

# NDK+OFM: Rapid Development of Accelerated Applications for FPGA SmartNICs

Jiří Matoušek, Daniel Kondys CESNET

8 May 2023 (at FCCM 2023 in Los Angeles, CA, USA)

# cesnet

## TUTORIAL ORGANIZERS

### Jiří Matoušek

- researcher at CESNET
- assistant professor at FIT BUT (Brno, Czech Republic)

# Daniel Kondys

developer in FPGA R&D team at CESNET

### Jakub Cabal

head of the FPGA R&D team at CESNET









- CESNET is an association of universities of the Czech Republic and the Czech Academy of Sciences
- Operates and develops the national e-infrastructure for science, research and education which encompasses
  - computer network
  - computational grids
  - data storage
  - collaborative environment





# LIBEROUTER TEAM

a.k.a. Security and Administration Tools Department at CESNET

- More than 60 people in R&D (about 28 FTE)
- Cooperation with universities (CTU, BUT) and students
- Applied research and tools based on unique technology
  - hardware acceleration
  - network monitoring & traffic analysis
  - Cyber Threat Intelligence
  - DDoS mitigation
  - configuration tools



# cesnet

# **RESULTS AND SUCCESS STORIES**

### One of the first 100 Gb FPGA acceleration cards

- Czech Head Prize 2016 Industrie award
- NETCONF tools awarded by ONF
  - now used by telco operators in the network infrastructure
- P4 compiler for FPGAs acquired by Intel
- Flowmon Networks spin-off acquired by Kemp
  - founded in 2007 based on the research results of Scampi and Geant projects (Netflow/IPFIX probes)
- DDoS mitigation system deployed by NIX











### WHAT ABOUT YOU?

### What is your experience with FPGA SmartNICs?

Why did you register for this tutorial?

What are your expectations?

What would you like to take away?





- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics

AGENDA



# **NDK Introduction**

[presentation]



# MOTIVATION

- Growing number of available/deployed FPGA SmartNICs
  - Intel, AMD/Xilinx, Cisco, Napatech, Silicom, BittWare, etc.
- Need for rapid development of applications targeting these devices
- Common low-level operations on FPGA SmartNICs
  - transmitting/receiving data via network interface (typically Ethernet)
  - transmitting/receiving data via host interface (typically PCIe)
  - writing/reading data via external memory interface (typically DDR)
- Implementation of low-level operations is difficult for users



### Network Development Kit (NDK)

- open-source framework for fast and easy development of accelerated applications for FPGA SmartNICs
- based on 20+ years of active research and development at CESNET

### NDK key features

- support for high-speed network interfaces (up to 400 GbE)
- support for high-speed host interfaces (up to PCIe gen5 x16)
- ready-to-use software stack (driver, library, tools) and API (C, Python)
- Data Plane Development Kit (DPDK) support
- single make command to create the entire FPGA bitstream



## NDK OVERVIEW

Software stack

### FPGA design architecture





## FPGA DESIGN ARCHITECTURE



### FPGA architecture modules

- user application (e.g., minimal) (not part of NDK)
- network module (e.g., F-Tile)
- DMA module (e.g., Medusa) (not open-source)
- PCIe module (e.g., R-Tile)
- memory controller (e.g., DDR4)
- timestamp generator
- C/S registers interconnect



## SUPPORTED FPGA CARDS

### NDK FPGA design architecture is divided into

- common part same for all supported FPGA cards (<u>ndk-core</u>)
- specific part for each supported FPGA card
- Supported FPGA cards
  - ReflexCES XpressSX AGI-FH400G (<u>ndk-card-agi-fh400g</u>)
  - Intel Stratix 10 DX FPGA Development Kit (<u>ndk-card-dk-dev-1sdx-p</u>)
  - Intel Agilex I-Series FPGA Development Kit (<u>ndk-card-dk-dev-agi027res</u>)
  - Silicom fb4CGg3@VU9P and fb2CGg3@VU9P (both in <u>ndk-card-fb4cgg3</u>)
  - Silicom fb2CGhh@KU15P (<u>ndk-card-fb2cghh</u>)
  - Bittware IA-420F card (ndk-card-ia-420f)



# FURTHER FIRMWARE PARTS OF NDK

Various basic modules for FPGA designs (<u>ofm</u>)

