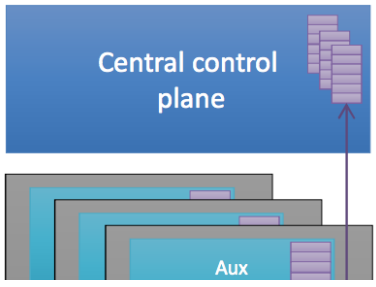
Next- Generation Telemetry Architecture

fine-grain, high-fidelity Telemetry for network elements

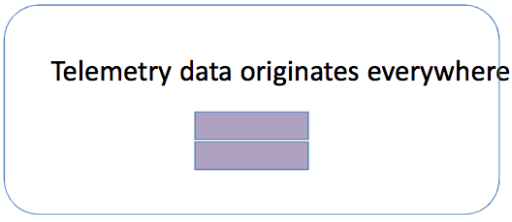
Rafal Jan Szarecki, Solution Architect, Routing, Juniper

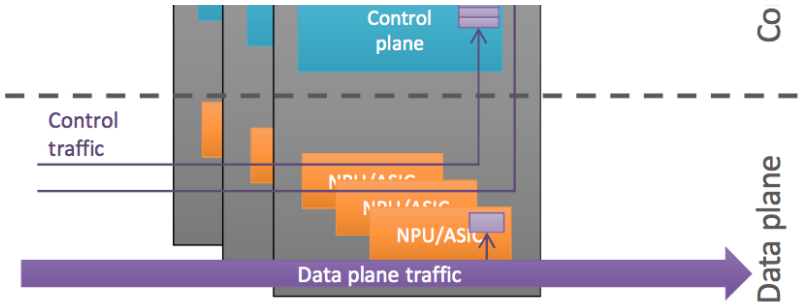
1

Network node model



ntrol plane





2

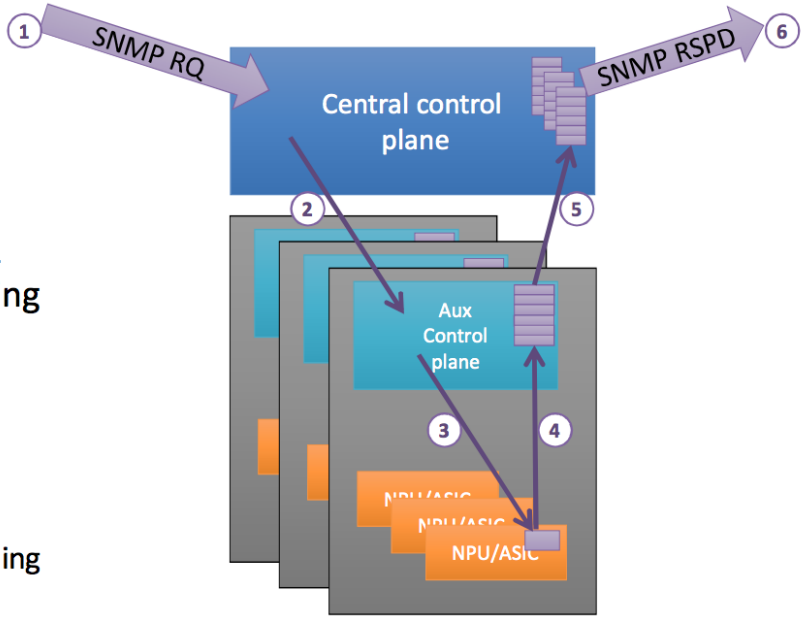
Traditional telemetry in packet network

- Data is pulled periodically by centralized system
- Provides “averaged” view of system’s over longer period of time
  - 5-30 minutes or longer intervals (5 minutes is “common standard”)
  - Obscure dynamic changes of states.
- Reasonable for long-term planning and optimization [months/years cycles]
- Known count of well known counters collected. Semantics' of counters is well known.

