

In-band Telemetry

ーデータプレーン遅延時間観測ー

Shishio Tsuchiya

shtsuchi@arista.com

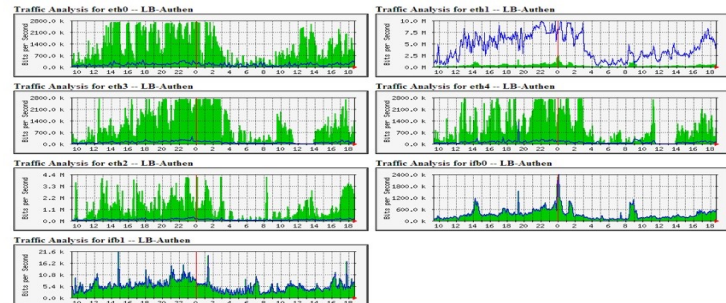
Agenda

- テレメトリーおさらい
- テレメトリーの実装状況
- これからのテレメトリー in-band Telemetry

従来のネットワーク運用の問題点

1. SNMPだと5分間隔ぐらいでしかデータを確認できない・・

MRTG Index Page



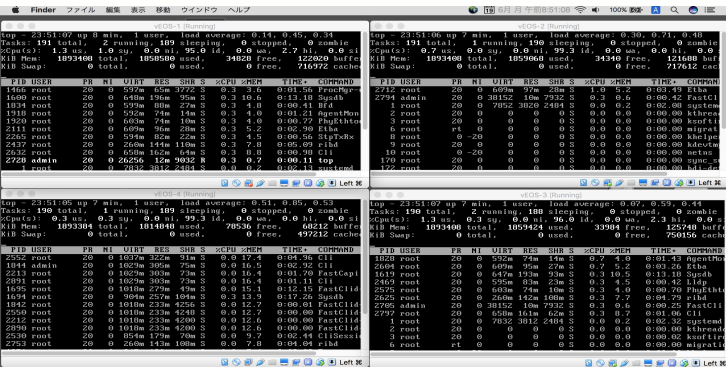
2. Syslogの確認、装置が複数に及ぶとログも膨大に・・

```

Feb 28 20:41:19 KS083-VM01r ProcMgr-worker: NPROCGR-6-PROCESS_STARTED: 'PowerSupplyDetector' starting with PID=1864 (PID=1439) --
executing '/usr/bin/PowerSupplyDetector'
Feb 28 20:41:19 KS083-VM01r ProcMgr-worker: NPROCGR-7-WORKER_WARMSTART_DONE: ProcMgr worker warm start done. (PID=1439)
Feb 28 20:41:19 KS083-VM01r Launcher: 39: NLAUNCHER-6-PROCESS_START: Configuring process 'FanDetector' to start in role
'AllSupervisors'
Feb 28 20:41:19 KS083-VM01r Launcher: 40: NLAUNCHER-6-PROCESS_START: Configuring process 'Xcvt' to start in role 'AllCells'
Feb 28 20:41:19 KS083-VM01r Launcher: 41: NLAUNCHER-6-PROCESS_WARMSTART: Initiating warm start of 'ProcMgr (worker)'
Feb 28 20:41:20 KS083-VM01r ProcMgr-worker: NPROCGR-6-WORKER_WARMSTART: ProcMgr worker warm start. (PID=1439)
Feb 28 20:41:20 KS083-VM01r Launcher: 42: NLAUNCHER-6-PROCESS_START: Configuring process 'Gn245' to start in role 'AllCells'
Feb 28 20:41:20 KS083-VM01r Launcher: 43: NLAUNCHER-6-PROCESS_START: Configuring process 'StrataAgent' to start in role 'AllCells'
Feb 28 20:41:20 KS083-VM01r ProcMgr-worker: NPROCGR-7-NEW_PROCESSES: New processes configured to run under ProcMgr control:
['FanDetector', 'Ln73', 'Max6558', 'Xcvt']
Feb 28 20:41:20 KS083-VM01r ProcMgr-worker: NPROCGR-7-PROCESSES_ADOPTE: ProcMgr (PID=1439) adopted running processes: (MlagTunnel,
PID=1472) (Mirroring, PID=1584) (Lag, PID=1474) (Acl, PID=1476) (PortSec, PID=1477) (PhyNetool, PID=1863) (Sdp, PID=1480) (TopoAgent,
PID=1482) (LedPolicy, PID=1485) (Arp, PID=1486) (Ucdm82, PID=1444) (Mlag, PID=1488) (Dns, PID=1489) (Ebr, PID=1587) (Thermostat,
PID=1494) (EventMon, PID=1496) (Rlb, PID=1498) (AgentMonitor, PID=1500) (Sysdb, PID=1448) (FastCid, PID=1441) (Fru, PID=1442)
(Launcher, PID=1445) (Mdp, PID=1588) (IpmSnmping, PID=1534) (Smbus, PID=1745) (Mdio, PID=1746) (Sdp, PID=1747) (PcBus, PID=1652)
(Sb828, PID=1653) (Lldp, PID=1465) (SuperServer, PID=1466) (PowerManager, PID=1467) (StpTopology, PID=1468) (Aaa, PID=1469)
Feb 28 20:41:19 KS083-VM01r ProcMgr-worker: NPROCGR-6-PROCESS_STARTED: 'PhyNetool' starting with PID=1863 (PID=1439) -- executing '/
usr/bin/PhyNetool'
Feb 28 20:41:20 KS083-VM01r ProcMgr-worker: NPROCGR-6-PROCESS_STARTED: 'FanDetector' starting with PID=1863 (PID=1439) -- executing '/
usr/bin/FanDetector'
Feb 28 20:41:19 KS083-VM01r ProcMgr-worker: NPROCGR-6-PROCESS_STARTED: 'Max6558' starting with PID=1882 (PID=1439) -- executing '/usr/
bin/Max6558'
Feb 28 20:41:20 KS083-VM01r ProcMgr-worker: NPROCGR-6-PROCESS_STARTED: 'Xcvt' starting with PID=1884 (PID=1439) -- executing '/usr/bin/
XcvtAgent'
Feb 28 20:41:20 KS083-VM01r Launcher: 44: NLAUNCHER-6-PROCESS_WARMSTART: Initiating warm start of 'ProcMgr (worker)'
Feb 28 20:41:20 KS083-VM01r ProcMgr-worker: NPROCGR-7-WORKER_WARMSTART_DONE: ProcMgr worker warm start done. (PID=1439)
Feb 28 20:41:20 KS083-VM01r ProcMgr-worker: NPROCGR-6-PROCESS_STARTED: 'Ln73' starting with PID=1886 (PID=1439) -- executing '/usr/bin/
Ln73'

```

3. CLI画面たくさん同時に開き、CPU使用率などを比較・・

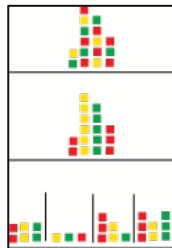
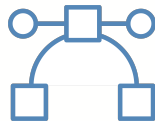


4. トラフィックフローの詳細を見るためにはflowコレクタが必要



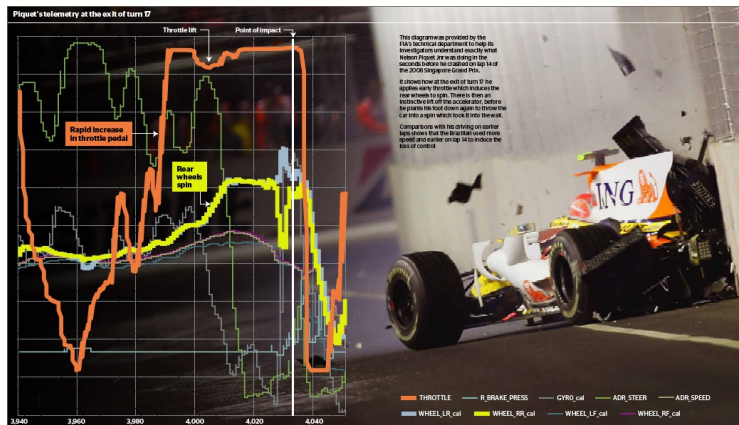
従来のネットワーク運用の問題点

- 精度が悪い/頻度が少ない
- 情報量が多すぎて整理ができない
- 全てのフォーマットが違う(SNMP/Syslog/xFlow/CLI)
 - 全てトランスポートが違う
 - 異なるタイムスタンプ
 - 解析ツールへの変換が難しい