- see <u>OFM Introduction</u> part of this tutorial
- Minimal/reference NDK application (<u>ndk-app-minimal</u>)
  - see <u>ndk-app-minimal Introduction</u> part of this tutorial
- Build system for easy simulation/verification, synthesis, implementation, and bitstream generation (<u>ofm/build</u>)
  - see <u>Advanced Topics</u> part of this tutorial

# cesnet

## SOFTWARE STACK

### Software stack components

- Linux kernel driver
   (nfb = NDK FPGA board)
- userspace library with C/Python API (libnfb)
- NDK SW tools
- DPDK driver
- user application (not part of NDK)





# **CONFIGURATION SW TOOLS**

- <u>nfb-tools</u> (nfb = NDK FPGA board)
- Tools to read/write configuration and status data via CSR bus
  - nfb-info basic info about card and current design
  - nfb-boot load bitstream into card
  - nfb-eth configure/read status of network module
  - nfb-dma configure/read status of DMA module
  - nfb-tsu configure timestamp generator
  - nfb-bus configure/read status of arbitrary register





## DATA TRANSFER SW TOOLS

- <u>ndp-tools</u> (ndp = NDK Data Plane)
- Tools to read/write packet data using DMA
  - ndp-read read packets from FPGA
  - ndp-generate send generated packets to FPGA
  - ndp-receive read packets from FPGA into PCAP
  - ndp-transmit send packets from PCAP to FPGA
  - ndp-loopback read packets from FPGA and send them back





### NDK SW PACKAGES

### RPM packages available via <u>COPR</u>

sudo dnf copr enable @CESNET/nfb-framework
sudo dnf install nfb-framework python3-nfb

- nfb driver
- libnfb library (including C and Python API)
- both nfb-\* and ndp-\* tools
- NDK software can also be built from sources (<u>ndk-sw</u>)
  - follow <u>Build instructions</u> in README.md
- DPDK driver available via mainline DPDK repository



# DESCRIPTION OF HW FOR SW

### Design in FPGA is characterized using <u>DeviceTree</u> (DT)

 build system composes DT string (dts) and translates it to DT blob (dtb), which is stored in the design and can be read from software

# DTS example

```
ref_name: my_comp {
    reg = <$BASE_ADDRESS 0x40>;
    compatible = "netcope,my_comp";
    version = <0x00010004>;
    type = "reduced";
};
```



# EXAMPLES OF NDK APPLICATIONS

### Precise network monitoring

- flow monitoring with deep packet inspection
- Network security applications
  - IDS/IPS acceleration Suricata pre-filter and bypass
  - Anti DDoS mitigation of volumetric attacks
- High-frequency trading
  - algorithmic trading with very low response delay
- And many others...



## RELATED PUBLICATIONS

- J. Cabal, J. Sikora, Š. Friedl, M. Špinler, and J. Kořenek, "<u>FPL Demo:</u> <u>400G FPGA Packet Capture Based on Network Development Kit</u>," FPL 2022, pp. 474-474, doi: 10.1109/FPL57034.2022.00090.
- J. Kubálek, J. Cabal, M. Špinler, and R. Iša, "<u>DMA Medusa: A</u> <u>Vendor-Independent FPGA-Based Architecture for 400 Gbps DMA</u> <u>Transfers</u>," FCCM 2021, pp. 258-258, doi: 10.1109/FCCM51124.2021.00045.
- L. Kekely, J. Cabal, V. Puš and J. Kořenek, "<u>Multi Buses: Theory and</u> <u>Practical Considerations of Data Bus Width Scaling in FPGAs</u>," DSD 2020, pp. 49-56, doi: 10.1109/DSD51259.2020.00020.



AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation] OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# ndk-app-minimal Introduction

[presentation]



### NDK-APP-MINIMAL

- <u>ndk-app-minimal</u> represents simple example of NDK-based FPGA application
  - possible starting point for user-defined FPGA applications based on NDK
- Does not process network packets, just routes them through application core
- Documentation
  - <u>The Application</u> section within NDK Architecture
  - Minimal NDK application

# cesnet

## THE APPLICATION

### ETH streams

- MVB+MFB interfaces
- DMA streams
  - MVB+MFB interfaces
- APP registers interface
  - MI interface
- EMIF interfaces
  - Avalon-MM interfaces
- Clock INs and OUTs





## THE APPLICATION

### The application interfaces

- details of data format can be found in <u>documentation</u>
- proprietary interfaces MVB, MFB, and MI will be introduced in <u>OFM</u>
   <u>Introduction</u> part of this tutorial
- DMA Streams Merger + DMA Chan Mod
  - deal with situation when ETH streams  $\neq$  DMA streams
- MI Async
  - clock-domain-crossing unit for MI interface

# cesnet

## MINIMAL NDK APPLICATION

- Single Base
   Application Subcore
   per ETH stream
- Single Memory Tester per EMIF interface
- Separate MI interface for each
  - Base Application Subcore
  - Memory Tester



# cesnet

### MINIMAL NDK APPLICATION

### MFB only TX ETH interfaces

 other interfaces utilize MVB+MFB



- TX direction channels (DMA→ETH)
  - Static mapping

- RX direction channels (ETH→DMA)
  - Dynamic mapping (configurable by user)



### <u>ndk-app-minimal</u> integrates all other NDK repositories

- app/ implementation of application
- build/ build scripts for supported FPGA cards
- conf/ build system configuration
- doc/ documentation system + top-level documentation
- ndk/ NDK platform sources (multiple repositories)
  - cards/ top-level designs for supported FPGA cards (e.g., ndk-card-agi-fh400g)
  - core/ common NDK core (ndk-core)
  - modules/ special modules and IPs (e.g., ndk-mod-dma-medusa)
  - ofm/ Open FPGA Modules basic components and build system (ofm)
- tests/ test scripts for Jenkins, etc.

note that NDK cannot be built without application core



AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation] OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# Playing with ndk-app-minimal

[hands-on]



# **CONNECTION SETUP**

- 1. Start provided VirtualBox VM
- 2. Open terminal
- 3. Connect via SSH to given "preklad" server using given credentials
  - ssh tutorial{1-5}@preklad{11-13, 21-24}.liberouter.org





# **CLONING REPOSITORIES**

- Read through "<u>How to start</u>" section of ndk-app-minimal README.md
- 5. Clone <u>ndk-app-minimal</u> repository with at least <u>ofm</u>, <u>core</u>, and <u>agi-fh400g</u> submodules
  - git clone https://github.com/CESNET/ndk-app-minimal.git
  - cd ndk-app-minimal
  - git submodule update --init ndk/ofm
  - git submodule update --init ndk/core
  - git submodule update --init ndk/cards/agi-fh400g



# **BUILDING DESIGN**

- 6. Read through "<u>How to start</u>" section of ndk-app-minimal documentation
- 7. Build default ndk-app-minimal design for agi-fh400g card
  - cd build/agi-fh400g
  - make
- In the meantime, study details of Gen Loop Switch module (next slide) and read through "<u>NDK testing</u>" section of ndk-app-minimal documentation



# GEN LOOP SWITCH (GLS)

### Module for easy debugging and measurements of FPGA designs

<u>documentation</u> available for instance as part of ndk-app-minimal documentation

### Basic features

- two configurable packet generation modules
- two configurable loopback paths
- four units for throughput measurements
- By default, GLS module is enabled in ndk-app-minimal
- GLS operation can be controlled using available <u>Python script</u>



# **BOOTING DESIGN**

- **9.** Once the demo server Solaris is available, ask Daniel Kondys for supervision
- Use nfb-boot tool to load default ndk-app-minimal design into 400G FPGA card in Solaris server
  - nfb-boot -f0 your\_ndk\_firmware.nfw
- 11. Use nfb-info tool for verification of successful design booting
  - nfb-info



### 12. Test R/W access to the scratch registers

- nfb-bus -p /firmware/mi\_bus0/mi\_test\_space <address>
- nfb-bus -p /firmware/mi\_bus0/mi\_test\_space <address> <data>
- nfb-bus -p /firmware/mi\_bus0/mi\_test\_space <address>
- 13. Enable Network Module in "PMA local loopback mode"
  - nfb-eth -Pc "+PMA local loopback"
  - nfb-eth -e1
- 14. Measure ndk-app-minimal throughput using GLS module
  - python3 gls\_mod.py 1



### AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# [lunch break]



### AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# **OFM** Introduction

[presentation]



### Open FPGA Modules (OFM)

- open-source library of basic FPGA modules for (not only) high-speed processing of data
- heavily utilized in NDK and ndk-app-minimal
- based on 20+ years of active research and development at CESNET
- Most common interfaces of modules within OFM repository
  - multi buses (MFB, MVB)
  - memory interface bus (MI)



# **MULTI BUSES**

- New generation of highly parametric proprietary buses
- Key benefit: support for multiple packets per clock cycle
  - necessary for processing 400G Ethernet data streams in FPGA
- Two different buses that are often used together
  - Multi-Frame Bus (MFB): designed for transfers of packet data
  - Multi-Value Bus (MVB): designed for transfers of (packet) meta-data
- Available documentation introduces MFB and MVB very well
  - MFB specification
  - MVB specification



# MEMORY INTERFACE BUS

### Memory Interface Bus (MI bus)

- low-performance bus for SW access to control/status registers and memories
- sometimes mentioned as MI32 bus (due to its default data width of 32 bits)
- Available documentation introduces MI bus very well
  - MI bus specification





# **BASIC COMPONENTS (TOOLS)**

set of basic components for use with common buses

- implementation: <u>comp/mfb\_tools</u>, <u>comp/mvb\_tools</u>, <u>comp/mi\_tools</u>, etc.
- documentation: <u>MFB Tools</u>, <u>MVB Tools</u>, <u>MI Tools</u>, etc.
- not all components are documented!
- similar components stored together within each group
  - data flow transformations (flow)
  - data storage (storage)
  - debugging (debug)
  - etc.





# **BASIC COMPONENTS (BASE)**

### set of basic components for general use

- implementation: <u>comp/base</u>
- documentation: <u>Basic Tools</u> (not all components are documented!)
- some components are optimized for specific FPGA architecture(s)
- similar components stored together
  - combinatorial logic (logic)
  - shift registers (shreg)
  - general (mem) and FIFO (fifo) memories



■ etc.



### AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# Implementing Custom NDK Application

[hands-on]



### HANDS-ON TASK





# HIGH-LEVEL ASSIGNMENT

- Extend ndk-app-minimal with 64-bit counter of RX MFB frames
  - use MFB\_FRAME\_CNT (not documented) from ofm/mfb\_tools directory
- Make the counter value accessible via MI bus at address offset 0x100 from Base Application Subcore base address
  - use MI\_SPLITTER\_PLUS\_GEN to split address space of Base Application
     Subcore between MVB\_CHANNEL\_ROUTER\_MI and MFB\_FRAME\_CNT
- Make sure to edit all relevant files
  - app/top/app\_subcore.vhd
  - app/top/Modules.tcl
  - app/top/DevTree.tcl



### AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# SW API Introduction

[presentation]



## NDK SW API

### NDK userspace library provides two APIs for user applications

- C API: native libnfb API
- Python API: Cython-based libnfb API
- Typical use cases of C API
  - applications with high-performance requirements
  - NDK SW tools (both nfb-tools and ndp-tools)
- Typical use cases of Python API
  - majority of user applications
  - testing/debugging scripts



### Comprises two sets of API functions

- <u>Base API</u>: basic operations over NFB (init/deinit, registers read/write, etc.)
- <u>NDP API</u>: manipulation of high-speed queues for packet transmissions

### Further materials

- <u>quick start guides</u> as well as <u>complete examples</u> for both Base and NDP API
- example use of Base and NDP API in <u>nfb-tools</u> and <u>ndp-tools</u>, respectively

# cesnet

# C API EXAMPLES

### Base API example

#include <nfb/nfb.h>

```
int main(int argc, char *argv[]) {
    struct nfb_device *dev = nfb_open("0");
```

```
int node = nfb_comp_find(dev, "netcope,rxmac", 0);
struct nfb_comp *comp = nfb_comp_open(dev, node);
```

```
const int RXMAC_EN = 0x20;
int en = nfb_comp_read32(comp, RXMAC_EN) & 0x01;
if (!en)
```

nfb\_comp\_write32(comp, RXMAC\_EN, 0x01);

```
nfb_comp_close(comp);
nfb_close(dev);
return 0;
```

