The BondMachine Toolkit A comprehensive approach to computing with FPGA

Mirko Mariotti

Department of Physics and Geology - University of Perugia INFN Perugia

Advanced Workshop on Modern FPGA Based Technology for Scientific Computing ICTP - Trieste 13-24 May 2019

Advanced Workshop on Modern FPGA-Based Technology for Scientific Computing





The BondMachine: a comprehensive approach to computing with FPGA.

In this presentation i will talk about:

- Technological background of the project
- The BondMachine Project: the Architecture
- The BondMachine Project: the Tools
- Use cases



Some topic will have an hands-on session.

For the hands-on sessions, some minimal Linux shell is required:

- directories
- less: used to show text file content
- anything you want: to edit a text file

Set the environment up with the command:

source bm.sh

3/116

Technological Background

FPGA workshop - ICTP - Trieste 13-24 May4/116

A field-programmable gate array (FPGA) is an integrated circuit whose logic is re-programmable. It's used to build reconfigurable digital circuits.

FPGAs contain an array of programmable logic blocks, and a hierarchy of reconfigurable interconnects that allow the blocks to be "wired together".

Logic blocks can be configured to perform complex combinational functions.





The FPGA configuration is generally specified using a hardware description language (HDL).



FPGA workshop - ICTP - Trieste 13-24 May 5/116

FPGA What is it ?

- A field-programmable gate array (FPGA) is an integrated circuit whose logic is re-programmable. It's used to build reconfigurable digital circuits.
 - FPGAs contain an array of programmable logic blocks, and a hierarchy of reconfigurable interconnects that allow the blocks to be "wired together".

Logic blocks can be configured to perform complex combinational functions.





The FPGA configuration is generally specified using a hardware description language (HDL).



FPGA workshop - ICTP - Trieste 13-24 May 5/116

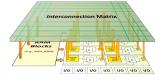
• A field-programmable gate array (FPGA) is an integrated circuit whose logic is re-programmable. It's used to build reconfigurable digital circuits.

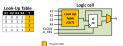
FPGAs contain an array of programmable logic blocks, and a hierarchy of reconfigurable interconnects that allow the blocks to be "wired together".

Logic blocks can be configured to perform complex combinational functions.

The FPGA configuration is generally specified using a hardware description language (HDL).

5/116





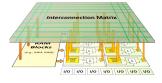


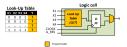
• A field-programmable gate array (FPGA) is an integrated circuit whose logic is re-programmable. It's used to build reconfigurable digital circuits.

FPGAs contain an array of programmable logic blocks, and a hierarchy of reconfigurable interconnects that allow the blocks to be "wired together".

Logic blocks can be configured to perform complex combinational functions.

The FPGA configuration is generally specified using a hardware description language (HDL).







FPGA workshop - ICTP - Trieste 13-24 May 5/116

The use of FPGA in computing is growing due several reasons:

can potentially deliver great performance via massive parallelism

6/116

can address payloads which are not performing well on uniprocessors (Neural Networks, Deep Learning)

can handle efficiently non-standard data types



The use of FPGA in computing is growing due several reasons:

can potentially deliver great performance via massive parallelism

6/116

 can address payloads which are not performing well on uniprocessors (Neural Networks, Deep Learning)

can handle efficiently non-standard data types



The use of FPGA in computing is growing due several reasons:

can potentially deliver great performance via massive parallelism

6/116

- can address payloads which are not performing well on uniprocessors (Neural Networks, Deep Learning)
- can handle efficiently non-standard data types



On the other hand the adoption on FPGA poses several challenges:

Porting of legacy code is usually hard.

Interoperability with standard applications is problematic.



On the other hand the adoption on FPGA poses several challenges:

Porting of legacy code is usually hard.

Interoperability with standard applications is problematic.



Computer Architectures

FPGA workshop - ICTP - Trieste 13-24 May 8/116

Today's computer architecture are:

• Multi-core, Two or more independent actual processing units execute multiple instructions at the same time.

- The power is given by the number of cores.
- Parallelism has to be addressed.

Heterogeneous, different types of processing units.

- Cell, GPU, Parallela, TPU.
- The power is given by the specialization.
- The units data transfer has to be addressed.
- The scheduling has to be addressed.



FPGA workshop - ICTP - Trieste 13-24 May 9/116

Today's computer architecture are:

Multi-core, Two or more independent actual processing units execute multiple instructions at the same time.

• The power is given by the number of cores.

Parallelism has to be addressed.

Heterogeneous, different types of processing units.

Cell, GPU, Parallela, TPU.

The power is given by the specialization.

The units data transfer has to be addressed.

The scheduling has to be addressed.



Today's computer architecture are:

9/116

• Multi-core, Two or more independent actual processing units execute multiple instructions at the same time.

- The power is given by the **number** of cores.
- Parallelism has to be addressed.

Heterogeneous, different types of processing units.

- Cell, GPU, Parallela, TPU.
- The power is given by the specialization.
- The units data transfer has to be addressed.
- The scheduling has to be addressed.



Today's computer architecture are:

• Multi-core, Two or more independent actual processing units execute multiple instructions at the same time.

- The power is given by the number of cores.
- Parallelism has to be addressed.

Heterogeneous, different types of processing units.

- Cell, GPU, Parallela, TPU.
- The power is given by the specialization.
- The units data transfer has to be addressed.
- The scheduling has to be addressed.



FPGA workshop - ICTP - Trieste 13-24 May 9/116

Today's computer architecture are:

• Multi-core, Two or more independent actual processing units execute multiple instructions at the same time.

- The power is given by the number of cores.
- Parallelism has to be addressed.

Heterogeneous, different types of processing units.

- Cell, GPU, Parallela, TPU.
- The power is given by the specialization.
- The units data transfer has to be addressed.
- The scheduling has to be addressed.



FPGA workshop - ICTP - Trieste 13-24 May 9/116

Today's computer architecture are:

Multi-core, Two or more independent actual processing units execute multiple instructions at the same time.

- The power is given by the **number** of cores.
- Parallelism has to be addressed.

Heterogeneous, different types of processing units.

- Cell, GPU, Parallela, TPU.
- The power is given by the specialization.
- The units data transfer has to be addressed.
- The scheduling has to be addressed.



FPGA workshop - ICTP - Trieste 13-24 May 9/116

Today's computer architecture are:

- **Multi-core**, Two or more independent actual processing units execute multiple instructions at the same time.
 - The power is given by the **number** of cores.
 - Parallelism has to be addressed.
- Heterogeneous, different types of processing units.
 - Cell, GPU, Parallela, TPU.
 - The power is given by the specialization.
 - The units data transfer has to be addressed.
 - The scheduling has to be addressed.



FPGA workshop - ICTP - Trieste 13-24 May 9/116

Today's computer architecture are:

- **Multi-core**, Two or more independent actual processing units execute multiple instructions at the same time.
 - The power is given by the **number** of cores.
 - Parallelism has to be addressed.
- Heterogeneous, different types of processing units.
 - Cell, GPU, Parallela, TPU.
 - The power is given by the specialization.
 - The units data transfer has to be addressed.
 - The scheduling has to be addressed.



FPGA workshop - ICTP - Trieste 13-24 May 9/116

Today's computer architecture are:

- **Multi-core**, Two or more independent actual processing units execute multiple instructions at the same time.
 - The power is given by the **number** of cores.
 - Parallelism has to be addressed.
- Heterogeneous, different types of processing units.
 - Cell, GPU, Parallela, TPU.
 - The power is given by the specialization.
 - The units data transfer has to be addressed.
 - The scheduling has to be addressed.



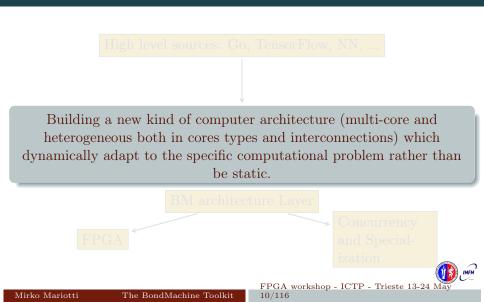
FPGA workshop - ICTP - Trieste 13-24 May 9/116

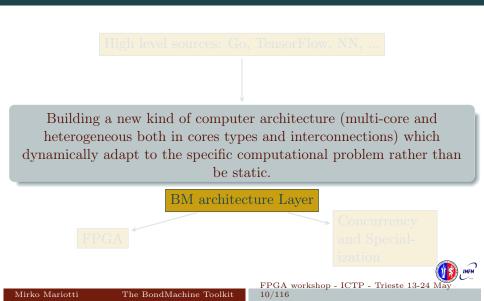
Today's computer architecture are:

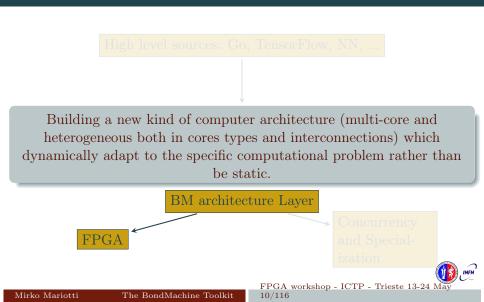
- Multi-core, Two or more independent actual processing units execute multiple instructions at the same time.
 - The power is given by the **number** of cores.
 - Parallelism has to be addressed.
- Heterogeneous, different types of processing units.
 - Cell, GPU, Parallela, TPU.
 - The power is given by the specialization.
 - The units data transfer has to be addressed.
 - The scheduling has to be addressed.

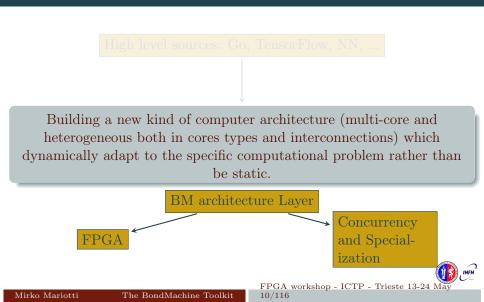


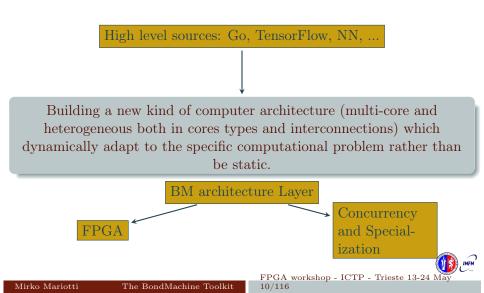
FPGA workshop - ICTP - Trieste 13-24 May

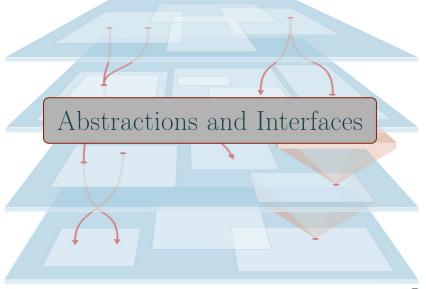












11/116



Layer, Abstractions and Interfaces Introduction

A Computing system is a matter of abstraction and interfaces. A lower layer exposes its functionalities (via interfaces) to the above layer hiding (abstraction) its inner details.

The quality of a computing system is determined by how abstractions are simple and how interfaces are clean.

