Subject: [patch 03/12] net namespace : share network ressources L2 with L3
Posted by Daniel Lezcano on Fri, 19 Jan 2007 15:47:17 GMT
View Forum Message <> Reply to Message

From: Daniel Lezcano <dlezcano@fr.ibm.com>

L3 namespace will use routes and devices belonging to its parent, so
the old network namespace structure is copied when allocating a new
one. By this way, hash value, dev list, routes are accessible from the
L3 namespaces. In case of L2 namespace, these values are overwritten
by the newly allocated values.

Signed-off-by: Daniel Lezcano <dlezcano@fr.ibm.com>

---
 include/linux/net_namespace.h | 14 +++++++++++++
 net/core/dev.c | 4 ++--
 net/core/net_namespace.c | 33 ++++++++++++++++++++++++++++++---------
 3 files changed, 34 insertions(+), 17 deletions(-)

Index: 2.6.20-rc4-mm1/net/core/net_namespace.c
===
--- 2.6.20-rc4-mm1.orig/net/core/net_namespace.c
+++ 2.6.20-rc4-mm1/net/core/net_namespace.c
@@ -37,7 +37,7 @@
  * Return ERR_PTR on error, new ns otherwise
  */
 static struct net_namespace *clone_net_ns(unsigned int level,
-      struct net_namespace *old_ns)
+   struct net_namespace *old_ns)
 {
  struct net_namespace *ns;

@@ -45,23 +45,26 @@
  if (current_net_ns->level == NET_NS_LEVEL3)
   return ERR_PTR(-EPERM);

- ns = kzalloc(sizeof(struct net_namespace), GFP_KERNEL);
+ ns = kmemdup(old_ns, sizeof(struct net_namespace), GFP_KERNEL);
  if (!ns)
   return NULL;

  kref_init(&ns->kref);
- ns->dev_base_p = NULL;
- ns->dev_tail_p = &ns->dev_base_p;
- ns->hash = net_random();
-
  if ((push_net_ns(ns)) != old_ns)

```
+
   BUG();
  if (level ==  NET_NS_LEVEL2) {
+   ns->dev_base_p = NULL;
+   ns->dev_tail_p = &ns->dev_base_p;
+   ns->hash = net_random();
+
 #ifdef CONFIG_IP_MULTIPLE_TABLES
    INIT_LIST_HEAD(&ns->fib_rules_ops_list);
 #endif
    if (ip_fib_struct_init())
     goto out_fib4;
+   if (loopback_init())
+    goto out_loopback;
  }

  if (level == NET_NS_LEVEL3) {
@@ -70,8 +73,6 @@
  }

  ns->level = level;
- if (loopback_init())
-  goto out_loopback;
  pop_net_ns(old_ns);
  printk(KERN_DEBUG "NET_NS: created new netcontext %p, level %u, "
   "for %s (pid=%d)\n", ns, (ns->level == NET_NS_LEVEL2) ?
@@ -127,15 +128,17 @@
  struct net_namespace *ns;

  ns = container_of(kref, struct net_namespace, kref);
- unregister_netdev(ns->loopback_dev_p);
- if (ns->dev_base_p != NULL) {
-  printk("NET_NS: BUG: namespace %p has devices! ref %d\n",
-    ns, atomic_read(&ns->kref.refcount));
-  return;
- }

- if (ns->level == NET_NS_LEVEL2)
+ if (ns->level == NET_NS_LEVEL2) {
   ip_fib_struct_cleanup(ns);
+   unregister_netdev(ns->loopback_dev_p);
+   if (ns->dev_base_p != NULL) {
+    printk("NET_NS: BUG: namespace %p has devices! ref %d\n",
+           ns, atomic_read(&ns->kref.refcount));
+    return;
+   }
+ }
+
```

```
  if (ns->level == NET_NS_LEVEL3)
   put_net_ns(ns->parent);

Index: 2.6.20-rc4-mm1/include/linux/net_namespace.h
===================================================================
--- 2.6.20-rc4-mm1.orig/include/linux/net_namespace.h
+++ 2.6.20-rc4-mm1/include/linux/net_namespace.h
@@ -56,6 +56,15 @@
 DECLARE_PER_CPU(struct net_namespace *, exec_net_ns);
 #define current_net_ns  (__get_cpu_var(exec_net_ns))

+static inline struct net_namespace *net_ns_l2(void)
+{
+ struct net_namespace *net_ns = current_net_ns;
+
+ if (net_ns->level == NET_NS_LEVEL3)
+  return net_ns->parent;
+ return net_ns;
+}
+
 static inline void init_current_net_ns(int cpu)
 {
  get_net_ns(&init_net_ns);
@@ -110,6 +119,11 @@

 #define current_net_ns   NULL

+static inline struct net_namespace *net_ns_l2(void)
+{
+ return NULL;
+}
+
 static inline void init_current_net_ns(int cpu)
 {
 }
Index: 2.6.20-rc4-mm1/net/core/dev.c
===================================================================
--- 2.6.20-rc4-mm1.orig/net/core/dev.c
+++ 2.6.20-rc4-mm1/net/core/dev.c
@@ -485,7 +485,7 @@
 struct net_device *__dev_get_by_name(const char *name)
 {
  struct hlist_node *p;
- struct net_namespace *ns = current_net_ns;
+ struct net_namespace *ns = net_ns_l2();

  hlist_for_each(p, dev_name_hash(name, ns)) {
   struct net_device *dev
```

```
@@ -768,7 +768,7 @@
  if (!err) {
   hlist_del(&dev->name_hlist);
   hlist_add_head(&dev->name_hlist, dev_name_hash(dev->name,
-       current_net_ns));
+              net_ns_l2()));
   raw_notifier_call_chain(&netdev_chain,
     NETDEV_CHANGENAME, dev);
  }

--
```

_____

Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers