

# Introducción a los FPGAs y el Cómputo Reconfigurable

Miguel Morales Sandoval  
INAOE, 2006

## ¿Qué son los FPGAs?

Matriz de bloques lógicos configurables (CLB) y una matriz de interconexión. Los bloques lógicos contienen memoria para configuración del bloque, usan LUT que implementa lógica combinacional, flip-flops y multiplexores a la entrada y salida del bloque lógico.

Hay otros bloques dedicados para las entradas y salidas del FPGA llamados IOB.

## ¿A que se refiere la granularidad de los FPGAs?

Al tamaño de la menor unidad funcional que las herramientas de ruteo pueden manipular en el FPGA. Granularidad fina utiliza unidades funcionales pequeñas que conllevan a tener alta flexibilidad pero también a tener mayores retardos cuando se implementan sistemas grandes.

## ¿Cómo se configuran los FPGAs?

Se descargan bits de configuración (todos están contenidos en un bitstream) para modificar la funcionalidad de los bloques lógicos y las interconexiones entre ellos. Generalmente los bits se almacenan en una memoria SRAM incluida en cada bloque lógico, que es quien controla la interconexión y funcionalidad de los elementos lógicos de cada bloque lógico.

## ¿Cuánto tarda reconfigurar un FPGA?

El tiempo de configuración es proporcional al tamaño del bitstream. Los FPGAs de grano fino tienen tiempos de configuración mayores, por la mayor cantidad de bloques funcionales que deben configurarse.

## ¿De cuantas formas puede configurarse un FPGA?

La tecnología permite diferentes formas de cambiar la configuración de los bloques y conexiones.

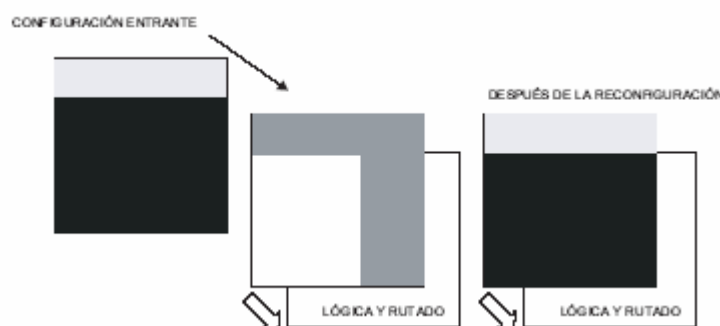
### **Reconfiguración total, en tiempo de compilación o reconfiguración estática:**

Cada vez que se realiza una nueva configuración, todo el FPGA se actualiza. Se tiene que detener la operación del FPGA, configurarlo todo y volver a ponerlo en marcha. Es para dispositivos con acceso secuencial a la memoria de configuración. Hay penalizaciones de tiempo de configuración.

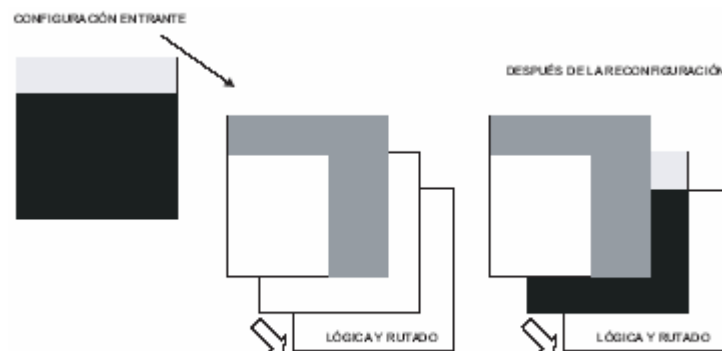
**Reconfiguración dinámica:** Partes del FPGA se configuran nuevamente sin detener la ejecución del resto del FPGA.

Existen tres variantes:

**Único contexto:** La configuración entrante sustituye completamente a la anterior. Aunque solo una pequeña parte del bloque vaya a cambiar, todo el bloque cambia.



**Multicontexto:** Existen varios bits de memoria de reconfiguración para cada bit de los elementos configurables. Los bits de memoria pueden considerarse como múltiples planos de información de configuración. Cada plano se configura totalmente como en el caso del contexto único. Se puede cargar una nueva configuración en un plano no activo mientras otro está activo. La reconfiguración implica un cambio de contexto, que se realiza de forma rápida.



