

Reconfigurable Network Processing: The FPGA Case

Carlos A. Zerbini and Jorge M. Finochietto

Laboratorio de Comunicaciones Digitales
Universidad Nacional de Córdoba, Facultad Ciencias Exactas, Físicas y Naturales

Departamento de Ingeniería Electrónica
Universidad Tecnológica Nacional, Facultad Regional Córdoba

12th AST Argentine Symposium on Technology
40th JAIIO
UTN Regional Córdoba

Agenda

Agenda

Requerimientos de procesamiento en Redes

Características de Tráfico

- Las redes de datos crecen en **Complejidad**: nuevas aplicaciones, multimedia
- Las redes de datos crecen en **Velocidad**: $n \times 100Gbps$ ($1Tbps@2015$)
- **Consolidación** de múltiples servicios sobre redes Ethernet
- Redes *Locales*, *Metropolitanas* y *Extensas* utilizan **Conmutación de Paquetes**
- Adopción de tecnologías para *virtualización* en redes y servidores

Procesamiento de Paquetes

- Los *enlaces* ofrecen alta capacidad. El *procesamiento* de paquetes es **crítico** y debe optimizarse
- El *rocesamiento a velocidad de línea*
- Paquete Ethernet mínimo = $64bytes \rightarrow 6nanosegundos/paquete$

Requerimientos de procesamiento en Redes

Características de Tráfico

- Las redes de datos crecen en **Complejidad**: nuevas aplicaciones, multimedia
- Las redes de datos crecen en **Velocidad**: $n \times 100Gbps$ (1Tbps@2015)
- **Consolidación** de múltiples servicios sobre redes Ethernet
- Redes *Locales*, *Metropolitanas* y *Extensas* utilizan **Conmutación de Paquetes**
- Adopción de tecnologías para *virtualización* en redes y servidores

Procesamiento de Paquetes

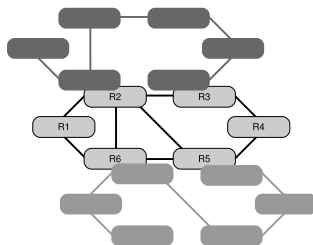
- Los *enlaces* ofrecen alta capacidad. El *procesamiento* de paquetes es **crítico** y debe optimizarse
- El rocesamiento *a velocidad de línea*
- Paquete Ethernet mínimo = 64bytes \rightarrow 6nanosegundos/paquete

Agenda

Soluciones

Granularidad

- **Actualmente** → **paquetes** de longitud **variable**
- Peor caso → mínima longitud (64 bytes en Ethernet)
- **Tendencia** → agregación de paquetes en **flujos**
- Ejemplos: Multi-protocol Label Switching (MPLS), VLANs (802.1Q)



Agregacion de flujos

Virtualizacion de redes

Tecnologías Actuales

Requerimientos → **flexibilidad, performance**

Circuitos de Propósito Específico (ASICs)

- Cientos de bloques especializados trabajando en paralelo
- Alto desempeño. No programables, alto costo y tiempo de desarrollo.

Network Processors (NPs)

- Múltiples elementos de procesamiento, buena performance para ciertas tareas. IXP(Intel), PowerNP (IBM)
- Difícil portabilidad, interfaces propietarias

Procesadores de Propósito General (GPPs)

- Arquitectura PC + Software especializado: *Click, Zebra/Xorp/Quagga*
- Alta flexibilidad, bajo costo. Limitación por transacciones con RAM y naturaleza secuencial

Tecnologías Actuales

Requerimientos → **flexibilidad, performance**

Circuitos de Propósito Específico (ASICs)

- Cientos de bloques especializados trabajando en paralelo
- Alto desempeño. No programables, alto costo y tiempo de desarrollo.

