System Design of Embedded Systems Running on an MPSoC Platform

Alexandru Andrei
Dept. of Computer and Information Science
Linköping University
SE-58 183 Linköping, Sweden

1 Introduction

This document is intended to be a guide through the design of embedded systems implemented on multiprocessor platforms. You will learn how to customize the hardware platform such that the resulting system is optimized for a given functionality. Most of the assignments will use as an example software application the GSM codec. You will work on a realistic multiprocessor system, composed of several ARM processor cores, private and shared memories, interconnected by an AMBA bus. This guide will present the hardware platform, the software development process as well as possible design alternatives.

A large fraction of the embedded systems are time constrained. We will introduce in this guide a set of tools used for the formal timing analysis of real-time applications.

2 MPARM Platform Description

In this section we will introduce the hardware and the software infrastructure.

2.1 Hardware Platform

The target architecture is a general template for a distributed multiprocessor system on a chip (MPSoC). The platform consists of computation cores, an AMBA AHB-compliant communication bus, private memories (one for each processor) and of a shared memory for interprocessor communication (see Fig. 1). The computation tiles are homogeneous and consist of ARM7 cores with instruction and data caches and of tightly coupled software-controlled scratch-pad memories.

The virtual platform environment provides power statistics for ARM cores, caches, on-chip memories and AMBA AHB bus, leveraging technology-homogeneous power models for a 0.13 \( \mu \text{m} \) technology provided by STMicroelectronics.

![Figure 1: MPARM Platform](image-url)
2.2 Software Platform

In this section we will describe with an example how to implement and run a small application in MPARM. Since MPARM is a cycle accurate simulator, any binary executable (actually a memory image obtained from the binary) for an ARM7 processor will run out of the box.

We will start with a very simple C application, show its particularities and then compile and execute it in the simulator.

```c
void main1() {
    int i;

    start_metric();
    pr("Starting Hello World", 0, PR_CPU_ID | PR_STRING | PR_NEWL );
    for(i=0;i<5;i++) {
        pr("", i, PR_CPU_ID | PR_DEC | PR_NEWL );
    }
    stop_metric();
    stop_simulation();
}
```

Let us have a look at this simple C program and notice a few particularities. The first particularity is that the main entry to a "classical" C program (_int main_), is replaced by a _void main1_. We also notice the function calls _start_metric(), stop_metric()_ and _stop_simulation()_. Please remember that the main benefit of having a simulator for a particular hardware platform is to collect various statistics about the hardware itself or about the software program running on it. _start_metric()_ must be called when we want to switch on the statistics collection and _stop_metric()_ at the end. _stop_simulation()_ must be called when we want to finish the simulation (you can see it as an equivalent to _exit()_).

Another useful function is _pr_. This is a very simple implementation of a _printf_ equivalent for the output of integers or strings on the console. The prototype is:

```c
void pr(char *text, unsigned long int x, unsigned long int mode)
```

The mode parameter specifies what to be printed, and can be the result of a bitwise OR of the following: _PR_CPU_ID_ (prints the id of the processor where the code runs), _PR_DEC_ (prints the integer variable _x_), _PR_STRING_ (prints the string _text_), _PR_NEWL_ (prints a newline). In order to use these functions, include _appsupport.h_ in your
programs. For the other lightweight functions available to the applications in the MPARM platform, please check the file:

/home/TDTS30/sw/mparm/MPARM/apps/support/simulator/appsupport.h

In order to be able to run this program, it must be first compiled. In order to make the compilation process easier, especially for large applications with many source files, we will use a makefile. The description of the make tool in general is outside the scope of this document. For further details, please refer to http://www.gnu.org/make. An example makefile, used for building our simple program, is presented in the following:

```plaintext
INCDIR = -I. -I${SWARMDIR}/core \
    -I${SWARMDIR}/../apps/support/simulator
OPT = -g -O3
CFLAGS = $(INCDIR) $(OPT)
all:
  arm-rtems-as -mfpu=softfp -o test.o \n    $(SWARMDIR)/../apps/support/simulator/test.s
  arm-rtems-gcc $(CFLAGS) -c -o appsupport.o \n    $(SWARMDIR)/../apps/support/simulator/appsupport.c
  arm-rtems-gcc $(CFLAGS) -c -o hello_world.o \n    hello_world.c
  arm-rtems-ld -T $(SWARMDIR)/../apps/support/simulator/test.ld \n    -o linked.o \n    test.o appsupport.o hello_world.o
  arm-rtems-objcopy -O binary linked.o TargetMem_1.mem
```

