

Power management in embedded multi-core architectures

Karim Ben Chehida, Raphaël David

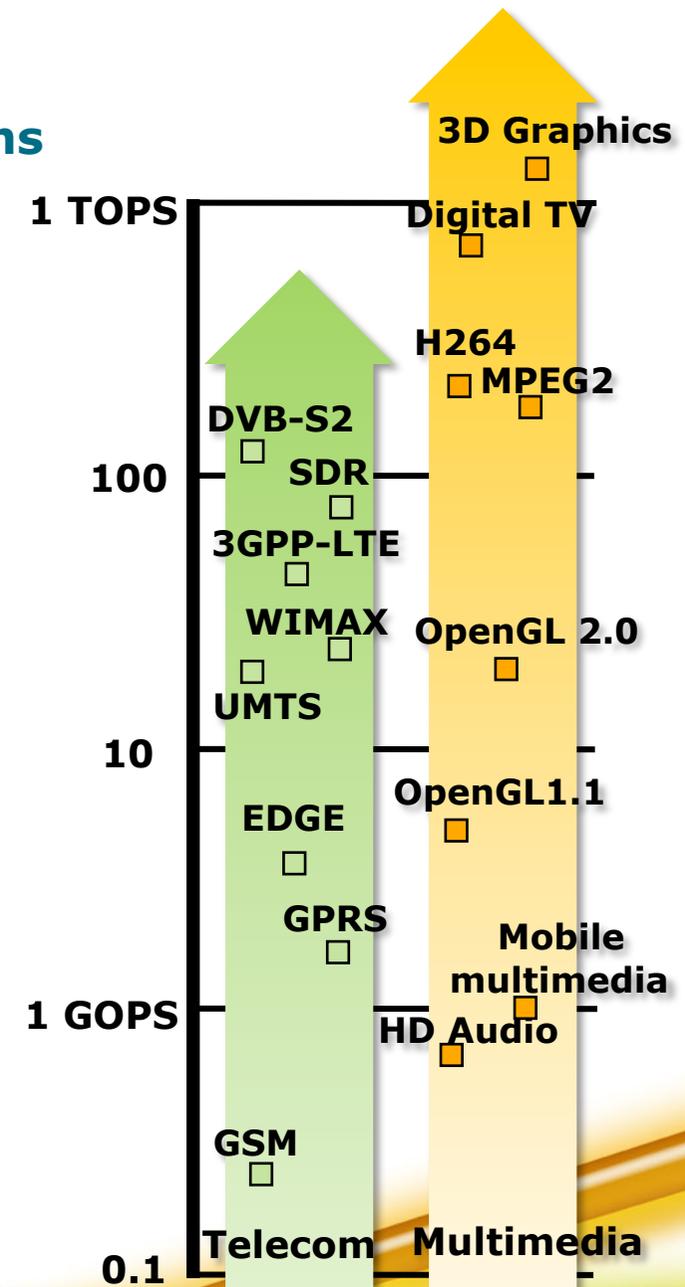
Karim.benchehida@cea.fr

CEA LIST, Saclay

- **Application and architectural trends**
- **The SCMP architecture**
- **Former LIST's power management solutions**
- **New challenges**

