# CONTENTS

1 Introduction ........................................ 3
   1.1 Overview ...................................... 3
   1.2 Manual Structure ................................ 3
   1.3 DCE Outlook .................................. 3
   1.4 Supported Features ............................... 3
   1.5 Tested Applications .............................. 4
   1.6 Tested Environment ............................... 4

2 Quick Start ........................................ 5
   2.1 Introduction .................................... 5
   2.2 Build DCE ...................................... 5
   2.3 Examples ....................................... 6

3 User’s Guide ....................................... 9
   3.1 Setup Guide .................................... 9
   3.2 Basic Use Cases ................................. 10
   3.3 Advanced Use Cases ......................... 26
   3.4 Technical Information ....................... 32
   3.5 DCE Python Scripts ............................ 37

4 Developer’s Guide ................................ 39
   4.1 Kernel Developer Information ................. 39
   4.2 DCE - POLL IMPLEMENTATION ............... 49
   4.3 Python Bindings ............................... 51

5 How It Works ..................................... 53
   5.1 Introduction ................................... 53
   5.2 Main classes and main data structures ....... 53
   5.3 Follow a very simple example ................ 55

6 Subprojects of DCE .............................. 59
   6.1 CCNx .......................................... 59
   6.2 Quagga ......................................... 59
   6.3 iperf ........................................... 59
   6.4 ping/ping6 ..................................... 60
   6.5 ip (iproute2 package) ......................... 60
   6.6 umip (Mobilt IPv6 daemon) .................... 60
   6.7 Linux kernel (from 2.6.36 to 3.14 version) . 60
   6.8 FreeBSD kernel (10.0.0 version) ............. 60
   6.9 thttpd ........................................ 60
   6.10 torrent ........................................ 60
This is the manual of Direct Code Execution (DCE).

- **Doxygen**: Documentation of the public APIs of the DCE
- **Manual**(this document) for the latest release and development tree

This document is written in reStructuredText for Sphinx and is maintained in the doc/ directory of ns-3-dce’s source code.
1.1 Overview

Direct Code Execution (DCE) is a module for ns-3 that provides facilities to execute, within ns-3, existing implementations of userspace and kernelspace network protocols or applications without source code changes. For example instead of using the pseudo application provided by ns-3 V4PingHelper you can use the true ping.

1.2 Manual Structure

This document consists of the following parts:

0. *Quick Start:* The document describes a quick instruction of DCE.


1.3 DCE Outlook

• To run an application using DCE, it is not necessary to change its sources. However you will need to recompile them.

• The simulation is executed wholly within a single process which greatly facilitates the debugging.

• DCE is very memory-efficient, thanks to the way it loads the executables similarly to shared libraries.

1.4 Supported Features

• Simulation with POSIX socket application (no manual modifications)

• C/C++ applications

• Simulation with Linux kernel implemented network protocol

• IPv4/IPv6

• TCP/UDP/DCCP

• running with POSIX socket applications and ns-3 socket applications

• configuration via sysctl-like interface

• multiple nodes debugging with single gdb interface
- memory analysis by single valgrind execution with multiple nodes
- Variance of network stacks
- ns-3 native stack (IPv4/IPv6, partially)
- Network simulation cradle network stack (IPv4 TCP only)
- Linux network stack (IPv4/IPv6/others)
- Per-node configuration/stdin input
- Per-node syslog/stdout/stderr files output

1.5 Tested Applications

- CCNx
- Quagga
- iperf
- ping/ping6
- ip (iproute2 package)
- Mobilt IPv6 daemon (umip)
- Linux kernel (from 2.6.36 to 3.7 versions)
- http server (thttpd)
- torrent (libtorrent from rasterbar + opentracker)

1.6 Tested Environment

Currently, DCE only supports Linux-based operating system. DCE has been tested on the following distributions:

- Ubuntu 10.04 64bit
- Ubuntu 12.04 32bit/64bit
- Ubuntu 12.10 64bit
- Ubuntu 13.04 64bit
- Ubuntu 13.10 64bit (new)
- Fedora 18 32bit
- CentOS 6.2 64bit

but you can try on the others (e.g., CentOS, RHEL). If you got run on another distribution, please let us know.
2.1 Introduction

The DCE ns-3 module provides facilities to execute within ns-3 existing implementations of userspace and kernelspace network protocols.

As of today, the Quagga routing protocol implementation, the CCNx CCN implementation, and recent versions of the Linux kernel network stack are known to run within DCE, hence allowing network protocol experimenters and researchers to use the unmodified implementation of their protocols for real-world deployments and simulations.

2.2 Build DCE

DCE offers two major modes of operation:

1. The basic mode, where DCE use the ns-3 TCP stacks,
2. The advanced mode, where DCE uses a Linux network stack instead.

2.2.1 Building DCE basic mode

First you need to download Bake using Mercurial and set some variables:

```
hg clone http://code.nsnam.org/bake bake
export BAKE_HOME=`pwd`/bake
export PATH=$PATH:$BAKE_HOME
export PYTHONPATH=$PYTHONPATH:$BAKE_HOME
```

then you must to create a directory for DCE and install it using bake:

```
mkdir dce
cd dce
bake.py configure -e dce-ns3-|version|
bake.py download
bake.py build
```

note that dce-ns3-1.9 is the DCE version 1.9 module. If you would like to use the development version of DCE module, you can specify `dce-ns3-dev` as a module name for bake.

the output should look likes this:
Installing selected module and dependencies. Please, be patient, this may take a while!

>> Downloading ccnx
>> Download ccnx - OK
>> Downloading iperf
>> Download iperf - OK
>> Downloading ns-3-dev-dce
>> Download ns-3-dev-dce - OK
>> Downloading dce-ns3
>> Download dce-ns3 - OK
>> Building ccnx
>> Built ccnx - OK
>> Building iperf
>> Built iperf - OK
>> Building ns-3-dev-dce
>> Built ns-3-dev-dce - OK
>> Building dce-ns3
>> Built dce-ns3 - OK

2.2.2 Building DCE advanced mode (with Linux kernel)

If you would like to try Linux network stack instead of ns-3 network stack, you can try the advanced mode. The difference to build the advanced mode is the different module name dce-linux instead of dce-ns3 (basic mode).

note that dce-linux-1.9 is the DCE version 1.9 module. If you would like to use the development version of DCE module, you can specify dce-linux-dev as a module name for bake.

2.2.3 Building DCE using WAF

While Bake is the best option, another one is the configuration and build using WAF. WAF is a Python-based framework for configuring, compiling and installing applications. The configuration scripts are coded in Python files named wscript, calling the WAF framework, and called by the waf executable.

In this case you need to install the single packages one by one. You may want to start with ns-3:

- HG_NS3= http://code.nsnam.org/ns-3-dev
- GIT_NS3= git@github.com:nsnam/ns-3-dev-git.git
- LAST_VERSION= ns-3.25

More detailed information can be found on the ns-3 wiki.

Then you can download and install net-next-sim and DCE (net-next-sim includes the linux stack module):

2.3 Examples

If you got succeed to build DCE, you can try an example script which is already included in DCE package.

2.3.1 Example: Simple UDP socket application

This example execute the binaries named udp-client and udp-server under ns-3 using DCE. These 2 binaries are written using POSIX socket API in order to send and receive UDP packets.

If you would like to see what is going on this script, please refer to the user’s guide.
This simulation produces two directories, the content of elf-cache is not important now for us, but files-0 is. files-0 contains first node’s file system, it also contains the output files of the dce applications launched on this node. In the /var/log directory there are some directories named with the virtual pid of corresponding DCE applications. Under these directories there is always 4 files:

1. cmdline: which contains the command line of the corresponding DCE application, in order to help you to retrieve what is it,
2. stdout: contains the stdout produced by the execution of the corresponding application,
3. stderr: contains the stderr produced by the execution of the corresponding application.
4. status: contains a status of the corresponding process with its start time. This file also contains the end time and exit code if applicable.

Before launching a simulation, you may also create files-xx directories and provide files required by the applications to be executed correctly.

### 2.3.2 Example: iperf

This example shows the usage of iperf with DCE. You are able to generate traffic by well-known traffic generator iperf in your simulation. For more detail of the scenario description, please refer to the user’s guide.

Once you successfully installed DCE with bake, you can execute the example using iperf.

```
cd source/ns-3-dce
./waf --run dce-iperf
```

As we saw in the previous example the experience creates directories containing the outputs of different executables, take a look at the server (node 1) output:

```
$ cat files-1/var/log/*/stdout
------------------------------------------------------------
Server listening on TCP port 5001
TCP window size: 124 KByte (default)
------------------------------------------------------------
[ 4] local 10.1.1.2 port 5001 connected with 10.1.1.1 port 49153
[ ID] Interval Transfer Bandwidth
[ 4] 0.0-11.2 sec 5.75 MBytes 4.30 Mbits/sec
```

the client (node-0) output bellow:

```
if you have already built the advanced mode, you can use Linux network stack over iperf.
```

```
cd source/ns-3-dce
./waf --run "dce-iperf --stack=linux"
```

the command line option --stack=linux makes the simulation use the Linux kernel stack instead of ns-3 network stack.

```
$ cat files-1/var/log/*/stdout
------------------------------------------------------------
Server listening on TCP port 5001
TCP window size: 85.3 KByte (default)
------------------------------------------------------------
[ 4] local 10.1.1.2 port 5001 connected with 10.1.1.1 port 60120
[ ID] Interval Transfer Bandwidth
[ 4] 0.0-11.2 sec 5.88 MBytes 4.41 Mbits/sec
```

2.3. Examples
$ cat files-0/var/log/*/stdout
------------------------------------------------------------
Client connecting to 10.1.1.2, TCP port 5001
TCP window size: 16.0 KByte (default)
------------------------------------------------------------
[ 3] local 10.1.1.1 port 60120 connected with 10.1.1.2 port 5001
[ ID] Interval       Transfer      Bandwidth
[ 3] 0.0- 1.0 sec 512 KBytes  4.19 Mbits/sec
[ 3] 1.0- 2.0 sec 640 KBytes  5.24 Mbits/sec
[ 3] 2.0- 3.0 sec 640 KBytes  5.24 Mbits/sec
[ 3] 3.0- 4.0 sec 512 KBytes  4.19 Mbits/sec
[ 3] 4.0- 5.0 sec 640 KBytes  5.24 Mbits/sec
[ 3] 5.0- 6.0 sec 512 KBytes  4.19 Mbits/sec
[ 3] 6.0- 7.0 sec 640 KBytes  5.24 Mbits/sec
[ 3] 7.0- 8.0 sec 640 KBytes  5.24 Mbits/sec
[ 3] 9.0-10.0 sec 640 KBytes  5.24 Mbits/sec
[ 3] 0.0-10.2 sec 5.88 MBytes  4.84 Mbits/sec

Interestingly, the two results between two network stacks are slightly different, though the difference is out of scope of this document.
CHAPTER THREE

USER’S GUIDE

This document is for the people who want to use your application in ns-3 using DCE.

Direct Code Execution (DCE) allows us to use POSIX socket-based applications as well as Linux kernel network stack.

3.1 Setup Guide

In order to install DCE you must follow the tutorial Build DCE.

Installation result

The result of the installation process is the creation of libraries from source of DCE and that of ns-3 and also some tools and sources of an optional Linux kernel if you have also chosen to use the stack of a Linux kernel. Below you will find the main directories:

|-- bakefile.xml Bake internal configuration file (generated by bake.py configure command).
|-- bakeSetEnv.sh Bake generated file used to configure environmental variable.
|-- build Target directory of |ns3| Core and DCE compilation.
    |-- bin
    |-- bin_dce
    |-- etc
    |-- include
    |-- lib
    |-- sbin
    |-- share
    |-- usr
    |--- var
|--- source Source directory during 'bake.py download'. Listed files below depend on the configuration of bake.
    |-- ccnx
    |-- ccnx-0.6.2.tar.gz
    |-- ns-3-dce
    |--- build Documentation source
    |--- doc
    |--- elf-cache
    |--- example Example scenarios using DCE
    |--- files-0
    |--- files-1
    |--- files-2
    |--- files-3
    |--- files-5
    |--- helper The source code directory for helper library
    |--- model The source code directory for DCE core
    |--- myscripts Sub-module and ad-hoc script directory
3.2 Basic Use Cases

3.2.1 Using your userspace protocol implementation

As explained in *How It Works*, DCE needs to relocate the executable binary in memory, and these binary files need to be built with specific compile/link options.

In order to this you should follow the two following rules:

1. Compile your objects using this gcc flag: `-fPIC` for example: `gcc -fPIC -c foo.c`
2. (option) Some application needs to be compile with `-U_FORTIFY_SOURCE` so that the application doesn’t use alternative symbols including `__chk` (like `memcpy_chk`).
3. Link your executable using this gcc flag: `-pie` and `-rdynamic` for example: `gcc -o foo -pie -rdynamic foo.o`
4. Verify the produced executable using `readelf` utility in order to display the ELF file header and to verify that your exe is of type `DYN` indicating that DCE should be able to relocate and virtualize it under *ns-3* virtual world and network. For example: `readelf -h foo | grep Type: ==> Type: DYN (Shared object file)`
5. Check also that your executable runs as expected outside of *ns-3* and DCE.

