Implementation of the Service Insights System at DE-CIX

Based on IPFIX and sFlow

Thomas King
R&D, DE-CIX
Agenda

» History
» Make or Buy?
» Building Blocks
» Demo
» Status and Outlook
» Lessons Learned
Tera/Exascale Platform:
- sFlow-Samples
- sFlow-Counter

XRS Platform:
- IPFIX
- sFlow-Counter

Make or Buy

» An extensive list of requirements was collected

» List of requirement was provided to a list of potential traffic statistics software vendors

» Three vendors qualified for a test run:
  » Invea-Tech Flowmon
  » Paessler PRTG
  » Isarnet Isarflow
Make or Buy II

» No product / vendor was able to fulfill the list of requirements completely

» No product was designed from scratch to be used in an IXP environment

» Major lack of features:
  » Scalability to up to 10 Tbit/s traffic
  » Scalability to over 1000 MAC addresses
  » Limited Layer-2 capabilities (e.g., MAC, VLAN)
  » Limited capabilities of configuration to reflect the constantly changing DE-CIX network topology
  » Limited capabilities of exporting the collected statistics
Make!

- elastic
- Java
- php
- python
- #!/bin/bash
- HIGHCHARTS
- TYPO3
- Scrum
Building Blocks

Data Collection: pmacct: IPFIX & sFlow

Data Storage: Elasticsearch - NoSQL DB

Data Presentation: Customer Portal
Data Collection - pmacct

» Open Source software
  » Vivid community with lots of support (thanks Paolo)
  » Enhancements have been made due to DE-CIX project:
    » Scability
    » IPFIX support enhanced
    » sFlow counter support added
    » Bugfixes and improvements

» Workflow:
  » Receives sFlow and IPFIX data from switches
  » Aggregation of network flows
    » Reduces data diversity
    » Precompiles statistics → reduces amount of data
  » Resolution of statistics
    » sFlow 15 seconds
    » IPFIX 5 minutes
  » Output in JSON format
Data Storage - Elasticsearch

» NoSQL document based database
» Easy initial setup
» Application development was challenging
» Data aging and timeline resolution

Installations:
» 11 Nodes:
  » Cluster with nine nodes storing data on SSDs
    » Indexing performance: about 50k/sec data records (current demand: 6k/sec (peak))
  » Two nodes serving queries and data aggregations
  » Nodes distributed over three locations in Frankfurt
    » One location can go offline without data loss

» Data: 14 TByte (capacity: 40 Tbyte)
» Documents: 8 billion (already)
Data Presentation - GUI

» Integration into DE-CIX website

» Different time-based resolutions
  » 5 minutes for 7 days
  » 1 hour for 1 month
  » 1 day up to several years

» Statistics available today
  » GlobePEER Service
    » IPv4, IPv6, Packets/s, Bits/s
  » Traffic Relationships
    » Packets/s, Bits/s
    » Compare up to 5 peers
Demo
Status and Outlook

» Status:
  » Statistics generated with IPFIX (packet sampling)
    » Interface / service statistics a.k.a GlobePEER service
    » Traffic Relationships between peers
  » Statistics at DE-CIX FRA, PMO, NYC

» Outlook:
  » Statistics with sFlow counters
    » MetroVlan service
    » Interface statistics with errors and drops
    » Interface statistics for resellers
  » Migration of internal applications
Lessons Learned

» Traffic statistic software vendors do not fit in the IXP world

» Make:
  » The concepts are easy, the implementation is often hard
  » The details is what has been complicated

» Open-Source software is great!
By joining DE-CIX, you become part of a universe of networks. Everywhere.

DE-CIX. Where networks meet.