BOOSTING KUBERNETES PERFORMANCE – HOW TO TUNE DNS RESOLUTION

Fabian Brundke | Senior Platform Engineer

REPLY LIQUID

COREDNS

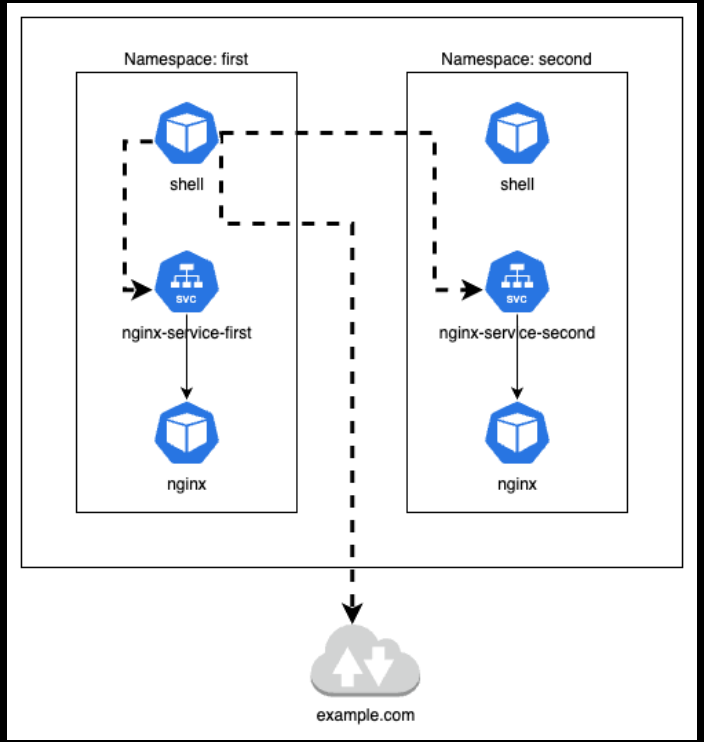
WHAT IS COREDNS

- DNS Server
- Responsible for Service Discovery: translates service names to cluster IP addresses
- Plugin Architecture

TIME FOR DEMO: COREDNS

DEMO OVERVIEW

CHECKING COREDNS LOGS OF VARIOUS REQUESTS



SUMMARY

DEFAULT DNS CONFIGURATION

DEFAULT DNS CONFIGURATION CAN RESULT IN **MANY EXTRA QUERIES** ESPECIALLY FOR EXTERNAL REQUESTS

Workaround: add “.” at end of domain to create FQDN.

There are better solutions.

```
[INFO] 10.244.2.7:60721 - 58400 "AAAA IN
example.com.first.svc.cluster.local. udp 53 false 512"
NXDOMAIN qr,aa,rd 146 0.000095036s
[INFO] 10.244.2.7:60721 - 58234 "A IN
example.com.first.svc.cluster.local. udp 53 false 512"
NXDOMAIN qr,aa,rd 146 0.000184987s
[INFO] 10.244.2.7:57404 - 40960 "AAAA IN
example.com.svc.cluster.local. udp 47 false 512"
NXDOMAIN qr,aa,rd 140 0.000069828s
[INFO] 10.244.2.7:57404 - 40752 "A IN
example.com.svc.cluster.local. udp 47 false 512"
NXDOMAIN qr,aa,rd 140 0.000033081s
[INFO] 10.244.2.7:50037 - 35220 "AAAA IN
example.com.cluster.local. udp 43 false 512" NXDOMAIN
qr,aa,rd 136 0.000040664s
[INFO] 10.244.2.7:50037 - 34839 "A IN
example.com.cluster.local. udp 43 false 512" NXDOMAIN
qr,aa,rd 136 0.000022874s
[INFO] 10.244.2.7:45201 - 39675 "AAAA IN example.com.
udp 29 false 512" NOERROR qr,rd,ra 68 0.058875555s
[INFO] 10.244.2.7:45201 - 39335 "A IN example.com. udp
29 false 512" NOERROR qr,rd,ra 56 0.068724547s
```