Network Processors (NPs)

- Múltiples elementos de procesamiento, buena performance para ciertas tareas. IXP(Intel), PowerNP (IBM)
- Difícil portabilidad, interfaces propietarias

Procesadores de Propósito General (GPPs)

- Arquitectura PC + Software especializado: *Click, Zebra/Xorp/Quagga*
- Alta flexibilidad, bajo costo. Limitación por transacciones con RAM y naturaleza secuencial

Tecnologías Actuales

Requerimientos → **flexibilidad, performance**

Circuitos de Propósito Específico (ASICs)

- Cientos de bloques especializados trabajando en paralelo
- Alto desempeño. No programables, alto costo y tiempo de desarrollo.

Network Processors (NPs)

- Múltiples elementos de procesamiento, buena performance para ciertas tareas. IXP(Intel), PowerNP (IBM)
- Difícil portabilidad, interfaces propietarias

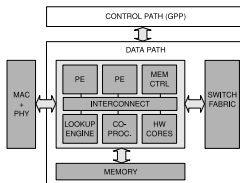
Procesadores de Propósito General (GPPs)

- Arquitectura PC + Software especializado: *Click, Zebra/Xorp/Quagga*
- Alta flexibilidad, bajo costo. Limitación por transacciones con RAM y naturaleza secuencial

Nuevas tecnologías

Dispositivos Lógicos Programables (FPGAs)

- Permiten *reconfiguración* y *reprogramación*, contando con librerías de *Open Hardware*.
- Su performance no es lejana a la de un ASIC. Fabricantes: Altera, Xilinx, Actel.
- Incorporación creciente de bloques *hardcore* especializados



Implementación con GPPs

Implementación con NPs

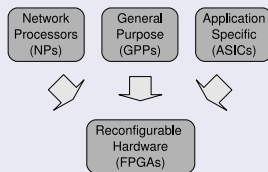
Implementación con FPGAs

Agenda

Investigación y Desarrollo

Tendencias actuales

- Arquitecturas **reconfigurables** y **reprogramables** para adaptarse a nuevas aplicaciones
- Arquitecturas **eficientes** para satisfacer exigencias de consumo y velocidad
- Extensión del *Open Software* mediante *Open Hardware*
- **Métricas a evaluar:** paralelismo vs. consumo de memoria/lógica, velocidad



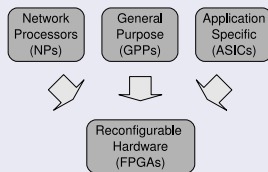
Propuestas

- Procesamiento de **flujos** de paquetes en arquitecturas **reconfigurables**
- Desarrollo de implementaciones **a medida** de la aplicación
- Integración en arquitecturas de procesamiento *asistido por hardware*

Investigación y Desarrollo

Tendencias actuales

- Arquitecturas **reconfigurables** y **reprogramables** para adaptarse a nuevas aplicaciones
- Arquitecturas **eficientes** para satisfacer exigencias de consumo y velocidad
- Extensión del *Open Software* mediante *Open Hardware*
- **Métricas a evaluar:** paralelismo vs. consumo de memoria/lógica, velocidad



Propuestas

- Procesamiento de **flujos** de paquetes en arquitecturas **reconfigurables**
- Desarrollo de implementaciones **a medida** de la aplicación
- Integración en arquitecturas de procesamiento *asistido por hardware*

Trabajos Relacionados

Click Modular Router (MIT)

- Librería de objetos **Software** libre, conectados mediante lenguaje propio
- Ejecutable a nivel de *usuario* o de *kernel*
- Sufre limitaciones por su naturaleza secuencial
- Migración automática a hardware → difícil de lograr

NaPPoX (Siemens)

- Plataforma Hardware basada en un FPGA Xilinx Virtex 5
- Versión 10 GbE liberada en 2010
- Sistema base *predefinido* para *incorporación* de funcionalidades

Trabajos Relacionados

Click Modular Router (MIT)

- Librería de objetos **Software** libre, conectados mediante lenguaje propio
- Ejecutable a nivel de *usuario* o de *kernel*
- Sufre limitaciones por su naturaleza secuencial
- Migración automática a hardware → difícil de lograr

