

The Architecture of the Next Generation DNS Server: BIND 10

Shane Kerr
shane@isc.org



DNS? BIND?

- DNS: convert *names* into *numbers*
 - `www.isc.org` → `149.20.64.42`
 - `www.isc.org` → `2001:4f8:0:2::d`
- BIND: DNS server software
 - Runs on about 80% of DNS servers



BIND 9. BIND 10!

- BIND 9: current version of BIND
 - 9.0.0 released in 2000-09-16
 - Now at 9.7.2-P2
- BIND 10: next version of BIND
 - 5 year project to re-implement BIND
 - Currently in year 2
 - Sponsored development



Today's Topic: Architecture

'To date there is still no agreement on the precise definition of the term "software architecture".'

*– Wikipedia article on
"Software architecture"*



Today's Topic: Architecture... Defined!

'I know it when I see it.'

*- Potter Stewart,
United States Supreme Court*

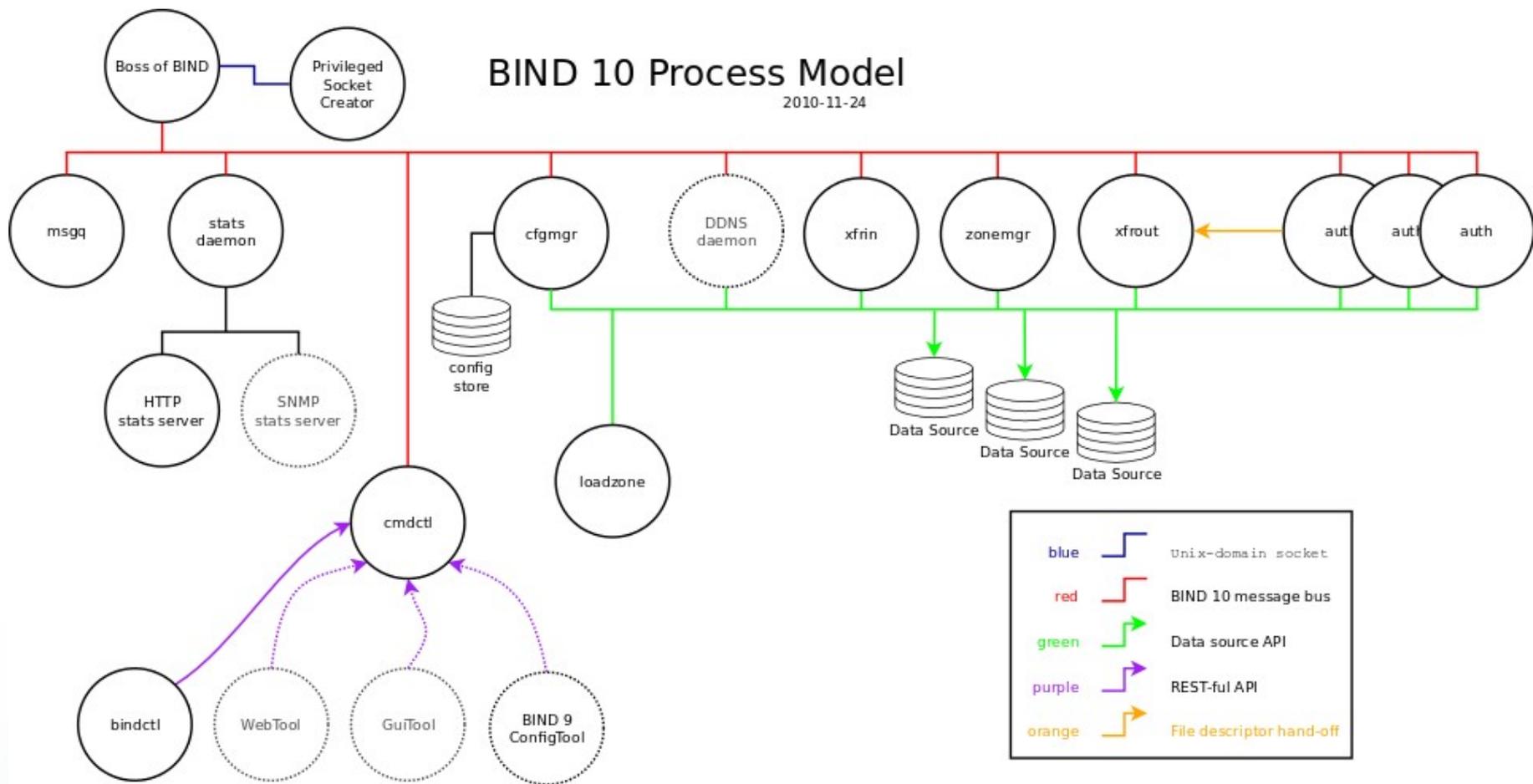


Goals that Affect Architecture I

- Customization “out-of-the-box”
 - authoritative-only, recursive-only
 - slave-only, master-only
 - enable/disable dynamic DNS
 - support favorite SQL backend
- Customization via code changes
 - non-ISC modules, or modifications
 - bespoke or in-house development

Goals that Affect Architecture II

- Scalability
 - BIND 8: single core (“CPU” then)
 - BIND 9: multiple cores (4-6 or so)
 - BIND 10: 10's or 100's of cores, multiple machines (clustered)
- Robustness
 - Reduce serious software bugs
 - Minimize impact of bugs
 - Reduce “fate sharing”



