

РАЗРАБОТКА ДРАЙВЕРА ПОД ОС LINUX ДЛЯ УПРАВЛЕНИЯ АППАРАТНОЙ НЕЙРОННОЙ СЕТЬЮ НА ОТЛАДОЧНОЙ ПЛАТЕ

Т.А. Ягунов, И.В. Зоев

Научный руководитель: Н.Г. Марков
Томский политехнический университет

E-mail: tay1@tpu.ru

Введение

Для обмена данными между ОС и внешними по отношению к ней устройствами используются драйверы, которые предоставляют возможность пользователю взаимодействовать с ними. Также, драйверы обеспечивают связь с аппаратной спецификой поддерживаемого оборудования, например, механизмы обработки аппаратных прерываний. Перечисленные возможности являются необходимыми для рациональной реализации системы, использующей нейросетевые технологии, критичным параметром которой является скорость работы системы.

Описание проекта

Имеется устройство технического зрения на основе платы Terasic SoCKit [1], содержащей программируемую логическую интегральную схему (ПЛИС) семейства Cyclone V SoC от Altera и производящей вычисления на основе нейросетевых технологий. Функциональная схема этого устройства представлена на рис. 1.

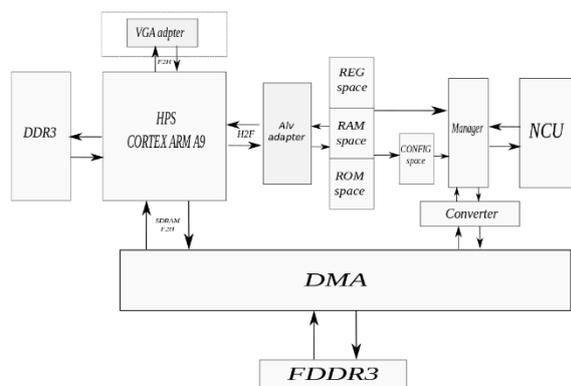


Рис. 1. Функциональная схема устройства из [1]

Представленный на рис. 1 блок процессорной системы работает под ОС Linux Ubuntu 16.04 версии, в которой и загружается драйвер, предоставляющий нейронной сети возможность взаимодействовать с различными чипами памяти, ПЛИС, процессором и т. д. Блок DDR3 памяти представляет чипы памяти, относящиеся к процессорной системе, блок FDDR3 (FPGA DDR3) представляет чипы памяти, относящиеся к ПЛИС. Остальные блоки схемы реализованы программным кодом средствами ПЛИС.

За управление взаимодействием между ПЛИС, процессором и памятью отвечает драйвер. Основные функции, которые выполняет разработанный драйвер в устройстве следующие:

- Выделение областей памяти необходимого объема;
- Чтение/Запись данных из/в памяти(-ть);
- Чтение/Запись данных из/в регистров(-ы) ПЛИС;
- Управление нейронной сетью;
- Управление отображением данных в памяти на экран VGA-монитора;
- Обработка прерываний.

Преимущества в использовании разработанного драйвера заключаются в получении гибкой и удобной в управлении системы, внедряемой в различные устройства.

Для удобства описания обозначим в разработанном драйвере 8 отдельных частей и назовем их следующим образом:

1. Main;
2. Dma;
3. Regs;
4. Vga;
5. Ctrl_fops;
6. Data_fops;
7. Data_vga_fops;
8. Ctrl_functions.

Main

Данная часть является основной «оболочкой» драйвера, которая отвечает за его регистрацию в ОС, регистрацию обработчика прерываний, создание char-devices (символьных устройств) [2] и предварительной настройкой взаимодействующих систем. Здесь объявляются и определяются основные структуры и глобальные переменные/константы.

Dma

В данной части запрограммирована функциональная часть DMA-устройств для взаимодействия с памятью DDR3. Сюда входит инициализация этих устройств, путем установки соответствующих значений в соответствующие регистры, выделение памяти для значений входов/выходов нейронной сети, весов связей нейронной сети и буфера данных для вывода на экран. Также в этой части реализована обработка прерываний для DMA-устройств.

Regs

Данная часть содержит функции, реализующие взаимодействие с регистрами.

Vga

Данная часть отвечает за инициализацию VGA-дисплея, путем установки соответствующих значений в соответствующие регистры, и выделение области памяти для данных, выводимых на экран

Ctrl_fops

Данная часть отвечает за реализацию функций символического устройства ctrl, которые будут вызываться при взаимодействии с соответствующим файлом устройства. Это устройство отвечает за управление разработанной системой (например, запуск нейронной сети или очистка буфера данных и т. д.). Ctrl_fops вместе с Ctrl_functions являются основной «начинкой» драйвера ввиду того, что в этих частях собран основной функционал по управлению разработанным устройством.