TUNING DNS RESOLUTION

/ETC/RESOLV.CONF

EXISTS FOR EVERY POD (AND NODE)

```
~ # cat /etc/resolv.conf
nameserver 10.43.0.10
search first.svc.cluster.local
svc.cluster.local cluster.local
options ndots:5
```

```
[INFO] 10.244.2.7:60721 - 58400 "AAAA IN
example.com.first.svc.cluster.local. udp 53 false 512"
NXDOMAIN qr,aa,rd 146 0.000095036s

[INFO] 10.244.2.7:60721 - 58234 "A IN
example.com.first.svc.cluster.local. udp 53 false 512"
NXDOMAIN qr,aa,rd 146 0.000184987s

[INFO] 10.244.2.7:57404 - 40960 "AAAA IN
example.com.svc.cluster.local. udp 47 false 512"
NXDOMAIN qr,aa,rd 140 0.000069828s

[INFO] 10.244.2.7:57404 - 40752 "A IN
example.com.svc.cluster.local. udp 47 false 512"
NXDOMAIN qr,aa,rd 140 0.000033081s

[INFO] 10.244.2.7:50037 - 35220 "AAAA IN
example.com.cluster.local. udp 43 false 512" NXDOMAIN
qr,aa,rd 136 0.000040664s

[INFO] 10.244.2.7:50037 - 34839 "A IN
example.com.cluster.local. udp 43 false 512" NXDOMAIN
qr,aa,rd 136 0.000022874s

[INFO] 10.244.2.7:45201 - 39675 "AAAA IN example.com.
udp 29 false 512" NOERROR qr,rd,ra 68 0.058875555s

[INFO] 10.244.2.7:45201 - 39335 "A IN example.com. udp
29 false 512" NOERROR qr,rd,ra 56 0.068724547s
```

/ETC/RESOLV.CONF

EXISTS FOR EVERY POD (AND NODE)

```
~ # cat /etc/resolv.conf
nameserver 10.43.0.10
search first.svc.cluster.local
svc.cluster.local cluster.local
options ndots:5
```

Search domains get appended if number of dots in domain is less than ndots.

Example:

example.com has 1 dot → 1 < 5 → search domains get appended

```
[INFO] 10.244.2.7:60721 - 58400 "AAAA IN
example.com.first.svc.cluster.local. udp 53 false 512"
NXDOMAIN qr,aa,rd 146 0.000095036s
[INFO] 10.244.2.7:60721 - 58234 "A IN
example.com.first.svc.cluster.local. udp 53 false 512"
NXDOMAIN qr,aa,rd 146 0.000184987s
[INFO] 10.244.2.7:57404 - 40960 "AAAA IN
example.com.svc.cluster.local. udp 47 false 512"
NXDOMAIN qr,aa,rd 140 0.000069828s
[INFO] 10.244.2.7:57404 - 40752 "A IN
example.com.svc.cluster.local. udp 47 false 512"
NXDOMAIN qr,aa,rd 140 0.000033081s
[INFO] 10.244.2.7:50037 - 35220 "AAAA IN
example.com.cluster.local. udp 43 false 512" NXDOMAIN
qr,aa,rd 136 0.000040664s
[INFO] 10.244.2.7:50037 - 34839 "A IN
example.com.cluster.local. udp 43 false 512" NXDOMAIN
qr,aa,rd 136 0.000022874s
[INFO] 10.244.2.7:45201 - 39675 "AAAA IN example.com.
udp 29 false 512" NOERROR qr,rd,ra 68 0.058875555s
[INFO] 10.244.2.7:45201 - 39335 "A IN example.com. udp
29 false 512" NOERROR qr,rd,ra 56 0.068724547s
```

POD.SPEC.DNSCONFIG

ALLOWS TO MODIFY /ETC/RESOLV.CONF

```yaml
apiVersion: v1
kind: Pod
spec:
  dnsConfig:
    nameservers:
      - 192.0.2.1
    searches:
      - ns1.svc.cluster-domain.example
      - my.dns.search.suffix
    options:
      - name: ndots
        value: "2"
```