3

Traditional telemetry architecture and implementation

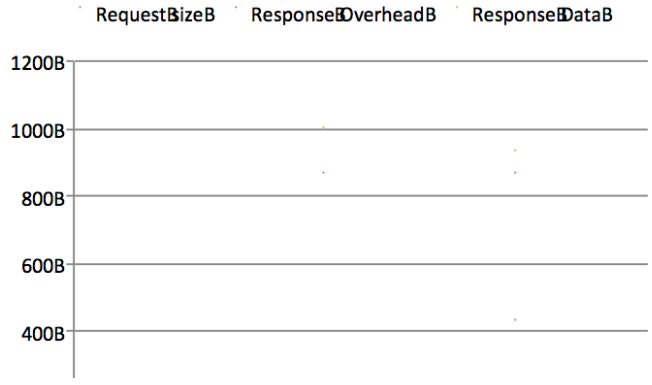
- SNMP is THE protocol
  - MIB data structures:
    - Not everything is a tree
    - Expensive translation
    - Complex modeling in SMI/ASN.1
  - inefficient and complex encoding
- Most data resides in data plane.
  - Nested pulling, or
  - Periodical update of Central control plane
    - Scalability (all possible data)
    - Internal interval put limit on polling interval.



4

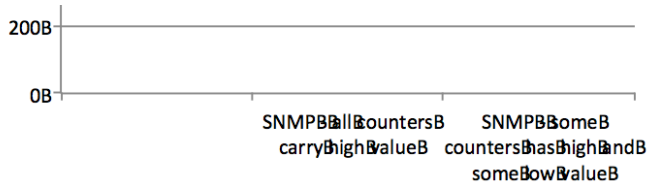
SNMP inefficiency - example

- Port level stats
  - ifIndex is dynamic
  - IF table lookup for name ifIndex
  - IF table for errors and traffic
  - IF CoS table
- Overhead in response
  - 4-8 times more data descriptions then actual data.
  - Often same description repeated



multiple times.

- Request == response's overhead
- SNMP manager to correlate data from multiple tables.



5

Optimization – driver for Telemetry

- IP Networking industry entering adulthood
  - Not as sexy as it was
  - Not as fast developing as it was
  - Its going to be more solid, reliable and efficient.
- We heading optimization phase of industry
  - Every optimization starts with data collection
  - More data – more optimization opportunities and higher optimization degree.

6

Next Generation Telemetry

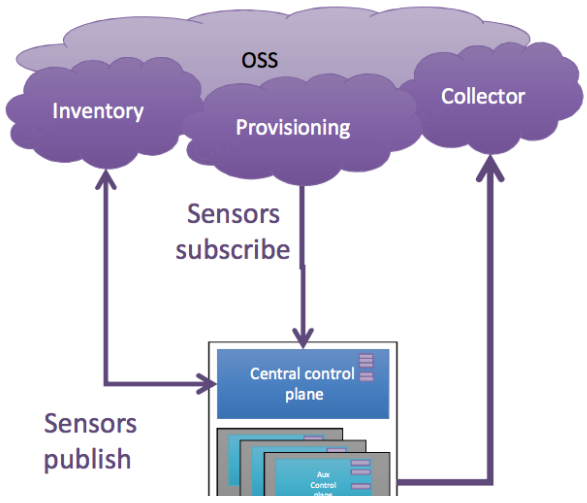
- Better suited for high fidelity event reporting.
  - Traffic Micro-bursts
  - Short-living states e.g. Signaling/routing messages peaks
  - In run-time
- Feeds Machine Learning [ML] network analytics
  - We do not know about network what we do not know.
  - Enables AI/ML/BigData analytics to locate patterns and relationships the operators haven't yet discovered. This requires a lot of data at high fidelity.
- Enables optimization:
  - usage-based provisioning and configuration.
  - Trend-based behavior prediction and proactive actions.