The actions executed by the makefile are captured by Fig. 2. In order to compile the C program, please type "make" in the directory holding your makefile (we assume that the code of this makefile is saved into the file with the name Makefile, located in the directory containing the application code). For setting up the correct paths for the MPARM simulator and crosscompiler, please run

```
source /home/TDTS30/sw/mparm/go_mparm.sh
```
from your terminal.

```plaintext
[sasa@lap-154 hello_world_mparm]$ ls
hello_world.c Makefile
[sasa@lap-154 hello_world_mparm]$ make
arm-rtems-as -mfpu=softfp -o test.o
/home/sasa/src/MPARM/swarm/.../apps/support/simulator/test.s
arm-rtems-gcc -I. -I/home/sasa/src/MPARM/swarm/core
    -I/home/sasa/src/MPARM/swarm/.../apps/support/simulator
    -g -03 -c -o appsupport.o
    /home/sasa/src/MPARM/swarm/.../apps/support/simulator/appsupport.c
arm-rtems-gcc -I. -I/home/sasa/src/MPARM/swarm/core
    -I/home/sasa/src/MPARM/swarm/.../apps/support/simulator
    -g -03 -c -o hello_world.o
    hello_world.c
arm-rtems-ld -T /home/sasa/src/MPARM/swarm/.../apps/support/simulator/test.ld -o linked.o
    test.o appsupport.o hello_world.o
arm-rtems-objcopy -O binary linked.o TargetMem_1.mem
```

Please note in the Makefile, that we do not use the standard GNU C compiler (gcc) available in any i386 Linux distribution. When developing applications for embedded systems, the code is written and debugged on desktop computers. The reasons are obvious: imagine writing an MPEG decoder on your mobile phone! The "standard" C compilers provided with Linux (Unix) distributions generate code for these desktop systems (Linux i386, Solaris
sparc, etc.). The embedded systems typically have different processors (ARM, PowerPC) and operating systems (RTEMS, RT Linux, or no operating system). Thus any code compiled with a "standard" C compiler will not run on the target embedded system platform. In order to compile the application code for an ARM7 processor, a special compiler is used: the crosscompiler. A crosscompiler is a compiler that is running on a system and generating binaries for another system. In the case of this lab, the crosscompiler runs on an i386 platform under Linux and generates code for ARM7 processors.

As a result, you should have now the memory image resulted from our C code (in the file TargetMem_1.mem). In order to run it in MPARM, please use the mpsim.x command:

```
[sasa@lap-154 hello_world_mparm]$ mpsim.x -c1
```

The option `-c1` indicates how many processors will be used. The maximum is 8 ARM7 processors. In this case, we have selected to run the program on one single processor. As a result, the file "stats.txt" will be created in the current directory and you will see the following output on the console:

```
SystemC 2.0.1 --- Jan 12 2005 11:34:16
Copyright (c) 1996-2002 by all Contributors
ALL RIGHTS RESERVED
Uploaded Program Binary: TargetMem_1.mem
Processor 0 starts measuring
Processor 0 - Hello World
Processor 0 - 0
Processor 0 - 1
Processor 0 - 2
Processor 0 - 3
Processor 0 - 4
Processor 0 stops measuring
Processor 0 shuts down
Simulation ended
SystemC: simulation stopped by user.
```

Let us try to start the same program, but with "-c2" parameter. First, we need to copy TargetMem_1.mem to TargetMem_2.mem.

```
[sasa@lap-154 hello_world_mparm]$ cp TargetMem_1.mem TargetMem_2.mem
[sasa@lap-154 hello_world_mparm]$ mpsim.x -c2
```

```
SystemC 2.0.1 --- Jan 12 2005 11:34:16
Copyright (c) 1996-2002 by all Contributors
ALL RIGHTS RESERVED
Uploaded Program Binary: TargetMem_1.mem
Uploaded Program Binary: TargetMem_2.mem
Processor 0 starts measuring
Processor 1 starts measuring
Processor 0 - 0
Processor 1 - 0
Processor 0 - 1
Processor 1 - 1
Processor 0 - 2
Processor 1 - 2
Processor 0 - 3
Processor 1 - 3
Processor 0 - 4
Processor 1 - 4
```

4
Processor 0 stops measuring
Processor 1 stops measuring
Processor 0 shuts down
Processor 1 shuts down
Simulation ended
SystemC: simulation stopped by user.