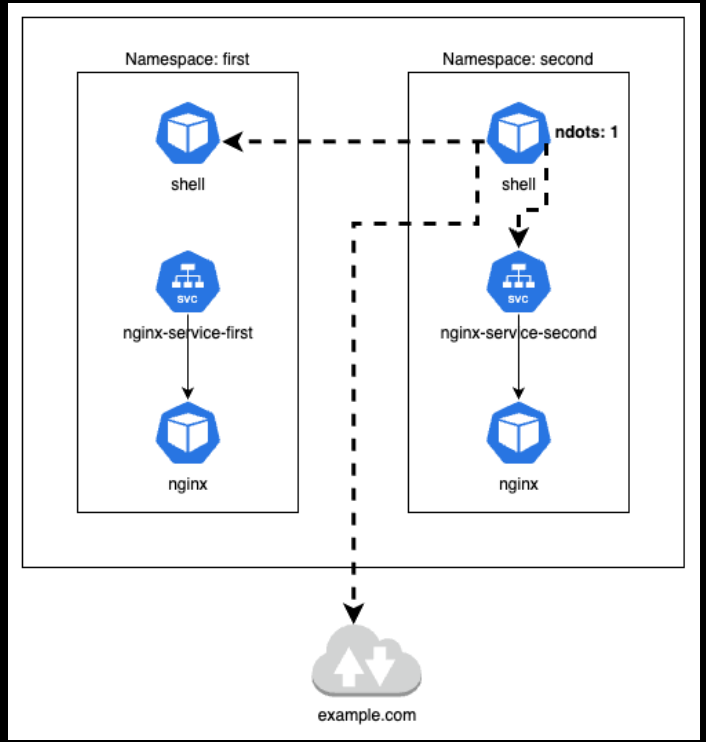
```
~ # cat /etc/resolv.conf
nameserver 10.43.0.10
search first.svc.cluster.local
svc.cluster.local cluster.local
options ndots:5
```

https://kubernetes.io/docs/concepts/services-networking/dns-pod-service/#pod-dns-config

TIME FOR DEMO: NDOTS

DEMO OVERVIEW

CHECKING COREDNS LOGS OF VARIOUS REQUESTS



SUMMARY
NDOTS

REDUCING NDOTS SETTING CAN
REDUCE NUMBER OF DNS QUERIES

“ndots: 1” is aggressive:

- can help with external requests
- is problematic for cross namespace requests

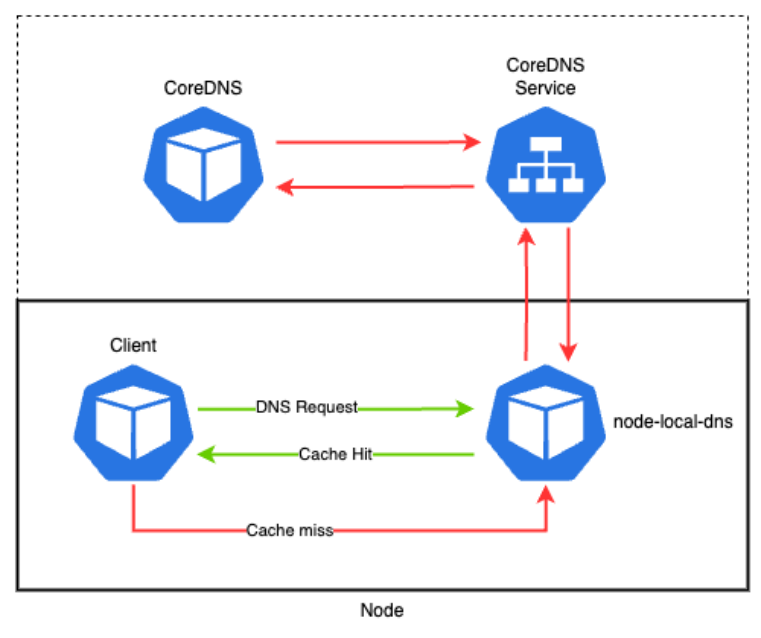
No “one size fits all” solution to ndots setting but “2” or “3” is good tradeoff.

```
[INFO] 10.42.0.6:44107 - 51102 "AAAA IN nginx-
service.same-namespace.svc.cluster.local. udp 63 false
512" NOERROR qr,aa,rd 156 0.000441955s


[INFO] 10.42.0.6:56669 - 61518 "A IN ninx-
service.other-namespace. udp 42 false 512" NXDOMAIN
qr,rd,ra 42 0.215348282s


[INFO] 10.42.0.6:37398 - 18497 "A IN example.com. udp
29 false 512" NOERROR qr,rd,ra 56 0.002901261s
```

NODELOCAL DNSCACHE

HOW DOES IT WORK

- Deployed in addition to CoreDNS
- CoreDNS optimized for caching and running on each node (DaemonSet)
- Creates network interface and iptables rules
- Small resource footprint
- Reduces load on DNS servers
- Reduces traffic node hopping

