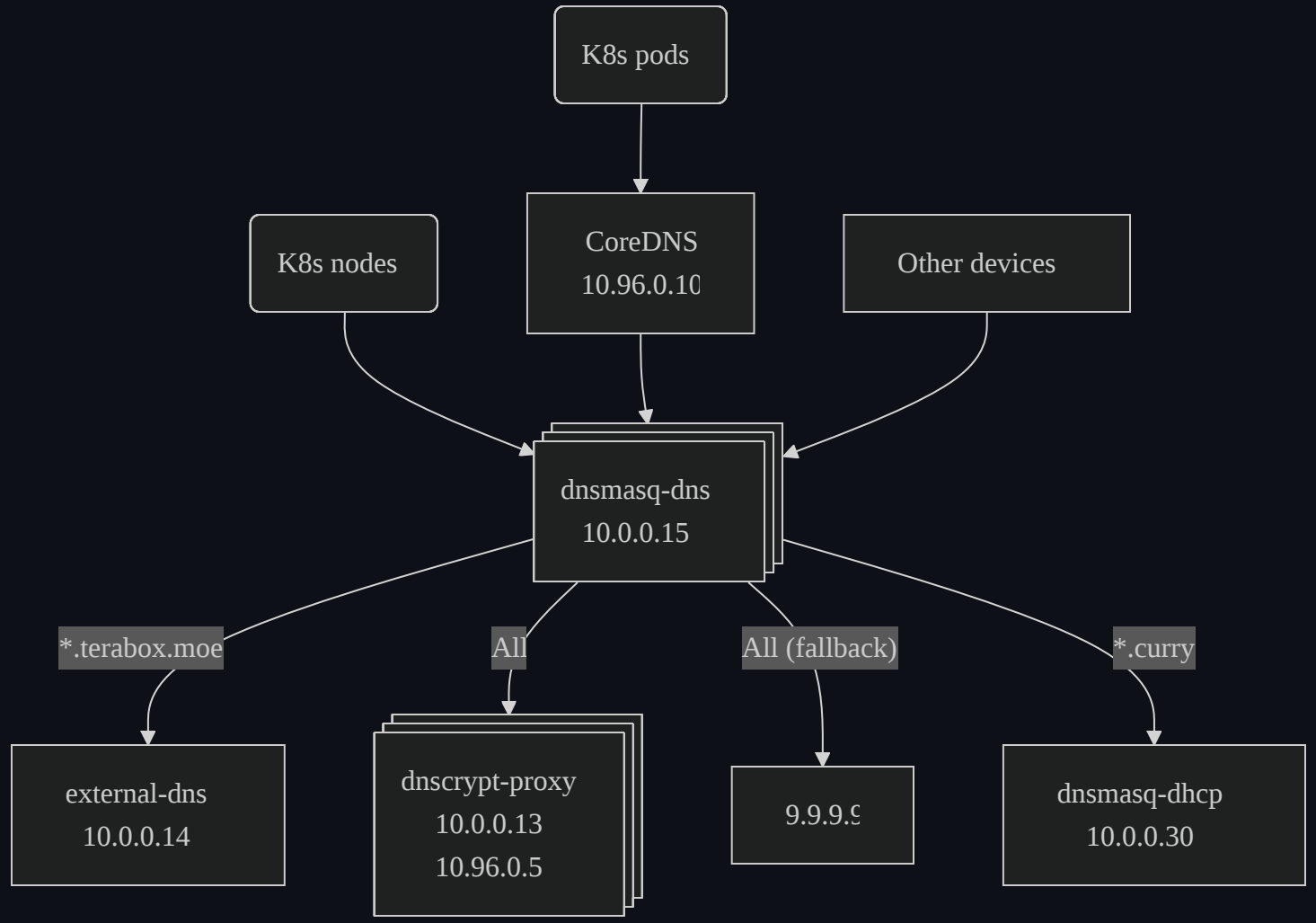


(c) https://pi-hole.net/

Wishlist

- Block ads
  - Exploitative
  - Malware vector
- Be fast
  - TTR weights a lot
- Be private
  - ISP DNS are questionable
  - Unencrypted DNS is dubious
- Be flexible
  - Unblock this or that
  - DHCP integration
  - Other NS integration
- Be reliable
  - DNS down = No internet
  - Updates need to happen

