

# Una metodologia di progettazione per hardware riconfigurabile

Enrico Giordano

## 1 Introduzione

Il flusso di progettazione standard di un sistema embedded si divide in tre step: analisi, progettazione e sintesi; questo flusso mira ad avere un sistema specifico per una situazione particolare e alla sintesi finale di tale sistema il meno generale possibile, in modo da renderlo massimamente performante per quel particolare ambiente.

Con questa tesi, si vuole proporre un approccio alternativo il cui mezzo è l'utilizzo di hardware metamorfico e quindi di una particolare tecnica di progettazione basata sulla "configurabilità parziale" della tecnologia delle FPGA. Usando questo mezzo, si vuole progettare un sistema che presenta aspetti embedded (quindi specifici per il target finale) e aspetti più generici e, calato in un contesto particolare, sarà in grado di modificarsi in base alle esigenze ed essere più performante per il contesto in cui viene utilizzato, risparmiando energia, computazioni e memoria. Basandosi sul concetto di "evoluzione", secondo cui un oggetto si modifica in base all'ambiente, si otterrà per ogni ambiente un particolare sistema embedded.

Si presenteranno quindi, a livello descrittivo, le diverse tecniche di configurabilità parziale, introducendo poi un modello di progettazione che mira ad un approccio standardizzato per i futuri sviluppi di questa idea, per poi provare il tutto su un caso di studio reale che possa dimostrare i benefici di questa tecnologia e tecnica di progettazione.

## 2 Tecnologia attuale del silicio

Attualmente esistono 3 tecnologie differenti riguardo l'utilizzo del silicio per scopi computazionali: Hardware General Purpose, Hardware Embedded e Hardware Reprogrammable. Queste 3 tipologie differiscono sia per la composizione, sia per la progettazione ma anche per il settore di utilizzo.

### 2.1 Hardware General Purpose

Questo tipo di Hardware lo si può trovare in processori e microprocessori utilizzati in strumenti di calcolo generici, come PC, moderni telefoni e televisori. La caratteristica peculiare di questi è la generalità: non essendo stati progettati per un sistema specifico, devono permettere di eseguire più operazioni possibili ed essere riprogrammati a piacimento, al limite delle loro potenzialità.

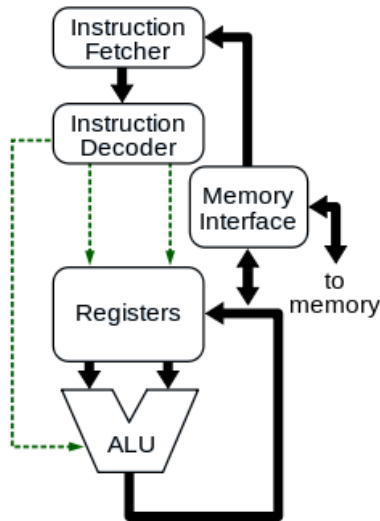


Figura 1: Architettura di una CPU

Questi sono composti da un'architettura standard, rappresentata nella figura 1, e viene implementata in ogni CPU, ovvero ciò che implementa questa tecnologia.

Il ciclo di vita di questo hardware è contraddistinto da 3 fasi: fetch, decode, exec; ogni fase è contraddistinta da un insieme di componenti interni utilizzati. La fase di “fetch” consiste nel caricare l'istruzione corrente per essere eseguita; la fase di “decode” fa in modo che tutti i valori siano disponibili per il calcolo e che l'istruzione possa essere eseguita correttamente impostando i registri interni; la fase di “exec” permette, tramite i componenti di memoria e di calcolo (registri e ALU), di eseguire l'istruzione.

Questo tipo di hardware è molto utile per implementare sistemi generici: la filosofia alla base di questa tecnologia è l'utilizzo più generico possibile, mettendo in secondo piano l'ottimizzazione e il consumo, oltre al costo che cerca di stare relativamente nelle politiche di mercato.

## 2.2 Hardware Embedded

Questo tipo di Hardware è contraddistinto dalla specificità dei suoi componenti interni in base al suo utilizzo; infatti si cerca di utilizzare questo hardware in ambienti o condizioni specifiche, in cui sono richieste operazioni particolari ma soprattutto è necessario ridurre i costi e il consumo dell'applicazione finale. La tecnologia che rappresenta questo tipo di hardware è una ASIC e difficilmente si trova un'architettura standard, proprio perché l'architettura è specifica per un certo comportamento. Esistono in commercio dei sistemi embedded basati su microprocessori RISC, come ad esempio i Cortex-M della ARM, che hanno un'architettura molto simile a quella delle CPU. In generale comunque ogni ASIC è diversa, in quanto viene mappato su silicio uno specifico algoritmo o un'architettura più complessa.

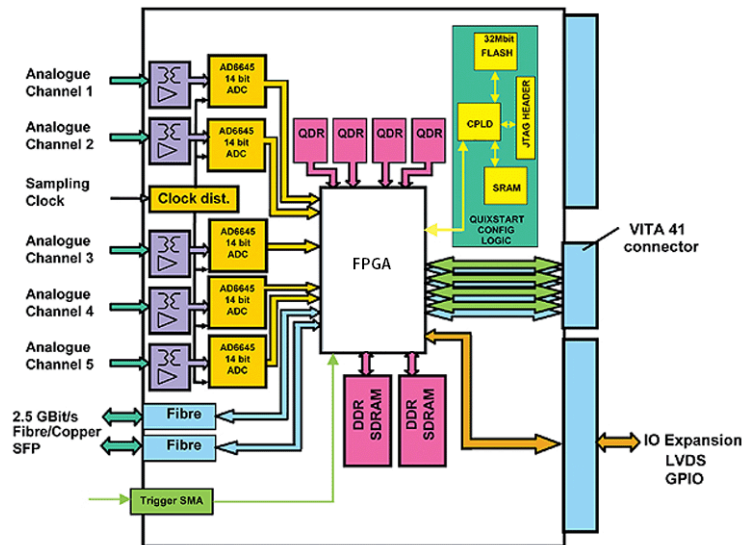


Figura 2: Architettura di una FPGA

Un' ASIC deve costare il meno possibile e deve consumare il meno possibile, in base alle esigenze richieste; ovviamente potrà eseguire solo il compito per cui è stata progettata; ha quindi una filosofia di progettazione opposta a quella dell'Hardware General Purpose.

### 2.3 Hardware Riprogrammabile

L'ultimo tipo di Hardware è quello riprogrammabile, ovvero permette di essere configurato a livello di celle di memoria più volte per implementare diversi comportamenti. Di solito viene utilizzato sia in ambito embedded, per permettere più libertà di progettazione e più efficienza, ma anche a scopo sperimentale, perchè è utile per testare un sistema in fase di sviluppo.