テレメトリーとは

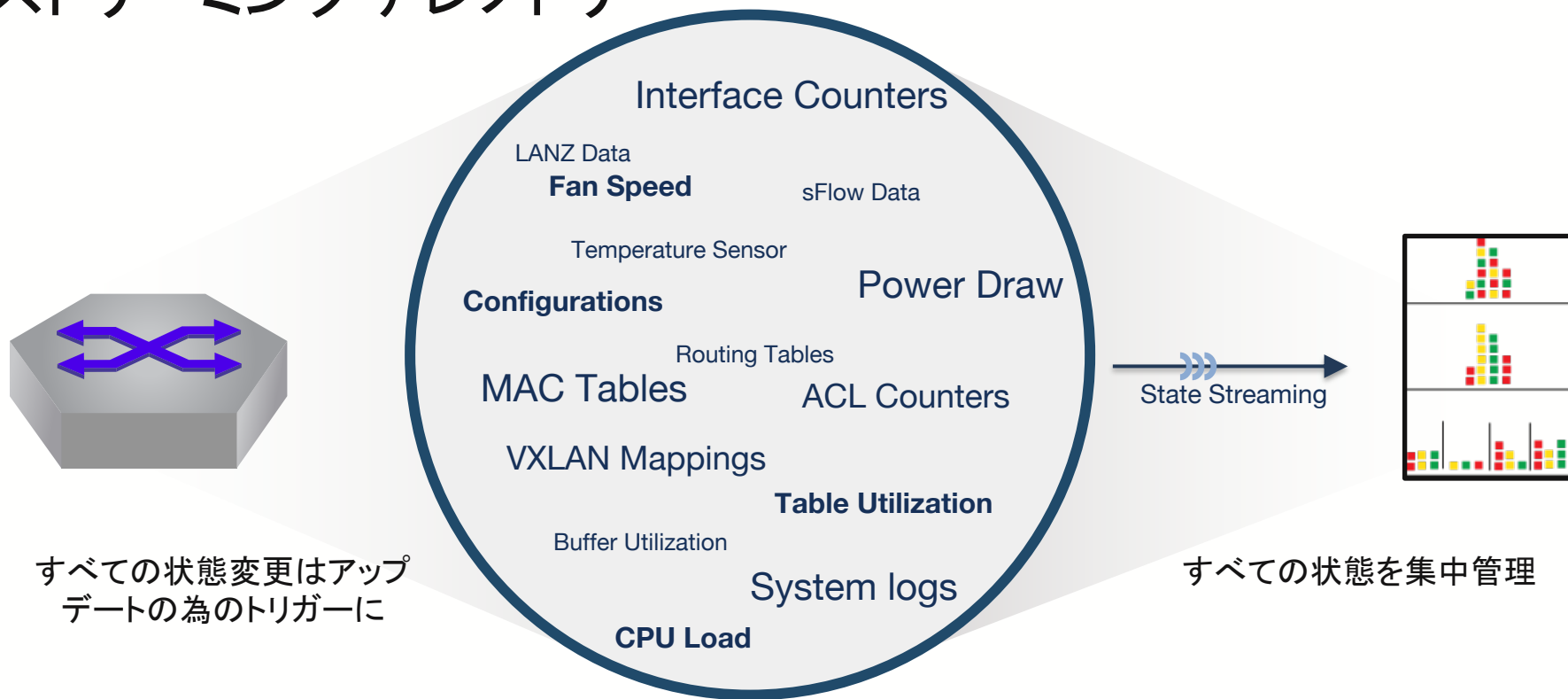
- テレメーター(遠隔計測装置)を使って、遠隔地の測定結果をコントロールセンターに送信すること。ガスや電気のメーター、自動販売機の売り上げ管理システムなどに用いられる。テレメトリング。テレメータリング。遠隔測定法。*デジタル大辞泉より



【例】F1でのテレメトリーデータ

ある選手が、レースでクラッシュする数秒前の行為を正確に理解するためにFIA技術部門が提供したもの。複数のデータを相関関係で分析。

ストリーミングテレメトリ



すべての状態変化をあらゆるデバイスから即座に

テレメトリーの標準化

- 加工がしやすいデータモデル

- YANGフォーマットで定義

- PublisherとSubscriberモデル

- トランスポート

- NETCONF/RESTCONF
- gNMI(gRPC Network Management Interface)

- 標準化

- IETF
- OpenConfig

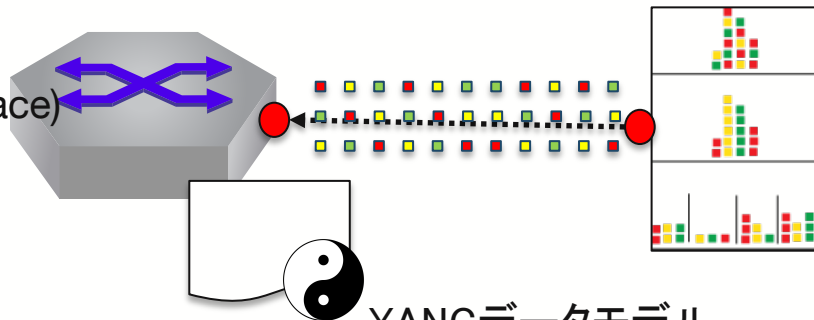
- Google/AT&T/Microsoft/British Telecom/Facebook/Comcast/Level 3/Cox Communications/Yahoo!/Apple/Jive Communications/Deutsche Telekom/Bell Canada/SK Telecom/Bloomberg/Netflix/Cloudflare/Oracle/Tencent/Baidu/Alibaba/Telefonica/GoDaddy/Viasat/LinkedIn/Digital Ocean



I E T F®



netconf/restconf/gNMI



YANGデータモデル

Agenda

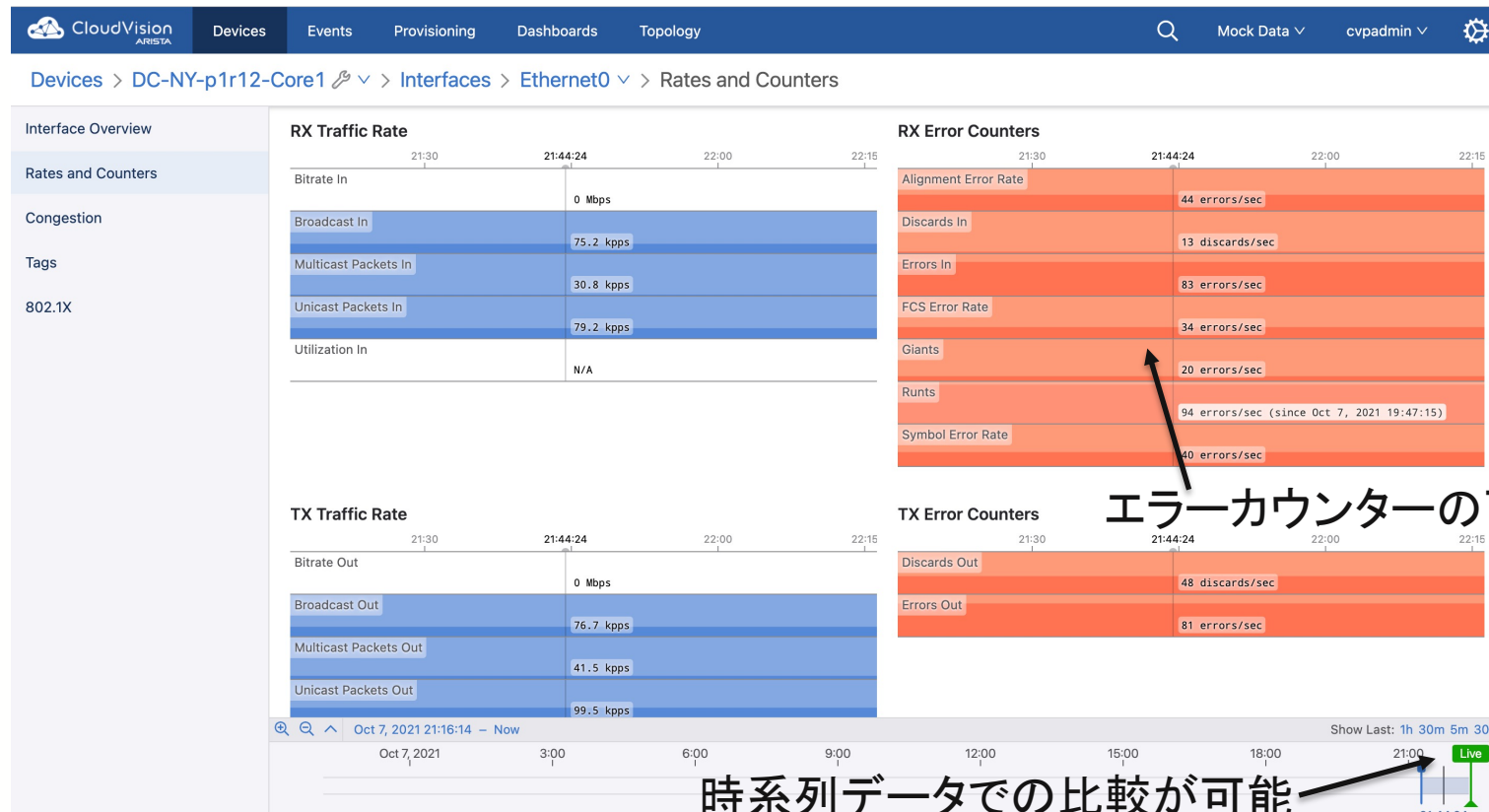
- テレメトリーおさらい
- **テレメトリーの実装状況**
- これからのテレメトリー in-band Telemetry