12/116



Layers, Abstractions and Interfaces

Mirko Mariotti

Programming language

User mode

Kernel mode

Processor

Transistors		
FPGA workshop - ICTP - Trieste 13-24 May		
The BondMachine Toolkit	13/116	

Layers, Abstractions and Interfaces

An example

Mirko Mariotti

Programming language

User mode

Kernel mode

Processor

 Register Machine

 Transistors

 FPGA workshop - ICTP - Trieste 13-24 May

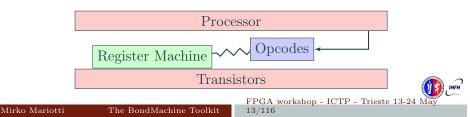
 The BondMachine Toolkit

Layers, Abstractions and Interfaces An example

Programming language

User mode

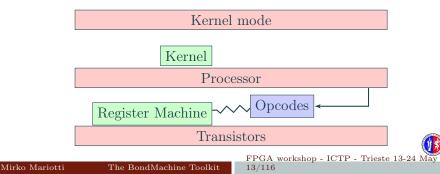
Kernel mode



Layers, Abstractions and Interfaces

Programming language

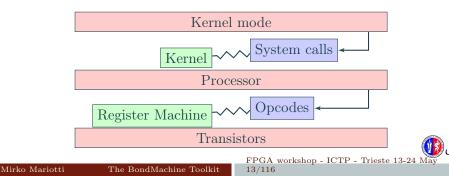
User mode



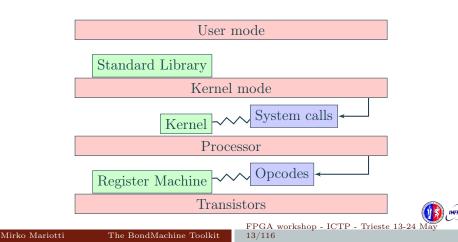
Layers, Abstractions and Interfaces An example

Programming language

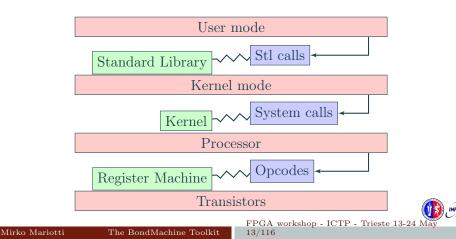
User mode

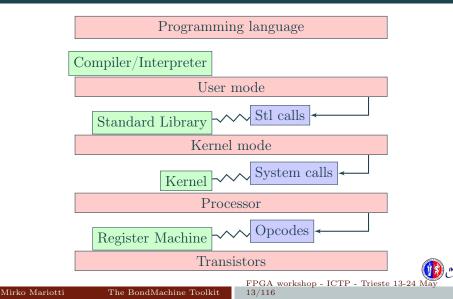


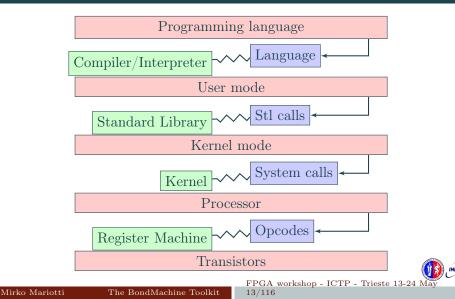
Programming language

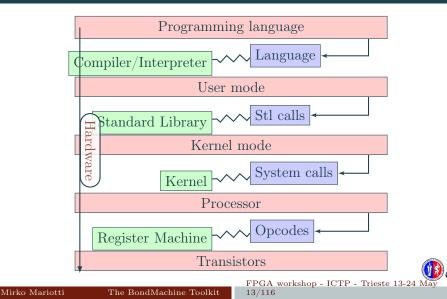


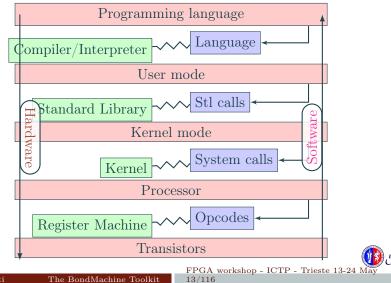
Programming language



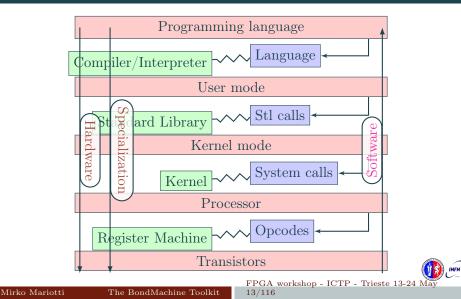


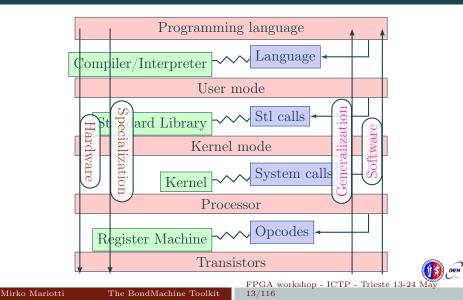


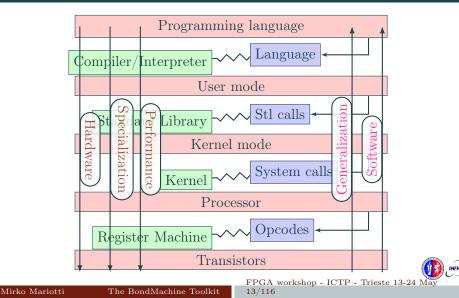




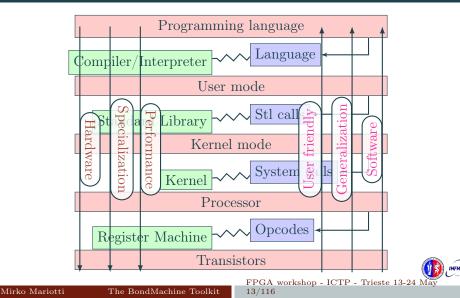
Mirko Mariotti







Layers, Abstractions and Interfaces



Layers, Abstractions and Interfaces The second idea

Build a computing system with a decreased number of layers resulting in a minor gap between HW and SW but keeping an user friendly way of programming it.



BondMachine



Mirko Mariotti

The **BondMachine** is a software ecosystem for the dynamic generation of computer architectures that:

- Are composed by many, possibly hundreds, computing cores.Have very small cores and not necessarily of the same type (different ISA and ABI).
- Have a not fixed way of interconnecting cores.
- May have some elements shared among cores (for example channels and shared memories).



The **BondMachine** is a software ecosystem for the dynamic generation of computer architectures that:

Are composed by many, possibly hundreds, computing cores.

Have very small cores and not necessarily of the same type (different ISA and ABI).

Have a not fixed way of interconnecting cores.

• May have some elements shared among cores (for example channels and shared memories).



The **BondMachine** is a software ecosystem for the dynamic generation of computer architectures that:

- Are composed by many, possibly hundreds, computing cores.Have very small cores and not necessarily of the same type (different ISA and ABI).
- Have a not fixed way of interconnecting cores.
- May have some elements shared among cores (for example channels and shared memories).



The **BondMachine** is a software ecosystem for the dynamic generation of computer architectures that:

- Are composed by many, possibly hundreds, computing cores.
- Have very small cores and not necessarily of the same type (different ISA and ABI).
- Have a not fixed way of interconnecting cores.
 - May have some elements shared among cores (for example channels and shared memories).



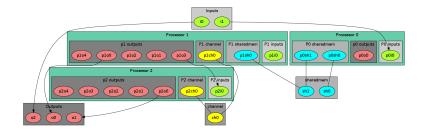
The **BondMachine** is a software ecosystem for the dynamic generation of computer architectures that:

- Are composed by many, possibly hundreds, computing cores.
- Have very small cores and not necessarily of the same type (different ISA and ABI).
- Have a not fixed way of interconnecting cores.
- May have some elements shared among cores (for example channels and shared memories).



The BondMachine

An example





Mirko Mariotti

The BondMachine Toolkit

The computational unit of the BM

The atomic computational unit of a BM is the "connecting processor" (CP) and has:

- Some general purpose registers of size Rsize.
- Some I/O dedicated registers of size Rsize.
- A set of implemented opcodes chosen among many available.
- Dedicated ROM and RAM.
- Three possible operating modes.

The computational unit of the BM

The atomic computational unit of a BM is the "connecting processor" (CP) and has:

- Some general purpose registers of size Rsize.
- Some I/O dedicated registers of size Rsize.
- A set of implemented opcodes chosen among many available.
- Dedicated ROM and RAM.
- Three possible operating modes.

General purpose registers

 2^{R} registers: r0,r1,r2,r3 ... r 2^{R}

The computational unit of the BM

The atomic computational unit of a BM is the "connecting processor" (CP) and has:

- Some general purpose registers of size Rsize.
- Some I/O dedicated registers of size Rsize.
- A set of implemented opcodes chosen among many available.
- Dedicated ROM and RAM.
- Three possible operating modes.

I/O specialized registers

N input registers: i0,i1 ... iN M output registers: o0,o1 ... oM

The computational unit of the BM

The atomic computational unit of a BM is the "connecting processor" (CP) and has:

- Some general purpose registers of size Rsize.
- Some I/O dedicated registers of size Rsize.
- A set of implemented opcodes chosen among many available.
- Dedicated ROM and RAM.
- Three possible operating modes.

Full set of possible opcodes

add, addf, addi, chc, chw, clr, cpy, dec, divf, dpc, hit, hlt, i2r, inc, j, je, jz, m2r, mult, multf, nop, r2m, r2o, r2s, rset, sic, s2r, saj, sub, wrd, wwr

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May18/116

The computational unit of the BM

The atomic computational unit of a BM is the "connecting processor" (CP) and has:

- Some general purpose registers of size Rsize.
- Some I/O dedicated registers of size Rsize.
- A set of implemented opcodes chosen among many available.
- Dedicated ROM and RAM.
- Three possible operating modes.

RAM and ROM

- \blacksquare 2^L RAM memory cells.
- \blacksquare 2^O ROM memory cells.

Mirko Mariotti

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May18/116

The computational unit of the BM

The atomic computational unit of a BM is the "connecting processor" (CP) and has:

- Some general purpose registers of size Rsize.
- Some I/O dedicated registers of size Rsize.
- A set of implemented opcodes chosen among many available.
- Dedicated ROM and RAM.
- Three possible operating modes.

Operating modes

- Full Harvard mode.
- Full Von Neuman mode.
- Hybrid mode.

Mirko Mariotti

The BondMachine Toolkit

Full set of possible opcodes

Opcode	Args	Description
add	reg_dst,reg_add	Add the values in reg_dst and reg_add writing the result in reg_dst
addf	reg_dst,reg_add	Add the values in reg_dst and reg_add writing the result in reg_dst (float32)
addi	reg_dst	Add the values of all the processor inputs in reg_dst
chc	reg_state, reg_op	Check for any channel operation, report the state and eventually which happened
chw	reg_op	Wait for any channel operation and report witch happened on reg_op
cir	reg	Set the register reg to 0
сру	reg_dst, reg_src	Copy the value of a register to another
dec	reg	Decrement a register by 1
di∨f	reg_dst,reg_di∨	Divide the values in reg_dst by reg_div writing the result in reg_dst (float32)
dpc	reg_dest	Decode the program counter into a register
hit	reg_state, barrier_name	Hit a barrier, report the state
hlt	none	Halt the processor
i2r	reg_dst, input_name	Copy the value from an input to a register
inc	reg	Increment a register by 1
j	rom_address	Jump to a given instruction
je	reg1, reg2, rom_address	Jump if the register are equals
jz	reg1, rom_address	Jump if a register is zero
m2r	reg_dest, ram_address	Copy data from the RAM to a register
mult	reg_dst,reg_mult	Multiply the values in reg_mult and reg_dest writing the result in reg_dst
multf	reg_dst,reg_mult	Multiply the values in reg_mult and reg_dest writing the result in reg_dst (float32)
nop	none	No operation
r2m	reg_source, ram_address	Copy data from a register to the RAM
r2o	reg_src, output_name	Copy the value from a register to an output
r2s	reg_source, ram_name, ram_address	Copy data from a register to a shared RAM
rset	reg_dst, numeric_value	Set a value for a register
sic	reg_dst, input_name	Stop until Input Changes accumulating on a register
s2r	reg_dest, ram_name, ram_address	Copy data from a shared RAM to a register
saj	rom or ram_address	Switch operating mode and jump
sub	reg_dst,reg_sub	Subtract the values in reg_sub from reg_dest writing the result in reg_dst
wrd	reg_dst, channel_name	Want read from a channel to a register (set flag)
wwr	reg_src, channel_name	Want write to a channel from a register (set flag)

Mirko Mariotti

The BondMachine Toolkit

The non-computational element of the BM

Alongside CPs, BondMachines include non-computing units called "Shared Objects" (SO).