<https://bind10.isc.org/wiki/DesignDiagrams>

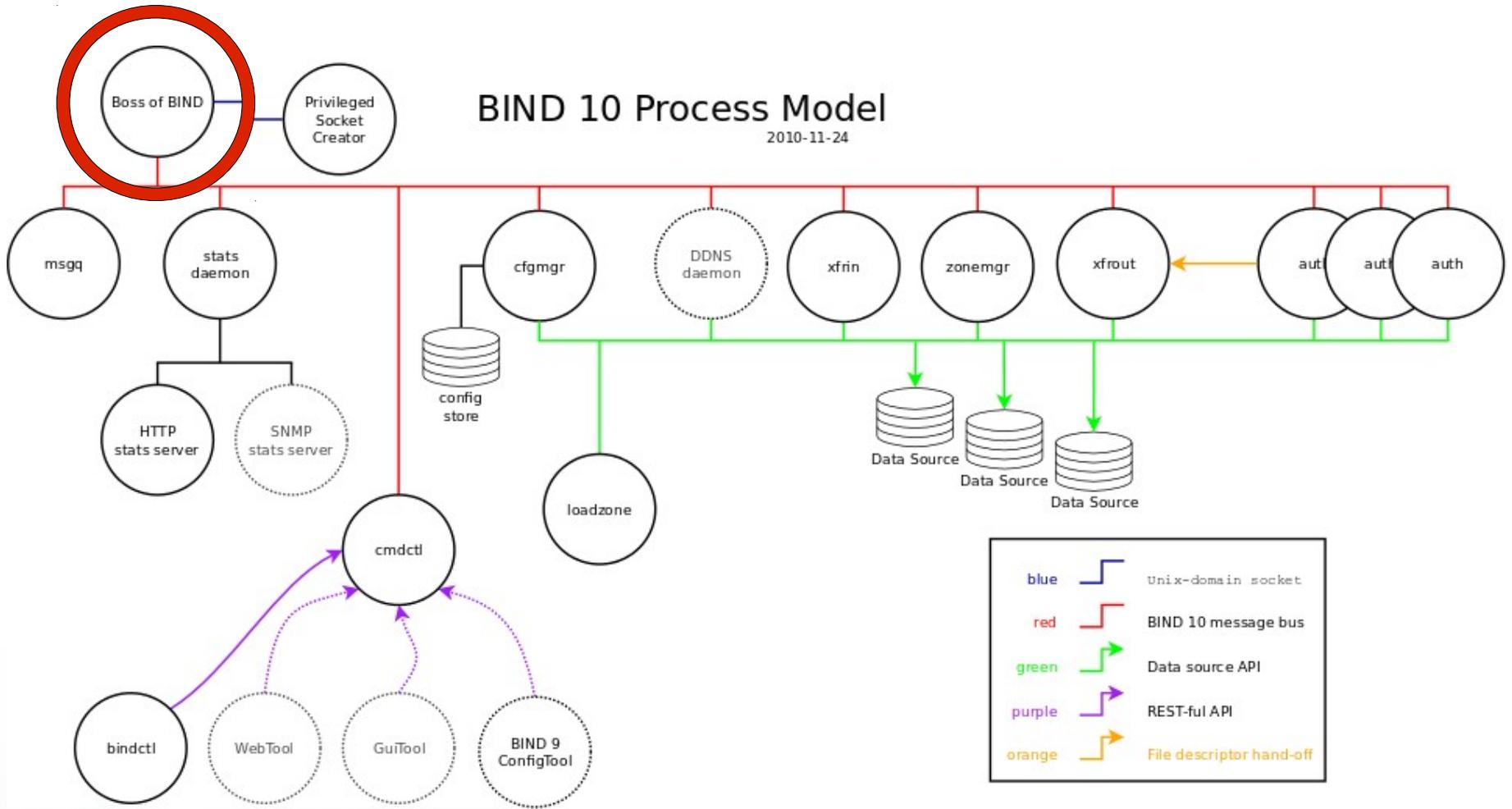
BIND 10

- Core
 - BoB
 - msgq
 - cfgmgr



BIND 10 Process Model

2010-11-24



Master^WBoss of BIND

- Handles startup, shutdown
- Restarts processes that die
- Written in Python



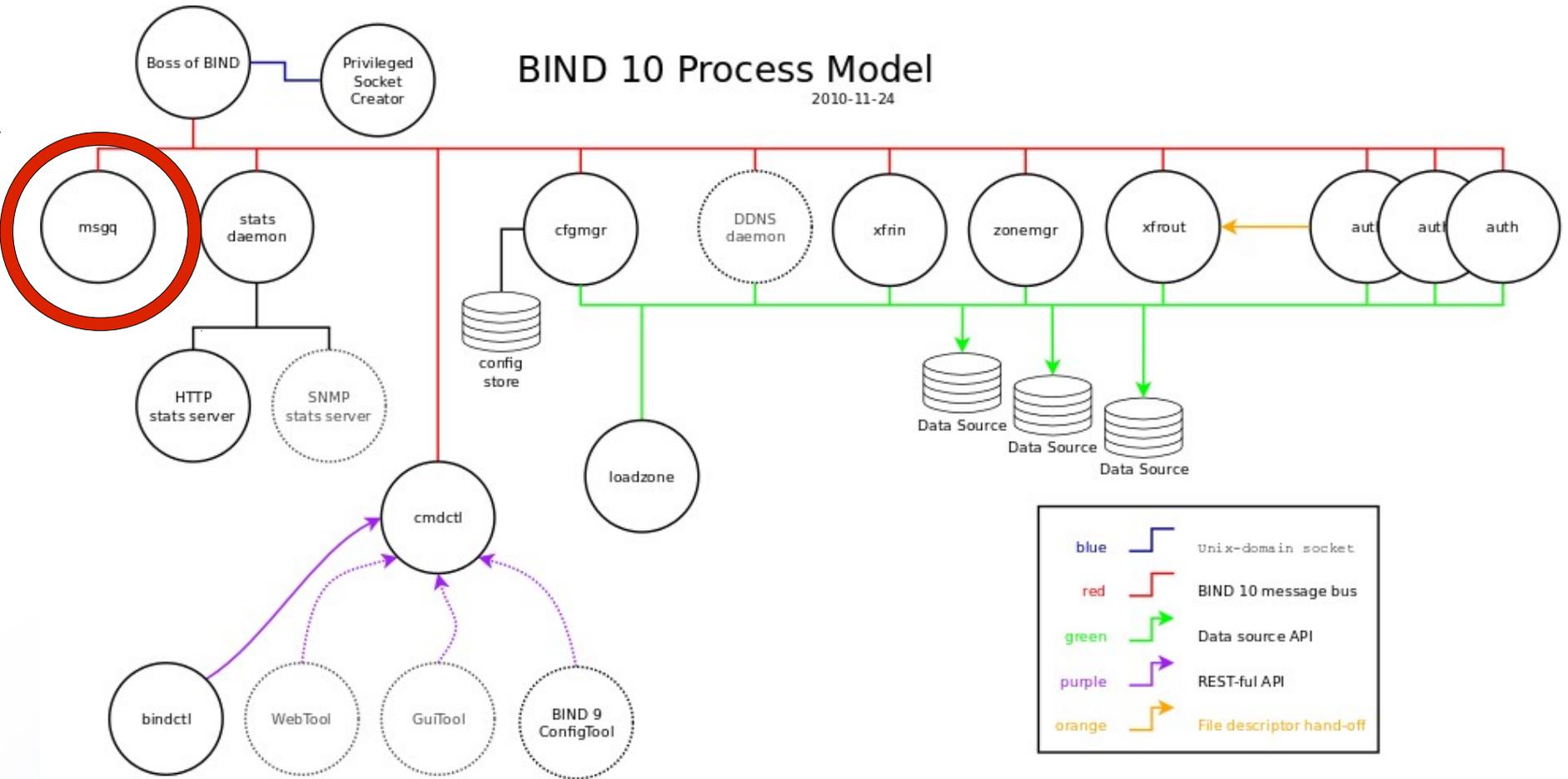
An Aside:

BIND 10 Languages

- C++ for performance critical parts
 - Modern compiled language
 - Widely used
- Python for... everything else
 - Modern scripting language
 - Widely used
 - Chose Python 3.x
 - Best. Language. Ever.