テレメトリーを利用したネットワークの見える化

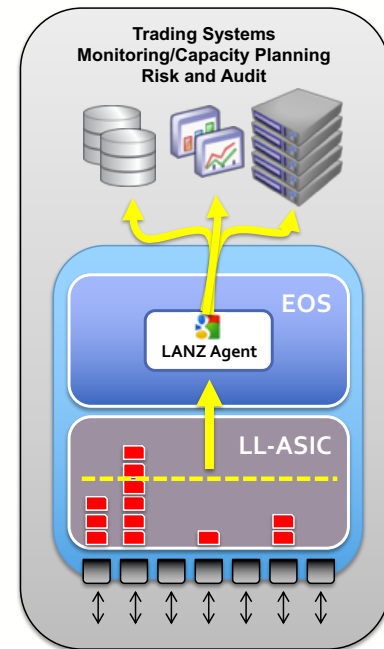
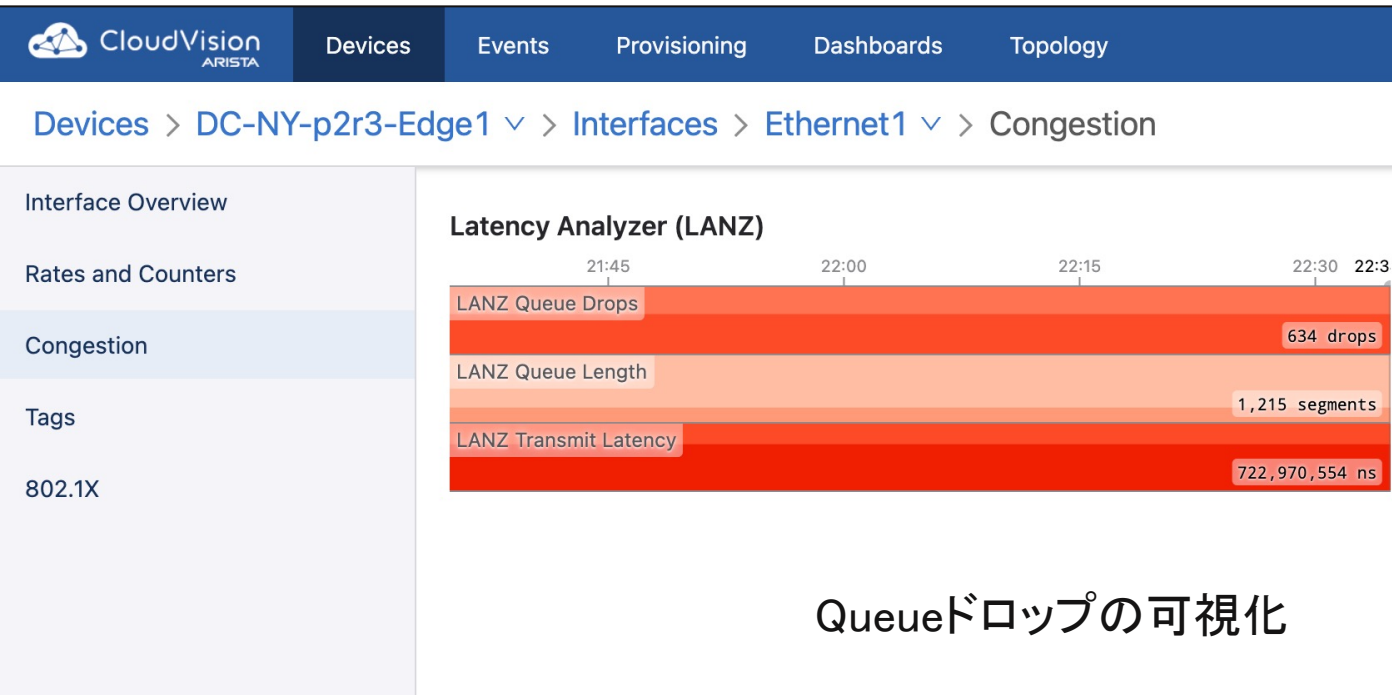


イベントのcorrelation(相関付)を実施
Informational/Warning/Errors/Criticalに分類

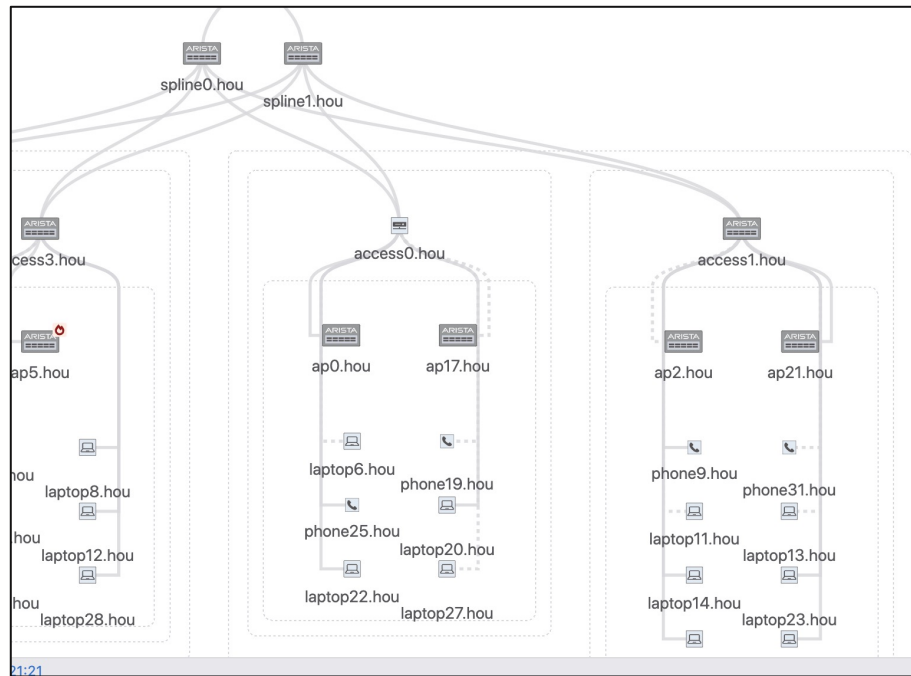
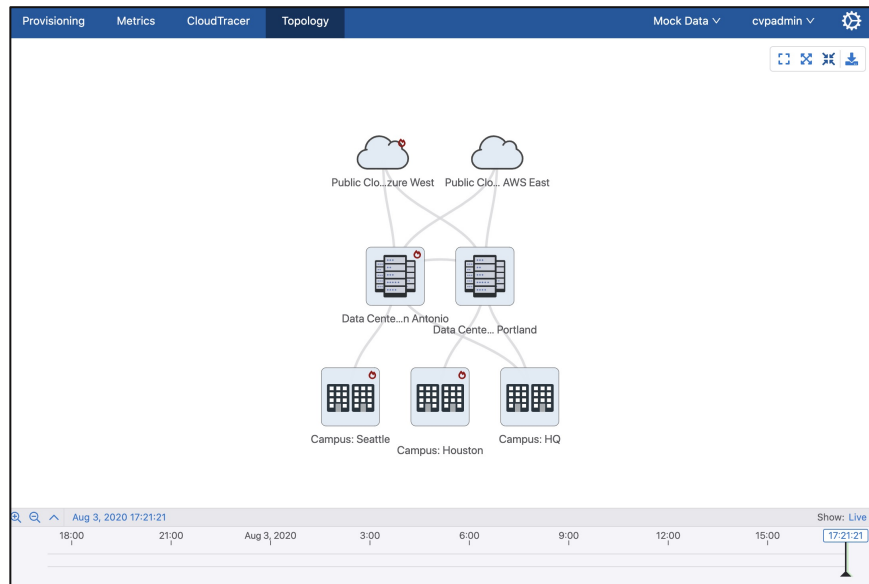
テレメトリの効果



