neutron-lan

SDN study environement @ home

github.com/alexanderplatz1999

Last update: May 29th, 2014

Background and Motivation

- Too many SDN definitions
 - I have been confused a lot.
 - OpenFlow, OVSDB, Netconf, BGP extensions such as FlowSpec...
 - The latest addition: OpFlex (DevOps-like)
- What's the real SDN?
 - Let's develop SDN by myself and examine every definition.
- But, wait! I need a SDN study environment at home.
 - I am a poor guy, so I cannot buy expensive SDN-capable switches from Cisco, Juniper...

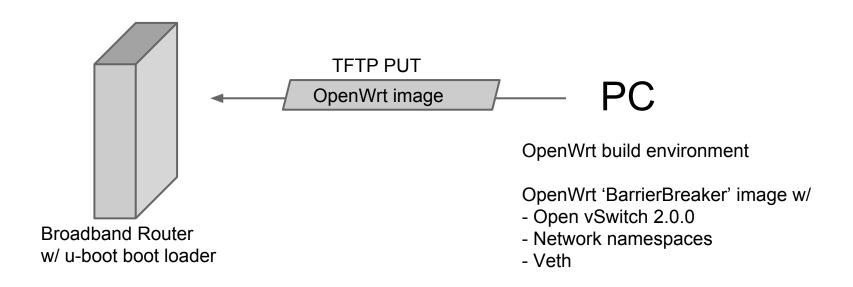
Strategy

- My budget is less than \$200.
- Switches/routers I purchased in Akihabara, Tokyo
 - Three \$40 broadband routers and one \$40 Raspberry Pi
- And I develop all the SDN software from scratch
 - But reuse existing networking software as much as possible, such as Open vSwitch
- Base knowledge/skills
 - SDN in the past: SIP and IP-PBX
 - OpenFlow, OpenStack neutron and SaltStack
 - Java and Python
 - HTML5 and CSS (a little)
- Let's develop neutron-like SDN for my home network ⇒ let's call it 'neutron-lan'

Project 'neutron-lan' characteristics

- Cheap routers as 'Baremetal Switch'
 - OpenWrt routers and Raspberry Pi
 - u-boot for installing new firmware
- OpenWrt kernet/kernel modules rebuilt(compiled) for supporting additional features
 - Open vSwitch 2.0.0, network namespaces(netns), veth and LXC.
- Home-made DevOps tool 'NLAN' for small routers w/ small memory/storage footprints
 - 100% Python implementation
 - YAML-based state rendering (Python's OrderedDict internally)
 - Model-driven service abstraction
 - Both imperative and declarative state definitions
 - OVSDB as a general-purpose config database
 - RFC7047 client library
 - With NLAN-specific OVSDB schema and schema extensions
 - MIME Mutipart messages to covery various kinds of info
 - CRUD/RPC Transaction outputs
 - Exceptions
 - Logging info
 - RFC7047 JSON-RPC transacions
- VXLAN-based edge-overlay for network virtualization
 - Broadcast tree per VNI
 - Distributed Virtual Routing topology as well as hub & spoke topology
- LXC for Network Functions Virtualization

Cheap routers as 'Baremetal Switch'



Test bed (cont'd)

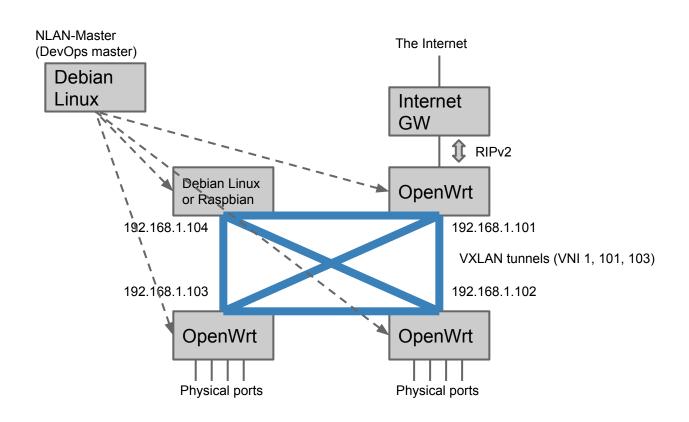


OpenWrt routers (and Home Gateway)

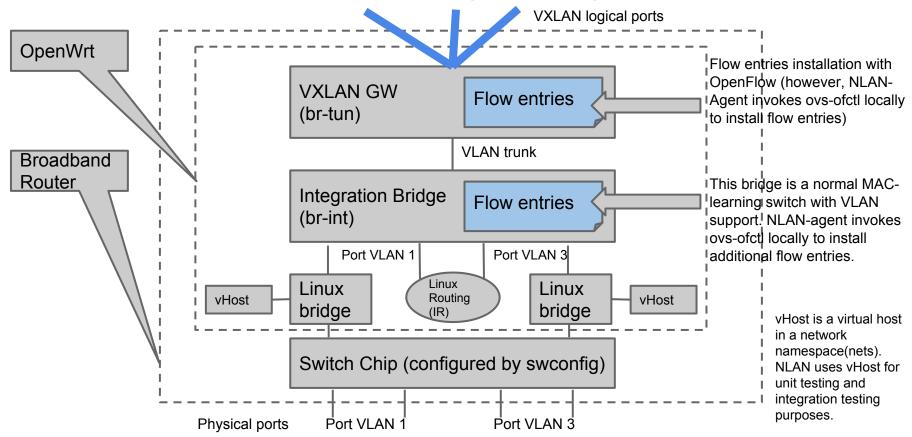


Raspberry Pi

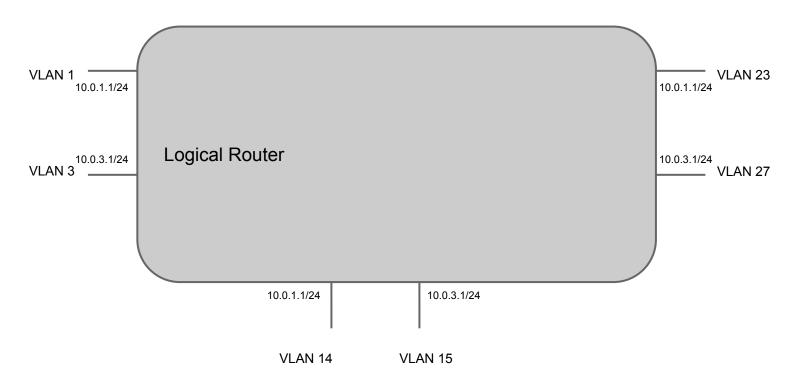
Test bed



OpenStack-neutron-like bridge configuration



Distributed Virtual Router (Logical view)