BIND 10 Process Model

2010-11-24



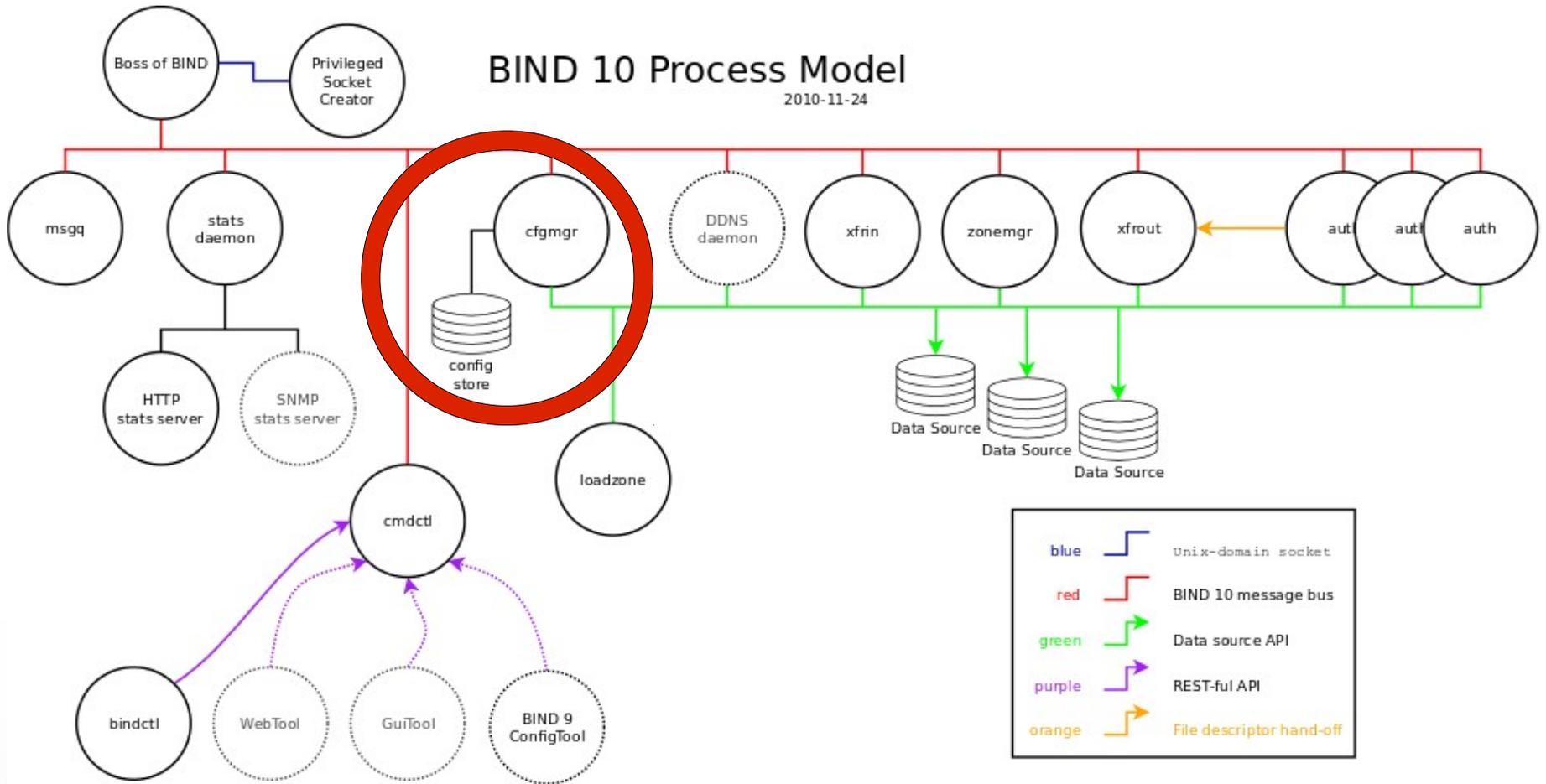
msgq

- Inter-process message bus
- Needed for extensibility
- Like d-bus, but also inter-machine
- Message format: JSON
- Unix domain socket connections
- No security



BIND 10 Process Model

2010-11-24



cfgmgr

- Configuration manager
- Never need to restart BIND 10
- Flexible, extensible configuration
- Not traditional Unix configuration
 - Changes are immediate & persistent
 - More like a router or an application



BIND 10

- Core

- BoB
- msgq
- cfgmgr

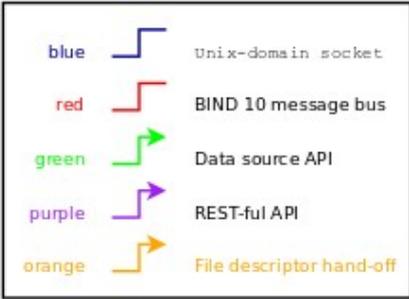
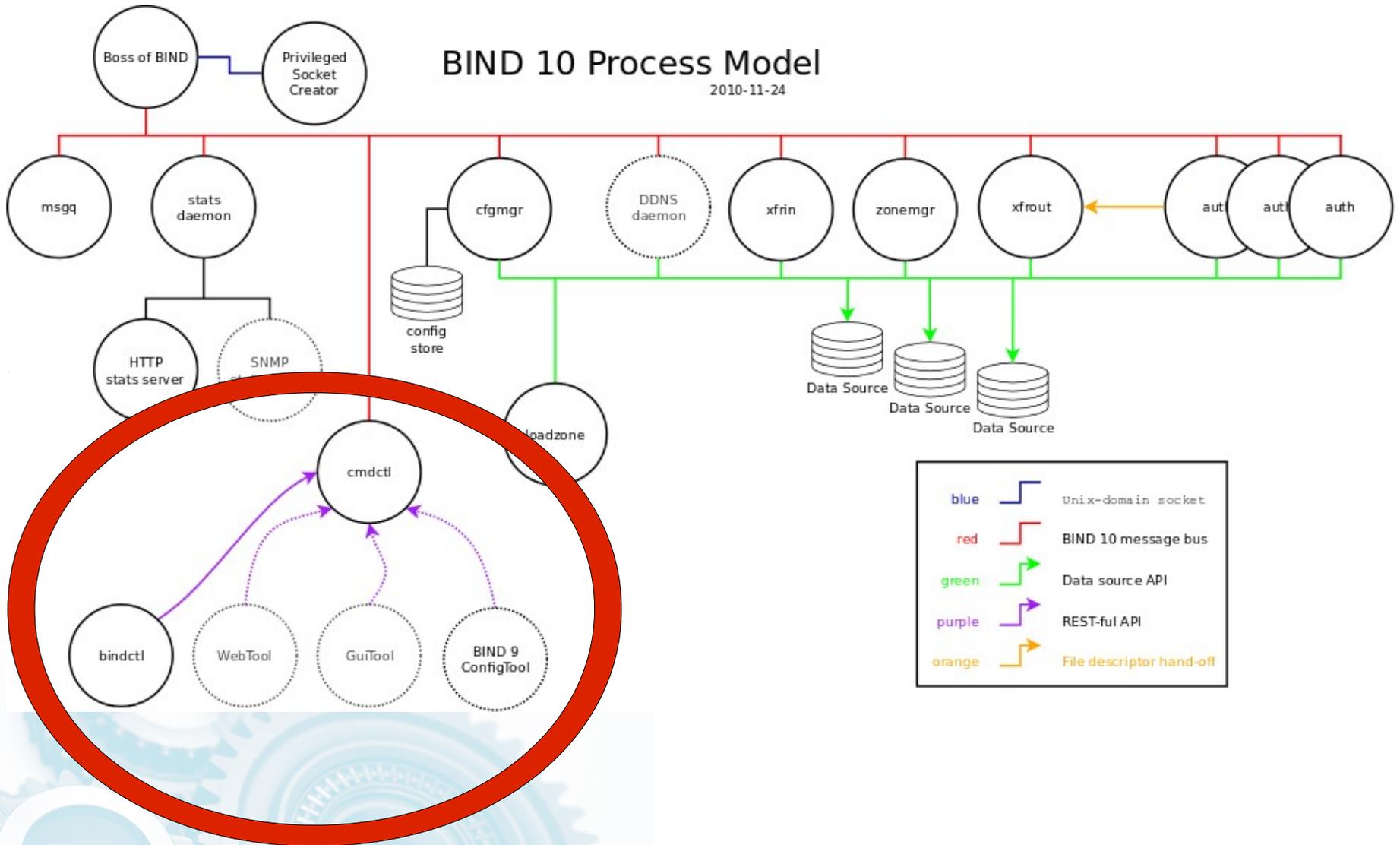
- Options