LANZ(Latency Analyzer)

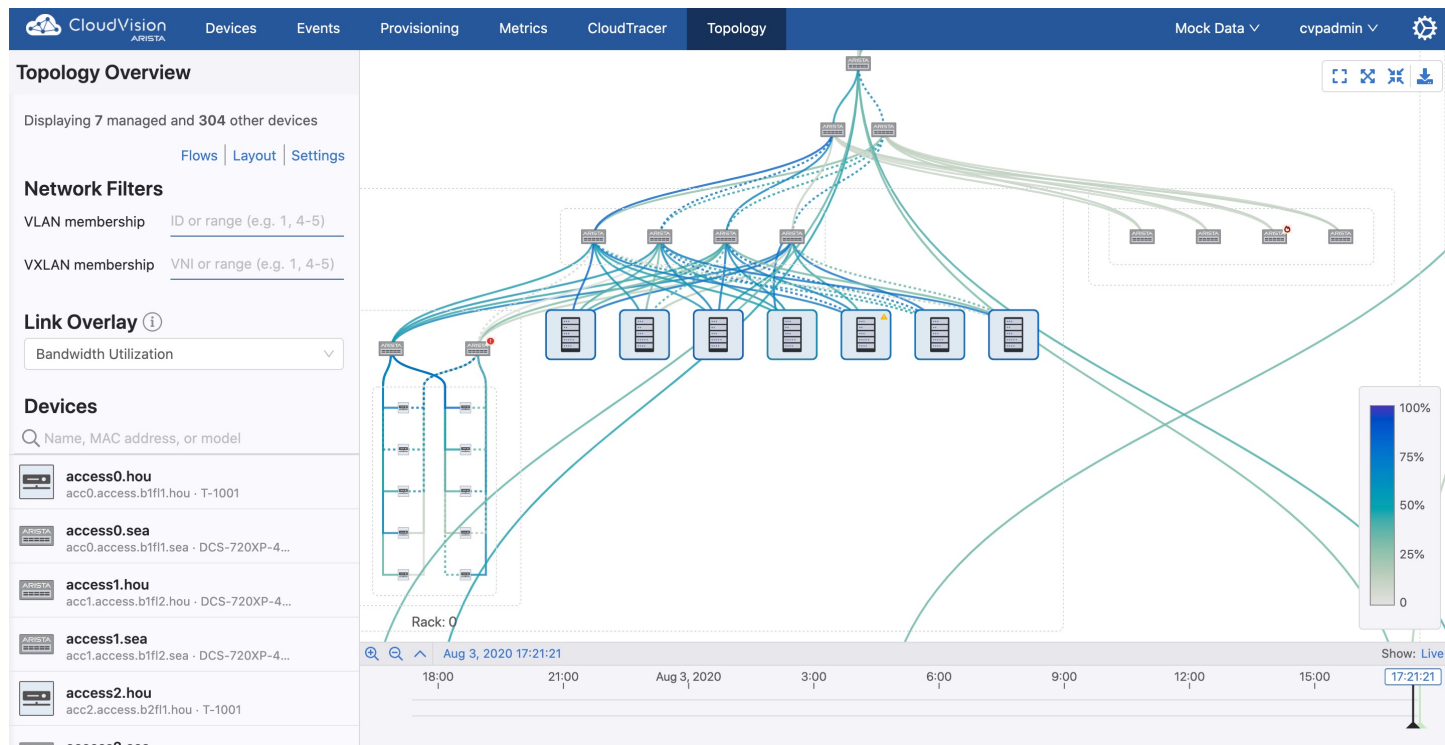


トポロジーマップと統合



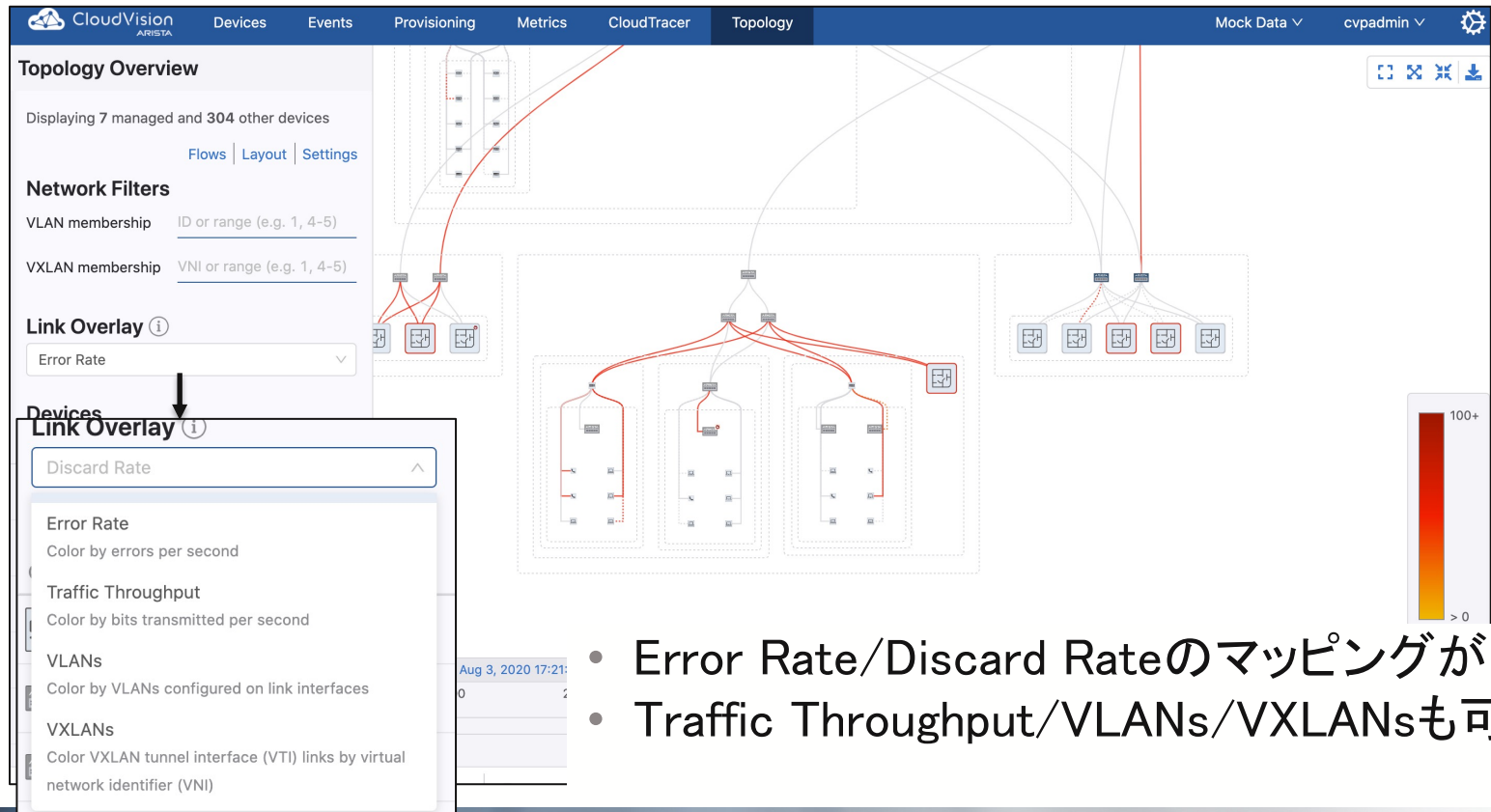
- LLDPの隣接情報を元にトポロジーを自動作成
- 標準プロトコルLLDPをサポートしていればタイプなどで判断し、トポロジーマップに登録される
- タイムスケールを持ち、時系列変更にも対応可能