NetFPGA (Stanford)

- Plataforma **Hardware** basada en un FPGA Xilinx Virtex 5
- Versión 10 GbE liberada en 2010
- Sistema base *predefinido* para *incorporación* de funcionalidades

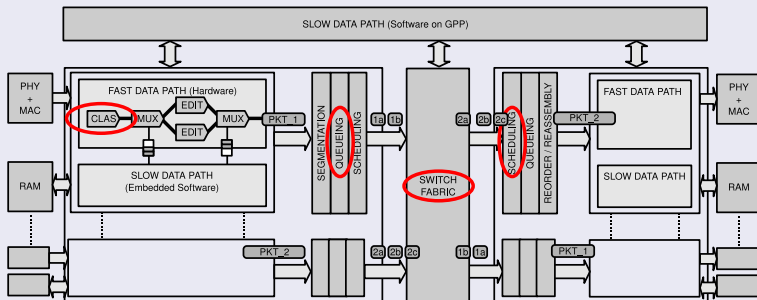


Agenda

Procesamiento en Redes

Introducción

- **Básico:** Routing → lookup, forwarding, queueing
- **Avanzado:** agregación de paquetes, modificación del *camino de datos*
- **Caminos de procesamiento**
 - *Fast Data Path:* operaciones a velocidad de línea
 - *Slow Data Path:* operaciones menos frecuentes
- **Bloques de interés:** Clasificación, Almacenamiento, Conmutación, Planificación

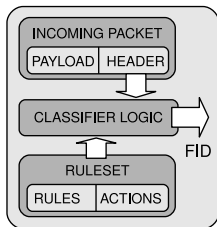


Agenda

Procesamiento en Redes de Datos

Clasificación

- Bloques **esenciales** para procesamiento de flujos diferenciados
- Exigencias particulares de *almacenamiento y velocidad*
- Extracción *configurable* de encabezados contra *uno o múltiples* campos *emph storage/rule*
- Matching *exacto* (Protocolo, puertos, MAC) o *mejor matching* (LPM, prefijos IP)
- Enfoques de diseño: *algorítmico* (Trees/decomp.) y *arquitectura* (TCAM)



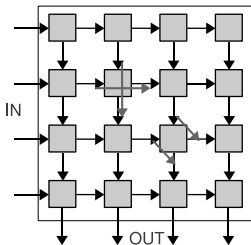
GPPs	FPGAs	ASICs
Many rules supported in large and slow SDRAM	Dedicated registers; low-latency limited internal SRAM; medium-latency external SDRAM	Limited ad-hoc internal memory; external SDRAM (if interface included)
Intensive algorithms at low speed	Parallel architectures at wire speed	Parallel architectures at wire speed
Fully reprogrammable rules	Fully reprogrammable/reconfigurable rules	Partially reprogrammable rules

Agenda

Procesamiento en Redes de Datos

Conmutación (Switching)

- Múltiples puertos de E/S → **Contención**
- **Time-division Switching:** bus o memoria compartidos. Soporta tráfico *agregado*.
- **Space-division Switching:** transferencias *simultáneas*. Limitado por efectos *físicos y bloqueo interno*.



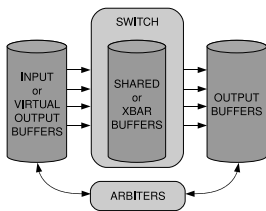
GPPs	FPGAs	ASICs
Shared bus/memory	Time/space division	Time/space division
Low/medium performance	High performance	Very high performance
Little/no optimization possible	Optimization limited by device family	Optimization at design time

Agenda

Procesamiento en Redes de Datos

Almacenamiento (Queueing)

- Junto con la *Conmutación*, determina desempeño y campo de aplicación
- Los buffers absorben *contención* y aplican políticas de QoS
- Pueden ubicarse en *entrada, salida, entrada/salida o integrada a la matriz de conmutación*
- Almacenamiento de *celdas o paquetes* → **segmentación/reensamblado**



