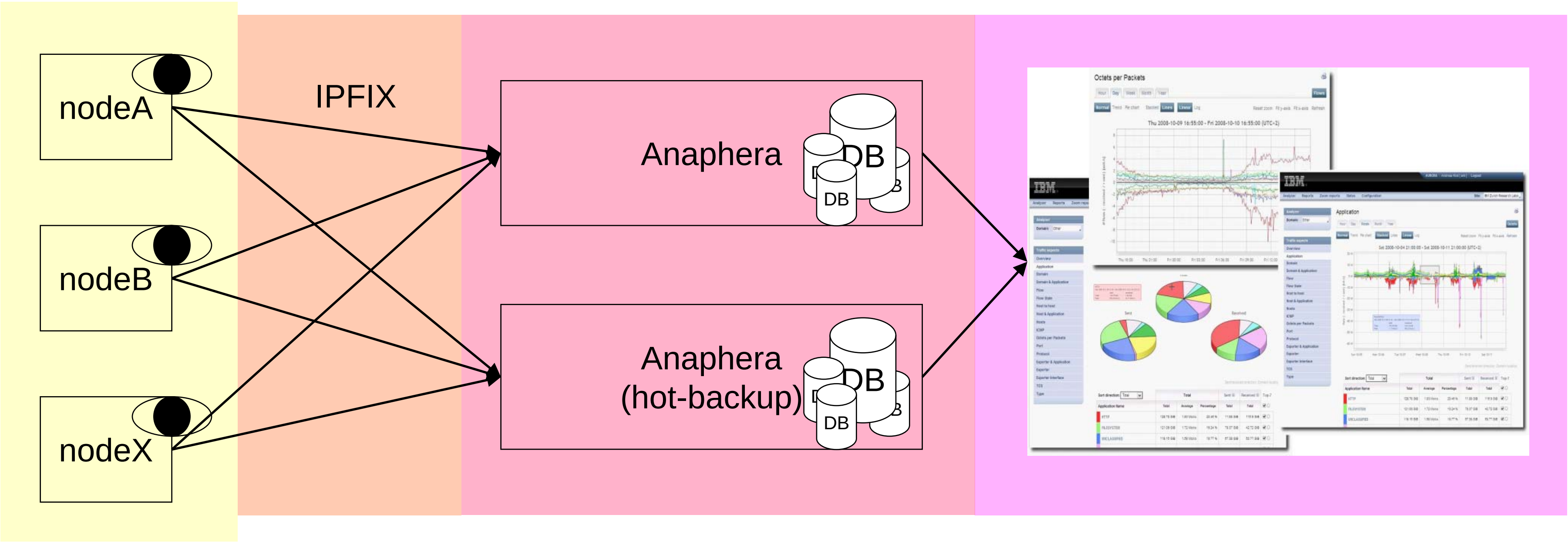
Using Flow For Other Things Than Network Data

Is the coke machine half empty or half full?

Jeroen Massar <jma@zurich.ibm.com>

Why are we doing this

- We have developed our own high-performance & scalable Flow Analyzer (Anaphera)
- First solely targeted at Network Traffic, which was our primary focus
- Does aggregation, correlation and anomaly detection



Why only look at network information?

- A number of IBM internal organizations saw our tool when used for network usage and where generally impressed with the speed flexibility and usability of the UI.

  The SONAS (Scale Out Network Attached Storage) team requested if we could also create a similar tool for their storage line of products.

- We know that IPFIX is a quite compact, easily parseable and generatable format and due to the Enterprise IDs and flexible Element IDs can easily be made useable for other data than network.

- We thus enhanced our tool to be able to analyze any kind of data
  - which is (partially) the idea behind IPFIX
  - and why not do it, same engine, just more data, more correlation

- Biggest advantage: a single parser for IPFIX

SNMP versus IPFIX

- SNMP = poll, IPFIX = push

- Problem with SNMP is that one has to poll all the devices
- Want measurements every *n* minutes, out of 100.000 meters
  - Great challenge in creating a tool that can poll that amount of meters
  - Especially when devices are not always online/reachable
  - TCP state complicates matters too, generally need to distribute collection over multiple machines

- With IPFIX, just configure those 100.000 devices to push their metrics out every *n* minutes
- Need a collector which can accept quite bursty traffic
- Could anycast collectors to spread load if really needed