(b) *Modelo multicontexto.*

**Reconfiguración parcial:** Algunos dispositivos la soportan. Se puede modificar una parte de la configuración mientras la otra sigue realizando la computación de forma no interrumpida.

El plano de configuración funciona como una memoria RAM. Las direcciones especifican una localización para que se reconfigure. Con esto se puede cargar nuevas configuraciones en áreas del FPGA sin necesitar cambiar el contexto.

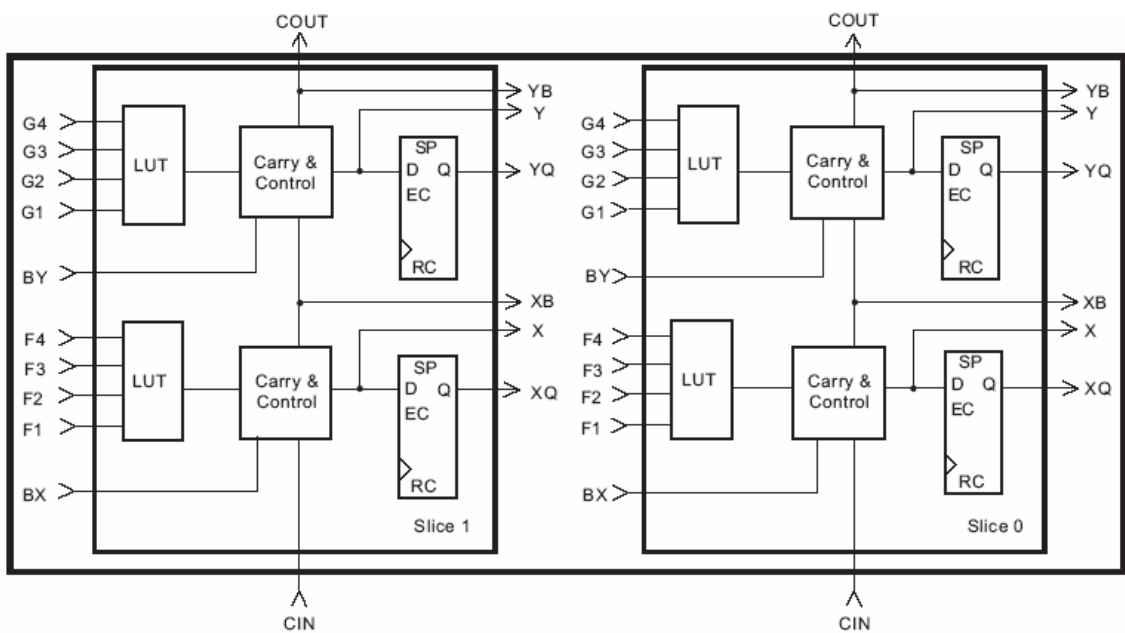
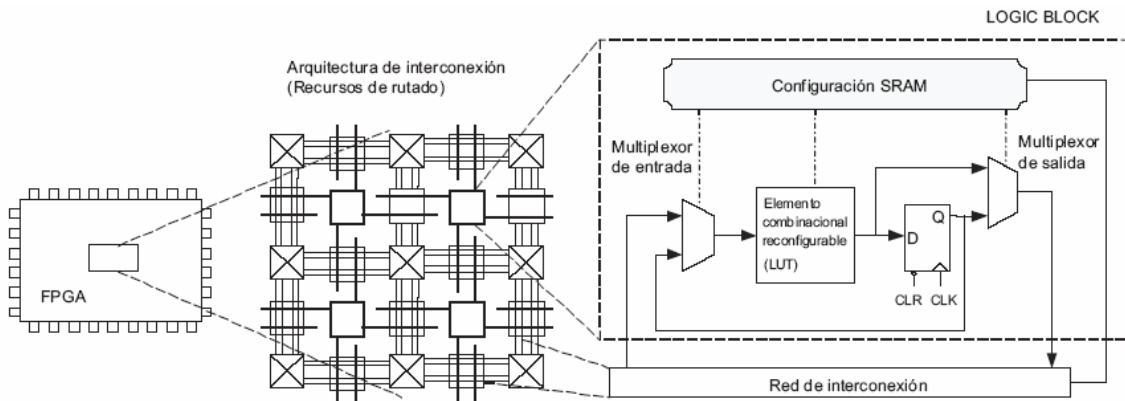
**Reconfiguración pipeline:** Técnica para reducir el tiempo de descarga de los bitstreams parciales.

