Network Performance Isolation in Data Centres using ConEx

draft-briscoe-conex-data-centre-00.txt

Bob Briscoe, BT
Murari Sridharan, Microsoft
IETF-84 ConEx Jul 2012
Multi-tenant data centre
Features of ConEx Solution

- Network performance isolation between tenants
- Zero (tenant-related) switch configuration
- No loss of LAN-like multiplexing benefits
  - work-conserving
- No change to existing switch implementations
  - if ECN-capable
- Simplest possible contract
  - per-tenant network-wide allowance
  - tenant can freely move VMs around without changing allowance
  - tenant can freely move allowance between virtual machines
- Transport-Agnostic
ConEx recap
basic signals and functional units

transport
sender

transport
receiver

f/b

ACKS

policy

audit

congested
network
element

DATA

Loss
/ECN

Re-Echo

infrastructure
multi-tenant data centre arrangement
topologically equivalent to ConEx

transport sender

policer

transport receiver

audit

feedback

ACKS

DATA

congested network element

Loss /ECN

Re-Echo

infrastructure
Arrangement of ConEx functions

- Per-node ‘congestion-policers’
  - policers created in hypervisor at VM boot
  - police all ConEx-enabled packets entering network

- Token buckets
  - congested-bit tokens, not bit tokens
  - drained by ConEx Re-Echo packets
- Filled from one single allowance (W) per tenant
one logical token bucket per tenant

- Any one sub-bucket can fill faster than others
- subject to
  - the total fill-rate allowance $W$
  - a maximum drain-rate per sub-bucket (not shown)

- if tokens represented bits
  - a big enough tenant could do unlimited harm to others
- but because tokens represent congested-bits
  - tokens drain faster the more a tenant harms others

- this* provides inherent performance isolation between tenants
- while giving each tenant maximum flexibility and minimum config hassle

* with max drain-rate per-sub-bucket constraint
Deployment

• Deploy all ConEx infrastructure under control of one administration
  except for sender (and receiver)
   – need ConEx in guest OS within virtual machine

• Alternative (cf Microsoft Seawall)
  – trusted feedback tunnel back to policer
  – under control of DC operator

• Hybrid
  – non-ConEx packets: feedback tunnel
  – ConEx packets: no tunnel
  – reward ConEx for being more efficient?
status & plans

working group input
Network Performance Isolation in Data Centres using ConEx

draft-briscoe-conex-data-centre-00.txt

Q&A