Bob's 00b -> 00c suggested changes shown as red text.
Toby, as you accept changes, please just reapply black font colour

A red arrow highlights something discussed in the covering email

ConEx Concepts and Uses

draft-moncaster-conex-concepts-uses-01

Toby Moncaster

John Leslie (JLC)

Bob Briscoe (BT)

Rich Woundy (Comcast)

draft status

draft-moncaster-conex-concepts-uses-01

- Individual draft
- Intended charter milestone: use-cases
- Intended status: Informational
- Intended next step: WG item

Overview

- The Problem
- Congestion Marking (ECN)
- Congestion Exposure
- Where do we stand?
- ConEx Use Cases
  - ConEx Conmponents
  - Use Case 1: Traffic management
    - *Targeted management*
    - *Incentivising better congestion control*
  - Use Case 2: DDoS protection
  - Looking to the Future
- Questions
- Next Steps
- Summary

The Problem

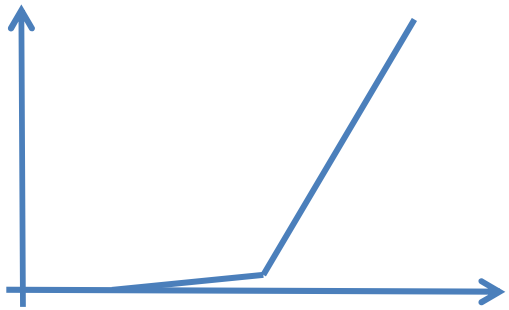
- The problem can be characterised in at least two ways:
  - Capacity Sharing – sharing limited resources between concurrent flows
  - Congestion Management – improving performance and delay for all
- Understanding congestion is definitely key
  - Too much traffic arriving too quickly = congestion
- Capacity sharing currently myopic:
  - In time (queues have no idea of past history of traffic)
  - In space (traffic may be causing problems elsewhere)
- Queues can only apply pressure by indicating congestion
  - Best signalled in forward direction (unlike Source Quench)
  - Requires honesty from receiver who wants the data as fast as possible
  - Needs sender to reduce rate in response, but it would rather send fast too
- Whole path congestion not visible at forwarding layer
  - Can't tell whether traffic is responsive to congestion

The Problem contd.

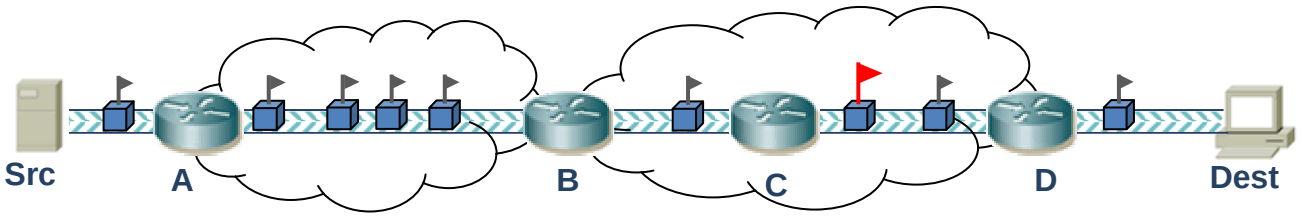
- Capacity sharing suffers from a key problem – how to measure it
- Current approaches (rate and volume) are bad:
  - They need to be measured over time
  - They don't reflect actual network conditions
- Congestion is a good measure of impact on other users
- Congestion-Volume is a better metric to measure this
  - Congestion-Volume = Volume x Congestion (units of bytes)
  - Congestion-Rate = Rate x Congestion (units of bps)
  - For a 1Mbps flow, 0.1% congestion = 125 bytes congestion-volume in 1 second
- Congestion-Volume is measure of how much excess traffic was in network over any sampling interval (millisec, minute, hour, month)
- Can be measured per-packet, per-flow, per-user, per-network, ...
- With ConEx can measure congestion-volume as easily as volume

Congestion Marking (ECN)

- Traditionally queues indicate congestion by dropping packets
  - Relies on stateful transport to spot gaps in data
  - Can lead to unwanted synchronisation effects
- RED improves this by dropping packets *before* queue overflows
  - Packets dropped probabilistically
  - Drop probability increases as the queue grows



