Route Servers at IXPs – Bugs and Scaling issues with Quagga

Mike Hughes

CTO

London Internet Exchange

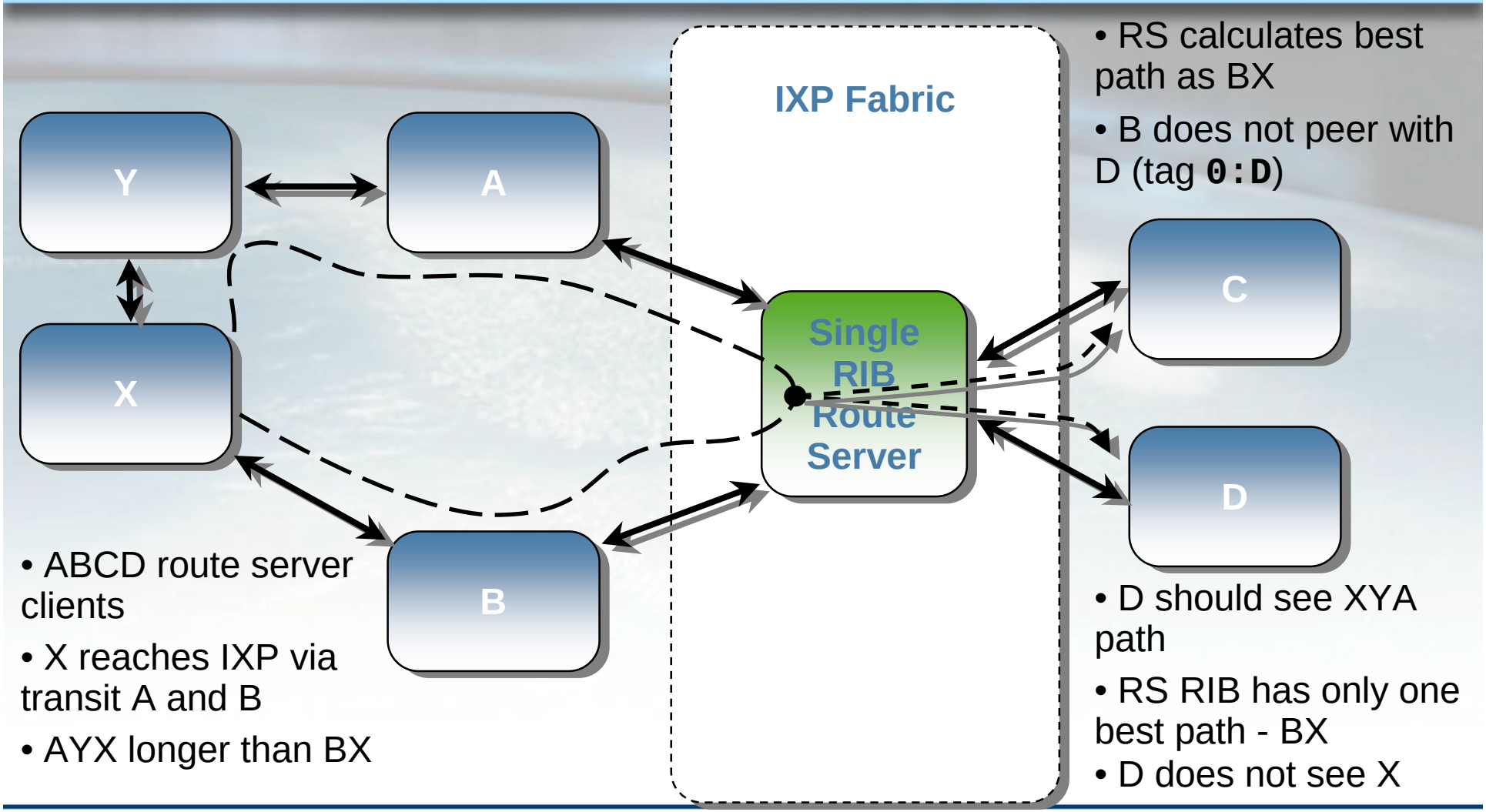
Agenda

- Requirements for IXP route servers
- Issues
- Stability Measures
- Alternatives
- Options
- Summing up

Why is Quagga used for IXP RS?

