

Report:

NANOG 39 BOF

Pushing the FIB limits, perspectives on the pressures confronting modern routers.

Rational:

The IETF IAB held a workshop addressing the scalability of the global routing system (Routing and Addressing Workshop) and what if anything could be done to reduce the amount of state that must be maintained in order to participate in the global Internet routing infrastructure. The outcome of that workshop were some recommendations brought on by a concern that the evolution of hardware development might not be adequate to sustain the growth in the global routing system. Splitting the location and identification roles currently held by IP addresses looked to be a possible solution to perceived scalability problems. The report from the raws workshop can be found here:

<http://tools.ietf.org/id/draft-iab-raws-report-00.txt>

While network operators have long-term concerns which are probably best addressed through the evolution of Internet routing standards, they have a more immediate problem with existing router platforms. The growth of prefixes in the Default Free Zone is approaching the point where the previous generations routing platforms are not likely to be able to keep up. This BOF was intended to draw in vendors and operators to examine the time-line of when those routers would have to be replaced and if current or near future routers would have a reasonable life-span given Current growth projections

Outcomes:

Participants were solicited from vendors and from the operator community to offer their perspectives. In total, 5 participants presented with healthy participation from the audience on specific points. Presenting were:

Greg Hankins – FORCE10

Bill Ryan – Foundry

Suran De Silva – Cisco

John Schudder - Juniper

Ran Atkinson – Extreme

While operators were solicited to present, none were willing to do so. Information about the hardware used inside a particular ISP and problems that it may be presenting to them, or techniques they're using to solve problems are all highly proprietary information at this point.

Broadly speaking, Greg Hankins, Bill Ryan, and Suran De Silva focused on the efforts of their vendors

to deliver products that would meet the needs of operators over the next 5 or so years. In the case of Cisco that means delivering switch routers with a capacity of about a million routes now. In the case of Foundry they are projecting that with some FIB aggregation techniques that switches capable of 512k fib entries will still be usable by 2014. Juniper is delivering new products (m120 mx960) with DRAM rather than TCAM/SRAM based FIB's with capacities on the order of 2 million ipv4 routes and they have no reason to expect that they couldn't deliver 10 million route FIB products in a few years given sufficient demand.

John Scudder and Ran Atkinson expended more focus on the areas pointed out in the the raws workshop. One approach suggested by Juniper involved the "BGP free core" which would relieve pressure on core routers albeit at the expense of pushing it out to the edges of a network. Ran did a treatment of some of the proposals on the table as possible solutions to the geometric growth of the DFZ, LISP (draft-farinacci-lisp-00.txt), classic ID locator splits etc.

Participants from the crowd expressed concern about the performance of recalculating and installing the new FIB as it continues to grow, and also with vendors choosing not to offer FIB optimization tools in favor of delivering routers with much larger FIB's. Some participants expressed a desire for the ability to replace the forwarding engine or upgrade the TCAM on line cards independently. One participant observed that modern software based routers had increased in performance substantially and were a viable alternative in a number of low-end cases.

Conclusions:

The Global DFZ for ipv4 will exceed 220,000 routes shortly, for ipv6, 850. It appears that hardware vendors are prepared to deliver today, hardware that will be usable for 5 or more years given current models of growth. Some large service provider's currently have substantially more internal routes due to the size of their networks or the use of MPLS etc than they carry from the DFZ. Their internal needs are driving the acquisition of new hardware. In any event, router platforms supporting ~200k FIB entries are approaching the end of their useful lifespan and while heavily filtering or FIB optimization techniques can extend it, those are thought to be marginal solutions.

It seems unlikely that the pressures driving the growth of the ipv4 DFZ will abate anytime soon. There is some concern that eventually hardware performance will not be able to keep up with the growth of the DFZ or that growth in the amount of time required for routing to converge will result in less overall reliability. For those reasons, it is worth pursuing a more complete solution. We have time (a couple of years). A concern is that the BGP and the global routing system as a whole is quite mature, until the pain associated with changing to something else is lower than the pain associated with continuing as we are, there will be heavy resistance to change.