Concurrent Hybrid Switching for Massively Parallel Systems on Chip: The CYBER Architecture

Francesca Palumbo, Danilo Pani, Andrea Congiu and Luigi Raffo

EOLAB - Microelectronics Lab
DIEE - Dept. of Electrical and Electronic Eng.
University of Cagliari - ITALY
### ON CHIP NETWORK COMMON STRATEGIES

<table>
<thead>
<tr>
<th>Thread Dependencies</th>
<th>Scalar Data</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Long Streaming Vector</td>
<td>CS</td>
</tr>
</tbody>
</table>

#### End to End Control
- **PS**: Load Balancing
- **CS**: Packet Switching

#### Load Balancing
- **PS**: Circuit Switching
- **CS**: Packet Switching

- **Packet Switching (PS) → Best-Effort (BE)**
  - Discontinuous traffic characteristics
  - Low Congestion

- **Circuit Switching (CS) → Guaranteed-Throughput (GT)**
  - Continuous traffic characteristics
  - High throughput
  - Set-up time required
Some studies (Prof. Li Shuan Phe, Princeton University) demonstrated that, for common MPPAs workloads, having classical packed switched NoC implementations can be ineffective.
Some studies (Prof. Li Shuan Phe, Princeton University) demonstrated that, for common MPPAs workloads, having classical packed switched NoC implementations can be ineffective.

**HYBRID SWITCHING:**

- better usage of the available bandwidth
- more complex hardware implementation of the router
In literature the implementation of the hybrid switching is exclusive: *circuit switched and packet switched data can not proceed together*.
In literature the implementation of the hybrid switching is exclusive: *circuit switched and packet switched data can not proceed together.*
In literature the implementation of the hybrid switching is exclusive: *circuit switched and packet switched data can not proceed together.*
In literature the implementation of the hybrid switching is exclusive: *circuit switched and packet switched data can not proceed together.*

• GT data have a higher priority with respect to the BE ones
  - Packet switched data, if a circuit has been already established, will have to wait in intermediate nodes
  - Tasks completion time will be delayed until all the data interdependences will be solved
• Introduction and Background

• The CYBER Architecture

• Performance Analysis

• Final Remarks and Future Trends
CYBER: OVERVIEW OF THE ARCHITECTURE

MIXED FLOW PROGRAM LOGIC

PKT FLW INPUT
- flit
- nack
- Input FSM

PKT FLW OUTPUT
- XBAR 5x10
- XR
- FSM0
- VC0
- VC1
- RD

CIR FLW INPUT
- Input REG

CIR FLW OUTPUT
- XBAR 5x5

CODE CHANNEL
- FSMW
- FIFO
- FSMR

DECODE CHANNEL
- FSMR
- FIFO
- FSMW

CIRCUIT CHANNEL
- bus a 32 bit
• **CYBER**: Concurrent **hYBrid Enhanced Router**
  
  **CONCURRENT**: guaranteed and best-effort services are implemented along separated links.

  **HIBRID**: parallel implementation of both circuit switching and packet switching.

  **ENHANCED**: prioritized best-effort support has been implemented

• Stall of PS communications for long time intervals due to CS communications is avoided.
  
  – **PRO**: Quick data interdependencies solving;
  
  – **PRO**: Robustness to traffic heterogeneity
  
  – **CONS**: Split of a wider link into two different parallel links driven by two different crossbars
• Round-Robin arbiter
• Virtual output queuing
  – P0: long vectors, load balancing and end-to-end control
  – P1: scalar data and small vectors
• Prioritized best-effort QoS
  – P0: wormhole
  – P1: credit based virtual cut through
• X-Y routing scheme semi-adaptive for P0 packets
PACKET FLOW – TRADITIONAL X-Y ROUTING
CIRCUIT FLOW – ADAPTIVE X-Y-X ROUTING
18 bits crossbar
- 16 bit data
- 1 bit valid
- 1 bit release
• 18 bits crossbar
  - 16 bit data
  - 1 bit valid
  - 1 bit release

• Traditional TDM (Time Division Multiplexing) implementations:
  - recording tables;
  - slot tables;
  - routing tables.

area-intensive and lack in flexibility
In order to save as much area as possible a simple shared 5 bits register has been instantiated.
The input channels of the PACKET HANDLING section access this shared register to check the availability of the required output in the CIRCUIT HANDLING section.
At the end of the communication over the circuit, a release signal reaches each router and requires access to the register to refresh the related bit on it.
• Accesses are managed in a RR fashion to guarantee fairness.
• Release signals are always served first: to be able to release an output port of the CIRCUIT HANDLING and re-grant it in two consecutive clock cycles.
Introduction and Background

The CYBER Architecture

Performance Analysis

Final Remarks and Future Trends
• Finding the appropriate communication support for modern application is still an open issue.

• CYBER is a concurrent enhanced hybrid switching NoC:
  - combining PS and CS
  - implementing prioritized BE support
  - first level of adaptivity at the communication level wrt circuits
CONCLUSIONS

- Finding the appropriate communication support for modern application is still an open issue.
- CYBER is a concurrent enhanced hybrid switching NoC:
  - combining PS and CS
  - implementing prioritized BE support
  - fist level of adaptivity at the communication level wrt circuits

AETHEREAL:
GT+BE traffic was a mistake:
- BE lower priority than GT one
- not possible to offer a low latency to BE communications, despite the extra cost ratio
CONCLUSIONS

- Finding the appropriate communication support for modern application is still an open issue.
- CYBER is a concurrent enhanced hybrid switching NoC:
  - combining PS and CS
  - implementing prioritized BE support
  - fist level of adaptivity at the communication level wrt circuits

AETHEREAL:
GT+BE traffic was a mistake:
- BE lower priority than GT one
- not possible to offer a low latency to BE communications, despite the extra cost ratio

CYBER:
overall performance results where very good:
- robustness and predictability are important
- cost ration is more affordable
Acknowledgements

The research leading to these results has received funding from:

- the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 248424, MADNESS Project
- the Region of Sardinia, Young Researchers Grant, PO Sardegna FSE 2007-2013, L.R.7/2007 “Promotion of the scientific research and technological innovation in Sardinia”
Concurrent Hybrid Switching for Massively Parallel Systems on Chip: The Cyber Architecture

Francesca Palumbo, Post-Doc.
francesca.palumbo@diee.unica.it
EOLAB - Microelectronics Lab
Dept. of Electrical and Electronics Eng.
University of Cagliari (ITALY)