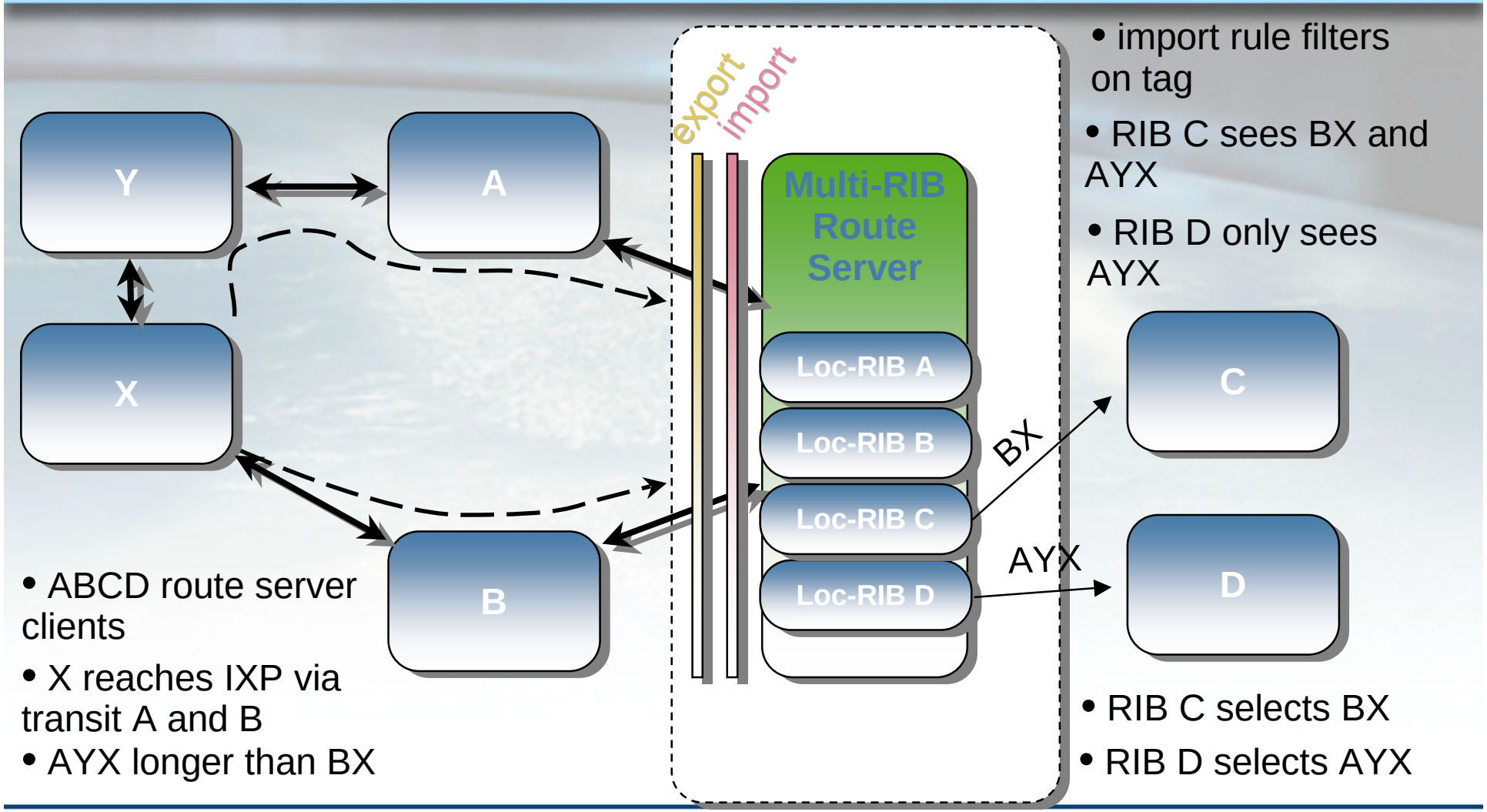
- Quagga forked from Zebra in 2003
- Open source, free, GPL
  - Free is good, right?
- Multiple LOC-RIB support
  - Maintains a view per client
  - Required by the different peering policies of participants at an IXP
- Limited viable alternatives