The output looks as expected. A copy of the program is running on each processor. So, one way of specifying the
code to be executed on each processor as above: each different program is written independently in C, compiled. The
memory image generated with objcopy is copied in the file TargetMem_i.mem, where i is the number of the processor.
The simulator is then started with the option -cn (n is the maximum number of processors used; so consequently there
have to be memory images for the n processors: TargetMem_1.mem, TargetMem_2.mem, ..., TargetMem_n.mem).
This coding style makes the process of porting existing applications to MPARM more difficult. Another disadvantage
is the lack of flexibility in exploring a new mapping.

Another alternative for specifying the mapping of the code to a particular processor, is to use the function
get_proc_id(). When using this method, we can take the code written as if it is targeted to run on a single
processor, and just mark in the source which block is executed on which processor. For example, in the following
small application, the first for loop (i=0 to 4) is executed on CPU1 and the second for loop (i=6 to 10) is executed on
CPU2. The disadvantage of this method is that we generate one big binary that contains the code to be executed on
all the processors.

```c
void main1() {
    int i;
    start_metric();

    if (get_proc_id()==1) { // executed on CPU1
        for(i=0;i<5;i++) {
            pr("", i, PR_CPU_ID | PR_DEC | PR_NEWL );
        }
    }

    if (get_proc_id()==2) { //executed on CPU2
        for(i=6;i<10;i++) {
            pr("", i, PR_CPU_ID | PR_DEC | PR_NEWL );
        }
    }

    stop_metric();
    stop_simulation();
}
```

2.3 Collecting Simulation Statistics

2.3.1 stats.txt

As explained in the previous section, after the simulation of the application has finished, MPARM dumps various
useful statistics in the file stats.txt. These statistics consist of data collected from the processors, caches, buses
and memories. By default, only performance related statistics are collected. If the simulator is started with the
-w command line argument (mpsim.x -c1 -w), the power/energy consumed by the system components is also
dumped. For the details regarding the precise data collected by the simulator, please read the file:
/home/TDTS30/sw/mparm/MPARM/doc/simulator_statistics.txt
During the design of an embedded system it is often interesting to analyze various statistics concerning individual tasks (snapshots of the system on various points during the execution of an application). It is possible to do that in MPARM, by calling the function `dump_light_metric(int x)` at different points of the application. The output is produced in the file `stats_light.txt` and consists of the number of clock cycles executed since the start of the simulation until `dump_light_metric(int x)` was called. Statistics about the interconnect activities are also collected. This is an example of output resulted from calling `dump_light_metric(int x)`:

```
-----------------------
Task 1
Interconnect statistics
-----------------------
Overall exec time = 287 system cycles (1435 ns)
Task NC = 287
1-CPU average exec time = 0 system cycles (0 ns)
Concurrent exec time = 287 system cycles (1435 ns)
Bus busy = 144 system cycles (50.17% of 287)
Bus transferring data = 64 system cycles (22.30% of 287, 44.44% of 144)
-----------------------
Task 2
Interconnect statistics
-----------------------
Overall exec time = 5554 system cycles (27770 ns)
Task NC = 5267
1-CPU average exec time = 0 system cycles (0 ns)
Concurrent exec time = 5554 system cycles (27770 ns)
Bus busy = 813 system cycles (14.64% of 5554)
Bus transferring data = 323 system cycles (5.82% of 5554, 39.73% of 813)
```

The per task statistics are the overall execution time (the number of clock cycles executed by the system since the application start), Task NC (the number of clock cycles executed by the current task), bus busy (the number of clock cycles the bus was busy), bus transferring data (the number of clock cycles when the bus was actually transferring data).

## Assignment 0: Getting Started

1) The MPARM cycle accurate simulation platform is available only under Linux. In order to be able to use it, you have to connect remotely to a Linux server. There are 7+1 recommended Linux servers: sy00-2.ida.liu.se, sy00-3.ida.liu.se, ..., sy00-7.ida.liu.se (130.236.186.40-130.236.186.46) and crabbofix.ida.liu.se. Use ssh to connect to one of these computers (for example `ssh sy00-7.ida.liu.se`).

2) Please copy the hello world code to your home directory from:

```
/home/TDTS30/sw/mparm/benches/hello_world
```

3) Run `source /home/TDTS30/sw/mparm/go_mparm.sh` (each time you ssh, in order to set up the correct paths for the simulator and compilers).

