Juniper BGP Convergence Performance

Richard A Steenbergen <ras@gt-t.net>

Problem Statement

- If you're like me, you may have noticed an ongoing issue with BGP convergence times on Juniper routers.

- Symptoms include:
    - Long periods of time (20+ minutes is not uncommon) where route installation completely "stalls", and the FIB is not updated.
    - Large numbers of BGP routes stuck in the "Pending" state.
    - Eventually followed by a sudden release of all the stuck routes.

- Many people have been seeing this for several years.
    - But it only seemed to affect "large" networks.
    - And Juniper has sworn up and down that they've been completely unable to reproduce or fix the issue.

Hallelujah

- We've finally identified and documented the issue!!!

Lab Environment

- To reproduce this issue, we set up the following:
  - Juniper MX480 with MPC uplinks, doing no traffic.
  - RE-S-1800x4-16G (1.8GHz Core 2 Quad RE w/16GB)
    - The current top of the line RE you can buy.
  - Participating in a decent sized IBGP mesh
    - 28 IBGP sessions across 4 BGP update groups.
    - Delivering 4.8 million paths to 434k IPv4 routes, 11k IPv6.
  - And 3 full tables towards simulated customers.
- Not even close to a real production network load.
  - No peering, no transit, minimal customers, simple policies.
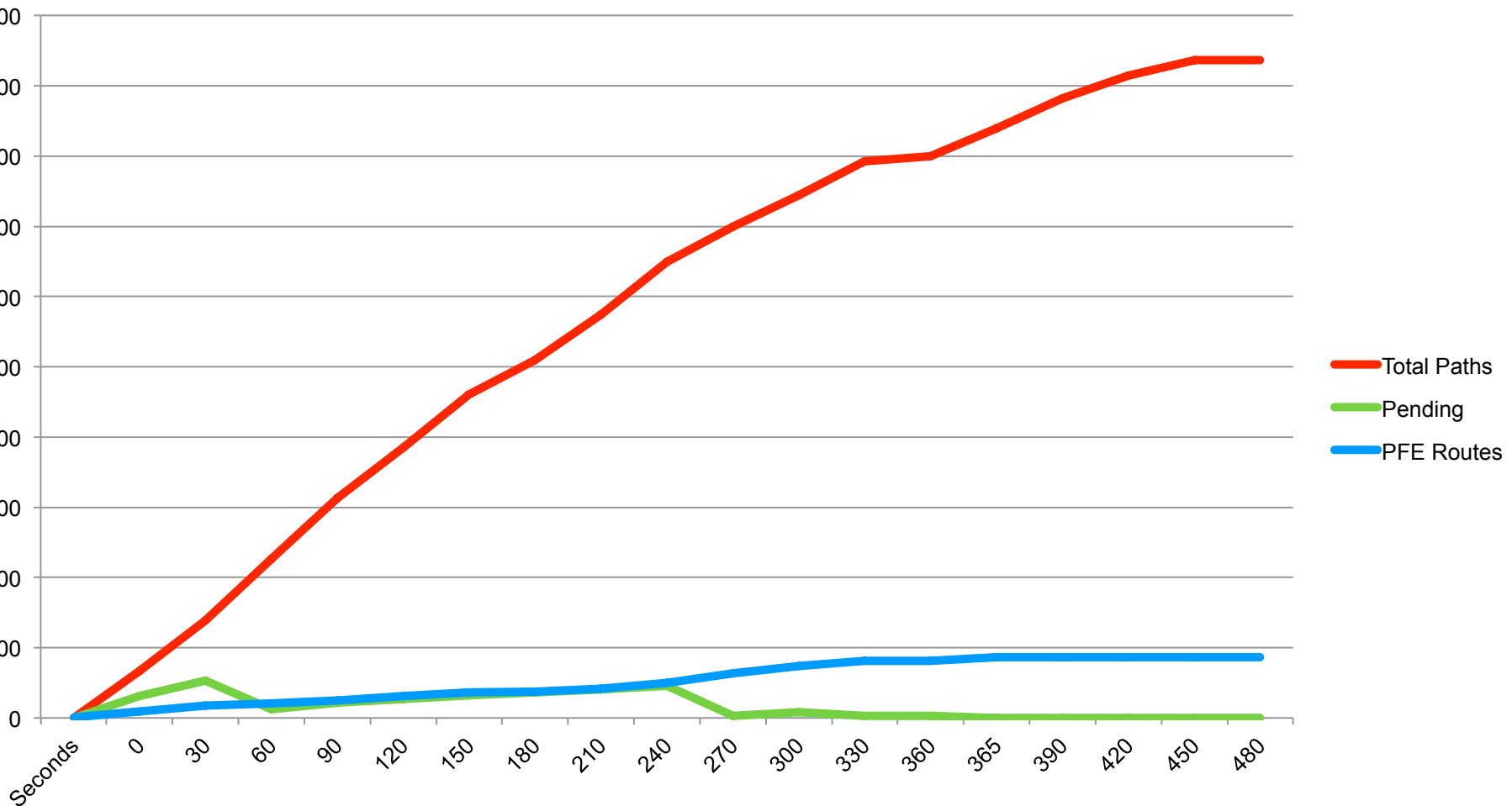 - But enough to study the issue in question.