**Install the target executable**

Copy the executable file produced in a specified directory in the variable environment DCE_PATH so that DCE can find it. DCE_PATH behaves like the variable PATH and can contain several directories such as

```
/home/USER/iproute2/ip:/home/USER/iperf3/src:/home/USER/iperf2/src
```
Write a ns-3 script

Now that you have compiled your executable you can use it within ns-3 script with the help of a set of DCE Helper Classes:

<table>
<thead>
<tr>
<th>HELPER CLASS NAME</th>
<th>INCLUDE NAME</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>DceManagerHelper</td>
<td>ns3/dce-manager-helper.h</td>
<td>A DceManager is a DCE internal class which manage the execution of the executable you will declare to run within ns-3. The DceManagerHelper is the tool you will use within your script to parameter and install DceManager on the ns-3 nodes where you plan to run binaries. You will use this helper in order to define which application you want to run within ns-3 by setting the name of the binary its optional arguments, its environment variables, and also optionally if it take its input from a file instead of stdin. This class can be derived if you need to do more preparation before running your application. Often applications need configuration file to work properly, for example if you look at the contents of the helper named CcnClientHelper you will see that his job is to create the key files needed for the operation of CCNx’s applications.</td>
</tr>
<tr>
<td>DceApplicationHelper</td>
<td>ns3/dce-application-helper.h</td>
<td>This class can be derived if you need to do more preparation before running your application. Often applications need configuration file to work properly, for example if you look at the contents of the helper named CcnClientHelper you will see that his job is to create the key files needed for the operation of CCNx’s applications.</td>
</tr>
<tr>
<td>LinuxStackHelper</td>
<td>ns3/linux-stack-helper.h</td>
<td>This helper is used to configure parameters of Linux kernel when we are using the advanced mode.</td>
</tr>
<tr>
<td>CcnClientHelper</td>
<td>ns3/ccn-client-helper.h</td>
<td>This helper is a subclass of DceApplicationHelper, its jobs is to create keys files used by ccnx executables in order to run them correctly within NS3.</td>
</tr>
<tr>
<td>QuaggaHelper</td>
<td>ns3/quagga-helper.h</td>
<td>This helper is a subclass of DceApplicationHelper. It will help you to setup Quagga applications.</td>
</tr>
</tbody>
</table>

Note that the table above indicates the name of includes, so you can look at the comments in them, but in reality for DCE use you need to include only the file ns3/dce-module.h.

The directory named myscripts is a good place to place your scripts. To create a new script you should create a new directory under myscripts, and put your sources and a configuration file for waf build system, this file should be named wscript. For starters, you may refer to the contents of the directory myscripts/ping.

For more detail, please refer DCE API (doxygen) document.

Compile the script

To compile simply execute the command waf. The result must be under the directory named build/bin/myscripts/foo/bar where foo is your directory and bar your executable according to the content of your wscript file.

Launch the script

Simply launch your script like any other program.

$ ./waf --run bar
Results

The execution of the apps using DCE generates special files which reflect the execution thereof. On each node DCE creates a directory /var/log, this directory will contain subdirectory whose name is a number. This number is the pid of a process. Each of these directories contains the following files: cmdline, status, stdout, stderr. The file cmdline recalls the name of the executable run followed arguments. The file status contains an account of the execution and dating of the start; optionally if the execution is completed there is the date of the stop and the return code. The files stdout and stderr correspond to the standard output of the process in question.

Example: DCE Simple UDP (dce-udp-simple)

The example uses two POSIX socket-based application in a simulation. Please take time to look at the source dce-udp-simple.cc:

```cpp
#include "ns3/network-module.h"
#include "ns3/core-module.h"
#include "ns3/internet-module.h"
#include "ns3/dce-module.h"

using namespace ns3;

int main (int argc, char *argv[]) {
  CommandLine cmd;
  cmd.Parse (argc, argv);

  NodeContainer nodes;
  nodes.Create (1);

  InternetStackHelper stack;
  stack.Install (nodes);

  DceManagerHelper dceManager;
  dceManager.Install (nodes);

  DceApplicationHelper dce;
  ApplicationContainer apps;

  dce.SetStackSize (1 << 20);
  dce.SetBinary ("udp-server");
  dce.ResetArguments ();
  apps = dce.Install (nodes.Get (0));
  apps.Start (Seconds (4.0));

  dce.SetBinary ("udp-client");
  dce.ResetArguments ();
  dce.AddArgument ("127.0.0.1");
  apps = dce.Install (nodes.Get (0));
  apps.Start (Seconds (4.5));

  Simulator::Stop (Seconds (1000100.0));
  Simulator::Run ();
  Simulator::Destroy ();

  return 0;
}
```
You can notice that we create a *ns-3* Node with an Internet Stack (please refer to *ns-3 doc.* for more info), and we can also see 2 new Helpers:

1. DceManagerHelper which is used to Manage DCE loading system in each node where DCE will be used.
2. DceApplicationHelper which is used to describe real application to be launched by DCE within *ns-3* simulation environment.

**Example: DCE with iperf(dce-iperf)**

The example uses iperf traffic generator in a simulation. The scenario is here:

```cpp
#include "ns3/network-module.h"
#include "ns3/core-module.h"
#include "ns3/internet-module.h"
#include "ns3/dce-module.h"
#include "ns3/point-to-point-module.h"
#include "ns3/applications-module.h"
#include "ns3/netanim-module.h"
#include "constant-position-mobility-model.h"

using namespace ns3;

NS_LOG_COMPONENT_DEFINE ("DceIperf");

// node 0 node 1
// +----------------+ +----------------+
// | | | |
// +----------------+ +----------------+
// | 10.1.1.1 | | 10.1.1.2 |
// +----------------+ +----------------+
// | point-to-point | | point-to-point |
// +----------------+ +----------------+
// |
// 5 Mbps, 2 ms

// 2 nodes : iperf client en iperf server ....

// Note: Tested with iperf 2.0.5, you need to modify iperf source in order to allow DCE to have a chance to end an endless loop in iperf as follow:
// in source named Thread.c at line 412 in method named thread_rest
// you must add a sleep (1); to break the infinite loop....

int main (int argc, char *argv[])
{
    std::string stack = "ns3";
    bool useUdp = 0;
    std::string bandwidth = "1m";
    CommandLine cmd;
    cmd.AddValue ("stack", "Name of IP stack: ns3/linux/freebsd.", stack);
    cmd.AddValue ("udp", "Use UDP. Default false (0)", useUdp);
    cmd.AddValue ("bw", "BandWidth. Default 1m.", bandwidth);
    cmd.Parse (argc, argv);

    NodeContainer nodes;
    nodes.Create (2);
}
```

### 3.2. Basic Use Cases
PointToPointHelper pointToPoint;
pointToPoint.SetDeviceAttribute ("DataRate", StringValue ("5Mbps"));
pointToPoint.SetChannelAttribute ("Delay", StringValue ("1ms"));

NetDeviceContainer devices, devices2;
devices = pointToPoint.Install (nodes);
devices2 = pointToPoint.Install (nodes);

DceManagerHelper dceManager;
dceManager.SetTaskManagerAttribute ("FiberManagerType", StringValue ("UcontextFiberManager"));

if (stack == "ns3")
{
    InternetStackHelper stack;
    stack.Install (nodes);
    dceManager.Install (nodes);
}
else if (stack == "linux")
{
#ifdef KERNEL_STACK
    dceManager.SetNetworkStack ("ns3::LinuxSocketFdFactory", "Library", StringValue ("liblinux.so")
    dceManager.Install (nodes);
    LinuxStackHelper stack;
    stack.Install (nodes);
#else
    NS_LOG_ERROR ("Linux kernel stack for DCE is not available. build with dce-linux module.");
    // silently exit
    return 0;
#endif
}
else if (stack == "freebsd")
{
#ifdef KERNEL_STACK
    dceManager.SetNetworkStack ("ns3::FreeBSDSocketFdFactory", "Library", StringValue ("libfreebsd.so")
    dceManager.Install (nodes);
    FreeBSDStackHelper stack;
    stack.Install (nodes);
#else
    NS_LOG_ERROR ("FreeBSD kernel stack for DCE is not available. build with dce-freebsd module.");
    // silently exit
    return 0;
#endif
}

Ipv4AddressHelper address;
address.SetBase ("10.1.1.0", "255.255.255.252");
Ipv4InterfaceContainer interfaces = address.Assign (devices);

// setup ip routes
Ipv4GlobalRoutingHelper::PopulateRoutingTables ();
#ifdef KERNEL_STACK
if (stack == "linux")
{
    LinuxStackHelper::PopulateRoutingTables ();
}
#endif

DceApplicationHelper dce;
ApplicationContainer apps;
dce.SetStackSize (1 << 20);

// Launch iperf client on node 0
dce.SetBinary ("iperf");
dce.ResetArguments ();
dce.ResetEnvironment ();
dce.AddArgument ("-c");
dce.AddArgument ("10.1.1.2");
dce.AddArgument ("-i");
dce.AddArgument ("-l");
dce.AddArgument ("--time");
dce.AddArgument ("10");
if (useUdp)
{
    dce.AddArgument ("-u");
    dce.AddArgument ("-b");
    dce.AddArgument (bandWidth);
}
apps = dce.Install (nodes.Get (0));
apps.Start (Seconds (0.7));
apps.Stop (Seconds (20));

// Launch iperf server on node 1
dce.SetBinary ("iperf");
dce.ResetArguments ();
dce.ResetEnvironment ();
dce.AddArgument ("-s");
dce.AddArgument ("-P");
dce.AddArgument ("1");
if (useUdp)
{
    dce.AddArgument ("-u");
}
apps = dce.Install (nodes.Get (1));
pointToPoint.EnablePcapAll ("iperf-" + stack, false);
apps.Start (Seconds (0.6));
setPos (nodes.Get (0), 1, 10, 0);
setPos (nodes.Get (1), 50, 10, 0);
Simulator::Stop (Seconds (40.0));
Simulator::Run ();
Simulator::Destroy ();
return 0;
}

This scenario is simple there is 2 nodes linked by a point 2 point link, the node 0 launch iperf as a client via the
command `iperf -c 10.1.1.2 -i 1 –time 10` and the node 1 launch iperf as a server via the command `iperf -s -P 1`. You can follow this to launch the experiment:

### 3.2.2 Using your in-kernel protocol implementation

There are a number of protocols implemented in kernel space, like many transport protocols (e.g., TCP, UDP, SCTP), Layer-3 forwarding plane (IPv4, v6 with related protocols ARP, Neighbor Discovery, etc). DCE can simulate these protocols with ns-3 and a kernel compiled as a library. This is not possible with the linux vanilla kernel hence you need to use a slightly modified kernel (`net-next-sim - deprecated` - or `libos` which is the successor of `net-next-sim`).

This document describes how to retrieve, configure, compile and use this custom kernel in DCE. As an example, it shows how to enable an in-tree but optional protocol - Stream Control Transmission Protocol (SCTP) - and an out of tree protocol: Multipath TCP (MPTCP). Although other protocols may not adapt these patterns as-is, you will see what’s needed to implement for your purpose.

1. Configure a kernel (make menuconfig in Linux)
2. Build a DCE compatible linux kernel
   2.1 With `net-next-sim` (deprecated)
   2.2 Build an MPTCP kernel with `net-next-sim`
   2.3 With `libos`
3. Write user space application to use this protocol
4. Write ns-3 scenario to use above applications.
5. run it!

#### 1. Configure a kernel (make menuconfig in Linux)

In Linux kernel, there is a configuration system in Linux kernel to enable/disable features. This is typically done by `make menuconfig` command, and it writes a file (.config) at the kernel source directory. Build system (e.g., make `bzImage`) refers the file which source files are compiled.

In our DCE Linux kernel module (i.e. `net-next-sim` available at github.com), we have `arch/sim/defconfig` file to store the default configuration of kernel features. You may need to add proper configuration parameters (e.g., `CONFIG_IP_SCTP`) to build the protocol by default.

For the Linux SCTP implementation, we need at least the following configuration parameters.

- `CONFIG_IP_SCTP=y`
- `CONFIG_SCTP_DEFAULT_COOKIE_HMAC_NONE=y`
- `CONFIG_CRYPTO_CRC32C=y`
- `CONFIG_CRC32=y`

These can be added to the file (`arch/sim/defconfig`) or added manually to `.config` when needed.

Note

These configuration options SHOULD be minimized at the beginning since adding new option may require new functions which DCE doesn’t support at the time and need to implement glue code.

#### 2. Build a DCE compatible linux kernel

DCE can simulate these protocols with ns-3 and a kernel compiled as a library named `liblinux.so`. This is not possible with the linux vanilla kernel hence you need to use a slightly modified kernel (`net-next-sim - deprecated` - or `libos`)
which is the successor of net-next-sim). The following section presents the two methods, net-next-sim being slightly easier to install.

### 2.1 With net-next-sim (deprecated)

To build the liblinux.so, DCE version of Linux kernel,

```bash
make defconfig ARCH=sim
make library ARCH=sim
```

You can use `make menuconfig` command (below) instead of editing the defconfig file. If everything is fine, you will see `liblinux.so` linked to `libsim-linuxv.y.z.so` file at the root directory of Linux kernel.

```bash
make menuconfig ARCH=sim
```

### 2.2 Build an MPTCP kernel with net-next-sim

DISCLAIMER: This is a transcript of Hajime’s guide https://plus.google.com/+HajimeTazaki/posts/1QUmR3n3vNA updated on a best effort basis. Hence it may be possible to compile newer versions, in which case patches are welcome. Build steps for DCE integration is also available as a script in `utils/mptcp-build.sh`.

