



Implementing the Advanced Switching Fabric Discovery Process

Antonio Robles-Gómez
Aurelio Bermúdez
Rafael Casado
Francisco J. Quiles

Universidad de
Castilla – La Mancha





Agenda

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Introduction to Advanced Switching
 - Architecture
 - Fabric Management
- Topology Discovery
 - Motivation
 - Serial Discovery
 - Parallel Discovery
- Performance Evaluation
 - Simulation Model
 - Simulation Methodology
 - Results
- Conclusions and Future Work



Introduction to Advanced Switching Overview

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Chip-to-chip and backplane interconnect switched fabric architecture
- Advanced Switching
 - Inherits most of the physical and link characteristics of PCI Express
 - Differing at the transaction layer
- Fabric management
 - High availability
 - Hot addition/removal of devices
 - Efficient handling of changes

Introduction to Advanced Switching

Fabric components

Advanced Switching

- Architecture
- Management

Discovery

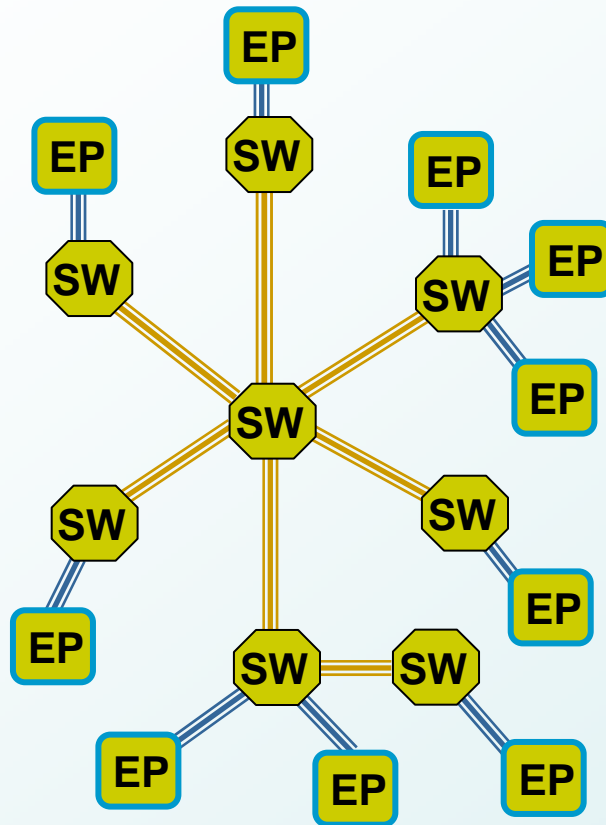
- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work



- Links
 - Communicate two devices
 - Minimum 2.5 Gbps
 - 2.0 Gbps by 8b/10b encoding
- Endpoints
 - Support up to 4 ports
 - Sources supply unicast packets with an explicit path
- Switches
 - Support up to 256 ports
 - Employ the packet routing header to calculate the output port

31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Header CRC								Turn Pointer				F	Credits Required				T	O	Traffic Class	P	P	Protocol Interface									
												E					S	O		C											
												C					N														
D	Turn Pool																														



Introduction to Advanced Switching Fabric Management

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- The specification provides some support for the fabric management
 - Set of configuration data structures and management packets
 - Management Entities
 - *Fabric Manager (FM)*
 - Configures and monitors the status of the fabric
 - *Device Manager (DM)*
 - Accesses to the local information in a device
- Management Mechanism
 - Detection of topological changes
 - Topology discovery
 - Path Computation and Distribution



Introduction to Advanced Switching

Device configuration space

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

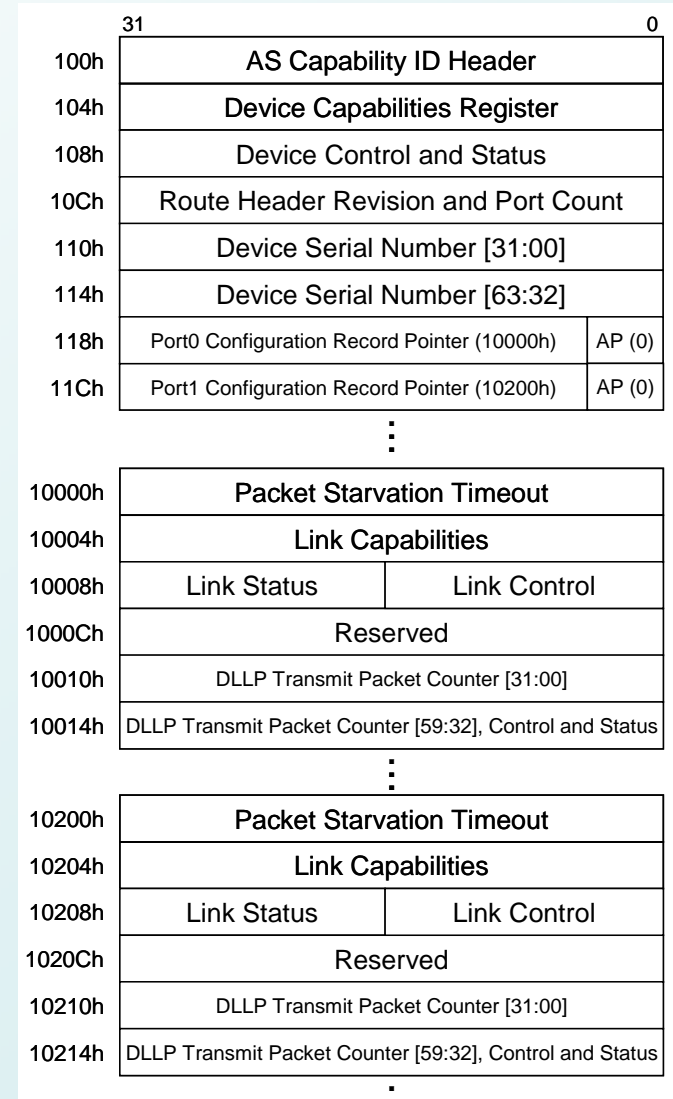
Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Set of structures, called capabilities
 - Organization detailed in the specification
 - Implemented in all fabric devices
- *Baseline device* - Device control and status information
 - General information
 - Device type
 - Serial number
 - Ports supported
 - Up to 256 pointers
 - Information about each particular port
 - Link speed and width
 - Current port state