XML Registry

IANA IPFIX Information Element registry http://www.iana.org/assignments/ipfix/ipfix.xhtml

```
<xml….
<registry….
…
 <record>
      <name>IBM_disk_reads</name>
      <ibm_title>Disk Reads</ibm_title>
      <ibm_type>uint</ibm_type>
      <ibm_related>
         <elementId>IBM_disk_writes</elementId>
         <elementId>IBM_cpu_load</elementId>
      </ibm_related>
      <group>IBM-Storage-Disk</group>
      <elementId>10001</elementId>
      <enterpriseId>2</enterpriseId>
      <description>
        <paragraph>
         CPU Usage, User part
        </paragraph>
      </description>
    </record>
```

The name of the component
Title for the graphs
The value is an integer
Related values

What group it belongs to
The IEID
The IBM Enterprise ID
Little description for humans

Data Types

- String (BPSL style)
- ISO Country Code (eg .ch)
- IP address (4 bytes it is IPv4, 16 it is IPv6)
- EUI48 (MAC Address)
- IE (Information Element)
- Hex
- Float
- Unsigned Integer
- Datetime
- Time
- Octets
- Packets
- Flows
- ASN
- FlowLabel
- Port
- Domain
- Interface
- FlowVersion
- VLan
- ICMP

Static Templates are cheap

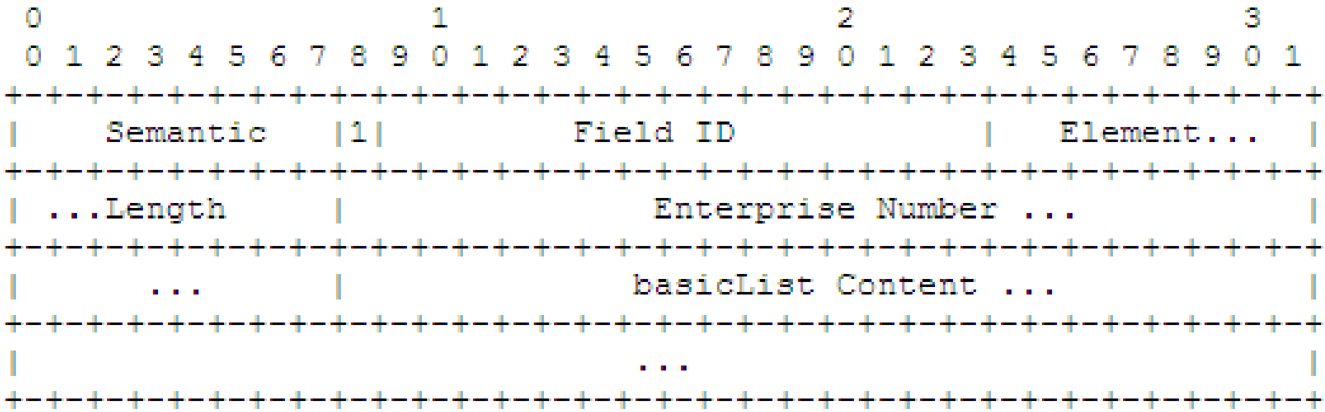
- Implementation wise, creating an IPFIX meter is ‘cheap’:
  - Define a static structure
  - Fill structure every <n> time with data
  - Export structure over the network
  - Once in a while send a template that describes the structure

- Can easily be done in silicon
- Watch out for endian issues ;)

Use of new IPFIX BasicLists

- https://datatracker.ietf.org/doc/draft-ietf-ipfix-structured-data/
- IETF Working Group item, but not finalized yet
- Defines a way to store repeating information into IPFIX records

- Useful for instance when one has multiple harddisks, multiple cpus, but also ASPaths



```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Semantic    |1|         Field ID            |   Element...  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| ...Length     |               Enterprise Number ...           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      ...      |              basicList Content ...            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              ...                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Aspects

Command format: aspect new <name> <type> [<components> …]

```
aspect new cpu tva ip_exp (*IBM_cpu_idle *IBM_cpu_iowait *IBM_cpu_system)
aspect set name “Host CPU Usage”
```

This configures an aspect called “cpu” with name “Host CPU Usage” which generates graphs for each host.

The keys will be generated from the IP address of the exporter (ip_exp) and the IEID (Information Element Identifier) of the components specified, the value will be what the IEID specifies.

The asterisk in front of a component name indicates that the name goes into the key and the value is used for the value. Normally, like for ip_exp above, the value is stored in the key.

The braces indicate a set of “or” components, eg to store both source and destination addresses one can use:

```
aspect new host tva (ip_src ip_dst)
```

IPFIX over Delay Tolerant Networking or SMTP

- Not all devices are connected 24/7
- DTN specifies two protocols for store-and-forward messaging (Licklider + Bundle)

- Can also use SMTP which is easier to setup, just have a local mailspool which gets flushed when the host dials in to the network / connects.

- Useful for retrieving metrics from nodes which are not always connected like sensors that are dropped around a place where the sensors don't have a lot of battery power

Storage

- Performance management is important in storage environments
- Can combine network trends with disk activity
- Instead of top talkers, figure out what files are “hot”, and in that case move those files/blocks of data to SSD for quicker access
- Can optimize LRU and MU caches based on data that is collected

Example statistics:

- NFS
- Samba/CIFS
- Disk Usage
- CPU load

In total >2500 separate metrics…

Electric cars & Windmills

EDISON: Electric vehicles in a distributed and integrated market using sustainable energy and open networks

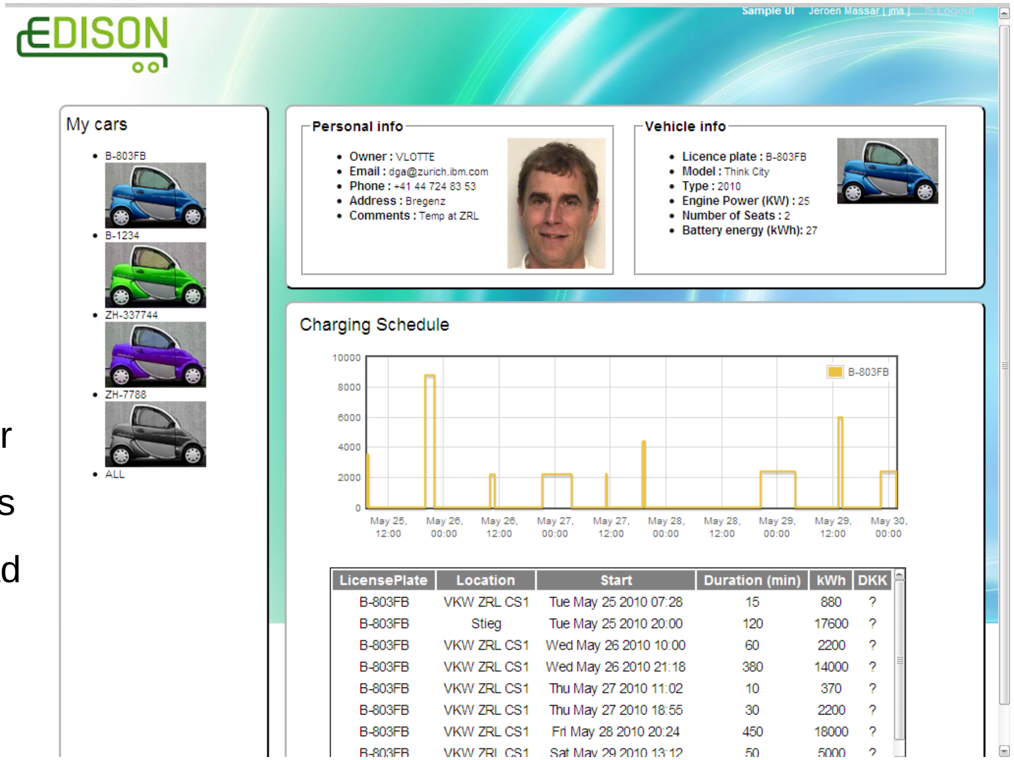
One part of this involves Electric Vehicles (EVs) and managing when these EVs re-charge, in a way to not overload the electrical network and using renewable resources as efficiently as possible.

When the cars charge, they can communicate with a central server.

We then send using IPFIX the averaged speed, drive duration, power consumption etc to the IPFIX collector.