1. Get linux kernel mptcp variant:

```bash
$ git clone mptcp_git_
$ git checkout -b sim3.14 mptcp_v0.89
```
1. Merge in mptcp code the changes required to make it work in DCE. Those changes are listed in the **net-next-sim** project

   % cd mptcp
   % git remote add dce git://github.com/direct-code-execution/net-next-sim.git
   % git fetch dce
   % git merge dce/sim-ns3-3.14.0-branch

2. Enable mptcp in the kernel configuration. There are 2 ways possible:
   - patch the kernel config manually (you can

     % cat >> arch/sim/defconfig <<END
     CONFIG_MPTCP=y
     CONFIG_MPTCP_PM_ADVANCED=y
     CONFIG_MPTCP_FULLMESH=y
     CONFIG_MPTCP_NDIFFPORTS=y
     CONFIG_DEFAULT_FULLMESH=y
     CONFIG_DEFAULT_MPTCP_PM="fullmesh"
     CONFIG_TCP_CONG_COUPLED=y
     CONFIG_TCP_CONG_OLIA=y
     END

   - or use menuconfig to enable these options as explained in [http://multipath-tcp.org/pmwiki.php/Users/DoItYourself](http://multipath-tcp.org/pmwiki.php/Users/DoItYourself)

     % make menuconfig ARCH=sim

3. Generate the kernel configuration:

   % make defconfig ARCH=sim

4. build kernel (as a shared library)

   $ make library ARCH=sim

If everything is going well, you can try to use it over ns-3

1. build ns-3 related tools

   % make testbin -C arch/sim/test

2. run an mptcp simulation !

   % cd arch/sim/test/buildtop/source/ns-3-dce
   % ./waf --run dce-iperf-mptcp

   you should see generated *.pcap files in your dce folder.

### 2.3 With libos

**Libos** is the successor of **net-next-sim**. There are attempts to merge **libos** within the linux kernel library **lkl** project but it will take quite some time before being able to run **lkl** in DCE.

1. Get **libos** code:

   $ git clone git@github.com:libos-nuse/net-next-nuse.git
   $ git checkout libos-v4.4
   $ cd net-next-nuse

2. Configure the kernel (you can refer to **configure**)
make defconfig ARCH=lib

3. Compile the kernel

make library ARCH=lib

This will download and compile git submodules from https://github.com/libos-nuse/linux-libos-tools and execute additional steps from the arch/lib/tools/Makefile, as for instance generate an additional libsim_linux.so. This is the shared library you need to load in DCE. By default DCE looks for "liblinux.so" so you should do:

DceManagerHelper dceManager;
dceManager.SetNetworkStack ("ns3::LinuxSocketFdFactory", "Library", StringValue ("libsim-linux.so"));

3. Write user space application to use this protocol

Then, we need to write userspace applications using new feature of kernel protocol. In case of SCTP, we wrote sctp-client.cc and sctp-server.cc.

Optional You may optionally need external libraries to build/run the applications. In this case, the applications need lksctp-tools, so that applications fully benefit the features of SCTP, rather than only using standard POSIX socket API.

Moreover, adding system dependency to bake configuration file (i.e., bakeconf.xml) would be nice to assist build procedure. The following is an example of lksctp-tools, which above applications use.

```xml
<module name="lksctp-dev">
    <source type="system_dependency">
        <attribute name="dependency_test" value="sctp.h"/>
        <attribute name="try_to_install" value="True"/>
        <attribute name="name_apt-get" value="lksctp-dev"/>
        <attribute name="name_yum" value="lksctp-tools-devel"/>
        <attribute name="more_information" value="Didn't find: lksctp-dev package; please install it."/>
    </source>
    <build type="none" objdir="no">
    </build>
</module>
```

4. Write ns-3 scenario to use above applications.

The next step would be writing ns-3 simulation scenario to use the applications you prepared. You need first to add to your DCE_PATH the path to the previously compiled liblinux.so:

```
:: $ export DCE_PATH="$HOME/net-next-sim:$DCE_PATH"
```

dce-sctp-simple.cc is the script that we prepared. In the script, you may need to load the applications by using DceApplicationHelper as follows.

```c
DceApplicationHelper process;
ApplicationContainer apps;

/* by default DCE, loads liblinux.so. If for some reason your library has a different name, either create a symlink or use these commands to change the name */
DceManagerHelper dceManager;
dceManager.SetNetworkStack ("ns3::LinuxSocketFdFactory", "Library", StringValue ("liblinux.so"));

process.SetBinary ("sctp-server");
process.ResetArguments ();
```
process.SetStackSize (1<<16);
apps = process.Install (nodes.Get (0));
apps.Start (Seconds (1.0));

process.SetBinary ("sctp-client");
process.ResetArguments ();
process.ParseArguments ("10.0.0.1");
apps = process.Install (nodes.Get (1));
apps.Start (Seconds (1.5));

5. run it!

./waf --run dce-simple-sctp

If you’re lucky, it’s done.

If you aren’t lucky, you may face errors of DCE, such as unresolved symbols in system calls (called by userspace applications) or missing kernel functions (used by newly added CONFIG_IP_SCTP option), or invalid memory access causing segmentation fault. In that case, adding missing functions, so called glue-code would be the next step.

3.2.3 How to add system calls?

Introduction

If your applications running with DCE are not able to run due to missing function symbols, you need to add the function call or system call to DCE by hand. The POSIX API coverage of DCE is growing day by day, but your contribution is definitely helpful not only for your case, but also for someone will use in future.

More specifically, if you faced the following error when you executed, you need to add a function call to DCE. In the following case, a symbol `strfry` not defined in DCE is detected during the execution of the simulation.

Types of symbol

There are two types of symbols that is defined in DCE.

- NATIVE
  
  NATIVE symbol is a symbol that DCE doesn’t care about the behavior. So this type of symbol is redirected to the one provided by underlying host operating system (i.e., glibc).

- DCE
  
  DCE symbol is a symbol that DCE reimplements its behavior instead of using the underlying system’s one. For instance, `socket()` call used in an application redirected to DCE to cooperate with ns-3 or Linux network stack managed by DCE. `malloc()` is also this kind.

In general (but not strictly), if a call is related to a kernel resource (like NIC, clock, etc), it should use DCE macro. Otherwise (like strcmp, atoi etc), the call should use NATIVE.

Files should be modified

In order to add function calls or system calls that DCE can handle, you need to modify the following files.

- model/libc-ns3.h
This is the first file that you need to edit. You may lookup the symbol that you’re going to add and once you can’t find it, add the following line.

NATIVE (strfry)

This is the case of the symbol `strfry()`, which we don’t have to reimplement. But you may need to add include file that defines the symbol (strfry()) at model/libc-dce.cc.

If the symbol needs to reimplemented for DCE, you may add as follows.

DCE (socket)

- model/dce-abc.cc

In case of DCE symbol, you’re going to introduce DCE redirected function. We use naming convention with prefix of `dce_` to the symbol (i.e., `dce_socket`) to define new symbol and add the implementation in a `cc` file. The following is the example of `dce_socket` implementation.

We implemented `dce_socket` function in the file `model/dce-fd.cc`.

```c
int dce_socket (int domain, int type, int protocol)
```

In the function, we carefully fill the function contents to cooperate with ns-3. The below line is creating DCE specific socket instance (i.e., ns-3 or DCE Linux) instead of calling system call allocating kernel space socket resources.

```c
UnixFd *socket = factory->CreateSocket (domain, type, protocol);
```

Other function calls such as file system related functions (e.g., read, fopen), time reheated features (e.g., gettimeofday, clock_gettime), signal/process utilities (e.g., getpid, sigaction), and thread library (e.g., pthread_create). All these functions should be DCE since DCE core reimplements these feature instead of using underlying host system.

- model/dce-abc.h

Once you got implemented the new redirected function, you may add the function prototype declaration to refer from other source files. `dce_socket()` is added to model/sys/dce-socket.h.

### 3.2.4 Creating your protocol implementation as a DCE sub-module

If your application has a configuration file to modify the behavior of applications, introducing a particular Helper class will be helpful to handle your application. In this section, we will give you an advanced way of using your application with DCE.

Some of existing submodule are following this way. You can find ns-3-dce-quagga and ns-3-dce-umip as examples to add sub-module.

#### Obtaining DCE sub-module template

First of all, you could start with referring sub module template available as follows.

```
hg clone http://code.nsnam.org/thehajime/ns-3-dce-submodule (your module name)
```

The template consists of, wscript, helper, test and documentation. You could rename all/some of them for your module. Then, put `ns-3-dce-submodule` directory under `ns-3-dce/myscripts/`. This will be required to build under ns-3-dce module as an extension (sub-module) of dce.
Writing wscript

Writing bakeconf.xml (optional)

Implementing helper class (optional)

Writing examples (optional)

3.2.5 Global DCE Configurations

Parameters

The DCE specifics variables are essentially two PATH like variables: so within them you may put paths separated by ‘:’ character.

**DCE_PATH** is used by DCE to find the executable you want to launch within *ns-3* simulated network. This variable is used when you reference the executable using a relative form like ‘ping’.

**DCE_ROOT** is similar to **DCE_PATH** but it is used when you use an absolute form for example ‘/bin/bash’.

Please pay attention that executables that you will place in the directories indicated in the previous variables should be recompiled accordingly to the rules defined in the next chapter.

(FIXME: to be updated)

Tweaking

DCE is configurable with NS3 Attributes. Refer to the following table:

<table>
<thead>
<tr>
<th>ATTRIBUTE</th>
<th>DESCRIPTION</th>
<th>VALUES</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiber-Manager-Type</td>
<td>The TaskManager is used to switch the execution context between threads and processes.</td>
<td>UcontextFiberManager the more efficient, PthreadFiberManager helpful with gdb to see the threads. This is the default.</td>
<td>--ns3::TaskManager::FiberManagerType=UcontextFiberManager dceManager.SetTaskManagerAttribute(&quot;FiberManager&quot;, StringValue(&quot;UcontextFiberManager&quot;)); --ns3::TaskManager::FiberManagerType=PthreadManagerType</td>
</tr>
<tr>
<td>Loader-Factory</td>
<td>The LoaderFactory is used to load the hosted binaries.</td>
<td>CoojaLoaderFactory is the default and the only one that supports fork. DlmLoaderFactory is the more efficient. To use it you have two ways: 1/ use dce-runner 2/ link using ldso as default interpreter.</td>
<td>--ns3::DceManagerHelper::LoaderFactory=ns3::CoojaLoaderFactory $ dce-runner my-dce-ns3-script OR gcc -o my-dce-ns3-script Wl,--dynamic-linker=PATH2LD$O/ldso ... $ my-dce-ns3-script --ns3::DceManagerHelper::LoaderFactory=ns3::DlmLoaderFactory dceManager.SetLoader(&quot;ns3::DlmLoaderFactory&quot;)</td>
</tr>
</tbody>
</table>

3.2.6 DCE Cradle

This document describes what DCE Cradle is, how we can use it, how we extend it.

Tutorials and how to reproduce the experiment of WNS3 2013 paper is available dce-cradle-usecase.
What is DCE Cradle?

DCE Cradle enables us to use Linux kernel via Direct Code Execution from the ns-3 native socket application. Applications can access it via ns-3 socket API. Currently (6th Jan. 2014) the following sockets are available:

- IPv4/IPv6 UDP
- IPv4/IPv6 TCP
- IPv4/IPv6 RAW socket
- IPv4/IPv6 DCCP
- IPv4/IPv6 SCTP

Installing DCE Cradle

DCE Cradle is already integrated in ns-3-dce module. You can just build and install DCE as instructed in the parent document.

How to use it

```c
OnOffHelper onoff = OnOffHelper("ns3::LinuxTcpSocketFactory",
        InetSocketAddress(interfaces.GetAddress (1), 9));
```

How to extend it

(To be added)

Article

- The project originally started during GSoC project 2012
- “DCE cradle: simulate network protocols with real stacks for better realism”, WNS3 2013, [PDF]

3.2.7 Aspect-based Tracing

Aspect-based tracing, provided by libaspect, allows us to use tracing facility with unmodified code.

One of contradictions when we use DCE is, tracing, how to put trace sources into unmodified code. While DCE gives an opportunity to use unmodified codes as simulation protocols, one might want to investigate which function is called or how many messages of a particular protocol are exchanged.

ns-3 originally has a nice feature of tracing with such a purpose, with on-demand trace connector to obtain additional information. Instead of inserting TraceSource into the original code, DCE gives dynamic trace points with this library, based on the idea of aspect-based tracing.

For more detail, see the Chapter 6.3.2 of the thesis.
Quick Start

To put trace sources without modifying the original code, aspcpp::HookManager gives trace hooks into arbitrary source codes and functions.

```cpp
#include <hook-manager.h>

HookManager hooks;
hooks.AddHookBySourceAndFunction ("ip_input.c", "::ip_rcv", &IpRcv);
hooks.AddHookByFunction ("::process_backlog", &ProcBacklog);
hooks.AddHookByFunction ("::arp_xmit", &ArpXmit);
```

The above examples specifies file name and functions with callback functions in the simulation script.

Limitations

- July 10th, 2013: aspect-based tracing (libaspect) is in the alpha release state. It might be updated frequently.
- Callback function has no argument that it can investigate the contents of buffer that each function handles.

3.2.8 FreeBSD kernel support with DCE

Overview

This module provides an additional network stack support for DCE with FreeBSD kernel.

<table>
<thead>
<tr>
<th>POSIX apps</th>
</tr>
</thead>
<tbody>
<tr>
<td>POSIX glue code</td>
</tr>
<tr>
<td>network stack (ns3/linux/freebsd)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

- This is a POSIX userspace application support over FreeBSD kernel over DCE, over ns-3 over Linux.
- This is **NOT** the DCE running on FreeBSD operating system (right now DCE only run on Linux).