トポロジーと使用帯域



- トポロジーと使用帯域のマッピングが可能
- ネットワークウェザーマップ同様な使い方が可能

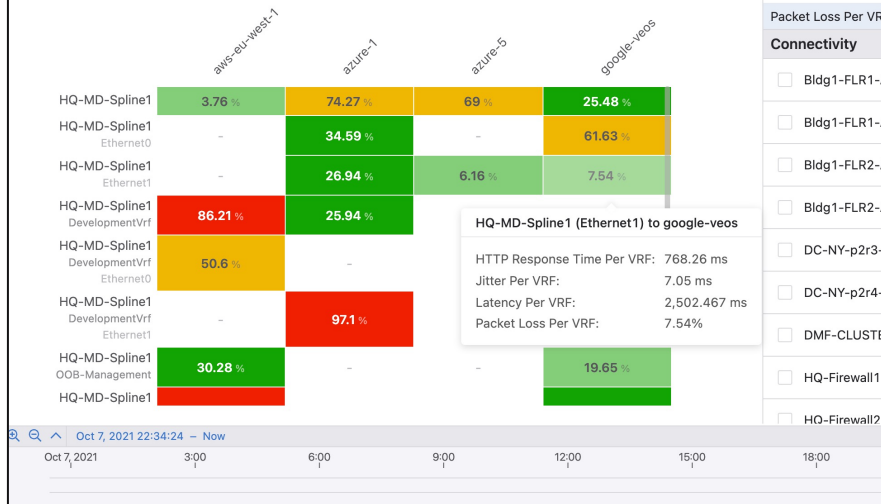
トポロジーとエラーレート/廃棄レート



- Error Rate/Discard Rateのマッピングが可能
- Traffic Throughput/VLANs/VXLANsも可能

Probeによるネットワーク遅延の可視化(HTTP/ICMP)

Viewing Packet Loss Per VRF for 19 connections



Metric

HTTP Response Time Per VRF

Jitter Per VRF

Latency Per VRF

Packet Loss Per VRF

Connectivity

☐ Bldg1-FLR1-AP1

☐ Bldg1-FLR1-AP2

☐ Bldg1-FLR2-AP1

☐ Bldg1-FLR2-AP2

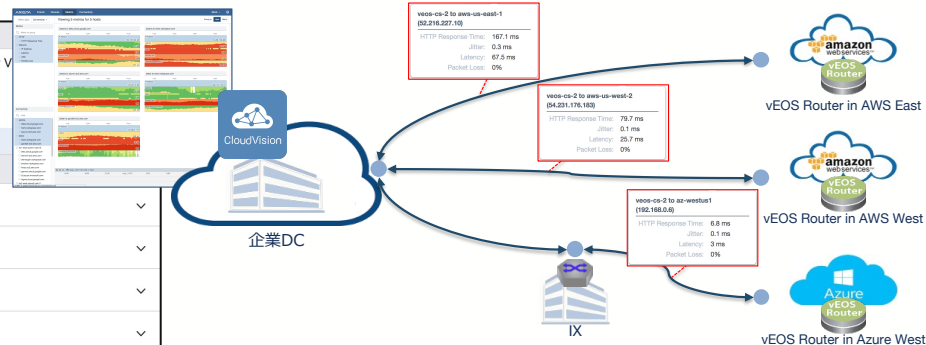
☐ DC-NY-p2r3

☐ DC-NY-p2r4

☐ DMF-CLUSTER

☐ HQ-Firewall1

☐ HQ-Firewall2



Viewing HTTP Response Time Per VRF for 19 connections



Metric

HTTP Response Time Per VRF

Jitter Per VRF

Latency Per VRF

Packet Loss Per VRF

Connectivity

☐ Bldg1-FLR1-AP1

☐ Bldg1-FLR1-AP2

☐ Bldg1-FLR2-AP1

☐ Bldg1-FLR2-AP2

☐ DC-NY-p2r3-Edge1

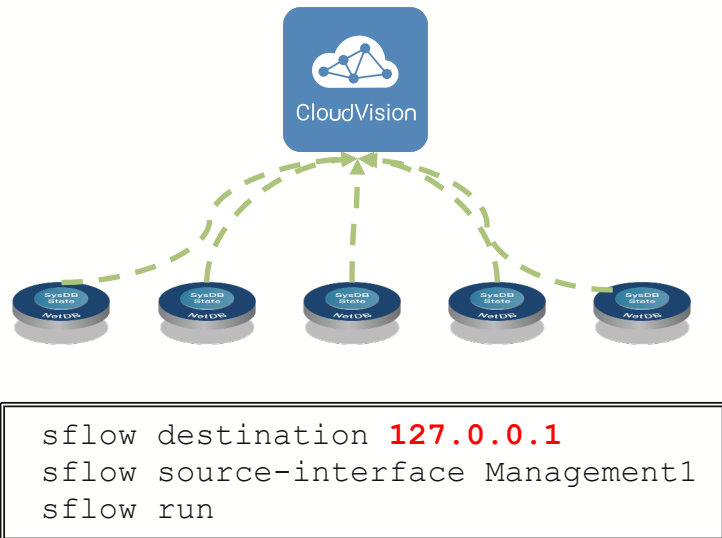
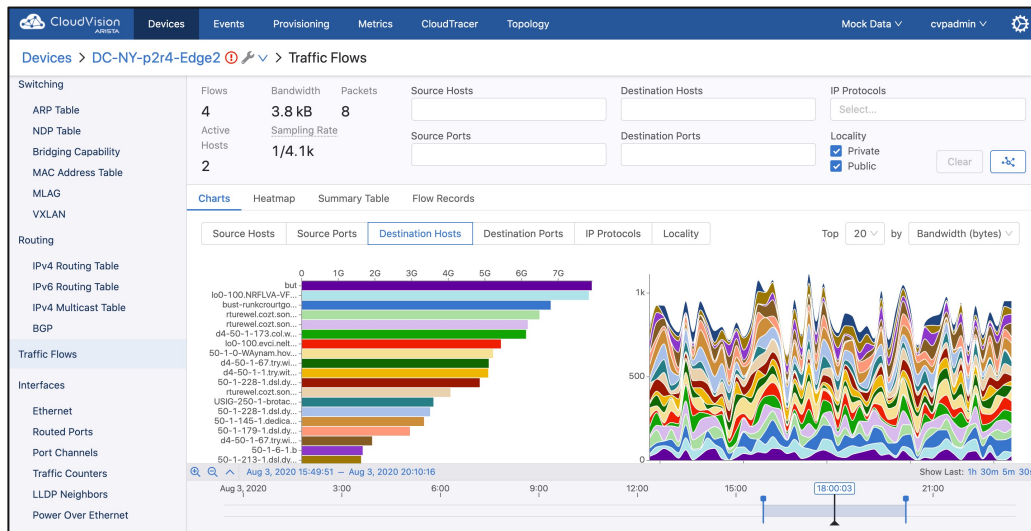
☐ DC-NY-p2r4-Edge2