High-Fidelity, Run-Time statistics and events collection

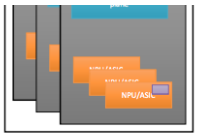
7

NG telemetry components

- Available sensor publication
  - Data models
  - NETCONF
- Subscription to data from sensor
  - Data models
  - NETCONF/CLI
- Sensor data stream



- Data models
- Export protocols



Sensors
Data
export

8

The sensor

- The “thing” in the node that is
  - Aware of all objects of particular type
  - And able to trace counters and states associated with that object
- If enabled, exports all data together
- Sensor instantiation point
  - Depends on type of object it monitors.
  - To cover whole node, multiple sensors of same type may be needed

Example:

- Physical interface sensor
  - Knows all physical interfaces of PFE it is instantiated
  - Trace and report following counters of all this interfaces
    - In/Out bytes
    - In/out packets
    - In/Out pps
    - In/Out bps
    - Max/Avr. Utilization
    - Ingres/Egress Queue stats:
      - Queued packets & bytes
      - Tail dropped packets
      - RED dropped packets & bytes
      - Average, current and peak occupancy
      - Allocated size

9

NG telemetry – publication and subscription

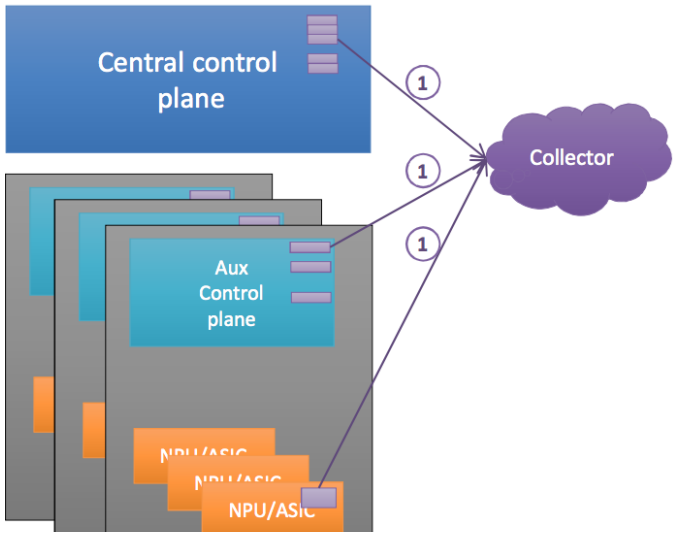
NG telemetry – publication and subscription

- NetConf / YANG is good for this
  - No much changes – static data
  - Rarely used – well… once.
  - Machine-Machine
  - Vendor-agonistic
- Human-Machine CLI for old school like me is an option

```
module: jvision
   +--rw sensor [sensor-name]
   |  +--rw sensor-name          string
   |  +--rw server-name?         string
   |  +--rw export-name?         string
   |  +--rw polling-interval?    uint32
   |  +--rw resource?            string
   |  +--rw resource-filter?     string
   +--rw streaming-server [server-name]
   |  +--rw server-name        string
   |  +--rw remote-address?    ipv4addr
   |  +--rw remote-port?       uint16
   +--rw export-profile [profile-name]
      +--rw profile-name       string
      +--rw local-address?     ipv4addr
      +--rw local-port?        uint16
      +--rw reporting-rate?    uint32
      +--rw format?            enumeration
      +--rw transport?         enumeration
      +--rw timestamp?         enumeration
```

