Contributions to the Linux Networking Subsystem

Vlad Dogaru <ddvlad@rosedu.org>
Supervisor: As. Dr. Ing. Răzvan Deaconescu

July 3, 2012

Computer Science
& Engineering
Department

Agenda

Network Interface Grouping

- Introduction
- Problem Factors
- Fixing the Problem
- Results and Further Development

Multipath TCP in Mobile Devices

- Multipath TCP Design
- Implementation
- Android Port

Computer Science
& Engineering
Department

Introduction

Problem Statement

- Linux kernel
- Large number of network devices (2048 used consistently for testing)
- Operations on these devices take a long time

Introduction

Problem Statement

- Linux kernel
- Large number of network devices (2048 used consistently for testing)
- Operations on these devices take a long time

Solution

- Group interfaces together
- Manipulate groups instead of interfaces

Problem Factors

```
ip link set dev eth0 down
```

1. create `iproute2` process
2. create Netlink socket
3. construct and send message
4. wait for response
5. close socket
6. terminate process

Problem Factors

```
ip link set dev eth0 down
```

1. create `iproute2` process
2. create Netlink socket
3. construct and send message
4. wait for response
5. close socket
6. terminate process

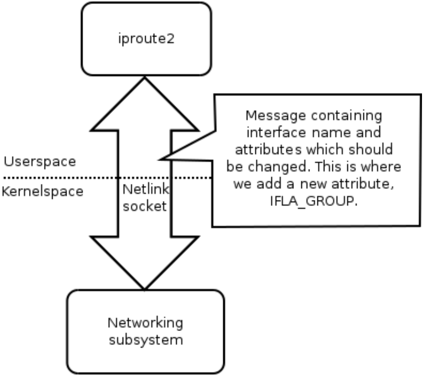
For each device

Device Groups

- Add field `group` to `struct net_device`
- Each device belongs to one group
- New devices are added in group 0 (*default*)
- Userspace can specify that an operation must apply to an entire group

Netlink Extension

- Add new message type: “change device group”
- Add new logic to existing messages: “this operates on an entire group”



Message containing interface name and attributes which should be changed. This is where we add a new attribute, IFLA_GROUP.

Userspace Modifications

Add `group` keyword to `iproute2`

- Change the group a device belongs to
  - `ip link set group 100 dev eth0`
- Change properties for an entire group
  - `ip link set group 100 mtu 1400`
  - `ip link set group 100 hwaddr 00:aa:bb:cc:dd:ee`
  - `ip address add 10.0.0.1/24 group ppp`

Computer Science
& Engineering
Department

Testing and Results

- KVM used to run a minimal Linux instance
- Change Maximum Transfer Unit of devices

Testing and Results

- KVM used to run a minimal Linux instance
- Change Maximum Transfer Unit of devices

	Without sysfs		With sysfs	
Interfaces	Group	No group	Group	No group
128	0.01	0.49	0.01	0.53
256	0.02	1.11	0.03	1.15
512	0.05	2.57	0.06	2.59
1024	0.17	7.02	0.20	7.51
2048	0.32	21.74	0.36	23.05

All times are in seconds.

Future Development

- Features integrated in upstream kernel
- Netfilter support has already been added
- Batching notification messages

Computer Science
& Engineering
Department

Multipath TCP in Mobile Devices

Computer Science
& Engineering
Department

MPTCP Design Goals

Better TCP support for multihomed devices (multi-NIC servers, mobile devices).

- **Redundancy**. Deliver all data as long as at least one connection is alive.
- **Throughput**. Be at least as efficient as TCP.
- **Backwards compatibility**. Fall back to TCP if problems arise.

MPTCP Operation

- TCP: `(protocol, local-addr, local-process, foreign-addr, foreign-process)`
- MPTCP: call that a *subflow*, use more than one subflow per multipath connection.
- Use TCP Options field to signal MPTCP capability.
- Subflow (TCP) -level sequence numbers are not enough, renumber segments with a flow-aware number.

Computer Science
& Engineering
Department

Linux Kernel Implementation

- Ongoing effort at UC Louvain
- Out-of-tree, small community
- Huge progress
- Binaries, test setup provided
- Kernel-like development model

MPTCP Goes Mobile

- Android operating system (infringes patents, etc.)
- Typical handset has wireless and 3G
- Use cases:
  - interface failover
  - interface preference
  - power saving

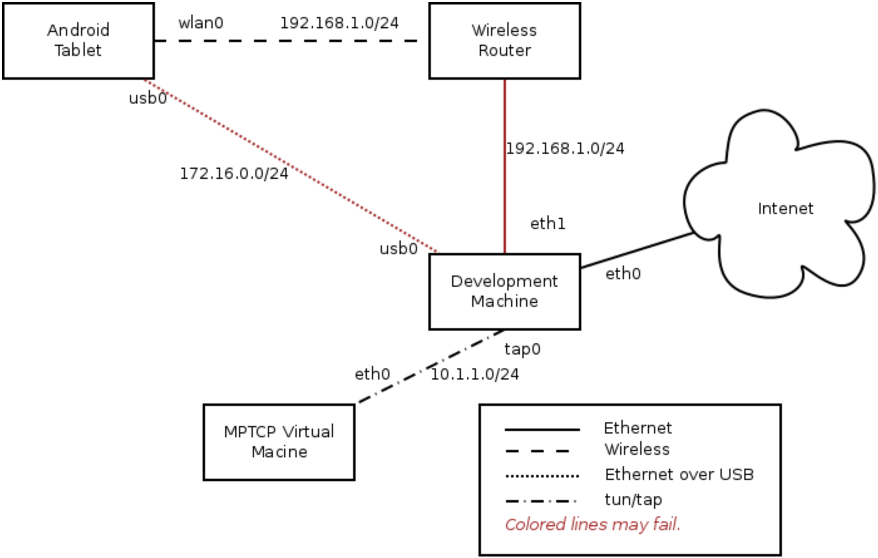
Computer Science
& Engineering
Department

Android Code Rebase

Task: rebase over 1400 patches to newer, modified kernel.
Challenges:

- Broken patch context
- Semantic changes in upstream
- Android-specific changes
- MPTCP tree was initially forked from different protocol (SHIM)

Testing Setup



Results

- All patches ported, semantics preserved
- Userspace changes by Andrei Maruseac
- Live demo on tablet, using SSH protocol and simulating network failures.

Computer Science
& Engineering
Department

Thank you!

Linux kernel, networking, interface grouping, Netlink, iproute2
Multipath TCP, (sub)flow, transparent fallback, mobile devices