Performance Under Stock Code

MX480 RE-S-1800x4 11.4R6.5
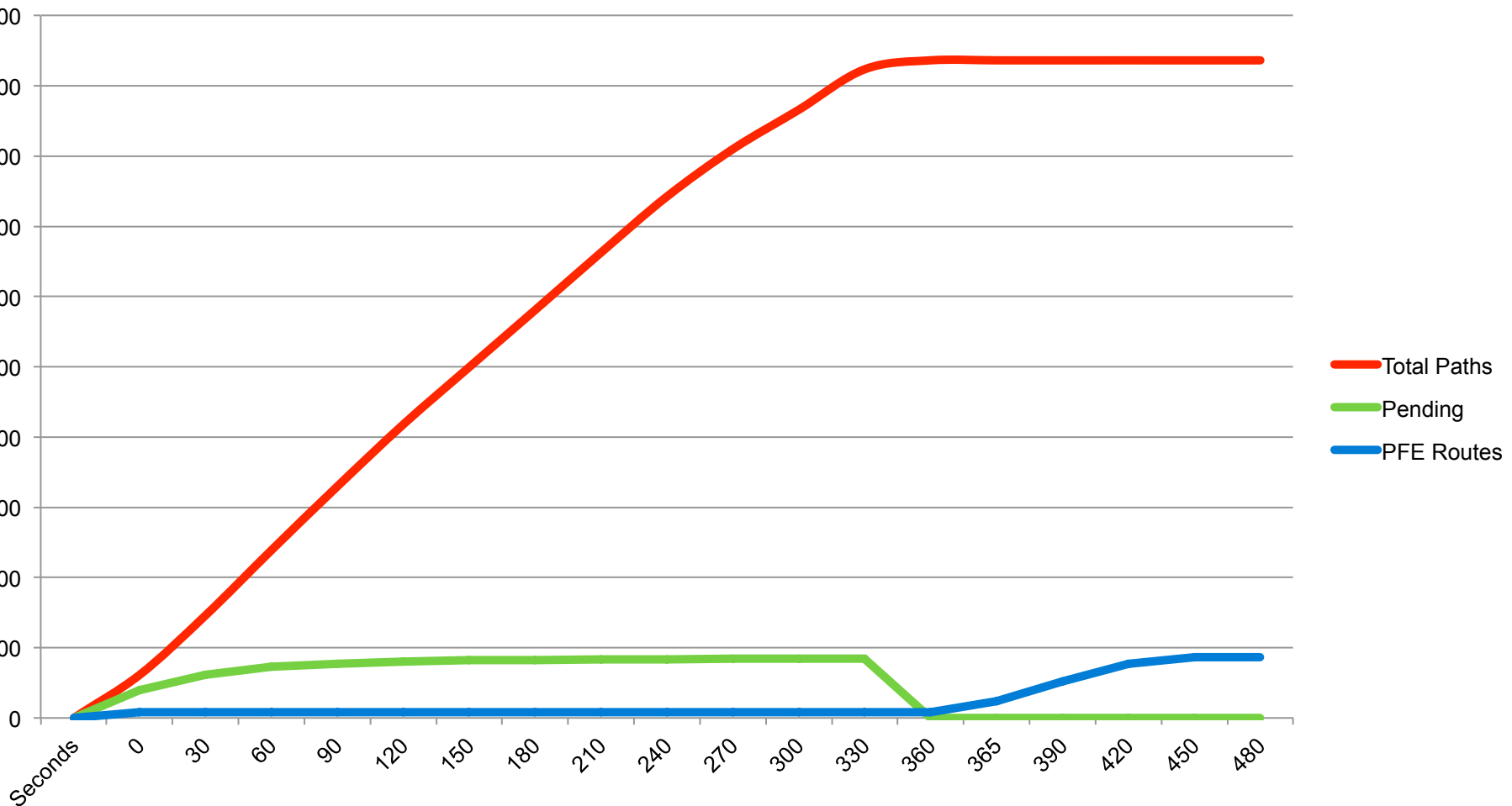
Performance With Proposed Fix

MX480 RE-S-1800x4 11.4I20121130_1932_jhaas

The Issue and Proposed Fix

- The Issue
    - Large amounts of BGP I/O (i.e. the receipt of lots of BGP updates) starves RPD of the ability to actually ACT on the routes that it has received.
    - After BGP routes are fully exchanged, routing updates can finally begin being serviced again, and the FIB installs.
- The Solution
    - Tweak the scheduler inside of RPD to prevent this type of starvation from occurring.
    - Total convergence time increases slightly, but at least you don't spend many minutes not installing any routes at all.

The Problem

- They don't actually want to fix it!
  - Even after finally documenting the issue that Juniper says they've been completely unable to reproduce for years, they still don't think that this is a serious issue.
  - I was told to go out and solicit feedback from other customers who think that not installing routes to the FIB for many minutes is actually a serious problem.

- If you use Juniper, go talk to your account team.
  - Mention PR836197, and tell them precisely what you think about the priority of resolving this issue.

Send questions, comments, complaints to:

Richard A Steenbergen   <ras@gt-t.net>