Un esempio di architettura (dipende molto dal modello che si vuole utilizzare) lo si può osservare nella figura 2: la parte riconfigurabile, denotata con FPGA, consiste in un blocco di celle riconfigurabili, che, in base alla descrizione hardware caricata nella memoria flash, possono cambiare struttura. Attorno alla FPGA sono presenti vari moduli standard che possono servire per il consueto utilizzo in ambito di controlli, ma non sempre sono necessarie, in quanto si possono descrivere all'interno della parte riconfigurabile. Parte necessaria per il funzionamento corretto di una FPGA è la logica di controllo della parte riprogrammabile: deve esserci un hardware dedicato per rimappare questa parte, in modo che il processo di riconfigurazione avvenga più velocemente possibile; questo si occuperà quindi di riconfigurare, in base al bitstream che rappresenta l'hardware, la parte riconfigurabile.

In base al suo utilizzo può assumere aspetti di un hardware general purpose e di un hardware embedded; essendo una tecnologia piuttosto costosa si cerca di utilizzarla il

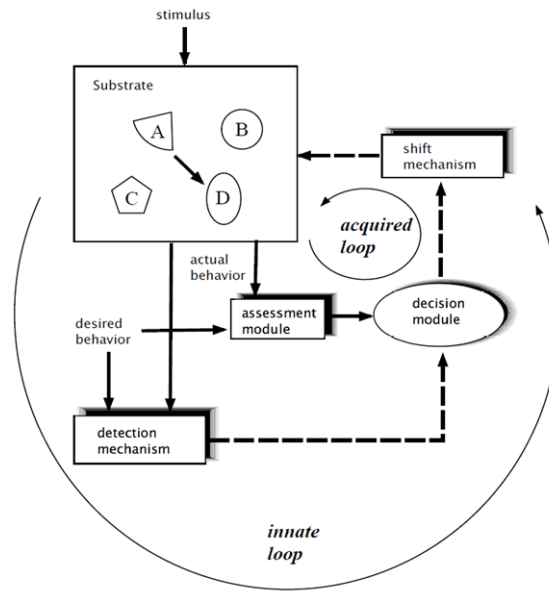


Figura 3: Ciclo evolutivo di un sistema metamorfico

meno possibile, o almeno solo in fase di test.

### 3 Il concetto di Hardware Metamorfico

Il concetto di “Hardware Metamorfico” nasce nei primi anni del 1990, per poi perdere parzialmente interesse in quanto veniva presentato come idea, non come tecnologia pratica. Con questo concetto si intende un tipo di Hardware capace di cambiare il proprio comportamento (“evolvere”) utilizzando un “metodo” di riconfigurazione: i termini “evolvere” e “metodo” non sono mai stati definiti in maniera standard, l’articolo che cerca di porre chiarezza su tutto questo ambito [1] focalizza l’attenzione sulle diverse sfaccettature del termine “evoluzione”, parlando di evoluzione “intrinseca”, “estrinseca” e “mixata”, rimandando la definizione ad un articolo scientifico precedente [2].

Questo tipo di Hardware è tipico del mondo naturale, in quanto tutti gli oggetti e gli esseri viventi sono in grado di modificarsi in base all’ambiente che li circonda.

Il ciclo di vita di un Hardware Metamorfico può essere rappresentato come nella figura 3: il superciclo che compone questo sistema, chiamato “ciclo innato”, ha un sotto-ciclo di acquisizione di stimoli, chiamato “ciclo di acquisizione”, che, tramite il sottostrato di algoritmi di esecuzione, cambia il funzionamento del sistema in base al “comportamento attuale” e il “comportamento desiderato”; con questi dati, tramite un modulo di valutazione degli stimoli e un meccanismo di investigazione (per valutare il comportamento desiderato rispetto al comportamento attuale), viene generato dal modulo di “decisione” il risultato che fa modificare l’intero sistema. Il ciclo innato successivamente fa rieseguire

questo procedimento, in modo da permettere, potenzialmente all'infinito, la metamorfosi del sistema.

## 4 Hardware riconfigurabile

Per rendere utile l'hardware metamorfico, è necessario trovare una tecnologia che lo rappresenti, in modo da rendere concreto l'utilizzo di tale idea. Per renderla concreta, è necessaria la presenza di un oggetto che possa riprodurre sia il comportamento di silicio, sia il comportamento evolutivo dell'hardware metamorfico. Nel 1985, la Xilinx aveva creato le prime FPGA, rilasciando in commercio il primo modello "XC2064", utilizzabile perlopiù per la prototipizzazione di circuiti programmabili. Questa tecnologia sembra prestarsi coerentemente con l'idea di Hardware metamorfico, in quanto è in grado di poter cambiare il proprio comportamento a livello hardware cambiando la propria struttura interna. Il suo carattere riconfigurabile però è stato progettato e utilizzato per la prototipizzazione, non per l'effettivo mutamento di comportamento durante il suo ciclo di vita.

Nei primi anni '90 nasce una tecnica di progettazione chiamata "Riconfigurabilità", da qui poi la tecnologia di applicazione, appunto le FPGA. Questa tecnica consiste nell'utilizzare Hardware con tecnologia riprogrammabile che, tramite la memoria messa a disposizione dentro il sistema, riesce a modificare il suo comportamento. Non è un comportamento appreso completamente dall'ambiente, ma è un comportamento previsto e quindi implementato che va a sostituire un comportamento esistente; il ciclo di vita quindi non sarà composto da un vero e proprio "apprendimento", ma sarà una modifica prevista dal progettista. Questo può sembrare limitante, in quanto non c'è piena libertà da parte del dispositivo, ma non avrebbe senso far prendere iniziativa a questo (per i principi della computabilità, non sarebbe nemmeno possibile).

Quindi, più precisamente, si può parlare di "riconfigurabilità parziale" e "riconfigurabilità run-time": il primo termine pone enfasi su la parziale capacità del dispositivo di modificarsi, in quanto ci sarà una parte progettata che non deve modificarsi ma soprattutto perché il dispositivo non si modificherà completamente, avrà delle modifiche consentite prestabilite; il secondo termine indica il fatto di potersi modificare durante il suo ciclo di vita, cambiando la propria configurazione di porte logiche ottenendo quindi una forma diversa.

### 4.1 Metodi di riconfigurazione

Essendo la Riconfigurabilità una tecnica di progettazione, prevede vari metodi per essere implementata, che possono essere scelti sia per la tecnologia sia in base alle risorse disponibili su essa. Questi metodi sono stati proposti dall'Altera [3] e dall'insegnante Dirk Koch, dell'Università di Manchester [4].

La prima tecnica è la "Riconfigurazione parziale dinamica", che consiste nel cambiare una parte della FPGA mentre il rimanente circuito continua il suo ciclo di vita. Per fare ciò, esiste una partizione riconfigurabile e in questa si carica il bitstream che corrisponde

al nuovo circuito da una memoria esterna. Per far comunicare la parte statica con quella dinamica, è posto un bridge fisico tra di esse.

Il bridge fisico consiste in un bus, solitamente a 32 bit, che permette lo spostamento di file sintetizzati che rappresentano la nuova configurazione hardware. Questi file si trovano o nella memoria esterna o nella memoria interna, quindi vengono caricati runtime in base a condizioni scelte dal progettista. Solitamente il bus è di tecnologia DMA in modo da essere il più veloce possibile e il trasferimento avviene tramite diversi protocolli gestiti a livello hardware (solitamente il protocollo PCAP a 32 bit, ma dipende dall'architettura ed è trasparente al progettista).