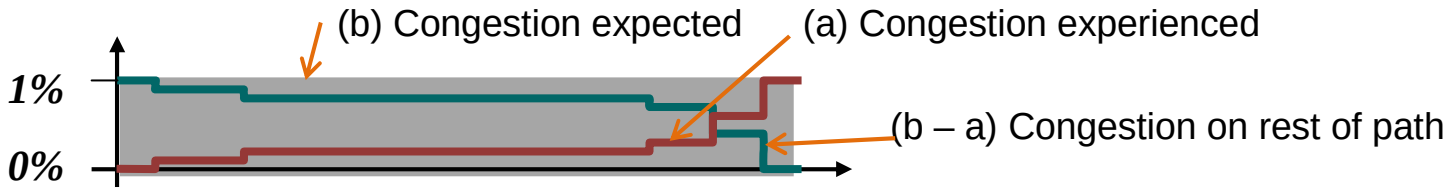
- ECN builds on RED
  - ECN marks packets instead of dropping them
  - Sender still responds as if there were a drop
  - But no data is lost so less re-transmission
- ECN shows how much congestion traffic has already experienced



- But can't see how much congestion traffic is going to encounter

Congestion Exposure

- Whole path congestion is hidden from network
  - Congestion is known to the end-systems (ECN marks or loss)
  - At any point, ECN reveals congestion so far
- What is needed is knowledge of congestion on rest of path
- ECN gives congestion experienced on every packet
- ConEx sender adds congestion expected for every packet
- ConEx enables packets to carry
  - ***a)*** ***Congestion experienced*** *(e.g. ECN markings)*
  - ***b)*** ***Congestion expected*** *(total congestion sender expects the packet to see)*
- subtracting a from b gives congestion on rest of path



- ConEx mechanism to be defined in later document

ConEx Design Requirements

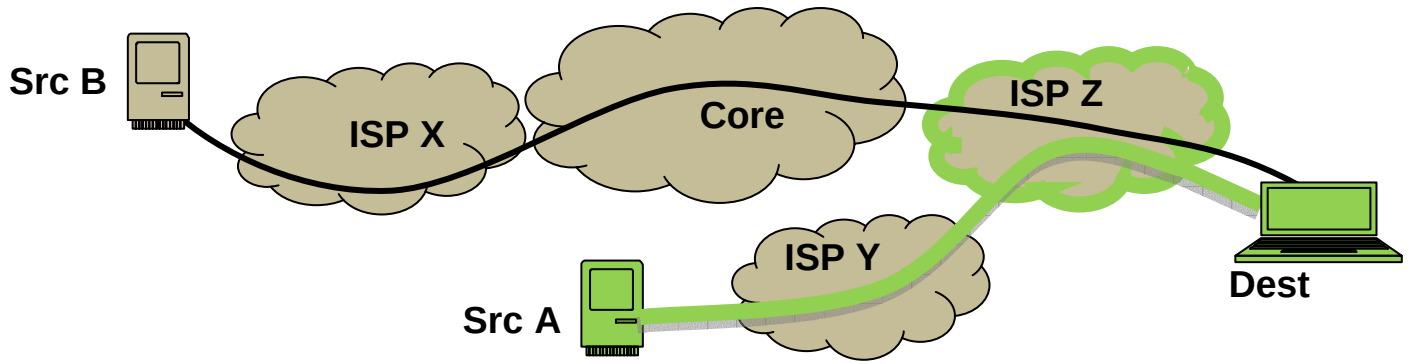
- Accuracy – ConEx info should be as accurate as possible.
  - Congestion is measured in fractions of a percent
  - Source must be trusted to correctly declare the expected congestion
  - Destination must feed back accurate whole-path congestion
  - Integrity of ConEx info must be verifiable locally
- Timeliness – ConEx info needs to be as recent as possible
  - design of network imposes min 1RTT delay
  - Transport protocol should seek to minimise delays
  - Feedback needs to be fast enough to prevent info going “stale”
- Visibility – ConEx must be visible at internetwork trust boundaries
  - ConEx must be visible in IP layer
  - ConEx markings need to survive tunneling, middleboxes, firewalls, etc

Where do we stand?