Usage

```
bake.py configure -e dce-freebsd-dev
bake.py download
bake.py build
```
Once you finished to build dce-freebsd module, you can write a simulation script that uses FreeBSD kernel as a network stack. All you need is to specify the library name with an attribute `Library` to `ns3::FreeBSDSocketFdFactory`, then install `FreeBSDStackHelper` to the nodes.

```cpp
DceManagerHelper processManager;
processManager.SetNetworkStack ("ns3::FreeBSDSocketFdFactory",
    "Library", StringValue ("libfreebsd.so"));
processManager.Install (nodes);
FreeBSDStackHelper stack;
stack.Install (nodes);
```

**How to use your kernel extension with DCE?**

No configuration support (like `make menuconfig` in Linux) right now. You need to add files into `sys/sim/Makefile`.

The following represents an example to add Multipath-TCP feature of FreeBSD by adding `mptcp_subr.o` to the `freebsd-sim/sys/sim/Makefile`. If you want to add your code into the kernel build, you may add object file names that is from your own codes.

```bash
diff --git a/sys/sim/Makefile b/sys/sim/Makefile
index 8115e3d..1b2feab 100644
--- a/sys/sim/Makefile
+++ b/sys/sim/Makefile
@@ -100,7 +100,7 @@ ip_divert.o tcp_hostcache.o ip_ecn.o tcp_input.o ip_encap.o
    tcp_lro.o ip_fastfwd.o tcp_offload.o ip_gre.o tcp_output.o \
    ip_icmp.o tcp_reass.o ip_id.o tcp_sack.o tcp_input.o tcp_subr.o \
    tcp_syncache.o ip_mroute.o tcp_timer.o ip_options.o tcp_timewait.o \
-    ip_output.o tcp_usrreq.o raw_ip.o udp_usrreq.o if_llatbl.o
+    ip_output.o tcp_usrreq.o raw_ip.o udp_usrreq.o if_llatbl.o mptcp_subr.o
```

**Limitations**

While this release gives a proof of concept to use FreeBSD kernel with DCE, there are tons of limitations (listed below) that have to be improved in near future.

- No sysctl
- No IPv6 address configuration
- No ifconfig
- No delete IP addresses support
- No route command
- No quagga, new FreeBSD-version glue code needed
- No extensive test
- No DCE Cradle
- No Poll implementation
- No getifaddr/routing socket implementations
- Socket options missing (inconsisitent defined value: SOL_SOCKET( Linux/FreeBSD)=20/0xffff)
TODO

- refactoring with LinuxStackHelper/FreeBSDStackHelper
- refactoring with wscript (–enable-kernel-stack uses both FreeBSD/Linux for now)

3.3 Advanced Use Cases

3.3.1 Using Alternative, Fast Loader

DCE optionally supports an alternative ELF loader/linker, so-called elf-loader, in order to replace system-provided
linker/loader module. The intention of the loader is to support unlimited number of instances used by dlmopen
call, which provides DCE to load a single ELF binary to multiple different memory space. dlmopen-based loader
(ns3::DlmLoaderFactory) is much faster than another default one (ns3::CoojaLOaderFactory), but few issues are re-
main so, this is optional.

To Speedup Run-time

In order to use DlmLoaderFactory, you can add command-line argument of waf.

./waf --run dce-tcp-simple --dlm

if you are in the ./waf shell mode, the following command should be used instead.

./build/bin/dce-runner ./build/bin/dce-tcp-simple

3.3.2 Tuning System Limits

When dealing with large or complex models, you can easily reach the limits of your system. For example, you cannot
open more than a fixed number of files. You can try the command “limit –a” to check them.

File limits: “Could not open ...”

You may see the following error:

msg="Could not open "/var"", file=../model/dce-manager.cc, line=149
terminate called without an active exception

This error masks error “24 Too many open files”. The cause of this is that the simulation process exceeded the limit
of open files per process. Check the limit of open files per process with “ulimit -n” To solve it, you can edit file
/etc/security/limits.conf and add the following lines at the end:

* hard nofile 65536
* soft nofile 65536

or

myuser hard nofile 65536
myuser soft nofile 65536
Processes limit: “Resource temporarily unavailable”

In this case you may see the an error like the following:

```c
assert failed. cond="error == 0", msg="error=Resource temporarily unavailable",
file=../model/pthread-fiber-manager.cc, line=321
terminate called without an active exception
```

`pthread-fiber-manager` invokes `pthread_create` this is what raises the “Resource temporarily unavailable”. This problem might be triggered because the maximum number of user processes is not big enough. Use “`ulimit -u`” to check this limit. To solve it, you can edit file `/etc/security/limits.conf` and add the following lines at the end:

```
* hard nproc 65536
* soft nproc 65536
```

or

```
myuser hard nproc 65536
```

Stack size

DCE directly manages the stack of the processes running on it, assigning it a default value 8192. For complex executables this value is too small, and may raise ‘stack overflow’ exceptions, or in other cases it may originate inconsistent values. For example, a value passed to a function changes without apparent reason when the program enters in that function. The value of the stack size can be changed with the `SetStackSize` instruction:

```c
DceApplicationHelper dce;
dce.SetStackSize (1<<20);
```

### 3.3.3 Debugging your protocols with DCE

**Gdb**

It is possible to use gdb to debug a script DCE/ns-3. As explained somewhere in the execution of a script is monoprocess, then you can put breakpoints in both sources of DCE and those of binaries hosted by DCE.

**Install**

Although it is not strictly necessary, it is recommended that you recompile a CVS Gdb for use with ns-3-dce. First, download:

```bash
cvs -d :pserver:anoncvs@sourceware.org:/cvs/src login {enter “anoncvs” as the password} cvs -d :pserver:anoncvs@sourceware.org:/cvs/src co gdb
```

Note that you might consider looking at [http://sourceware.org/gdb/current/](http://sourceware.org/gdb/current/) to obtain more efficient (cpu/bandwidth-wise) download instructions.

Anyway, now, you can build:

```bash
cd gdb
./configure
make
```

And, then, invoke the version of gdb located in `gdb/gdb` instead of your system-installed gdb whenever you need to debug a DCE-based program.
Using

If you use gdb (a CVS or stable version), do not forget to execute the following command prior to running any DCE-based program:

```
(gdb) handle SIGUSR1 nostop
Signal StopPrintPass to programDescription
SIGUSR1 NoYesYesUser defined signal 1
(gdb)
```

An alternate way to do this and avoid having to repeat this command ad-nauseam involves creating a .gdbinit file in your ns-3-dce directory and putting this inside:

```
handle SIGUSR1 nostop
```

or it can be put on the command line using the “-ex” flag:

```
./waf --run SCRIPT_NAME_HERE --command-template="gdb -ex 'handle SIGUSR1 nostop noprint' --args %s <args>"
```

Setup Eclipse Remote Debugging Environment

To remotely debug a DCE script you can use gdbserver as in the following example, changing the host name and port (localhost:1234):

```
./waf --run dce-httpd --command-template="gdbserver localhost:1234 %s <args>"
```

Then you can point a gdb client to your server. For example, in the following figure is reported an Eclipse debug configuration:

![Image of Eclipse debug configuration]

Once you start the debug session, you can use the usual Eclipse/gdb commands.

Helpful debugging hints

There are a couple of functions which are useful to put breakpoints into:

- `ns3::DceManager::StartProcessDebugHook`
Put a breakpoint in a specific node in a simulation

If you got a trouble in your protocol during interactions between distributed nodes, you want to investigate a specific state of the protocol in a specific node. In a usual system, this is a typical case of using distributed debugger (e.g., ddt, or `mpirun xterm -e gdb -args xxx`), but it is annoying task in general due to the difficulty of controlling distributed nodes and processes.

DCE gives an easy interface to debug distributed applications/protocols by the single-process model of its architecture.

The following is an example of debugging Mobile IPv6 stack (of Linux) in a specific node (i.e., home agent). A special function `dce_debug_nodeid()` is useful if you put a break condition in a gdb session.

```
(gdb) b mip6_mh_filter if dce_debug_nodeid()==0
Breakpoint 1 at 0x7ffff287c569: file net/ipv6/mip6.c, line 88.
<continue>
(gdb) bt 4
#0  mip6_mh_filter (sk=0x7ffff7f69e10, skb=0x7ffff7cde8b0)
at net/ipv6/mip6.c:109
#1  0x00007fff2831418 in ipv6_raw_deliver (skb=0x7ffff7cde8b0, nexthdr=135)
at net/ipv6/raw.c:199
#2  0x00007fff2831697 in raw6_local_deliver (skb=0x7ffff7cde8b0, nexthdr=135)
at net/ipv6/raw.c:232
#3  0x00007fff27e6068 in ip6_input_finish (skb=0x7ffff7cde8b0)
at net/ipv6/ip6_input.c:197
(More stack frames follow...)
```

Valgrind

(FIXME: simple session using valgrind)

### 3.3.4 Testing your protocols with DCE

Since DCE allows protocol implementations to expose network conditions (packet losses, reordering, and errors) with the interactions among distributed nodes, which is not easily available by traditional user-mode virtualization tools, exercising your code is easily done with a single simulation script.

**Coverage Test**

Improving code coverage with writing test programs is one of headache; - writing test program is annoying, - preparing test network tends to be short-term, and - the result is not reproducible.

This text describes how to measure code coverage of protocol implementations with DCE.

1. build target implementations (applications, kernel stack) with profile option
2. run test program with DCE
3. parse the result of test coverage

**Setup**

First, you need to compile your application with additional flags. **-fprofile-arcs -ftest-coverage** is used for a compilation flag (CFLAGS/CXXFLAGS), and **-fprofile-arcs** is used for a linker flag (LDFLAGS).
Write Test Program

Next, write a test program like ns-3 simulation script for your application (i.e., newapp).

```cpp
$ cat myscripts/dce-newapp.cc

int main (int argc, char *argv[]) {
    CommandLine cmd;
    cmd.Parse (argc, argv);

    NodeContainer nodes;
    nodes.Create (2);

    InternetStackHelper stack;
    stack.Install (nodes);

    DceManagerHelper dceManager;
    dceManager.Install (nodes);

    DceApplicationHelper dce;
    ApplicationContainer apps;

    // application on node 0
    dce.SetBinary ("newapp");
    dce.ResetArguments();
    apps = dce.Install (nodes.Get (0));
    apps.Start (Seconds (4.0));

    // application on node 1
    dce.SetBinary ("newapp");
    dce.ResetArguments();
    apps = dce.Install (nodes.Get (1));
    apps.Start (Seconds (4.5));

    Simulator::Stop (Seconds (100.0));
    Simulator::Run ();
    Simulator::Destroy ();

    return 0;
}
```

Run Test

Then, test your application as normal ns-3 (and DCE) simulation execution.

```
./waf --run dce-newapp
```

If you successfully finish your test, you will see the coverage data files (i.e., gcov data files) with a file extension .gcda.

```bash
$ find ./ -name "*.gcda"
```
Parse Test Result

We use lcov utilities as a parse of coverage test result.

Put the compiler (gcc) generated files (*.gcno) in the result directory,

cp *.gcno ./files-0/home/you/are/here/ns-3-dce/
cp *.gcno ./files-1/home/you/are/here/ns-3-dce/

then run the lcov and genhtml command to generate coverage information of your test program.

lcov -c -d . -b . -o test.info
genhtml test.info -o html

You will see the following output and generated html pages.

Reading data file test.info
Found 8 entries.
Writing .css and .png files.
Generating output.
Processing file ns-3-dce/example/udp-server.cc
genhtml: Use of uninitialized value in subtraction (-) at /usr/bin/genhtml line 4313.
Processing file ns-3-dce/example/udp-client.cc
genhtml: Use of uninitialized value in subtraction (-) at /usr/bin/genhtml line 4313.
Processing file /usr/include/c++/4.4/iostream
Processing file /usr/include/c++/4.4/ostream
Processing file /usr/include/c++/4.4/bits/ios_base.h
Processing file /usr/include/c++/4.4/bits/locale_facets.h
Processing file /usr/include/c++/4.4/bits/char_traits.h
Processing file /usr/include/c++/4.4/bits/basic_ios.h
Writing directory view page.
Overall coverage rate:
  lines.......: 49.3% (35 of 71 lines)
  functions..: 31.6% (6 of 19 functions)

Fuzz Test

(TBA, about integration of trinity)
Regression Test

(TBA)

3.4 Technical Information

3.4.1 DCE in a Nutshell

File System

To start a program in the world of ns-3 you must indicate on which node it will be launched. Once launched this program will have access only to the file system corresponding to the node that corresponds to a directory on your machine called file-X where X is the decimal number of the corresponding node. The file-X directories are created by DCE, only when they do not already exist. Also note that the contents of this directory is not cleared when starting the script. So you can copy the files required for the operation of your executables in the tree nodes. If possible it is best that you create these files from the script itself in order to simplify maintenance. DCE provides some helpers for creating configuration files necessary to the execution of certain apps like CCNx and Quagga.

Network

Your program running in a ns-3 node views the network defined by the script for this node.

Time

Time perceived by your executable is the simulated time of ns-3. Also note that DCE supports real time scheduler of ns-3 with the same limitations.

3.4.2 Limitations

- Currently the POSIX API (libc) is not fully supported by DCE. However there are already about 400 methods supported. As the goal of DCE is to allow to execute network applications, many methods related to the network are supported for example socket, connect, bind, listen, read, write, poll, select. The next chapter list the applications well tested using DCE.
- Some methods are not usable with all options of DCE. For more details refer to chapter Coverage API that lists all the supported methods.
- The scheduler is not as advanced as that of a kernel, for example if an infinite loop in a hosted application, DCE can not get out, but this should not happen in applications written correctly.