### NDP API example

#include <nfb/nfb.h>
#include <nfb/ndp.h>

int main(int argc, char \*argv[]) {
 struct nfb\_device \*dev = nfb\_open("0");

struct ndp\_queue \*txq = ndp\_open\_tx\_queue(dev, 0);
ndp\_queue\_start(txq);

ndp\_close\_tx\_queue(txq); nfb\_close(dev); return 0;



# **PYTHON API**

### Work-in progress, but already useable for real applications

- basic operations over NFB
- manipulation of high-speed queues for packet transmissions
- ethernet port manipulation
- Aims to provide at least the same functionality as C API
- Documentation is to be published in the near future, however
  - <u>complete examples</u> are already available within ndk-sw repository
  - OFM contains handful or real use cases: <u>Memory Tester</u>, <u>Rate Limiter</u>, etc.

# cesnet

# PYTHON API EXAMPLES

### Basic example

### import nfb

dev = nfb.open()

node = dev.fdt\_get\_compatible("netcope,eth")[0]
phandle = node.get\_property("rxmac").value
node = dev.fdt\_get\_phandle(phandle)

```
comp = dev.comp_open(node)
```

RXMAC\_EN = 0x20 en = comp.read(RXMAC\_EN) & 0x01 if (not en): comp.write(RXMAC\_EN, 0x01)

### Data transfer example

import nfb

dev = nfb.open() ndp = dev.ndp

txq = ndp.tx[0] txq.start()

pkt = bytes([0]\*128) hdr = bytes([0]\*16) txq.send([pkt], hdrs=hdr)





### AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# Controlling NDK Application from SW

[hands-on]



# CONTROLLING VIA EXISTING TOOLS

- Test extended ndk-app-minimal using GLS module in mode 1
  - use design from "Implementing Custom NDK Application" hands-on session
  - follow the same steps as in "<u>Playing with ndk-app-minimal</u>" hands-on session
- Read out the number of packets generated in GLS module
  - nfb-bus -p /firmware/mi\_bus0/dbg\_gls0/mfb\_gen2eth 0x20
- Read out the number of packets observed by counter of RX MFB frames in Base Application Subcore
  - nfb-bus -p /firmware/mi\_bus0/application/app\_core\_minimal\_0/<name> 0x0
- Compare numbers read out from GLS module and RX MFB frame counter



# CONTROLLING VIA PYTHON API

 Implement Python script that performs the same operations as specified on the previous slide

- Useful Python API examples
  - <u>01-basics.py</u>
  - <u>03-eth.py</u>
- Not all operations can be done via Python API
  - where necessary, use module <u>os</u> allowing to execute command in subshell

import os
os.system(<command>)



### AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# **Advanced Topics**

[presentation]



# **BUILD SYSTEM**

- Set of scripts for simple simulation, verification, synthesis, and implementation of individual modules as well as full applications
  - Tcl scripts for controlling EDA tools
  - Makefile scripts for tasks automation
- Available within OFM repository
  - implementation
  - documentation
- Key elements from user's perspective
  - definition of module's sources within <u>Modules.tcl</u> file
  - scripts for individual modules/applications (directories sim, uvm, synth, etc.)



# FUNCTIONAL VERIFICATION

### Simulation based on advanced testing techniques

- constrained pseudo-random generation of test vectors
- analysis of functional coverage
- invariants definition and checking
- scoreboarding
- OFM provides infrastructure for functional verification according to UVM (Universal Verification Methodology)
  - implementation
  - documentation



# DOCUMENTATION

- Documentation of NDK platform and OFM modules is compiled from sources using <u>Sphinx</u>
- VHDL sources are compiled using Sphinx-vhdl extension
  - implementation
  - documentation
- Two modes of documentation building
  - automatic (default)
  - manual



### AGENDA

- [presentation] NDK Introduction
- [presentation] ndk-app-minimal Introduction
- [hands-on] Playing with ndk-app-minimal
- [lunch break]
- [presentation]
   OFM Introduction
- [hands-on] Implementing Custom NDK Application
- [presentation] SW API Introduction
- [hands-on] Controlling NDK Application from SW
- [presentation] Advanced Topics



# [Q&A session]



# THANK YOU FOR YOUR ATTENTION