Subject: [patch 00/12] net namespace : L3 namespace - introduction
Posted by Daniel Lezcano on Fri, 19 Jan 2007 15:47:14 GMT
View Forum Message <> Reply to Message

This patchset provide a network isolation similar at what
Linux-Vserver provides. It is based on the L2 namespaces and relies on
the mechanisms provided by the namespace. This L3 namespaces does not
aim to bring full virtualization for the network, it provides an IP
isolation which can be reused for Linux-Vserver, jailed application or
application containers.

A L3 namespace are always L2 s' childs and they can not create more
network namespaces, furthermore, they lose their NET_ADMIN
capability. They share their parent's network ressources. From the
parent namespace, IP addresses are created and assigned to the
different L3 childs. From this point, L3 namespaces can use their
assigned IP address and all computed broadcast addresses.

Because the L3 namespace relies on the L2 virtualization mechanisms,
it is possible to have several L3 namespaces listening on
INADDR_ANY:port without conflict, that's allow to run several server
without modifying the network configuration.

The loopback is a shared device between all L3 namespaces. To ensure
the 127.0.0.1 address isolation, the sender store its namespace into
the packet, so when the packet arrives, the destination namespace is
already set, because "source" == "destination". By this way, it is
easy to disable the loopback isolation and let the application to talk
with application outside of the namespace via the 127.0.0.1 because we
consider them trusted (like portmap).

The ifconfig / ip commands will only show IP addresses assigned to the
L3 namespace. When a L3 namespace dies, the assigned IP address is
released to its parent.

At the IP level, when a packet arrives, the L3 network namespace
destination is retrieved from the destination address.

At the bind time, the address is checked against the assigned IP
address.

--