Questa tecnica è molto efficace quando si ha a disposizione poca memoria per memorizzare il bitstream che rappresenta l'hardware e può essere utilizzata in quelle architetture che non dispongono di sufficiente memoria per caricare al completo tutto il sistema descritto.

La seconda tecnica è la "Rilocazione parziale di bitstream", che consiste nello "spostare" i moduli compilati che descrivono l'hardware (che rappresentano la parte riconfigurabile) da un'area della FPGA all'altra con le stesse dimensioni e proprietà. Questa tecnica è più limitante della precedente, un quanto si pone il vincolo di avere due moduli con stesse dimensioni e proprietà, però può essere utile per questioni di ottimizzazione di codice, in quanto i due moduli possono rappresentare la stessa unità ma al loro interno eseguono lo stesso algoritmo in maniera differente.

## 4.2 Readback

Dal modello "Virtex 6" delle FPGA Xilinx, è possibile leggere il contenuto della FPGA e quindi ottenere la configurazione attuale del sistema. Esistono due tipi di lettura: la "readback verify", che consiste nella lettura di tutte le celle di memoria, e la "readback capture", che permette di leggere le celle di memoria e lo stato dei registri, eseguendo un effettivo dump totale. Questa informazione è di particolare interesse per questo progetto, in quanto permette di ottenere la configurazione ottenuta dopo la metamorfosi, quindi ottenere un nuovo codice. Questa proprietà è stata in parte resa negativa dal punto di vista della sicurezza informatica, in quanto permette di estrapolare codice per poi potenzialmente farne un cattivo uso, però è molto utile per sapere l'andamento della metamorfosi.

## 4.3 Metodi di reallocazione

La parte cruciale quindi di queste tecniche è proprio la reallocazione dei moduli, in quanto la riconfigurabilità consiste proprio nello "spostare" un modulo da una sezione attiva a una non attiva. Esistono due modalità di reallocazione: usando dei tools offerti dalla Xilinx e usando il bus.

I tools della Xilinx sono dei software che fungono da middleware per gestire lo spostamento da un'area all'altra e vengono dati in dotazione in base al sistema acquistato. Le direttive del bus sono delle macro da utilizzare durante l'esecuzione del codice per istruire il bus prima di far comunicare i diversi moduli e collegarli. Entrambe le modalità

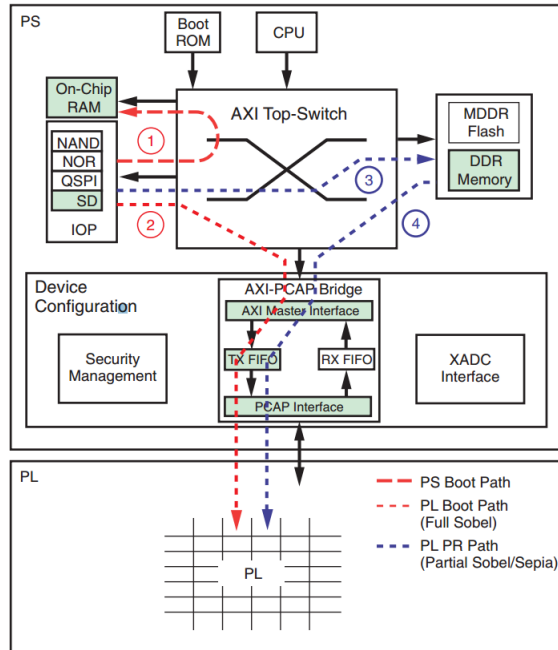


Figura 4: Flusso di riconfigurazione

creano un “ponte” tra la zona statica e la zona dinamica della logica programmabile, in modo da interconnetterle per generare una nuova descrizione hardware.

Entrambi i metodi si basano sulla riconfigurazione runtime della parte di logica programmabile tramite lo stesso hardware della FPGA che si occupa della riconfigurazione statica (quindi durante la progettazione da parte del progettista), ma in maniera automatica: il modulo di controllo di riconfigurazione sceglie quale può essere la zona di logica programmabile corretta (la prima che ha abbastanza spazio per contenere la nuova logica) e genera il segnale di riconfigurazione; al momento del segnale, si esegue la procedura di riconfigurazione come da prassi ma il bitstream da caricare non deriva da un programma esterno via JTAG bensì dalla memoria interna. Quindi ciò che viene caricato non viene generato runtime, è già presente all’interno della memoria, ma il comportamento globale del sistema cambia in quanto viene caricata una porzione di bitstream che andrà ad alterare il circuito generato dal bitstream statico (che non cambia mai) e quello dinamico (caricato runtime).

#### 4.4 Aree di reallocazione

Il processo di reallocazione visto finora è riassumibile con la figura 4, in cui si può vedere come la parte di “processing system” (PS) configuri l’informazione da inviare alla parte di “programmable logic” (PL) tramite il bridge; caricato il file che rappresenta la nuova parte sintetizzata di programma dalla memoria, si invia alla parte da configurare PL utilizzando il protocollo del bridge. Le aree precedentemente descritte quindi si

trovano all'interno di PL, dove risiede l'effettiva parte logica. Queste aree normalmente possiedono una frammentazione interna ed esterna, in quando il codice non sempre occupa tutti i settori; proprio per questo si può deframmentare e ottenere dei settori vuoti in cui inserire runtime il codice.

Spostare un nuovo settore dentro la parte di logica programmata significa quindi configurare hardware precedentemente inattivo, dando quindi una nuova configurazione hardware e quindi un diverso comportamento del sistema. In base alla grandezza delle aree da spostare, possono coesistere all'interno dell'area PL, ma non avrebbe senso perché potrebbe significare avere hardware non utilizzato e configurato, quindi si spreca silicio.

## 5 La progettazione

Poiché non esiste un metodo standard per progettare, sviluppare o semplicemente simulare questi sistemi, si vuole creare un nuovo flusso di progettazione per gestire al meglio la fase di sviluppo. Poiché questi sistemi fanno parte di un sottoinsieme dei Sistemi Embedded generici, si utilizzerà ciò che è noto durante la fase di sviluppo di un normale Sistema Embedded per poi aggiungere ciò che serve per specializzare tale sistema in uno Metamorfico.

Il punto di partenza della progettazione non dovrà essere l'analisi dell'ambiente in cui calare il sistema da progettare, bensì sarà scindere cosa deve essere statico e cosa deve essere dinamico: una business logic, ad esempio, può essere una parte statica, mentre un insieme di range su cui fare misure può essere una parte dinamica perché dipende dall'ambiente. Una volta scelte le due parti, si procede con la progettazione come da prassi di un sistema embedded, progettando contemporaneamente queste parti: in questo modo, durante la progettazione, si può capire se le scelte adottate sono state corrette. Infine si configura l'hardware in modo da prepararlo per la sua futura metamorfosi, impostando la parte riconfigurabile, per poi osservare come cambia durante l'acquisizione delle informazioni.