Spoiler

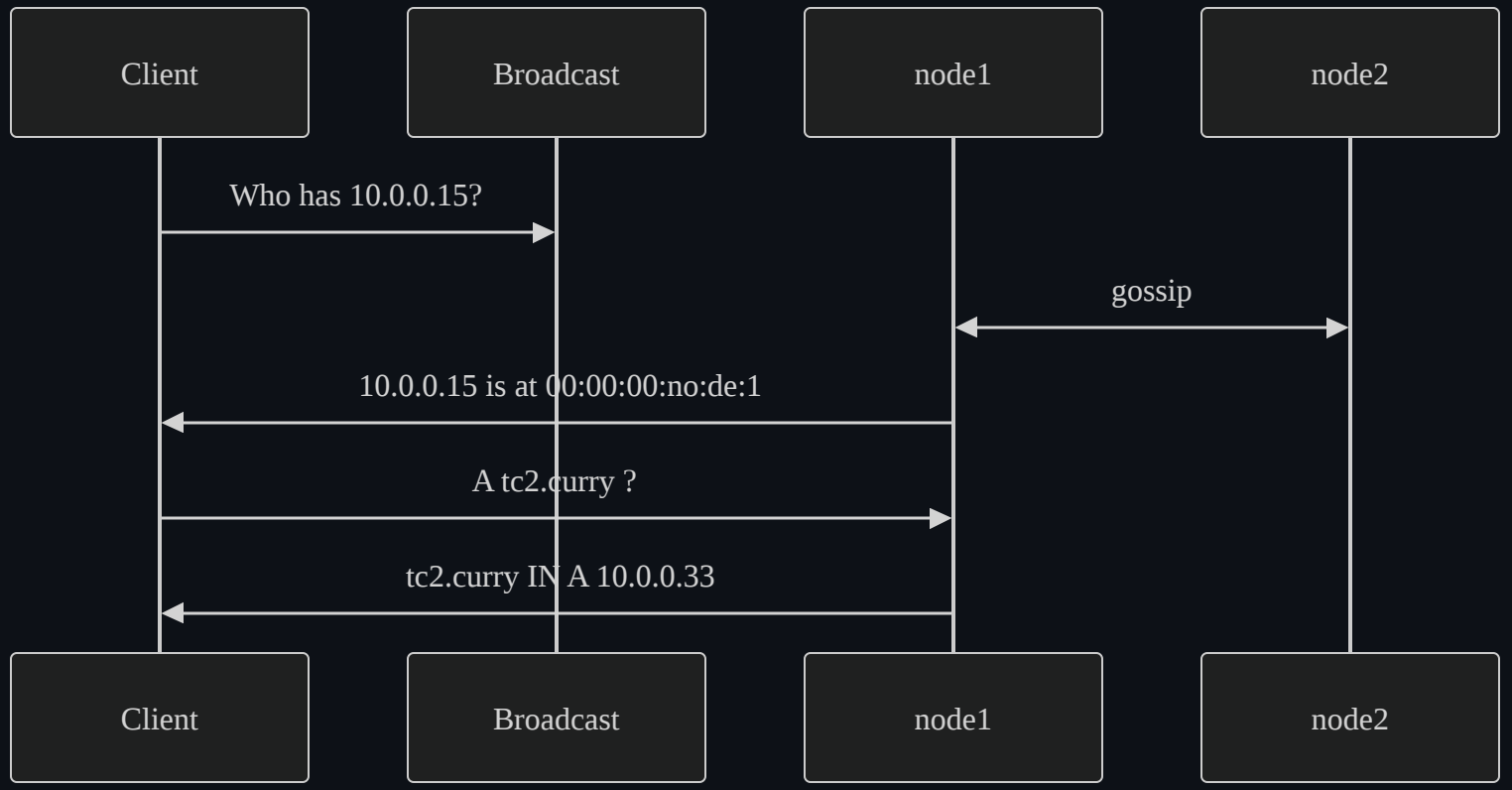

K8s pods
K8s nodes
CoreDNS
10.96.0.10
Other devices
dnsmasq-dns
10.0.0.15
*.terabox.moe
All
All (fallback)
*.curry
external-dns
10.0.0.14
dnscrypt-proxy
10.0.0.13
10.96.0.5
9.9.9.9
dnsmasq-dhcp
10.0.0.30

Why ☸ ?

