

Re: [PATCH 0/3] netfilter : 3 patches to boost ip_tables performance

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2005-10/1845.html>

From: Andi Kleen (ak_at_suse.de)

Date: 10/07/05

To: Patrick McHardy <kaber@trash.net>

Date: Fri, 7 Oct 2005 19:21:39 +0200

On Friday 07 October 2005 19:08, Patrick McHardy wrote:

> *There are lots of other hooks and conntrack/NAT already have a
> quite large negative influence on performance. Do you have numbers
> that show that enabling this actually causes more than a slight
> decrease in performance? Besides, most distributors enable all
> these options anyway, so it only makes a difference for a small
> group of users.*

I don't know about other distributions but SUSE at some point found that some web benchmarks dramatically improved in the default configuration when local conntrack was off. It was off then since ever.

> > *Perhaps there would be other ways to fix this problem without impacting
> > performance unduly? Can you describe it in detail?*
>
> *When an ICMP error is send by the firewall itself, the inner
> packet needs to be restored to its original state. That means
> both DNAT and SNAT which might have been applied need to be
> reversed. DNAT is reversed at places where we usually do
> SNAT (POST_ROUTING), SNAT is reversed where usually DNAT is
> done (PRE_ROUTING/LOCAL_OUT). Since locally generated packets
> never go through PRE_ROUTING, it is done in LOCAL_OUT, which
> required enabling NAT in LOCAL_OUT unconditionally. It might
> be possible to move this to some different hook, I didn't
> investigate it.*

This sounds wrong anyways. You shouldn't be touching conntrack state for ICMPs generated by routers because they can be temporary errors (e.g. during a routing flap when the route moves). Only safe way to handle this is to wait for the timeout which doesn't need local handling. And the firewall cannot be an endhost here.

-Andi

Linux-Kernel: Re: [PATCH 0/3] netfilter : 3 patches to boost ip_tables performance

—

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to majordomo@vger.kernel.org

More majordomo info at <http://vger.kernel.org/majordomo-info.html>

Please read the FAQ at <http://www.tux.org/lkml/>