☐ DMF-CLUSTER

☐ HQ-Firewall1

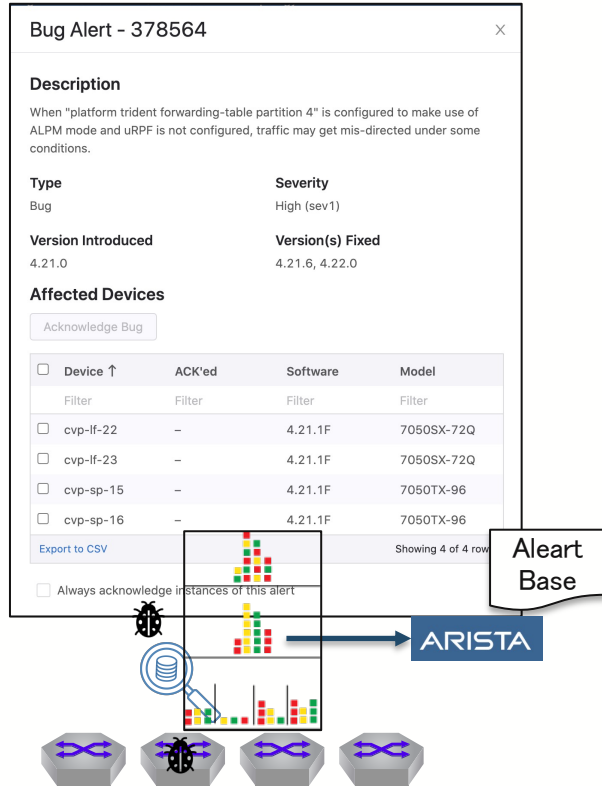
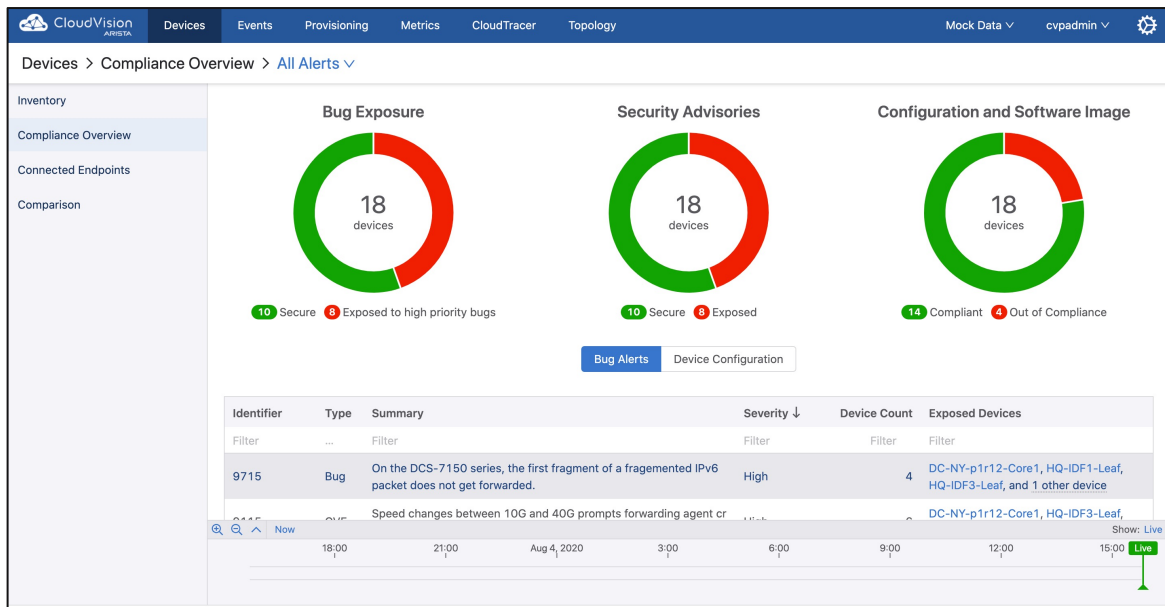
☐ HQ-Firewall2

フローデータをtelemetryで送る



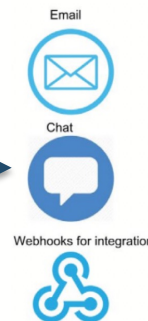
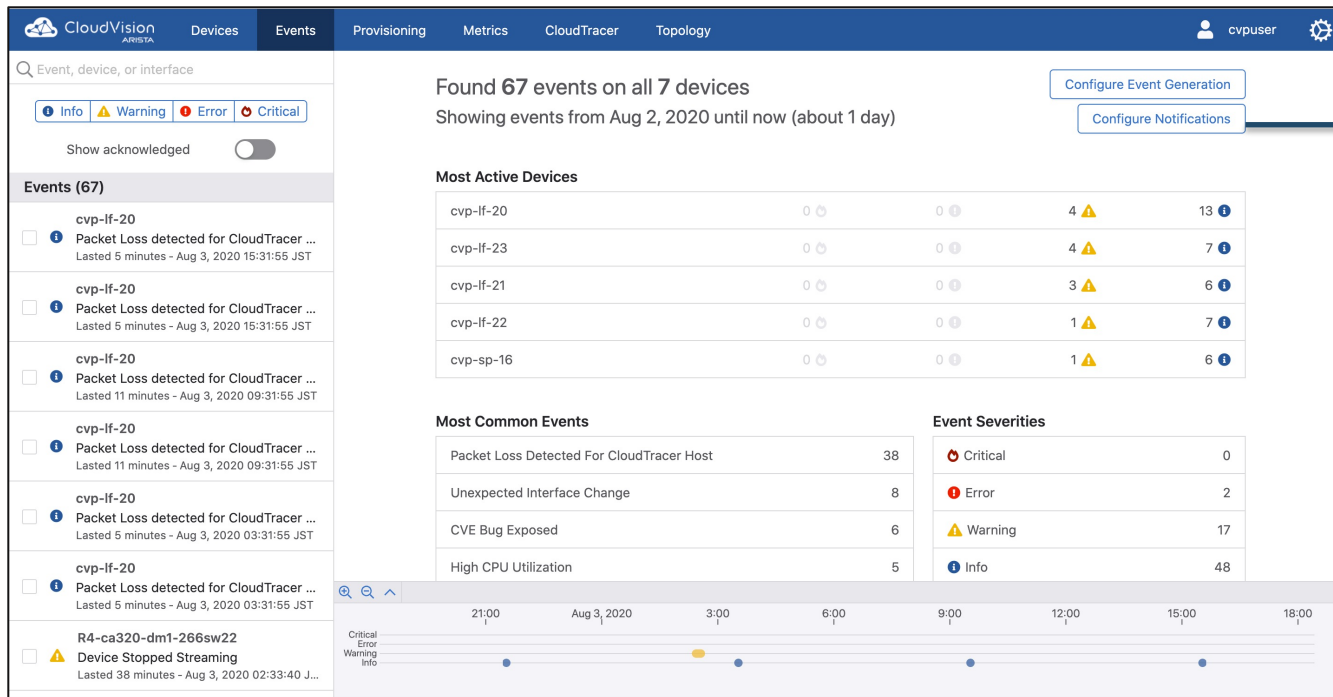
- sFlow/IPFIXでExportされたトラフィックフローデータを可視化
- フローをローカルに転送、フロー情報をテレメトリデータとして変換し、コレクタに送る

リスクアナリシス



- 脆弱性やバグ情報が記載されたベンダーのデータベースの情報を元にテレメトリデータのコンフィグやバージョン情報と比較を実施し、CVPが管理しているネットワークでの影響のある不具合を調査する

外部システムとの連携



- Event情報をメール/SlackEmail/Slack/VictorOps/OpsGenie/PagerDuty/WeChatなどのツールに通知可能

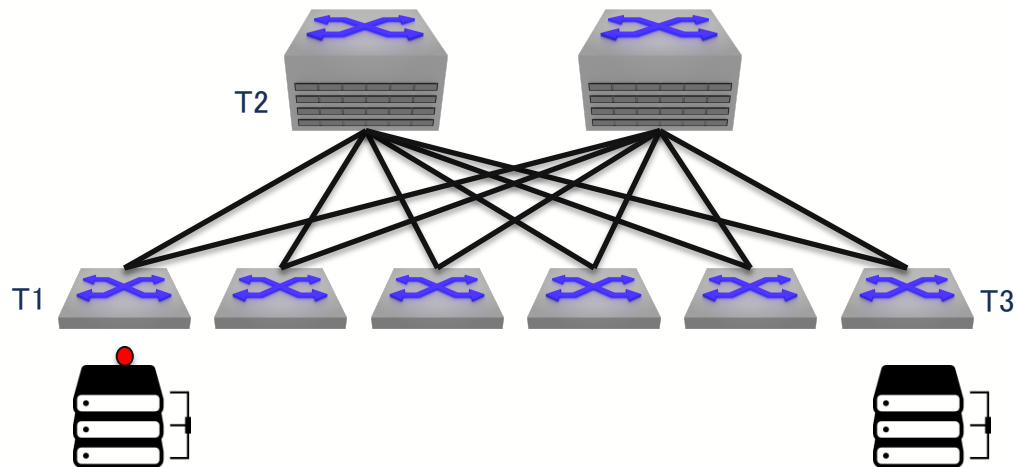
テレメトリーで出来ている事

- インターフェースカウンターの可視化
- エラーパケットなどの可視化
- Queueの輻輳や廃棄状態の可視化
- プローブによるパケット到達性遅延の可視化
- ネットワーク機器を通るフローの可視化
- テレメトリーデーターを使った予兆検知/外部連携