Introduction to Advanced Switching

Node configuration and control protocol (PI-4)

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

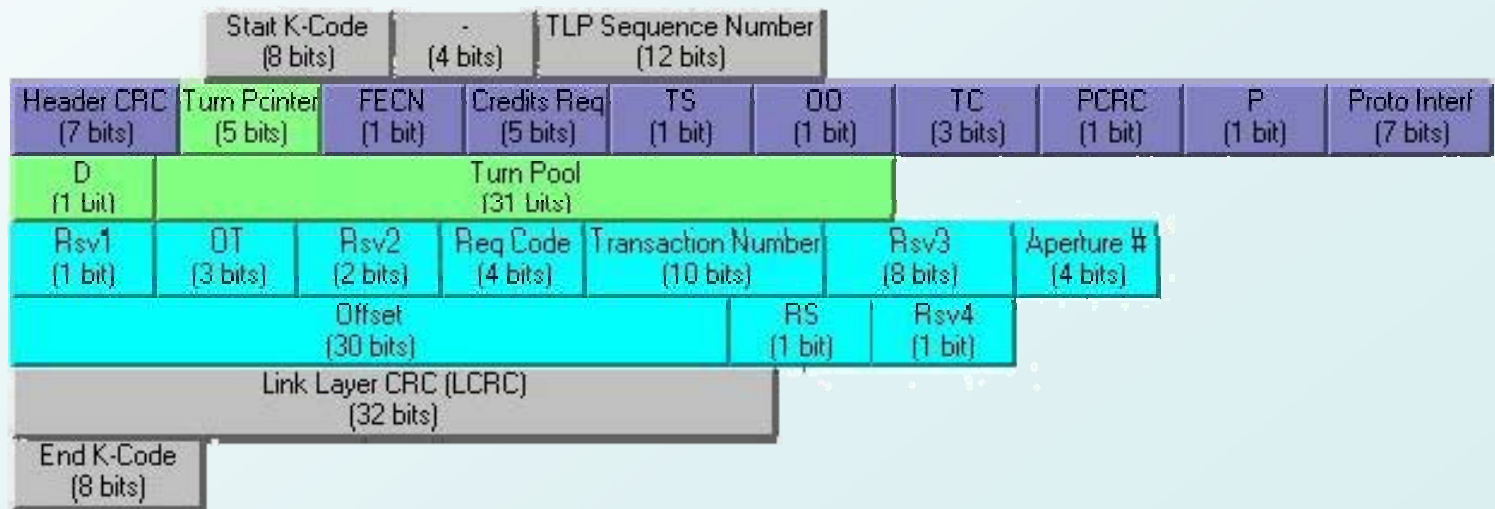
Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Used for the communication between the FM and device managers
 - Access to the capability structures
 - Types
 - *Read request*: Obtains information from a capability
 - *Read completion*: Contains the information requested by the FM
 - *Write*: Allows the FM to modify any data in the device



Introduction to Advanced Switching

Event reporting (PI-5)

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

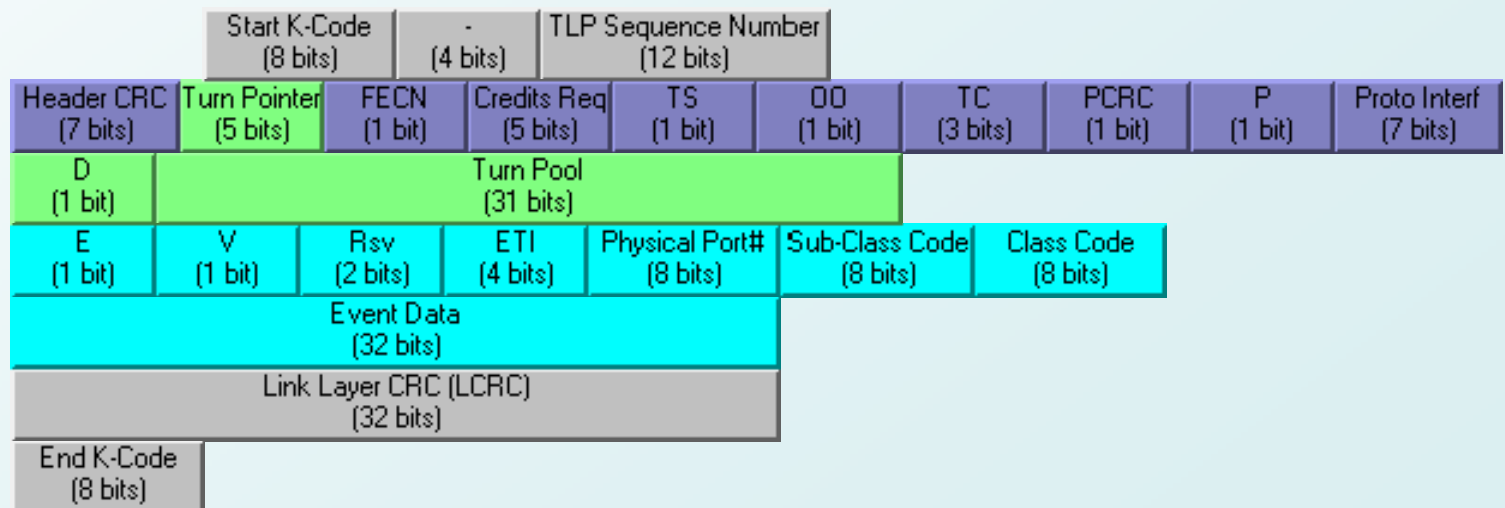
Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- The specification provides an event-reporting mechanism to notify topology changes
 - A particular DM can report the FM about a change in the state of a local port, through a PI-5 packet
 - The FM must update the set of fabric routes