Otras técnicas estudiadas para reducir el tiempo de reconfiguración son: prefetching de configuración, compresión de la configuración, uso de caché de configuraciones en el dispositivo (se reduce la cantidad de información transmitida). En este último caso se aplican los conceptos de localidad espacial y temporal con el fin de decidir que configuraciones mantener en las memorias caché.

Hay arquitecturas que integran varios dispositivos a fin de experimentar reconfiguración para varias aplicaciones potentes. Algunos ejemplos de estas arquitecturas son Címera y Piperech. Otras arquitecturas son los FPGAs de grano fino, que permiten reconfiguración dinámica.

Elementos que conforman los FPGAs de Xilinx.

- *Elementos reconfigurables: CLB e IOBs.* Elemento básico del CLB es la celda lógica. Cada CLB puede tener varias de estas celdas lógicas (CL). Los CLB de Virtex tienen cuatro CL distribuidas en dos slices.



CLB del FPGA Virtex. 4 celdas lógicas en dos slices

La celda lógica incluye:

1. Generador de funciones (circuito programable LUT). Puede funcionar como circuito combinatorial de 4 entradas o como memoria RAM 16x1. Las memorias pueden ser de varios tamaño combinando varias LUTs. También puede funcionar como registro de corrimiento de 16 bits.
  2. Lógica de acarreo.
  3. Elemento de memoria (para almacenar el resultado de la LUT).
- *Matriz de ruteo global*. Rutas horizontales y verticales con switches situados en la intersección entre las rutas. Para interconectar diferentes CLB internamente.
  - *I/Os*: Para interconectar los pines del FPGA y los CLBs.
  - *Rutas locales*: Para interconectar CLBs localmente y para conectar los CLBs con la matriz de ruteo global.
  - *Dispositivos para la gestión del reloj*: DDL en Spartan y DCMS en Virtex.
  - *Memoria RAM*: No emplean recurso del CLB. Los bloques de memoria se distribuyen en las columnas cubriendo toda la altura del dispositivo. También hay recursos de ruteo para conectar estos bloques con los CLBs. El contenido se fija mediante la memoria SRAM de configuración, incluyéndose en el bitstream inicial o bitstream parciales posteriores. Estos bloques pueden funcionar como memoria ROM.
  - *Buffers de 3 estados*: Para conectar los CLBs a las rutas horizontales globales. Se llaman Tbufs.

Todos estos elementos se configuran mediante valores almacenados en celdas de memoria SRAM. Cambiar la funcionalidad del dispositivo equivale a cambiar los valores de esta memoria.

### **Formas de almacenar los valores (bitstream) en la memoria.**

**Configuración serie:** El bitstream se lee desde una memoria PROM serie exterior. Dos variantes:

Serie Master: El FPGA genera el reloj de la carga

Serie Slave: El reloj lo genera otro dispositivo externo. Permite que la carga se realice desde un microprocesador externo.

**Configuración en paralelo:** Configuración a través de un puerto paralelo de 8 bits *SelectMap*. La memoria de configuración está dividida en frames. A través del *SelectMap* se introducen frames concatenados en el bitstream. Aunque es un puerto de 8 bits, la configuración se realiza mediante paquetes de 32 bits.

El *SelectMap* dispone de varios registros internos. Registro de comandos, registro de control y registro de dirección. Los paquetes de configuración de 32 bits pueden incluir información para estos registros de forma que se controlan operaciones de lectura de la memoria de configuración o la reconfiguración parcial dinámica.

**Configuración *Boundary scan* (JTAG).** La configuración se realiza mediante el protocolo serie JTAG. La configuración es a través de la interfaz *Boundary scan*, que esta activa y accesible en todo momento. Es el sistema utilizado en la fase de desarrollo y depuración.

