Subject: Re: [PATCH 2.6.21-rc6] [netfilter] early_drop imrovement
Posted by vaverin on Fri, 06 Apr 2007 10:26:02 GMT
View Forum Message <> Reply to Message

Eric Dumazet wrote:
> On Fri, 06 Apr 2007 12:00:29 +0400
> Vasily Averin <vvs@sw.ru> wrote:
>
>> When the number of conntracks is reached ip_conntrack_max limit, early_drop() is
>> called and tries to free one of already used conntracks in one of the hash
>> buckets. If it does not find any conntracks that may be freed, it
>> leads to transmission errors.
>> However it is not fair because of current hash bucket may be empty but the
>> neighbour ones can have the number of conntracks that can be freed. With the
>> following patch early_drop() will search conntracks in all hash buckets.
>
> Have you tested your patch in a DOS situation ?
> Some machines have a huge ip_conntrack_max.
> A single scan of the whole table might take 1000 ms or even more.

No, I've not investigated this scenario. It looks like you are right and my
patch can leads to a long delays.

In my experiments I've decreased ip_conntrack_max lower than number of hash
buckets and got the "table full, dropping packet" errors in logs. I've looked on
the conntrack list and found a huge number of conntracks that can be freed.
However my hash bucket was empty and therefore I even did not have any chances
to free something. That's why I would like to check the other hash buckets too.

Ok, let's limit the number of conntracks that can be checked inside
early_drop(). What do you prefer: some round number (for example 100) or
fraction of ip_conntrack_max (for example 1%)?

Thank you,
 Vasily Averin