Agenda

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Introduction to Advanced Switching
 - Architecture
 - Fabric Management
- Topology Discovery
 - Motivation
 - Serial Discovery
 - Parallel Discovery
- Performance Evaluation
 - Simulation Model
 - Simulation Methodology
 - Results
- Conclusions and Future Work



Topology Discovery

Motivation

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Gathering the topological information
 - Fabric initialization
 - Detection of a topological change
- General behavior
 - If the FM receives general information about device, it must inject more packets to obtain information about each port
 - If it is discovered a new active port, the FM must start discovering the new device
- Algorithms evaluated
 - *Serial Packet* (ASI-SIG)
 - *Serial Device*
 - *Parallel Device*

Topology Discovery

Serial Discovery

Advanced Switching

- Architecture
- Management

Discovery

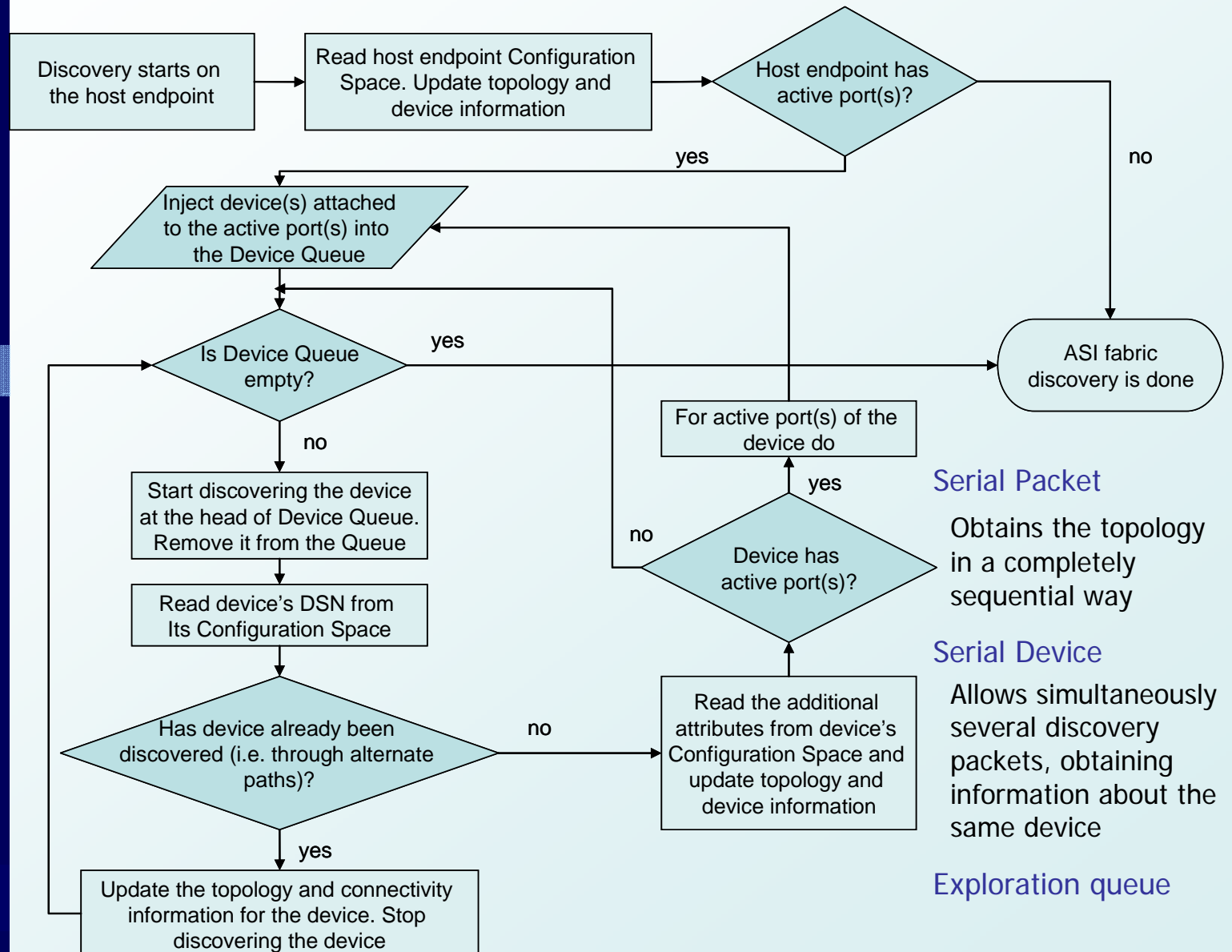
- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work