The driver can indicate what kind of trips will be undertaken and when the car should be fully charged. Various algorithms then instruct the car when it is cheapest to charge and at which times

It is preferred to charge itself due to network load



Personal info

- Owner : VLOTTE
- Email : dga@zurich.ibm.com
- Phone : +41 44 724 83 53
- Address : Bregenz
- Comments : Temp at ZRL

Vehicle info

- Licence plate : B-803FB
- Model : Think City
- Type : 2010
- Engine Power (KW) : 25
- Number of Seats : 2
- Battery energy (kWh): 27

LicensePlate	Location	Start	Duration (min)	kWh	DKK
B-803FB	VKW ZRL CS1	Tue May 25 2010 07:28	15	880	?
B-803FB	Stieg	Tue May 25 2010 20:00	120	17600	?
B-803FB	VKW ZRL CS1	Wed May 26 2010 10:00	60	2200	?
B-803FB	VKW ZRL CS1	Wed May 26 2010 21:18	380	14000	?
B-803FB	VKW ZRL CS1	Thu May 27 2010 11:02	10	370	?
B-803FB	VKW ZRL CS1	Thu May 27 2010 18:55	30	2200	?
B-803FB	VKW ZRL CS1	Fri May 28 2010 20:24	450	18000	?
B-803FB	VKW ZRL CS1	Sat May 29 2010 13:12	50	5000	?

Road Traffic

- System which can identify license plates
- Record speed at point X

  => Send using IPFIX: license plate, color and speed
- Record speed at point Y

  => Send using IPFIX: license plate, color and speed

Collector can average the measurements out, toss the license plate

Add a road topology to the mix and you gain insight on what routes cars take and where there are a lot of cars, where congestion Happens what changes in speed there are during congestion etc.

Open Issues / Future Work