- Long process leading up to chartering
- ConEx chartered in June 2010 with limited scope
- Concentrates on one usage scenario:
  - end hosts and receiving network are ConEx enabled (other networks might not be enabled)
  - note difference between *Use Case* and *Usage Scenario*
- Can consider other use cases:
  - "Experiments on use cases are encouraged and the WG will solicit feedback from such deployments. “
- This draft covers Milestone 1 “Use Cases Description” (info)
- Several use cases explored. Some go beyond charter, but demonstrate how powerful ConEx can be

ConEx Use Cases Introduction

- Lots of use cases for ConEx
- Charter asks for use cases to focus on the following scenario:

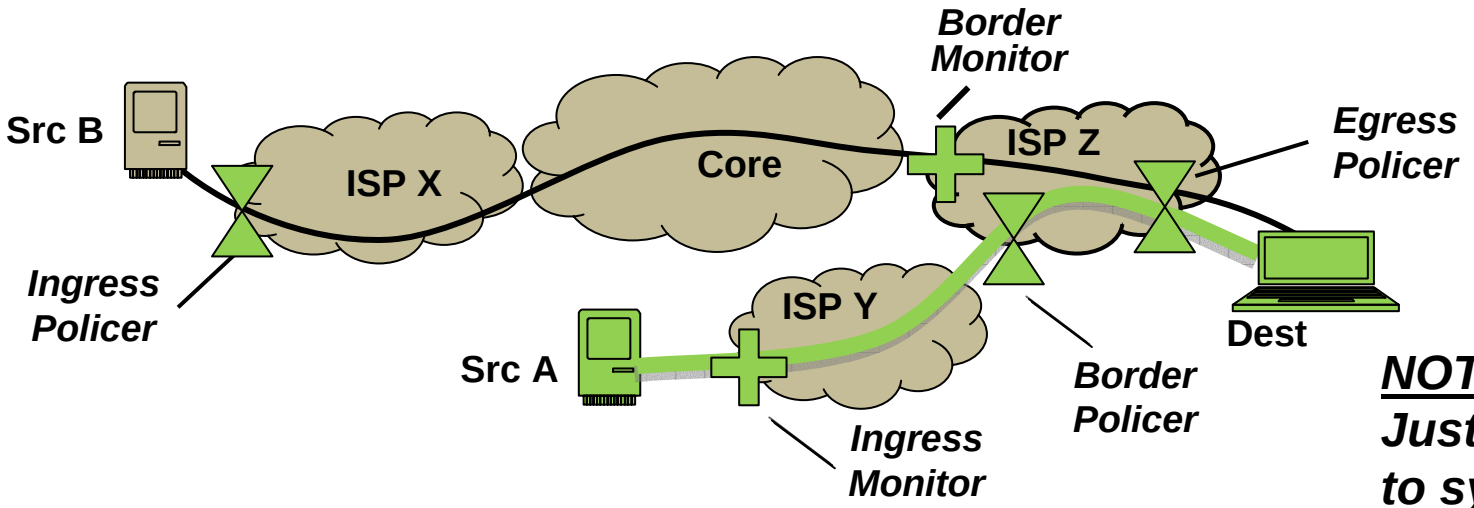


Green elements ConEx-Enabled. Grey elements not Enabled

- NB: the symmetry of most networks implies that **ISP Z** can be a ConEx-Enabled ***source*** network for any traffic that **Dest** sends into the network
- Following slides show 2 of the main use cases for ConEx
  - Traffic management
  - DDoS protection

ConEx Components

- Two new network components defined:
  - **ConEx Monitor** – a node that uses ConEx markings to measure/report the Congestion Volume that it forwards
  - **ConEx Policer** – A node that uses ConEx markings to change the queuing priority it gives a flow, or to actively control the Congestion Volume it forwards
- Policers and Monitors can be at Ingress, Egress or Border