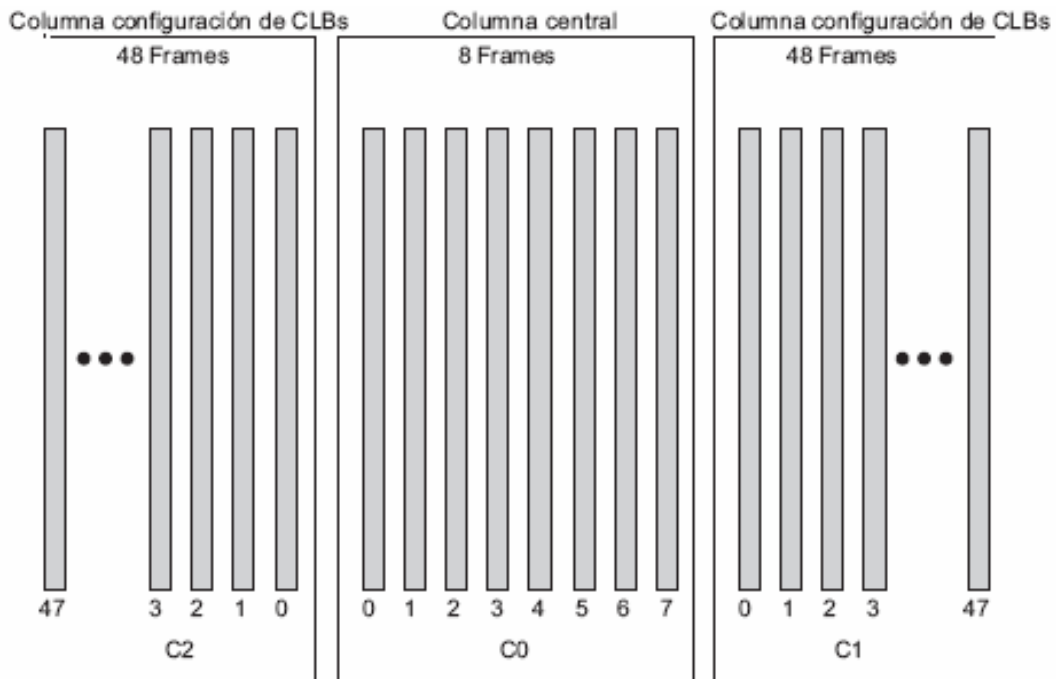
## Reconfiguración en FPGAs de XILINX

La mínima sección del FPGA que puede cambiar esta definida por un frame. El frame es una unidad de la memoria SRAM de configuración. Un frame tiene un bit de anchura y abarca todo el alto del FPGA

Hay BRAM Frames y CLB Frames.

Los frames se agrupan en unidades superiores denominadas frame columns.

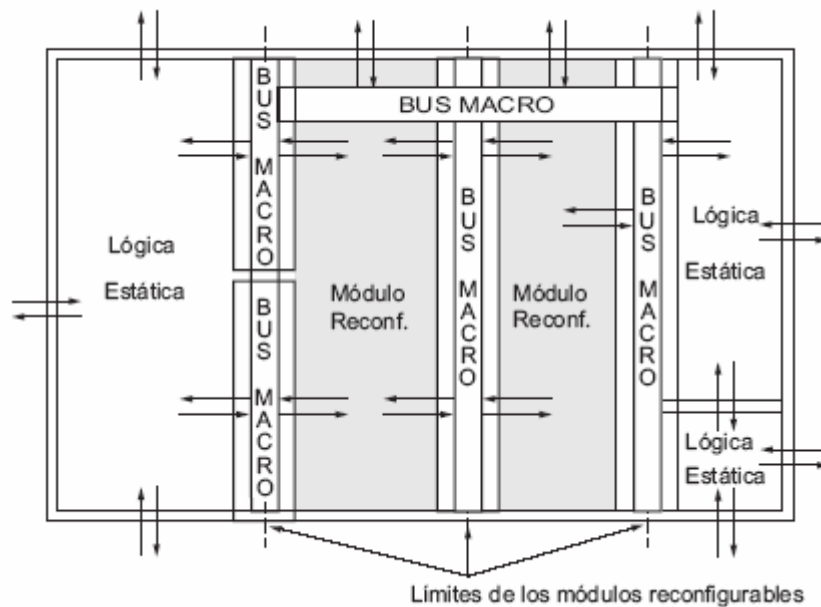
- 1 columna central de 8 frames
- columnas de 48 frames para configurar los CLBs.
- 2 columnas de 27 frames que especifican la conexión de las BRAM
- 2 columnas cada una de 54 frames para configurar los IOBs.



Dirección de un frame: Dirección del column frame (mayor address) y el numero de frame en esa columna (minor address).

Restricciones generales:

- La altura de una sección reconfigurable = la altura de la matriz de recursos lógicos.
- La anchura ocupada por el modulo reconfigurable debe ser multiplo de 4 slices.
- La lógica definida en la anchura ocupada por un modulo reconfigurable pertenece a ese módulo.
- Existen frames separados para configurar la lógica del reloj.
- Los IOBs se configuran mediante CLB frames.
- Los límites definidos un módulo reconfigurable no se pueden cambiar dinámicamente.
- La comunicación entre distintos módulos reconfigurables se realiza utilizando elementos fijos llamados Bus-Macros.