- cmdctl
- auth
- xfrin
- xfrout
- zonemgr
- stats



BIND 10 Process Model

2010-11-24



cmdctl

- Used to control the server
- Authenticates users
- Interacts with cfgmgr
 - Gets per-module options
- Interacts with modules
 - Commands like “refresh zone”
- Current client: bindctl (CLI)
- Future clients: web, GUI, new CLI

Supporting Library: Data Sources

- Idea stolen from PowerDNS
- Back-end for authoritative DNS
- Currently SQLite
- Plans:
 - MySQL, PostgreSQL
 - Berkeley DB
 - In-memory data structure(s)
- Used by auth, xfrin, xfrout, loadzone

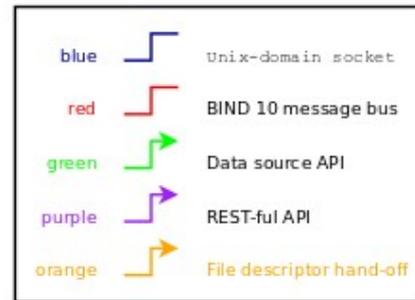
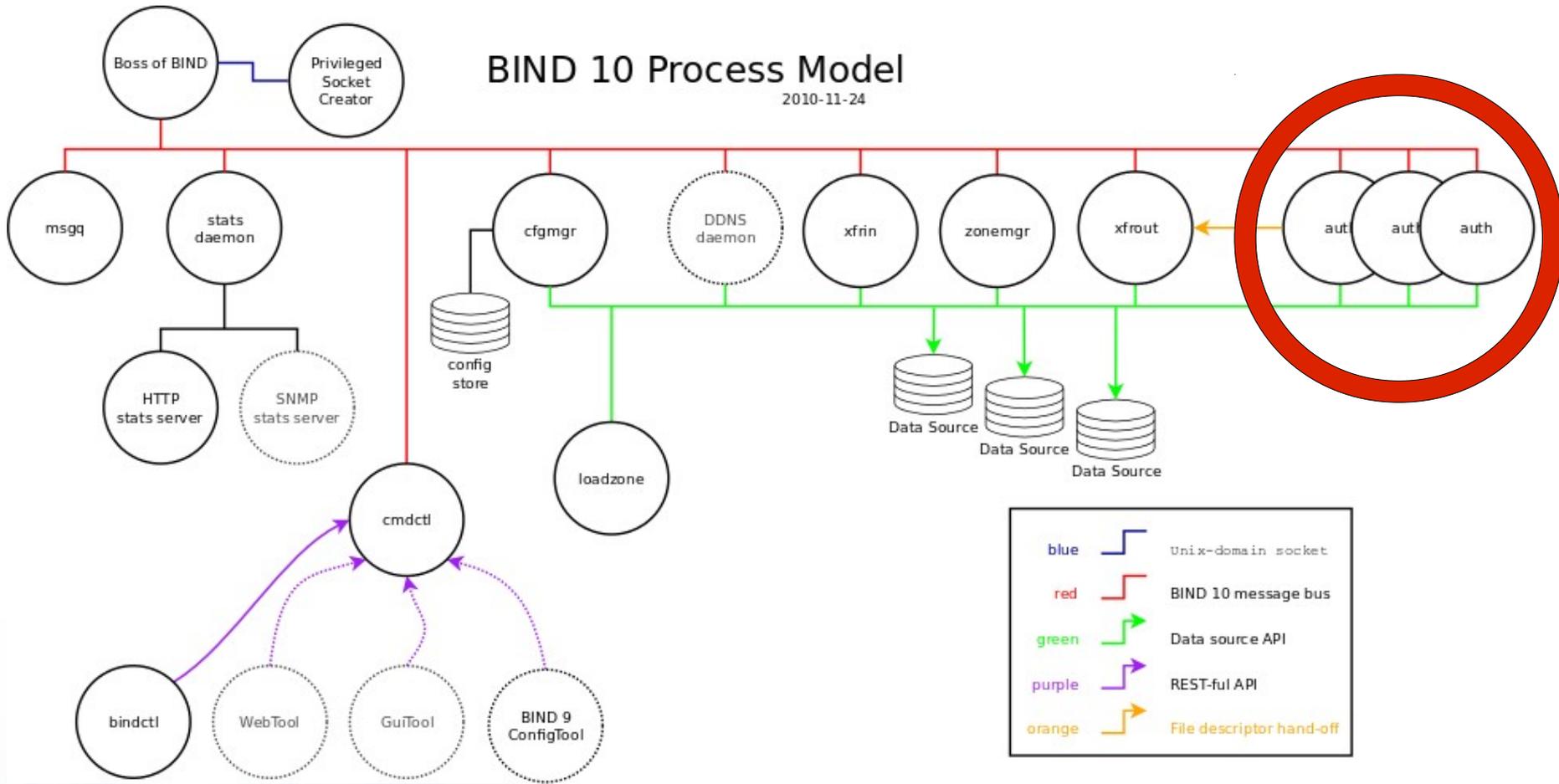
Supporting Library: DNS Message