Topology Discovery

Parallel Discovery

Advanced Switching

- Architecture
- Management

Discovery

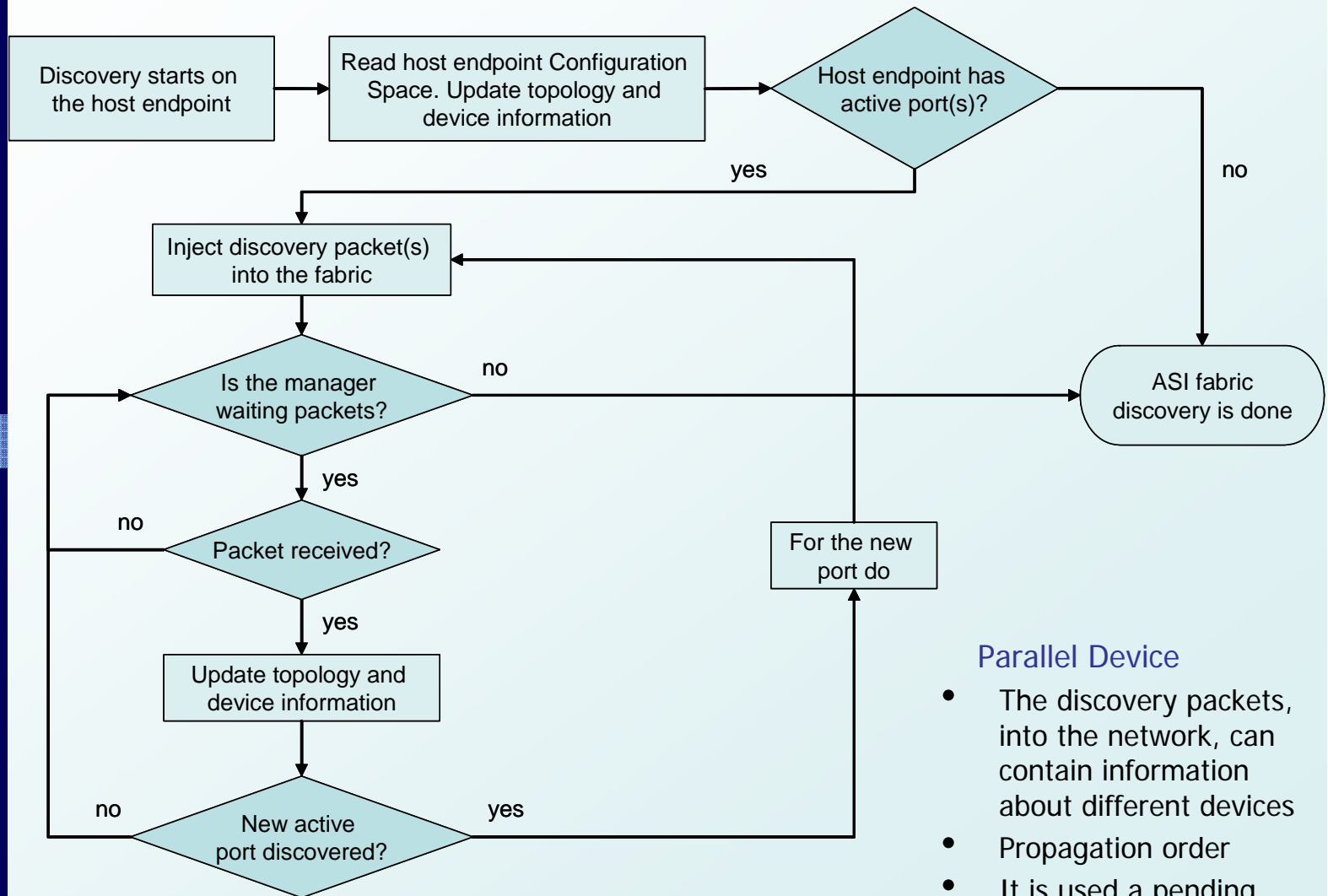
- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work



Parallel Device

- The discovery packets, into the network, can contain information about different devices
- Propagation order
- It is used a pending packets table

Topology Discovery

Serial and parallel ideal behaviors

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

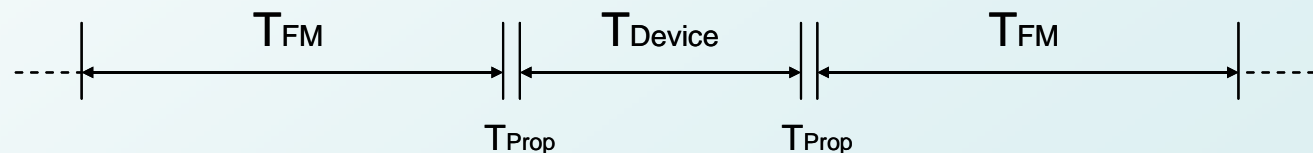
- Model
- Methodology
- Results

Conclusions

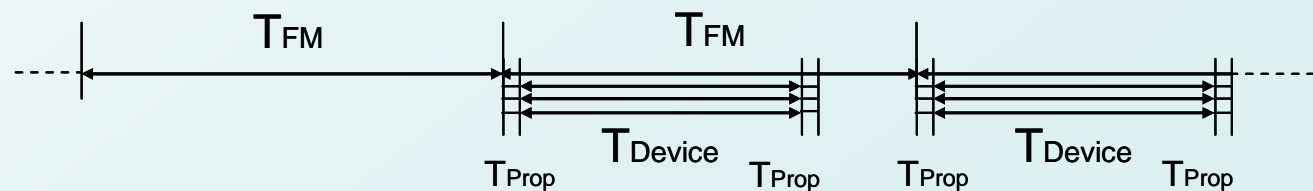
- Conclusions
- Future Work

- T_{FM} and T_{Device} : Processing time of each packet received
- T_{Prop} : Time to transmit a request/response packet
- *Serial behavior*: the FM is idle while it is waiting for the corresponding packet response
- *Parallel behavior*: while the devices process request packets, the FM can process a response packet