Lo scopo di questa tesi sarà comunque capire come sviluppare al meglio la progettazione di questo tipo di sistema, comprendendo il grado di libertà che si può raggiungere per la parte dinamica e capire quale tipo di riconfigurabilità è più conveniente usare, oltre al fatto di essere sicuri che sia conveniente sfruttare la metamorfosi per risparmiare tempo e denaro durante la progettazione di un sistema embedded.

## 6 Simulazione e sintesi di sistemi embedded

Esistono vari metodi per progettare un sistema embedded; l'approccio ingegneristico più avanzato vuole che si progetti in maniera astratta il sistema, per poi calarlo in un contesto di descrizione tramite HDL per poi eseguire la sintesi.



## 6.1 Simulazione

Per la descrizione ad alto livello e simulazione, solitamente si utilizza SystemC a diversi livelli, passando da una descrizione mirata a osservare il comportamento dei diversi moduli durante la loro interazione (TLM), fino alla descrizione a livello di EFSM cycle accurate o analogica (RTL e AMS). I vantaggi dell'utilizzo di SystemC sono molteplici, soprattutto grazie allo scheduler interno si riesce a simulare correttamente i vari comportamenti a livello di segnali e attivazione dei moduli.

## 6.2 Scheduler di SystemC

Lo scheduler di SystemC è stato progettato in modo da rispettare una gerarchia di wakeup dei moduli, in quanto l'hardware solitamente è pilotato sia da un clock interno sia da segnali provenienti da diversi moduli, quindi è stato necessario modellare questo parallelismo tramite uno scheduler. Lo schedule eseguito dallo scheduler è diviso in passi:

1. Elaborazione: vengono create le strutture dati e viene fatto il binding delle porte;
2. Inizializzazione: vengono inizializzate porte e segnali e vengono fatti eseguire tutti i processi fino alla loro terminazione o al primo wait, in modo da creare una coda d'esecuzione iniziale;
3. Evaluation: viene fatto eseguire il primo processo pronto in coda d'esecuzione fino alla sua terminazione o ad un wait. Se vengono notificati eventi (con delle notify()), i processi interessati vengono messi in coda d'esecuzione. Se ci sono altri processi in coda pronti ad essere eseguiti ripeto le operazioni precedenti, altrimenti proseguo;
4. Update: vengono aggiornate porte e segnali ed aggiunti in coda d'esecuzione i processi sensibili ad essi;
5. Delta: vengono aggiunti alla coda d'esecuzione tutti i processi sensibili ad eventuali delta notification (notify(SC\_ZERO\_TIME)) Se ci sono altri processi in coda pronti ad essere eseguiti ripeto le operazioni precedenti a partire dalla fase di evaluation, altrimenti proseguo;
6. Timed: Vengono aggiunti alla coda d'esecuzione tutti i processi sensibili alla prima timed notification (notify(SC\_TIME(n,SC\_NS))). Aggiorno il tempo di simulazione (clock). Se ci sono altri processi in coda pronti ad essere eseguiti ripeto le operazioni precedenti a partire dalla fase di evaluation, altrimenti termino la simulazione.

Grazie a questo scheduler, si ha uno strumento che simula correttamente un sistema embedded, rispettando anche i tempi accuratamente di esecuzione, osservabili tramite timestamp.

### 6.3 Sintesi

Una volta preparato il modello del sistema in SystemC e simulato correttamente, si procede con la descrizione hardware in HDL, simulazione ed infine sintesi. Solitamente si utilizza VHDL e Verilog, due linguaggi di descrizione hardware che permettono sia di simulare sia di sintetizzare. Solitamente se viene descritto correttamente il sistema in SystemC, il passaggio ad una descrizione HDL è quasi automatico, in quanto si rappresentano gli stessi concetti semanticamente; infatti questo processo dovrebbe essere una traduzione sintattica tra linguaggi.

## 7 Simulazione e sintesi di sistemi metamorfici

Poichè non esiste un metodo standard per progettare sistemi metamorfici, bisogna definire *ex novo* un metodo efficace. Intuitivamente sarebbe consigliato utilizzare la stessa metodologia di progettazione di sistemi embedded standard, in modo da rendere facile l'apprendimento a chi è già abituato a progettare sistemi embedded consuetudinari, oltre a sfruttare le potenzialità di un metodo già assodato e verificato nel tempo.

### 7.1 Simulazione con SystemC ReChannel

Per la simulazione si vorrebbe utilizzare SystemC, in quanto offre una descrizione fedele dell'hardware da simulare e sfrutta la conoscenza di C++ per descrivere i sistemi. Attualmente esiste una libreria di SystemC per la descrizione di sistemi riconfigurabili chiamata ReChannel.

ReChannel è una libreria sviluppata dall'Università di Bonn dal Prof. Armin Felke nel 2008, ceduta a GreenSocs con licenza open-source l'anno successivo e attualmente non più mantenuta, a seguito degli aggiornamenti di SystemC 2.0 che non permettevano la compilazione. La libreria è stata aggiornata e riscritta durante la stesura di questa tesi ed ora è compilabile ed utilizzabile.

La libreria si basa sul concetto di **switch**, ovvero l'elemento cardine di un sistema riconfigurabile: deve esserci la possibilità di cambiare un modulo con un altro, possibilmente di stessa dimensione e con la stessa quantità di input e output. Lo switch, in questa libreria, viene rappresentato come una sorta di multiplexer, che tiene collegati i diversi moduli della parte riconfigurabile con il modulo della parte statica; al momento opportuno, deciso dal progettista, sia dinamicamente sia staticamente in fase di compilazione, si imposta in modo che attivi il canale di comunicazione tra il modulo attivato e la parte statica, mentre disattiva i moduli della parte dinamica che vengono sostituiti. Lo switch permette la comunicazione sia da parte statica a parte dinamica sia viceversa: la prima viene chiamata *portal*, la seconda viene chiamata *exportal*.

Lo switch permette quindi la comunicazione tra i moduli della parte statica e della parte dinamica; la comunicazione può essere semplice (a un solo valore) o multivalore (coda fifo). Questo modella correttamente il concetto di bridge descritto precedentemente, ossia quel componente che permette la comunicazione tra le parti della FPGA.

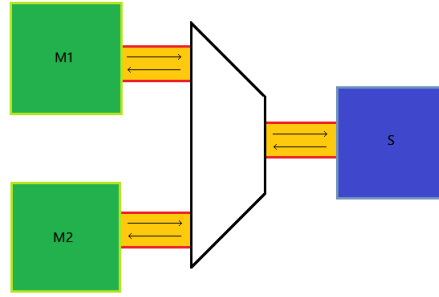


Figura 5: Rappresentazione dello switch

Questo si può vedere come una sorta di bus, che dirige il traffico point to point tra i diversi moduli.

L'utilizzo di questo componente per la comunicazione è un aspetto che si distanzia molto dalla solita progettazione di hardware in SystemC, poiché solitamente si definiscono i moduli con vari input e output e si effettua un collegamento punto-punto all'interno del main; con questa libreria invece si definiscono le modalità di comunicazione tra i moduli di diversa natura, per poi far gestire a questo componente i diversi collegamenti. Questo fatto modella correttamente le FPGA riconfigurabili moderne.