Examples of their purposes are:

- Data storage (Memories).
- Message passing.
- CP synchronization.

A single SO can be shared among different CPs. To use it CPs have special instructions (opcodes) oriented to the specific SO.

Four kind of SO have been developed so far: the Channel, the Shared Memory, the Barrier and a Pseudo Random Numbers Generator.



The non-computational element of the BM

Alongside CPs, BondMachines include non-computing units called "Shared Objects" (SO).

Examples of their purposes are:

- Data storage (Memories).
- Message passing.
- CP synchronization.

A single SO can be shared among different CPs. To use it CPs have special instructions (opcodes) oriented to the specific SO.

Four kind of SO have been developed so far: the Channel, the Shared Memory, the Barrier and a Pseudo Random Numbers Generator.



The non-computational element of the BM

Alongside CPs, BondMachines include non-computing units called "Shared Objects" (SO).

Examples of their purposes are:

- Data storage (Memories).
- Message passing.
- CP synchronization.

A single SO can be shared among different CPs. To use it CPs have special instructions (opcodes) oriented to the specific SO.

Four kind of SO have been developed so far: the Channel, the Shared Memory, the Barrier and a Pseudo Random Numbers Generator.



The non-computational element of the BM

Alongside CPs, BondMachines include non-computing units called "Shared Objects" (SO).

Examples of their purposes are:

- Data storage (Memories).
- Message passing.
- CP synchronization.

A single SO can be shared among different CPs. To use it CPs have special instructions (opcodes) oriented to the specific SO.

Four kind of SO have been developed so far: the Channel, the Shared Memory, the Barrier and a Pseudo Random Numbers Generator.



The non-computational element of the BM

Alongside CPs, BondMachines include non-computing units called "Shared Objects" (SO).

Examples of their purposes are:

- Data storage (Memories).
- Message passing.
- CP synchronization.

A single SO can be shared among different CPs. To use it CPs have special instructions (opcodes) oriented to the specific SO.

Four kind of SO have been developed so far: the Channel, the Shared Memory, the Barrier and a Pseudo Random Numbers Generator.



The non-computational element of the BM

Alongside CPs, BondMachines include non-computing units called "Shared Objects" (SO).

Examples of their purposes are:

- Data storage (Memories).
- Message passing.
- CP synchronization.

A single SO can be shared among different CPs. To use it CPs have special instructions (opcodes) oriented to the specific SO.

Four kind of SO have been developed so far: the Channel, the Shared Memory, the Barrier and a Pseudo Random Numbers Generator.



Channel

The Channel SO is an hardware implementation of the CSP (communicating sequential processes) channel.

It is a model for inter-core communication and synchronization via message passing.

CPs use channels via 4 opcodes wrd: Want Read. wwr: Want Write. chc: Channel Check. chw: Channel Wait.

21/116

Channel

The Channel SO is an hardware implementation of the CSP (communicating sequential processes) channel.

It is a model for inter-core communication and synchronization via message passing.

J
(M) (MEN

21/116

Channel

The Channel SO is an hardware implementation of the CSP (communicating sequential processes) channel.

It is a model for inter-core communication and synchronization via message passing.

CPs use channels via 4 opcodes	
wrd: Want Read.	
wwr: Want Write.	
■ chc: Channel Check.	
Channel Wait.	
	FPGA workshop - ICTP - Trieste 13-24 May

Shared Memory

The Shared Memory SO is a RAM block accessible from more than one CP.

Different Shared Memories can be used by different CP and not necessarily by all of them.

CPs use shared memories via 2 opcodes

s2r: Shared memory read.

| r2s: Shared memory write.



FPGA workshop - ICTP - Trieste 13-24 May 22/116

Mirko Mariotti The I

The BondMachine Toolkit

Shared Memory

The Shared Memory SO is a RAM block accessible from more than one CP.

Different Shared Memories can be used by different CP and not necessarily by all of them.

CPs use shared memories via 2 opcodes s2r: Shared memory read.

1 r2s: Shared memory write.



FPGA workshop - ICTP - Trieste 13-24 May The BondMachine Toolkit 22/116

Shared Memory

The Shared Memory SO is a RAM block accessible from more than one CP.

Different Shared Memories can be used by different CP and not necessarily by all of them.

CPs use shared memories via 2 opcodes

- s2r: Shared memory read.
 - r2s: Shared memory write.



Mirko Mariotti The Bo

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May 22/116



The Barrier SO is used to make CPs act synchronously.

When a CP hits a barrier, the execution stop until all the CPs that share the same barrier hit it.

CPs use barriers via 1 opcode

hit: Hit the barrier.

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May23/116



The Barrier SO is used to make CPs act synchronously.

When a CP hits a barrier, the execution stop until all the CPs that share the same barrier hit it.

CPs use barriers via 1 opcode

hit: Hit the barrier.

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May23/116



The Barrier SO is used to make CPs act synchronously.

When a CP hits a barrier, the execution stop until all the CPs that share the same barrier hit it.

CPs use barriers via 1 opcode

hit: Hit the barrier.

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May23/116

Multicore and Heterogeneous

First idea on the BondMachine

The idea was:

Having a multi-core architecture completely heterogeneous both in cores types and interconnections.

The BondMachine may have many cores, eventually all different, arbitrarily interconnected and sharing non computing elements.

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May24/116

Architectures Handling

25/116



Mirko Mariotti

The BM computer architecture is managed by a set of tools to:

- build a specify architecture
- modify a pre-existing architecture
- simulate or emulate the behavior
- Generate the Register Tranfer Code (RTL)

Processor Builder

Selects the single processor, assembles and disassembles, saves on disk as JSON, creates the RTL code of a CP

BondMachine Builder

Connects CPs and SOs together in custom topologies, loads and saves on disk as JSON, create BM's RTL code

Simulation Framework

Simulates the behaviour, emulates a BM on a standard Linux workstation



The BM computer architecture is managed by a set of tools to:

- build a specify architecture
- modify a pre-existing architecture
- simulate or emulate the behavior
- Generate the Register Tranfer Code (RTL)

Processor Builder

Selects the single processor, assembles and disassembles, saves on disk as JSON, creates the RTL code of a CP

BondMachine Builder

Connects CPs and SOs together in custom topologies, loads and saves on disk as JSON, create BM's RTL code

Simulation Framework

Simulates the behaviour, emulates a BM on a standard Linux workstation



The BM computer architecture is managed by a set of tools to:

- build a specify architecture
- modify a pre-existing architecture
- simulate or emulate the behavior
- Generate the Register Tranfer Code (RTL)

Processor Builder

Selects the single processor, assembles and disassembles, saves on disk as JSON, creates the RTL code of a CP

BondMachine Builder

Connects CPs and SOs together in custom topologies, loads and saves on disk as JSON, create BM's RTL code

Simulation Framework

Simulates the behaviour, emulates a BM on a standard Linux workstation



FPGA workshop - ICTP - Trieste 13-24 May 26/116

The BM computer architecture is managed by a set of tools to:

- build a specify architecture
 - modify a pre-existing architecture
 - simulate or emulate the behavior
 - Generate the Register Tranfer Code (RTL)

Processor Builder

Selects the single processor, assembles and disassembles, saves on disk as JSON, creates the RTL code of a CP

BondMachine Builder

Connects CPs and SOs together in custom topologies, loads and saves on disk as JSON, create BM's RTL code

Simulation Framework

Simulates the behaviour, emulates a BM on a standard Linux workstation



FPGA workshop - ICTP - Trieste 13-24 May 26/116

Procbuilder is the CP manipulation tool.



Examples

(32 bit registers counter machine)

procbuilder -register-size 32 -opcodes clr,cpy,dec,inc,je,jz

(Input and Output registers)

proc
builder -
inputs 3 -outputs $2\ \dots$

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May27/116

Procbuilder is the CP manipulation tool.

CP Creation CP Load/Save CP Assembler/Disassembler CP RTL

(Loading a CP) procbuilder -load-machine conproc.json ...

(Saving a CP)

procbuilder -save-machine conproc.json ...

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May27/116

Procbuilder is the CP manipulation tool.

CP Creation CP Load/Save CP Assembler/Disassembler CP RTL

(Assembiling) prochuilder -input-assembly program.asm ...

(Disassembling)

procbuilder -show-program-disassembled ...

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May27/116

Procbuilder is the CP manipulation tool.

CP Creation CP Load/Save CP Assembler/Disassembler CP RTL

Examples

(Create the CP RTL code in Verilog) procbuilder -create-verilog ...

(Create testbench)

proc
builder -create-verilog-test
bench test.v \ldots

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May27/116

Goals are:

- To create a simple processor
- To assemble and disassemble code for it
- To produce its RTL code



Bondmachine is the tool that compose CP and SO to form BondMachines.

BM CP insert and remove

BM SO insert and remove BM Inputs and Outputs BM Bonding Processors and/or IO BM Visualizing or RTL



Bondmachine is the tool that compose CP and SO to form BondMachines.

BM CP insert and remove

BM SO insert and remove

BM Inputs and Outputs

BM Bonding Processors and/or IO

BM Visualizing or RTL

(Add a Shared Object) bondmachine -add-shared-objects specs ... (Connect an SO to a processor) bondmachine -connect-processor-shared-object ...

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May29/116

Bondmachine is the tool that compose CP and SO to form BondMachines.

BM CP insert and remove BM SO insert and remove BM Inputs and Outputs BM Bonding Processors and/or IO BM Visualizing or RTL

Examples

(Adding inputs or outputs) bondmachine -add-inputs ... ; bondmachine -add-outputs ...

(Removing inputs or outputs)

bondmachine -del-input ... ; bondmachine -del-output ...

Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May29/116

Bondmachine is the tool that compose CP and SO to form BondMachines.

BM CP insert and remove BM SO insert and remove BM Inputs and Outputs BM Bonding Processors and/or IO BM Visualizing or RTL

Examples				
	(Bonding p	rocessor)		
bondmachine -add-bond p0i2,p1o4				
(Bonding IO)				
bondmachine -add-bond i2,p0i6				
Mirko Mariotti		FPGA workshop - ICTP - Trieste 13-24 May 29/116		

Bondmachine is the tool that compose CP and SO to form BondMachines.

BM CP insert and remove BM SO insert and remove BM Inputs and Outputs BM Bonding Processors and/or IO BM Visualizing or RTL

Examples				
	(Visualizing)			
bondmachine -emit-dot				
(Create RTL code)				
bondmachine -create-verilog				
Mirko Mariotti	FPGA workshop - ICTP - Trieste 13 The BondMachine Toolkit 29/116	3-24 May		

Goals are:

- To create a single-core BondMachine
- To attach an external output
- To produce its RTL code



A set of toolchains allow the build and the direct deploy to a target device of BondMachines.

Bondgo Toolchain example

A file local.mk contains references to the source code as well all the build necessities.

make bondmachine creates the JSON representation of the BM and assemble its code.

make show displays a graphical representation of the BM.

make simulate start a simulation.

make videosim create a simulation video.

make flash the device into the destination target.