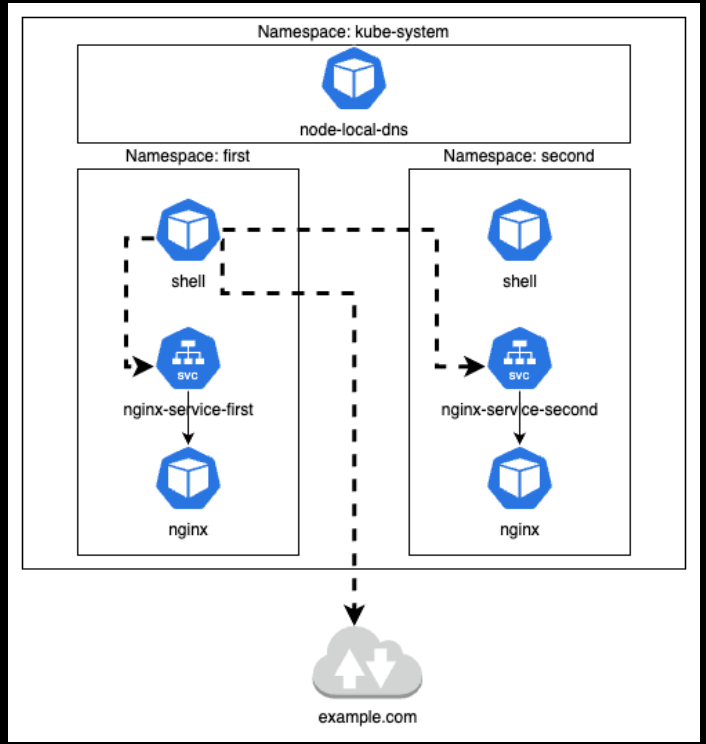


Detailed diagram: https://kubernetes.io/docs/tasks/administer-cluster/nodelocaldns/#architecture-diagram

TIME FOR DEMO: NODELOCAL DNSCACHE

DEMO OVERVIEW

CHECKING COREDNS LOGS OF VARIOUS REQUESTS



SUMMARY
NODELOCAL DNSCACHE

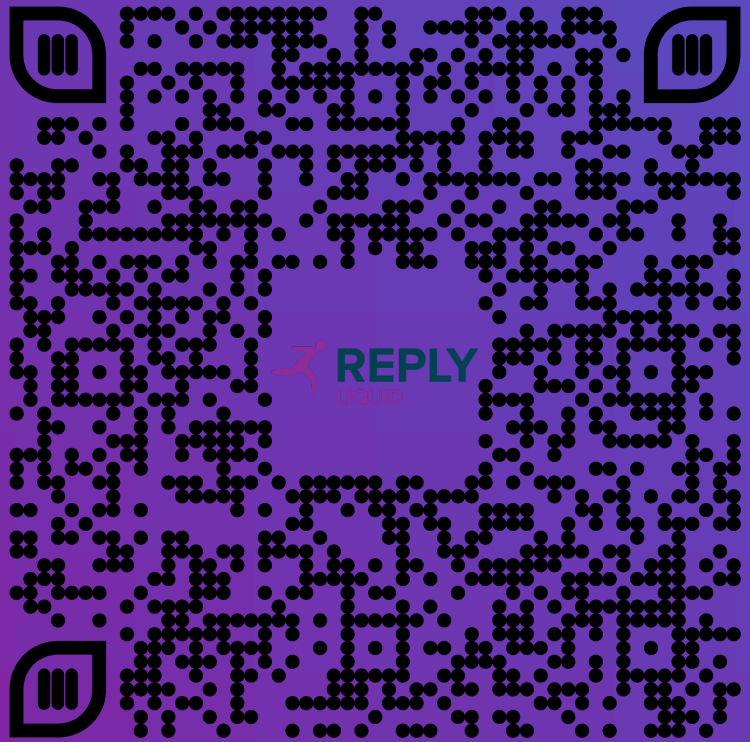
- Significant reduction in queries
- Works for cluster internal and external requests
- No need to adjust any application
- Cache based on record TTL (default 5s – can be set via CoreDNS)

SUMMARY

MULITPLE TUNING OPTIONS

- Default configuration should be adjusted for high traffic scenarios
- Use `pod.spec.dnsConfig` to adjust /etc/resolv.conf per application
  - `search`: finetune search domains
  - `ndots`: value of 2 or 3 reduces DNS queries
- Deploy NodeLocal DNSCache to improve latency, reduce load on DNS servers and potentially reduce cloud bill

Presentation & Material



Presentation & Material