Serial behavior



Parallel behavior





Agenda

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Introduction to Advanced Switching
 - Architecture
 - Fabric Management
- Topology Discovery
 - Motivation
 - Serial Discovery
 - Parallel Discovery
- Performance Evaluation
 - Simulation Model
 - Simulation Methodology
 - Results
- Conclusions and Future Work

Performance Evaluation

Device Model

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

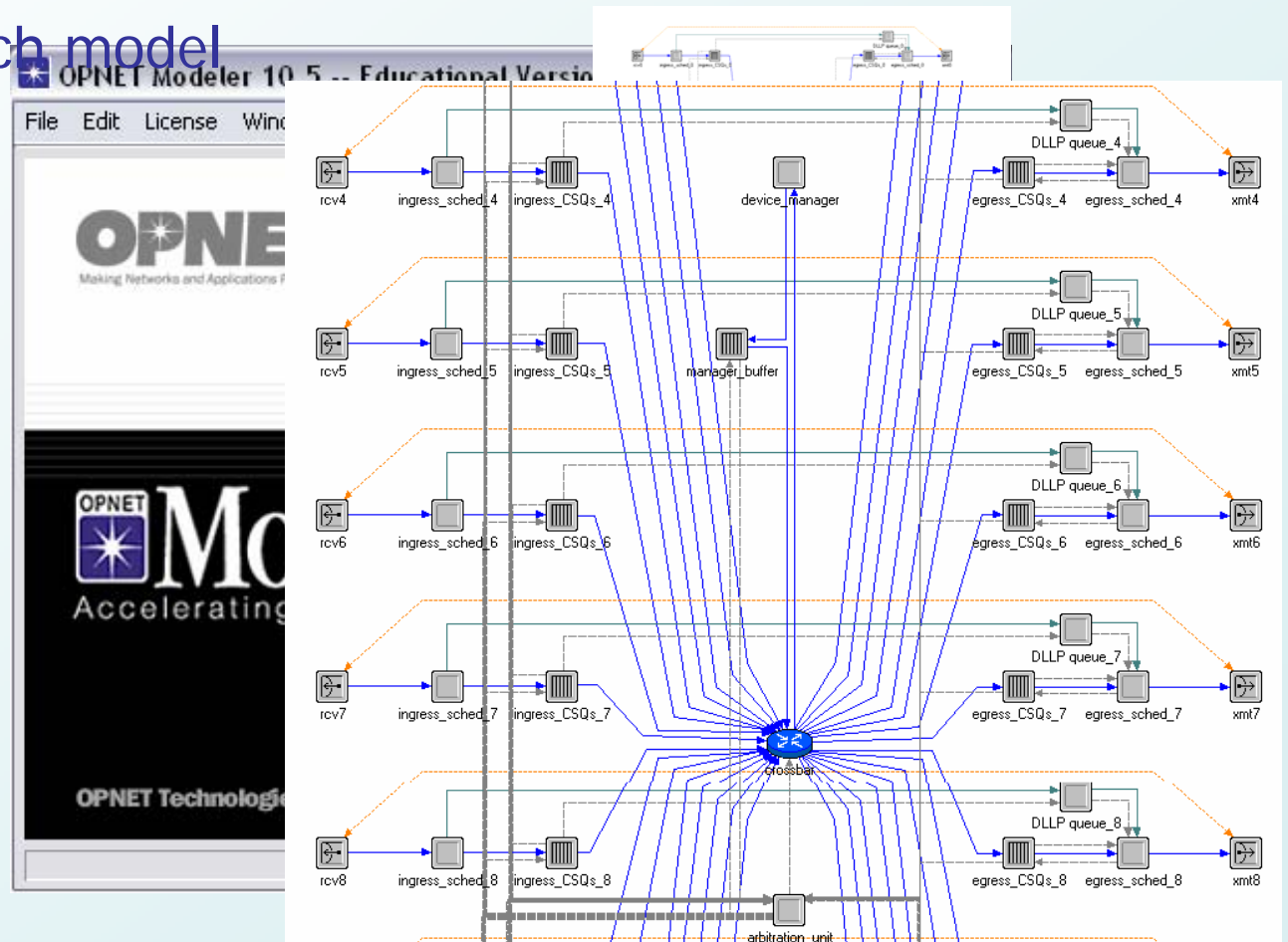
Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Model developed with the OPNET Modeler
 - Development environment of communications fabrics and distributed systems
- Switch model



