Kea – DHCP servers in BIND10

Open Source BoF

RIPE66, Dublin, Ireland

16 May 2013

Tomek Mrugalski <tomasz@isc.org>

Agenda

1. About presenter and ISC

2. DHCP in BIND10 (codename Kea)
   - Reasons
   - Status

3. Performance

4.Kea Roadmap and plans

Who is Tomek?

- M.Sc., Ph.D from Gdansk University of Technology
- Primary author of Dibbler
  - Portable DHCPv6 implementation (srv, cli, relay)
  - Supports Win 2k-Win8, Linux, BSD, Solaris
  - Confirmed use in 34 countries
- 7 years at Intel (Network Quality Labs, chipsets group)
- 2 years at ISC
  - Lead Developer of BIND10 DHCP (Kea)
  - Occasional contributor to ISC-DHCP
- Active IETF participant since 2009
  - DHC WG co-chair
  - 2 RFCs, 15+ drafts

What is ISC?

Internet Systems Consortium, Inc. (ISC) is a non-profit 501(c)(3) public benefit corporation dedicated to supporting the infrastructure of the universal connected self-organizing Internet - and the autonomy of its participants - by developing and maintaining core production quality software, protocols, and operations.

BIND10 DHCP

Kea

Why DHCP rewrite?

- Existing code is 17 years old
- Hardware changed (many cores)
- Networks changed
- DHCP landscape changed
- New software development techniques
- Lacking performance
- Monolithic
- Documentation is lacking
- Complex code, difficult to extend

BIND10 DHCP
Codename Kea

- Common infrastructure with BIND10 DNS
  - On-line configuration
  - Logging
  - Statistics
- Performance is essential
- IPv6 is a first class citizen, not add-on
- C++ as a language of choice
- Multi-core support
- Switchable backends (mem+file, SQLite, MySQL, …)
- Hooks
- Modular
- Resiliency (fault isolation and recovery)

ISC

Kea: Current status (1)

DHCPv4 server
(b10-dhcp4)

- Supports DORA
- Direct/Relayed traffic

DHCPv6 server
(b10-dhcp6)

- Supports SARR
- Direct/Relayed traffic

- Address assignment, renewal, release, expiration
- On-line configuration (common for all BIND10)
- Switchable backends: MySQL, memfile
- Custom option definitions
  - Standard options
  - Custom formats
  - Nested options
  - Option namespaces

Kea: Current status (2)

DHCPv4 server
(b10-dhcp4)

DHCPv6 server
(b10-dhcp6)

Perfdhcp
Performance Tool
(stand alone)

libdhcp++
- general purpose DHCP library
- v4/v6 packet parsing/assembly
- v4/v6 options parsing/assembly
- interface detection (Linux, other OSes planned)
- socket management

Kea: Work to Date (2)

- Documentation
  - BIND10 Guide
  - BIND10 Developer's Guide
  - Man pages
- Designs
  - Hooks
  - Lease/database design
  - Option Definition Design

http://bind10.isc.org/docs/bind10-guide.html
http://bind10.isc.org/wiki/Kea

DHCP Performance

DHCP Performance
Problem Space

Vendors often provide performance results, why measure it again?

- Marketing data is always trustworthy, right?
- Your HW may differ from reference HW (CPU, disk, fs, OS,…)
- Your traffic model may differ