3.4.3 API Coverage

Below there is the list of the systems calls supported by DCE, the column named Type represents how the system call is implemented ie:

1. DCE the method is fully rewritten,
2. NATIVE the real corresponding system call is used.
### Table 3.1: API Coverage

<table>
<thead>
<tr>
<th>System Call Name</th>
<th>Domain</th>
<th>Include</th>
</tr>
</thead>
<tbody>
<tr>
<td>gettimeofday</td>
<td>Date &amp; Time</td>
<td>sys/time.h</td>
</tr>
<tr>
<td>time</td>
<td>Date &amp; Time</td>
<td>time.h</td>
</tr>
<tr>
<td>asctime, ctime, gmtime, localtime</td>
<td>Date &amp; Time</td>
<td>time.h</td>
</tr>
<tr>
<td>asctime_r, ctime_r, gmtime_r, localtime_r, mktime, strftime</td>
<td>Date &amp; Time</td>
<td>time.h</td>
</tr>
<tr>
<td>clock_gettime</td>
<td>Date &amp; Time</td>
<td>time.h</td>
</tr>
<tr>
<td>read</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>write</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>writev</td>
<td>IO</td>
<td>sys/uio.h</td>
</tr>
<tr>
<td>clearerr</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>setbuf, setbuffer, setlinebuf, setvbuf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fseek, ftell, rewind, fgetpos, fsetpos</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>printf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fprintf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>sprintf, snprintf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>asprintf, vasprintf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>dprintf, vdprintf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fgetc, fgetc_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>getc, getc_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>getchar, getchar_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>_IO_getc</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fputc, fputc_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>putc, putc_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>putchar, putchar_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>_IO_putc</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fgetls, fgetls_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fputls, fputls_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>puts</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>ungetc</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fclose</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fcloseall</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fopen, fopen64, fdopen, freopen</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fwrite, fwrite_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>ferror, ferror_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>feof, feof_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fileno, fileno_unlocked</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>perror</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>vprintf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>vfprintf, vsprintf, vsprintf</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>fnct1</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>dup, dup2</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>open, open64</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>close</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>unlink</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>remove</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>mkdir</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>rmdir</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>select</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>isatty</td>
<td>IO</td>
<td>stdio.h</td>
</tr>
<tr>
<td>System Call Name</td>
<td>Domain</td>
<td>Include</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>ioctl</td>
<td>IO</td>
<td>sys/ioct.h</td>
</tr>
<tr>
<td>poll</td>
<td>IO</td>
<td>poll.h</td>
</tr>
<tr>
<td>getcwd, getwd, get_current_dir_name</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>chdir, fchdir</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>alphasort, alphasort64, versionsort</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>umask</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>truncate, fruncate</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>ttyname</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>lseek</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>umask</td>
<td>IO</td>
<td>sys/stat.h</td>
</tr>
<tr>
<td>chdir, fchdir</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>alphasort, alphasort64, versionsort</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>getcwd, getwd, get_current_dir_name</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>chdir, fchdir</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>getcwd, getwd, get_current_dir_name</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>chdir, fchdir</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>alphasort, alphasort64, versionsort</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>umask</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>truncate, fruncate</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>ttyname</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>lseek</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>euidaccess, eaccess</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>pathconf</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>opendir, fdopendir</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>readdir, readdir_r</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>closedir</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>dirfd</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>rewinddir</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>scandir</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>unlinkat</td>
<td>IO</td>
<td>dirent.h</td>
</tr>
<tr>
<td>pread, pwrite</td>
<td>IO</td>
<td>unistd.h</td>
</tr>
<tr>
<td>uname</td>
<td>IO</td>
<td>Kernel</td>
</tr>
<tr>
<td>allocated, malloc, free, realloc</td>
<td>IO</td>
<td>Memory Allocation</td>
</tr>
<tr>
<td>mmap, mmap64, munmap</td>
<td>IO</td>
<td>Memory Mappings</td>
</tr>
<tr>
<td>htonl</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>htons</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>htonl</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>ntohl</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>socket</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getsockopt, setsockopt</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getsockopt, setsockopt</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getsockname</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getpeername</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>bind</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>connect</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>inet_aton, inet_addr, inet_network, inet_ntoa, inet_makeaddr, inet_lnaof, inet_neto</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>inet_ntop</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>inet_aton, inet_addr, inet_network, inet_ntoa, inet_makeaddr, inet_lnaof, inet_neto</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>inet_ntop</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getsockopt, setsockopt</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>listen</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>accept</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>shutdown</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>send, sendto, sendmsg</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>recv, recvfrom, recvmsg</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>gethostbyname, gethostbyname2</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getsockname, gethostbyname</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>gethostbyname, gethostbyname2</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getaddrinfo, freearcinfo, gai_strerror</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>gethostent, gethostent, gethostset, gethostent</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>herror</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getprotoent, getprotobyname, getprotoent, getprotobyname2</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getservent, getservbyname, getservbyport, getservent, endservent</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>if_nameindex</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td>getnameinfo</td>
<td>Networking</td>
<td>Networking</td>
</tr>
</tbody>
</table>

Continued on next page...
<table>
<thead>
<tr>
<th>System Call Name</th>
<th>Domain</th>
<th>Include</th>
</tr>
</thead>
<tbody>
<tr>
<td>ether_aton_r</td>
<td>Networking</td>
<td>netif.h</td>
</tr>
<tr>
<td>atexit</td>
<td>Process</td>
<td>stdlib.h</td>
</tr>
<tr>
<td>getpid, getppid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>getuid, geteuid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>setuid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>setgid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>seteuid, setegid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>setreuid, setregid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>setresuid, setresgid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>sched_yield</td>
<td>Process</td>
<td>sched.h</td>
</tr>
<tr>
<td>exit</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>getenv</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>putenv</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>setenv, unsetenv</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>clearenv</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>fork</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>abort</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>exec, execp, execle, execv, execvp, execve</td>
<td>Process</td>
<td>sys/unistd.h</td>
</tr>
<tr>
<td>wait, waitpid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>sbk</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>getpagesize</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>getgid, getegid</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>gethostname</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>getpgrp</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>pipe</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>__sigsetjmp, siglongjmp</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>getdtablesize</td>
<td>Process</td>
<td>unistd.h</td>
</tr>
<tr>
<td>random, random</td>
<td>Random</td>
<td>stdlib.h</td>
</tr>
<tr>
<td>rand, srand</td>
<td>Random</td>
<td>stdlib.h</td>
</tr>
<tr>
<td>drand48, erand48, lrand48, mrand48, jrand48, srandom, seed48, lcng48</td>
<td>Random</td>
<td>stdlib.h</td>
</tr>
<tr>
<td>rand48_r, erand48_r, lrand48_r, mrand48_r, jrand48_r, srandom_r, seed48_r, lcng48_r</td>
<td>Random</td>
<td>stdlib.h</td>
</tr>
<tr>
<td>signal</td>
<td>Signal</td>
<td>signal.h</td>
</tr>
<tr>
<td>sigaction</td>
<td>Signal</td>
<td>signal.h</td>
</tr>
<tr>
<td>sigemptyset, sigfillset, sigaddset, sigdelset, sigismember</td>
<td>Signal</td>
<td>signal.h</td>
</tr>
<tr>
<td>sigprocmask</td>
<td>Signal</td>
<td>signal.h</td>
</tr>
<tr>
<td>qsort</td>
<td>Sort</td>
<td>string.h</td>
</tr>
<tr>
<td>strerror, strerror_h</td>
<td>String</td>
<td>local.h</td>
</tr>
<tr>
<td>setlocale</td>
<td>String</td>
<td>local.h</td>
</tr>
<tr>
<td>strcoll</td>
<td>String</td>
<td>local.h</td>
</tr>
<tr>
<td>uselocale</td>
<td>String</td>
<td>local.h</td>
</tr>
<tr>
<td>newlocale</td>
<td>String</td>
<td>local.h</td>
</tr>
<tr>
<td>wcscat</td>
<td>String</td>
<td>wchar.h</td>
</tr>
<tr>
<td>btowc</td>
<td>String</td>
<td>wchar.h</td>
</tr>
<tr>
<td>memset</td>
<td>String</td>
<td>string.h</td>
</tr>
<tr>
<td>memccpy</td>
<td>String</td>
<td>string.h</td>
</tr>
<tr>
<td>bcopy</td>
<td>String</td>
<td>string.h</td>
</tr>
<tr>
<td>memcmp</td>
<td>String</td>
<td>string.h</td>
</tr>
<tr>
<td>memmove</td>
<td>String</td>
<td>string.h</td>
</tr>
<tr>
<td>memcmp</td>
<td>String</td>
<td>string.h</td>
</tr>
<tr>
<td>strcpy, strncpy</td>
<td>String</td>
<td>string.h</td>
</tr>
<tr>
<td>strcat, strncat</td>
<td>String</td>
<td>string.h</td>
</tr>
</tbody>
</table>

3.4. Technical Information
<table>
<thead>
<tr>
<th>System Call Name</th>
<th>Domain</th>
<th>Include file</th>
<th>Type</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>strcmp, strncmp</td>
<td>String</td>
<td>string.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strlen, strnlen</td>
<td>String</td>
<td>string.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strspn, strcspn</td>
<td>String</td>
<td>string.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strchr, strchr</td>
<td>String</td>
<td>string.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strcasecmp, strncasecmp</td>
<td>String</td>
<td>string.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strdup, strndup</td>
<td>String</td>
<td>string.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>getopt, getopt_long</td>
<td>String</td>
<td>unistd.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>atoi, atol, atoll, atof</td>
<td>String</td>
<td>stdlib.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strtol, strtoll</td>
<td>String</td>
<td>stdlib.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>strtoul, strtoull</td>
<td>String</td>
<td>stdlib.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>strtod</td>
<td>String</td>
<td>stdlib.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>toupper, tolower</td>
<td>String</td>
<td>ctype.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>index, rindex</td>
<td>String</td>
<td>strings.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strtok, strtok_r</td>
<td>String</td>
<td>string.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>sscanf</td>
<td>String</td>
<td>stdio.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>basename, dirname</td>
<td>String</td>
<td>libgen.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>bindtextdomain, textdomain, gettext</td>
<td>String</td>
<td>libintl.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>mbrlen</td>
<td>String</td>
<td>wchar.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>strtoimax, strtoumax</td>
<td>String</td>
<td>inttypes.h</td>
<td>NATIVE</td>
<td></td>
</tr>
<tr>
<td>openlog, syslog, closelog, vsyslog, setlogmask</td>
<td>Syslog</td>
<td>syslog.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_create</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_exit</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_self</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_once</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_getspecific, pthread_setspecific</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_key_create</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_key_delete</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_mutex_destroy, pthread_mutex_init</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_mutex_destroy, pthread_mutex_init</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_mutexattr_destroy, pthread_mutexattr_init</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_mutexattr_settype</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_cancel</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_kill</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_join</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_detach</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_cond_destroy, pthread_cond_init</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_cond_destroy, pthread_cond_init</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_cond_broadcast, pthread_cond_signal</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_cond_timedwait, pthread_cond_wait</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>pthread_condattr_destroy, pthread_condattr_init</td>
<td>Thread</td>
<td>pthread.h</td>
<td>DCE</td>
<td></td>
</tr>
<tr>
<td>sem_init</td>
<td>Thread Synchronization</td>
<td>semaphore.h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sem_destroy</td>
<td>Thread Synchronization</td>
<td>semaphore.h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sem_post</td>
<td>Thread Synchronization</td>
<td>semaphore.h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sem_wait, sem_trywait, sem_timedwait</td>
<td>Thread Synchronization</td>
<td>semaphore.h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sem_getvalue</td>
<td>Timer</td>
<td>unistd.h</td>
<td>Timer</td>
<td></td>
</tr>
<tr>
<td>sleep, usleep</td>
<td>Timer</td>
<td>unistd.h</td>
<td>Timer</td>
<td></td>
</tr>
<tr>
<td>nanosleep</td>
<td>Timer</td>
<td>unistd.h</td>
<td>Timer</td>
<td></td>
</tr>
<tr>
<td>getitimer, setitimer</td>
<td>Users &amp; Groups</td>
<td>grp.h</td>
<td>Users &amp; Groups</td>
<td></td>
</tr>
<tr>
<td>timerfd_create, timerfd_settime, timerfd_gettime</td>
<td>Users &amp; Groups</td>
<td>sys/time.h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>getgrnam</td>
<td>Users &amp; Groups</td>
<td>sys/time.h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>getusage</td>
<td>Users &amp; Groups</td>
<td>sys/time.h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.5 DCE Python Scripts

Currently DCE includes an experimental support to the Python language. To enable it, you may need to recompile it with the flags:

`--with-pybindgen=HERE_THE_PYBINDGEN_PATH`

indicating the path to an existing Pybindgen source tree to use. Or in case waf didn’t find the interpreter, you can try to use the flags:

`--with-python=HERE_THE_PYTHON_PATH`

The first thing you may want to do is to import the DCE module. For example a minimal DCE script in Python could be:

```python
from ns.DCE import *
print "It works!"
```

3.5.1 A first example

In this example, DCE executes a program running ten seconds on a single node.

```python
# DCE import
from ns.DCE import *
# ns-3 imports
import ns.applications
import ns.core
import ns.network

# Increase the verbosity level
ns.core.LogComponentEnable("Dce", ns.core.LOG_LEVEL_INFO)
ns.core.LogComponentEnable("DceManager", ns.core.LOG_LEVEL_ALL)
ns.core.LogComponentEnable("DceApplication", ns.core.LOG_LEVEL_INFO)
ns.core.LogComponentEnable("DceApplicationHelper", ns.core.LOG_LEVEL_INFO)

# Node creation
nodes = ns.network.NodeContainer()
nodes.Create(1)

# Configure DCE
dceManager = ns.DCE.DceManagerHelper()
dceManager.Install (nodes);
dce = ns.DCE.DceApplicationHelper()

dce.SetBinary ("tenseconds")
# dce.SetStackSize (1<<20)
# dce.Install returns an instance of ns.DCE.ApplicationContainer
apps = dce.Install (nodes )
apps.Start ( ns.core.Seconds (4.0))

# Simulation
ns.core.Simulator.Stop (ns.core.Seconds(20.0))
ns.core.Simulator.Run ()
ns.core.Simulator.Destroy ()
print "Done."
```
You can then run the example with “waf --pyrun ...”

```
./waf --pyrun PATH_TO_YOUR_SCRIPT_HERE
```

or attach gdb to the python script:

```
./waf shell
gdb python -ex "set args PATH_TO_YOUR_SCRIPT_HERE" -ex "handle SIGUSR1 nostop noprint"
```

