Initial ConEx Deployment Examples

draft-briscoe-conex-initial-deploy-02.txt

Bob Briscoe, BT

Dirk Kutscher, NEC

IETF-83 ConEx Mar 2012

draft status

- individual draft

- draft-briscoe-conex-initial-deploy-02.txt

- work in progress

- intended status: informational

- new co-author, Dirk Kutscher

- immediate intent:

  - WG feedback on scenarios

  - adopt as WG item?

three example network arrangements
that have incentives for unilateral ConEx deployment

1. single receiving access network

   • presented in Taipei

2. (new) mobile network

   • simple scenario for single operator mobile network
     – problems that ConEx addresses in a mobile/cellular network
     – arrangement of ConEx functions
     – deployment incentives
     – [kutscher-conex-mobile] covers more scenarios and details

3. (new) multi-tenant data centre

   • network performance isolation
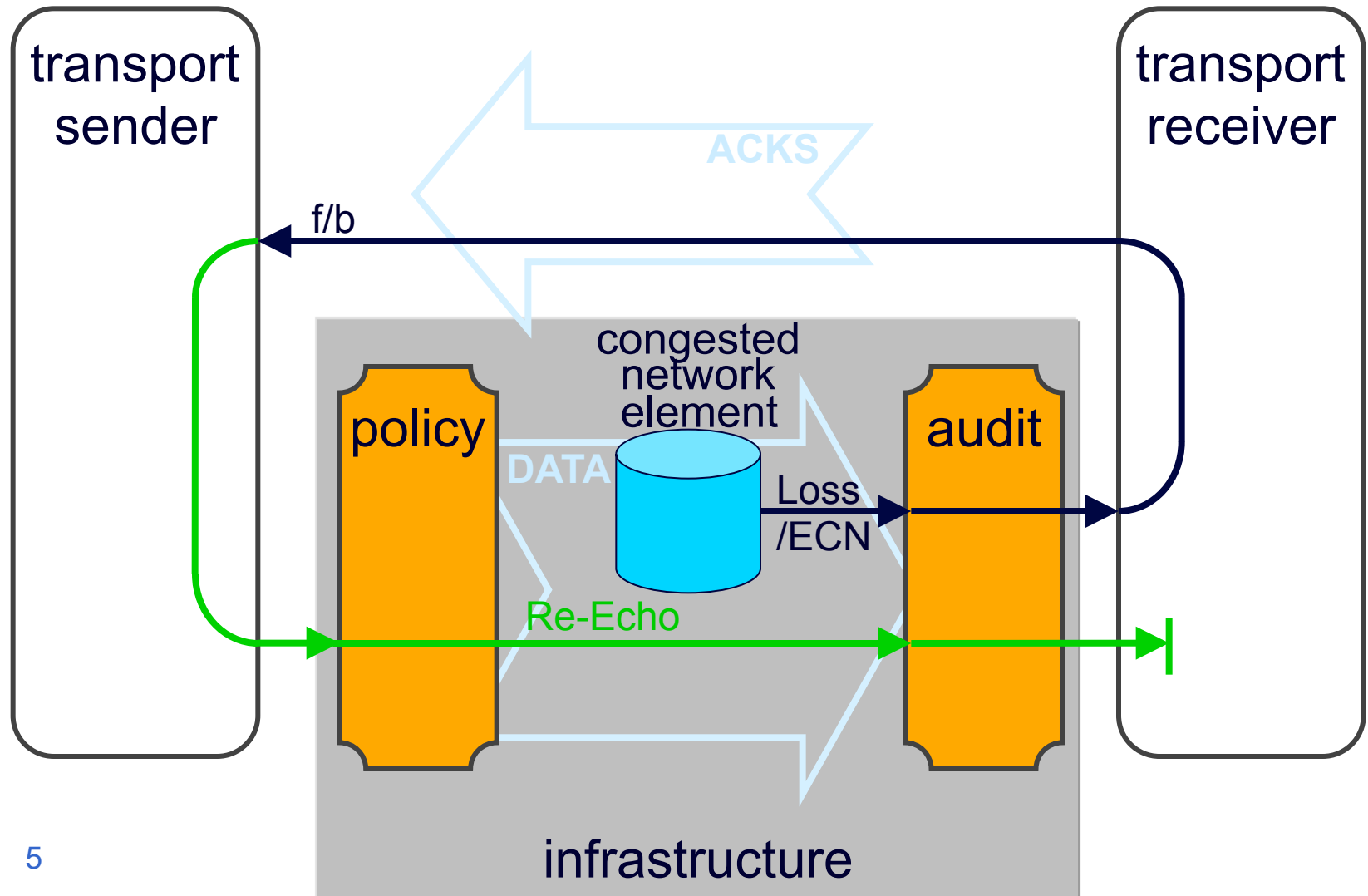
   • the subject of the body of this talk