Goals are:

To explore the toolchain

To flash the board with the code from the previous example

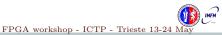


Goals are:

■ To build a BondMachine with a processor and a shared object

33/116

To flash the board



Mirko Mariotti

The BondMachine Toolkit

Goals are:

- To build a dual-core BondMachine
- To connect cores
- To flash the board



Mirko Mariotti

BondMachine web front-end

Operations on BondMachines can also be performed via an under development web framework

B	Test I/O and Bonds Processors Prozes Navagement Octputs Management Reads Management	Shared Objects
Ochine	Bonds	Layout
Tools Processors Bordinachnes Examples Brog Prog Projects Text	Bits Bitson 1 tan Bitson 2 tan Bitson 2 tan 1 1 1 1 Date tan 1 1 1 Date tan Date tan 1 1 1 Date tan Date tan 1 1 Date tan Date tan Date tan	Processor 2 Processor 2 Proce
	Consep	on a Gaper Water 126 - Gaparger 19 203 - Main Menter - Those may be fair (225 - Manine)



FPGA workshop - ICTP - Trieste 13-24 May 35/116

The BondMachine Toolkit

Mirko Mariotti

An important feature of the tools is the possibility of simulating BondMachine behavior.

An event input file describes how BondMachines elements has to change during the simulation timespan and which one has to be be reported.

The simulator can produce results in the form of:

- Activity log of the BM internal.
- Graphical representation of the simulation.
- Report file with quantitative data. Useful to construct metrics



An important feature of the tools is the possibility of simulating BondMachine behavior.

An event input file describes how BondMachines elements has to change during the simulation timespan and which one has to be be reported.

The simulator can produce results in the form of:

- Activity log of the BM internal.
- Graphical representation of the simulation.
- Report file with quantitative data. Useful to construct metrics



An important feature of the tools is the possibility of simulating BondMachine behavior.

An event input file describes how BondMachines elements has to change during the simulation timespan and which one has to be be reported.

The simulator can produce results in the form of:

- Activity log of the BM internal.
- Graphical representation of the simulation.
- Report file with quantitative data. Useful to construct metrics



Activity log example:

IIIIIII
[discovery]> /home/mirko/Projects/comproc/tests/asm2sim % bondmachine -register-size 8 -bondmachine-file asmtest05.json -sim -sim
Loading simbox rule: config:show_pc
Loading simbox rule: config:show_ticks
Loading simbox rule: config:show_instruction
Loading simbox rule: config:show_disasm
Loading simbox rule: config:show_proc_io_pre
Loading simbox rule: config:show_proc_io_post
Loading simbox rule: config:show_proc_regs_pre
Loading simbox rule: config:show_proc_regs_post
Loading simbox rule: config:show_io_post
Loading simbox rule: config:show_io_pre
Loading simbox rule: absolute:1:set:10:2
Absolute tick:0
Pre-compute I0: 10: 00000000 o0: 00000000
Proc: 0
PC: 0
Instr: 00000
Disasm: i2r r0 i0
Pre-compute ID; i0: 00000000 a0: 00000000 Pre-compute Regs: r0: 00000000 r1: 0000000
Pre-compute regg: P0: 0000000 P1: 0000000 Post-compute ID: 10: 0000000 00000000
Post-compute 10: 0000000 01: 0000000 Post-compute Ress: r0: 0000000 r1: 0000000
Post-compute [0: 00: 000000 0: 0000000
Absolute tick:1
Pre-compute ID: 10: 00000010 o0: 00000000
Proc: 0
PC: 1
Instr: 00000
Disasm: i2r r0 i0
Pre-compute I0: 10: 00000010 o0: 00000000
Pre-compute Regs: r0: 00000000 r1: 00000000
Post-compute ID: 10: 00000010 o0: 00000000
Post-compute Reas: r0: 0000010 r1: 0000000
Post-compute I0: 10: 00000010 o0: 00000000
Absolute tick:2

A graphical example:



The BondMachine Toolkit



Simulation hands-on Hands-on N.6

Goals are:

■ To show the simulation capabilities of the framework



Mirko Mariotti

The BondMachine Toolkit



The same engine that simulate BondMachines can be used as emulator.

Through the emulator BondMachines can be used on Linux workstations.



Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May 39/116

Architectures Molding



FPGA workshop - ICTP - Trieste 13-24 May 40/116

Molding the BondMachine

As stated before BondMachines are not general purpose architectures, and to be effective have to be shaped according the specific problem.

Several methods (apart from writing in assembly and building a Bond-Machine from scratch) have been developed to do that:

- bondgo: A new type of compiler that create not only the CPs assembly but also the architecture itself.
- A set of API to create BondMachine to fit a specific computational problems.
- An Evolutionary Computation framework to "grow" BondMachines according some fitness function via simulation.
 - A set of tools to use BondMachine in Machine Learning.

FPGA workshop - ICTP - Trieste 13-24 May41/116

Molding the BondMachine

As stated before BondMachines are not general purpose architectures, and to be effective have to be shaped according the specific problem.

Several methods (apart from writing in assembly and building a Bond-Machine from scratch) have been developed to do that:

- **b**ondgo: A new type of compiler that create not only the CPs assembly but also the architecture itself.
- A set of API to create BondMachine to fit a specific computational problems.
- An Evolutionary Computation framework to "grow" BondMachines according some fitness function via simulation.
 - A set of tools to use BondMachine in Machine Learning.

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May41/116

Molding the BondMachine

As stated before BondMachines are not general purpose architectures, and to be effective have to be shaped according the specific problem.

Several methods (apart from writing in assembly and building a Bond-Machine from scratch) have been developed to do that:

- bondgo: A new type of compiler that create not only the CPs assembly but also the architecture itself.
- A set of API to create BondMachine to fit a specific computational problems.
- An Evolutionary Computation framework to "grow" BondMachines according some fitness function via simulation.
 - A set of tools to use BondMachine in Machine Learning.

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May41/116

Molding the BondMachine

As stated before BondMachines are not general purpose architectures, and to be effective have to be shaped according the specific problem.

Several methods (apart from writing in assembly and building a Bond-Machine from scratch) have been developed to do that:

■ bondgo: A new type of compiler that create not only the CPs assembly but also the architecture itself.

• A set of API to create BondMachine to fit a specific computational problems.

An Evolutionary Computation framework to "grow" BondMachines according some fitness function via simulation.

A set of tools to use BondMachine in Machine Learning.

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May41/116

Molding the BondMachine

As stated before BondMachines are not general purpose architectures, and to be effective have to be shaped according the specific problem.

Several methods (apart from writing in assembly and building a Bond-Machine from scratch) have been developed to do that:

- bondgo: A new type of compiler that create not only the CPs assembly but also the architecture itself.
- A set of API to create BondMachine to fit a specific computational problems.
- An Evolutionary Computation framework to "grow" BondMachines according some fitness function via simulation.

A set of tools to use BondMachine in Machine Learning.

FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May41/116

Molding the BondMachine

As stated before BondMachines are not general purpose architectures, and to be effective have to be shaped according the specific problem.

Several methods (apart from writing in assembly and building a Bond-Machine from scratch) have been developed to do that:

- bondgo: A new type of compiler that create not only the CPs assembly but also the architecture itself.
- A set of API to create BondMachine to fit a specific computational problems.
- An Evolutionary Computation framework to "grow" BondMachines according some fitness function via simulation.
- A set of tools to use BondMachine in Machine Learning.

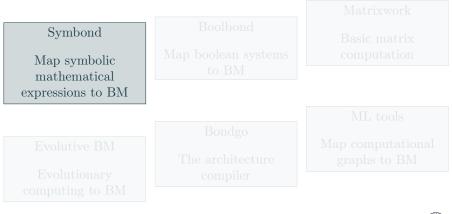
FPGA workshop - ICTP - Trieste $13\mathchar`-24$ May41/116

Mapping specific computational problems to BMs



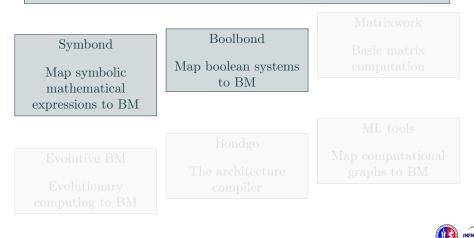
FPGA workshop - ICTP - Trieste 13-24 May 42/116

Mapping specific computational problems to BMs

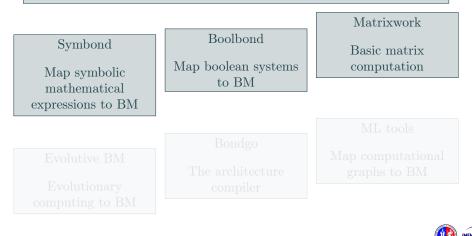




Mapping specific computational problems to BMs

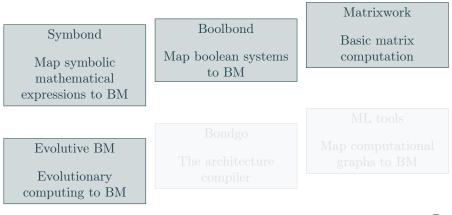


Mapping specific computational problems to BMs



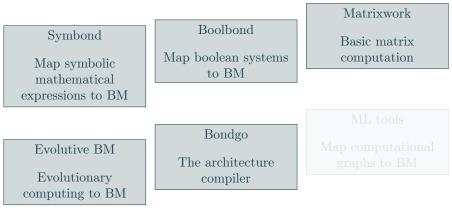
FPGA workshop - ICTP - Trieste 13-24 May 42/116

Mapping specific computational problems to BMs



42/116

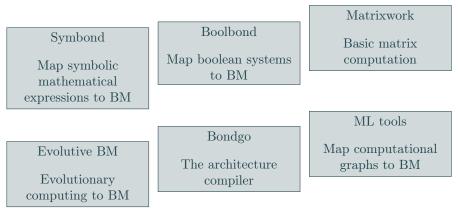
Mapping specific computational problems to BMs





FPGA workshop - ICTP - Trieste 13-24 May 42/116

Mapping specific computational problems to BMs





FPGA workshop - ICTP - Trieste 13-24 May 42/116



FPGA workshop - ICTP - Trieste 13-24 May43/116

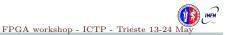


The major innovation of the BondMachine Project is its compiler.

Bondgo is the name chosen for the compiler developed for the BondMachine.

The compiler source language is Go as the name suggest.

44/116





This is the standard flow when building computer programs



Mirko Mariotti

FPGA workshop - ICTP - Trieste 13-24 May 45/116



This is the standard flow when building computer programs

high level language source

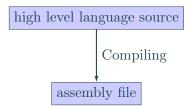


Mirko Mariotti





This is the standard flow when building computer programs

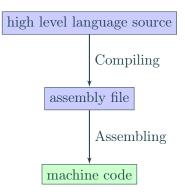


45/116





This is the standard flow when building computer programs



bondgo loop example

```
package main
import ()
func main() {
   var reg_aa uint8
  var reg_ab uint8
   for reg_aa = 10; reg_aa > 0; reg_aa-- {
      reg_ab = reg_aa
      break
}
```

bondgo loop example in asm

clr aa clr ab		
rset ac 10		
cpy aa ac		
cpy ac aa		
jz ac 11		
cpy ac aa		
cpy ab ac		
j 11		
dec aa		
j 4		
		FPGA workshop - ICTP - Trieste 13-24 May
Mirko Mariotti	The BondMachine Toolkit	46/116



Bondgo does something different from standard compilers ...



Mirko Mariotti

FPGA workshop - ICTP - Trieste 13-24 May 47/116





Bondgo does something different from standard compilers ...