- Low-level DNS messages (packets)
- New C++ implementation
- Python wrapper



BIND 10 Process Model

2010-11-24

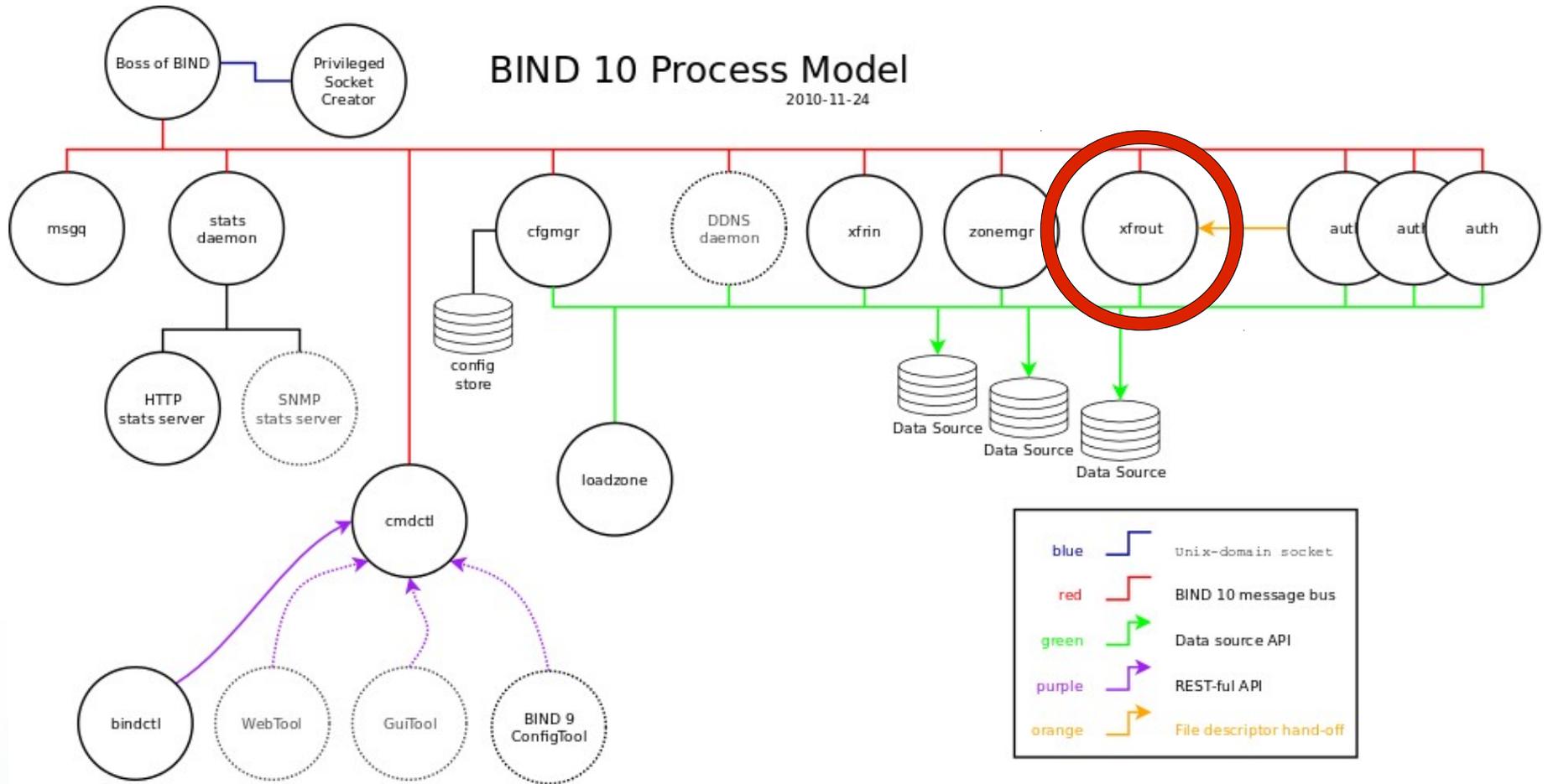


auth

- Authoritative DNS server
- DNS library
 - + data sources
 - + I/O
 - + bit of logic
- Scales via multiple processes
 - Idea stolen from NSD

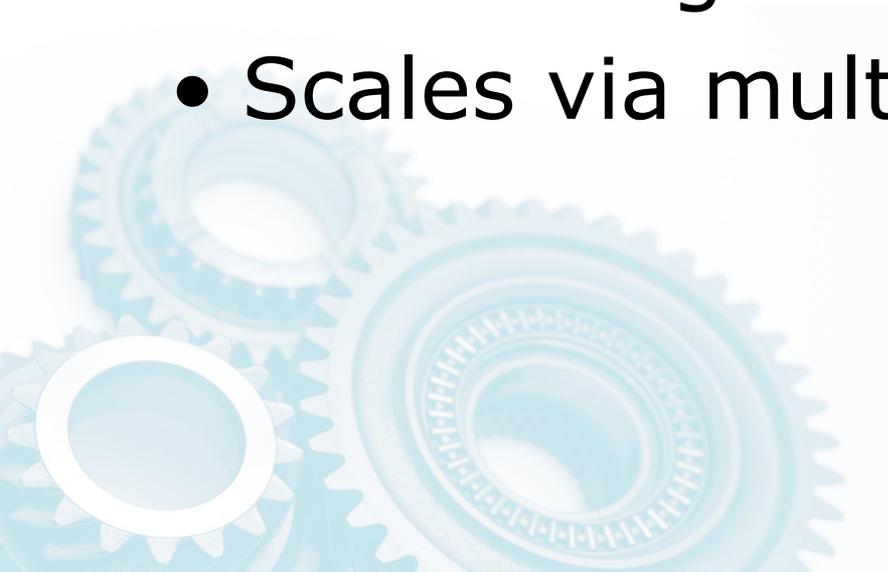
BIND 10 Process Model

2010-11-24



xfrout

- AXFR out, to act as a master server
- DNS library
 - + data sources
 - + I/O
 - + bit of logic
- Scales via multiple threads



