Subject: Re: [NETFILTER] early_drop() imrovement (v3)
Posted by Patrick McHardy on Tue, 26 Jun 2007 13:27:30 GMT
View Forum Message <> Reply to Message

Vasily Averin wrote:
> Patrick McHardy wrote:
>> I don't like the NF_CT_PER_BUCKET constant. First of all, each
>> conntrack is hashed twice, so its really only 1/2 of the average
>> conntracks per bucket. Secondly, its only a default and many
>> people use nf_conntrack_max = nf_conntrack_htable_size / 2, so
>> using this constant for early_drop seems wrong.
>>
>> Perhaps make it 2 * nf_conntrack_max / nf_conntrack_htable_size
>> or even add a nf_conntrack_eviction_range sysctl.
>>
>
> IMHO The number of conntracks checked in early_drop() have following restrictions:
> - it should be not too low -- to decrease chances of transmission failures,
> - it should be limited by some reasonable value -- to prevent long check delays.

Agreed.

> Also I believe it makes sense to have it constant (how about NF_CT_EVICTION
> name?) -- to have the same behaviour on various nodes. However I doubt strongly
> that anybody will want to change this value. Do you think it is really required?
>

I don't know. The current behaviour will on average scan 16 entries.
For people manually tuning their hash to saner settings it will scan
a single entry. So we have a quite wide range of values already.
The single entry with sane hash settings is too little IMO, maybe use
some middle-ground, make it 8 by default as you did and rename the
constant. NF_CT_EVICTION_RANGE sounds fine.