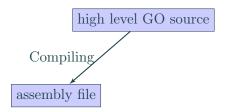
high level GO source



Mirko Mariotti

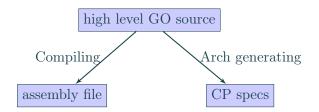
47/116



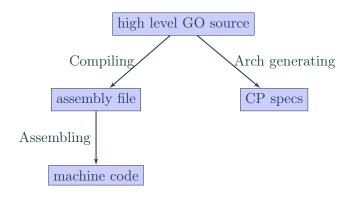






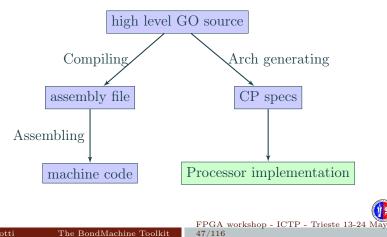




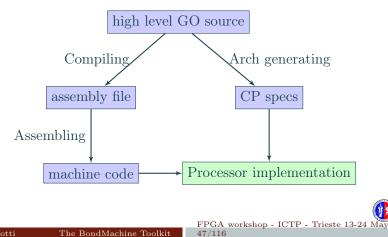


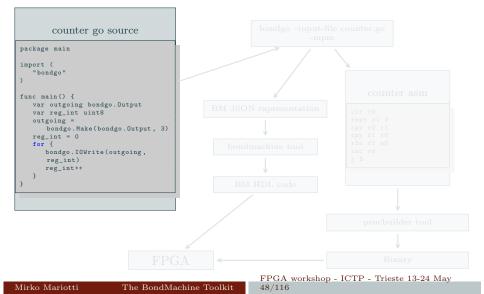




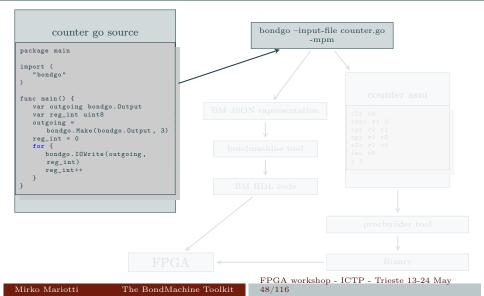


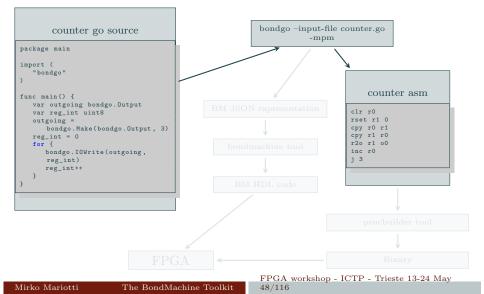


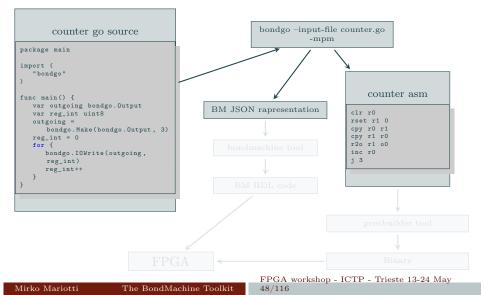


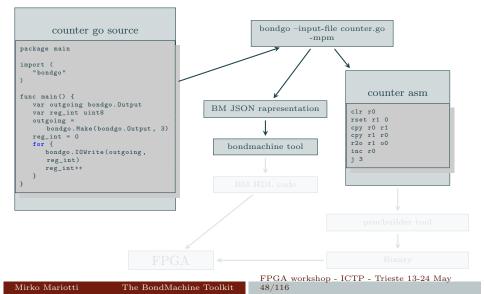


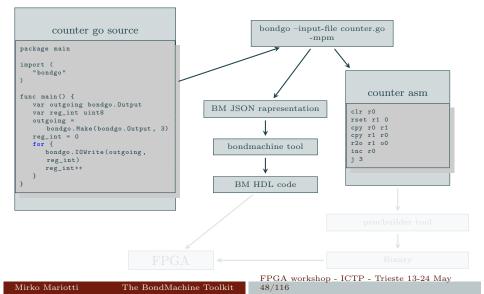
Bondgo A first example

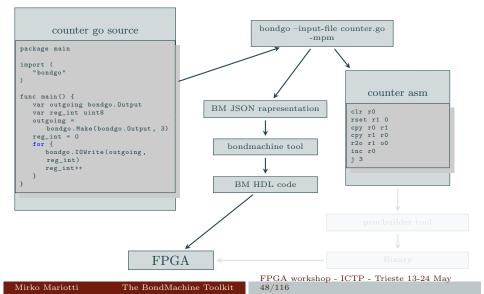


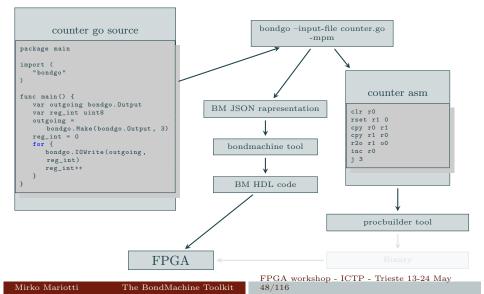


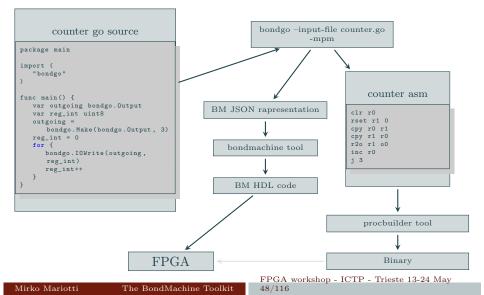


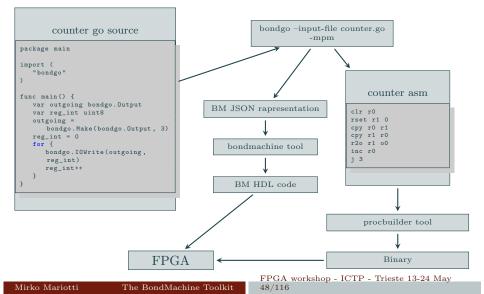












49/116

Bondgo hands-on

Goals are:

- To create a BondMachine from a Go source file
- To build the architecture
- To build the program
- To create the firmware and flash it to the board





\dots bond go may not only create the binaries, but also the CP architecture, and \dots

50/116





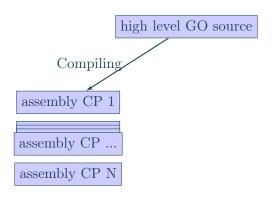




... it can do even much more interesting things when compiling concurrent programs.

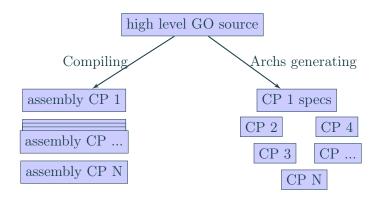
high level GO source



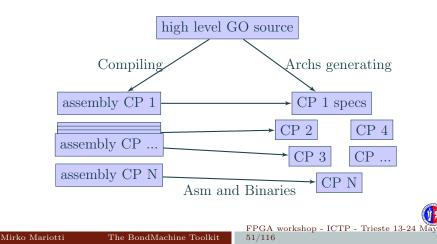




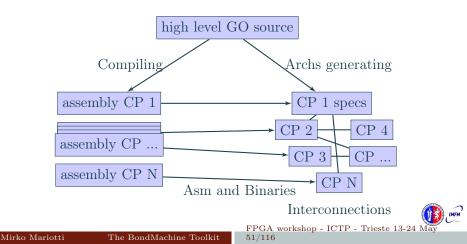
Bondgo



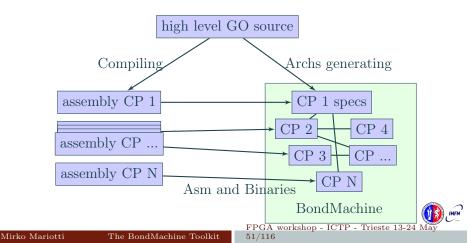
Bondgo



Bondgo



Bondgo



multi-core counter

```
package main
import (
   "bondgo"
func pong() {
   var inO bondgo.Input
   var out0 bondgo.Output
   in0 = bondgo.Make(bondgo.Input, 3)
   out0 = bondgo.Make(bondgo.Output, 5)
   for {
      bondgo.IOWrite(out0, bondgo.IORead(in0)+1)
func main() {
   var inO bondgo.Input
   var out0 bondgo.Output
   in0 = bondgo.Make(bondgo.Input, 5)
   out0 = bondgo.Make(bondgo.Output, 3)
device 0:
   go pong()
   for {
      bondgo.IOWrite(out0, bondgo.IORead(in0))
```

```
Mirko Mariotti
```

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May 52/116

Compiling the code with the bondgo compiler:

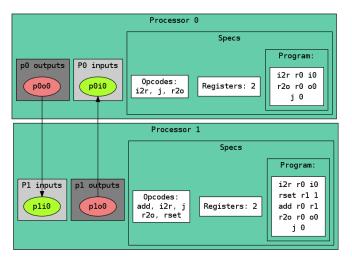
```
bondgo -input-file ds.go -mpm
```

The toolchain perform the following steps:

- Map the two goroutines to two hardware cores.
- Creates two types of core, each one optimized to execute the assigned goroutine.
- Creates the two binaries.
- Connected the two core as inferred from the source code, using special IO registers.

The result is a multicore BondMachine:





FPGA workshop - ICTP - Trieste 13-24 May 54/116

Goals are:

To use bondgo to create a chain of interconnected processors

55/116

To flash the firmware to the board



Mirko Mariotti

The BondMachine Toolkit

Compiling Architectures

One of the most important result

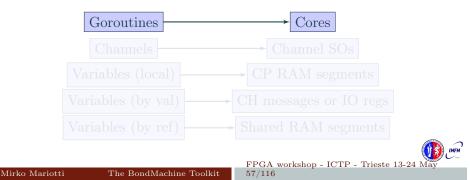
The architecture creation is a part of the compilation process.

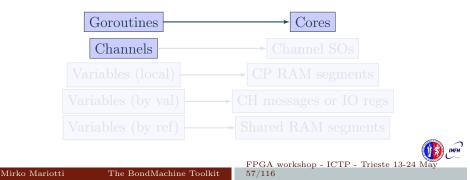


FPGA workshop - ICTP - Trieste 13-24 May 56/116

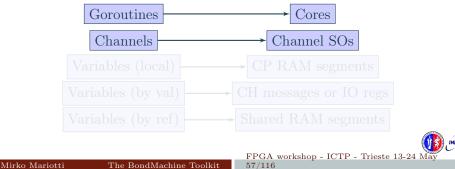




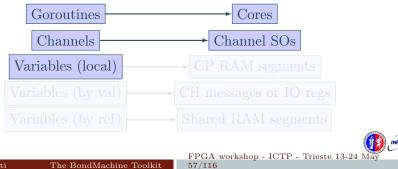




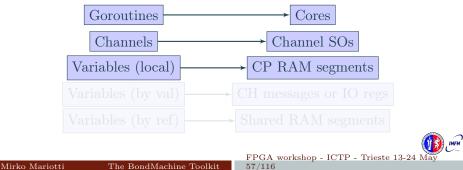
High level Go source code is directly mapped to interconnected processors without Operating Systems or runtimes.



The BondMachine Toolkit

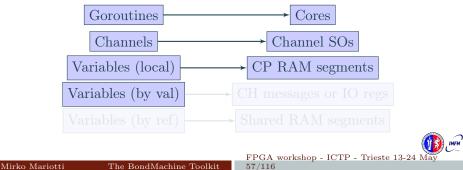


High level Go source code is directly mapped to interconnected processors without Operating Systems or runtimes.

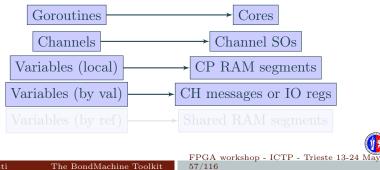


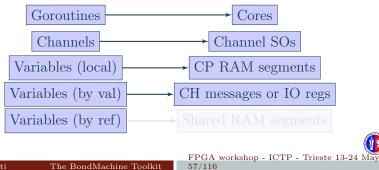
The BondMachine Toolkit

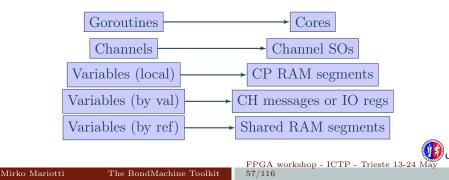
High level Go source code is directly mapped to interconnected processors without Operating Systems or runtimes.