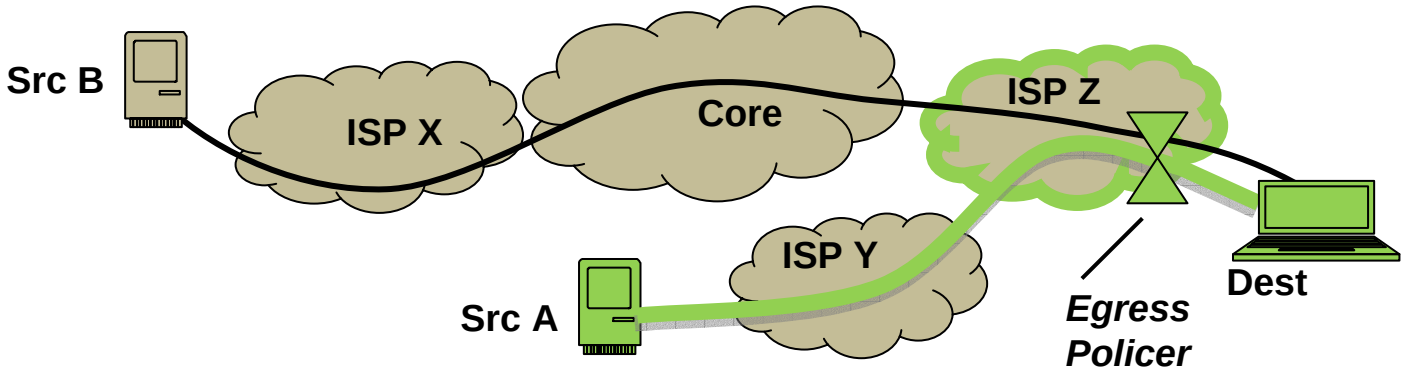


NOTE:
Just a key to symbols. Not a scenario!

- Border might have both a Policer and a Monitor
  - policing to prevent serious congestion
  - monitoring against a traffic contract to deter unnecessary congestion

Use Case 1a – Traffic Management

- ISPs often perform traffic management:
  - Aim is to give majority of users an adequate service at peak times
  - Users targeted based on application, traffic rate, volume transferred, etc
- ConEx policers offer an alternative:
  - Each sender is declaring the congestion they expect to cause
  - This can be used to control the impact they have on others
- Egress ConEx policer at the last IP node before backhaul/access can:
  - Identify the heaviest users – in terms of congestion-volume
  - Prioritise traffic depending on congestion it has declared
  - Penalise traffic that has caused excessive congestion



*Egress Policer can use ConEx info to prioritise traffic from **Src A**.*

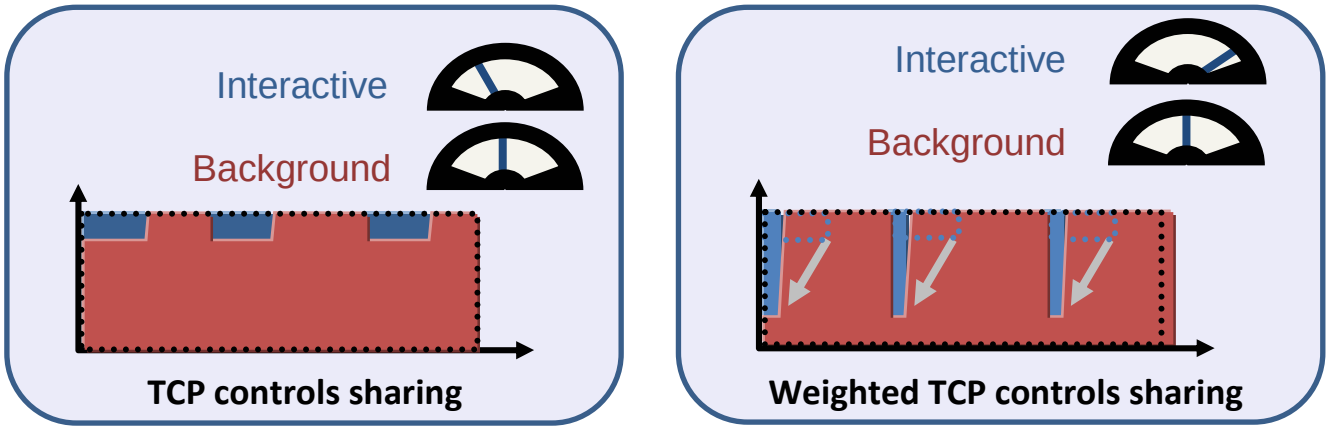
*Traffic from **Src B** might be prioritised by volume/rate/app*