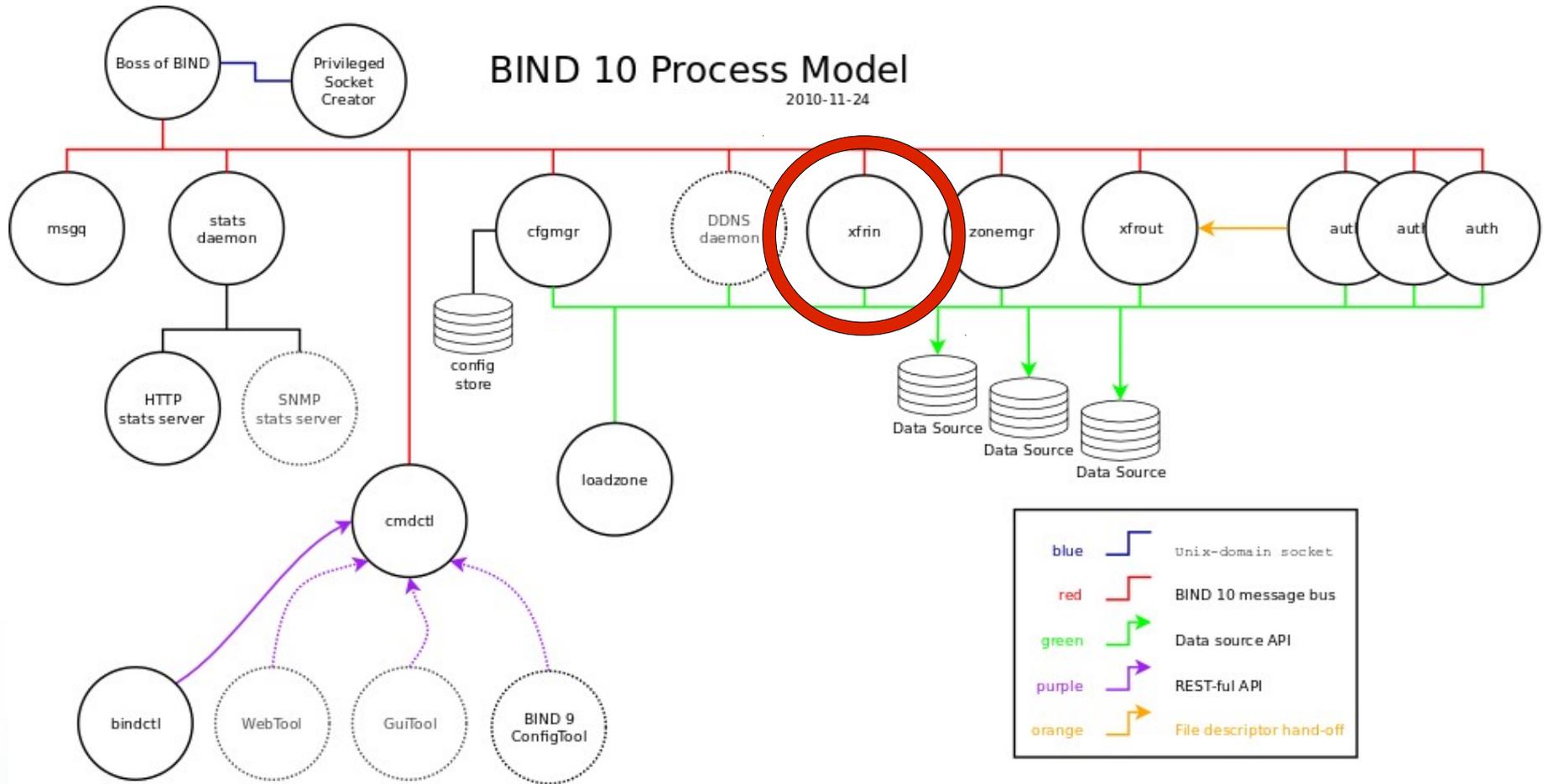
A Final Aside:

Passing Around Open Files

- AXFR messages come to auth
- UDP packets can be forwarded
- TCP connections must go to xfrout
- Send file descriptor via `sendmsg()` :
 - `SOL_SOCKET, SCM_RIGHTS`
- Works on Linux, Solaris, BSD