Il problema principale ora, conoscendo lo scheduler di SystemC, è gestire le modalità di schedule dei moduli della parte riconfigurabile; è necessario quindi estendere lo scheduler nativo con delle nuove opzioni.

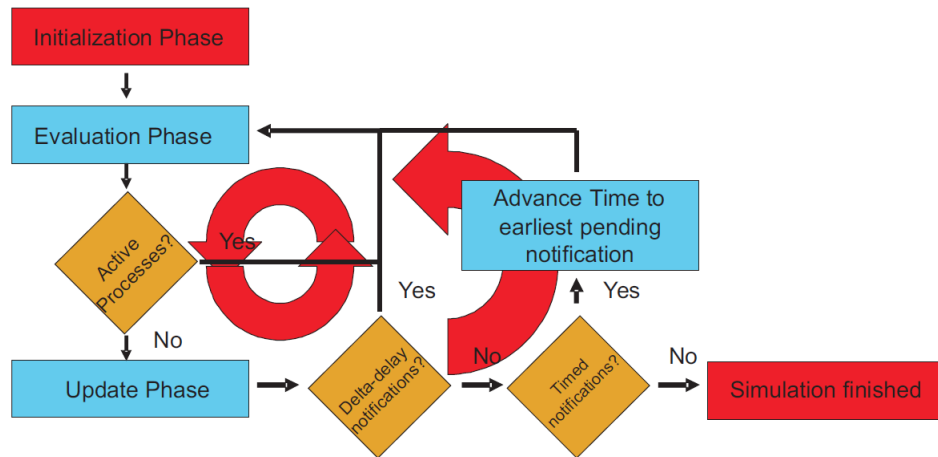


Figura 6: Scheduler esteso con ReChannel

Come si può notare con la figura 6, è stato esteso lo scheduler dando priorità alla riconfigurazione, quindi nella fase di Evaluation si aggiunge un'ulteriore fase, ovvero il controllo sui processi attivi, oltre ad aggiungere passaggi nella fase di inizializzazione.

Durante la fase di inizializzazione, vengono istanziati tutti i moduli riconfigurabili e disattivati; è compito del progettista decidere quale modulo inizialmente deve essere attivato, ma non è necessario che ci siano moduli attivi. Successivamente, un modulo può essere attivato in qualunque momento; lo scheduler lo attiverà appena entrerà nella parte finale della fase di Evaluation, controllando che ci siano processi attivi pendenti (quindi da attivare).

Solitamente si utilizza un modulo statico che si occupa di fare da transattore tra la parte statica e la parte dinamica; solitamente viene chiamato processo “Top” e si occupa sia di descrivere lo switch sia di effettuare la commutazione dei moduli della parte dinamica. Anche in questo caso, viene modellata correttamente l’architettura riconfigurabile, in quando deve esserci un’entità del sistema che si occupa di effettuare la riconfigurazione, che può essere il processore nella board della FPGA o lo stesso progettista se si descrive una riconfigurazione statica.

Un ultimo aspetto molto importante che simula correttamente la realtà è la possibilità di impostare il tempo necessario per la riconfigurazione: ogni modulo, in base alla FPGA target, può metterci tempi differenti per essere attivato e disattivare gli altri, quindi è necessario un meccanismo per modellare il tempo di riconfigurazione. Questa libreria modella il tempo di riconfigurazione in base al caso pessimo di tempistiche, ovvero permette al progettista di impostare il tempo massimo che impiega il modulo ad essere attivato (infatti non sempre impiega lo stesso tempo per essere riconfigurato, per questo bisogna ragionare per caso pessimo). Può essere deciso questo tempo secondo le specifiche della FPGA target proporzionali all’area che il circuito ricopre (si possono effettuare delle stime in base a quante e quali operazioni vengono eseguite).

## 7.2 Sintesi

La sintesi di questi sistemi è molto simile alla sintesi dei normali sistemi embedded, poiché si utilizza HDL compilato e sintetizzato tramite tool automatici.

La traduzione da SystemC Rechannel a VHDL è molto semplice: si rappresentano i diversi moduli come moduli statici, contrassegnando i moduli dinamici con un nome riconoscibile ed incrementale (es: `reconfigurable_module1`, `reconfigurable_module2`, ecc...), ci si assicura che tutto sia stato scritto correttamente e si importano i file nel programma di sintesi prescelto.

Solitamente i programmi di sintesi vengono offerti direttamente dai produttori della FPGA target (ad esempio per la Virtex 4 danno in dotazione il programma) e permettono di impostare facilmente i diversi moduli scritti in HDL tramite interfaccia grafica. Tutti i tool automatici partono quindi dalla descrizione in HDL e, tramite un’ulteriore sforzo del progettista nell’impostare il comportamento dei moduli, acquisiscono la particolarità di essere statici o dinamici.

Come si può notare con la figura 7, il tool automatico identifica i sorgenti come moduli che devono essere sintetizzati; possiamo notare i diversi insiemi di moduli tra cui il modulo “Top”, già visto precedentemente con SystemC ReChannel, che identifica il transattore tra la parte statica e la parte dinamica del sistema. Una volta sintetizzati i sorgenti, si crea una Netlist, ovvero si creano i collegamenti interni ed esterni tra moduli,

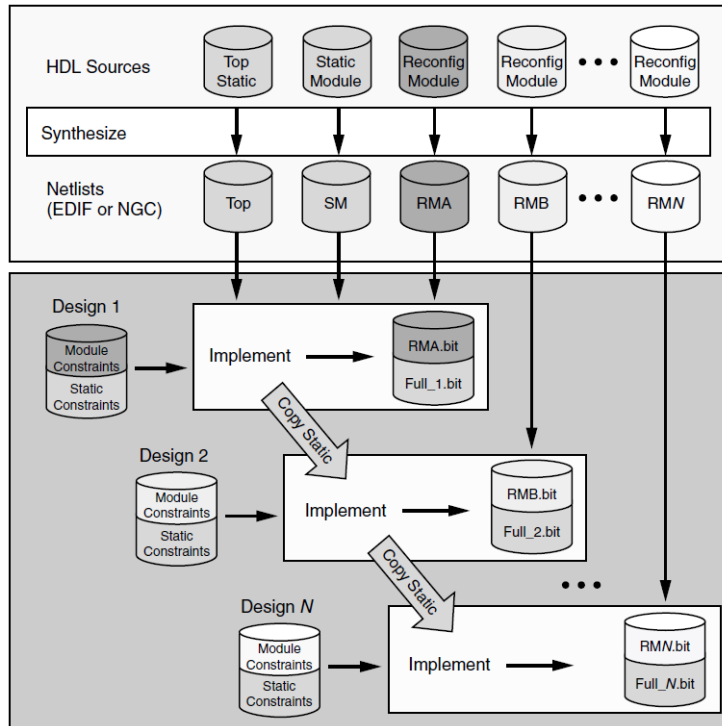


Figura 7: Processo di sintesi eseguito da un tool automatico

in modo da definire il metodo di comunicazione. Fatto ciò, il tool identifica in automatico tutti i differenti design che si creeranno durante il ciclo di vita del sistema, in base alle configurazioni possibili definite nella Netlist; ogni Design è una copia statica presente in memoria, quindi una volta sintetizzata esistono le configurazioni in memoria già pronte per essere nuovamente sintetizzate in hardware (solitamente risiedono nella memoria Flash).