Data_fops

Данная часть отвечает за реализацию функций символического устройства data, которые будут вызываться при взаимодействии с соответствующим файлом устройства. Это устройство отвечает за предоставление (вывод) информации из области памяти, выделенной драйвером.

Data_vga_fops

Данная часть отвечает за реализацию функций символического устройства data_vga, которые будут вызываться при взаимодействии с соответствующим файлом устройства. Это устройство отвечает за предоставление (вывод) информации VGA-дисплею.

Ctrl_functions

Данная часть содержит в себе реализацию всех основных функций по управлению разработанным устройством, которые вызываются из части Ctrl_fops.

Помимо самого драйвера, в операционной системе регистрируются устройства (в данном случае символические), для которых данный драйвер и реализуется. Взаимодействие с устройством обеспечивается взаимодействием с соответствующим зарегистрированным файлом устройства. Например, устройству можно передать данные, путем записи этих данных в файл устройства, или, прочитав файл устройства, можно прочитать полученные данные от него.

Результаты

В качестве результата работы приведен скриншот вывода информации драйвером операционной системе при его инициализации о состоянии конфигурации устройства (рис. 2).

```
[ 52.295602] [etn_print_rom_reg]
[ 52.298752] #####
[ 52.302830] MAIN ARCH DEFINE PARAM
[ 52.306346] #####
[ 52.310433] [ENABLE_DDR3]: 0;
[ 52.313581] [ENABLE_CNN]: 1;
[ 52.316633] [DATA_W]: 16;
[ 52.319418] [CR_CNT]: 9;
[ 52.322114] [CONFIG_PARAM]: 35;
[ 52.325471] [VENDOR]: ALTERA
[ 52.328545] [VERSION]: 1.4.3d
[ 52.331674] #####
[ 52.335787] PARAMS NCU
[ 52.338229] #####
[ 52.342307] [COUNT_UNIT_CONV]: 10;
[ 52.345800] [COUNT_UNIT_POOL]: 10;
[ 52.349275] [LIMIT_CORE]: 16;
[ 52.352401] [LIMIT_LAYER]: 50;
[ 52.355631] [ADDR_W_RAM]: 11;
[ 52.358757] #####
[ 52.362840] CALCULATE UNIT
[ 52.365766] #####
[ 52.369851] [NEURAL_DATA_W]: 14;
[ 52.373150] [FLOAT_POINT]: 1;
[ 52.376342] [CUSTOM_FLOAT_POINT]: 1;
[ 52.379991] [DIF_NEURAL_DATA_W]: 2;
[ 52.383576] [NEURAL_DATA_EXP_W]: 5;
[ 52.387132] [BASIS]: 15;
[ 52.389826] [FULL_NEURAL_DATA_W]: 16;
[ 52.393578] [FIXED_POINT]: 0;
[ 52.396713] [BITS_INT_PART]: 0;
[ 52.399938] [BITS_REAL_PART]: 0;
[ 52.403280] #####
[ 52.407376] BITNESS NCU
[ 52.409899] #####
[ 52.414005] [INPUT_ADDR_W]: 7;
[ 52.417225] [WEIGHT_ADDR_W]: 10;
[ 52.420533] [SDRAM_DATA_W]: 128;
[ 52.423953] [SDRAM_ADDR_W]: 32;
[ 52.427262] [SDRAM_BE_W]: 16;
[ 52.430388] [FDDR3_DATA_W]: 128;
[ 52.433796] [FDDR3_ADDR_W]: 32;
[ 52.437100] [FDDR3_BE_W]: 16;
[ 52.440226] [H2F_DATA_W]: 64;
[ 52.443372] [H2F_ADDR_W]: 32;
[ 52.446500] [H2F_BE_W]: 8;
```

Рис. 2. Вывод информации драйвером ОС на плате Terasic SoCKit

Заключение

В результате работы был разработан драйвер под ОС Linux Ubuntu 16.04 для управления аппаратной нейронной сетью, с помощью которого в дальнейшем планируются проводиться работы и тестирование устройств на основе различных плат.

Благодарности

Исследования были поддержаны грантом РФФИ № 18-47-700010 p_a.

Список использованных источников

1. Зоев И.В. Аппаратная реализация на ПЛИС свёрточных нейронных сетей для распознавания объектов на изображениях И.В. Зоев, Н.Г. Марков, А.П. Береснев, Т.А. Ягунов; // GraphiCon: сборник трудов 28-ой Международной конференции по компьютерной графике и машинному зрению, г. Томск, 24-27 сентября 2018 г.: — Томск: Изд-во ТПУ, 2018. — [С. 188-191];
2. A. Rubini и J. Corbet, «Linux Device Drivers, 2nd Edition» (перевод: Князев Алексей) [Электронный ресурс] — URL: <http://knzsoft.ru/ldd2-ru> (дата обращения 05.08.2018);