The BondMachine Toolkit







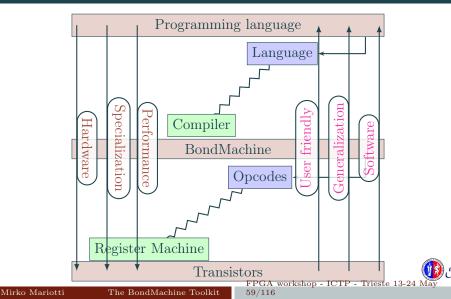
Go in hardware

The idea was: Build a computing system with a decreased number of layers resulting in a lower HW/SW gap. This would raise the overall performances vet keeping an user friendly way of programming.

Between HW and SW there is only the processor abstraction, no Operating System nor runtimes. Despite that programming is done at high level.

FPGA workshop - ICTP - Trieste 13-24 May 58/116

Layers, Abstractions and Interfaces and BondMachines



bondgo stream processing example

```
package main
import (
    "bondgo"
)
func streamprocessor(a *[]uint8, b *[]uint8,
    c *[]uint8, gid uint8) {
    ((c)[gid] = (*a)[gid] + (*b)[gid]
}
func main() {
    a := make([]uint8, 256)
    b := make([]uint8, 256)
    c := make([]uint8, 256)
    c := make([]uint8, 256)
    // ... some a and b values fill
    for i := 0; i < 256; i++ {
        go streamprocessor(&a, &b, &c, uint8(i))
    }
}
```

The compilation of this example results in the creation of a 257 CPs where 256 are the stream processors executing the code in the function called streamprocessor, and one is the coordinating CP. Each stream processor is optimized and capable only to make additions since it is the only operation requested by the source code. The three slices created on the main function are passed by reference to the Goroutines then a shared RAM is created by the Bondgo compiler available to the generated CPs.



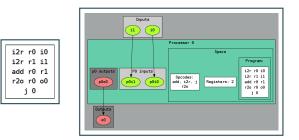
Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May 60/116

The Assembly language for the BM has been kept as independent as possible from the particular CP.

Given a specific piece of assembly code Bondgo has the ability to compute the "minimum CP" that can execute that code.



These are Building Blocks for complex BondMachines.

61/116



Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May62/116

With these Building Blocks

Several libraries have been developed to map specific problems on BondMachines:

63/116

- **Symbond**, to handle mathematical expression.
- Boolbond, to map boolean expression.
- Matrixwork, to perform matrices operations.
- Neuralbond, to use neural networks.



With these Building Blocks

Several libraries have been developed to map specific problems on BondMachines:

Symbond, to handle mathematical expression.

- Boolbond, to map boolean expression.
- Matrixwork, to perform matrices operations.
- Neuralbond, to use neural networks.



Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May 63/116

With these Building Blocks

Several libraries have been developed to map specific problems on BondMachines:

Symbond, to handle mathematical expression.

- Boolbond, to map boolean expression.
- Matrixwork, to perform matrices operations.
- Neuralbond, to use neural networks.



Mirko Mariotti

The BondMachine Toolkit

FPGA workshop - ICTP - Trieste 13-24 May 63/116

With these Building Blocks

Several libraries have been developed to map specific problems on BondMachines:

63/116

- **Symbond**, to handle mathematical expression.
- Boolbond, to map boolean expression.
- Matrixwork, to perform matrices operations.
- Neuralbond, to use neural networks.



With these Building Blocks

Several libraries have been developed to map specific problems on BondMachines:

- **Symbond**, to handle mathematical expression.
- Boolbond, to map boolean expression.
- Matrixwork, to perform matrices operations.
- Neuralbond, to use neural networks.



A mathematical expression, or a system can be converted to a BondMachine:

sum(var(x), const(2))

Boolbond

symbond -expression "sum(var(x), const(2))" -save-bondmachine bondmachine.json

Resulting in:

64/116



A mathematical expression, or a system can be converted to a BondMachine:

sum(var(x), const(2))

Boolbond

symbond -expression "sum(var(x), const(2))" -save-bondmachine bondmachine.json

Resulting in:

64/116

FPGA workshop - ICTP - Trieste 13-24 May



A mathematical expression, or a system can be converted to a BondMachine:

sum(var(x), const(2))

Boolbond

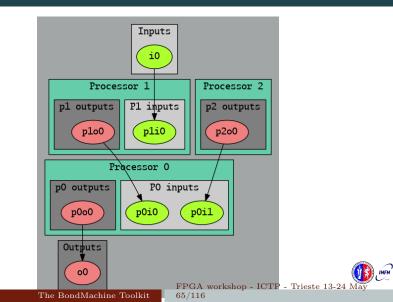
symbond -expression "sum(var(x), const(2))" -save-bond machine bond machine.json

Resulting in:

64/116

FPGA workshop - ICTP - Trieste 13-24 May

Builders API



Mirko Mariotti

A system of boolean equations, input and output variables are expressed as in the example file:

```
var(z)=or(var(x),not(var(y)))
var(b)=or(and(var(x),var(y)),var(z))
var(l)=and(xor(var(x),var(y)),var(t)
i:var(x)
i:var(x)
o:var(z)
o:var(t)
o:var(l)
```

Boolbond

boolbond -system-file expression.txt -save-bondmachine bondmachine.json

Resulting in:



A system of boolean equations, input and output variables are expressed as in the example file:

```
var(z)=or(var(x),not(var(y)))
var(t)=or(and(var(x),var(y)),var(z))
var(l)=and(vor(var(x),var(y)),var(t))
i:var(x)
i:var(y)
o:var(z)
o:var(z)
o:var(t)
o:var(l)
```

Boolbond

boolbond -system-file expression.txt -save-bondmachine bondmachine.json

Resulting in:



A system of boolean equations, input and output variables are expressed as in the example file:

```
var(z)=or(var(x),not(var(y)))
var(t)=or(and(var(x),var(y)),var(z))
var(l)=and(var(x),var(y)),var(t))
i:var(x)
i:var(y)
o:var(z)
o:var(t)
o:var(l)
```

Boolbond

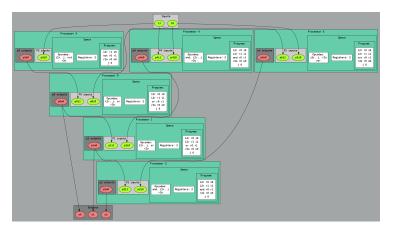
boolbond -system-file expression.txt -save-bond machine bond machine.json

Resulting in:





Builders API Boolbond





Boolbond hands-on

Goals are:

■ To create complex multi-cores from boolean expressions

68/116



Mirko Mariotti

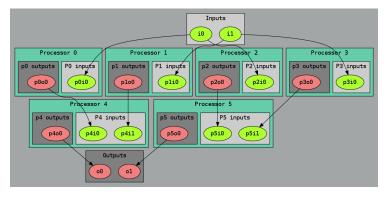
The BondMachine Toolkit

Builders API

Matrixwork

Matrix multiplication

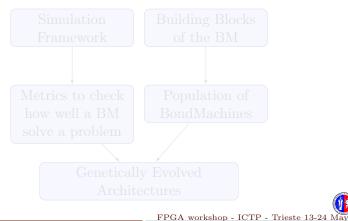
if mymachine, $ok := matrixwork.Build_M(n, t)$; ok == nil ...





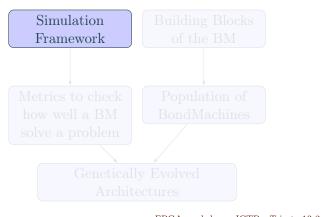


Find an architecture that solve a problem



Mirko Mariotti

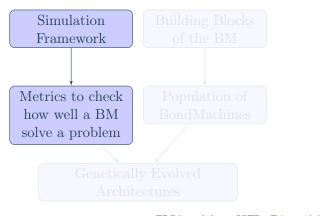
Find an architecture that solve a problem



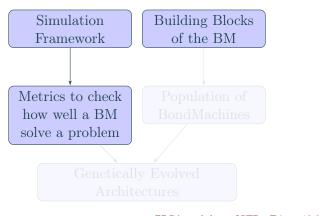
Mirko Mariotti

The BondMachine Toolkit

Find an architecture that solve a problem

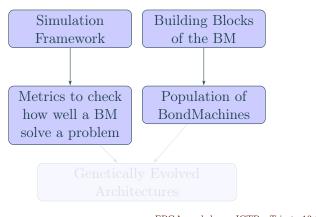


Find an architecture that solve a problem



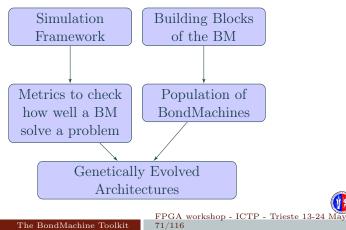
Mirko Mariotti

Find an architecture that solve a problem



Mirko Mariotti

Find an architecture that solve a problem



Mirko Mariotti

Machine Learning



73/116

Machine Learning with BondMachine

Architectures with multiple interconnected processors like the ones produced by the BondMachine Toolkit are a perfect fit for Neural Networks and Computational Graphs.

Several ways to map this structures to BondMachine has been developed:

- A native Neural Network library
- A Tensorflow to BondMachine translator
- An NNEF based BondMachine composer



73/116

Machine Learning with BondMachine

Architectures with multiple interconnected processors like the ones produced by the BondMachine Toolkit are a perfect fit for Neural Networks and Computational Graphs.

Several ways to map this structures to BondMachine has been developed:

- A native Neural Network library
 - A Tensorflow to BondMachine translator
- An NNEF based BondMachine composer



Machine Learning with BondMachine Native Neural Network library

The tool neuralbond allow the creation of BM-based neural chips from an API go interface.

Neurons are converted to BondMachine connecting processors.Tensors are mapped to CP connections.

<pre>layers := []ist(2, 5, 2) weights := nake(]heuralbond.Weight, 0) if save.bondmakhins != " { for save.bondmakhins != " { save.bondmakhins != " { save.bondmakhins); os.la%etKint(farr { f.err := os.ftat(farve_bondmakhins); check(farr) check(farr) defor 1.Close() } }</pre>	
}	



$\underset{tf2bm}{\mathrm{TensorFlow}^{\mathsf{TM}}} \text{ to Bondmachine }$

$$\label{eq:compared} \begin{split} \text{TensorFlow}^{\mathbb{M}} \text{ is an open source software library for numerical} \\ \text{ computation using data flow graphs.} \end{split}$$

Graphs can be converted to BondMachines with the tf2bm tool.





Machine Learning with BondMachine NNEF Composer

Neural Network Exchange Format (NNEF) is a standard from Khronos Group to enable the easy transfer of trained networks among frameworks, inference engines and devices

The NNEF BM tool approach is to descent NNEF models and build BondMachine multi-core accordingly

This approch has several advandages over the previous:

- It is not limited to a single framework
- NNEF is a textual file, so no complex operations are needed to read models



Hardware



Mirko Mariotti

Hardware implementation FPGA

The RTL code for the BondMachine is written in Verilog and System Verilog, and has been tested on these devices/system:

- Digilent Basys3 Xilinx Artix-7 Vivado.
- Kintex7 Evaluation Board Vivado.
- Digilent Zedboard Xilinx Zynq 7020 Vivado.
- Linux Iverilog.
- Terasic De10nano Intel Cyclone V Quartus

Within the project other firmwares have been written or tested:

- Microchip ENC28J60 Ethernet interface controller.
- Microchip ENC424J600 10/100 Base-T Ethernet interface controller.

78/116

FPGA workshop - ICTP - Trieste 13-24 May

ESP8266 Wi-Fi chip.

The Prototype

The project has been selected for the participation at MakerFaire 2016 Rome (The Europen Edition) and a prototype has been assembled and presented.