Performance Evaluation

Simulation methodology

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

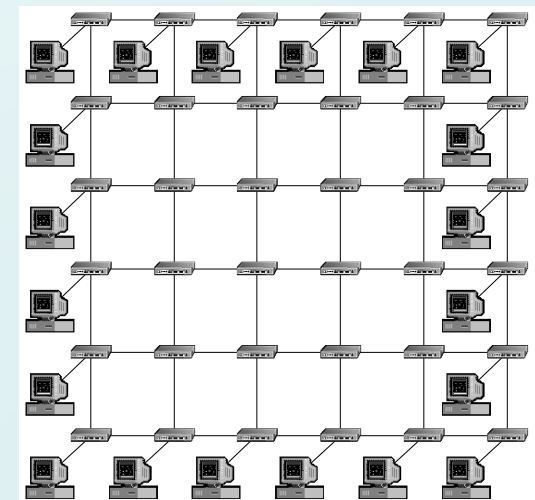
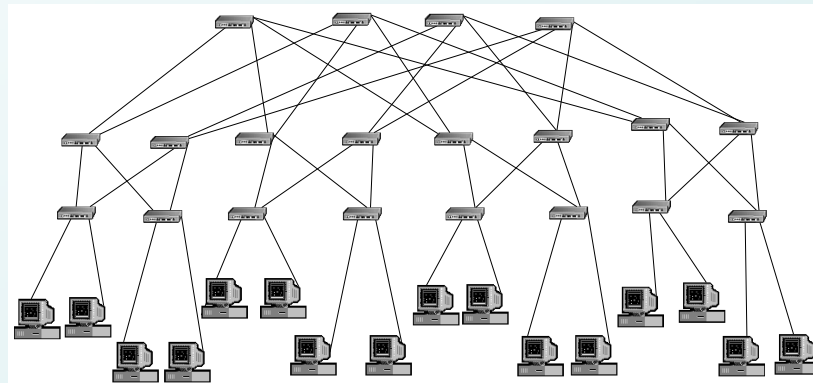
Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Evaluation of several regular topologies
 - Meshes, tori, and fat-trees
- For each simulation
 - A switch is randomly activated or deactivated
 - Some information gathered
 - Discovery time
 - Number of packets used
 - ...



Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

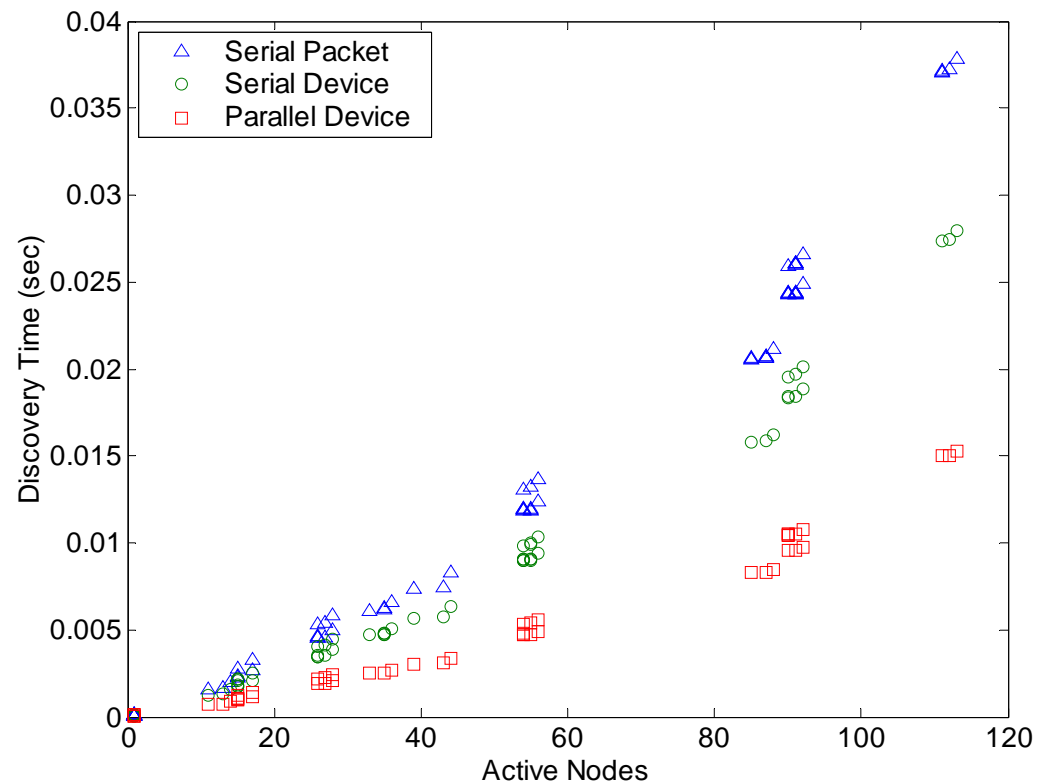
- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

Performance Evaluation Results

- Discovery time versus active nodes
 - Results show a significant improvement in the time required to discover the topology
 - This improvement is scalable
 - The behavior does not depend on the type of topology



Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

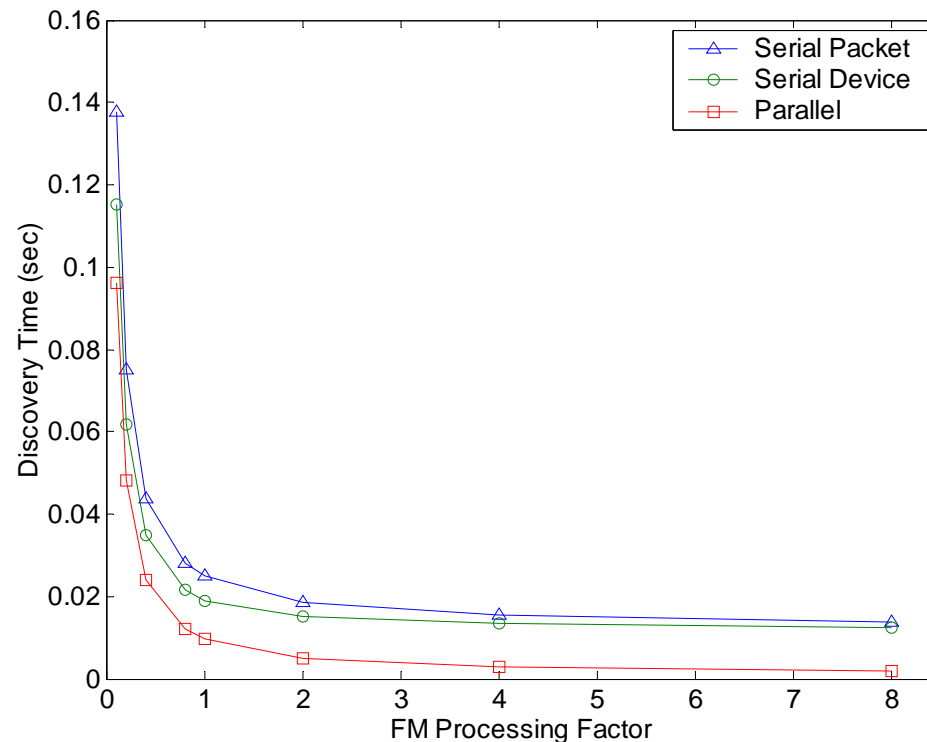
- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

