ZeroSDN: A Highly Flexible and Modular Architecture for Full-range Network Control Distribution

Thomas Kohler, Frank Dürr, Kurt Rothermel

October 12th, 2017
Introduction – Control Plane as a Distributed System

- SDN paradigm
  - Separation of control plane and data plane
  - Logically centralized control
    - Global view
    - Distribution transparency
Introduction – Evolution of SDN Controller Architectures

- **Control Plane (CP)**
  - Monolithic
  - Replication support

- **Data Plane (DP)**
  - Fixed controller-instance assignment (master / slave)
Introduction – Evolution of SDN Controller Architectures

- **CP**
  - Modular
  - Monolithic
  - Replication support

- **DP**
  - Fixed controller instance assignment (master / slave)

---

**CP**
- Modular
- Monolithic
- Replication support

**DP**
- Fixed controller instance assignment (master / slave)
Introduction – Evolution of SDN Controller Architectures

- CP
  - Modular
  - Distributed
    - Partitioning
    - Replication support

- DP
  - CF partitioning
  - controller mapping
  - Dynamism

DevoFlow, Kandoo
Introduction – Evolution of SDN Controller Architectures

• Observations & Shortcomings

1. Fully (logically) centralized control model
   ▪ No control communication between switches
   ▪ No switch-local logic (decision making)

2. Heavyweight controllers
   ▪ Modularization frameworks typically heavy-weight
   ▪ Lightweight controllers do not have distribution capabilities

3. Tight switch-controller coupling
   ▪ Fixed mapping
   ▪ No module dynamism
Outline

• Introduction – Evolution of SDN Controller Architectures

• **Architecture for Full-range Network Control Distribution**

• Highly Flexible Control Plane Distribution

• Implementation & Evaluation

• Roadmap to a Highly Scalable Holistic System Control Plane

• Conclusion
Contributions

- SDN Controller Architecture with high flexibility in distribution
  - From logically centralized to fully decentralized control

- Micro-Kernel controller architecture for distributed light-weight controller modules (controllets)

- Pushing down network control to switches (local logic)
  - While leveraging global knowledge

- Decoupling of controllets through a message bus
  - Content-based filtering of data- & control plane events
Architecture for Full-range Network Control Distribution

- μ-Kernel architecture
- Modularity
- Full distribution
  - Local & remote controllets
- Replication
- Message bus
- Event-based network control
Message Bus – Decoupling Event-based Controllets

• Data- / Control- plane events
  ◦ Packets or state changes of switches and data plane end-hosts (DPE)
  ◦ Control coordination messages and control state changes (CPE)

• Message bus paradigm
  ◦ Routing of events to subscribers w/ content-based filtering
  ◦ Emulation of P2P, multicast communication patterns
  ◦ Transparently implement arbitrary delivery semantics

Ravana: Controller Fault-tolerance in SDN; Rexford et al. SOSR‘15
Outline

- Introduction – Evolution of SDN Controller Architectures
- Architecture for Full-range Network Control Distribution
- **Highly Flexible Control Plane Distribution**
  - Fully Distributed Control
  - Local Logic: Local Data Plane Event Processing (LDPEP)
- Implementation & Evaluation
- Roadmap to a Highly Scalable Holistic System Control Plane
- Conclusion
Highly Flexible Control Plane Distribution – Fully Distributed Control

- Network control distribution
  - Top end:
    - full centralization (majority of current SDN controller architectures)
  - Bottom end:
    - full distribution of network control
- Our approach: offer full range of distribution
Highly Flexible Control Plane Distribution – Local Data Plane Event Processing

- So far: enhancing switch capabilities

- Local Data Plane Event Processing (LDPEP)
  - Running controller (decision making) on switch hardware
  - Increased resilience against controller-failures, inherent load balancing
  - Ideally: entirely local decision making $\rightarrow$ most timely reaction
  - .. while having access to neighbours and global knowledge $\rightarrow$ optimality
  - Scale scope of local control with available computing resources
  - Trade-offs (*intermediate procedures*)

---

InSPired Switches
SOSR‘15

OpenState
SIGCOMM-CCR‘14

P4
SIGCOMM-CCR‘14

---

Optimality
Resources
Time
Local Data Plane Event Processing

$S_i$ switch

white-box

Decision Process

Control Logic

query/publish

event notif.

query/modify

add/mod. flow

forward packet

Cache

TCP

TCP (OpenFlow)