NG telemetry - data streaming

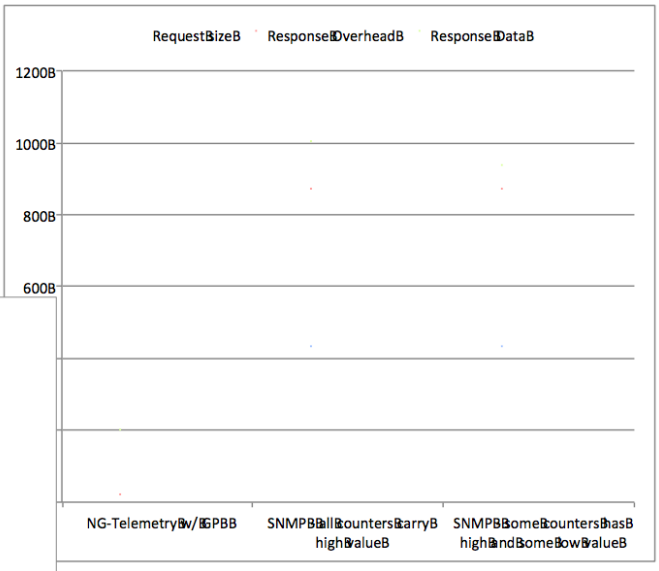
- Pushed reports
  - Periodical (**sub-second**)
  - At time of state change (events).
- Distributed sensors. Reports directly from the location of the data
  - No nested pull
  - Sensor and export profile provisioned once.
- Open export protocol
  - Efficient
    - On wire
    - On resources - ASIC
  - Compatible with IT systems in use
  - Easily extensible

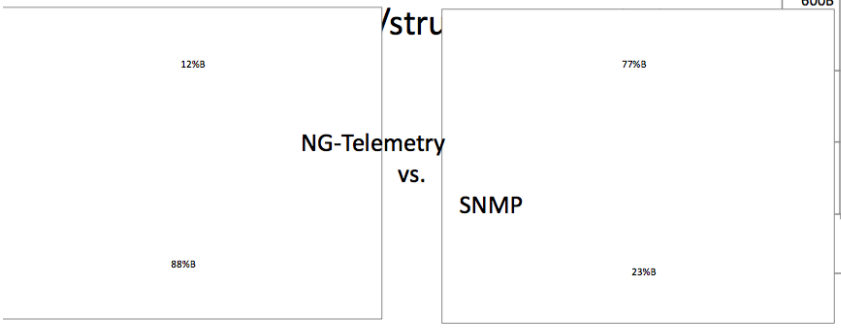


11

Binary streaming is efficient

- No request
  - No-Data
  - No processing on node
- N:1 data- to-overhead ratio in report
  - N >> 1





