Control Plane Modularization and Decentralization for BGP – Design and Performance Evaluation

Tomas Klockar, Markus Hidell, Lenka Carr-Motyčková, Peter Sjödin
Overview

- Router
- BGP
- Solution
- Performance
Introduction

- Cooperation between LTU and KTH
- KTH
  - Router technology
- LTU
  - Algorithm properties of protocols
    + BGP
Motivation

- ForCES
  - Control element
  - Forwarding element
  - Cost

- BGP
  - Lack of overview within AS
  - Designed for monolithically system
Router architectural difference

• Normal router (Now)
  - Monolithic unit (one-router one location)
    + Control (one)
    + Forwarding (one)
  - Several routers might be needed for a AS
  - Several routers = several configurations

• Distributed router (Future)
  - One router several locations
    + Control (one or few)
    + Forwarding (several)
  - One router is enough for a AS
  - One configuration
  - Cheaper?
Monolithic router

- Control and forwarding
- Several router for an AS
Distributed router

- Control element
  - Protocol
- Forwarding element
  - Forwarding table
- Internal networks
  - Forwarding
  - Control
**BGP - Border Gateway Protocol**

- De-facto routing protocol on Internet
- AS – Autonomous System
- Inter-domain
  - E-BGP
- Intra-domain
  - I-BGP
- Path vector protocol
  - E.g. [AS1, AS2, AS4]
BGP - Details

- RIB-In
- Loc-RIB
- RIB-Out
- Policies
- Decision process
- BGP session
BGP - Problems

• Inter-domain
  - Policy disputes
  - Policy inconsistencies
  - Oscillation

• Intra-domain
  - Oscillation
  - Forwarding loops

• General
  - Session recovery
Statistics: Prefix distribution

- Algorithms
  - Contiguous
  - Round Robin
  - Fixed size
- Heritage
- Class-full
  - A(8-bit), B(16-bit), C (24-bit)
- Class-less
  - 8-32
- 160000 prefixes
Solution: Overview

- Forwarding table
- Decision process
  - RIB, Policy

AS

Winternet, 18 Aug 2005

EISLab, LTU
Solution: Overview

- **Session manager**
  - Communicates with BGP peers
  - One-to-many

- **Decision process**
  - BGP decision process
  - RIBs

- **Forwarding elements**
  - Forward without control

- **Control elements**
Solution: FIB

- Distributed calculation
- Backup
Experiment setup

- Peer
- Session manager
- Decision Processes
  - 1-14
- Gigabit network
- Linux machines
  - Zebra
Performance: No route processing

- Data
  - 1 prefix/update
  - 5 prefixes/update
  - 20 Prefixes/update
Performance: Algorithms

- Algorithms
  - Round robin
  - Contiguous

- Data
  - Ideal
  - Real
Future work

- Test robustness and failover of the BGP routing
Questions?
Statistics: Size of updates

- Burst size
- Many small updates
Statistics: Interval between updates

- 4-5 seconds
- Different days
- <30 sec EBGP timer
- 5 sec IBGP timer