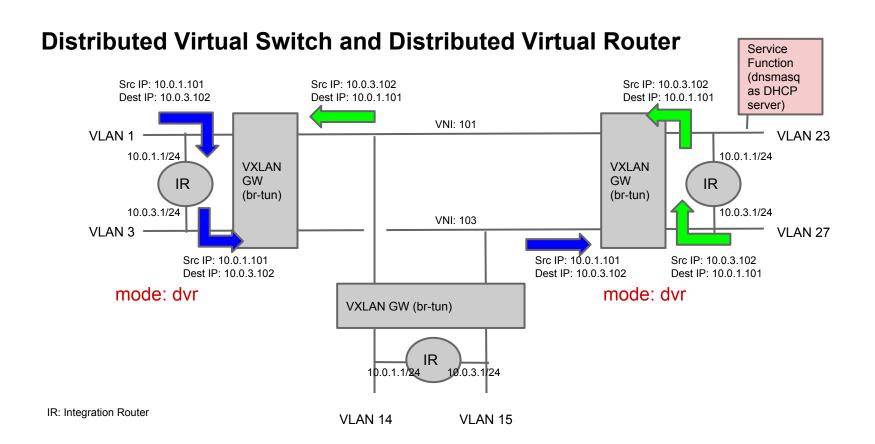


Virtual network topologies

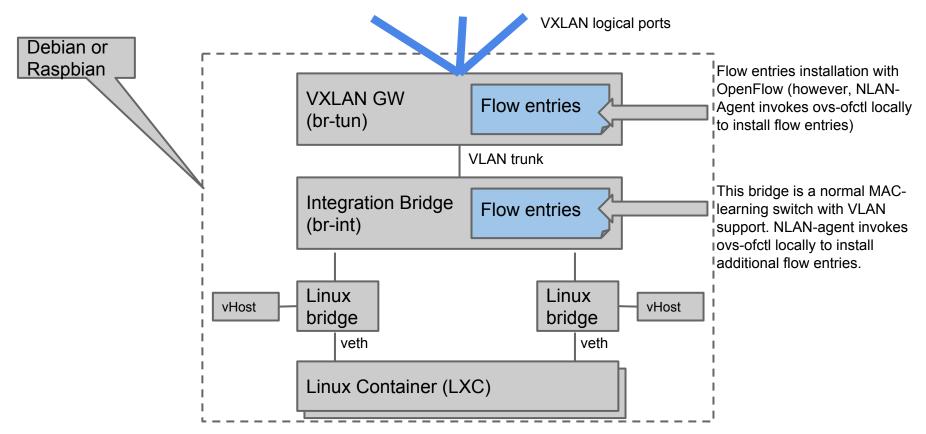
NLAN node operation mode	Virtual Network Topology
dvr	Distributed Virtual Router
hub	Hub & Spoke
spoke	Hub & Spoke
spoke_dvr	Mixture of DVR and Hub & Spoke

NLAN "subnets" state and its parmeters in YAML

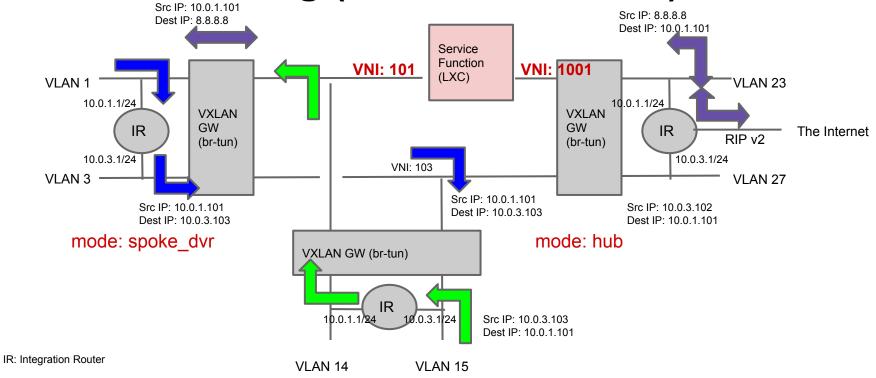
```
subnets:
- vid: 1
vni: 1001
ip_dvr:
addr: '10.0.1.1/24'
mode: hub mode can be 'dvr', 'hub', 'spoke' or 'spoke_dvr'
dhcp: enabled
ip_vhost: '10.0.1.101/24'
ports: [eth0.1]
peers: <peers>
```



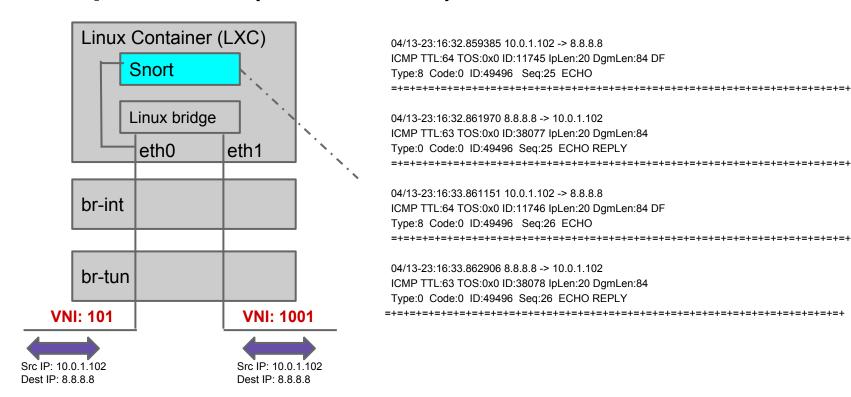
Service Function in Linux Container



Service Chaining (Service Insertion)