12

Sensor's object models and data structures

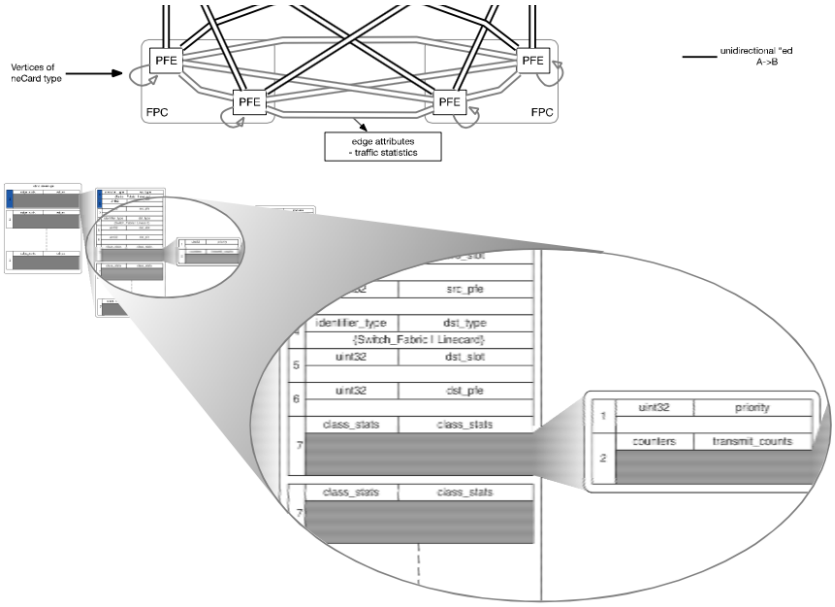
structures

Formal, abstraction model

- Unambiguous
- Implementation (vendor) agnostic

Data structure

- Follows model
- Efficient (optional fields)
- Extendable

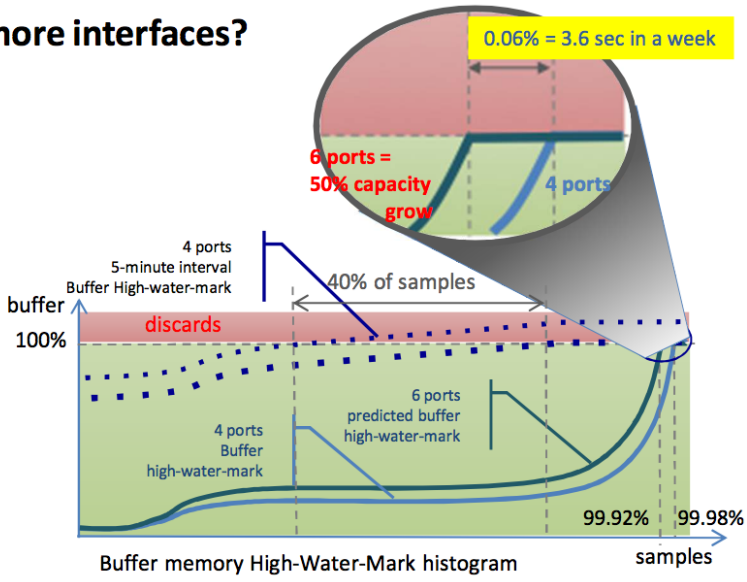


NG telemetry use case– example 1

What would be the impact if I connect more interfaces?

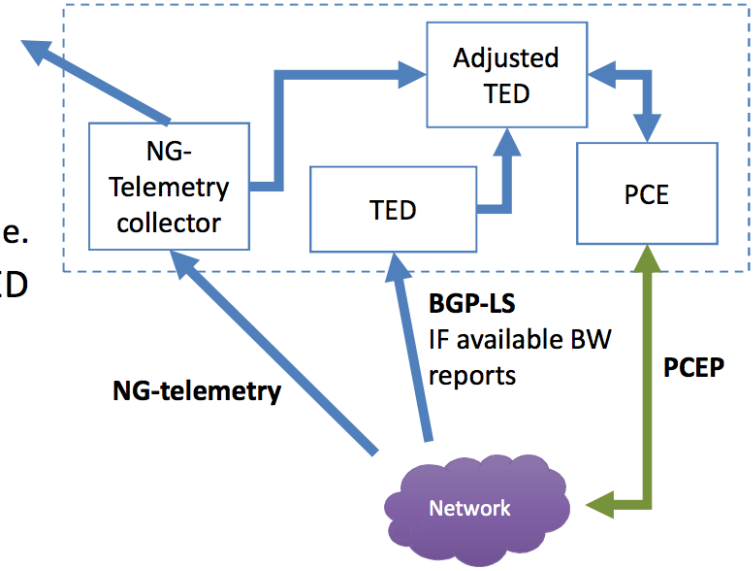
- Network device with 4 active interfaces. **What would be impact if I connect another 2 ports of given ASIC to increase BW?**

- Every 2sec NG Telemetry reports "peak buffer occupancy"
  - 1 week – 220,000 samples
  - Predicted buffer utilization - increase 1.5 times
  - 0.06% of samples reach 100%
- **Predicted impact - loss rate increase by 0.06% or less**
- 5min interval (SNMP typical) - Predicted impact - loss rate increase by 40% or less.