#3 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
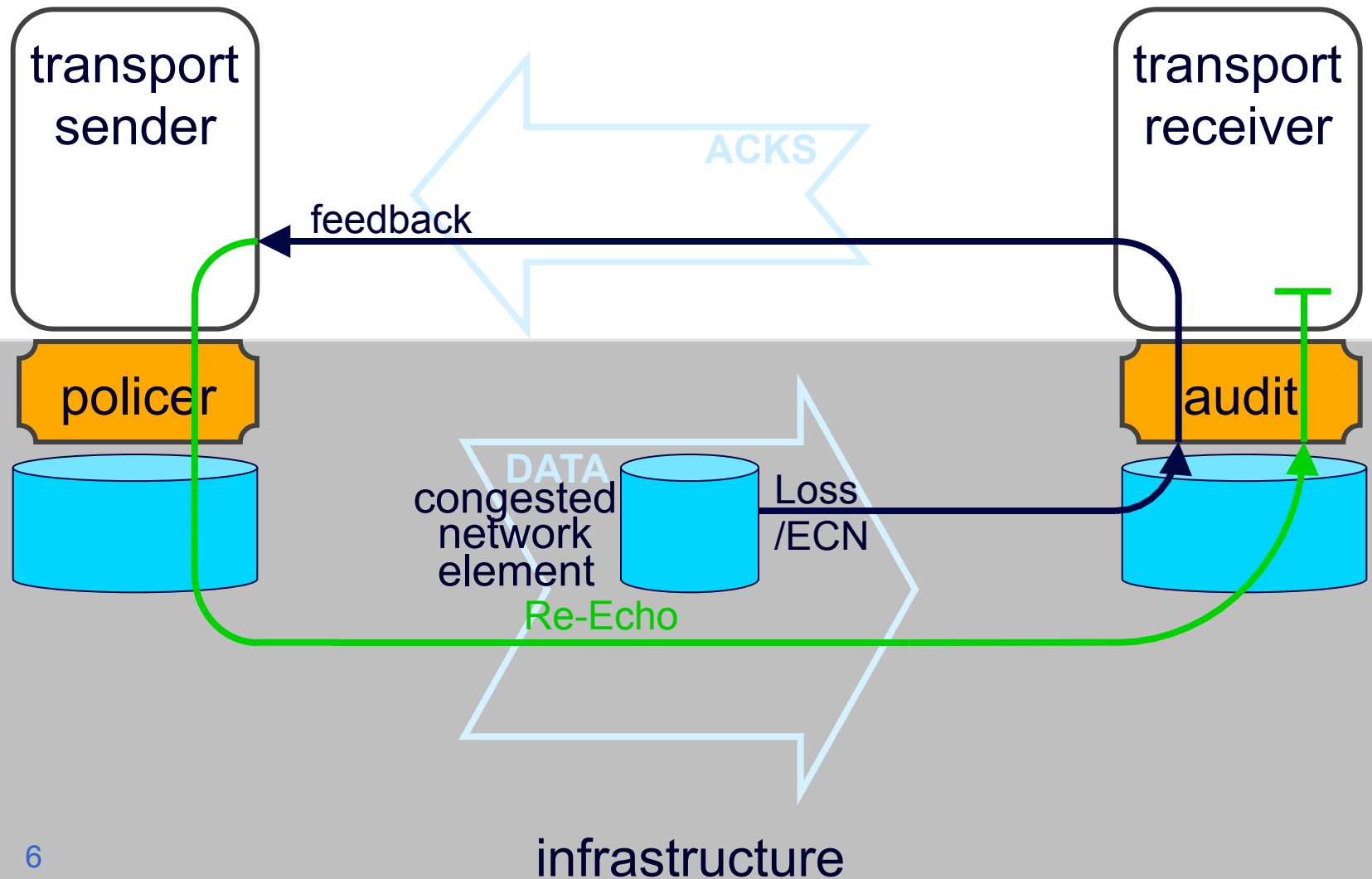