- Failover & HA for free
  - Zero downtime rolling updates
  - For both containers and nodes
- Gitops for free
  - DNS records versioned and as text
- Easy o11y
  - Toss exporter, prometeheus SD
- Self-documenting
  - 4 nameservers talking to each other
- Already there

MetalLB ❤️

- Layer 3 failover
- Nameserver becomes a VIP
- HA for free
- ☑️ Be reliable



dnsmasq for DHCP

- Every network needs a DHCP server
- We're using dnsmasq somewhere else
- dnsmasq is nice
- DHCP/dns integration

- Some paper cuts
  - `/var/lib/misc/dnsmasq.leases`
  - `hostNetwork`
- Pinned to a node

DNSCrypt proxy as upstream

- Presents as a DNS server (UDP/53)
- Queries upstreams using dnscrypt
- Self-refreshing list of resolvers
- Automatically picks fastest

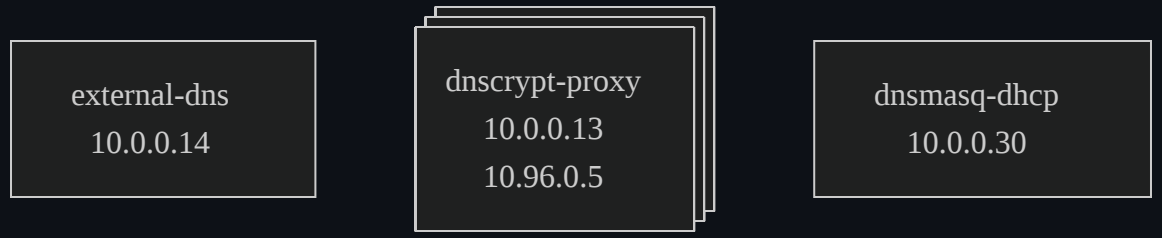
https://github.com/DNSCrypt/dnscrypt-proxy

Stateles-dns cluster services outside of the cluster

- External DNS meets PowerDNS
  - And it's stateless!
  - Ephemeral SQLite DB
- Authoritative NS for out-of-cluster clients
  - Exposes Ingresses and LoadBalancers
- txqueuelen/stateless-dns