NG telemetry use case– example 2

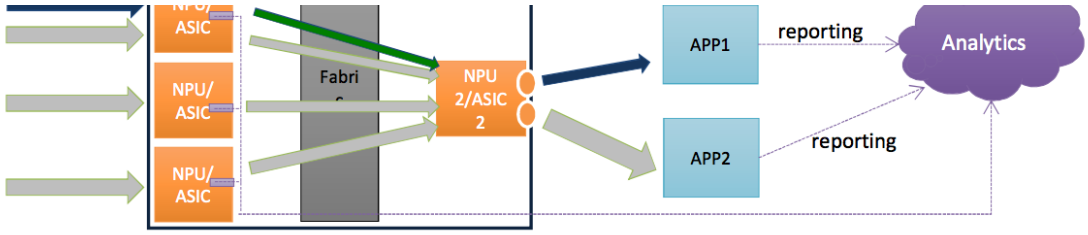
- TED carries only RSVP view of interface BW
  - Total BW (== IF BW in most cases)
  - Current available BW
  - Traffic of native IP and LDP invisible.
- Telemetry data allows to adjust TED used by PCE.
  - Fix accuracy of existing metrics (IF BW)
  - New innovative metrics (e.g. # transit LSP, Switch Fabric load)
- PCE decision more accurate.



15

NG telemetry use case– example 2

Interface do not show packet loss, but Application experience performance loss

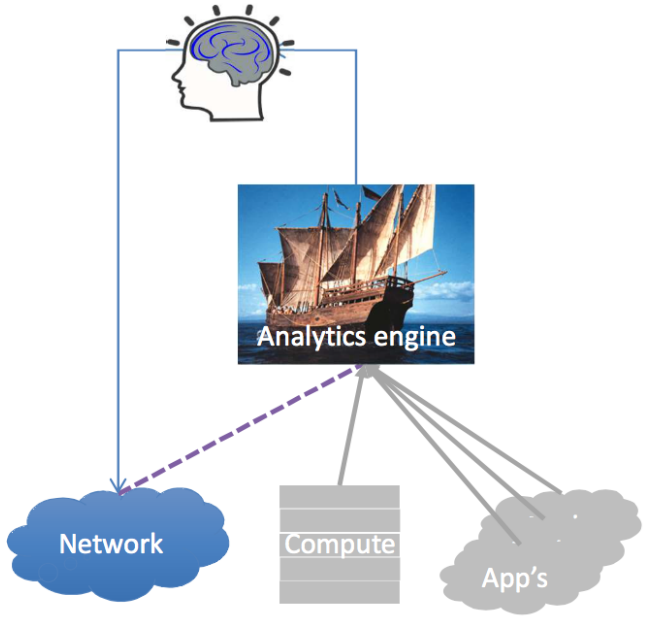


- Analytics locates a pattern – App complains reports align with NPU1 packet drops report
- NetOp investigate how they are connected
  - NPU2 From-Fabric overloaded
  - Traffic spike on another NPU toward NPU2 (green)
- Resolution - Re-locating some systems connected to NPU2 to other NPU

16

NG telemetry App – Big Data

- We do not know what we do not know
  - That's Columbus adventure - journey to the unknown
- You may discover **Eldorado**
  - But most probably just America.'
- Analytics engines is a todays *"Caravels"* – a vehicle to bring us to the new territory.
- Human brain power, would still be required to provide:
  - Practical needs
  - Interpretations



17

Early Field Experiments

- Web2.0 company and Juniper
- Export format - GPB
  - Easy to program, compiles w/ multiple languages, open source library
  - Binary data format
  - Fit's well into Company's Analytics infrastructure.

18

Experiments – LSP monitoring

- Using RSVP-TE with auto-bandwidth for network self-optimization.
- Report
  - state change of each end every LSP (rsvp) followed by

```
// Schema for communicating properties of
LSPs
message lsp_monitor_data_property {
    optional int64 bandwidth           = 1;
    optional string path_name          = 2;
```

- New BW reservation
- New ERO and RRO

- Reporting is pushed by router immediately
- Purpose:
  - Network state rewind – troubleshooting
  - Bandwidth demands trend analysis

```
    optional int32 metric                = 3;
    optional float max_avg_bw            = 4;
    optional ero_type ero                = 5;
    optional rro_type rro                = 6;
}
```

19

Experiments – interface stats monitoring

- 1000 of interfaces per router to be monitored
  - 19 counters each
  - 64bit counters
- Independent sensor per line card
  - Load distribution
  - Not all data to be exported as same time
  - Better scalability
- Purpose: Performance management – Higher fidelity – shorter intervals.