Single RIB Route Server



LINX

Per-Client RIB Route Server



- ABCD route server clients
- X reaches IXP via transit A and B
- AYX longer than BX

- import rule filters on tag
- RIB C sees BX and AYX
- RIB D only sees AYX

- RIB C selects BX
- RIB D selects AYX

Issues from Per-Client RIB use

- Increase in memory and CPU needs
  - No longer scales linearly with prefixes
  - But as a function of route server peers and prefixes per peer
- Quagga also single-threaded
  - Multiple cores per CPU won't help you
  - Can mitigate with separate v4/v6 daemons
- But works okay in small scenarios

How big is big? LINX scenario...

	Server	Prefixes	Configured Peers	Actual Peers
Foundry LAN v4	RS1	22231	206	152
	RS3	22366	210	142
Extreme LAN v4	RS2	9100	134	94
	RS4	9331	141	92
Foundry LAN v6	RS1	135	43	34
	RS3	127	37	30
Extreme LAN v6	RS2	82	21	13
	RS4	85	21	15

Problems Experienced

- Route server would crash when
  - max-prefix was hit
  - a peer was removed
- Race conditions and CPU hogging
  - During periods of high churn
  - During turn-up of large amounts of peers
- Mostly issues with rs-client code
  - Due to structure locking issues
- Hit by long 4-byte ASN path bug

Stability Measures

- Separating IPv4 and IPv4 daemons have improved stability
- Moving to Quagga 0.99.11 plus long AS-4 path patch has helped
- No crashes or CPU hog conditions on LINX route servers since this update
- Doubt Quagga 0.99.12 will be a panacea though – still more to do

Fixing Quagga

- Some interest in fixing Quagga issues
- Some fixes in 0.99.12
- Further useful patch contributed on 21st May – [quagga-dev 6600]
  - From a developer at Vyatta
  - Fixes some work queue locking
- But it won't fix itself…

Euro-IX Quagga Action

- Euro-IX IXPs are users of Quagga
- Now working on…
  - Implementable feature wishlist
  - Useful bug list and debugging environment
  - Active development by active developers
  - Useful bug reports
    - Tracebacks, coredumps, configs, etc.
  - Sourcing funding for above

Alternatives

- Free
  - BIRD – http://bird.network.cz/
    - Being actively developed with an IXP
  - OpenBGPd – http://www.openbgpd.org/
    - Missing per-client LOC-RIB support
    - Portability issues – part of OpenBSD suite
- Commercial
  - Typical router vendors: Cisco, Juniper, etc.
    - Cost of an unneeded forwarding plane
- Any others?

Options

- LINX are researching many options
  - Fixing Quagga
  - Viability of BIRD and others
  - Evaluating a commercial bespoke BGPd
  - Been approached by programmers saying "I can write you a BGPd"
- Whatever is chosen needs to have widespread appeal
  - Sustainability and peer support

Summing Up

- Sporadic development and maintenance of Quagga is a concern
- IXPs are the primary (only?) user of rs-client feature
- Recent renewed interest in fixing bugs among Quagga community
- Whatever is chosen needs to be widely deployed to ensure sustainable support

Thanks

- Nick Hilliard @ INEX
  - Single vs. Per-Client RIB diagrams, explanation of problem space
  - Euro-IX Quagga maintenance wishlist
- Questions?