Use Case 1a contd. – targeted management

- Lots of debate about traffic management
  - Current approaches tend to be relatively unfocused
  - Assumptions made about when “peak time” happens
  - Often targets specific applications - big problem for Net Neutrality camp
- ConEx approach is better
  - Only targets traffic that has caused congestion
  - Because it monitors actual congestion will always know when peak time is
  - Entirely application-agnostic – only cares about impact of traffic in the network
- Overall this is better for ISP and its user community
  - Less damaging to customer relationships
  - Offers reasonable freedom to differentiate bandwidth without QoS in the net
  - No need for expensive flow-aware kit in backhaul or access

Use Case 1b. – Encouraging Better CC

- Lots of current work looking at better congestion control
- LEDBAT introduced idea of highly reactive congestion control
  - Designed for bulk data transfers which don't care about instantaneous rate
  - As soon as queues start to build it backs off
  - In effect it reacts to congestion before other transports need to
- MulTCP and related work introduced weighted congestion control
  - Application chooses how much to react to congestion
  - High priority apps don't back off much, low priority back off more
  - Logical extension is fully weighted congestion control



Use Case 1b contd. – Encouraging Better CC

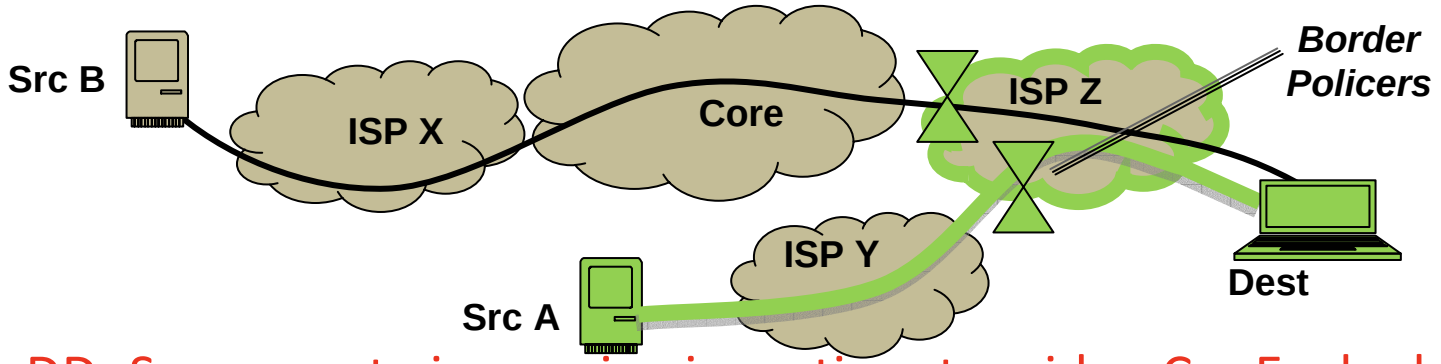
- Current traffic management disincentivises use of LEDBAT
  - LEDBAT still transfers high volumes, so is still targeted
  - LEDBAT used for applications like P2P, so is still targeted
  - LEDBAT can still reach high data rates, so is still targeted
- ConEx encourages LEDBAT-like transports
  - ConEx based traffic management brings correct incentives
  - Traffic is controlled based on congestion it causes
  - LEDBAT causes less congestion so gets less control
- ConEx encourages use of weighted congestion controls
  - Applications can choose their weight
  - Interactive applications can afford to cause more congestion to go faster
  - Background applications can back off more
  - What matters is overall Congestion-Volume...

Use Case 1c – Targeted Provisioning

- Outcome of traffic management:
  a) users — avoid gratuitous congestion
  b) protocol designers — avoid gratuitous congestion
- Any remaining congestion (need not be high)
  – represents real demand for more capacity
- Congestion-volume stats
  – loss/marks at each interface card over month
  – even small upward trend can target provisioning
- Without ConEx traffic management
  – don't know whether demand for more capacity is real
- co-authors plan to add this use-case in next rev

Use Case 2 – Raising the DDoS Bar

- DDoS is a serious problem – currently no robust solution
- ConEx Border Policers could help raise the bar
  - ConEx Policers naturally limit traffic rate towards congestion hot-spots
  - Policers can rate-limit non-ConEx traffic routing towards same hot-spot
- ConEx Border Monitors could help raise the bar too
  - ConEx DDoS traffic shows ultra-high congestion, so obvious anomaly at border
  - ConEx info closer to attack sources is a foundation for other solutions