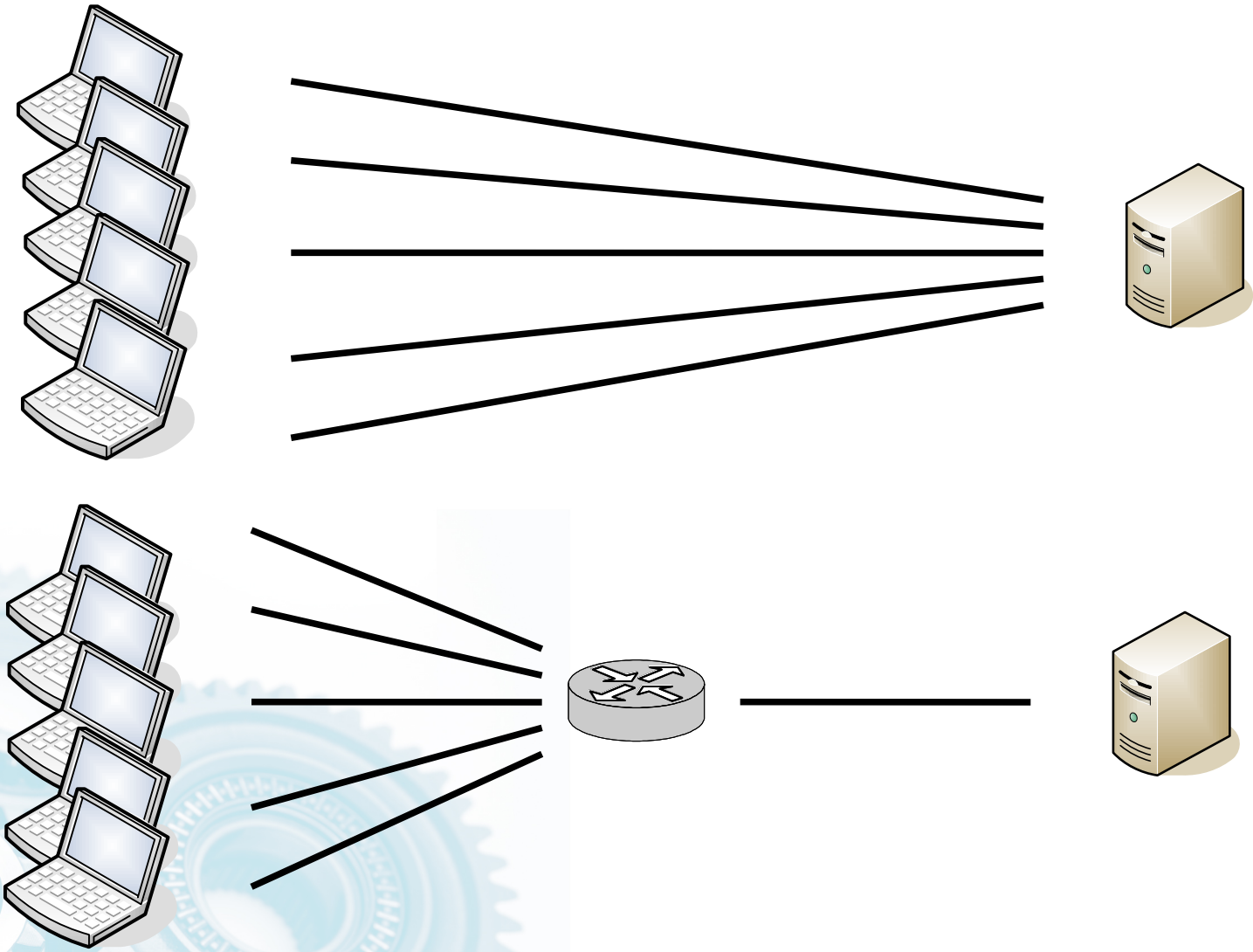
Conclusion:

The most reliable measurements

are ***your own***.

DHCP Performance Measurements (1)

DHCP Performance Measurements (2)



Perfdhcp :: Overview

- No feasible alternatives
  - are outdated (e.g. v4 only)
  - commercial (dedicated test HW is $$$$)
- Need a tool that is:
  - Flexible (lots of options and knobs)
  - Portable (Linux, BSD, perhaps Solaris)
  - Test any conformant implementation
  - Free (open source)
- Started project on our own

Perfdhcp :: Status

- Open source (ISC), currently Linux, but BSD and Solaris planned
- DHCPv4 & DHCPv6 (2-way & 4-way exchanges)
- Support for packet template files (optional)
- Server/interface selection (multicast/unicast)
- Parameterized traffic/test
  - # of clients,
  - # of transactions/sec,
  - best effort test,
  - test duration,
  - number of requests,
  - max number/% of drops …
- Diagnostics selector
- Measurements

```
Rate: 986.6 exchanges/second, expected rate:

***Statistics for: SOLICIT-ADVERTISE***
sent packets: 9866
received packets: 9866
drops: 0
orphans: 0

min delay: 0.168 ms
avg delay: 0.263 ms
max delay: 0.655 ms
std deviation: 0.039 ms
```

```
perfdhcp
[-hv] [-4|-6] [-r<rate>] [-t<report>] [-R<range>] [-b<base>]
[-n<num-request>] [-p<test-period>] [-d<drop-time>] [-D<max-drop>]
[-l<local-addr|interface>] [-P<preload>] [-a<aggressivity>]
[-L<local-port>] [-s<seed>] [-i] [-B] [-c] [-1]
[-T<template-file>] [-X<xid-offset>] [-O<random-offset]
[-E<time-offset>] [-S<srvid-offset>] [-I<ip-offset>]
[-x<diagnostic-selector>] [-w<wrapped>] [server]
```

Perfdhcp :: Roadmap

2013: No specific plans (unfunded)