- Standardize the types and the extra information in the
- Central/Global registry where every organization can register their Information Elements most likely IANA will be appropriate for this as the default IPFIX IEs are also there

Is the coke machine half empty or half full?

Sometimes you want a drink

Sometimes the vending machine is empty

Do you want to walk over to find out if it is empty, or do you want to just stay in your chair?

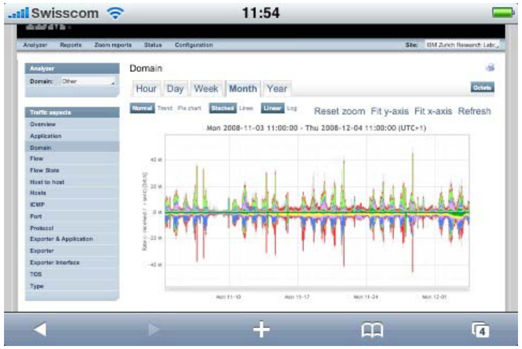
=> Instrument the vending machine

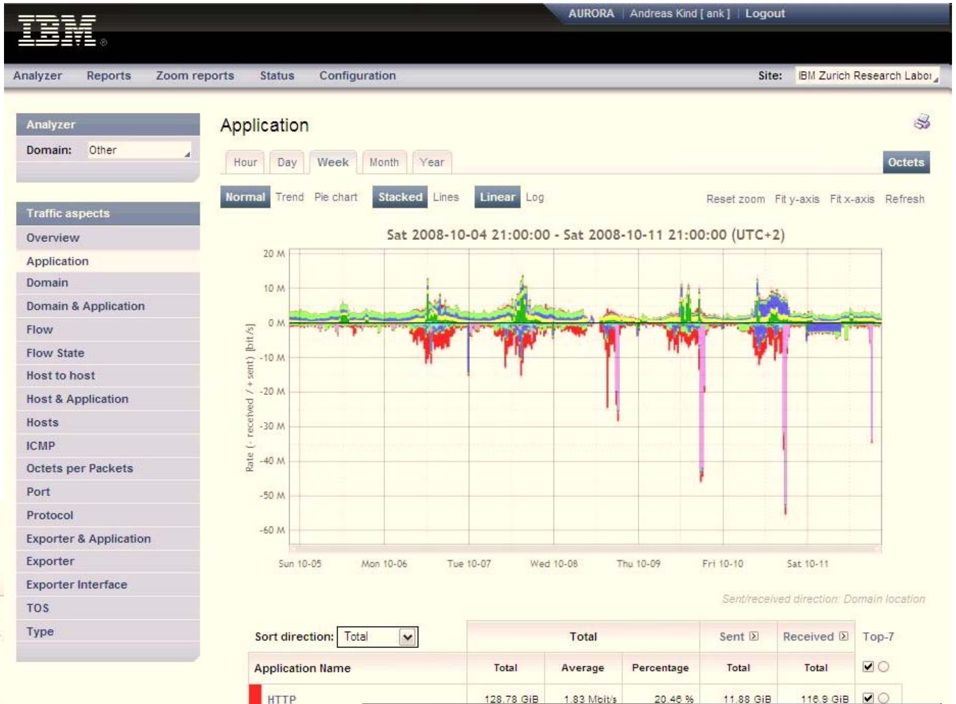
- Vending machine has a payment protocol
- Cards contain an ID, credit is centrally administered.
- Tap into the serial protocol between the vending machine and the credit machine
- Let the sniffer generate IPFIX packets, solely on the part of the protocol acknowledging payment and the type of product bought.

Questions?

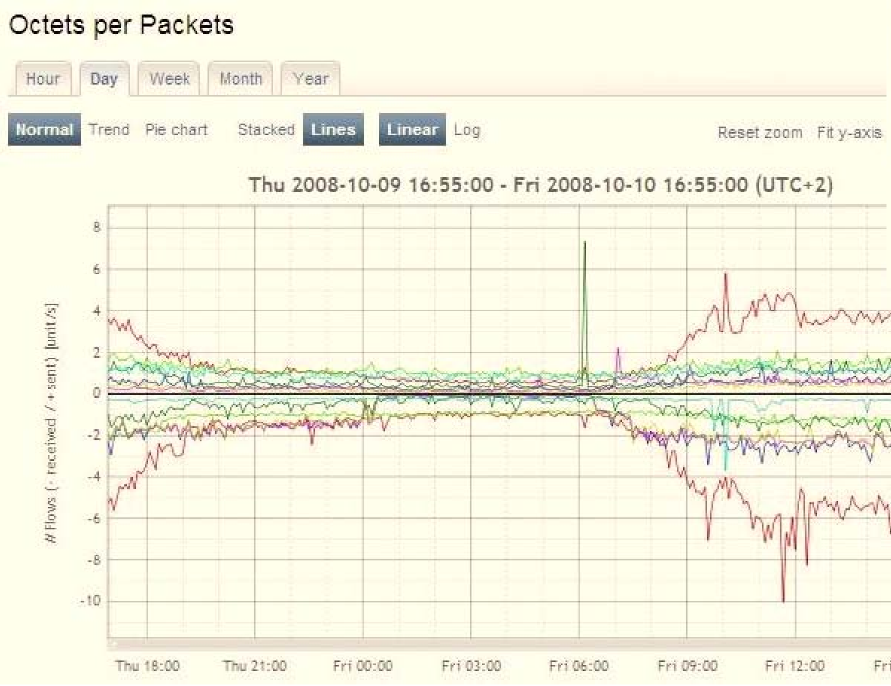
Jeroen Massar <jma@zurich.ibm.com>

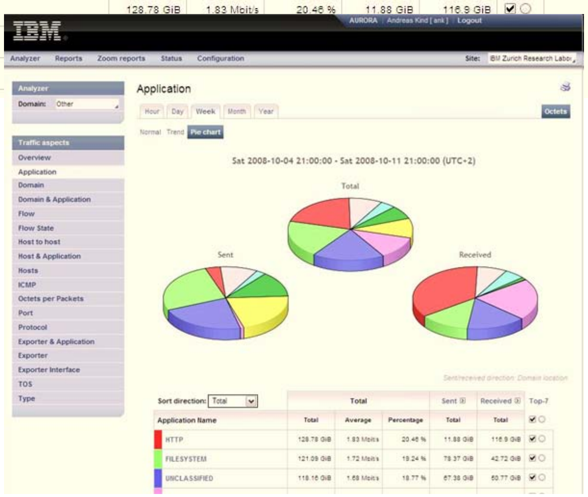
Screenshots





	Total			Sent	Received	Top-7
Application Name	Total	Average	Percentage	Total	Total	☑ ○
HTTP	128.78 GiB	1.83 Mbit/s	20.46 %	11.88 GiB	116.9 GiB	☑ ○
FILESYSTEM						
UNCLASSIFIED						





	Total			Sent	Received	Top-7
Application Name	Total	Average	Percentage	Total	Total	☑ ○
HTTP	128.78 GiB	1.83 Mbit/s	20.46 %	11.88 GiB	116.9 GiB	☑ ○
FILESYSTEM	121.09 GiB	1.72 Mbit/s	19.24 %	78.37 GiB	42.72 GiB	☑ ○
UNCLASSIFIED	118.16 GiB	1.68 Mbit/s	18.77 %	67.38 GiB	50.77 GiB	☑ ○