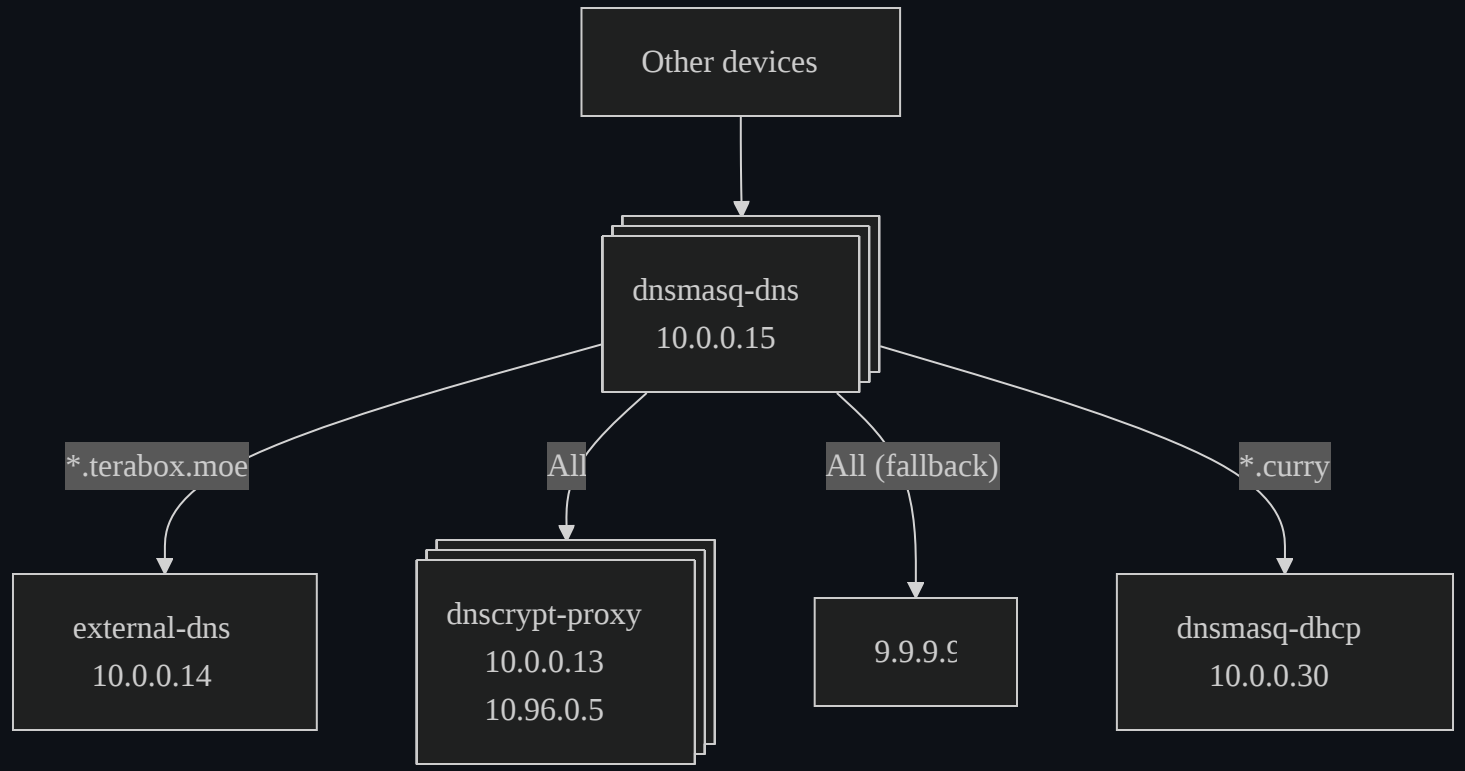
POWERDNS

What we got so far



dnsmasq for caching

- dnsmasq is a caching nameserver
  - Flexible and efficient
- Can be fed host denylists
- Can route requests to other NSs



dnsmasq for caching

```
# Never forward k8s names
local=/local/
local=/svc/
local=/svc.cluster/
local=/svc.cluster.local/

server=/curry/10.0.0.30#5353
server=/terabox.moe/10.0.0.14
server=/use-application-dns.net/

# Use upstream servers in strict priority order
strict-order
# dnscrypt-proxy svc with fixed clusterIP
server=10.96.0.5
# quad9 as fallback
server=9.9.9.9
```

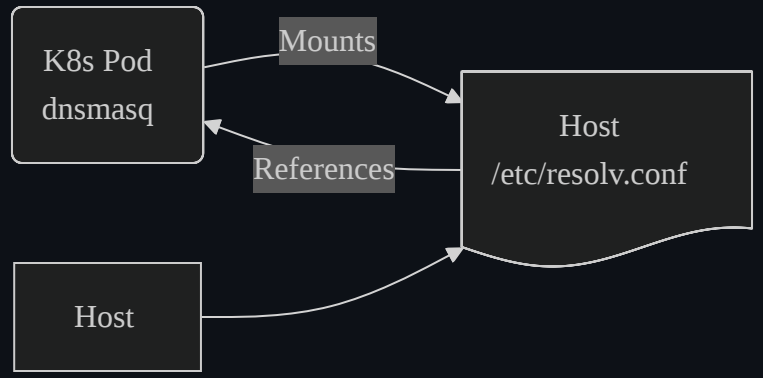
```
# Aggressively retry if we don't get a response
# within 100ms, up to 1000ms.
# dnsmasq applies exponential backoff.
fast-dns-retry=100,1000

cache-size=16384
no-negcache

# Be kind
bogus-priv
domain-needed

# K8s trash
no-hosts
# Include blocklist
addn-hosts=/dnsmasq/adblock.hosts
```