### 3.5.2 Limitations

The DCE Python bindings does not currently match completely the C++ API of DCE. The following classes are supported:

<table>
<thead>
<tr>
<th>Class</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>DceApplication</td>
<td>GetPid, SetArguments, SetBinary, SetEgid, SetEnvironment, SetEuid, SetGid, SetStackSize, SetStdinFile, SetUid, AddArgument, AddArguments, AddEnvironment, Install, InstallInNode, ParseArguments, ResetArguments, ResetEnvironment, SetBinary, SetEgid, SetEuid, SetGid, SetStackSize, SetStdinFile, SetUid, GetPid</td>
</tr>
<tr>
<td>DceApplicationHelper</td>
<td>GetCmdLine, GetExitCode, GetNode, GetPid, GetRealDuration, GetRealEndTime, GetRealStartTime, GetSimulatedDuration, GetSimulatedEndTime, GetSimulatedStartTime, GetProcStatus, GetVirtualPath, Install, SetAttribute, SetVirtualPath</td>
</tr>
<tr>
<td>ProcStatus</td>
<td></td>
</tr>
<tr>
<td>DceManagerHelper</td>
<td></td>
</tr>
<tr>
<td>Ipv4DceRoutingHelper</td>
<td>Create</td>
</tr>
<tr>
<td>LinuxStackHelper</td>
<td>Install, InstallAll, RunIp, SetRoutingHelper, SysctlGet, SysctlSet</td>
</tr>
</tbody>
</table>
This document is for the people who want to develop DCE itself.

4.1 Kernel Developer Information

This technical documentation is intended for developers who want to build a Linux kernel in order to use with DCE. A first part will describe the architecture and the second will show how we went from a net-next kernel 2.6 has a Linux kernel-stable 3.4.5.

4.1.1 Prerequisite

You must be familiar with ns-3, DCE and Linux Kernel Development.

4.1.2 Download

The source code can be found in the following repository: http://code.nsnam.org/furbani/ns-3-linux. You must use mercurial to download the source code.

4.1.3 Goal

The goal of this work is to use the real implementation of the Linux Network Stack within the Simulation environment furnished by ns-3.

4.1.4 Solution

The solution chosen was to use the Linux kernel source, compile the Net part and make a dynamic library and interface the result with DCE.

The following schema shows the different parts between a software user space application and the hardware network.
The following schema show the same application running under DCE and ns-3 and using a real kernel network stack:

The green parts are implemented in ns-3-linux source files, the grays parts comes from the Linux kernel sources and
are not modified at all or with only few changes. Application should not be modified at all.

4.1.5 Concepts

If you need a more theoretical documentation you can read the chapter 4.5 of this Ph.D. thesis Experimentation Tools for Networking Research.

4.1.6 List of files and usage

After doing the cloning of the source,

```
$ hg clone http://code.nsnam.org/furbani/ns-3-linux
destination directory: ns-3-linux
requesting all changes
adding changsets
adding manifests
adding file changes
added 62 changesets with 274 changes to 130 files
updating to branch default
125 files updated, 0 files merged, 0 files removed, 0 files unresolved
```

Below are the files delivered under the directory ns-3-linux:

```
$ ls ns-3-linux/
generate-autoconf.py generate-linker-script.py kernel-dsmip6.patch kernel.patch Makefile Makefile.print processor.mk README sim
```

The main file is the Makefile its role is to recover the kernel source, compile the NET part of the kernel and all that is necessary for the operation, the end result is a shared library that can be loaded by DCE.

```
$ ls ns-3-linux/sim
cred.c glue.c Kconfig pid.c random.c seq.c sim-socket.c softirq.c tasklet.c
defconfig hrtimer.c Makefile print.c sched.c sim.c slab.c syscall.c tasklet-hrtimer.c
time.c
```

```
$ ls ns-3-linux/sim/include
asm generated sim-assert.h sim.h sim-init.h sim-printf.h sim-types.h
```

These directories contains the architecture specific code and the code doing the interface between the kernel and DCE and ns-3. Ideally we should not change a line of code outside the kernel arch portion, but in practice we make small changes: see the patches files.

Recall: the code of Linux source is mainly C so it is very easy to port to new architecture, the architecture specific code is contained in a specific directory under arch/XXX directory where XXX name recall the processor used. In our case we have chosen to create a special architecture for our environment NS3 + DCE, we called sim.

4.1.7 Interfaces between Kernel and DCE

In order to install a kernel on a Node DCE do the following steps:

1. Load the shared library containing the kernel compilation result,
2. Call the init function called sim_init, this method is located in the just loaded library,
3. This sim_init method is called with a parameter which is a struct containing functions pointers to DCE methods able to be callable from the kernel part,
4. in return the `sim_init` fill a struct containing function pointers in kernel part which will be used by DCE to interact with the kernel part.

5. before finish `sim_init` must initialize the kernel to put it in a running state ready to be usable.

**Kernel -> DCE**

Methods (there is also one variable) of DCE called by the kernels are the following:

- LinuxSocketFdFactory::Vprintf
- LinuxSocketFdFactory::Malloc
- LinuxSocketFdFactory::Free
- LinuxSocketFdFactory::Memcpy
- LinuxSocketFdFactory::Memset
- LinuxSocketFdFactory::Random
- LinuxSocketFdFactory::EventScheduleNs
- LinuxSocketFdFactory::EventCancel
- CurrentNs
- LinuxSocketFdFactory::TaskStart
- LinuxSocketFdFactory::TaskWait
- LinuxSocketFdFactory::TaskCurrent
- LinuxSocketFdFactory::TaskWakeup
- LinuxSocketFdFactory::TaskYield
- LinuxSocketFdFactory::DevXmit
- LinuxSocketFdFactory::SignalRaised
- LinuxSocketFdFactory::PollEvent

there are located in the source file `linux-socket-fd-factory.cc` of DCE.

**DCE -> Kernel**

Methods of Kernel (sim part) called by DCE are the following:

- task_create
- task_destroy
- task_get_private
- sock_socket
- sock_close
- sock_recmmsg
- sock_sendsmsg
- sock_getsockname
- sock_getpeername
• sock_bind
• sock_connect
• sock_listen
• sock_shutdown
• sock_accept
• sock_ioctl
• sock_setsockopt
• sock_getsockopt
• sock_poll
• sock_pollfreewait
• dev_create
• dev_destroy
• dev_get_private
• dev_set_address
• dev_set_mtu
• dev_create_packet
• dev_rx
• sys_iterate_files
• sys_file_read
• sys_file_write

the corresponding sources are located in the sim directory.

4.1.8 Build net-next 2.6 kernel

All build operations are done using the make command with the Makefile file under the directory ns-3-linux.

Make Setup

First you should call make setup in order to download the source of the kernel:

$ make setup
  git clone git://git.kernel.org/pub/scm/linux/kernel/git/davem/net-next.git net-next-2.6; \
  cd net-next-2.6 && git reset --hard fed66381d65a35198639f564365e61a7f256bf79
  Cloning into net-next-2.6...
  remote: Counting objects: 2441000, done.
  remote: Receiving objects: 100% (2441000/2441000), 493.28 MiB | 28.45 MiB/s, done.
  remote: Total 2441000 (delta 2043525), reused 2436782 (delta 2039307)
  Resolving deltas: 100% (2043525/2043525), done.
  Checking out files: 100% (33319/33319), done.
This sources correspond to a specific version well tested with DCE the net-next 2.6 and git tag =
fed63b21653da5198639f564365e61a7f256bf79.
Now the directory net-next-2.6 contains the kernel sources.

Make Menuconfig

Use make menuconfig to configure your kernel, note that modules are not supported by our architecture so options
chosen as modules will not be included in the result kernel.

Build

Finally make will compile all the needed sources and produce a file named libnet-next-2.6.so: this is the library
contains our net-next kernel suitable for DCE usage.

Usage

To use this kernel you should:

1. configure DCE in order to compile using the includes under sim directories to have the good interfaces between
DCE and the kernel. For this you should give to the waf configure the path to the ns-3-linux directory i.e.:

   $ ./waf configure --enable-kernel-stack=/ABSOLUTE-PATH-TO/ns-3-linux

2. In your ns-3 scenario you should indicate the good kernel file: (the file should be located in a directory presents
   in the DCE_PATH env. variable)

dceManager.SetNetworkStack("ns3::LinuxSocketFdFactory", "Library",
   StringValue("libnet-next-2.6.so"));

Test

Use DCE unit test:

   $ ./waf --run "test-runner --verbose"

   PASS process-manager 9.470ms
   PASS Check that process "test-empty" completes correctly. 0.920ms
   PASS Check that process "test-sleep" completes correctly. 0.080ms
   PASS Check that process "test-pthread" completes correctly. 0.110ms
   PASS Check that process "test-mutex" completes correctly. 0.200ms
   PASS Check that process "test-once" completes correctly. 0.070ms
   PASS Check that process "test-pthread-key" completes correctly. 0.070ms
   PASS Check that process "test-sem" completes correctly. 0.080ms
   PASS Check that process "test-malloc" completes correctly. 0.060ms
   PASS Check that process "test-malloc-2" completes correctly. 0.060ms
   PASS Check that process "test-fd-simple" completes correctly. 0.070ms
   PASS Check that process "test-strerror" completes correctly. 0.070ms
   PASS Check that process "test-stdio" completes correctly. 0.240ms
   PASS Check that process "test-string" completes correctly. 0.060ms
   PASS Check that process "test-netdb" completes correctly. 3.940ms
   PASS Check that process "test-env" completes correctly. 0.050ms
   PASS Check that process "test-cond" completes correctly. 0.160ms
   PASS Check that process "test-timer-fd" completes correctly. 0.060ms
   PASS Check that process "test-stdlib" completes correctly. 0.060ms
   PASS Check that process "test-fork" completes correctly. 0.120ms
PASS Check that process "test-select" completes correctly. 0.320ms
PASS Check that process "test-nanosleep" completes correctly. 0.070ms
PASS Check that process "test-random" completes correctly. 0.090ms
PASS Check that process "test-local-socket" completes correctly. 0.820ms
PASS Check that process "test-poll" completes correctly. 0.320ms
PASS Check that process "test-exec" completes correctly. 0.380ms
PASS Check that process "test-iperf" completes correctly. 0.070ms
PASS Check that process "test-name" completes correctly. 0.080ms
PASS Check that process "test-pipe" completes correctly. 0.160ms
PASS Check that process "test-dirent" completes correctly. 0.070ms
PASS Check that process "test-socket" completes correctly. 0.270ms
PASS Check that process "test-bug-multi-select" completes correctly. 0.260ms
PASS Check that process "test-tsearch" completes correctly. 0.080ms

All is OK.

4.1.9 net-next 2.6 to linux-stable 3.4.5

Now we will try to use a more recent linux kernel. We start with a fresh clone of the ns-3-linux sources.

Makefile

First we need to modify the makefile in order to change the kernel downloaded. For that we need to modify the value of 2 variables:

1. KERNEL_DIR=linux-stable
2. KERNEL_VERSION=763c71b1319c56272e42cf6ada6994131f0193a7

Also we need to remove the patch target named .target.ts because the patch will not pass for this newer version of kernel.

First Build

Now we can try to build:

$ make defconfig
$ make menuconfig
$ make
mkdir -p sim
cc -O0 -g3 -D_KERNEL__ -Wall -Wstrict-prototypes -Wno-trigraphs -fno-inline \  
-iwiprefix ./linux-stable/include -DKBUILD_BASENAME="clnt" \  
-fno-strict-aliasing -fno-common -fno-delete-null-pointer-checks \  
-fno-stack-protector -DKBUILD_MODNAME="nsc" -DMODVERSIONS \  
-DEXPORT_SYMTAB -include autoconf.h -U_FreeBSD__ -D_linux__=1 \  
-Dlinux=1 -D__linux=1 -I./sim/include -I./linux-stable/include \  
-fpic -DPIC -D_DEBUG \  
-I/home/furbani/dev/dce/dev/etude_kernel/V3/ns-3-linux \  
-DCONFIG_64BIT -c sim/fs.c -o sim/fs.o

In file included from ./linux-stable/include/asm-generic/bitops.h:12:0, from ./sim/include/asm-generic/bitops.h:4, from ./linux-stable/include/asm/bitops.h:22, from ./linux-stable/include/linux/thread_info.h:52, from ./linux-stable/include/linux/preempt.h:9,
Ok now we will try to fix the compilation errors trying not to change too the kernel source. In the following we will list the main difficulties encountered.

First Error

Recall: the linux source directory include/asm-generic contains a C reference implementation of some code that should be written in assembly language for the target architecture. So this code is intended to help the developer to port to new architectures. So our sim implementation uses many of these include files. The first warning show that our code redefine a method defined elsewhere in kernel sources, so the fix is to remove our definition of this function in our file named sim/include/asm/irqflags.h.

Second Error

The file asm/barrier.h is missing, we just create under sim/include/asm directory and the implementation is to include the generic one ie: include/asm-generic/barrier.h.