Metodología de reconfiguración basada en Java (JBits)

- JBits posibilita el acceso a los datos de configuración del dispositivo (bitstream).
- Usa el modelo de dispositivo reconfigurable basado en una matriz de CLBS.
- Usa secciones de la matriz dinámicamente y las ve como objetos java.
- El conexionado de objetos es lo mas complejo de JBits.

Metodología de reconfiguración con manipulación directa del bitstream

- Extrae dinámicamente la parte del circuito del bitstream
- Con la parte extraída, se compone un bitstream parcial

- Existen dos flujos de diseño: Manipulación de bits (usa el FPGA editor) y basado en módulos (varios bitstreams parciales, uno total para carga inicial, requiere de los BUS Macro).

## Cómputo reconfigurable

El cómputo reconfigurable es mucho más que utilizar FPGAs para implementar sistemas. Un FPGA solo sirve para hacer un prototipo de un sistema que finalmente será implementado en un ASIC.

Los FPGAs están siendo utilizados en el mercado para implementar una variedad de aplicaciones, ya que mantienen los beneficios del hardware y del software. Un protocolo de red en un FPGA permite futuras actualizaciones, ya que el dispositivo puede cambiar su funcionalidad de acuerdo a actualizaciones del protocolo para resolver algún bug en particular. Simplemente “programa” con la nueva configuración.

El computo reconfigurable va un paso mas allá, involucra manipular la lógica dentro del FPGA en tiempo de ejecución.

El diseño del FPGA cambia en respuesta a las demandas de la aplicación en tiempo de ejecución.

El CPU (unidad central de procesamiento) en una PC actúa como un bloque que ejecuta diferentes programas (cómputo temporal). Un FPGA (RPU, Unidad reconfigurable de procesamiento) puede verse como un bloque que permite la ejecución de varias funciones en hardware (computo espacial).

El cómputo reconfigurable es un enfoque para diseñar sistemas que hagan mas, , cuesten menos, y tengan ciclos de diseño e implementación cortos.

Ventajas del Cómputo Reconfigurable:

- Mayor funcionalidad con diseños hardware simples
- Menores costos de diseño e implementación
- Menor tiempo de mercado.

Computo configurable: Reprogramar todo el FPGA, no se puede mantener una parte.

Cómputo Reconfigurable: Programar partes del FPGAs mientras otras continúan ejecutándose.

La tecnología permite aplicaciones: reconfiguración en tiempo de ejecución con el beneficio de reducción en el consumo de potencia.

- FPGAs de AMTEL
- FPGAs Virtex de Xilinx.

Objeto hardware (componente hardware funcional o lógico): Puede ejecutarse en alguna parte de un RPU.

- Tiene restricciones como tamaño y forma que limitan el número de posibles posiciones dentro de la RPU. La forma debe ser rectangular y la altura de los bordes debe ser múltiplo de alguna unidad de longitud.
- Se debe definir una forma estándar de interconectar los objetos hardware (look and feel).
- La interface de un bloque hardware con otros módulos (otra RPU, memoria, procesador, etc.) debe ser una abstracción.

El cómputo reconfigurable necesita de software para:

- Decidir que objetos hardware ejecutar y cuando.
- Intercambiar objetos hardware dentro y fuera de RPU.
- Realizar el ruteo entre el objeto hardware y la abstracción (interface).

Estas tareas implican tener un microprocesador embebido o un microcontrolador. El software que corre en este procesador es el equivalente al sistema operativo de un CPU de una PC.

Xilinx describe dos flujos de implementación de reconfiguración dinámica:

- Ambos modos: slave SelectMAP y Boundary Scan (JTAG).

Como se configuran los FPGAs Virtex, VirtexII y VIRTEX II Pro (Sedcole 2006).

- La configuración se almacena en una memoria bidimensional SRAM
- Se puede leer y escribir a esta memoria sin detener el dispositivo.
- Un frame es la parte mas pequeña que se puede leer o escribir. Es una columna de la memoria de un bit de ancho.
- Un Frame se extiende a lo alto del FPGA.
- El árbol de relojes es una consideración importante en la reconfiguración modular parcial.