dnsmasq for caching

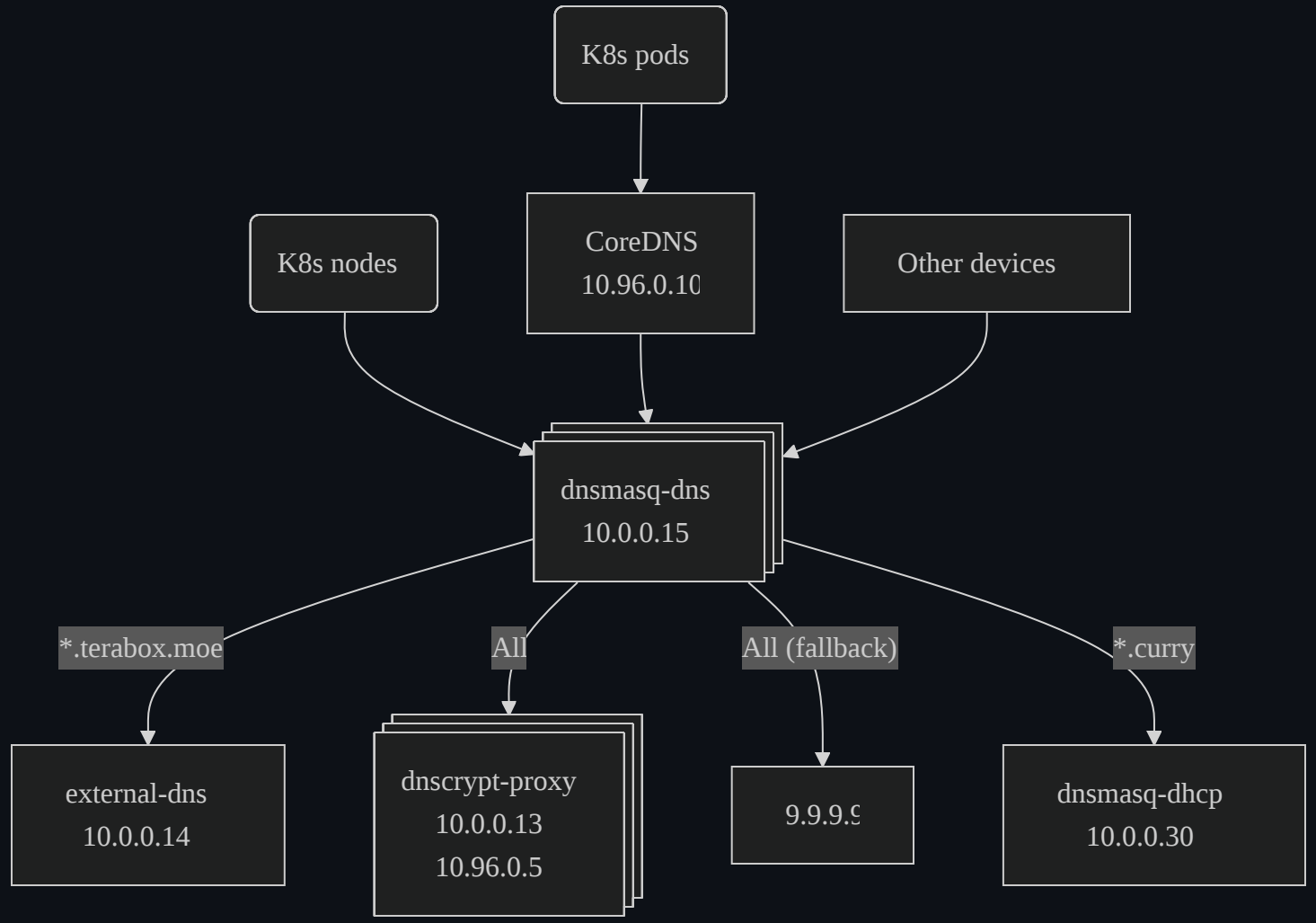
- Can dnsmasq download blocklists from the internet?

```
containers:
- name: adblock-downloader
  image: ghcr.io/nadiamoe/dnsmasq:v2.91-r0-bdd07fc
  command:
    - /bin/sh
  args:
    - -c
    - |-
      while [ true ]; do
        wget --no-verbose -O - "$ABD_URL" > /dnsmasq/adblock.hosts \
          && killall -HUP dnsmasq \
          && sleep 8h || ( echo "Failed to download blocklist, retrying..." && sleep $(( 5 + $RANDOM % 10 )) )
      done
```