Agenda

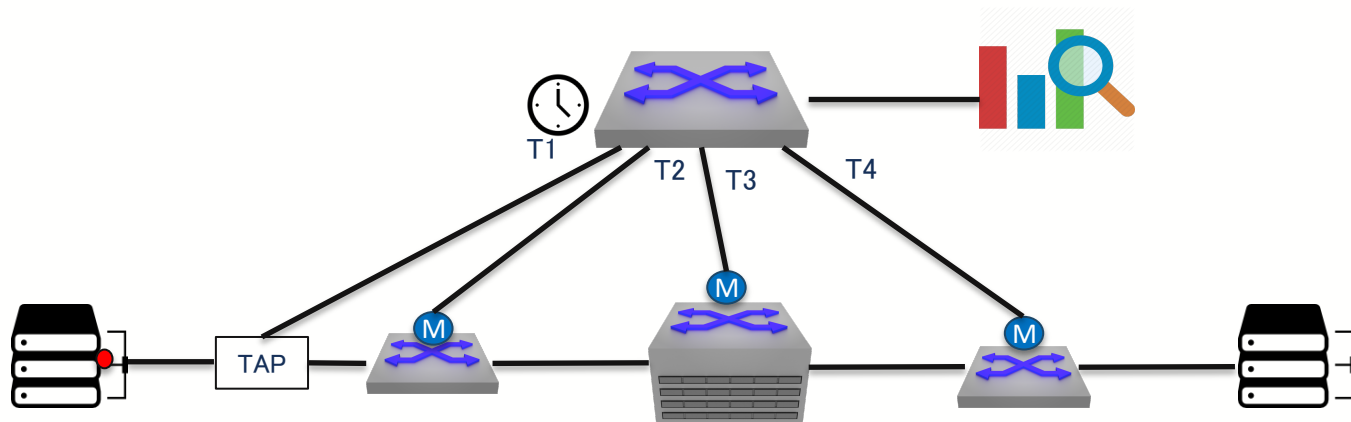
- テレメトリーおさらい
- テレメトリーの実装状況
- **これからのテレメトリー in-band Telemetry**

出来ない事



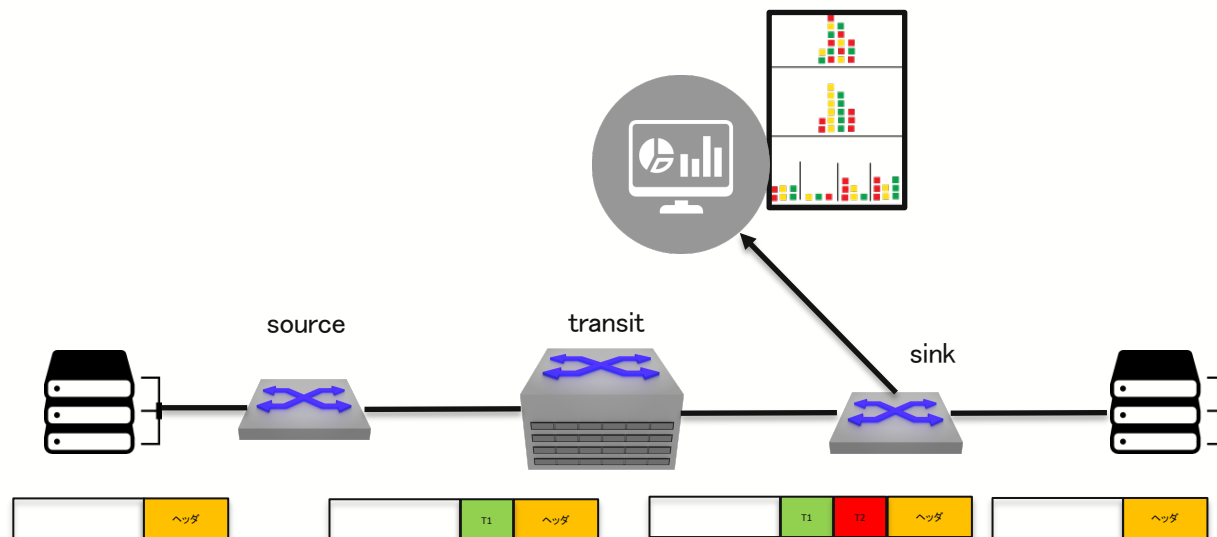
- 実際にネットワークを通過するパケット/フローの通過時間(T1,T2,T3)の可視化
- In-band Network Telemetry

今まであったソリューション



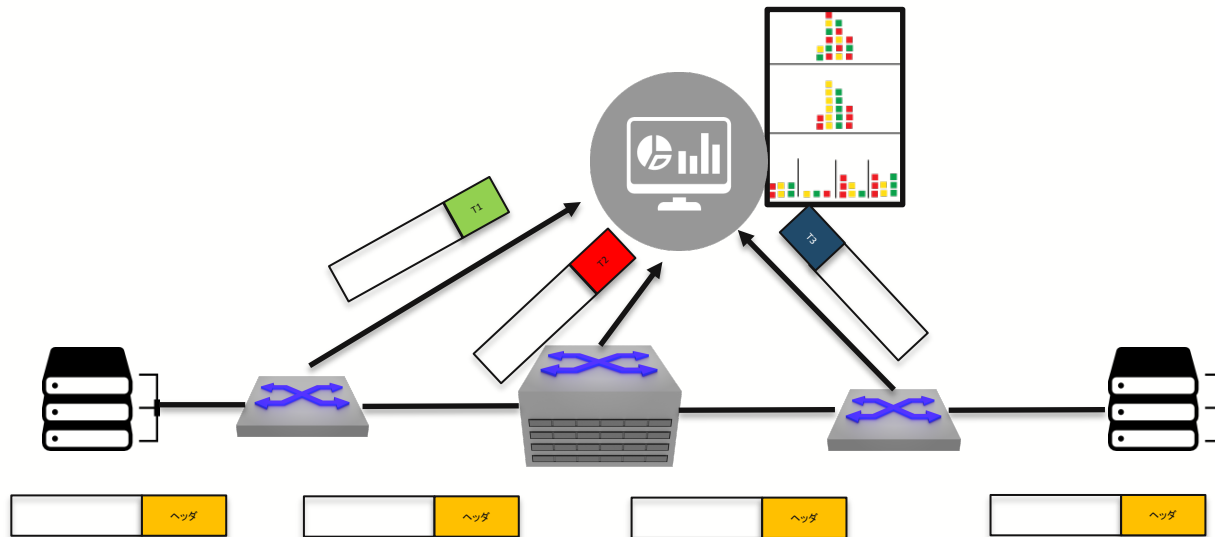
- TAPやミラーリングしたパケットにTAPアグリゲーション機器でタイムスタンプを付与
- ツールポートにてパケットをキャプチャし、時刻情報を読み込む
- out-bandモニタリング