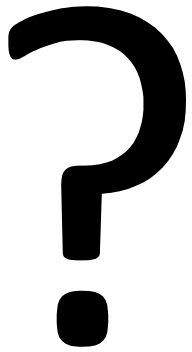
- DDoS can create increasing incentives to widen ConEx deployment
  - protection strengthens as deployment proceeds
- Details are important – but beyond scope of this presentation

Use Cases – Looking to the Future

- Current charter calls for focus on ConEx-enabled destination network :
- Obvious Usage scenarios:
  - *CDN distributing e.g. Movies;*
  - *User watching VoD;*
- Easy to add ingress policing for traffic heading other way (from this network)
  - *End user transferring P2P;*
  - *User doing live video chat with remote user via relay server;*
- ConEx for QoS (builds on weighted CC) – allows user to prioritise their traffic with no network involvement. Makes sense with ingress policing
- Congestion accounting: works best with full deployment. But deployment at sender allows any operator to monitor congestion-causing traffic
  - incentives may translate roughly through non-ConEx networks that approximate traffic costs with other metrics
- Usage monitoring Needs ingress and egress monitoring. Ideal in Mobile (e.g. 3G).
- Others listed on mailing list

Questions

- Did we pick a reasonable set of use cases?
- Should we add a non-commercial use case like campus, corporate, etc?



Next Steps

- Believe this is ready for adoption as first WG draft
- Lots of work already done
- Lots of discussion already on ML
  - Need to tweak layout
  - Might add more use cases from those suggested on mailing list
  - Expand “Other Issues” section
- Some questions remain – what is ConEx for, in a few words:
  - A way to reduce overall congestion?
  - A metric to improve capacity sharing?
  - A metric to allow better traffic management?
  - All the above and more?

Conclusions

- This draft describes some of the use cases for ConEx
- By no means exhaustive – this is a radical idea that will generate some truly innovative uses
- Included a brief description of a possible mechanism as readers need that to understand the use cases
- Congestion Volume is the key metric for controlling capacity sharing
- Introduced the ConEx Monitor and the ConEx Policer
- Highlighted several use cases, concentrated on 2 main ones

ConEx Concepts and Uses

spare slides

ConEx verifier

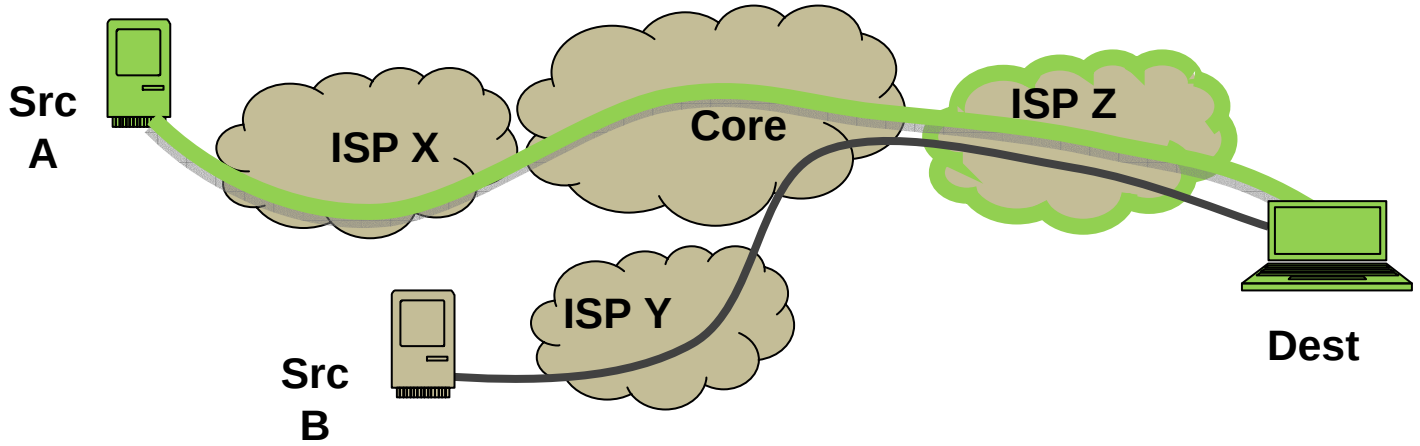
mediating between modern cc’s

frozen scenario slides

The following use Toby's original topology
I kept a copy before changing it
but they still show the other suggested changes