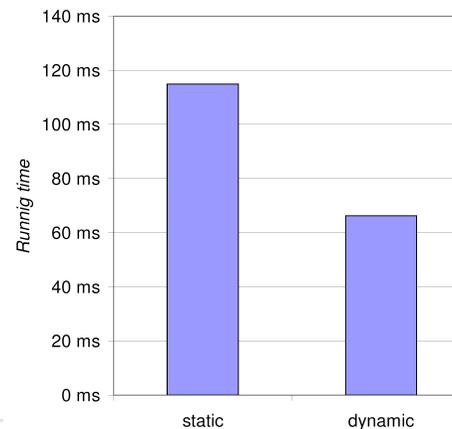
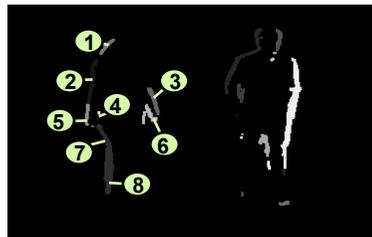
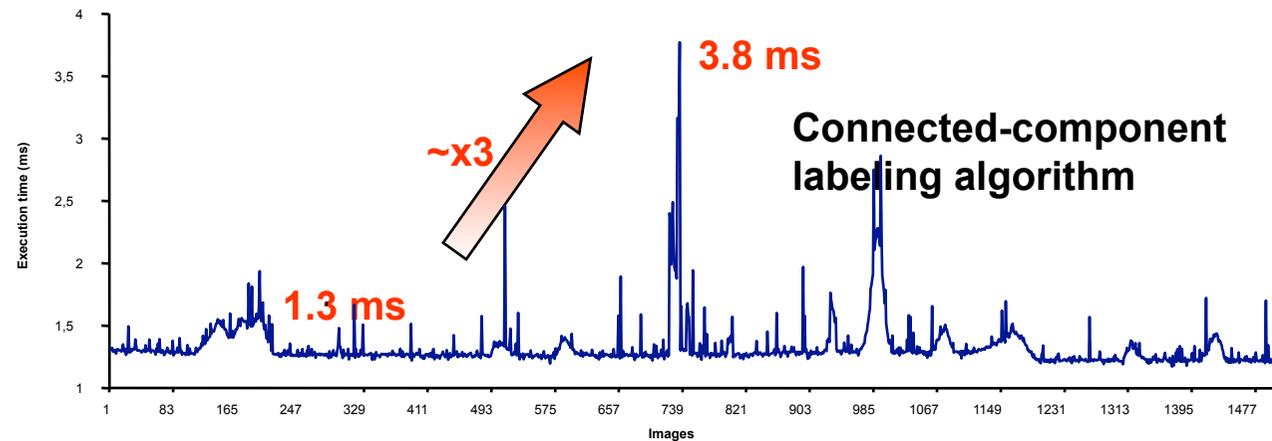
Embedded Applications trends

- Embedded systems must support various applications
- They need more computing power



Task management in MPSOC architectures for Embedded applications

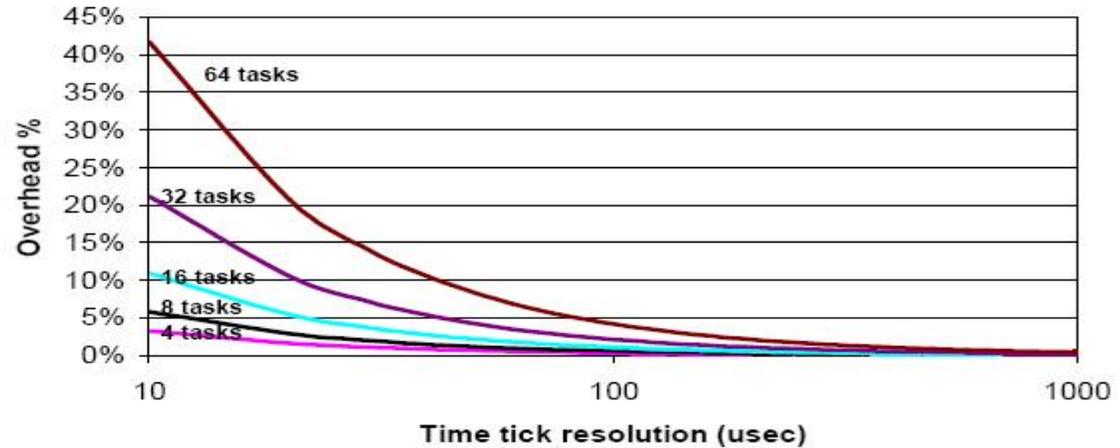
- **Applications are more and more dynamic**
 - Dynamic control flow
 - Data-dependent processing
- **System becomes less and less predictable**
 - Variability
 - Defects
 - Aging
- **Dynamic Load balancing is needed to improve processor utilization rate**