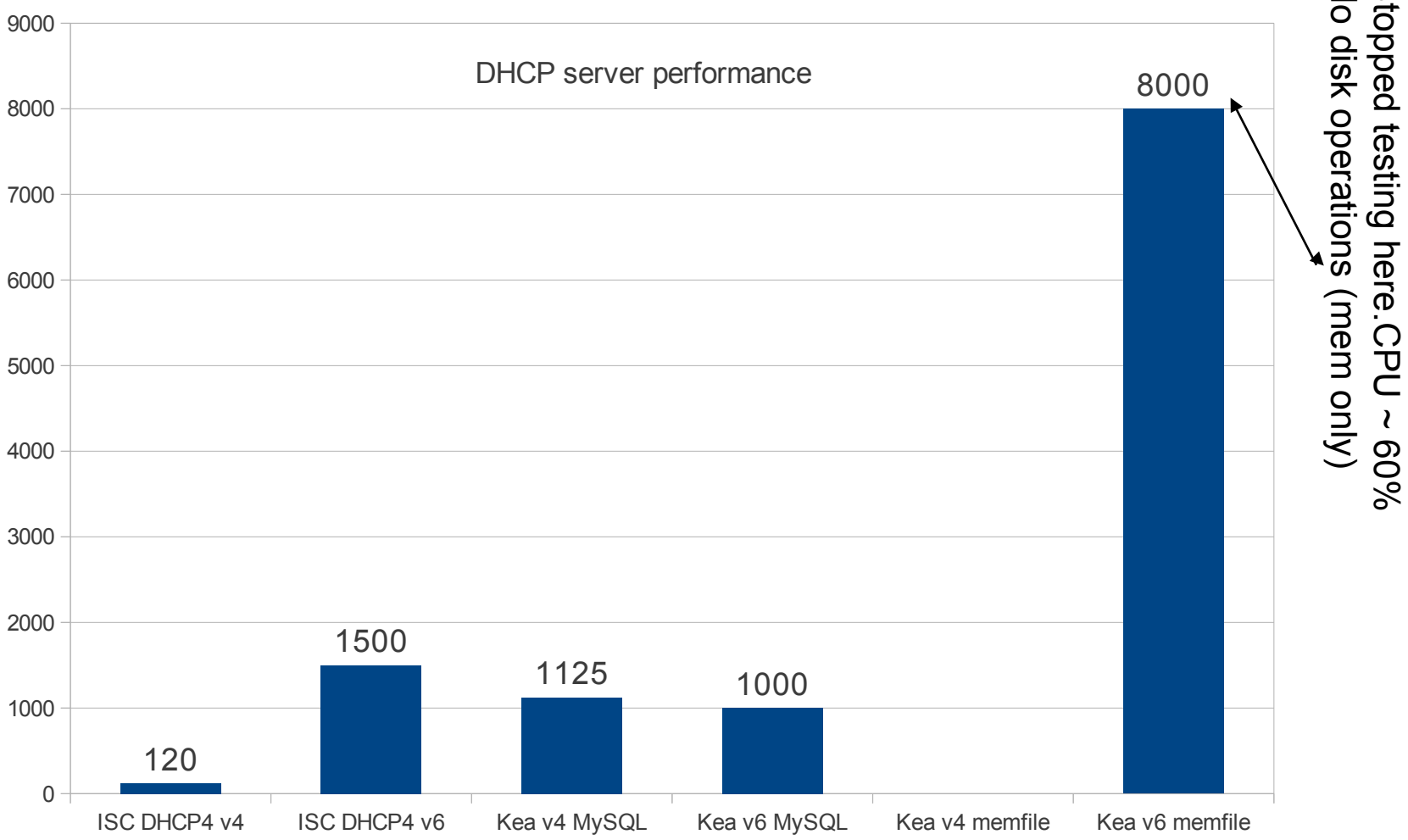
- Implement support for Prefix Delegation
- Relays
  - Traffic via relays
  - Relay options (subscriber-id, remote-id,...)
  - DOCSIS3.0 options
- Expand customization
- Improve response validation

Long term: maintain and develop

Kea :: Performance Results(1)

- DHCP server run on a beefy hardware
  - HP ProLiant DL360 G7
  - Intel(R) Xeon(R) CPU E5649 (24 logical CPUs)
  - 72GB ram
  - HP Smart Array P410i + 10k rpm disks
- Client traffic generated by perfdhcp
- Performance may go...
  - ...up (optimizations, mutli-core)
  - ...down (new features)

Kea :: Performance Results (2)