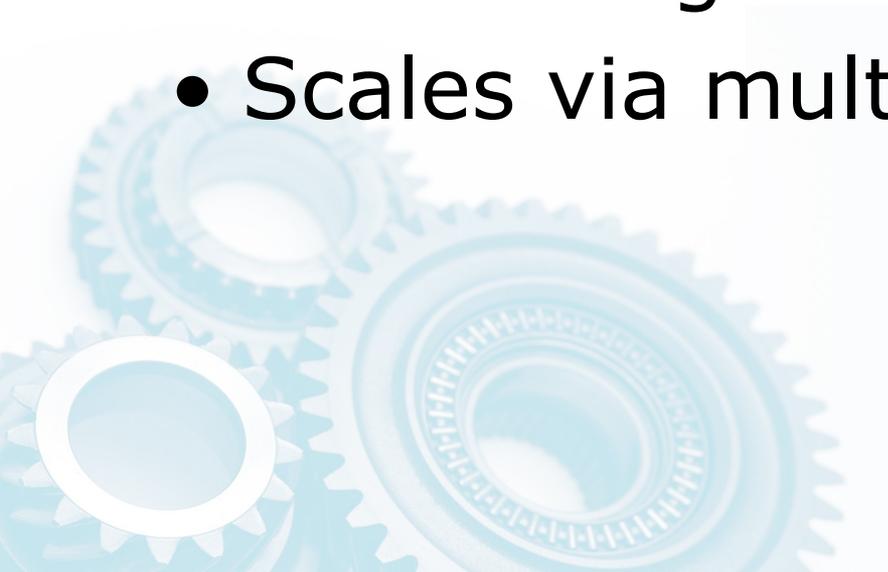
BIND 10 Process Model

2010-11-24



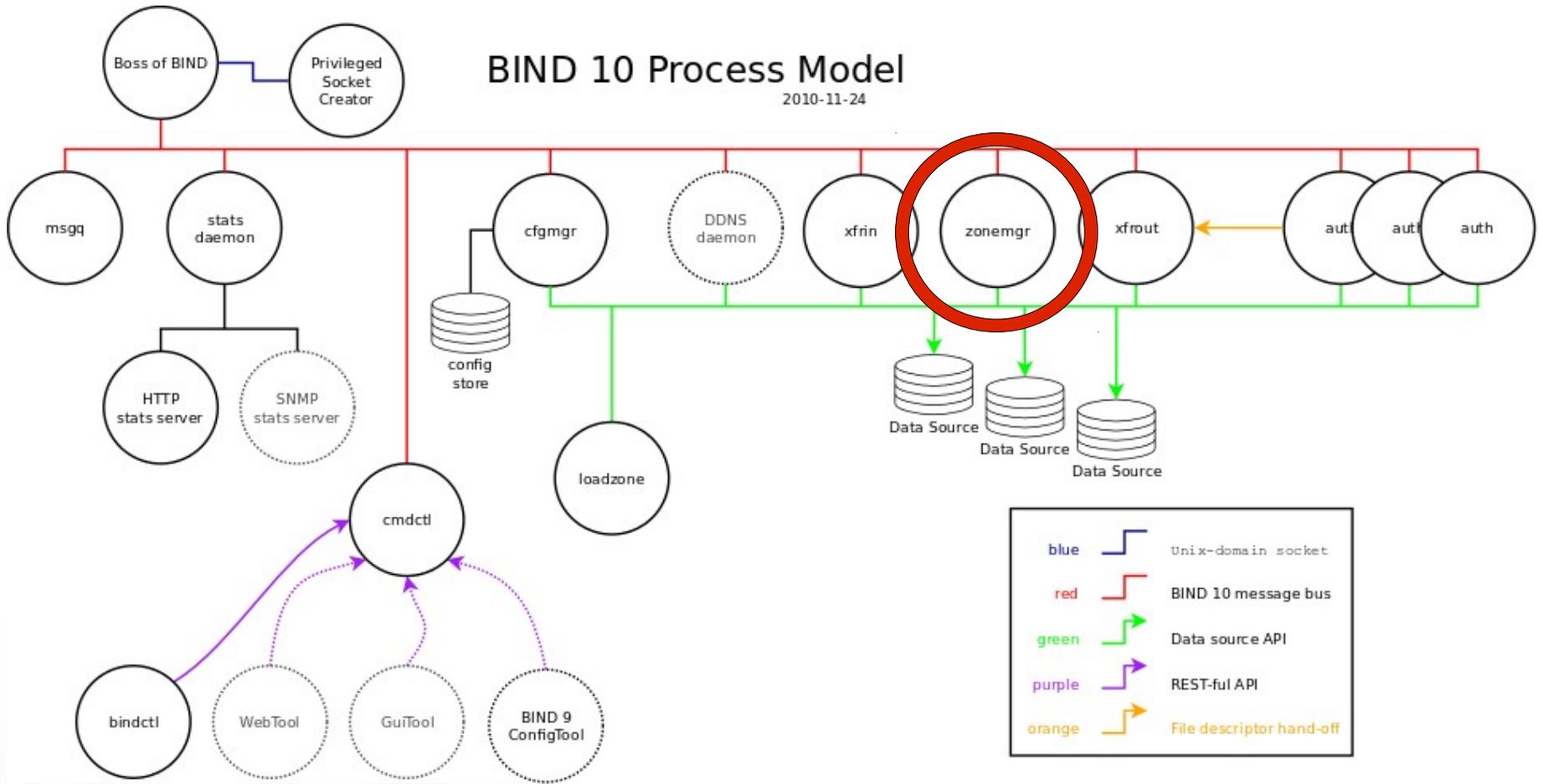
xfrin

- AXFR in, to act as a slave server
- DNS library (Python)
 - + data sources
 - + I/O
 - + bit of logic
- Scales via multiple threads



BIND 10 Process Model

2010-11-24



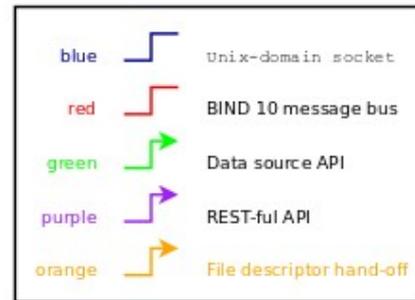
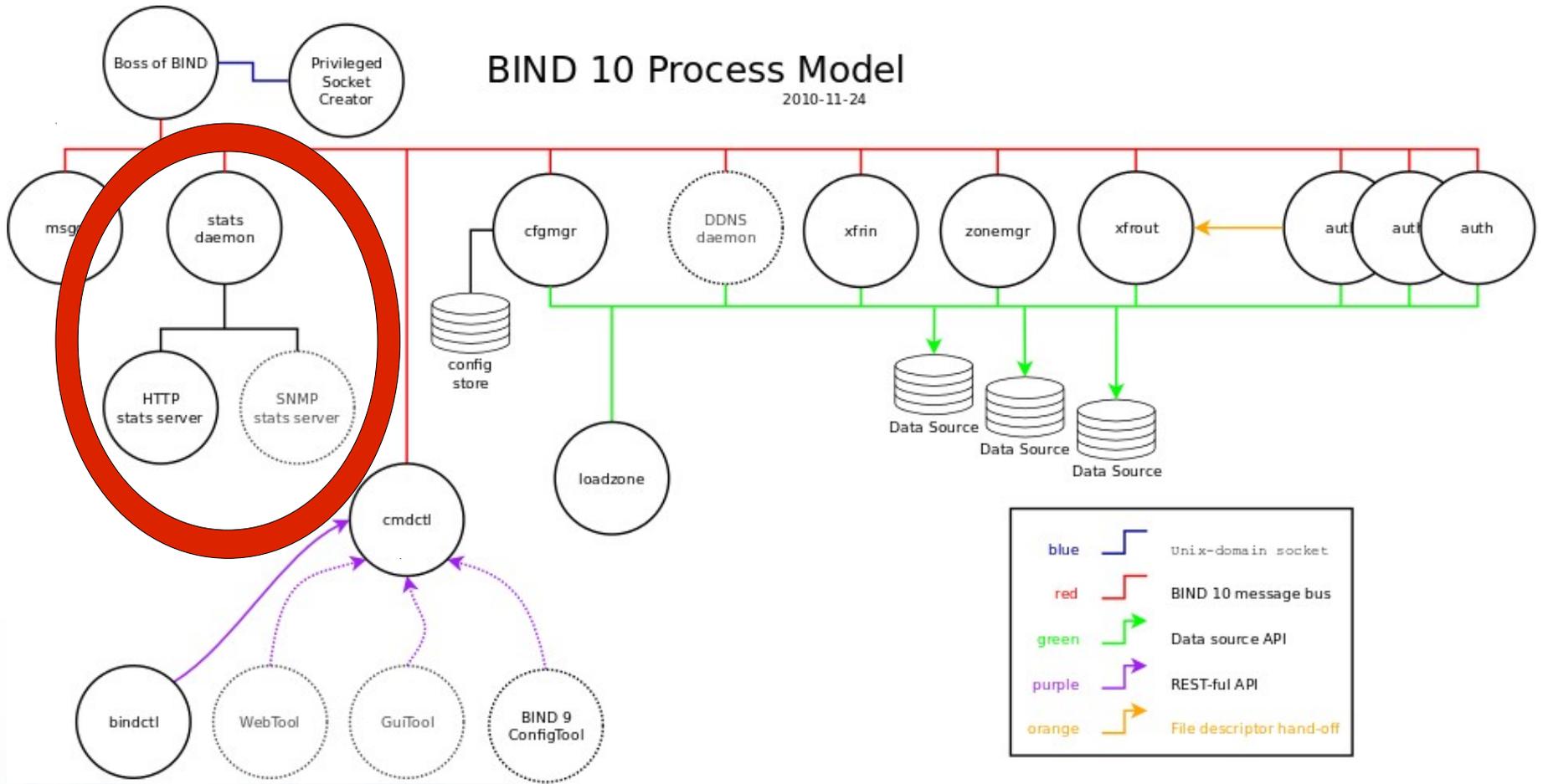
zonemgr

- Zone manager, times slave refresh
- Data sources
- That's it!



