Subject: [PATCH 12/12 net-2.6.25] [NETNS]: Add namespace for ICMP replying code.
Posted by den on Wed, 23 Jan 2008 07:46:27 GMT
View Forum Message <> Reply to Message

All needed API is done, the namespace is available when required from the
device on the DST entry from the incoming packet. So, just replace init_net
with proper namespace.

Other protocols will follow.

Signed-off-by: Denis V. Lunev <den@openvz.org>
---
 net/ipv4/icmp.c | 21 ++++++++++++++--------
 net/ipv4/ip_output.c | 2 +-
 2 files changed, 14 insertions(+), 9 deletions(-)

```
diff --git a/net/ipv4/icmp.c b/net/ipv4/icmp.c
index 052b278..a6c092c 100644
--- a/net/ipv4/icmp.c
+++ b/net/ipv4/icmp.c
@@ -404,7 +404,7 @@ static void icmp_reply(struct icmp_bxm *icmp_param, struct sk_buff
*skb)
        .tos = RT_TOS(ip_hdr(skb)->tos) } },
          .proto = IPPROTO_ICMP };
    security_skb_classify_flow(skb, &fl);
-   if (ip_route_output_key(&init_net, &rt, &fl))
+   if (ip_route_output_key(rt->u.dst.dev->nd_net, &rt, &fl))
      goto out_unlock;
   }
   if (icmpv4_xrlim_allow(rt, icmp_param->data.icmph.type,
@@ -436,9 +436,11 @@ void icmp_send(struct sk_buff *skb_in, int type, int code, __be32 info)
   struct ipcm_cookie ipc;
   __be32 saddr;
   u8  tos;
+  struct net *net;

   if (!rt)
     goto out;
+  net = rt->u.dst.dev->nd_net;

   /*
    * Find the original header. It is expected to be valid, of course.
@@ -514,7 +516,7 @@ void icmp_send(struct sk_buff *skb_in, int type, int code, __be32 info)
    struct net_device *dev = NULL;

    if (rt->fl.iif && sysctl_icmp_errors_use_inbound_ifaddr)
-     dev = dev_get_by_index(&init_net, rt->fl.iif);
```

```
+   dev = dev_get_by_index(net, rt->fl.iif);

   if (dev) {
    saddr = inet_select_addr(dev, 0, RT_SCOPE_LINK);
@@ -569,7 +571,7 @@ void icmp_send(struct sk_buff *skb_in, int type, int code, __be32 info)
   struct rtable *rt2;

   security_skb_classify_flow(skb_in, &fl);
-  if (__ip_route_output_key(&init_net, &rt, &fl))
+  if (__ip_route_output_key(net, &rt, &fl))
    goto out_unlock;

   /* No need to clone since we're just using its address. */
@@ -591,14 +593,14 @@ void icmp_send(struct sk_buff *skb_in, int type, int code, __be32 info)
   if (xfrm_decode_session_reverse(skb_in, &fl, AF_INET))
    goto out_unlock;

-  if (inet_addr_type(&init_net, fl.fl4_src) == RTN_LOCAL)
-   err = __ip_route_output_key(&init_net, &rt2, &fl);
+  if (inet_addr_type(net, fl.fl4_src) == RTN_LOCAL)
+   err = __ip_route_output_key(net, &rt2, &fl);
   else {
    struct flowi fl2 = {};
    struct dst_entry *odst;

    fl2.fl4_dst = fl.fl4_src;
-   if (ip_route_output_key(&init_net, &rt2, &fl2))
+   if (ip_route_output_key(net, &rt2, &fl2))
     goto out_unlock;

    /* Ugh! */
@@ -666,6 +668,9 @@ static void icmp_unreach(struct sk_buff *skb)
  int hash, protocol;
  struct net_protocol *ipprot;
  u32 info = 0;
+ struct net *net;
+
+ net = skb->dst->dev->nd_net;

  /*
   * Incomplete header ?
@@ -696,7 +701,7 @@ static void icmp_unreach(struct sk_buff *skb)
         "and DF set.\n",
           NIPQUAD(iph->daddr));
    } else {
-    info = ip_rt_frag_needed(&init_net, iph,
+    info = ip_rt_frag_needed(net, iph,
          ntohs(icmph->un.frag.mtu));
```

```
     if (!info)
      goto out;
@@ -734,7 +739,7 @@ static void icmp_unreach(struct sk_buff *skb)
   */

  if (!sysctl_icmp_ignore_bogus_error_responses &&
-     inet_addr_type(&init_net, iph->daddr) == RTN_BROADCAST) {
+     inet_addr_type(net, iph->daddr) == RTN_BROADCAST) {
    if (net_ratelimit())
     printk(KERN_WARNING "%u.%u.%u.%u sent an invalid ICMP "
           "type %u, code %u "
diff --git a/net/ipv4/ip_output.c b/net/ipv4/ip_output.c
index 6a5b839..4fad239 100644
--- a/net/ipv4/ip_output.c
+++ b/net/ipv4/ip_output.c
@@ -1377,7 +1377,7 @@ void ip_send_reply(struct sock *sk, struct sk_buff *skb, struct
ip_reply_arg *ar
         .dport = tcp_hdr(skb)->source } },
         .proto = sk->sk_protocol };
    security_skb_classify_flow(skb, &fl);
-   if (ip_route_output_key(&init_net, &rt, &fl))
+   if (ip_route_output_key(sk->sk_net, &rt, &fl))
     return;
  }

--
1.5.3.rc5
```