🥴 `no-resolv` or die recursing

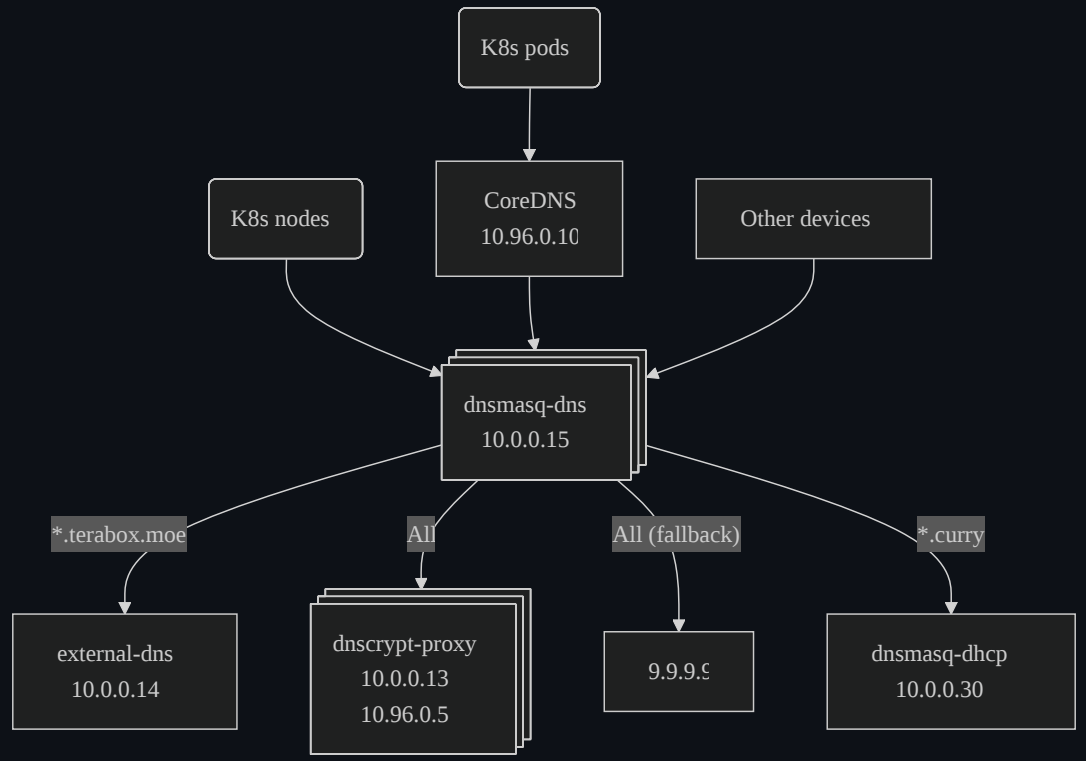


The full picture again



Dependency loop?

- Does this setup bootstrap cold?
  - No
- Does it matter?
  - Also no (probably)
- Container image cache does a great job
  - `imagePullPolicy: Always` is a crime
- Redundancy helps a lot
  - spegel-org/spegel
- CP toleration/affinity



Takeaways

- DNS is really composable, and you can make cool setups
- Kubernetes ecosystem offers HA and failover at a low cost
- Self-documenting, git-versioned infrastructure is priceless
- Brain is finite, reusing knowledge is a superpower

Thank you! ❤️