BIND 10 Process Model

2010-11-24



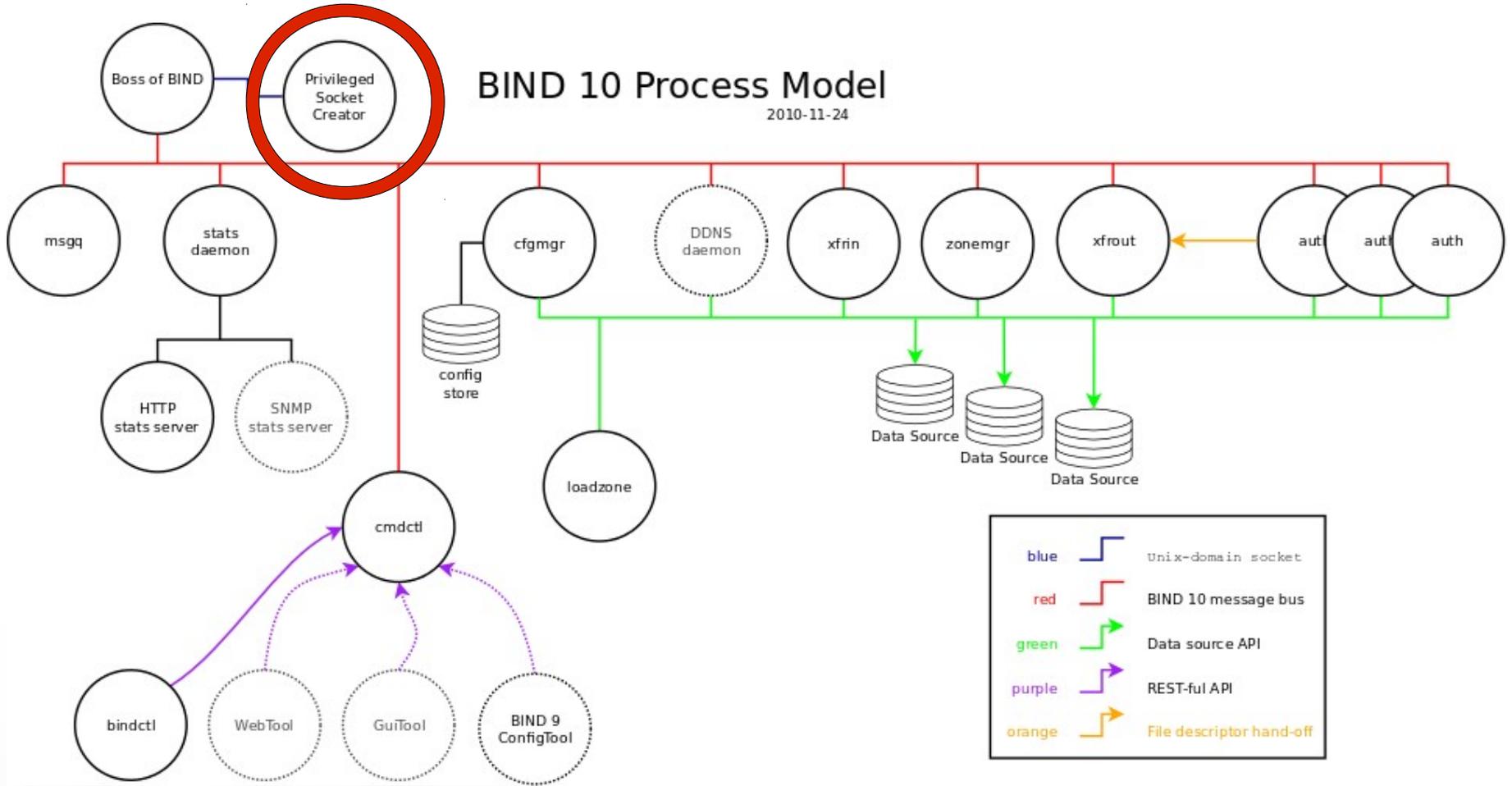
Statistics

- Modules report stats
- Collected by stats daemon
- Present in various ways:
 - Via bindctl
 - XML over HTTP (BIND 9 style)
 - SNMP



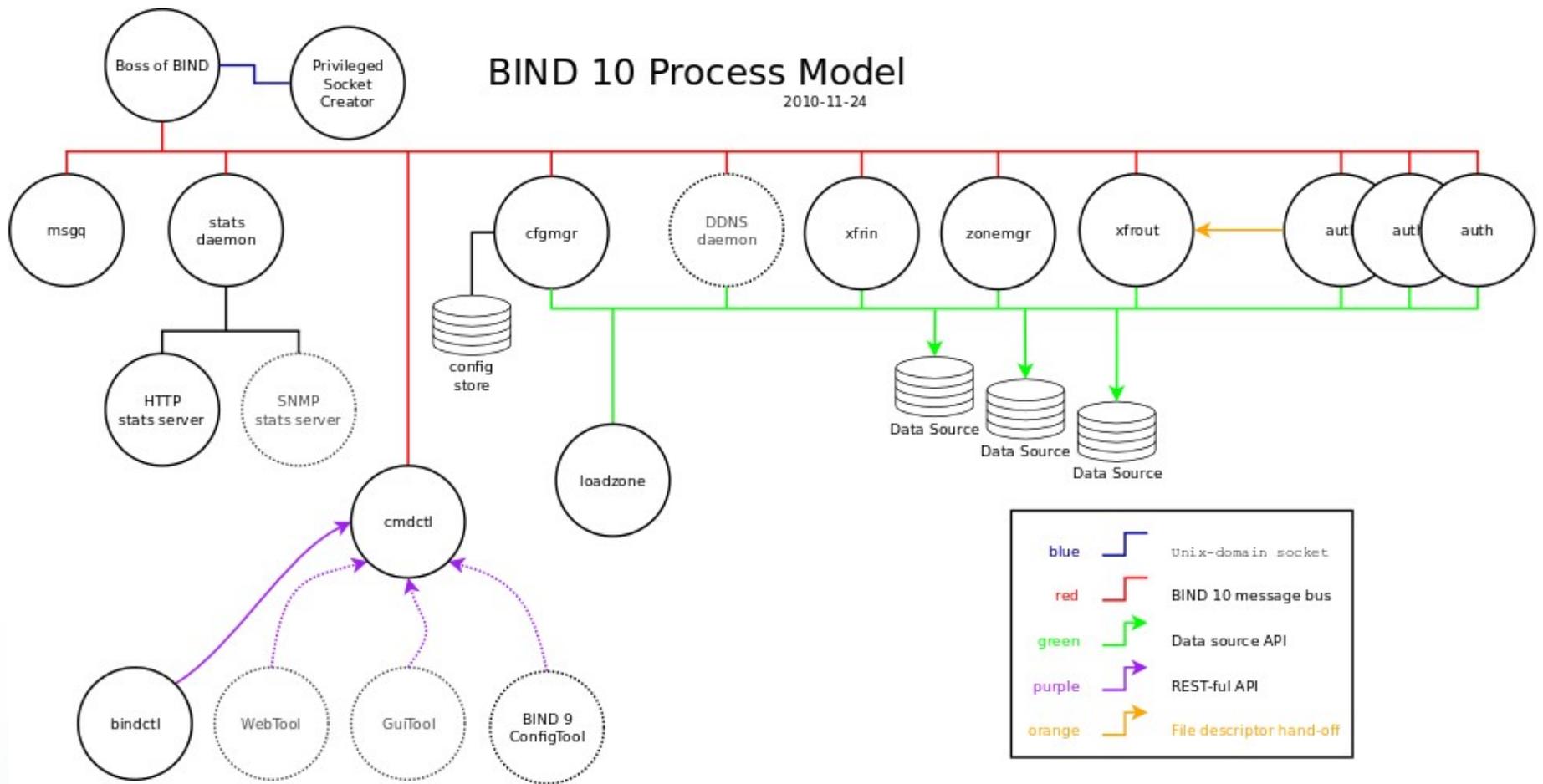
BIND 10 Process Model

2010-11-24



Privileged Socket Creator

- Ports < 1024 restricted to root
 - DNS runs on port 53
- We want to drop permissions ASAP
- We want port 53 at any time!
 - Administrator may reconfigure
- Solution: Privileged Socket Creator
 - Small, single purpose C++ program
 - Uses file descriptor trick to send sockets around



<https://bind10.isc.org/wiki/DesignDiagrams>

Sponsors