- Transport-Agnostic
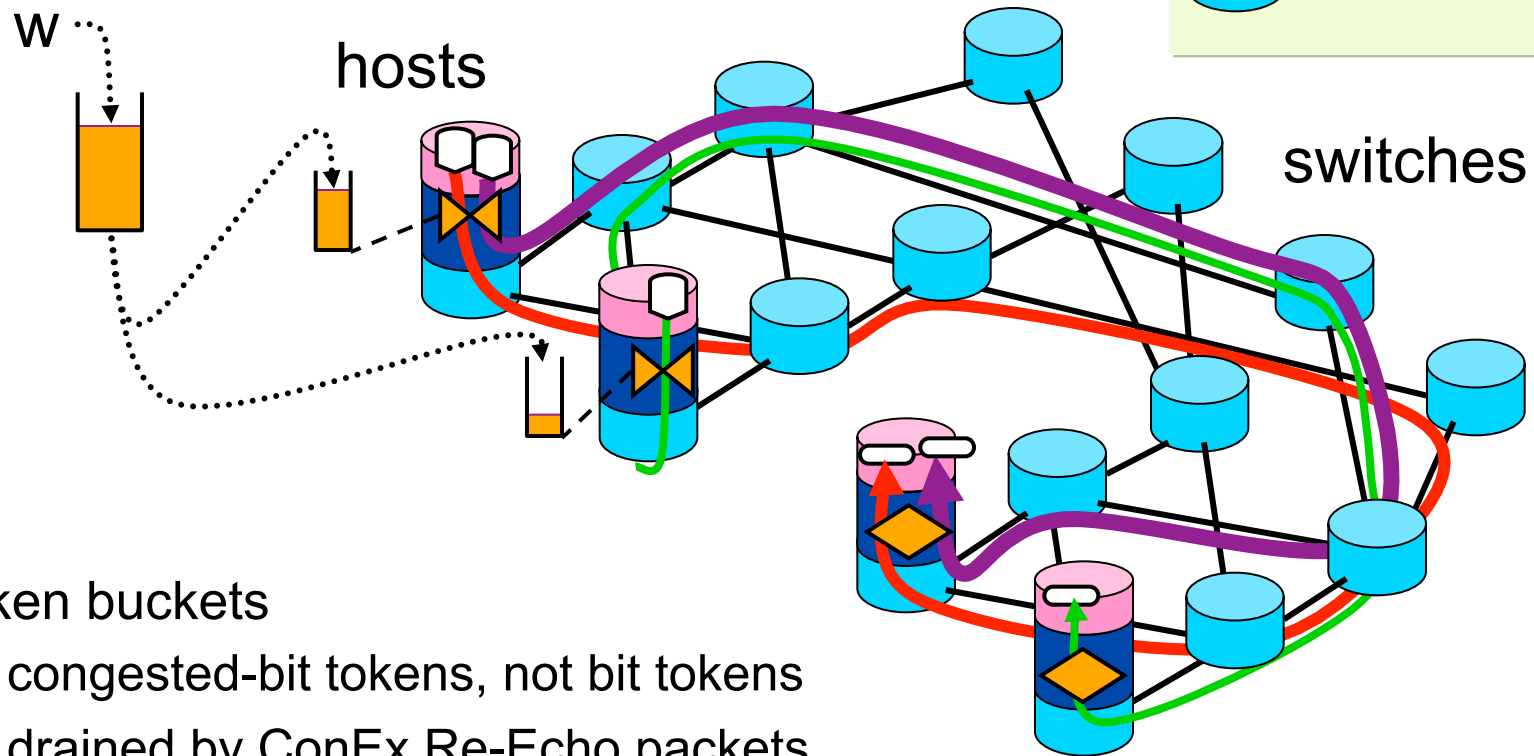
ConEx recap
basic signals and functional units