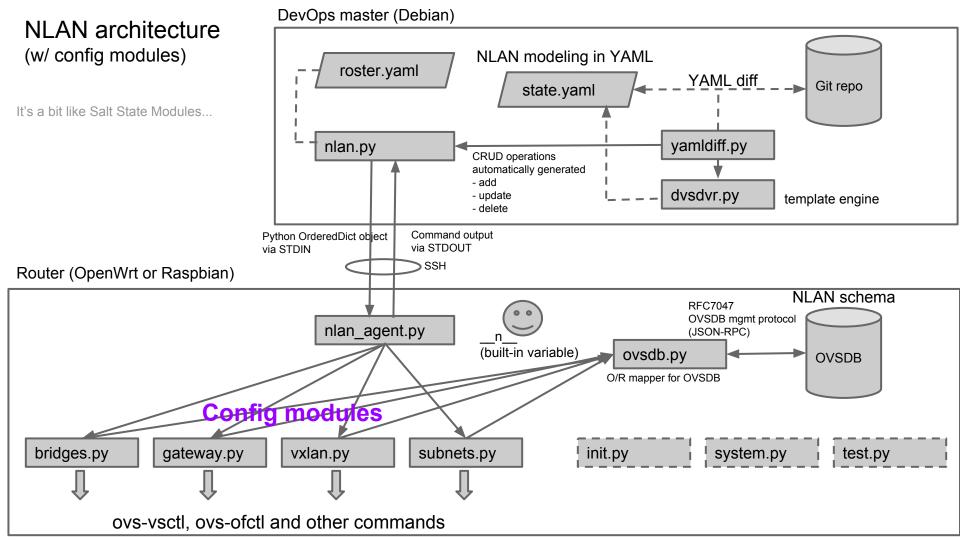
Example: Snort (in IPS mode) as Service Function

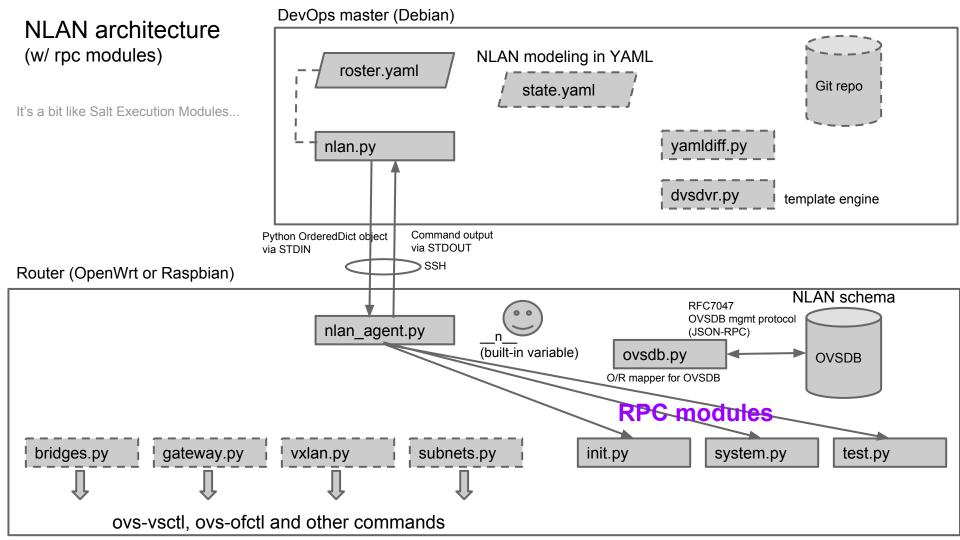


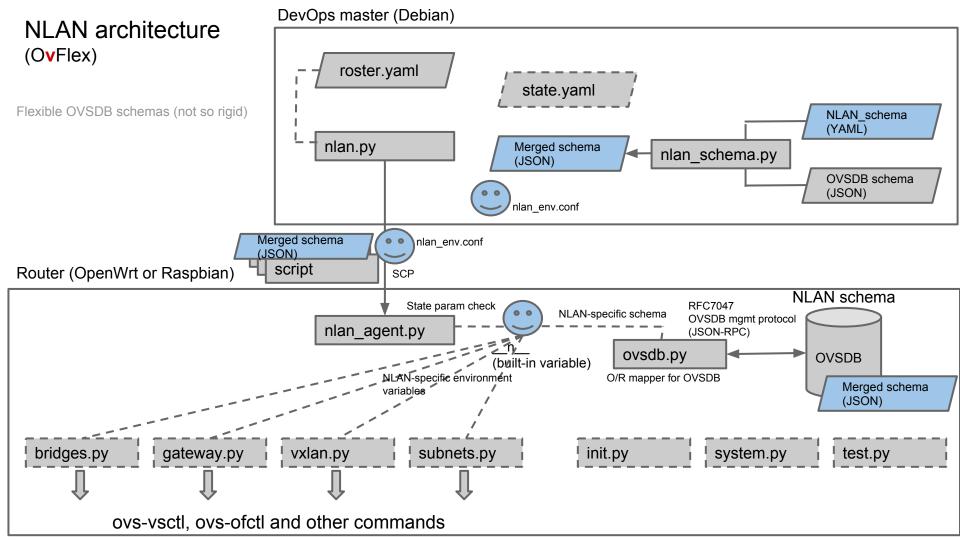
Snort is a free and open source IPS/IDS software.

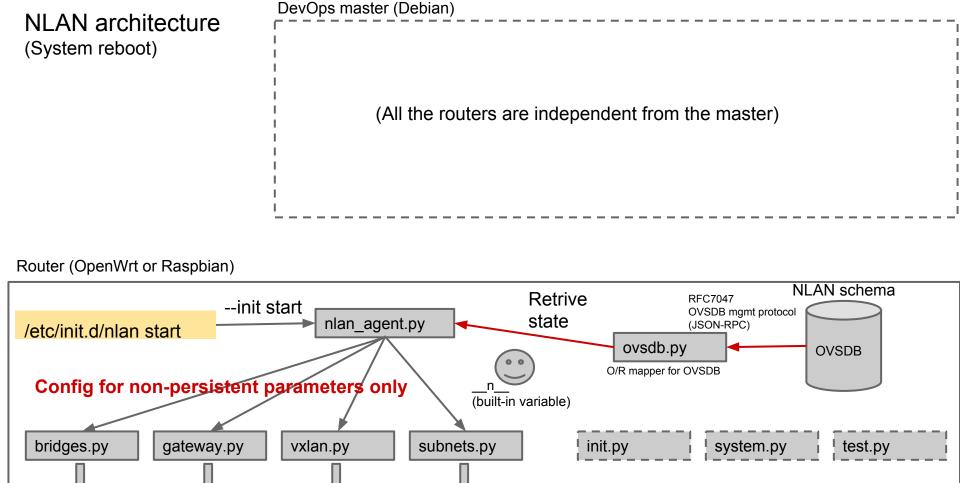
Home-made DevOps tool 'NLAN'