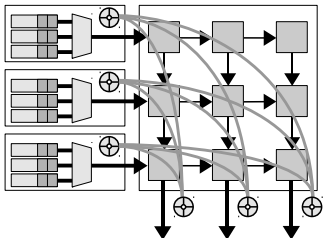
GPPs	FPGAs	ASICs
System SDRAM	Internal SRAM or external SDRAM	Internal SRAM or external SDRAM
High latency and capacity	Low latency or high capacity	Low latency or High capacity
Centralized	Centralized/distributed	Centralized

Agenda

Procesamiento en Redes de Datos

Planificación (Scheduling)

- Surge ante el problema de *contención*
- Distribución Equitativa → algoritmo *Round-Robin*
- Se combina usualmente con el efecto de *policers* (entrada) y *shapers* (salida)
- Fuertemente condicionado por la función de *Switching*



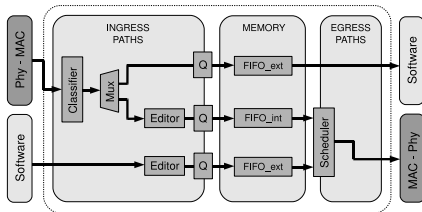
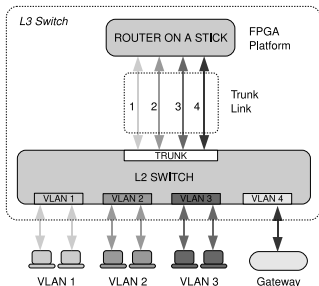
GPPs	FPGAs	ASICs
Programmable	Configurable/ Programmable	Fixed
Slow	Medium	Fast

Agenda

Caso de Estudio

Arquitectura planteada

- **Premisa:** utilizar sólo *lo necesario* en la *ubicación necesaria*
- Escenario de aplicación → **Conmutación entre LAN Virtuales (VLANs)**
- Configuración *Router on a Stick*: múltiples flujos (VLANs) se transportan mediante un *enlace troncal* único
- Los puertos de un *Switch* de Capa 2 se agrupan en LANs virtuales. La conmutación *entre VLANs* se realiza en un *Router* de Capa 3 (FPGA)



Agenda

Caso de Estudio

Resultados

Stage	Comb. ALUTs	Registers
Classifier (8 rules)	70	231
Queuers (3)	179	84
Scheduler	180	139
Editors (2)	77	327
<i>Total processing stages</i>	506	781
MAC IP Core and glue logic	1906	1792
<i>Total resources utilization</i>	2412	2573

Agenda

Conclusiones

Experiencia realizada

- La arquitectura conceptual implementada consume menos de 1 % de recursos en un FPGA *Altera Stratix IIGX*
- Trabajo *sin optimizaciones* a velocidad 10 GbE
- Sin embargo, su escalabilidad es limitada → campo para **optimizaciones**

Expectativas

- Contrastación de posibles optimizaciones y evaluación de otros escenarios
- Integración con plataforma *software* para migración parcial de funciones
- La evolución de esta tecnología acompañará el trabajo de optimización
- **Factor limitante:** largos ciclos de implementación → **Lenguajes de Alto Nivel (HLLs)**

Conclusiones

Experiencia realizada

- La arquitectura conceptual implementada consume menos de 1 % de recursos en un FPGA *Altera Stratix IIGX*
- Trabajo *sin optimizaciones* a velocidad 10 GbE
- Sin embargo, su escalabilidad es limitada → campo para **optimizaciones**

Expectativas

- Contrastación de posibles optimizaciones y evaluación de otros escenarios
- Integración con plataforma *software* para migración parcial de funciones
- La evolución de esta tecnología acompañará el trabajo de optimización
- **Factor limitante:** largos ciclos de implementación → **Lenguajes de Alto Nivel (HLLs)**

Gracias por su atención,
preguntas?

jfinochietto@efn.uncor.edu
czerbini@electronica.frc.utn.edu.ar