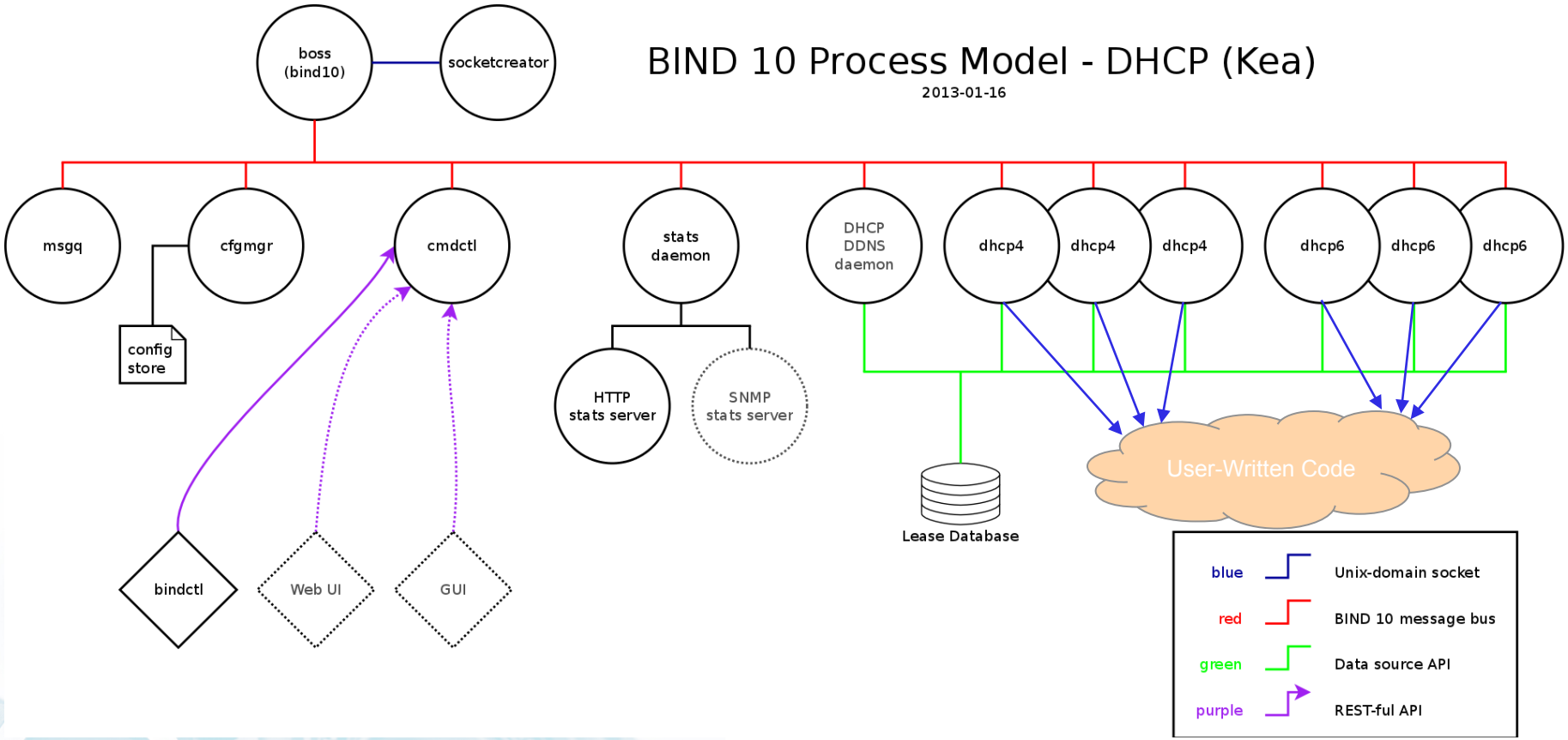
* initial data. Your mileage may vary.

Kea Plans

2013 Kea Roadmap

- Support for directly-connected IPv4 clients
- Support for IPv6 relays
- DDNS daemon
- Hooks
- Separate DNS/DHCP build in BIND10
- Optional: Extend OS support to BSD, Solaris (?)

DHCP Hooks (1)


BIND 10 Process Model - DHCP (Kea)
2013-01-16
boss (bind10)
socketcreator
msgq
cfgmgr
config store
cmdctl
stats daemon
HTTP stats server
SNMP stats server
DHCP DDNS daemon
dhcp4
dhcp4
dhcp4
dhcp6
dhcp6
dhcp6
Lease Database
User-Written Code
bindctl
Web UI
GUI
blue
Unix-domain socket
red
BIND 10 message bus
green
Data source API
purple
REST-ful API

DHCP Hooks (2)

- Set of hooks to be included in the code:
  - Call out to user code at defined points in packet processing
  - Replaces “conditional” configuration processing in DHCP4
- API design in progress
- Comments are more than welcome

http://bind10.isc.org/wiki/DhcpHooks

Possible features

Unfunded ideas

- Mutli-core support
- Prefix delegation in DHCPv6
- DHCPv6 failover
- DHCPv4 failover
- Different backends (Postgres? Cassandra?)
- CPE market
- …

Interested?

Fully open source model

- ISC license (BSD-like), available for free
- GIT repo, trac tickets available
- Test, report bugs
- Submit patches

Contribute

- We are looking for sponsors (money and developers)
- Development contracts
- Review design documents (e.g. requirements)

Questions?

Thank you

isc.org