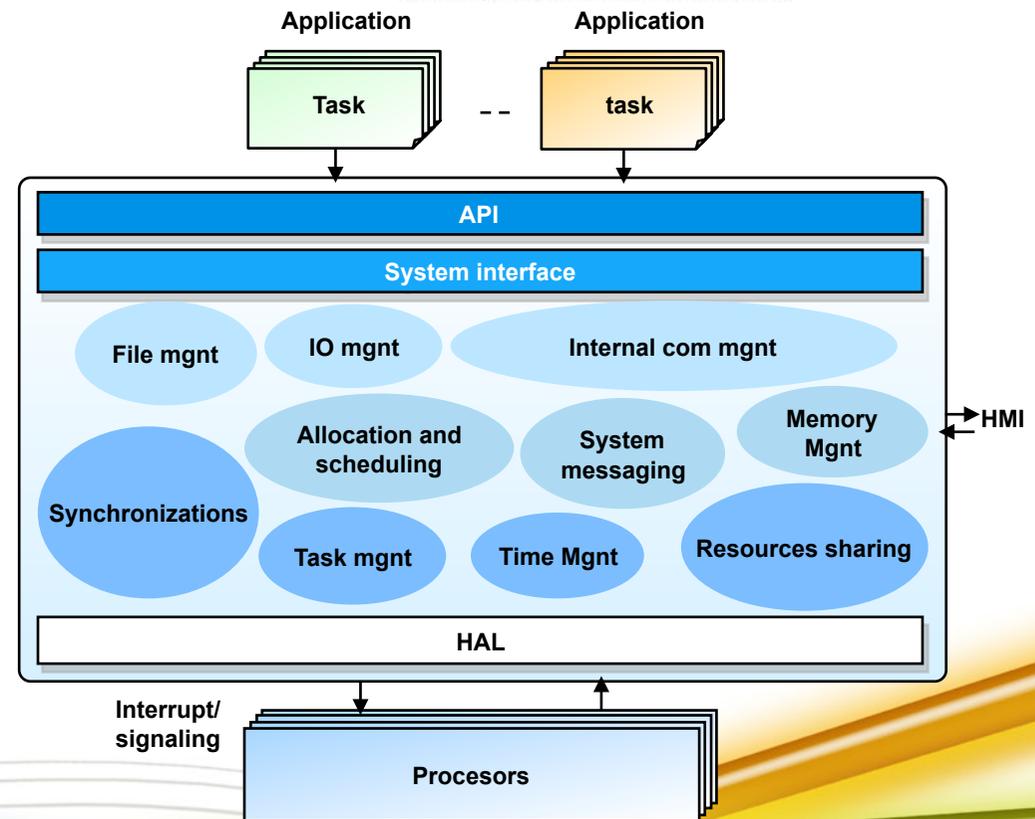
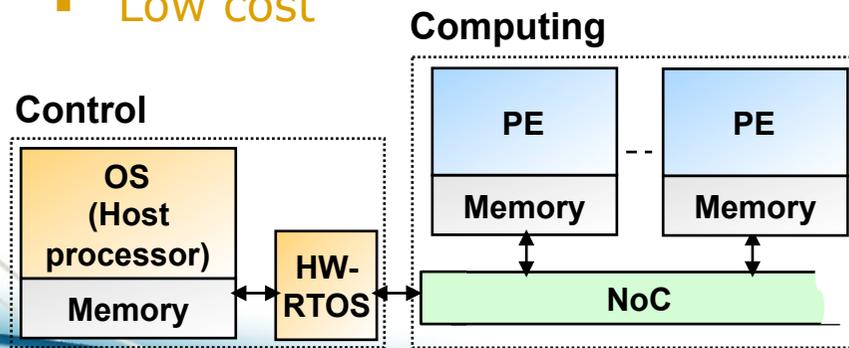
Connected component labeling execution time for a parallelization on 8 processors (128 tasks)

Task management implementation issues

- **Software management of tasks is however not for free**
 - Low reactivity
 - Low transistor and silicon efficiency
 - Overhead hardly predictable because of its dependencies regarding workload