- 100% Python implementation
- Borrwed a lot of ideas from SaltStack
 - Model-driven approach
 - YAML-based state rendering w/ a simple template engine
 - Imperative/declarative state rendering
- Works with OpenWrt with minimal Python
 - opkg install python-mini
 - opkg install python-json
 - sshd
- OVSDB as a local config mgmt database
- State schema defined in YAML
 - merged with OVSDB schema in JSON



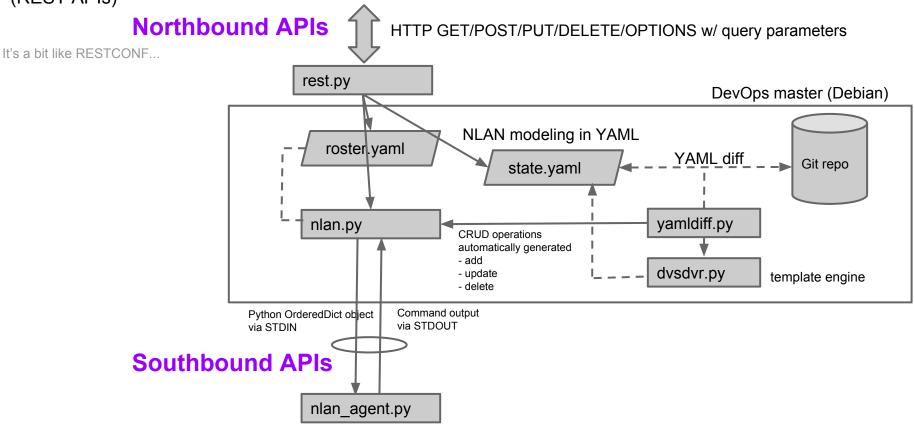






ovs-vsctl, ovs-ofctl and other commands (ovs-vsctl is skipped, since all the config is read from OVSDB when booting)

NLAN architecture (REST APIs)



OvFlex -- Flexible OVSDB schemas

OpFlex:

http://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-731302.html

OvFlex:

"It uses dynamic, flexible schemas for interaction with devices, effectively increasing the network to a higher common denominator feature set. The Open vSwitch Database (OVSDB) management protocol allows configuration of high-level abstract data models as well as basic primitives such as ports and bridges, and can support SDN geeks' innovations"

OVSDB schema (Open_vSwitch database)

NLAN schema in YAML

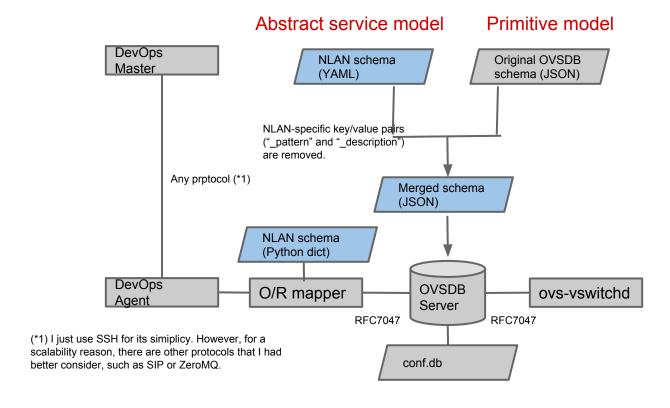
```
NLAN Subnet:
          columns:
          vni:
          type:
          key: {type: integer, minInteger: 0, maxInteger: 16777215}
          min· 1
          max· 1
          description: "Virtual network identifier"
          vid:
          key: {type: integer, minInteger: 0, maxInteger: 4095}
          max· 1
          description: "VLAN ID"
          ip dvr:
          key: {type: string, enum: [set, [addr, mode, dhcp]]}
          value: {type: string, pattern: {addr: ipv4 prefix, mode: dvr mode, dhcp:
string}}
          min: 0
          max: 3
          description: "Distributed Virtual Router setting"
          ip vhost:
          key: {type: string, pattern: ipv4 prefix}
          min: 0
          description: "Virtual host in a linux network namespace"
          default gw:
         type:
          key: {type: string, pattern: ipv4 address}
          min: 0
          max· 1
          description: "Default GW address for this subnet"
```

Original OVSDB schema in JSON

```
{"name": "Open vSwitch",
"version": "7.4.2",
"cksum": "951746691 20389",
"tables": {
  "NLAN": {
        "columns": {
        "bridges": {
        "type": {"key": {"type": "uuid",
                          "refTable": "NLAN Bridges"},
                 "min": 0, "max": 1}},
        "services": {
        "type": {"key": {"type": "uuid",
                          "refTable": "NLAN Service"},
                 "min": 0, "max": "unlimited"}},
        "gateway": {
        "type": {"key": {"type": "uuid",
                          "refTable": "NLAN Gateway"},
                 "min": 0, "max": 1}},
        "vxlan": {
        "type": {"key": {"type": "uuid",
                         "refTable": "NLAN VXLAN"},
                 "min": 0, "max": 1}},
        "subnets" · {
        "type": {"key": {"type": "uuid",
                          "refTable": "NLAN Subnet"},
                 "min": 0, "max": "unlimited"}}},
        "isRoot": true,
        "maxRows": 1},
```

Merging schemas

OVSDB schema can also express more abstract data models.

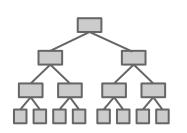


NLAN states in **OVSDB**

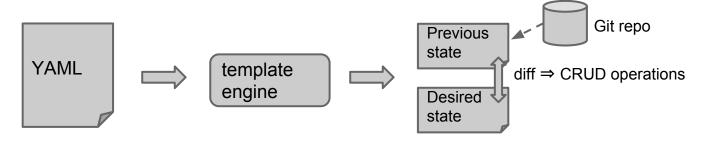
(ovsdb-client dump Open_vSwitch)