In-band Telemetry



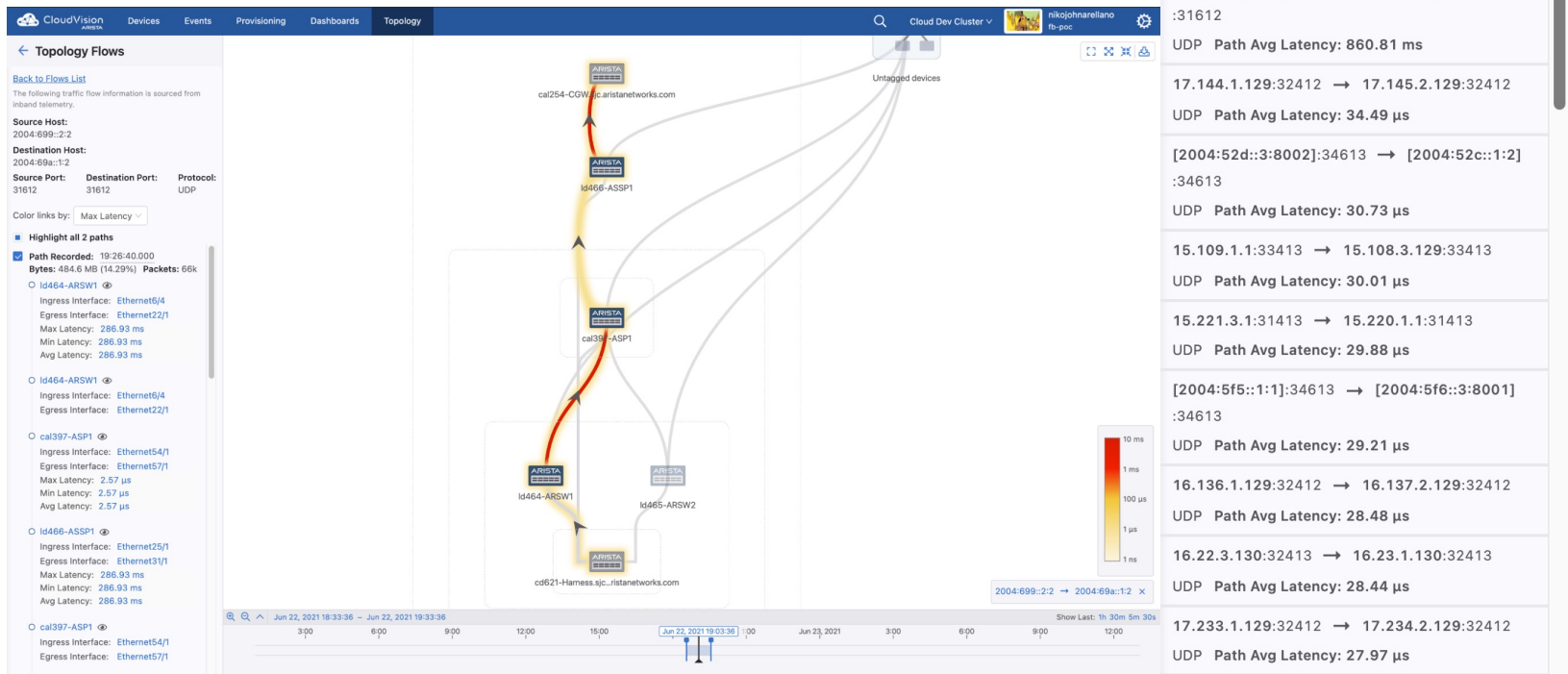
- パスポートモードとも呼ばれる。
- 各ノード毎にフロー/Queue/Drop情報を付与、取り外すノード(sink)でレポートする
- P4.orgで定義されたeXport Data (INT-XD)を解析可能なIntel Deep Insight Network Analyticsソフトウェアなどで解析する
- Inband Network Telemetry (INT) Support on 7170
 - <https://eos.arista.com/eos-4-26-0f/inband-network-telemetry-int-support-on-7170/>
- Inband Telemetry
 - <https://eos.arista.com/eos-4-26-2f/inband-telemetry/>

Postcard Telemetry



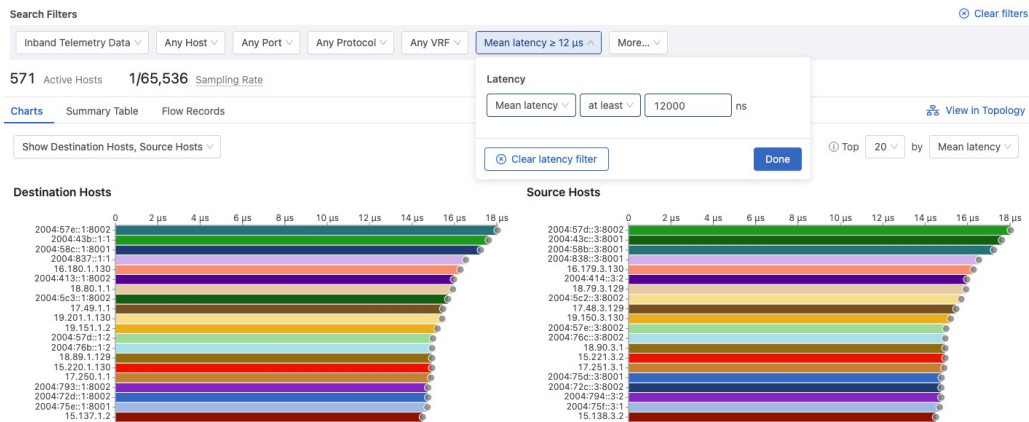
- ポストカードモード。
- 各ノードがそれぞれフロー/Queue/Drop情報をレポートする
- Postcard Telemetry
 - <https://eos.arista.com/eos-4-26-2f/inband-telemetry/>

トポロジーとIn-band Telemetryデータの統合



- データプレーンの遅延時間とトポロジーをマッピングし可視化

トラフィックフローとIn-band telemetryのマッピング



- 特定のフローでの
遅延時間がわかる

Charts

Summary Table

Flow Records

View in Topology

Hiding 4 columns

Top20byMean latency

Start Time	End Time	Source Host	Destination Host	Hops	Mean Latency	Bytes	Packets			
Jul 22, 2021 16:26:28	Jul 22, 2021 16:26:28	2004:57d::3:8002	2004:57e::1:8002	Id465-ARSW2, cal39...	24.48 μs	577.1 MB	65.5k			
Jul 22, 2021 16:33:11	Jul 22, 2021 16:33:11	2004:58b::3:8001	2004:58c::1:8001	Id465-ARSW2, cal39...	23.44 μs	542.0 MB	65.5k			
Jul 22, 2021 17:08:54	Jul 22, 2021 17:08:54	2004:43c::3:8001	2004:43b::1:1	Id465-ARSW2, cal39...	23.04 μs	420.5 MB	65.5k			
Jul 22, 2021 17:51:47	Jul 22, 2021 17:51:47	18.79.3.129	18.80.1.1	Latency RSW2, cal39...	22.57 μs	533.8 MB	65.5k			
Jul 22, 2021 17:38:06	Jul 22, 2021 17:38:06	Id465-ARSW2	Ethernet7/3	Ethernet23/1	64	15.71 μs	15.71 μs	22.49 μs	542.8 MB	65.5k
Jul 22, 2021 16:35:33	Jul 22, 2021 16:35:33	cal397-ASP1	Ethernet55/1	Ethernet58/1	63	2.04 μs	2.04 μs	2.04 μs	540.7 MB	65.5k
Jul 22, 2021 16:58:54	Jul 22, 2021 16:58:54	Id466-ASSP1	Ethernet26/1	Ethernet31/1	63	3.1 μs	3.1 μs	3.1 μs	547.3 MB	65.5k
Jul 22, 2021 16:58:54	Jul 22, 2021 16:58:54	Id466-ASSP1	Ethernet31/1	Ethernet26/1	61	1.01 μs	1.01 μs	1.01 μs	547.3 MB	65.5k
Jul 22, 2021 17:58:27	Jul 22, 2021 17:58:27	cal397-ASP1	Ethernet58/1	Ethernet56/1	60	—	—	—	509.9 MB	65.5k
Jul 22, 2021 17:57:28	Jul 22, 2021 17:57:28	2004:5c2::3:8002	2004:5c3::1:8002	Id465-ARSW2, cal39...	21.86 μs	538.1 MB	65.5k			

DeviceIngress InterfaceEgress InterfaceTTLMinMeanMax

Export to CSV

Showing 1 to 10 of 20 rows

まとめ

- P4のキラーアプリケーションの一つであるIn-band Telemetryの実装も進んできた
- In-band Telemetryの手法もそれぞれ (Postcard/Passport)
- データプレーンの遅延時間の可視化も可能になり、よりトラブルシューティングが容易となる。

議論したい事

- まだ実運用で足りない事ある？
 - 機器の対応=>気持ちはわかるけどとりあえず抜きで。。
- 実際に使いたいケースある？
 - 想定アプリケーション/顧客
 - 過去に困った事
- Postcard/Passportどちらが良い？どちらでも良くない？



Thank You

www.arista.com