20

Experiments – findings

- Interface statistic interval was decreased by **order of magnitude** times without any impact on Router

- Time needed to collect all LSP states and attributes reduced by order of magnitude.
  - No need to pull nodes for attributes after trap
  - No need to parse attributes

21

Summary

- Telemetry is critical for Networking industry in optimization phase

- Higher scale network and nodes requires NG Telemetry architecture
  - High Amount of monitored objects on each node
  - High Amount of monitored nodes
  - Resource optimization w/o compromising quality

- NG Telemetry initially augment SNMP where high scale and high fidelity is required.

- True power of NG telemetry is yet to be discovered w/ analytics
  - The modern approach already in place by most IT organizations.
  - This may be the beginning of journey.

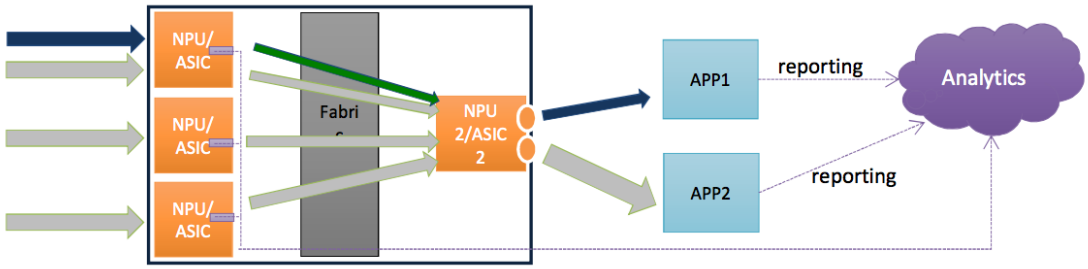
22

Thank You

23

NG telemetry use case– example 2

Interface do not show packet loss, but Application experience performance loss

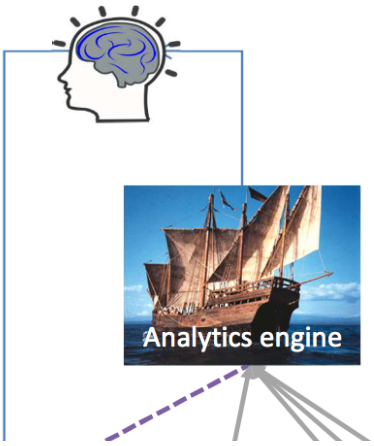


- Analytics locates a pattern – App complains reports align with NPU1 packet drops report
- NetOp investigate how they are connected
  - NPU2 From-Fabric overloaded
  - Traffic spike on another NPU toward NPU2 (green)
- Resolution - Re-locating some systems connected to NPU2 to other NPU

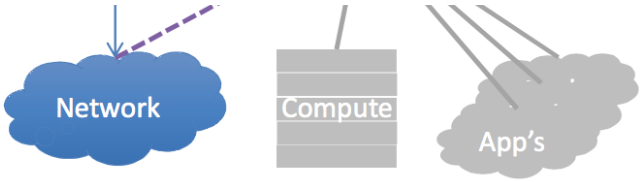
24

NG telemetry App – Big Data

- We do not know what we do not know
  - That's Columbus adventure - journey to the unknown
- You may discover **Eldorado**
  - But most probably just America.'
- Analytics engines is a todays *"Caravels"* – a vehicle to bring us to the new territory



- Human brain power, would still be required to provide:
  - Practical needs
  - Interpretations



25