NLAN table _uuid	bridges		gateway vxlan	sei	rvices s	subnets				
0fc74b76-a3cb-40d7-9de5-26ea5d93e908 8, 6a607cd4-d0d0-4c7b-9b02-bf26c2acd						0ea1caea-5	7ab-4c29-913a-b91	6e9032	231	
NLAN_Bridges table _uuid	controller	ovs_bridges								
ae90946b-b00c-4e39-9ef0-2eda7ed89d66	[]	enabled								
NLAN_Gateway table _uuid	network rip	р								
dd3a47e2-b22c-45c6-a9f2-59f70f01eab7	"eth2" ena	abled								
NLAN_Service table _uuid chain name										
NLAN_Subnet tableuuid	default_gw	ip_dvr		ip_vhost	peers			port	s v i	d vni
0ea1caea-57ab-4c29-913a-b916e9032318 abeeafb2-bedf-41f6-bf03-acc1ab773574 6a607cd4-d0d0-4c7b-9b02-bf26c2acd343	[]	{addr="192.168.100.1/24", {addr="10.0.1.1/24", dhcp-{addr="10.0.3.1/24", dhcp-	=enabled, mode=hub}		["192.1	.68.1.104"]	-	[]	1	1001
NLAN_VXLAN table _uuid	local_ip	remote_ips								
4b0a1be2-1149-4665-943e-76088c20f271	"192.168.1	.101" ["192.168.1.102", "1	92.168.1.103", "192.	168.1.104"]						

Model-driven service abstraction (cont'd)



Step1: define network service model

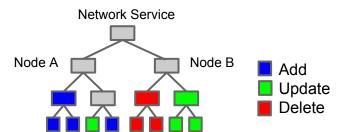


Step2: write the mode as "desired state" in YAML format w/ some placeholders for a template engine

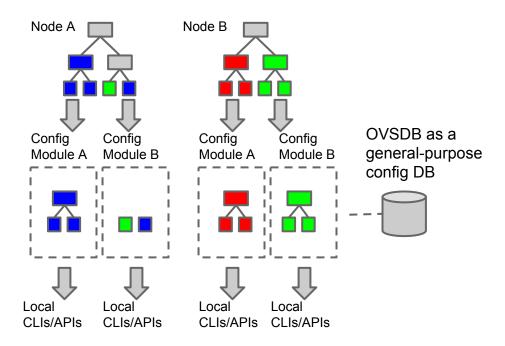
Step3: write a template engine to fill out the placeholders automatically.

Step4: NLAN-Master (nlan. py and yamldiff.py) generates CRUD operations comparing the desired state with the previous state

Model-driven service abstraction



Step5: Now CRUDoperations (= difff between the previous and the desired states) are in the form of Python OrderedDict object Step6: NLAN-Agent (nlan_agent.py) routes the CRUD operations to corresponding nodes/modules



Template and placeholders example

```
#!template.dvsdvr
                                                                                openwrt1:
                                                template engine
openwrt1:
                                                "template.
                                                                                  vxlan:
                                                dvsdvr
 vxlan:
                                                                                        local ip: '192.168.1.101'
                                       - generates a local ip address
        local ip: <local ip>
                                                                                        remote ips: ['192.168.1.102', '192.168.1.103', '192.168.1.104']]
                                       - generates VXLAN remote ip addresses
                                       - generates broadcast tree per VNI
        remote_ips: <remote_ips>
                                                                                  subnets:
                                       - automatically resolves dependencies
 subnets:
                                                                                        - vid: 1
                                       among parameters
        - vid: 1
                                                                                        vni: 101
        vni: 101
                                                                                        ip dvr: {addr: '10.0.1.1/24', mode: dvr}
        ip dvr: {addr: '10.0.1.1/24'. mode: dvr}
                                                                                        ip vhost: '10.0.1.101/24'
        ip vhost: '10.0.1.101/24'
                                                                                        ports:
                                                                                        - eth0.1
        ports:
        - eth0.1
                                                                                        peers: ['192.168.1.102', '192.168.1.104']
        peers: <peers>
```

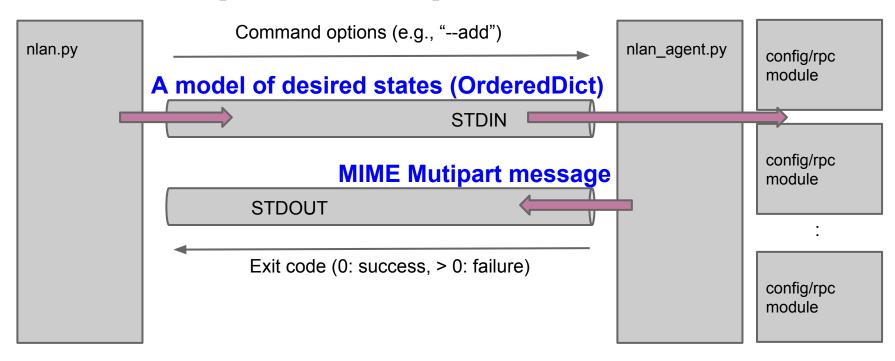
Model of desired state

- NLAN-Master sends Python OrderedDict to NLAN-Agent via ssh STDIN.
- To be exact, string form of an OrderedDict object (sort of object serialization).

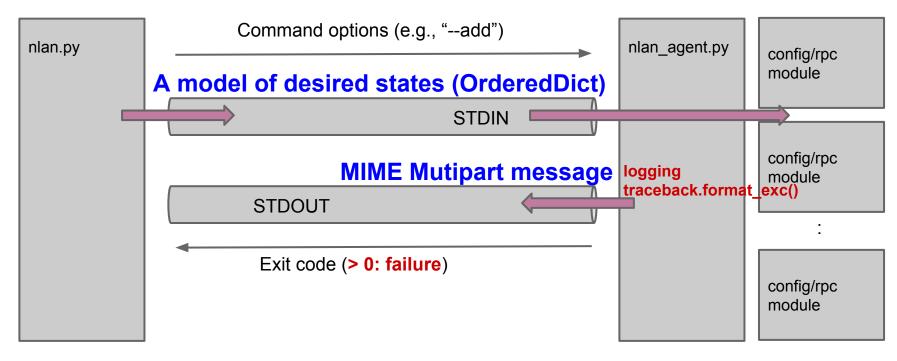
```
"OrderedDict([('bridges', {'ovs_bridges': 'enabled'}), ('gateway', {'network': 'eth2', 'rip': 'enabled'}), ('vxlan', {'remote_ips': ['192.168.1.103', '192.168.1.102', '192.168.1.104'], 'local_ip': '192.168.1.101'}), ('subnets', [{'peers': ['192.168.1.102', '192.168.1.103'], 'vid': 2, '_index': ['vni', 1], 'ip_vhost': '192.168.100.101/24', 'vni': 1, 'ip_dvr': OrderedDict([('addr', '192.168.100.1/24'), ('mode', 'dvr')])}, {'peers': ['192.168.1.102', '192.168.1.103'], 'vid': 3, '_index': ['vni', 103], 'ip_vhost': '10.0.3.101/24', 'vni': 103, 'ip_dvr': OrderedDict([('addr', '10.0.3.1/24'), ('mode', 'dvr')])}, {'peers': ['192.168.1.104'], 'vid': 1, '_index': ['vni', 1001], 'ip_vhost': '10.0.1.101/24', 'vni': 1001, 'ip_dvr': OrderedDict([('addr', '10.0.1.1/24'), ('mode', 'hub')])}])])"
```

- Imperative/declarative state representation.
- I don't use JSON, since NLAN is 100% Python implementation.