Performance Evaluation Results

- Varying the Performance of the FM (*8x8 mesh*)
 - As the processing factor grows up, the discovery time decreases
 - The difference between the serial and parallel implementations briefly increases



Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

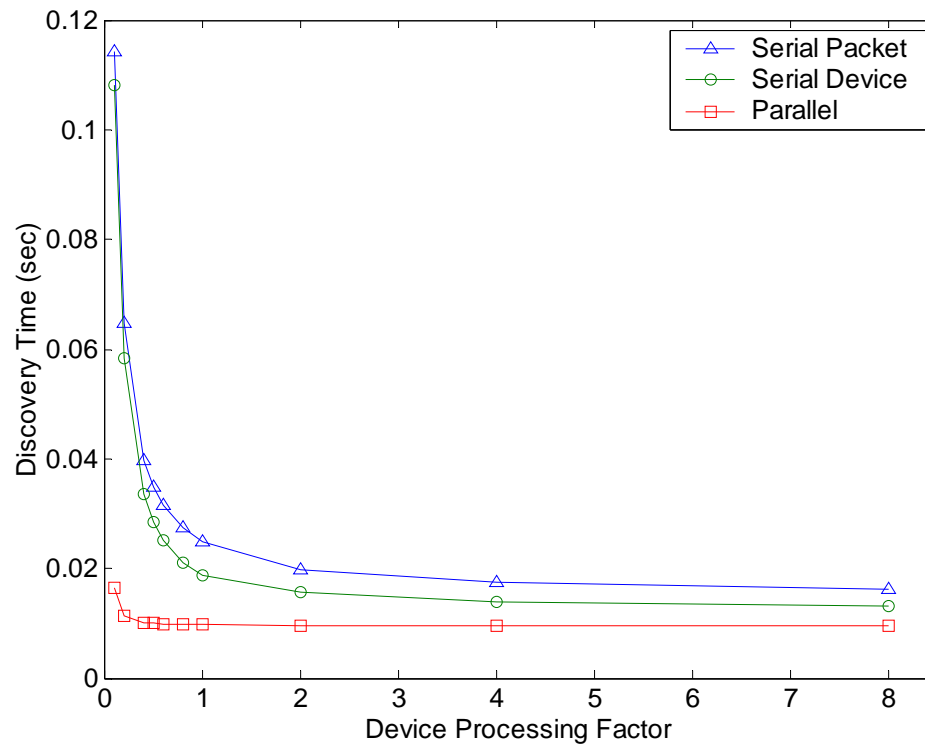
- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

Performance Evaluation Results

- Varying the Performance of the DM (8x8 mesh)
 - The *Parallel Device* algorithm is not affected by the time consumed by the devices



Performance Evaluation Results

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

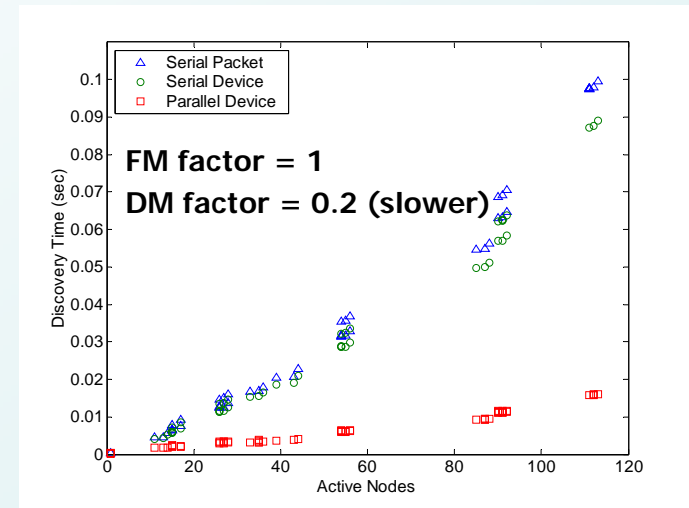
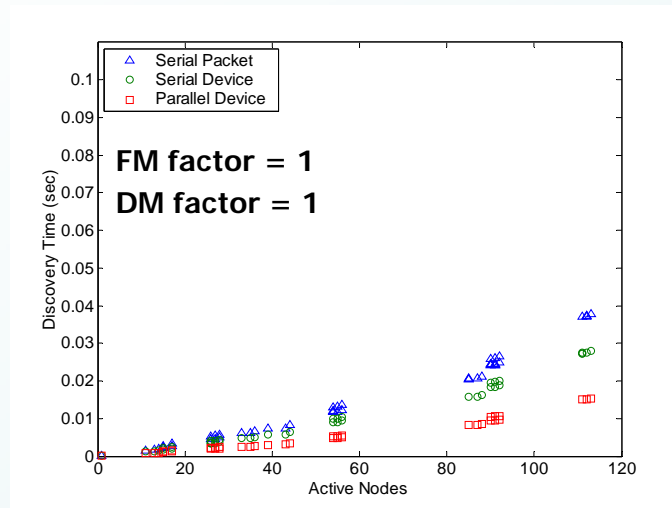
Evaluation