4) Compile the code with: `make -f Makefile`.

5) Run it with various simulator flags and study the statistics collected by the simulator. For a complete list of the command line parameters, type `mpsim.x -h` or study

```
/home/TDTS30/sw/mparm/MPARM/doc/simulator_switches.txt
```

## Simulation-Based Design Space Exploration for Energy Minimization
4.1 Introduction

One of the advantages of having a simulation platform is that it can be integrated early in the design flow. Traditionally, as we can see from Fig. 3, the final validation is performed at the end of design flow, after the first prototype is available. In order to obtain a correct and efficient product, the system model must be very accurate. If a cycle accurate simulator of the target hardware platform is available in the early design stages, the system model may be complemented or even replaced by this simulator, as in Fig. 3. In this way, we would gain accuracy at the expense of using a slow simulator as opposed to fast models.

We will illustrate with this assignment the usage of the MPARM simulation platform for a design space exploration problem. The goal is to optimize a GSM encoder/decoder application, given as a source code written in C. This application will run on an instance of the MPARM platform with one ARM7 processor. The size of the cache, as well as the cache associativity and the frequency of the processors are variable. Find an assignment of the cache size, associativity and frequency for the processor, such that the energy consumed by the system is optimized.

The frequency of the processor is a key parameter, affecting the speed of the application and the power consumption. Typically, in the power consumption is proportional with the frequency and with the square of supply voltage: $P = f \cdot C_{eff} \cdot V_{dd}^2$. In most systems, choosing a lower frequency translates in using a lower supply voltage as well. So the power consumed by the processor is reduced cubically. Actually, for battery powered systems, we are interested in the energy consumption. The energy is defined as the product of the power over time:

$$E = P \cdot t = f \cdot C_{eff} \cdot V_{dd}^2 \cdot \frac{NC}{f} = C_{eff} \cdot V_{dd}^2 \cdot NC$$

where $t$ is the execution time in seconds and $NC$ is the execution time expressed in number of clock cycles. So, if we would only scale the frequency, without scaling the supply voltage, although the power is reduced linearly, the energy would remain constant.

It is important to note that by scaling down the frequency (and thus the supply voltage), only the energy consumed by the processor is reduced. The power consumption of the rest of the components other then the CPU cores (such as caches, memories) remains constant. We could express this mathematically as follows:

$$E_{system} = \sum_{CPU_i} P_{CPU_i}(f) \cdot t(f) + \sum_{nonCPU_i} P_{nonCPU_i} \cdot t(f)$$
The energy consumed by a processor is a monotonically decreasing function of the frequency (to a lower frequency it corresponds a lower energy). However, the energy consumed by the other components is an increasing function of the frequency (the power does not depend on the processor frequency, while the execution time increases if we decrease the frequency). To conclude, the total energy of the system is achieved at an optimal frequency, depending on the application. Scaling the frequency below that value will result in an increase in the energy due to components other than the processor cores (caches, memories, bus). We have studied the frequency scaling for an MP3 decoder running on one processor in MPARM. The results are plotted in Fig. 4(a). On the X axis, we have varied the frequency from 200 MHz to 50 MHz in four steps: 200, 100, 66, 50 MHz. The values for the total energy obtained for each frequency are represented on the Y axis. Clearly, running the MP3 decoder at 100 MHz provides the best energy. Scaling down the frequency below this value results in an increase in the total energy. Fig. 4(b) presents the energy of the processor core, when the frequency is scaled. For the processor core, the lowest frequency provides the lowest energy.

Instruction and data caches are used to improve the performance of applications by achieving a faster execution time. From the power perspective, an issue particularly important for embedded systems, the cache hardware cannot be neglected. The power consumption of the cache and its efficiency depend on the associativity (direct mapped, k-associative, fully associative) and size. Intuitively, a large cache is more efficient. However, it will consume more power. An interesting aspect is the interaction in terms of energy between the cache and the CPU. A larger cache, while consuming more power, could reduce the execution time of the application and thus the energy consumed by the processor. Similar to frequency scaling, we have performed experiments on the MP3 decoder observing the influence of the size of the instruction cache on the total energy of the system and speed of the application. The results are presented in Fig. 4(c) and (d). We have considered a split instruction and data cache. Clearly, a big instruction cache translates in a shorter execution time. This can be observed in Fig. 4(d). The more interesting result is the one on energy, presented in Fig. 4(c), showing that very small instruction caches are bad for energy as well. Big caches, provide a good execution time at the expense of a higher energy. We notice that the energy curve has an optimum when the instruction cache size is 4096 bytes.