NLAN Request/Response over SSH



Exception handling and debugging



Trouble shooting becomes much easier by sending traceback.format_exc() and raw command outputs (test/plain) to nlan.py.

MIME Multipart example

```
root@debian:~# nlan -t openwrt1 --get bridges ovs bridges -M
*** Response from router:openwrt1,platform:openwrt
From nobody Thu May 29 03:16:42 2014
MIME-Version: 1.0
Content-Type: multipart/mixed; boundary="NKwDX7NAzowkZDZymawAVVwY3TcnvS1D"
--NKwDX7NAzowkZDZymawAVVwY3TcnvS1D
Content-Type: test/plain
Content-Description: Default out
X-NLAN-Type: default out
--NKwDX7NAzowkZDZymawAVVwY3TcnvS1D
Content-Type: application/x-nlan
Content-Description: CRUD get
X-NLAN-Type: crud response
OrderedDict([('bridges', {'ovs bridges': 'enabled'})])
--NKwDX7NAzowkZDZymawAVVwY3TcnvS1D
Content-Type: application/x-nlan
Content-Description: NLAN Response
X-NLAN-Type: nlan response
OrderedDict([('message', 'Execution completed'), ('exit', 0)])
--NKwDX7NAzowkZDZymawAVVwY3TcnvS1D--
```

Just use Python's "email" package for parsing the message.

Showing exception from remote routers

```
root@debian:~# nlan -t openwrt1 --add bridges ovs bridges=enabled
*** Response from router:openwrt1,platform:openwrt
exception: CmdError
message: Command execution error
traceback: Traceback (most recent call last):
 File "/opt/nlan/nlan agent.py", line 158, in route
       call()
  File "/opt/nlan/config/bridges.py", line 35, in add
       cmdp('ovs-vsctl add-br br-int')
  File "/opt/nlan/cmdutil.py", line 115, in check cmdp
                                                                                        traceback.format exc()
       return cmd('check call', True, *args)
  File "/opt/nlan/cmdutil.py", line 20, in cmd
       return cmd2(check=check, persist=persist, args=cmd args)
  File "/opt/nlan/cmdutil.py", line 62, in cmd2
       raise CmdError(argstring, e.returncode, out)
CmdError: Command execution error
command: ovs-vsctl add-br br-int
stdout: None
exit: 1
operation: add
progress: [('bridges', False)]
```

Showing debug information from remote routers

```
--37axIIOrg4QKTUh6L12jpcwxt3HC3Vm2
Content-Type: text/plain
Content-Description: [DEBUG] 2014-05-29 10:23:50, 057 module:cmdutil,
function:log, router:openwrt3
X-NLAN-Type: logger
cmd: ip netns exec ns1 ip route add default via 192.168.100.3 dev eth0
--37axIIOrq4QKTUh6L12jpcwxt3HC3Vm2
Content-Type: text/plain
Content-Description: [DEBUG] 2014-05-29 10:23:50, 060 module:cmdutil,
function:log, router:openwrt3
X-NLAN-Type: logger
cmd: ip addr add dev int dvr1 192.168.100.3/24
--37axIIOrq4QKTUh6L12jpcwxt3HC3Vm2
Content-Type: text/plain
Content-Description: [DEBUG] 2014-05-29 10:23:50, 065 module:cmdutil,
function:log, router:openwrt3
X-NLAN-Type: logger
cmd. route
Kernel IP routing table
Destination
               Gateway
                               Genmask Flags Metric Ref Use Iface
10.0.2.0
                               255.255.255.0 U 0 0
                                                                       0 eth0
192 168 1 0
                               255.255.255.0 U
                                                    0
                                                             0
                                                                      0 eth2
                               255.255.255.0 U 0 0
192 168 56 0
                                                                      0 eth1
                               255 255 255 0 II
192 168 100 0 *
                                                                       0 int dvr1
```

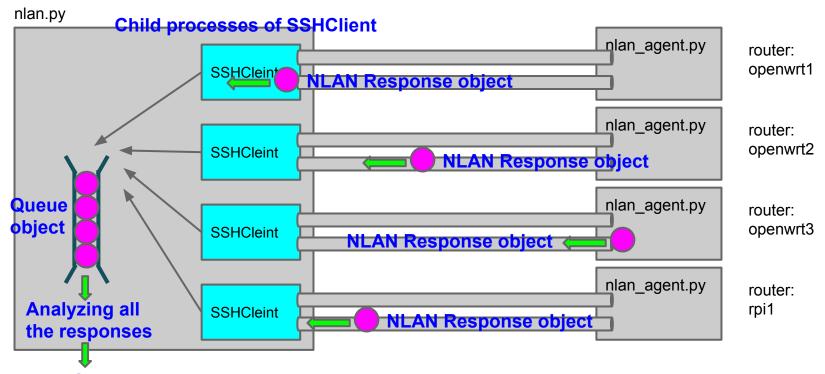
NLAN's logger can print every command executed at remote routers.

Showing transactions with OVSDB at remote routers

```
[DEBUG] 2014-05-29 03:45:53,322 module:ovsdb,function:crud,router:rpi1
CRUD operation (add): {'peers': ['192.168.1.102', '192.168.1.103'], 'ports': ['mz.101'], 'vni': 101, 'vid': 1}
[DEBUG] 2014-05-29 03:45:53,323 module:ovsdb, function: send, router:rpi1
request:
 id: 89846
 method: transact
 params:
 - Open vSwitch
 - op: insert
         peers:
         - set
         - - 192.168.1.102
         - 192.168.1.103
         ports:
         - set
         - - mz.101
         vid: 1
         vni: 101
         table: NLAN Subnet
         uuid-name: temp 515801
 - mutations:
          - - subnets
         - insert
         - - set
         - - - named-uuid
         - temp 515801
         op: mutate
         table: NLAN
         where: []
response:
 error: null
 id: 89846
 result:
 - uuid:
          - cb6bc9a6-0a86-415f-a82c-19f6a5e841d0
 - count: 1
```

NLAN can show RFC7047 JSON-RPC transactions in the form of YAML.