L'ultimo passaggio del tool automatico è la creazione delle partizioni della FPGA, perché deve essere possibile identificare dove caricare la parte riconfigurabile e dove deve rimanere staticamente il codice; viene creato quindi il file delle partizioni, solitamente chiamato *XPARTITION.xml*, che risiede in memoria e viene consultato ogni volta che deve avvenire una riconfigurazione.

## 8 Metodologia di progettazione proposta

Lo scopo di questo documento è di proporre una metodologia efficiente di progettazione e simulazione per questa particolare architettura. Attualmente non esiste un metodo standard per la progettazione di sistemi basati su questa architettura, quindi si è cercato di proporre un metodo più efficiente possibile per garantire anche a chi non conosce approfonditamente questa tecnologia di poterla sfruttare al meglio. Per decidere uno

stile di progettazione, si è deciso di basarsi su uno stile già noto nei sistemi consuetudinari embedded, ovvero la progettazione di macchine a stati finiti. La tecnica delle macchine a stati finiti garantisce chiarezza di progettazione e un flusso chiaro di istruzioni che vengono eseguite durante la “vita” del sistema, evidenziando anche i possibili blocchi ed anomalie che possono esserci durante l’esecuzione. Inoltre ci si è serviti del noto concetto di “macchine a stati finiti non deterministiche”: è possibile rappresentare un sistema a stati che può comportarsi “contemporaneamente” in molteplici modi, quindi avere più comportamenti tramite lo stesso input in ingresso (solitamente evitato durante la progettazione delle macchine a stati finiti, in quanto non è implementabile).

La progettazione proposta quindi si differenzia in diverse fasi:

- *analisi del sistema*, evidenziando cosa potrebbe essere trattato come sistema standard e sistema riconfigurabile; si analizza quindi cosa deve rimanere statico per tutta l’esecuzione di un algoritmo e cosa può cambiare durante l’esecuzione;
- *analisi degli stati*, ovvero si comincia a creare il sistema a stati evidenziando come evolve il sistema dinamico dividendolo in fasi; si deve quindi creare una macchina a stati generale creando il flusso di esecuzione del sistema, ma non rappresentando ciò che succede all’interno degli stati (perché si suppone che per diverse configurazioni ci saranno diversi comportamenti all’interno delle suddette fasi);
- *definizione degli stati*, ovvero si “disegnano” gli stati del sistema riconfigurabile, evidenziando le diverse macchine a stati che vengono prodotte: in questa fase si devono ottenere più macchine a stati duplicate, con lo stesso numero di stati il cui contenuto però cambia tra macchine, perché il flusso di esecuzione sarà lo stesso ma ciò che avviene negli stati deve essere differente;
- *definizione del Top*, il componente che deve guidare lo svolgimento delle attività della parte riconfigurabile; questo deve avere anch’esso la stessa macchina a stati dei moduli riconfigurabili, ma sarà l’unico componente che può decidere in quale stato deve trovarsi il sistema;
- *implementazione e simulazione*, tramite il tool SystemC Rechannel;
- *validazione e sintesi*.

Il sistema finale dovrebbe essere simile alla figura 8: si può notare che il sistema risulta modulare e ben diviso tra parte statica e parte dinamica: questo, oltre a facilitare la comprensibilità del progetto, garantisce anche una buona analisi in fase di simulazione e validazione, in quanto è possibile analizzare nel dettaglio le singole parti del sistema e garantire l’analisi di ogni stato e di ogni modulo. La divisione netta tra parte statica e parte dinamica garantisce anche il parallelismo di esecuzione tra le diverse parti del sistema: potenzialmente la parte statica può eseguire codice in maniera del tutto asincrona rispetto alla parte dinamica, garantendo quindi più potenza di calcolo in termini di velocità. Infine, rappresentare il sistema con macchine a stati finiti permette di effettuare tutte le ottimizzazioni note per diminuire il numero degli stati del sistema,

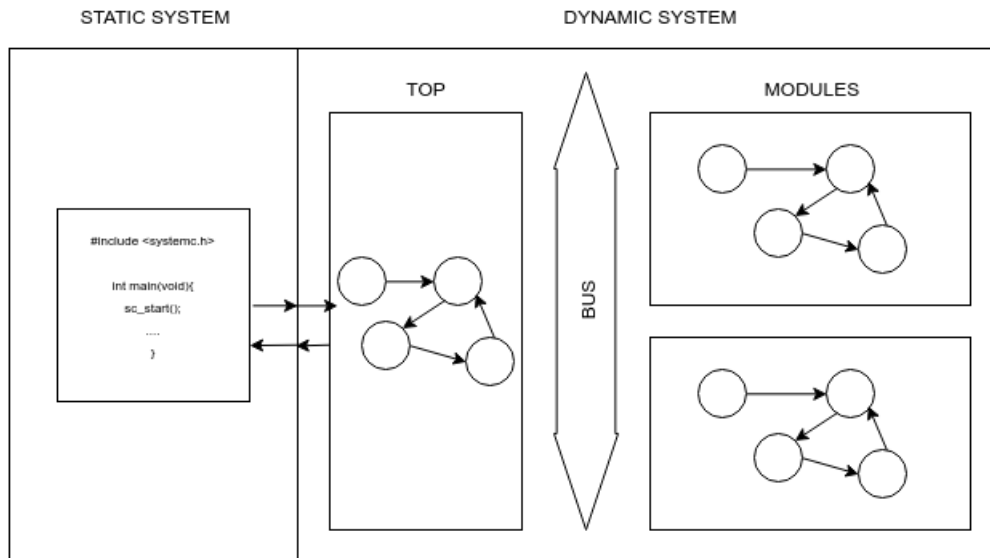


Figura 8: Struttura finale del sistema progettato

guadagnando ancor più spazio. Si vedrà nel dettaglio con il caso di studio il guadagno ottenuto da questa architettura proposta in contesto di vincoli realtime hard.

Se da un lato si hanno degli ottimi benefici, con questa configurazione si ottengono anche degli svantaggi dal punto di vista implementativo: la parte dinamica e la parte statica necessitano di comunicazione tramite il bus, però essendo entrambi asincroni necessitano di meccanismi di sincronizzazione per evitare di perdere dati. Inoltre, visto che questo tipo di sistemi viene calato in contesti critici, essendo sistemi embedded, bisogna essere sicuri di progettare il sistema in modo che non presenti anomalie e quindi perdere l'efficacia dell'utilizzo di questa architettura. È quindi necessario garantire che il sistema sia sempre in stato "safe": come tutti i sistemi asincroni, se sono presenti le 3 condizioni necessarie (mutua esclusione, hold and wait, no preemption, attesa circolare), il sistema può andare in deadlock, quindi è necessario adottare tutte le strategie necessarie per evitare ciò. Fino ad ora la garanzia di stato safe del sistema dal punto di vista di asincronicità era lasciata al progettista, con questo progetto si propone invece un tool automatico per generare il software per SystemC Rechannel adottando questo metodo di progettazione, in modo da lasciare al progettista solo il compito di pensare e progettare il comportamento dei moduli del sistema.