4.2 Assignment 1: Simulation Based Design Space Exploration for Energy Minimization

In this assignment you will use the MPARM platform in order to optimize a multimedia application. You are given the source code of GSM voice codec.

1) Please copy it to your home directory from:
/home/TDTS30/sw/mparm/benches/gsm-good/single-newmparm/gsm-1.0-pl10-one-task-mparm (use the shell command cp -r to copy recursively the entire directory).

2) Compile the code. Assuming you are in the directory
$HOME/TDTS30/gsm-1.0-pl10-one-task-mparm,
please use the command make -f Makefile.mparm.
You should have now the memory image file
$HOME/TDTS30/gsm-1.0-pl10-one-task-mparm/bin/TargetMem_1.mem. The GSM codec is running on one processor. You are allowed to set the cache type (split instruction and data cache, or unified cache), associativity (direct mapped, k-way set associative) and size as well as the frequency of the processor. One encoding of a GSM frame has to finish in 20ms. Find a configuration for the processor such that the energy is minimized. Report the
best configuration obtained, as well as the ones that you have tried (at least 6). Explain your results. Which parameter has the greatest impact on energy consumption, according to your experiments?

Important: There are 2 known bugs in the simulator.
1) You will not be able to simulate the system with a unified cache (corresponding to the -u command line flag).
2) There is no power model associated to a fully associative cache (corresponding to the –it 0 or –dt 0 options).

5 Assignment 2: Communication in Multiprocessor Systems

An important aspect during the design of multiprocessor systems is the exchange of data between tasks running on different processors. We will present two communication alternatives in this section. Another important issue addressed in this section and which is tightly coupled with the communication, is the synchronization.

5.1 Shared Memory Based Communication

Passing the data through the shared memory is the most common way of interprocessor communication. An example is illustrated in Fig. 5. Let us assume that the task running on processor CPU2 (consumer) needs some data that was produced by the task running on CPU1 (producer). The easiest way to communicate this data is to store it on a common resource (the shared memory). Let us look closer at the following code example:

```c
void main1() {
    int i,j;
    if (get_proc_id()==1) { //the producer: task T1 on processor CPU1
        int *x;
        x=(int *)SHARED_BASE; //the address in the shared memory
        //where to store the variable
        //produce the message
        for(j=0;j<5;j++) {
            x[j]=5+j;
        }
    }
    else { //the consumer: task T2 on processor CPU2
        int *x;
        x = (int *)SHARED_BASE; //the address in the shared memory
        //where to read the variable
        for(j=0;j<5;j++) pr("", x[j], PR_CPU_ID | PR_DEC | PR_NEWL );
    }
}
```

Task T1 writes in the shared memory starting from the address \texttt{SHARED\_BASE} the values for 5 integer variables. T2 reads these variables and displays their value on the console. Two interesting question are raised by this implementation:
1) Task T1 writes at a certain address in the shared memory. How does T2 know from what address should read the data? This question is left as an exercise (see assignment 2.1 in section 5.4).

2) What happens if the task T2 starts reading from the shared memory before the task T1 has finished to write? We will answer this question in the next section.

(The \texttt{SHARED\_BASE} definition can be used by MPARM programs to denote the start address of the shared memory, if \texttt{appsupport.h} is included.)

5.2 Synchronization

In this section we will describe the synchronization mechanism from MPARM. We start by coming back to question 2 from the end of the previous section. What happens if the task T2 starts reading from the shared memory before the task T1 has finished to write? In the previous example, the value of variable x from task T2 is undefined. Task T2 reads values from a certain memory address. When running that example several times, the values displayed in task T2 could be different. The correct functionality would be achieved if task T2 would wait until task T1 finishes updating the shared memory with the correct x values. The MPARM platform provides a set of hardware semaphores (that can be accessed via the software using the \texttt{*lock} variable), as well a simple software API with the functions \texttt{WAIT(int lock_id)} and \texttt{SIGNAL(int lock_id)}. The semantic of \texttt{WAIT} and \texttt{SIGNAL}, operating both on the same hardware lock identified by the parameter \texttt{lock_id} is the following: the task that calls \texttt{WAIT(lock_id)} will stop its execution and wait until another task running on a different processor calls \texttt{SIGNAL(lock_id)}. A simple shared memory based communication example is presented in the following example code:

```c
extern volatile int *lock;

lock[0]=1;//lock[0] is taken
if (get_proc_id()==1) {
    int *x;
    x=(int *)SHARED_BASE;
    //produce the message
    for(j=0;j<5;j++) {
        x[j]=5+j;
    }
    //signal data ready to the consumer
    SIGNAL(0);
}

if (get_proc_id()==2) {
    int *x;
    x = (int *)SHARED_BASE;
    WAIT(0);
    for(j=0;j<SIZE;j++) pr("", x[j], PR_CPU_ID | PR_DEC | PR_NEWL );
}
```