Parallel SSH sessions



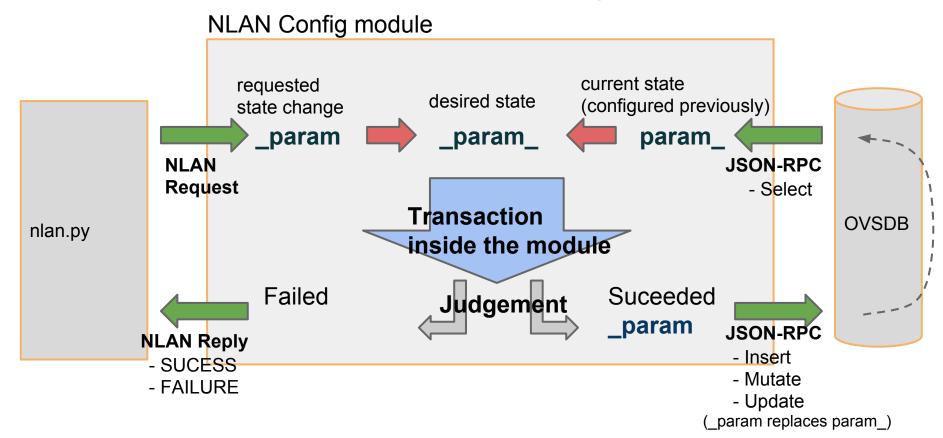
Transaction Summary output

Transaction Summary output

Start Time: 2014-04-22 20:48:31.835552

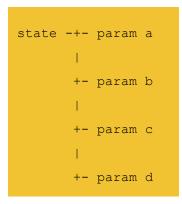
Router	Result	Elapsed Time					
openwrt1	:-)	2.88(sec)					
openwrt3	:-)	2.99(sec)					
openwrt2	:-)	3.00(sec)					
rpi1	:-)	3.08(sec)					

CRUD operations inside NLAN config modules



Global variables (__dict__) generated by CRUD.params()

model



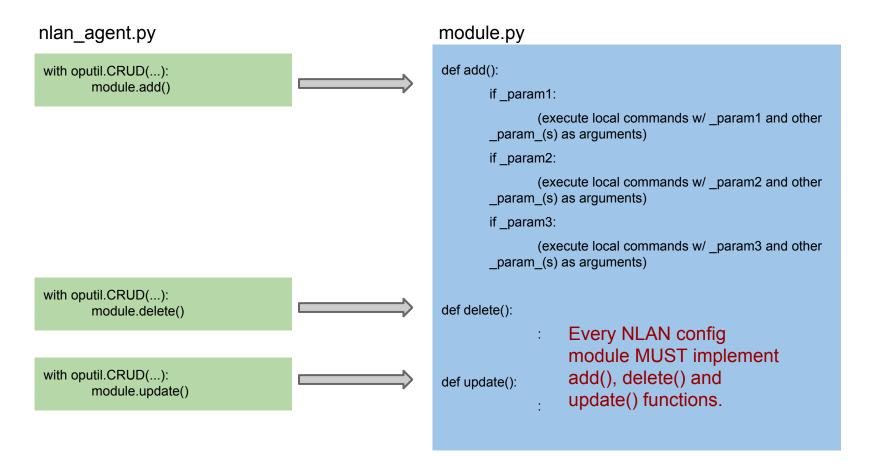
State parameters (for add/update operations)

	ctate parameters (for addrapates operations)					
			Global variables generated by Model.params()			
State params	Requested by Master	OVSDB	_param (Requested change)	_param_ (Desired state)	param_ (Current state in OVSDB)	Operation
а	1	None	_a=1	_a_=1	a_=None	add
b	2	1	_b=2	_b_=2	b_=1	update
С	None	1	_c=None	_c_=1	c_=1	
d	None	None	_d=None	_d_=None	d_=None	

State parameters (for delete operations)

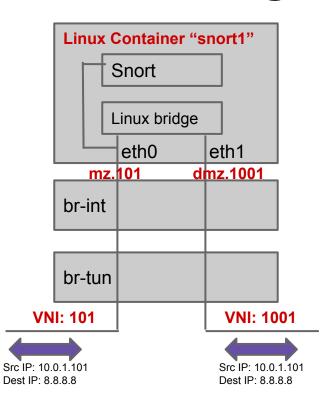
			Global variables generated by Model.params()			
State params	Requested by Master	OVSDB	_param (Requested change)	_param_ (Desired state)	param_ (Current state in OVSDB)	Operation
а	1	None	_a=1	_a_=***	a_=None	(Never exists)
b	1	1	_b=1	_b_=None	b_=1	delete
С	None	1	_c=None	_c_=1	c_=1	
d	None	None	_d=None	_d_=None	d_=None	

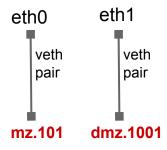
Python conding in NLAN Config modules



Service Function Chaining in YAML

```
rpi1:
  bridges:
   ovs_bridges: enabled
  services: # Service Functions
   - name: snort1
   chain: [mz.101, dmz.1001]
 vxlan:
   local ip: <local ip>
   remote ips: <remote ips>
  subnets:
   - vid: 111
   vni: 1001
   peers: <peers>
   ports: <sfports>
   - vid: 1
   vni: 101
   peers: <peers>
   ports: <sfports>
```





NLAN command usage (nlan.py)

Copy NLAN-agent-side modules to all the target routers (incl. NLAN/OVSDB schema and Linux init.d scripts):

\$ nlan.py --scpmod

Initialize states at all the target routers:

\$ nlan.py init.run

Ask all the target routers to transit to the desired states

\$ nlan.py -G deploy

Rollback to the previous config

\$ nlan.py init.run

\$ nlan.py -R deploy

Command line CRUD (add/get/update/delete) operations

\$ nlan.py -t openwrt1 --add subnets index=101 vid=1 ip dvr=mode:dvr,addr:10.0.1.1/24 ip vhost=10.0.1.101/24