ConEx Use Cases Introduction

- Lots of use cases for ConEx
- Charter asks for use cases to focus on the following scenario:

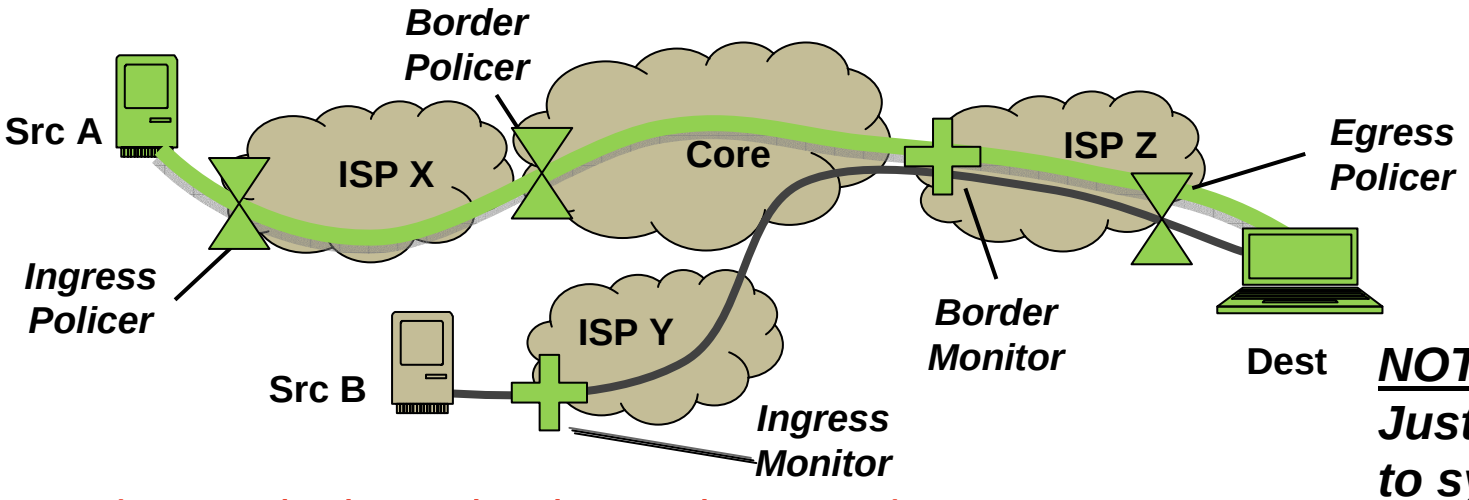


Green elements ConEx-Enabled. Grey elements not Enabled

- NB: the symmetry of most networks implies that **ISP Z** can be a ConEx-Enabled ***source*** network for any traffic that **Dest** sends into the network
- Following slides show 2 of the main use cases for ConEx
  - Traffic management
  - DDoS protection

ConEx Components

- Two new network components defined:
  - **ConEx Monitor** – a node that uses ConEx markings to measure/report the Congestion Volume that it forwards
  - **ConEx Policer** – A node that uses ConEx markings to change the queuing priority it gives a flow, or to actively control the Congestion Volume it forwards
- Policers and Monitors can be at Ingress, Egress or Border