5.3 Allocating Variables on the Shared Memory

MPARM provides a replacement of the \texttt{malloc} function that handles the shared memory allocation. The function is \texttt{(void *)shared\_alloc(int size)}. The function is synchronized internally, so when several allocations are running concurrently, each task receives another memory block with the requested size. On the other hand, an implementation of a function for freeing the unused memory, is not very easy to implement. At this point it is interesting to note that \texttt{(void *)shared\_alloc(int size)} as well as the functions used for displaying messages and profiling (\texttt{pr, dump\_light\_metric}) do not require an operating system.
5.4 Assignment 2.1: Implementation of a Generic Shared Memory Based Communication Infrastructure

In this assignment you are required to implement a shared memory based communication API, that can be used for passing data from a task running on one processor (producer) to a task running on another processor (consumer). The consumer does not know the address in the shared memory used by the producer to store the data. Write a small test application with two tasks running on two processors to test your API.

5.5 Distributed Message Passing

We have introduced in section 5.1 a widely used interprocessor communication method, namely the usage of shared memory. In order to safely access the shared memory we have shown that semaphores must be used. We will present in this section another approach for interprocessor communication. Your assignment is to compare the two communication alternatives using for this purpose a GSM voice codec application running in MPARM and implemented using these two communication alternatives. The comparison will include the bus utilization, the overall application runtime and the system energy consumption.

Let us look closer at the shared memory based communication. The main drawback here is the bus traffic generated due to the synchronization. MPARM has a hardware semaphore device, implemented as a small memory that can be accessed by the processors via the bus. Such an architecture is depicted in Fig. 6(a). The task \( t_1 \) running on CPU1 produces some data and writes it in a shared memory buffer that is later read by the task running on CPU2. We assume that the buffer size is 1 and the two tasks repeatedly produce and respectively consume this data. Before \( t_1 \) writes a new value to the buffer, it has to make sure that \( t_2 \) has read the previous value. A semaphore \( buffer\_full \) is used for this purpose. On the other hand, before reading, \( t_2 \) has to wait until new data is available in the buffer, and uses the semaphore \( buffer\_empty \) to do that. Waiting on a semaphore is performed by repeatedly checking the semaphore (polling) until it becomes free, potentially causing a lot of traffic on the bus. An obvious reason why this is bad is power consumption: this traffic on the bus consumes energy from the battery. Another system parameter that is indirectly affected by this polling is the execution time of the application. The semaphore polling generates traffic on the bus, interfering for example with the traffic due the cache misses (in case of a cache miss, the information is retrieved in the cache from the memory). In this way the miss penalty increases and the processor is stalled waiting for the cache refills to finish.

Shared memory based communication is not the only available design choice. MPARM provides a direct processor to processor communication infrastructure with distributed semaphores. This alternative is depicted in Fig. 6(b). Instead of allocating the semaphores on a shared hardware, the two semaphores are each allocated on a scratchpad memory directly connected to each processor (ie. in order for a processor to access data on its scratchpad memory it does not need to go via the bus; the scratchpad is small and fast, compared to a RAM). In our example, \( buffer\_full \) is allocated on Scratchpad_CPU1, while \( buffer\_empty \) is allocated on Scratchpad_CPU2. Consequently when
tasks $t_1$ and $t_2$ are doing busy waiting, they are not generating any bus traffic. The scratchpad memories have to be connected to the bus, in order to be accessible to the other processors. In this way, when for example, $t_2$ running on CPU2 releases the semaphore $buffer\_full$ it can access the corresponding location on the Scratchpad_CPU1.

