

Re: Solaris 10 IP Multipathing

Source: <http://unix.derkeiler.com/Newsgroups/comp.unix.solaris/2006-06/msg00146.html>

- *From:* "Colin B." <cbigam@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>
 - *Date:* Sat, 03 Jun 2006 09:58:07 GMT
-

Mike Jacobs <michael.jacobs@xxxxxxxxxxx> wrote:

Hello, I would appreciate anyone's comments on how IP Multipathing has worked out for you on Solaris 10. I work with a VoIP product running Solaris 8 and 9. Last year we gave IPMP a try and didn't have good success due to links being taken out of service because pings to the router failed. (We found that some routers when under a load quit responding to pings). I looked through this board and found that there are several posts of encountering the same type of issue.

We've had this problem in the past—our Nortel core routers are notorious for obscure CPU spikes which drop ICMP echo packets and lead to ICMP failovers inappropriately.

I have looked through the Sun docs for the Solaris 10 implementation of IPMP and it appears that they have addressed this issue in a couple of ways by allowing one to specify a "pinging" target list by setting up static routes as well as disable the pinging altogether by not specifying a test address.

Has this new implementation worked well for you or are there still some more "gotchas" with it?

This was our solution, and it works like a charm. At a fundamental level, all you're doing is adding static routes to hosts. There are a few caveats, but as long as you realise them, then it makes ICMP incredibly robust, and fails to the secondary only when the actual link goes down.

The caveats are:

- 1) All ping targets must be in the same subnet as the primary interface.
- 2) Adding the static routes causes them to immediately be included in the ping round-robin, up to a maximum of five.
- 3) Since the default router is used as the first target, you have to stop and restart ICMP to eliminate it from the list if you've added hosts 'live'. However, if static routes are added at boot time, default route will not be used as a ping target.

Re: Solaris 10 IP Multipathing

4) The rules about when failover occurs remain unchanged (ten consecutive ping failures, I believe). With the maximum of five targets, that means that each machine will have to fail to return two packets in a row in order to cause a failover—a nearly impossible feat, unless the interface has actually failed for some reason.

The only real downside to it is that it's something of an administrative headache, since you have to manage the targets individually for each machine. That said, it not only makes ICMP more robust, but it can substantially reduce the traffic to the routers.

After some testing, we rolled it out to our production clusters. Since then we've had router CPU spikes which caused failovers in 'default' configuration ICMP systems, but the clusters quietly recognised that it was the target that was the problem, not the NIC.

Oh yeah, final note—if you snoop for pings, you can watch all of this happen in realtime as you add and remove static routes. It gives a very good feeling for how things actually work.

Colin

.