XFR does not scale

06.09.2023 · Klaus Darilion · Head of Operations

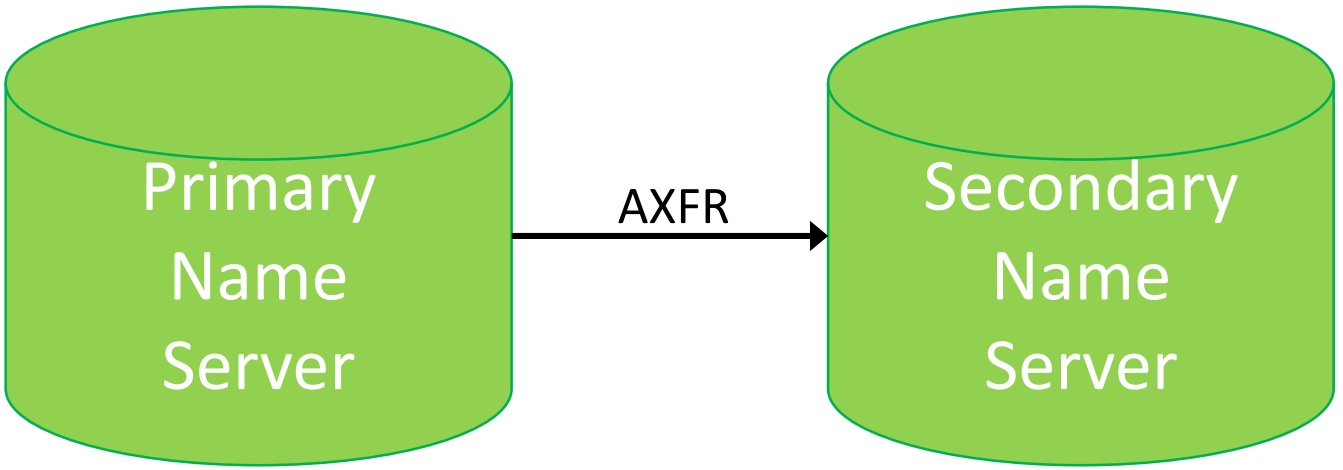
Klaus Darilion

- nic.at Head of Operations
- Linux Sysadmin
- RcodeZero DNS from the beginning: 13 years

Motivation of this Talk

- See my talk from Yesterday:
  - Migrate from PowerDNS to Knot with help of Catalog Zones

DNS Zone Replication: XFR (AXFR or IXFR)



XFR: When?

- „Refresh“ operation detects stale serial
  - SOA query to get zone‘s SERIAL on Primary
  - If Primary‘s SERIAL > Secondary‘s SERIAL -> XFR
  - (or immediate IXFR instead of SOA)

- Refresh operation is triggered by:
  - NOTIFY sent by Primary to Secondary, or
  - REFRESH interval elapsed