- Model
- Methodology
- Results

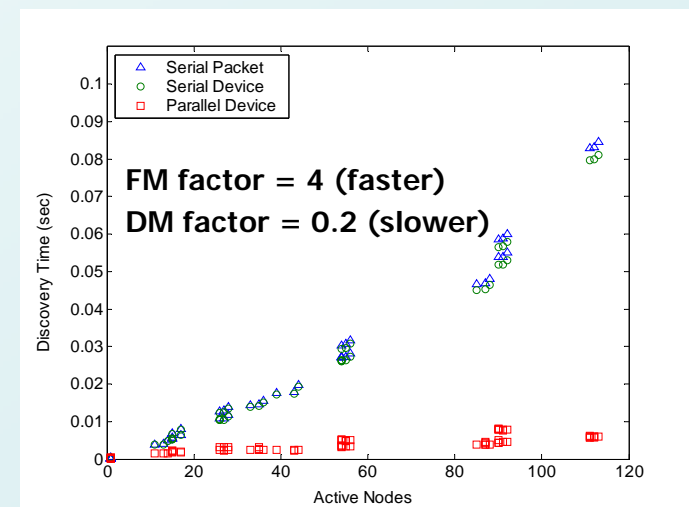
Conclusions

- Conclusions
- Future Work

- Varying the processing factors of the entities



- First plot
Same results than before
- For slower DMs
The improvement increases
- For any FM processing factor
The improvement is maintained





Agenda

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Introduction to Advanced Switching
 - Architecture
 - Fabric Management
- Topology Discovery
 - Motivation
 - Serial Discovery
 - Parallel Discovery
- Performance Evaluation
 - Simulation Model
 - Simulation Methodology
 - Results
- Conclusions and Future Work



Conclusions and Future Work

Conclusions

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Several mechanisms to discover the topology compared
 - The *Parallel Device* algorithm obtains the expected improvement compared with the serial ones
 - The improvement is more noticeable as the performance of fabric devices are slower
 - The improvement is maintained independently of the performance of the FM



Conclusions and Future Work

Future Work

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- Design of a complete management mechanism for the *Advanced Switching* technology
 - Path computation and distribution (*developing*)
- Tackling the problem of dynamically distributing new paths to fabric endpoints after the occurrence of a change
 - Reducing the negative effects of applying an static mechanism

Questions?

Advanced Switching

- Architecture
- Management

Discovery

- Motivation
- Serial
- Parallel

Evaluation

- Model
- Methodology
- Results

Conclusions

- Conclusions
- Future Work

- ***Antonio Robles-Gómez***

- arobles@dsi.uclm.es

- ***Aurelio Bermúdez***

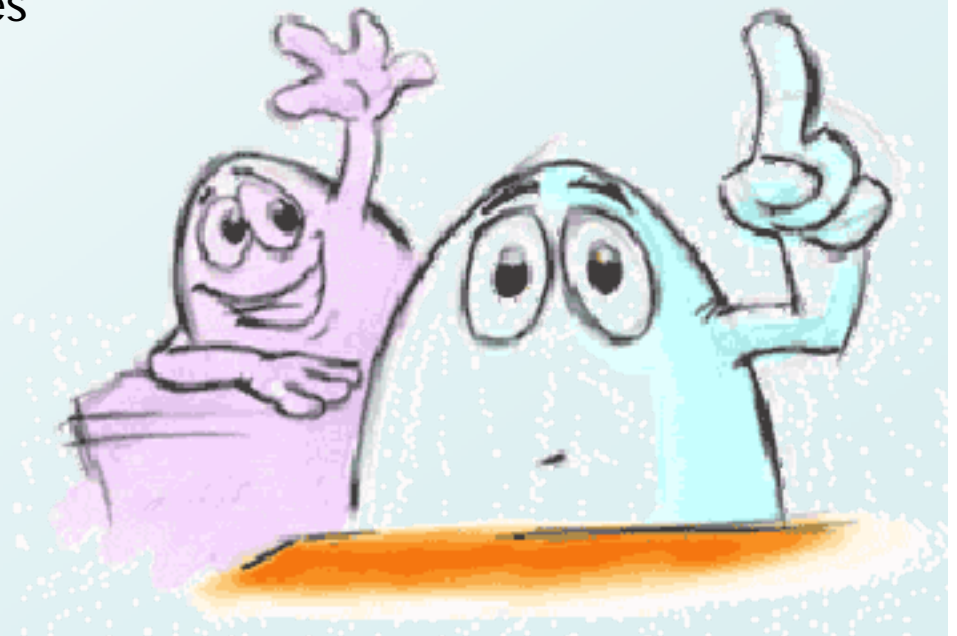
- abermu@dsi.uclm.es

- ***Rafael Casado***

- rcasado@dsi.uclm.es

- ***Francisco J. Quiles***

- paco@dsi.uclm.es





Implementing the Advanced Switching Fabric Discovery Process

Antonio Robles-Gómez
Aurelio Bermúdez
Rafael Casado
Francisco J. Quiles

**Universidad de
Castilla – La Mancha**