Con questo progetto quindi non si propone esclusivamente una metodologia di progettazione, ma anche un tool automatico per progettare correttamente questi sistemi.

## 9 Tool automatico

TODO

## 10 Caso di studio

Per verificare le potenzialità di questa architettura e del metodo di progettazione proposto, è stato implementato un progetto complesso che comprende molteplici aspetti avanzati dell'applicazione dell'informatica; in questo modo, si è proposto un esempio robusto per far notare i pregi (anche dal punto di vista della complessità di progettazione) di questa architettura. Il caso di studio consiste in un robot *intelligente e riconfigurabile* che deve essere in grado di affrontare un percorso ad ostacoli per raggiungere una fonte di calore, quindi calato in un contesto di robot di salvataggio che è in grado di ricercare persone in ambienti avversi o identificare fonti di incendi e tentare di domare le fiamme. Il tutto è stato implementato con diversi tool di simulazione che sono stati fatti comunicare e rispettando i vincoli temporali di un sistema reale (quindi utilizzando clock, tempi fisici realistici con tutti i problemi di hard realtime annessi). Sono stati utilizzati quindi i seguenti tool di simulazione:

- *Blender* per la progettazione e implementazione della parte fisica ed estetica del robot;
- *V-REP* per la simulazione del robot implementato in Blender, utilizzando il più possibile una fisica realistica;
- *SystemC Rechannel* per la simulazione dell'hardware riconfigurabile;
- *Linux Ubuntu* per simulare il programma di intelligenza artificiale che dovrebbe essere eseguito su un qualsiasi processore e che deve interfacciarsi con la parte riconfigurabile (implementata in SystemC Rechannel).

Il funzionamento del sistema è il seguente: il robot può avere 2 differenti configurazioni, ovvero configurazione a quadrupede e configurazione a doppio serpente; queste configurazioni vengono pilotate e mosse dal sistema riconfigurabile in base al percorso che il robot deve eseguire. Il robot è sensibile alle fonti di calore, ovvero tramite sensore di visione di calore, è in grado di percepire dove si trova una fonte di calore e capire quanto è distante da esso. Per capire quanto è distante dalla fonte di calore, viene fatto uno studio statistico sulla quantità di pixel rossi che appaiono nell'immagine acquisita. Il robot non conosce l'ambiente in cui si trova, si deve solo basare sui sensori di distanza posti su di esso, uno per ogni lato, e sul sensore di visione; non deve interagire con l'ambiente ma deve solo essere in grado di non essere ostacolato nei suoi movimenti. Il task del robot consiste nel raggiungere o di avvicinarsi il più possibile alla zona di calore e, per farlo, deve riuscire a trovare il percorso migliore (ovvero quello fattibile). La strategia del robot consiste nel "provare" quale delle 2 sue configurazioni può essergli utile per raggiungere l'obiettivo: durante il suo tragitto, potrebbe trovare un ostacolo che gli impedisce di proseguire, quindi deve essere in grado di oltrepassarlo o passandoci sotto oppure aggirandolo. Deve inoltre essere in grado di capire se la scelta di strategia per oltrepassare l'ostacolo è corretta: è ammesso che il robot esegua degli sbagli durante le sue scelte, in quanto non ha una mappatura dell'ambiente, ma deve essere in grado



di tornare indietro nelle sue scelte, capire quando è avvenuto l'errore della sua scelta e quindi adottare una strategia diversa per raggiungere il target. La scelta del tragitto viene sempre attuata in base a dove si trova la fonte di calore: l'euristica è appunto quella di minimizzare il tragitto andando in direzione sempre dove la concentrazione di pixel rossi è maggiore. Ovviamente in questo modo il robot non sa dove si trovano gli ostacoli in quanto è in grado di "vedere" soltanto le fonti di calore, quindi il suo tragitto sarà pieno di scelte.

Si vedranno ora nel dettaglio le diverse implementazioni dei componenti.

## 10.1 Hardware riconfigurabile

Il sistema riconfigurabile svolge la funzione di moto, riconfigurazione e sensoristica del robot: tutta la parte di cinematica viene delegata a questa parte, con i relativi controlli sui sensori. Si potrebbe paragonare ad un sistema nervoso adattivo: esso è in grado di adattarsi all'ambiente percependolo, ma non comprendendolo. La comprensione di tutti i segnali viene delegata alla parte di intelligenza artificiale, questa parte si occupa solo di acquisire dati e di inviarli al "sistema di controllo e comprensione". Questo è stato implementato in simulazione su SystemC Rechannel ed è composto da 2 entità:

- *Top*, che contiene la parte di controllo del sistema riconfigurabile che si interfaccia direttamente con i moduli riconfigurabili; è in grado di dare il comando di riconfigurazione ai moduli riconfigurabili e di acquisire i segnali dai sensori del robot per trasformarli in dati comprensibili alla parte di intelligenza;
- *Moules*, che contiene i moduli riconfigurabili che eseguono gli algoritmi di movimento.

L'algoritmo di funzionamento è il seguente: il Top controlla i sensori del robot e li invia al sistema di intelligenza artificiale che li elabora; successivamente il Top riceve il comando in base ai valori dei sensori e il Top, in base alla sua macchina a stati, invia l'azione da eseguire al modulo riconfigurabile attualmente attivo. Parallelamente il Top controlla se deve essere eseguita una riconfigurazione: se deve essere riconfigurato il sistema, dealloca il modulo riconfigurabile e attiva quello prescelto, in modo che venga eseguita la prossima azione su quel modulo. Contemporaneamente il modulo riconfigurabile attualmente attivo esegue l'azione che gli è stata precedentemente inviata dal Top, per poi attendere il prossimo comando.

Il Top è composto quindi dai seguenti componenti:

- *activator*, che si occupa di attivare e disattivare i moduli riconfigurabili;
- *producer*, che invia i dati ai moduli riconfigurabili;
- *controller*, che si occupa di tenere traccia della riconfigurazione attuale e di quella futura;
- *monitor*, che si occupa di leggere i dati di risposta dei moduli riconfigurabili.

Si ha quindi anche parallelismo interno al modulo Top, in quanto deve essere in grado di gestire contemporaneamente i diversi aspetti della riconfigurazione. Questo parallelismo è gestito da SystemC Rechannel, però lo scambio di messaggi (che risulta bloccante secondo l'architettura Client-Server rispettivamente dei moduli riconfigurabili e del Top) viene gestito dal progettista.

## 10.2 Robot riconfigurabile

Il robot è stato progettato appositamente per questo progetto, riprendendo il modello del robot serpente di esempio del simulatore V-REP. Questo non presenta una riconfigurabilità totale, in quanto, a causa del simulatore, non era possibile rappresentare un robot che presentasse giunti semovibili. In ogni caso, il robot è in grado di riconfigurarsi grazie alle direttive dell'hardware riconfigurabile.