The following code is a simple example of exchanging data between two tasks running on different processors, using this distributed message passing infrastructure.

```c
scratch_queue_autoinit_system(0,0);

if (get_proc_id()==1) {
    int *buffer;
    SCRATCH_QUEUE_PRODUCER* prod=scratch_queue_autoinit_producer(2,1,MAX_QUEUE_SIZE,OBJ_SIZE,0);
    //ask for space in the queue buffer
    buffer=(int*)scratch_queue_getToken_write(prod);
    for(j=0;j<SIZE;j++) {
        buffer[j]=5+j;
    }
    //signal data ready to the consumer
    scratch_queue_putToken_write(prod);
}

if (get_proc_id()==2) {
    int *buffer_copy;
    int i;
    SCRATCH_QUEUE_CONSUMER* cons=scratch_queue_autoinit_consumer(1,0);
    buffer_copy=(int*)scratch_queue_read1(cons);
    for(j=0;j<SIZE;j++) pr(NULL, buffer_copy[j], PR_DEC | PR_NEWL);
}
```

### 5.6 Assignment 2.2: Shared Memory vs. Distributed Message Passing

In this assignment you will compare the efficiency of two communication approaches: shared memory and distributed message passing. The comparison is based on simulation results of the GSM voice codec, implemented using the two alternatives.

1) Please copy the GSM code to your home directory from:
   `/home/TDTS30/sw/mparm/benches/gsm-good/mproc/queues/gsm-mparm-multiproc-map1`
   `/home/TDTS30/sw/mparm/benches/gsm-good/mproc/shared/gsm-mparm-multiproc-map1`

   Check the code related to the communication, located in the file `src/toast.c` in the function `process_codec()`.

2) Compile the code. Please use for both versions (shared memory and scratchpad queues) the corresponding `makefile` `Makefile.mparm`.

3) Run the two versions. In both cases, the application is mapped on 3 processors. Use the `-c3` flag when you run the simulator. The distributed message passing version needs to allocate data on the scratchpad memories. By default these are disabled in MPARM. When running the GSM with distributed message passing (from the `queues/gsm-mparm-multiproc-map1` directory), use the `-C` option of the simulator.

   Report the simulation results that you think are relevant to this comparison.

4) Try to reduce the amount of traffic due to the synchronization in the shared memory implementation by frequency selection. The frequency for a processor can be selected statically (with the `-F` command line option), or changed dynamically, while the application is running. At runtime, while the application is executing, the function `scale_device_frequency(unsigned short int divider, int ID)` can be called to change the frequency of a processor (the `-f -intc=s` command line options are needed when using this function). Other functions related to frequency selection are in
   `/home/TDTS30/sw/mparm/MPARM/apps/support/simulator/appsupport.h.`
6 Mapping and Scheduling in Multiprocessor Systems

The purpose of this assignment is to illustrate two key design issues related to the implementation of software applications on multiprocessor systems: the mapping and scheduling of the software tasks. This assignment is based on the GSM voice codec (encoder and decoder), implemented on the MPARM platform.

6.1 Introduction to Mapping and Scheduling

During the design of embedded systems composed of software applications and a multiprocessor hardware platform, an important decision is how to assign the various software components (tasks) to the processors (mapping) and in which order to execute the software tasks running on each of the processors (scheduling). This decision can be made at the design time, because embedded systems have a dedicated, (known) functionality, as opposed to general purpose computers that work with a variety of (unknown) applications. We will introduce in the following a simple embedded design flow (see Fig. 7(a)), emphasizing certain aspects related to the mapping and scheduling of the application on the target hardware platform.

The input of our design flow is the software application, specified in the C programming language, and the hardware platform. The specification of the application consists of code that can be compiled and executed on an a single processor. For example, in case of the GSM codec, the code is composed of two distinct parts, the encoder and the decoder, that can be independently compiled and executed on a single processor instance of MPARM. The first step of the design flow is to extract from the application task graph.

The functionality of applications can be captured by task graphs, $G(T, C)$. An example task graph is depicted in Fig. 7(b). Nodes $\tau \in T$ in these directed acyclic graphs represent computational tasks, while edges $\gamma \in C$ indicate data dependencies between these tasks (i.e. communications).