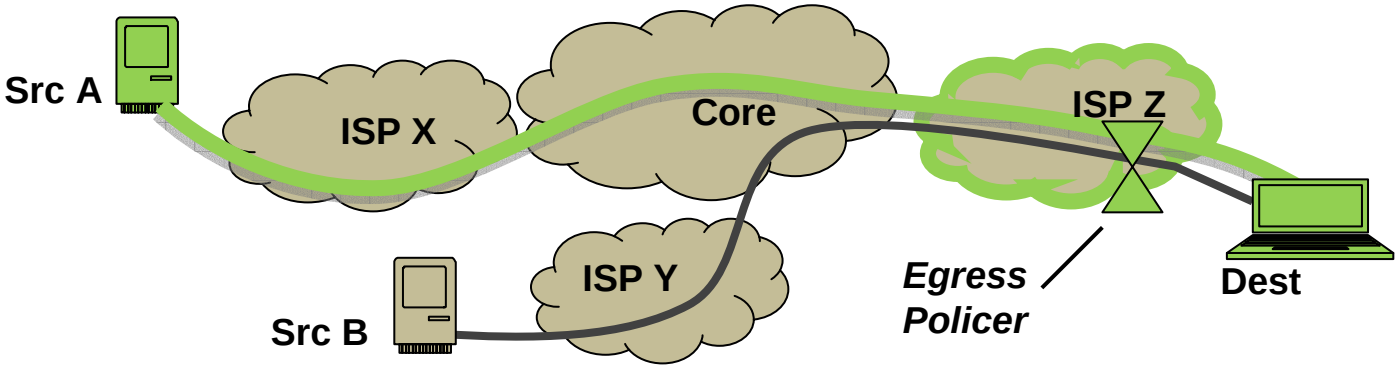


NOTE:
Just a key to symbols. Not a scenario!

- Border might have both a Policer and a Monitor
  - policing to prevent serious congestion
  - monitoring against a traffic contract to deter unnecessary congestion

Use Case 1a – Traffic Management

- ISPs often perform traffic management:
  - Aim is to give majority of users an adequate service at peak times
  - Users targeted based on application, traffic rate, volume transferred, etc
- ConEx policers offer an alternative:
  - Each sender is declaring the congestion they expect to cause
  - This can be used to control the impact they have on others
- Egress ConEx policer at the last IP node before backhaul/access can:
  - Identify the heaviest users – in terms of congestion-volume
  - Prioritise traffic depending on congestion it has declared
  - Penalise traffic that has caused excessive congestion

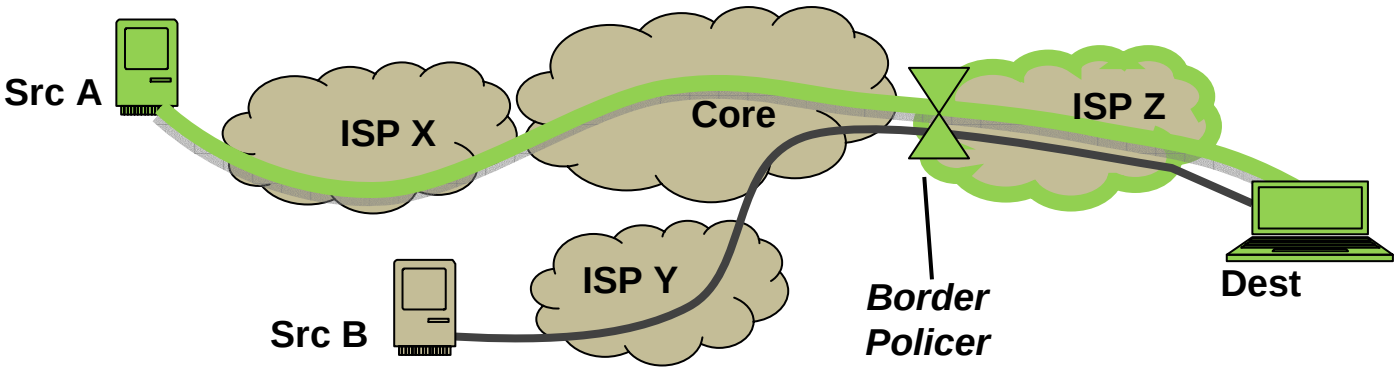


*Egress Policer can use ConEx info to prioritise traffic from **Src A**.*

*Traffic from **Src B** might be prioritised by volume/rate/app*

Use Case 2 – Raising the DDoS Bar

- DDoS is a serious problem – currently no robust solution
- ConEx Border Policers could help raise the bar
  - ConEx Policers naturally limit traffic rate towards congestion hot-spots
  - Policers can rate-limit non-ConEx traffic routing towards same hot-spot
- ConEx Border Monitors could help raise the bar too
  - ConEx DDoS traffic shows ultra-high congestion, so obvious anomaly at border
  - ConEx info closer to attack sources is a foundation for other solutions



- A lot more to DDoS use-case than one slide can do justice