PCI-E

OF-Agent

PKT_IN

FLOW_MOD

PKT_OUT

Flow Table

TCAM

query

update

$p_0$ fast-path

$p_i$

Host$_{src}$

Host$_{dst}$

ASIC - DP CP - CPU

μ-Kernel

MSG

BUS

subscr./notify

aggregate policies

Decision Logic

TCP

query/modify

add/mod. flow

forward packet

Cache

TCP

TCP (OpenFlow)

PCI-E

Flow Table

TCAM

query

update

$p_0$

Host$_{src}$

Host$_{dst}$

ASIC - DP CP - CPU

μ-Kernel

MSG

BUS

subscr./notify

aggregate policies

Decision Logic

TCP

query/modify

add/mod. flow

forward packet

Cache

TCP

TCP (OpenFlow)

PCI-E

Flow Table

TCAM

query

update

$p_0$

Host$_{src}$

Host$_{dst}$

ASIC - DP CP - CPU

μ-Kernel

MSG

BUS

subscr./notify

aggregate policies

Decision Logic

TCP

query/modify

add/mod. flow

forward packet

Cache

TCP

TCP (OpenFlow)

PCI-E

Flow Table

TCAM

query

update

$p_0$

Host$_{src}$

Host$_{dst}$

ASIC - DP CP - CPU

μ-Kernel

MSG

BUS

subscr./notify

aggregate policies

Decision Logic

TCP

query/modify

add/mod. flow

forward packet

Cache

TCP

TCP (OpenFlow)

PCI-E

Flow Table

TCAM

query

update

$p_0$

Host$_{src}$

Host$_{dst}$

ASIC - DP CP - CPU

μ-Kernel

MSG

BUS

subscr./notify

aggregate policies

Decision Logic

TCP

query/modify

add/mod. flow

forward packet

Cache

TCP

TCP (OpenFlow)

PCI-E

Flow Table

TCAM

query

update

$p_0$

Host$_{src}$

Host$_{dst}$
InFEP – Lightweight Virtualization of Distributed Control on White-box Networking Hardware, Kohler et al., ManSDN/NFV - CNSM'17
Outline

- Introduction – Evolution of SDN Controller Architectures
- Architecture for Full-range Network Control Distribution
- Highly Flexible Control Plane Distribution
- Implementation & Evaluation
- Roadmap to a Highly Scalable Holistic System Control Plane
- Conclusion
Implementation

- **ZMF**: modular execution framework
  - Dependency and life-cycle management
  - Message bus: communication middleware: ZeroMQ

- **ZeroSDN**: distributed SDN controller application
  - Message de-/serialization: Google Protocol Buffers
  - Core controllets implemented (OF 1.0 - 1.3)

http://zerosdn.github.io/
Evaluation – Raw Controller Performance – Methodology

- Raw controller performance
  - 12 nodes (Xeon, 4 x 3.4GHz), 1GbE \textit{local}
  - 6 node pairs \textit{remote}

- cbench emulates switches
- \# emulated switches \( s = 1 \)
- \# end hosts \( e = 10000 \)
- Observation time \( T = 12\text{min} \)
Evaluation – Raw Controller Performance – Results

- Distribution impact on throughput
  - factor 0.5

- Moderate locality impact on throughput

- Remote vs. local latency
  - factors 2 to 6
Evaluation – Hardware Switch – Methodology

- Performance on a Whitebox-Networking Hardware Switch

- Provoke packet processing in switch control plane

- Measure data plane RTT
Evaluation – Hardware Switch – Results

- Remote TCP connection latency impact factors:
  2 (ZSDN-AFC), 1.9 (NOX), 1.5 (ZSDN)
Outline

• Introduction – Evolution of SDN Controller Architectures
• Architecture for Full-range Network Control Distribution
• Highly Flexible Control Plane Distribution
• Implementation & Evaluation
• Roadmap to a Highly Scalable Holistic System Control Plane
• Conclusion
Roadmap to a Highly Scalable Holistic System Control Plane (Future Work)

• Complex network event processing
  ◦ Predicate logic for evaluation of events
    ▪ logical operators
    ▪ timing constraints
    ▪ sequences of events absence of events
  ◦ Distributed evaluation of complex events in controllets, subscribing to relevant simple events (DPE & CPE)

• Holistic system control plane:
  Application layer events & end-host control
  ◦ NFVs, end-systems (including VMs), applications
  ◦ Execute complex system management workflows
Conclusion

- Highly flexible & modular architecture for full-range network control distribution

- Lightweighted micro-kernel approach

- Decoupling of controllets through a message bus
  - Content-based filtering and hybrid processing of control plane and data plane events

- Exploiting locality of switches along with global knowledge
Thanks for your attention

Any Questions?

Contact & further information:

https://goo.gl/tYWSgW