まずはスイッチを静かに

fae mlxi2c set_fan /FAN1/FAN 1 30 fae mlxi2c set_fan /PS1/FAN 1 30 fae mlxi2c set_fan /PS2/FAN 1 30



NVIDIA。 RIVERMAXとNVIDIAスイッチ を使用した、PIM-SSMと OPENFLOWのハイブリッド 構成デモ

Speaker, Makoto Taguchi Date 2020, 10/20





AGENDA

PIM-SSM

Openflow

Rivermax

ConnectXとPTP同期



SDN? OPENFLOW?

ライブ・放送インフラには大きく分けてIGMPを使ったネットワーク、SDNでコントロール するネットワーク、の2種類があります。

現在、SDNの使用を検討したり、IGMP制御のネットワークとどのように結合するのか、で きるのかという疑問をお持ちの方が多いと思います。

今回のセッションでは、NVIDIA/MellanoxのスイッチSN2410を使用して、IGMPと SDN(Openflow)がハイブリッドで動作する様子をご覧いただき、検討の材料としてお使い いただくために、デモを交えたご紹介を行います。





PIM-SSMCDUT

PIM=Protocol Independent SSM=Source Specific Multicast

L3レイヤーのプロトコル、つまり、スイッチ上のソフトウ ェアスタックがプロトコルを実行している

特徴

-マルチキャスト受信者は、マルチキャストの送信元IPアドレスを指定してIGMPV3 Joinを送信する

-ユニキャストルーティングテーブルを参照して送信元の所在 が特定できる。

-このメカニズムにより、PIM-SMで必要となる待ち合わせ場所 (RP=ランデブーポイント)の必要性が無い

-IGMPv2ではストリームに参加できない



ONYX SOFTWARE ARCHITECTURE



/ 😢 I IVIDIA.



	##
	## PIM conf
コンフィグ	## protocol p interface interface
送信、受信テスト	ip pim vrf
IGMPV3 Join by Rivermax viewer	172.16 vla
IGMPv2 join by Rivermax viewer	1/50



figuration

pim vlan 100 ip pim sparse-mode vlan 200 ip pim sparse-mode f default ssm range standard



設定上の注意

IGMP Snoopinは、端末側だけ有効化する

マルチキャストルーター(PIMを有効にしたノード)間にIGMP snoopingを有効化した場合、Joinを出す端末が存在しないためフ ローが通らない。

リモート間で一方だけPIMを運用する構成IGMP Snoopingは設定しない

右の図のような構成の場合(たまにある)、PIMが動いていない スイッチ側でIGMP Snoopingを動作させてしまうと、取れない フローが出てきたりする。





OPENFLOW

HYBRID DATAPATH



OpenFlow 1.3

📀 NVIDIA.

ハイブリッド動作

初期状態:マルチキャストフローを流しても、show ip mrouteでフローが表示されない。 理由:ベースのルールが設定されていないから。

ベースルール

openflow add-flows 1 priority=1, Actions=NORMAL

Show ip mroute

->フローが表示される

PTPとも共存が可能



ACTIONの種類 (SPECTRUM2以降)

- Permit/Deny
- VLAN manipulations (push, pop, replace)
- QoS (DSCP, PCP, DEI, EXP, ECN, internal priority, color)
- Set egress queue
- Police
- Count
- Mirror
- Trap/Mirror to CPU Set cookie Sampled mirroring
- Set user token / Metadata Manipulate LAG / ECMP hash value
- Set SMAC / DMAC
- Set VRF
- Set L4 ports
- Set/Dec TTL
- Set VNI (Enter VXLAN tunnel)
- MPLS (push, pop, swap)
- Bind next ACL table
- OF 1.4 actions (according to TTPs)
- Set configurable field
- Fields manipulations Copy Add Shift left / right AND / XOR between fields

フロー設定 BY ONYX CLI

ONYX CLIでの設定

openflow add-flows 10 priority=100, ip, dl_vlan=200, nw_dst=232.20.3.3, actions=pop_vlan, push_vlan: 0x8100, set_field: 41 96->vlan_vid,set_field:ec:0d:9a:fa:15:d4->eth_src,output:Eth1/51 アクション マッチ条件

<u>Priority: 数値が高い方が、優先的に使われる。</u>

VLAN設定、L3設定されたインタフェース間でフロー設定を行う場合は、この様に全てのヘッダの 整合性を合わせて出力する必要がある。

仕様:出力側VLANの設定値は、VID+4096という値に設定する必要あり

NAT

openflow add-flows 10

priority=100,ip,dl_vlan=200,nw_dst=232.20.3.3,actions=pop_vlan,push_vlan:0x8100,set_fi eld:4196->vlan_vid,set_field:ec:0d:9a:fa:15:d4->eth_src, set_field:232.20.3.10->ip_dst ,set_field:01:00:5e:16:03:0a->eth_dst, output:Eth1/51



フロー設定 BY OPENFLOWコントローラ(参考)

Openflowコントローラー (**RYU**)を使った、フロー設定

スイッチ側設定 openflow controller-ip x.x.x.x [tcp-port y]

ryu-manager インストールはpip install ryu

ryu-manager ryu.app.ofctl_rest

別端末から

curl -X POST -d '{"dpid": "40175254667200","cookie":"2","table_id":"0", "priority":"100", "match": {"in_port":81,"vlan_vid":200},"actions": [{"type":"POP_VLAN"},{"type":"PUSH_VLAN", "ethertype":" 33024},{"type":"SET_FIELD", "field":"vlan_vid","value":4196},{"type":"SET_FIELD", "field":"eth_src","value":"24:8A:07:C8:80:08"},{"type":"OUTPUT", "port": 83}]}' http://x.x.x.x:8080/stats/flowentry/add

dpidや出力のポート番号はスイッチ上show openflowで取得可能







RIVERMAXとは?

- ST2110だけでなく、あらゆるデータのストリーミングに使用できるライ ブラリ。
- サンプルアプリが付属し、ST2110-20の1080i59.94をはじめ全てのフレームサイズ(2K, 4K, 8K)とレート(59.94, 50, 60, 29.97, 30, ...)と方式(インターレース、プログレッシブ)をサポート。
- ► GPU-Directをサポートしているので、低レイテンシで画像処理や画像解 析を加えたりが可能。
- ► ESXi、KVMなどの仮想プラットフォーム対応
 - Vidmeet onlinでは実際にデーセンターで運用している





Full HD	20bpp	25fps	IPG 11,000 nSec	Can be achieved with software packet pacing
4K UHD	20bpp	60fps	IPG 1,150 nSec	Cannot be achieved with Software packet pacing
4K UHD	24bpp	120fps	IPG 469 nSec	Cannot be achieved with Software packet pacing
8K UHD	36bpp	60fps	IPG 156 nSec	Cannot be achieved with Software packet pacing



PTP WITH CONNECTX-5

CONNECTX NICによるHWタイムスタンピング

LINUXPTPとConnectXに搭載されているHWタイムスタンピング機能 使って、時刻を高精度に同期させることが可能。

1PPSの入出力が可能なNICをリリース予定

Linuxptpを使った、2ノード間の同期実験