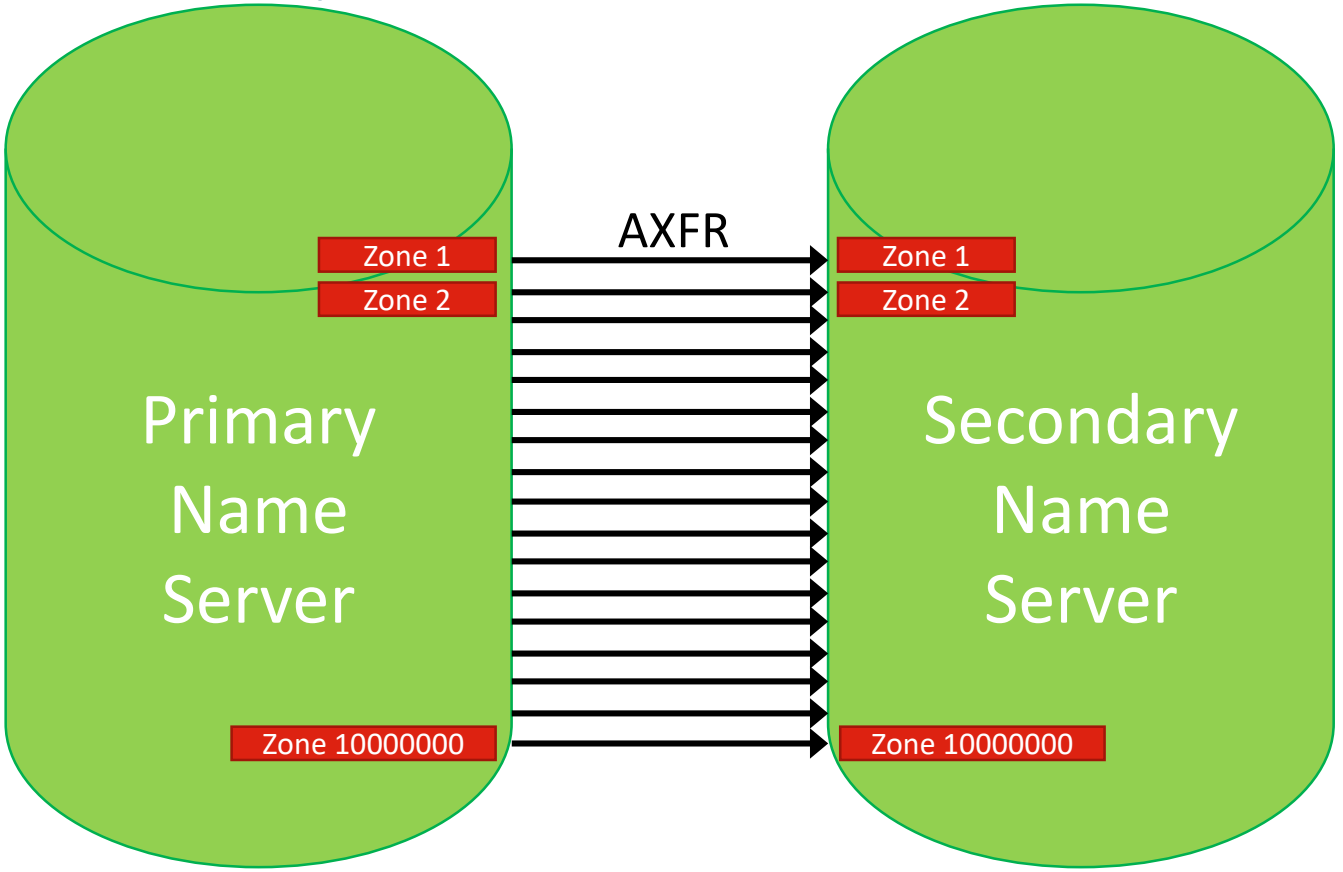
- NOTIFY failed? Primaries do not care
  - Secondary should detect higher serial during next Refresh-Check

- XFR failed? Secondaries do not care
  - Next try during next Refresh-Check

Is Secondary in Sync with Primary?

- Who knows it? Nobody, neither Primary nor Secondary
- Out of protocol monitoring required
  - SOA checks against Primary and Secondaries, or
  - get SERIAL by other means (nsd-control, knotc)
- Monitoring Interval?
  - Once per minute? Once per second?
  - Depends on SLAs
  - 1 Zone, 1 Minute -> 2 queries/minute

DNS Zone Replication with 1 Million Zones



1 Million Replications!!!

1 Million Zones

- 1 Million „Replications“
- 1 Million independent Refresh checks
- 1 Million Replications to monitor (zone serial)

Monitoring by DNS

- 1 Zone, 1 Minute -> 2 queries/minute
- 1 mio Zones, 1 Minute, 50 Secondaries -> 833.000 queries/second
- Zones are changing
  - Temporarily Different Serials are not an error
  - Some zones change every few seconds (never OK)
    - Monitoring by DNS must be „statefull“
- → Zone propagation delay monitoring is CPU consuming and complex
- → That does not scale
  - XFR itself is not the problem

Alternatives to NOTIFY/Refresh Checks

- SERIAL property in catalog zones
  - Klaus Darilion
  - Was in „catalog zones“ draft, but removed to speed up RFC finalizing

- ZONEMD in catalog zones
  - Leo Vandewoestijne
  - https://archive.fosdem.org/2020/schedule/event/dns_catz/attachments/slides/3704/export/events/attachments/dns_catz/slides/3704/catalog_zones.pdf

- Dedicated Meta-Zone for Serials and other stuff
  - Petr Špacek
  - https://chat.dns-oarc.net/community/pl/ofox8fxntiyt9enmbaqfpzaqfw

Alternatives to DNS-based Replication

- PowerDNS with SQL-Backend
  - Serial, single Stream Replication by Database (Mysql Master-Slave, Postgresql Logical Replication …)
  - Easy to monitor: just monitor DB replication lag
- PowerDNS with LMDB-Backend and „Lighting Stream“
  - LMDB data files replicated from/to S3 storage
  - Serial, single Stream
  - Monitoring: ??? (Tracking last sync)
- Cloudflare, Google DNS, Route 53?
  - Probably Proprietary Solutions, not DNS based

Conclusion

- Monitoring of >1 Mio Replications is not easy
- All zones should use a single replication
  - Or a single „status“ to easily fetch the syncronicity status

- Can there be a standardized solution?
- Or how do you solve this issue?
  - No monitoring? ☺
- Is there interest in starting a discussion?

Klaus Darilion · Head of Operations

klaus.darilion@nic.at