First run: https://youtube.com/embed/hukTrGxTb7A FPGA workshop - ICTP - Trieste 13-24 May



The BondMachine Toolkit



Clustering



Mirko Mariotti

So far we saw:

- An user friendly approach to create processors (single core).
- Optimizing a single device to support intricate computational work-flows (multi-cores) over an heterogeneous layer.

Interconnected BondMachines

What if we could extend the this layer to multiple interconnected devices ?



Mirko Mariotti

So far we saw:

An user friendly approach to create processors (single core).

Optimizing a single device to support intricate computational work-flows (multi-cores) over an heterogeneous layer.

Interconnected BondMachines

What if we could extend the this layer to multiple interconnected devices ?



So far we saw:

- An user friendly approach to create processors (single core).
- Optimizing a single device to support intricate computational work-flows (multi-cores) over an heterogeneous layer.

Interconnected BondMachines

What if we could extend the this layer to multiple interconnected devices ?



So far we saw:

- An user friendly approach to create processors (single core).
- Optimizing a single device to support intricate computational work-flows (multi-cores) over an heterogeneous layer.

Interconnected BondMachines

What if we could extend the this layer to multiple interconnected devices ?

81/116



Mirko Mariotti

The same logic existing among CP have been extended among different BondMachines organized in clusters.

Protocols, one ethernet called etherbond and one using UDP called udpbond have been created for the purpose.

FPGA based BondMachines, standard Linux Workstations, Emulated BondMachines might join a cluster an contribute to a single distributed computational problem.



The same logic existing among CP have been extended among different BondMachines organized in clusters.

Protocols, one ethernet called etherbond and one using UDP called udpbond have been created for the purpose.

FPGA based BondMachines, standard Linux Workstations, Emulated BondMachines might join a cluster an contribute to a single distributed computational problem.

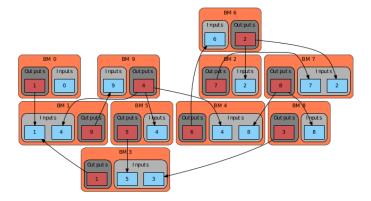


The same logic existing among CP have been extended among different BondMachines organized in clusters.

Protocols, one ethernet called etherbond and one using UDP called udpbond have been created for the purpose.

FPGA based BondMachines, standard Linux Workstations, Emulated BondMachines might join a cluster an contribute to a single distributed computational problem.





A distributed example

distributed counter

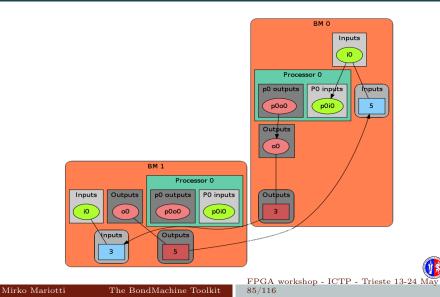
```
package main
import (
   "bondgo"
func pong() {
   var inO bondgo.Input
   var out0 bondgo.Output
   in0 = bondgo.Make(bondgo.Input, 3)
   out0 = bondgo.Make(bondgo.Output, 5)
   for {
      bondgo.IOWrite(out0, bondgo.IORead(in0)+1)
func main() {
   var inO bondgo.Input
   var out0 bondgo.Output
   in0 = bondgo.Make(bondgo.Input, 5)
   out0 = bondgo.Make(bondgo.Output, 3)
device 1:
   go pong()
   for {
      bondgo.IOWrite(out0, bondgo.IORead(in0))
```

The BondMachine Toolkit

Redeployer

BondMachine Clustering

A distributed example



Redeployer

BondMachine Clustering A distributed example

The result is: https://youtube.com/embed/g9xYHKOzca4

A general result

Parts of the system can be redeployed among different devices without changing the system behavior (only the performances).



Results

User can deploy an entire HW/SW cluster starting from code written in a high level description (Go, NNEF, etc)

Workstation with emulated BondMachines, workstation with etherbond drivers, standalone BondMachines (FPGA) may join these clusters.



Mirko Mariotti

The BondMachine Toolkit

BondMachine Clustering

Results

User can deploy an entire HW/SW cluster starting from code written in a high level description (Go, NNEF, etc)

• Workstation with emulated BondMachines, workstation with etherbond drivers, standalone BondMachines (FPGA) may join these clusters.







Two use cases in Physics experiments are currently being developed:

Real time pulse shape analysis in neutron detectorsbringing the intelligence to the edge

Test beam for space experiments (DAMPE, HERD)increasing testbed operations efficiency

Physics

The operation of the new generation of high-intensity neutron sources like SNS, JSNS and European Spallation Source (ESS, Lund, Sweden), now under construction, are introducing a new demand for neutron detection capabilities.

These demands yield to the need for new data collection procedures and new technology based on solid state Si devices.

We are trying to use BondMachines to make the real time shape analysis in this kind of detecting devices.

90/116

Courtesv of Prof. F.Sacchetti



Test beam for space experiments (DAMPE, HERD) Trigger logic for test beams

In test beams, the DAQ system relies on the trigger system for data tacking (sensor signal digitization) during

- Calibration (random trigger or "off-spill" trigger)
- On spill data taking

Minimum elements used for trigger system:

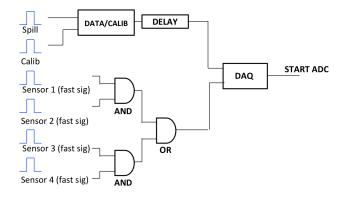
- Clock, pulser
- Logic gates (AND, OR,...)
- Delays

Trigger system implemented using NIM crates and DAQ machines

Courtesy of V.Vagelli and M.Duranti



Test beam for space experiments (DAMPE, HERD) Trigger logic for test beams



Courtesy of V.Vagelli and M.Duranti

Physics

Test beam for space experiments (DAMPE, HERD) Trigger logic for test beams

We are trying to explore the possibility of using BondMachine to handle efficiently this kind of operations.

93/116



The BondMachine could be used in several types of real world applications, some of them being:

IoT and CyberPhysical systems.

Computer Science educational applications.

Computing Accelerator

Our effort is now in enabling the possibility of building computing accelerators to be used from within standard (Linux) applications.



The BondMachine could be used in several types of real world applications, some of them being:

IoT and CyberPhysical systems.

Computer Science educational applications.

Computing Accelerator

Our effort is now in enabling the possibility of building computing accelerators to be used from within standard (Linux) applications.



The BondMachine could be used in several types of real world applications, some of them being:

IoT and CyberPhysical systems.

Computer Science educational applications.

Computing Accelerator

Our effort is now in enabling the possibility of building computing accelerators to be used from within standard (Linux) applications.



The BondMachine could be used in several types of real world applications, some of them being:

IoT and CyberPhysical systems.

Computer Science educational applications.

Computing Accelerator

Our effort is now in enabling the possibility of building computing accelerators to be used from within standard (Linux) applications.



Accelerators

Real world applications Accelerators

A BM may be used as an hardware accelerator so that one can mix all together CPU and BM threads, that is one can off-load a task or a function using the BM (i.e. the FPGA)

The resulting accelerator would the advantage of being better suited to the specific problem than generic accelerators (GPU)



Accelerators

The BondMachine can be used (emulated or on FPGA) as a single stand-alone computing device.

It can be used spawned on multiple devices (emulated, on FPGA or both) $% \left({{\rm both}} \right)$

It can be used as computing accelerator

BondMachine can be created and used from within standard (Linux) applications.



The BondMachine can be used (emulated or on FPGA) as a single stand-alone computing device.

It can be used spawned on multiple devices (emulated, on FPGA or both)

97/116



Accelerators

The BondMachine can be used (emulated or on FPGA) as a single stand-alone computing device.

It can be used spawned on multiple devices (emulated, on FPGA or both) $\,$

It can be used as computing accelerator

BondMachine can be created and used from within standard (Linux) applications.

97/116



We are currently working to enable the use the BM as accelerator in two directions:

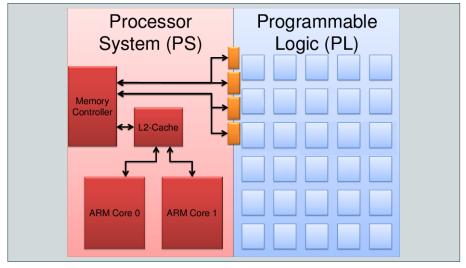
98/116

Using standard processor/FPGA hybrid chips
 Zynq, Cyclone V

Using PCI-express FPGA evaluation boards
 Kintek 7 Evaluation board



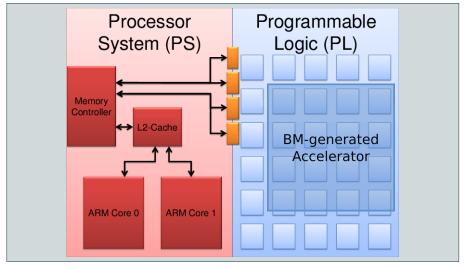
Accelerators



Mirko Mariotti

The BondMachine Toolkit

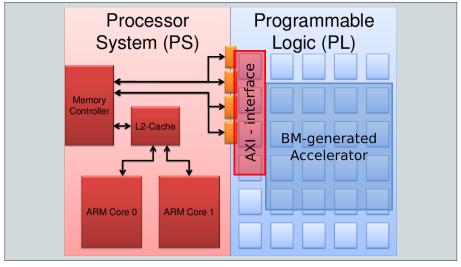
Accelerators _{Hybrid chips}



Mirko Mariotti

The BondMachine Toolkit

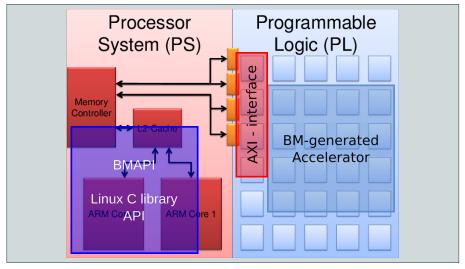
Accelerators _{Hybrid chips}



Mirko Mariotti

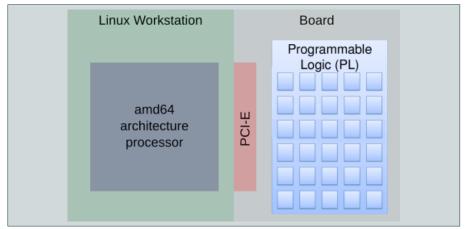
The BondMachine Toolkit

Accelerators _{Hybrid chips}

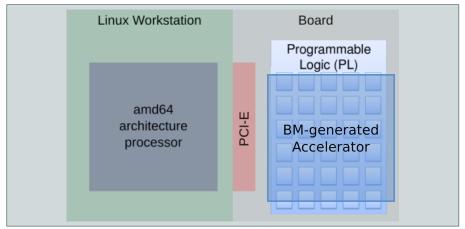


Mirko Mariotti

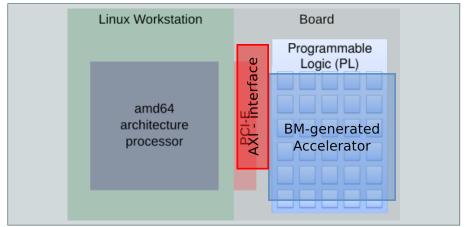
The BondMachine Toolkit





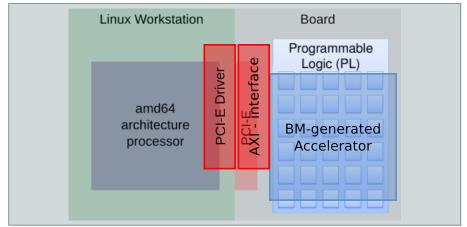




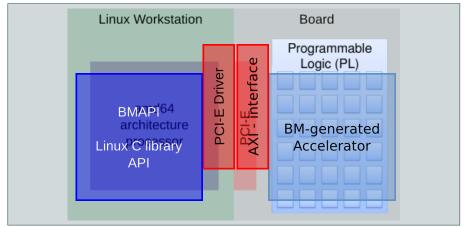




FPGA workshop - ICTP - Trieste 13-24 May 100/116









FPGA workshop - ICTP - Trieste 13-24 May 100/116

Accelerators

Hardware

Digilent Zedboard



Zynq-7000 SoC XC7Z020 512 MB DDR3 Up to 667 MHz Hybrid chips Xilinx ZC702

and the second

Zynq-7000 SoC XC7Z020 1GB DDR3 85k cells - 220 DSP slices

Terasic DE10Nano



Intel Cyclone V 1GB DDR3 SDRAM 110K LEs

PCI-Express board Xilinx KC705



Kintex-7 FPGAs 1GB DDR3 SODIM 326k cells - 840 DSP slices

Mirko Mariotti

The BondMachine Toolkit

FPGA accelerators can be used in the cloud:

- Several public cloud providers offers solution of VM connected to FPGAs (Amazon, Nimbix)
- FPGAs can be inserted in private clouds infrastructures

102/116



FPGA accelerators can be used in the cloud:

- Several public cloud providers offers solution of VM connected to FPGAs (Amazon, Nimbix)
- FPGAs can be inserted in private clouds infrastructures

To be used a firmware has to be uploaded to the accelerated VM FPGA

102/116



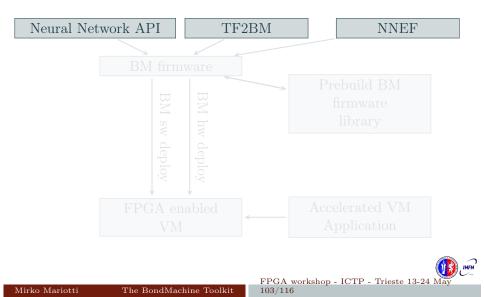
FPGA accelerators can be used in the cloud:

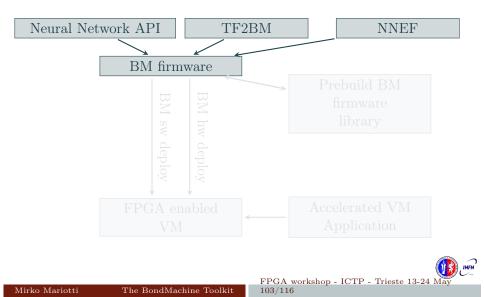
- Several public cloud providers offers solution of VM connected to FPGAs (Amazon, Nimbix)
- FPGAs can be inserted in private clouds infrastructures

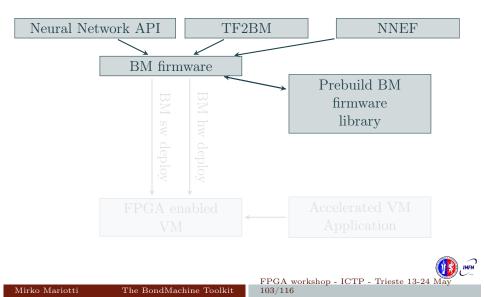
To be used a firmware has to be uploaded to the accelerated VM FPGA

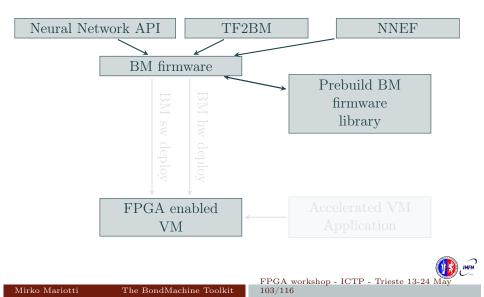
The BondMachine toolkit can be used to build such firmware

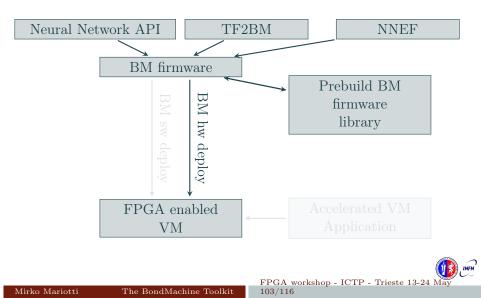


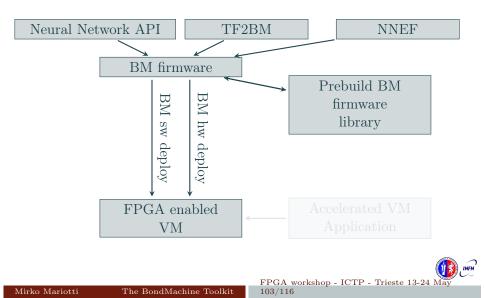




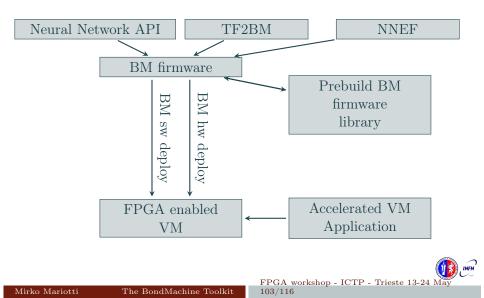








Accelerated ML in the Cloud



FPGA workshop - ICTP - Trieste 13-24 May104/116

Started in May 2015 as a Verilog "garage" experiment, with the idea of creating a processor on an FPGA, so completely bottom-up.

A prototype in every aspects.



Mirko Mariotti

FPGA workshop - ICTP - Trieste 13-24 May 105/116

- May 2016 Idea presented at INFN-computing and Networking Commission Workshop 2016.
- September 2016 The first prototype is built.
- October 2016 It is Selected and the prototype is presented at "Makerfaire 2016 Rome (The European edition)".
- November 2016 Presented at "Umbria Business Match 2016".
- March 2017 First tests for Physics applications.
- November 2017 Presented at "Umbria Business Match 2017".

106/116

December 2107 - Submitted at InnovateFPGA 2018



- May 2016 Idea presented at INFN-computing and Networking Commission Workshop 2016.
- September 2016 The first prototype is built.
- October 2016 It is Selected and the prototype is presented at "Makerfaire 2016 Rome (The European edition)".
- November 2016 Presented at "Umbria Business Match 2016".
- March 2017 First tests for Physics applications.
- November 2017 Presented at "Umbria Business Match 2017".

December 2107 - Submitted at InnovateFPGA 2018



- May 2016 Idea presented at INFN-computing and Networking Commission Workshop 2016.
- September 2016 The first prototype is built.
- October 2016 It is Selected and the prototype is presented at "Makerfaire 2016 Rome (The European edition)".
- November 2016 Presented at "Umbria Business Match 2016".
- March 2017 First tests for Physics applications.
- November 2017 Presented at "Umbria Business Match 2017".

106/116

December 2107 - Submitted at InnovateFPGA 2018



Feb 2018 - Reached the EMEA Semifinal.

Jun 2018 - Reached the EMEA Regional final.

Jul 2018 - EMEA Silver Award, Reached the Grand Final.

Aug 2018 - Presented at Intel Campus, Santa Jose (CA).

Aug 2018 - Won the Iron Award in the Grand Final.



Feb 2018 - Reached the EMEA Semifinal.

Jun 2018 - Reached the EMEA Regional final.

- Jul 2018 EMEA Silver Award, Reached the Grand Final.
- Aug 2018 Presented at Intel Campus, Santa Jose (CA) .
- Aug 2018 Won the Iron Award in the Grand Final.



FPGA workshop - ICTP - Trieste 13-24 May 108/116

Feb 2018 - Reached the EMEA Semifinal.
Jun 2018 - Reached the EMEA Regional final.
Jul 2018 - EMEA Silver Award, Reached the Grand Final.
Aug 2018 - Presented at Intel Campus, Santa Jose (CA) .
Aug 2018 - Won the Iron Award in the Grand Final.



FPGA workshop - ICTP - Trieste 13-24 May 109/116



Feb 2018 - Reached the EMEA Semifinal.
Jun 2018 - Reached the EMEA Regional final.
Jul 2018 - EMEA Silver Award, Reached the Grand Final.
Aug 2018 - Presented at Intel Campus, Santa Jose (CA) .
Aug 2018 - Won the Iron Award in the Grand Final.



FPGA workshop - ICTP - Trieste 13-24 May 110/116



Aug 2018 - Won the Iron Award in the Grand Final.



Mirko Mariotti



The BondMachine Toolkit





Conclusions

FPGA workshop - ICTP - Trieste 13-24 May112/116

The BondMachine is a new kind of computing device made possible in practice only by the emerging of new re-programmable hardware technologies such as FPGA.

The result of this process is the construction of a computer architecture that is not anymore a static constraint where computing occurs but its creation becomes a part of the computing process, gaining computing power and flexibility.

Over this abstraction is it possible to create a full computing Ecosystem, ranging from small interconnected IoT devices to Machine Learning accelerators.



- Improve the use of BondMachines as accelerators, integrating them into the ecosystem
- Start making benchmarks (the real missing piece)
- Find a way to sustain the project
 - Move all the code to github
 - Integrate low and trans-precision instructions (Architectures and Algorithms for Energy-Efficient IoT and HPC Applications, Perugia (Italy), September 3rd to 6th, 2019)



- Improve the use of BondMachines as accelerators, integrating them into the ecosystem
- Start making benchmarks (the real missing piece)
- Find a way to sustain the project
 - Move all the code to github
 - Integrate low and trans-precision instructions (Architectures and Algorithms for Energy-Efficient IoT and HPC Applications, Perugia (Italy), September 3rd to 6th, 2019)



- Improve the use of BondMachines as accelerators, integrating them into the ecosystem
- Start making benchmarks (the real missing piece)
- Find a way to sustain the project
 - Move all the code to github
 - Integrate low and trans-precision instructions (Architectures and Algorithms for Energy-Efficient IoT and HPC Applications, Perugia (Italy), September 3rd to 6th, 2019)



- Improve the use of BondMachines as accelerators, integrating them into the ecosystem
- Start making benchmarks (the real missing piece)
- Find a way to sustain the project
- Move all the code to github
 - I Integrate low and trans-precision instructions (Architectures and Algorithms for Energy-Efficient IoT and HPC Applications, Perugia (Italy), September 3rd to 6th, 2019)



- Improve the use of BondMachines as accelerators, integrating them into the ecosystem
- Start making benchmarks (the real missing piece)
- Find a way to sustain the project
- Move all the code to github
- Integrate low and trans-precision instructions (Architectures and Algorithms for Energy-Efficient IoT and HPC Applications, Perugia (Italy), September 3rd to 6th, 2019)



The project is at the stage of a working prototype, so work has to be done in several areas:

- Include new processor shared objects and currently unsupported opcodes.
- Extend the compiler to include more data structures.
- Improve the networking including new interconnection firmwares.

What would an OS for BondMachines look like ?



The project is at the stage of a working prototype, so work has to be done in several areas:

■ Include new processor shared objects and currently unsupported opcodes.

Extend the compiler to include more data structures.

Improve the networking including new interconnection firmwares.

What would an OS for BondMachines look like?



The project is at the stage of a working prototype, so work has to be done in several areas:

- Include new processor shared objects and currently unsupported opcodes.
- Extend the compiler to include more data structures.
- Improve the networking including new interconnection firmwares.

What would an OS for BondMachines look like ?



The project is at the stage of a working prototype, so work has to be done in several areas:

- Include new processor shared objects and currently unsupported opcodes.
- Extend the compiler to include more data structures.
- Improve the networking including new interconnection firmwares.

What would an OS for BondMachines look like ?



The project is at the stage of a working prototype, so work has to be done in several areas:

- Include new processor shared objects and currently unsupported opcodes.
- Extend the compiler to include more data structures.
- Improve the networking including new interconnection firmwares.

What would an OS for BondMachines look like ?





If you have question/curiosity on the project:

Mirko Mariotti mirko.mariotti@unipg.it http://bondmachine.fisica.unipg.it

FPGA workshop - ICTP - Trieste 13-24 May 116/116