Such a task graph is extracted (manually, partially or fully automatically) from the source code of the application.
The task graph must capture two important aspects: dependencies (tasks that can be executed in a certain order) and parallelism (tasks that can be executed in parallel, potentially on different processors). Partitioning the application code into tasks is not an exact science. For example, the granularity of the tasks can be finer or coarse. A fine granularity can offer flexibility for various optimizations. There are, however, drawbacks in having a fine granularity: the overhead due to context switches, the amount of communication or the complexity of analyzing and performing optimizations can be very high. For example, we present in Fig. 8 the task graph of the GSM voice codec. Another possible task graphs for the decoder and encoder, with a finer granularity are given in Fig. 9 and 10. Please note, that in case of the encoder, there are 16 tasks in 8 and 53 tasks in 10.

Given a task graph (Fig. 7(c)) and a target hardware platform (Fig. 7(b)), the designer has to map and schedule the tasks on the processors. Mapping is the step in which the tasks are assigned for execution to the processors and the communications to the bus(es). We have depicted a possible mapping for the task graph from Fig. 7(c) in Fig. 7(d). The next step is to compute a schedule for the system, i.e., to decide in which order to run the tasks mapped on the same processor. One important set of constraints that have to be respected during mapping and scheduling are the precedence constraints given by the dependencies in the task graph. An example schedule is depicted in Fig. 7(e). Please note that task $\tau_2$ starts only after task $\tau_1$ and the communication $\gamma_{1-2}$ have finished. Most embedded applications must also respect the real-time constraints, such as the application deadline.

Computing the task mapping and schedule is in general a complex problem and exact solutions usually belong to the NP class. Nevertheless there exist tools that perform mapping and scheduling with different objectives. For example, given a task graph with dependencies, each task being characterized by an execution time and a deadline, there exist algorithms for computing a mapping and a schedule such that the precedence constraints and deadlines are respected. If the objective is the energy minimization, besides the precedence and deadline constraints, there exist algorithms that compute the mapping, schedule and task speed that provide the minimum energy. The details of these mapping and scheduling algorithms are beyond the scope of this lab. The purpose of the next assignment is to study, using the MPARM cycle accurate simulator, the impact of the mapping and schedule on the timing of the application and on the energy consumed by the system.

### 6.2 Mapping and Scheduling in MPARM

In this section we will discuss the way to specify the mapping and the schedule for applications running on MPARM without operating system support. The specification of the mapping was discussed previously, in section 2.2. To summarize, you can use the function `get_proc_id()` for this purpose. This function returns the ID of the processor...
executing the code. Let us consider the following example:

```c
a();
if (get_proc_id()==1) {
    b();
    c();
}
if (get_proc_id()==2) {
    d();
    e();
}
```

In this example, the function `a()` will be executed by all the processors. Functions `b()` and `c()` will execute only on processor 1, while `d()` and `e()` only on processor 2. So, in other words, `b()` and `c()` are mapped on processor 1 and `d()` and `e()` are mapped on processor 2.

The schedule is given by the order of execution. For example `b();c();` is the schedule from processor 1 in the previous example. If we change the order of the function calls, we obtain another schedule.

### 6.3 Assignment 3.1: Extracting the Task Execution Times for the GSM Voice Codec with MPARM

1) In the first step, match the task graph of the GSM encoder and decoder from Fig. 8 with the corresponding code from functions `process_encode` and `process_decode` in the file `/home/TDTS30/sw/mparm/benches/gsm-good/single-newmparm/gsm-1.0-pl10-one-task-mparm/src/toast.c`. The single processor version of the GSM executes either the `process_encode` function or `process_decode`. In order to choose between one or the other, you have to set the variable `f_decode` (located in the file `toast.c`) accordingly: `f_decode=0` selects `process_encode`, while `f_decode=1` selects `process_decode`. 2) Extract and report the task execution times, using the `dump_light_metric()` function.

### 6.4 Assignment 3.2: Mapping and Scheduling of the GSM Voice Codec on MPARM

1) Identify and report the mapping and schedule used for the GSM codec implemented with distributed message passing:
Figure 10: GSM voice encoder task graph
Both the encoder and decoder tasks are located in the function `process_codec()` from the file `toast.c`.

2) Find another (better, with a lower energy) schedule, given this mapping (the application is mapped on 3 processors).

3) Find another mapping (with a lower energy)

7 Summary of the assignments

1) Simulation-based design space exploration for energy minimization.
2) Communication in MPSoCs.
   2.1) Implementation of a shared memory communication API.
   2.2) Comparison between shared memory and distributed message passing.
3) Mapping & Scheduling.
   3.1) Extract execution times for the GSM codec.
   3.2) Find a better schedule given a mapping and find a better mapping.