Change in sim method

Another problem arise the function named kern_mount_data defined in sim/fs.c do not compile any more. So we need to investigate about this function:

1. Where this function is located in the real code: in linux/fs/namespace.c

2. Why it is reimplemented in sim/fs.c: if you look at our Makefile why try to not compile all the kernel we focus on the net part only, you can see this line in the Makefile:

dirs=kernel/ mm/ crypto/ lib/ drivers/base/ drivers/net/ net/

in fact we include only this directories. So at this time we can comment the failing line and insert a sim_assert (false);

in order to continue to fix the compilation errors, and then when we will do the first run test we will see if this method is called and if yes we will need to do a better fix. Remark: sim_assert (false); is a macro used to crash the execution, we often place it in functions that we need to emulate because required by the linker but that should never be called.

Change in our makefile

After we have the following problem while compiling sim/glue.c the macro IS_ENABLED is not defined. After some search we found that we need to include linux/kconfig.h in many files. So we modify our makefile to fix like this:

```bash
-fno-stack-protector \ 
-DKBUILD_MODNAME="nsc" -DMODULEVERSIONS -DEXPORT_SYMTAB \ 
-include autoconf.h \
```
Change in kernel source

Our sim/slab.c do not compile, in this case we want to use our implementation of memory allocation and to do this it is easier to modify slightly an include file in the kernel sources include/linux/slab.h :

```c
--- a/include/linux/slab.h
+++ b/include/linux/slab.h
@@ -185,6 +185,8 @@ size_t ksize(const void *
    #include <linux/slub_def.h>
    #elif defined(CONFIG_SLOB)
    #include <linux/slob_def.h>
-#elif defined(CONFIG_SIM)
-#include <asm/slab.h>
    #else
    #include <linux/slab_def.h>
    #endif
```

As we have already written we do not recommend to change the kernel sources to facilitate future upgrades.

First Launch

After a few corrections we finally get a library containing the kernel named liblinux-stable.so. At this moment we need to try it using DCE. For the beginning we will try with test-runner executable.

```
./test.py
assert failed. cond="handle != 0", msg="Could not open elf-cache/0/libnet-next-2.6.so elf-cache/0/liblinux-stable.so: undefined symbol: noop_llseek", file=../model/cooja-loader-factory.cc, line=225
terminate called without an active exception
Aborted (core dumped)
```

We can see that a symbol is not defined : noop_llseek. We find this symbol defined in the kernel source named fs/read_write.cc. We need to choose a way to add this symbol in our kernel library, we can:

- rewrite it in a source under our sim directory,
- or add it in our makefile.

In this case we choose the second solution so we need to modify our makefile, first we see that the directory fs is not present in the dirs entry, so we need to add it in the write order (order is the same as found in the kernel Makefile defined by the variable named vmlinux-main); we also need to indicate that we want only the object read_write.o:

```
@@ -51,7 +52,7 @@
    AUTOCONF=autoconf.h
    # note: the directory order below matters to ensure that we match the kernel order
    -dirs=kernel/ mm/ crypto/ lib/ drivers/base/ drivers/net/ net/
+dirs=kernel/ mm/ fs/ crypto/ lib/ drivers/base/ drivers/net/ net/
    empty:=
    space:= $(empty) $(empty)
```

4.1. Kernel Developer Information
Fake Function

We continue to try our kernel library, now another symbol is missing `generic_file_aio_read`, this symbol is defined in the source `mm/filemap.cc`, it is referenced at least by `read_write.c`. In this case we decided to create a fake function because the source `mm/filemap.cc` is voluminous and we do not want to take all the kernel sources. So we create a new source under `sim` directory named `sim/filemap.c` the body of the function is `sim_assert (false)`; so if this function called sometimes we will be warned and we will write a more accurate version.

Assert

Later we meet again the function `kern_mount_data`, thanks to the presence of the `sim_assert`:

```
0x00007ffff5c8c572 in kern_mount_data (fs=<optimized out>, data=<optimized out>) at sim/fs.c:52
52 sim_assert (false);
```

So this function is called by the initialisation, we must provide an implementation for it:

Here we do not want to integrate all the code namespace.c, so we copy and paste the function named `kern_mount_data`. This solution has the advantage of minimizing code size, the disadvantage is that it can introduce problems if the next version of the kernel need changes in this function.

4.1.10 Conclusion

We will not describe the rest of the port here. But after some iteration we end up with a version that works correctly. Sometimes we should not hesitate to use `gdb` to trace the actual execution and correct accordingly code. The rules that we can gain from this experience’s are as follows:

1. Be patient,
2. Try to not modify the kernel sources,
3. Be pragmatic,
4. Try to not import all the kernel code into our library,
5. Do not hesitate to go back and test other alternatives.
4.2 DCE - POLL IMPLEMENTATION

4.2.1 Introduction

The implementation of the poll system call is inspired by the Linux kernel, therefore we will first study the kernel poll implementation.

Kernel implementation

Firstly in the kernel every type of file descriptor (file, socket, pipe ...) must provide a function named poll and conform to this prototype:

```
int poll(struct file *file, poll_table *pwait);
```

Where `file` is a pointer to a structure representing the file (it looks like `this` in C++) and `pwait` is a pointer to a poll table. `pwait` may be NULL, we will see later why. The return integer of this function is a mask of poll events which have already occurred on the corresponding file descriptor. The behavior of this function is as follows: 1. It is not a blocking function, it immediately returns the mask of events regardless of event desired by the caller of poll. 2. if `pwait` is not NULL then it adds `pwait` in the wait queue of `file`, and secondly a pointer to the wait queue is stored in the poll table `pwait`.

Thus an event on the file will ascend to the poll, and in the opposite direction when the poll ends it can de-register itself from the wait queue of the file.

Now that we know the function `poll` of `file`, we can study the `poll` system call, here the following pseudo code commented:

```
POLL( .... )
{
    poll_table table; // This table will contain essentially the list of wait queues that need to wake me,
    // and also information about the current thread in order to be awakened.
    poll_table *pwait=&table; // pointer to current poll table.

    while (true)
    {
        foreach( fd ) // For each file descriptor ...
        {
            file *file = get_file(fd); // Retrieve **file** data structure corresponding to fd.

            if (!file)
                mask = POLLNVAL; // fd does not correspond to an open file.
            else
                mask = file->poll (file, pwait); // During the first loop pwait is not NULL.

            if (mask)
            {
                count++; // Increases the number of responses
                pwait = NULL; // Once we have at least one response POLL should not be blocking,
                // so we nullify the pointer to the poll table in order to not register the
                }
            pwait = NULL; // For the next loops we must not re-register to the wait queue of files.

            if (count) break; // we have a result.
            if (timeout) break; // it is too late.

            Wait(timeout); // Put to sleep until awakening from a file or because of the time limit.
        }
    }
}
```
As we have already seen, the poll will look like that of the kernel. Firstly we create a virtual method named Poll in the class UnixFd. This method will do the same job that the function poll seen early in the struct file of the kernel linux implementation. Before writing the function dce_poll which is our implementation of poll we need to create some classes for mimic the role of the poll table and the wait queues.

So we add 2 sources files named wait-queue.h and wait-queue.cc in order to implements poll table and wait queue.

It is also on this occasion that I deleted all the objects used to wait which was allocated on the stack, and I replaced by objects allocated in the heap. Concerning the dce-poll function it looks like the kernel one with some differences. The more important difference is that the PollTable cannot be allocated on the stack so it cannot be a local variable, so the PollTable object is allocated with a C++ new. I guess you’re wondering why the poll table cannot be allocated on the stack, it is because of the fork implementation of DCE. Indeed, if a process makes a fork, this creates another stack which use the same memory addresses, thus another thread of the same process cannot use an object allocated on this stack, and when a event of a file want to wake up the poll thread it will use especially this poll table. So allocating the Poll Table in the heap generates a side effect which is that we need to release this memory if another thread call exit while we are within the dce-poll. So we need to register the Poll Table somewhere in a DCE data, and the DCE place choosen is the thread struct (in file model/process.h), because each thread can be in doing a poll. Thus there is a new field in struct thread which is:

PollTable *pollTable; // No 0 if a poll is running on this thread

There is another reason to have this field, this reason arises from the fact that a file descriptor can be shared by multiple processes (thanks to dup fork ...), thus when a process exit while doing a poll, we need to deregister from the corresponding wait queues referred by the poll table.

Concerning the kernel implementation the dce-poll method is the same but the difference comes from the Poll method specialized implementation of the class inherited from UnixFd and which correspond to a File Descriptor open with the help of the Kernel Linux. For example the class LinuxSocketFd represents a socket which is opened in the kernel, therefore the method poll of LinuxSocketFdFactory will do much work.

Now look at the interface between DCE and the kernel, in the direction DCE to kernel, we use 2 functions which are sock_poll and sock_pollfreewait, and in the other direction there is sim_poll_event. sock_poll obviously has the same semantics as the kernel poll. sock_poll has the following signature:

void sock_poll (struct SimSocket *s, void *ret);

where s represents the socket int the kernel and ret is a pointer to a data structure of type struct poll_table_ref:

struct poll_table_ref
{
    int ret;
    void *opaque;
};

This structure allows the kernel to pass a reference to the poll table DCE via the opaque field. This reference will be used by the kernel only to warn DCE that event just happened on socket, this using the function sim_poll_event (void
*ref). In return this function modifies the value of opaque and assign it a pointer to a core structure which represents an entry in the wait queue of the socket. This value will be used by DCE for unregister it from the wait queue using the function `sock_pollfreewait function (void *ref)`. The field `ref` is also affected in return and it contain the mask of poll events which have already occurred on the corresponding socket. Most of the kernel code is in the file sim-socket.c it consists of two structures, and the following functions:

<table>
<thead>
<tr>
<th>NAME</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>sim_pollwake</td>
<td>Function called by the kernel when the arrival of an event on the socket, if the event is expected by DCE, the function forwards it to DCE.</td>
</tr>
<tr>
<td>sim_pollwait</td>
<td>Function called by the kernel, its role is to register the poll table in the wait queue.</td>
</tr>
<tr>
<td>sim_sock_poll</td>
<td>Function called by DCE, it is the interface between the DCE’s poll and the kernel’s poll.</td>
</tr>
<tr>
<td>sim_sock_pollfreewait</td>
<td>Function called by DCE allows it to unregister from the wait queue.</td>
</tr>
<tr>
<td>poll_table_ref struct</td>
<td>This is the same struct as that of DCE.</td>
</tr>
<tr>
<td>sim_ptable_entry struct</td>
<td>This is used for the entry in the wait queue of the socket.</td>
</tr>
</tbody>
</table>

# TODO add example, gdb breakpoint to follow the behavior in live

### 4.3 Python Bindings

This section describes how generate the DCE Python bindings. The intended audience are the DCE developers, not the users willing to make simulations. People not interested in adding new public API can ignore this section, and read the chapter *DCE Python Scripts*.

Waf configuration scripts generate the Python bindings in a semi-automatic way. They use PyBindGen to generate a Python script template. This script needs to be manually updated. This intermediate script will generate a C++ source file that can be then compiled as a shared object, installed and imported by the Python scripts.

#### 4.3.1 Step by Step

**Step 1: Api scan**

In this step, *Waf* calls PyBindGen to analyze the DCE public reference headers, in order to generate a temporary template Python file, that will generate a C++ file.

```
./waf --apiscan
```

**Step 2: Pybindgen script update**

The Python script generated in the first step needs to be adjusted and renamed as `ns3_module_dce.py`. In particular to reduce the coupling with the ns-3 installation the waf configuration has been simplified. For this reason you need to paste the correct references for the ns-3 exported symbols, as for example:

```
module.add_class('Object', import_from_module='ns.core')
module.add_class('Node', import_from_module='ns.network')
module.add_class('NodeContainer', import_from_module='ns.network')
```

**Step 3: C++ source code generation**

`ns3_module_dce.py` can directly generate a C++ file for the Python module: `ns3_module_dce.cpp`. It should be added to the code revision system and included in the distribution. It just requires an installed recent version
of Pybindgen.

Step 4: Compilation

In the compilation phase, wscript (the Waf installation module), compiles ns3_module_dce.cpp and generate a shared object called dce.so. This is the Python binding module.

./waf
./waf install

4.3.2 Configuration parameters

waf configuration flags:

--disable-python   Don’t build Python bindings.
--apiscan  Rescan the API for Python bindings. Needs working GCCXML / pygccxml environment.
--with-pybindgen=WITH_PYBINDGEN Path to an existing Pybindgen source tree to use.
--with-python=WITH_PYTHON Path to an existing Python installation
If you are interested to know why DCE exists and how it can work, you should read this document *Experimentation Tools for Networking Research* (and in particular the chapter 4) written by the principal author of DCE, Mathieu Lacage. Also you can read the sources too.

### 5.1 Introduction

You know that the goal of DCE is to execute actual binaries within *ns-3*. More precisely a binary executed by DCE perceives time and space of *ns-3*, rather than the real environment.

To do this, DCE does a job equivalent to an operating system like:

1. DCE loads in memory the code and data of executable,
2. DCE plays the role of intermediary between the executable and the environment through the systems functions called by executables,
3. DCE manages and monitors the execution of processes and handles liberate the memory and close open files when the stop process.
4. DCE manages the scheduling of the various virtual processes and threads.

### 5.2 Main classes and main data structures

#### 5.2.1 DceManager

The **DceManager** is somewhat the entry point of DCE. It will create virtual processes and manages their execution. There is an instance of **DceManager** which is associated to each node which need to virtualize the execution of a process. In reality, the developer uses the classes **DceManagerHelper** and **DceApplicationHelper**.