\$ nlan.py -t openwrt1 --update subnets _index=101 ip_vhost=10.0.1.109/24

\$ nlan.py -t openwrt1 --delete subnets _index=101 ip_vhost=10.0.1.109/24

\$ nlan.py -t openwrt1 --get subnets _index=101 vid ip_dvr

Reboot all the target routers:

\$ nlan.py system.reboot

NLAN command usage (nlan_agent.py)

Initialize states:

\$ nlan_agent.py init.run

Command line CRUD (add/get/update/delete) operations

\$ nlan_agent.py --add subnets _index=101 vid=1 ip_dvr=mode:dvr,addr:10.0.1.1/24 ip_vhost=10.0.1.101/24

\$ nlan_agent.py --update subnets _index=101 ip_vhost=10.0.1.109/24

\$ nlan_agent.py --delete subnets _index=101 ip_vhost=10.0.1.109/24

\$ nlan_agent.py --get subnets _index=101 vid ip_dvr

Reboot all the target routers:

\$ nlan.py system.reboot

REST APIs

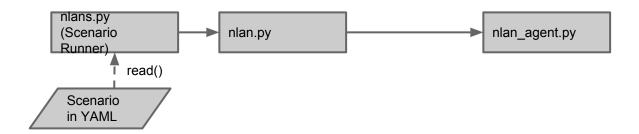
HTTP Method	NLAN CRUD operations	NLAN RPC operations
GET	CRUD: get URL: / <router>/config/<module>/<_index></module></router>	URL: / <router>/rpc/<module>/<func></func></module></router>
POST	CRUD: add URL: / <router>/config/<module>/<_index></module></router>	-
PUT	CRUD: update URL: / <router>/config/<module>/<_index></module></router>	-
DELETE	CRUD: delete URL: / <router>/config/<module>/<_index></module></router>	-
OPTIONS	Get NLAN schemas URL: none	-

REST APIs example

HTTP Method	URL	Query parameters
POST	/_ALL/rpc/test/echo	params=Hello!
OPTIONS	(none)	params=subnets
POST	/openwrt1/rpc/init/run	(none)
POST	/openwrt1/config/bridges	ovs_bridges=enabled
POST	/openwrt1/config/vxlan	local_ip=192.168.1.101&remote_ips=192.168.1.102,192.168.56.103
PUT	/openwrt1/config/vxlan	remote_ips=192.168.1.102,192.168.56.104
GET	/openwrt1/config/vxlan	params=remote_ips
POST	/openwrt1/config/subnets/101	vni=101&vid=1&ip_dvr=addr:10.0.1.1/24,mode:dvr
DELETE	/openwrt1/config/subnets/101	params=ip_dvr
POST	/openwrt1/rpc/db/state	(none)

Scenario Runner

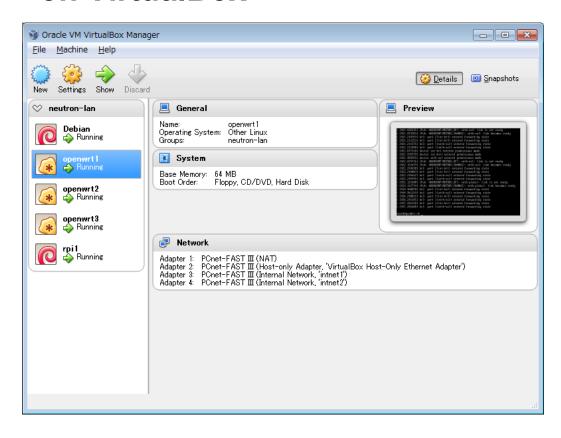
- nlans.py -- reads test scenarios and executes each test
- Test scenarios written in YAML
- Automatic test result confirmation
 - Inspired by Python's "unittest"
 - "assert"
 - "asserRaises"
 - ("assertOutputs" to be supported)



Simple RPC library

```
import rpc
# calls test.kwargs test(...) at router 'openwrt1'
rpc = rpc.RPC(module='test', func='kwargs test', target='openwrt1')
result = rpc(1, b='Hello', d='World!')
print result
# calls test.echo(*args) at all routers on the roster
rpc = rpc.RPC(module='test', func='echo')
results = rpc('Hello World!')
for I in results:
      print [['router'], [['stdout']
```

NLAN Software Development environment on VirtualBox

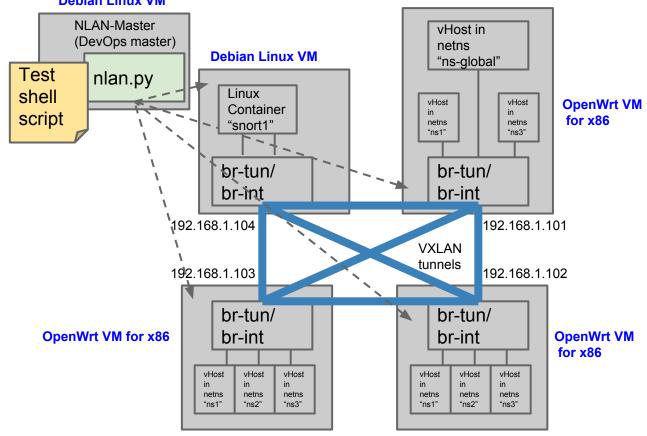


- Five VMs running on one Win7 PC.
 - Two Debian VMs
 - Three OpenWrt VMs
- OpenWrt image for x86
 - I built the kernel with Open vSwtich 2.0.0 and netns/veth/LXC support
 - Very light-weight Linux supporting Open vSwitch 2.0.0 ⇒ An alternative to mininet 2.0.0
- Network adapters setting
 - Internet access: "NAT"
 - Management: "Host-Only"
 - NLAN underlay: "Internal"

Integration Test environment on VirtualBox

(running the test script every day)

Debian Linux VM



- Open vSwitch-based network more realisic than mininet 2.0
 - Every vSwitch with full-fledged(?) Linux
 - netns-based virtual hosts
- Mimics "Beremetal Switch"
- Integration Test scipt running on Debian Linux VM
 - Makes use of "nlan. py"

Future work

- NLAN enhancements
 - NLAN Agent in a Linux container (lxc-excute)
 - Multi-generation config rollback
 - Full-automatic integration testing
- Studying NETCONF and OpenDaylight
 - I have understood that modeling is the core part of SDN and DevOps, and NETCONF supposed to be a perfect language for modeling network.
 - o I want to see if it is really 'an idealized SDN controller' as O'Reilly "SDN" book tells me.