transport
sender

transport
receiver

ACKS

f/b

congested
network
element

policy

DATA

Loss
/ECN

audit

Re-Echo

infrastructure

5

ConEx recap
basic signals and functional units

transport sender

transport receiver

ACKS

feedback

policer

audit

DATA

congested network element

Loss /ECN

Re-Echo

infrastructure

Arrangement of ConEx functions
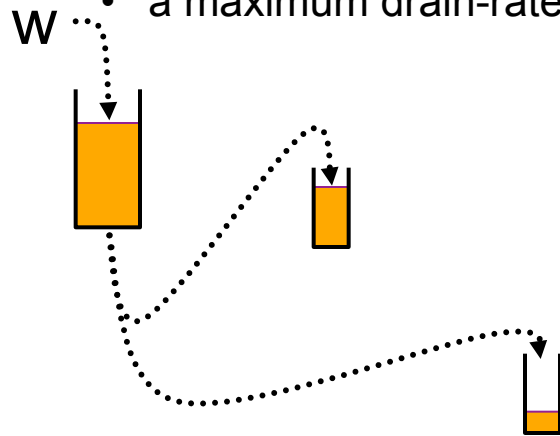
- Per-node 'congestion-policers'
  - policers created in hypervisor at VM boot
  - police all ConEx-enabled packets entering network

w

hosts

switches

VM sender
VM receiver
congestion policer
audit

guest OS
hypervisor
switching

- 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)

W

- 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
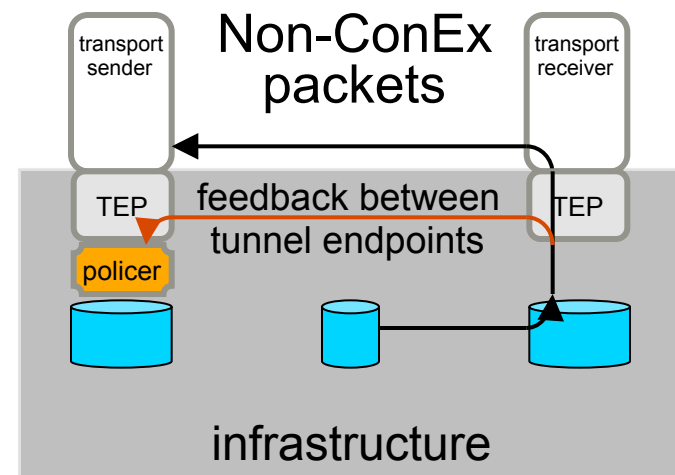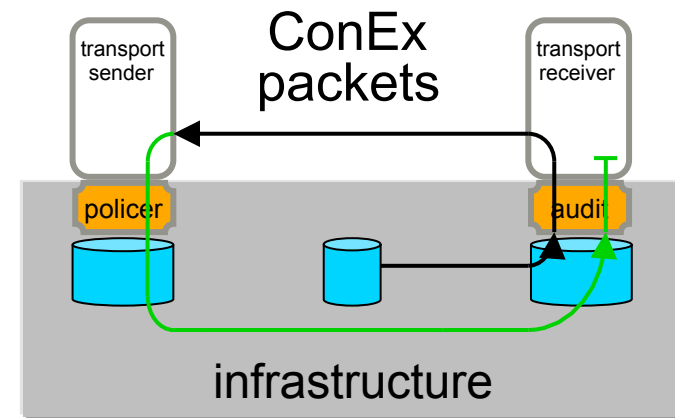- while giving each tenant maximum flexibility and minimum config hassle

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

Deployment

ConEx packets

- 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?

Non-ConEx packets

status & plans

- relationship to conex-mobile
    - mobile section in initial-deploy is for general ConEx audience
    - conex-mobile is ConEx entry-point for mobile audience
    - both hoped to become WG items
- plan – finish the document
    - re-organise to describe incentives up front
    - complete empty sections (e.g. tail pieces)

working group input

- more reviews please
- WG feedback on choice of scenarios?
- ready to be adopted as WG item of work?

Initial ConEx Deployment Examples
draft-briscoe-conex-initial-deploy-02.txt

Q&A