I invite you to look at the source code `dce-manager.cc` and `dce-manager.h` and particularly the public methods **Start**, **Stop**, **Exit**, **Wakeup**, **Wait** and **Yield**; the following private methods are also important: **CreateProcess**, **PrepareDoStartProcess**, **DoStartProcess**, **AllocatePid**, **TaskSwitch**, **CleanupThread** and **LoadMain**. The **Start** method is called when starting the executable, if you look at, it begins by initializing an object of type **struct Process**. **struct Process** is very important, it contains information about the virtual processes that DCE creates, this type is described below. **Start** then also initializes a structure of type **struct thread**, it represents the principal thread in which the **main** entry of the executable will run. Finally **Start** asks the **TaskManager** to create a new **Task** and to start this one using the method **DceManager::DoStartProcess** as the entry point.

Class **TaskManager** is a major class of DCE, it is described below.
5.2.2 Process

`struct process` contains everything you need to describe a process, I invite you to study the source code in the file `process.h`.

This structure contains references to standard objects for example a list of open files via a vector of `FILE *`, but it contains also the references to objects useful to manage DCE threads, the memory allocated by the process...

Field `openFiles` represents the open file descriptors of the process, the key is the fd and the value a pointer to an object of type DCE `FileUsage`. The field `threads` contains all the threads in the process see description below. Field `loader` is a pointer to the `Loader` used to load the corresponding code.

The `alloc` field is important it is a pointer to the memory allocator used to allocate the memory used by this process, at the end of the process it will liberate all the memory allocated so as simple and efficient.

5.2.3 Thread

`struct thread` represents a thread. It contains various fields including a pointer to the `process` to which it belongs, and also a pointer to `Task` object described later.

5.2.4 Taskmanager and Task

The `TaskManager` manages the Tasks, ie the threads of virtualized processes by DCE. It allows you to create new task. It controls the activity of the task by the following methods: `Stop`, `Wakeup`, `Sleep` and `Yield`. A `Task` possesses a stack which contains the call stack functions. There is one instance of `TaskManager` per node. The implementation of `TaskManager` is based on a class of type `FiberManager` described below.

5.2.5 FiberManager

`FiberManager` actually implements the creation of the execution contexts. These contexts are called fiber. `FiberManager` offers the following:

1. Create a context, it returns a fiber
2. Delete a fiber
3. Yield hand to another fiber

DCE provides two implementations:

1. `PthreadFiberManager`, which is based on the pthread library,
2. `UcontextFiberManager` which is based on the POSIX API functions offered by ucontext.h: `makecontext`, `getcontext` and `setcontext`.

I invite you to watch the corresponding man.

5.2.6 LoaderFactory and Loader

The `Loader` is a very important object of DCE. A DCE `Loader` loads the executable code in memory of a special way, load several times the same executable, while isolating each of the other executable. The `Loader` must link the executable loaded with the 3 emulated libraries, i.e., lib C, lib pthread and lib rt. The same way the libraries used by the executable must also be linked with the emulated libraries.

DCE offers several actually Loader:
1. **CoojaLoader**: it has the following characteristics: it loads into memory only a copy of the code, by cons it duplicates data (i.e., global variables and static). For each change of context there are 2 memory copies: backup data of the current context then restoration of context memory that will take control. Comment: it is rather reliable, the size of the copied memory size depends on the total static and global variables, and in general there is little, in a well-designed executable.

2. **DlmLoader**: Uses a specialized loader to not duplicate the code but only the data but without special operations to do when changing context. Comment: offers the best performance in memory and cpu, but not very reliable especially during the unloading phase.

### 5.3 Follow a very simple example

After theory, do a bit of practice. Follow the execution of very simple example.

You can find the used sample under the directory named `myscripts/sleep`. This executable used by the scenario do only a sleep of ten seconds:

```c
#include <unistd.h>

int main(int c, char ***v)
{
    sleep (10);

    return 1;
}
```

The *ns-3/DCE* scenario execute **tenseconds** one time starting at time zero:

```cpp
#include "ns3/core-module.h"
#include "ns3/dce-module.h"

using namespace ns3;

int main (int argc, char *argv[])
{
    NodeContainer nodes;
    nodes.Create (1);

    DceManagerHelper dceManager;
    dceManager.Install (nodes);

    DceApplicationHelper dce;
    ApplicationContainer apps;
    dce.SetStackSize (1<<20);
    dce.SetBinary ("tenseconds");
    dce.ResetArguments ();
    apps = dce.Install (nodes.Get (0));
    apps.Start (Seconds (0.0));

    Simulator::Stop (Seconds(30.0));
    Simulator::Run ();
    Simulator::Destroy ();
}
```

First we can launch the binary with `$ ./build/bin_dce/tenseconds`. After ~10 seconds you retrieve the prompt.
Then we can try the DCE scenario: .. code-block:: sh

    $ ./build/myscripts/sleep/bin/dce-sleep

This time the test is almost instantaneous, because the scenario is very simple and it uses the simulated time.

Same test by activating logs:

We can see that an event occurs at 30s it is the end of the simulation corresponding to the line:

    Simulator::Stop (Seconds(30.0));

We can also see that at 10s an event occurs, this is the end of our sleep(10).

Now we do the same experiment using the debugger:

You can notice that:

1. We have two breakpoints.
2. After run the first stop is in **ns3::DceManager::Start**.
3. At this time there is only one thread.
4. We are currently processing an event, this event was scheduled by the call **apps::Start (Seconds (0.0))**; of our scenario.

Now we continue our execution:

    (gdb) continue
Continuing.

    [New Thread 0x7ffff65fc700 (LWP 8159)]
    [Switching to Thread 0x7ffff65fc700 (LWP 8159)]

Breakpoint 3, ns3::DceManager::DoStartProcess (context=0x633d50) at ../model/dce-manager.cc:274
274 Thread *current = (Thread *)context;

(gdb) info thread

* 2 Thread 0x7ffff66fc700 (LWP 8159) "dce-sleep" ns3::DceManager::DoStartProcess (context=0x633d50)
  1 Thread 0x7ffff6600740 (LWP 7977) "dce-sleep" pthread_cond_wait@@GLIBC_2.3.2 () at ../nptl/sysdeps/unix/sysv/linux/x86_64/pthread_cond_wait.S:162

(gdb) bt

# 0 ns3::DceManager::DoStartProcess (context=0x633d50) at ../model/dce-manager.cc:274
# 1 0x000000000077d21b90 in ns3::TaskManager::Trampoline (context=0x633bd0) at ../model/task-manager.cc:250
# 2 0x000000000077d166a84 in ns3::PthreadFiberManager::Start (arg=0x634040) at ../model/pthread-fiber-manager.cc:402
# 3 0x000000000034be206cc6 in start_thread (arg=0x7ffff65fc700) at pthread_create.c:301
# 4 0x000000000034bd6e0c2d in clone () at ../sysdeps/unix/sysv/linux/x86_64/clone.S:115

You can notice that:

1. Now there is a second thread
2. Gdb break execution in **ns3::DceManager::DoStartProcess** in the context of the second thread

This second thread is the thread corresponding to the main thread of our hosted executable **tenseconds**, if you look at **ns3::DceManager::DoStartProcess** you can notice that we are on the point of calling the main of **tenseconds**.

You can also see that the pointer to the **main** is the result of the method **ns3::DceManager::PrepareDoStartProcess**. Now we can put a breakpoint before the sleep of **tenseconds** and follow the code of sleep:

    (gdb) break tenseconds.c:5
    (gdb) continue
Breakpoint 1, main (c=1, v=0x630b30) at ../myscripts/sleep/tenseconds.c:5
5 sleep (10);
    (gdb) list
    (gdb) step
sleep () at ../../../model/libc-ns3.h:193
193 DCE (sleep)
(gdb) step
dce_sleep (seconds=10) at ../../../model/dce.cc:226
226 Thread *current = Current ();
(gdb) list
224 unsigned int dce_sleep (unsigned int seconds)
225 {
226    Thread *current = Current ();
227    NS_LOG_FUNCTION (current << UtilsGetNodeId ());
228    NS_ASSERT (current != 0);
229    current->process->manager->Wait (Seconds (seconds));
230    return 0;
231 }
(gdb) bt
#0  dce_sleep (seconds=10) at ../../../model/dce.cc:226
 #1 0x00007ffff62cdcb9 in sleep () at ../../../model/libc-ns3.h:193
 #2 0x00007ffff5c36725 in main (c=1, v=0x630b30) at ../../../myscripts/sleep/tenseconds.c:5
 #3 0x00007ffff7c9b0bb in ns3::DceManager::DoStartProcess (context=0x633d50) at ../../../model/dce-manager.cc:281
 #4 0x00007ffff7d21b90 in ns3::TaskManager::Trampoline (context=0x633bd0) at ../../../model/task-manager.cc:267
 #5 0x00007ffff7da87 in ns3::PthreadFiberManager::Run (arg=0x634040) at ../../../model/pthread-fiber-manager.cc:402
 #6 0x00000034be206ccb in start_thread (arg=0x7ffff65fc700) at pthread_create.c:301
 #7 0x00000034bb6e0c2d in clone () at ../../../sysdeps/unix/sysv/linux/x86_64/clone.S:115
(gdb) info thread
Id Target Id Frame
* 2 Thread 0x7ffff6600740 (LWP 15233) "dce-sleep" dce_sleep (seconds=10) at ../../../model/dce.cc:226
1 Thread 0x7ffff6600740 (LWP 15233) "dce-sleep" pthread_cond_wait@@GLIBC_2.3.2 () at ../../../nptl/sysdeps/unix/sysv/linux/x86_64/clone.S:310
(gdb)
We can notice that sleep call dce_sleep which call Wait, this Wait method is from the class TaskManager. TaskManager is a major class of DCE and we will detail it below. Basically Wait schedules and event in ns-3 event queue (in order to be woken up after sleep time) and give the control to another Task. Now we can put a breakpoint in ns3::DefaultSimulatorImpl::ProcessOneEvent and see the time advance up to 10s:

gdb) b ns3::DefaultSimulatorImpl::ProcessOneEvent
Breakpoint 2 at 0x7ffff7619207: file ../../core/model/default-simulator-impl.cc, line 131.
(gdb) c
Continuing.
[Switching to Thread 0x7ffff6600740 (LWP 3942)]
Breakpoint 2, ns3::DefaultSimulatorImpl::ProcessOneEvent (this=0x6308e0) at ../../core/model/default-simulator-impl.cc:131
warning: Source file is more recent than executable.
131 Scheduler::Event next = m_events->RemoveNext ();
(gdb) n
133 NS_ASSERT (next.key.m_ts >= m_currentTs);
(gdb) n
134 m_unscheduledEvents--;
(gdb) n
136 NS_LOG_LOGIC ("handle " << next.key.m_ts);
(gdb) n
137 m_currentTs = next.key.m_ts;
(gdb) n
138 m_currentContext = next.key.m_context;
(gdb) p m_currentTs
$1 = 10000000000
(gdb)
This next event will wake the thread 2 will therefore complete the sleep of our scenario.
In summary we saw briefly that DCE uses the events of ns-3 to schedule the execution between different tasks.

5.3. Follow a very simple example

57
Below are the list of the tested applications with DCE.

- CCNx
- Quagga
- iperf
- ping/ping6
- ip (iproute2 package)
- umip (Mobilt IPv6 daemon)
- Linux kernel (from 2.6.36 to 3.14 version)
- FreeBSD kernel (10.0.0 version)
- thttpd
- torrent

### 6.1 CCNx

CCNx is an open source implementation of **Content Centric Networking**. All the C written executables are supported and none of the Java ones. The versions tested are those between 0.4.0 and 0.6.1 included.

For more detail document, see [dce-ccnx](#).

### 6.2 Quagga

Quagga is a routing software suite, providing implementations of OSPFv2, OSPFv3, RIP v1 and v2, RIPng and BGP-4 for Unix platforms, particularly FreeBSD, Linux, Solaris and NetBSD. More information.

### 6.3 iperf

iperf from the following archive [http://walami.googlecode.com/files/iperf-2.0.5.tar.gz](http://walami.googlecode.com/files/iperf-2.0.5.tar.gz) as been tested. It is the exception that proves the rule. That is to say that this particular example requires a change in its code. In the source file named Thread.c at line 412 in the function named `thread_rest` you must add a `sleep(1)` in order to help DCE to break the infinite loop.
6.4 ping/ping6

Ping from the following archive http://www.skbuff.net/iputils/iputils-s20101006.tar.bz2 is supported.

6.5 ip (iproute2 package)

6.6 umip (Mobilt IPv6 daemon)

The umip (Usagi-Patched Mobile IPv6 stack) support on DCE enables the users to reuse routing protocol implementations of Mobile IPv6. UMIP now supports Mobile IPv6 (RFC3775), Network Mobility (RFC3963), Proxy Mobile IPv6 (RFC5213), etc, and can be used these protocols implementation as models of network simulation.

For more information, see the latest support document.

6.7 Linux kernel (from 2.6.36 to 3.14 version)

Linux kernel support is built with a separate ‘dce-linux’ module, available on github. Many protocols implemented in kernel space such as TCP, IPv4/IPv6, Mobile IPv6, Multipath-TCP, SCTP, DCCP, etc, are available with ns-3.

6.8 FreeBSD kernel (10.0.0 version)

FreeBSD kernel support is based on Linux kernel module of DCE. A few protocols implemented in kernel space such as TCP, IPv4, etc, are available with ns-3.

6.9 thttpd

(TBA)

6.10 torrent

(TBA)
The creator of DCE is Mathieu Lacage.
The current maintainers and developers are Hajime Tazaki and Frédéric Urbani.

7.1 Contacts

tazaki AT sfc.wide.ad.jp
mathieu.lacage AT cutebugs.net