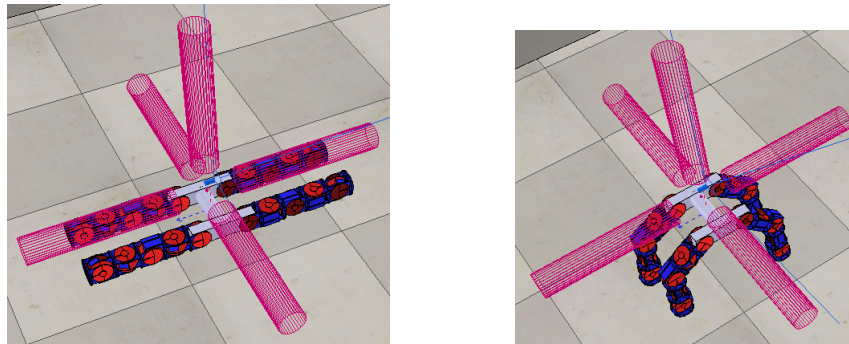


Figura 9: Robot nelle sue 2 configurazioni: a serpente doppio e a quadrupede

Dal punto di vista di controllo, il robot viene controllato in base alle posizioni dei giunti (controllo di posizione), quindi ogni istante della simulazione viene inviato un comando di posizione a tutti i giunti del robot per eseguire un movimento. A causa della natura del sistema che lo comanda (hardware riconfigurabile), ogni posizione è discreta, quindi il moto non è stato descritto con un'equazione nel continuo ma da una serie di posizioni dei giunti.

Il robot è composto dalle seguenti parti:

- diversi componenti locomotori connessi in maniera opposta l'uno dall'altro, che uniti tra loro permettono di eseguire molteplici movimenti;
- un corpo centrale a forma di "H", in cui risiede tutta la circuiteria e permette di avere una base solida per alcuni tipi di movimenti.

Ogni componente pesa 40g e deve avere l'attrito necessario per permettere il movimento; sfortunatamente il tool di simulazione non permette di impostarlo e quindi alcuni movimenti non rappresentano correttamente la realtà.

Per le diverse configurazioni, sono stati presi diversi modelli naturali per rappresentare il moto: la configurazione a serpente doppio ha preso spunto dal moto del serpente

(moto ondulatorio verticale), mentre la configurazione a quadrupede ha preso spunto dalla marcia del cavallo (passo).

### 10.2.1 Configurazione a serpente doppio

In questa configurazione, il robot si trova sdraiato, completamente a contatto con il terreno. Per muoversi, deve eseguire un moto ondulatorio verticale, come alcuni tipi di serpenti. Questo tipo di moto può essere descritto con queste equazioni:

$$y(t_x) = (j_{v1} \cdot (1 - s) + j_{v2} \cdot s) \cdot \sin(t \cdot (v_{v1} \cdot (1 - s) + v_{v2} \cdot s) + i \cdot (j_{v1} \cdot (1 - s) + j_{v2} \cdot s))$$

$$y(t_{x+1}) = (j_{v1} \cdot (1 - s) + j_{v2} \cdot s) \cdot \cos(t \cdot (v_{v1} \cdot (1 - s) + v_{v2} \cdot s) + i \cdot (j_{v1} \cdot (1 - s) + j_{v2} \cdot s))$$

la variabile  $y$  indica l'output della funzione,  $t_x$  indica un istante di tempo generico mentre  $t_{x+1}$  indica l'istante di tempo successivo al precedente. Le variabili  $j_{v1}, j_{v2}$  indicano le posizioni dei giunti verticali, indipendentemente da quanti giunti esistono (come nel caso del tempo, si indicano semplicemente giunti successivi tra loro). Le variabili  $v_{v1}, v_{v2}$  indicano invece le velocità dei giunti successivi tra loro. Si deve pensare che queste 2 equazioni si devono eseguire ciclicamente in modo da avere un movimento continuo ed ondulatorio, per permettere di strisciare.

A causa della natura del robot, poichè la parte centrale è statica, si deve eseguire il movimento con più potenza del necessario, in modo da essere in un'istantanea fase di volo che può dare la spinta necessaria per avanzare (e quindi spostare la parte statica del robot).

Per la rappresentazione spaziale, è stato adottato il modello di *Chirikjian - Burdick*, in cui si prende come sistema di riferimento fisso la "cod" del serpente, cioè l'ultimo giunto, e tutti i calcoli relativi ai successivi giunti del corpo serpentoide del robot verranno eseguiti in base a tale scelta. In particolare, il sistema viene visto come un insieme di 4 serpenti collegati ad un corpo rigido: in questo modo i movimenti vengono semplicemente ripetuti in maniera speculare o come prolungamento di altri serpenti, semplicemente riferendosi ai giunti sequenziali del robot.

Poichè ogni posizione di ogni giunto è nota durante l'esecuzione, è stata utilizzata la cinematica diretta per determinare la posizione e il movimento di ogni componente. Inoltre, poiché il numero dei giunti era molto elevato, è stata utilizzata la convenzione di *Denavit-Hartenberg*, rappresentando una trasformazione geometrica nello spazio euclideo tridimensionale con 4 parametri. Calcolando questi valori, è stato possibile quindi calcolare la posizione dei giunti in maniera ottimale per il corretto movimento del robot.

Quindi, implementando ogni posizione dei giunti, si è riusciti a far muovere il robot con il moto ondulatorio verticale.

In questa configurazione è anche possibile muoversi lateralmente, rovesciando il robot lateralmente 2 volte consecutive. Per fare ciò, è necessario piegare contemporaneamente 2 giunti dello stesso lato, in modo che il robot si possa rovesciare nel lato opposto, per poi rieseguire lo stesso movimento con gli altri 2 giunti.

### 10.3 Intelligenza Artificiale

### 10.4 Considerazioni

## 11 Conclusione

### Riferimenti bibliografici

- [1] Garrison W. Greenwood, Senior Member, IEEE and Andy M. Tyrrell, Senior Member, IEEE, “*Metamorphic Systems: A New Model for Adaptive System Design*”, 2010
- [2] Sekanina, L. ; Brno Univ. of Technol., Brno ; Martinek, T. ; Gajda, Z., *Extrinsic and Intrinsic Evolution of Multifunctional Combinational Modules*, 2006
- [3] Altera, *FPGA Run-Time Reconfiguration: Two Approaches*, 2008
- [4] Dirk Koch, *Partial Reconfiguration on FPGAs*, 2013
- [5] Raabe A., Tech. Comput. Sci., Univ. of Bonn, Felke A., *A SYSTEMC language extension for high-level reconfiguration modelling*, 2008
- [6] Erlangung des Doktorgrades, Mathematisch-Naturwissenschaftlichen Fakultät, Rheinischen Friedrich-Wilhelms-Universität Bonn, *Describing and Simulating Dynamic Reconfiguration in SystemC Exemplified by a Dedicated 3D Collision Detection Hardware*, 2008