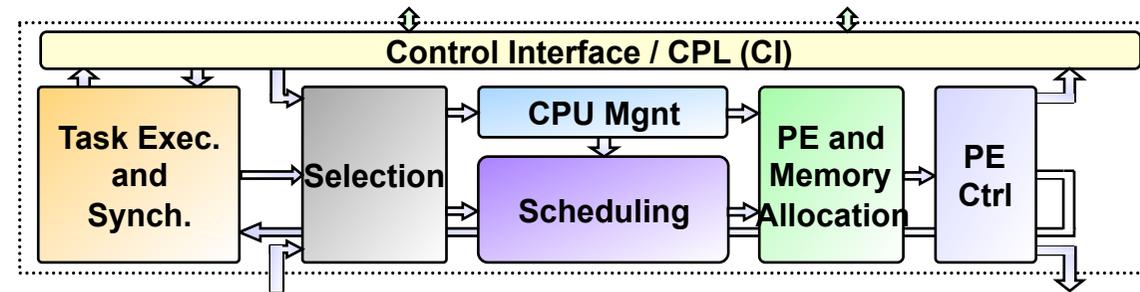
- **Benefits of hardware acceleration**
 - Overlapping between control and computation activities
 - Determinism
 - Reactivity
 - Low cost



Hardware support for task management

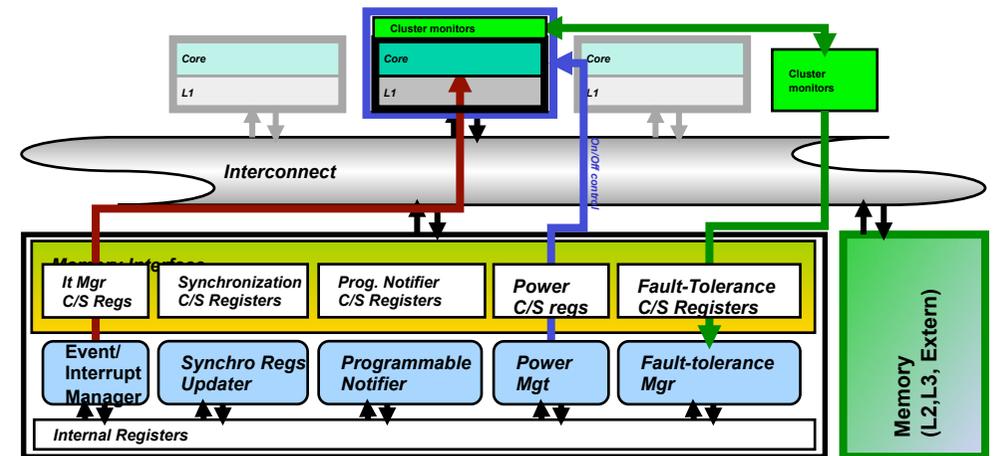
- **Full Hardware solution**

- For asymmetric approaches
 - May need several 100kgate but support very aggressive real time scheduling approaches



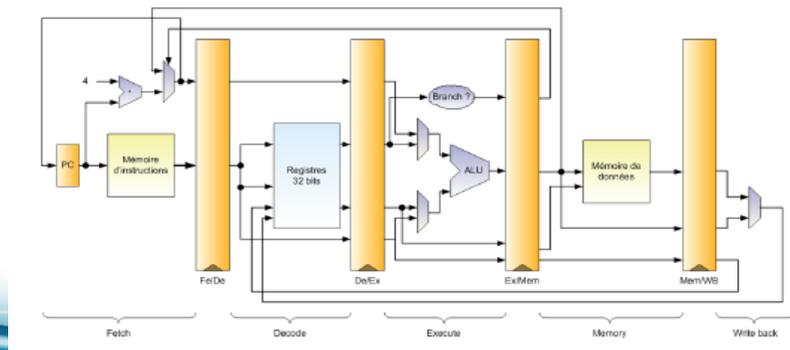
- **HW accelerated SW solution**

- For SMP systems
 - Less than 100kgate
 - Not so smart but allows secured sharing of system information and centralized signalization schemes



- **Mixed approach**

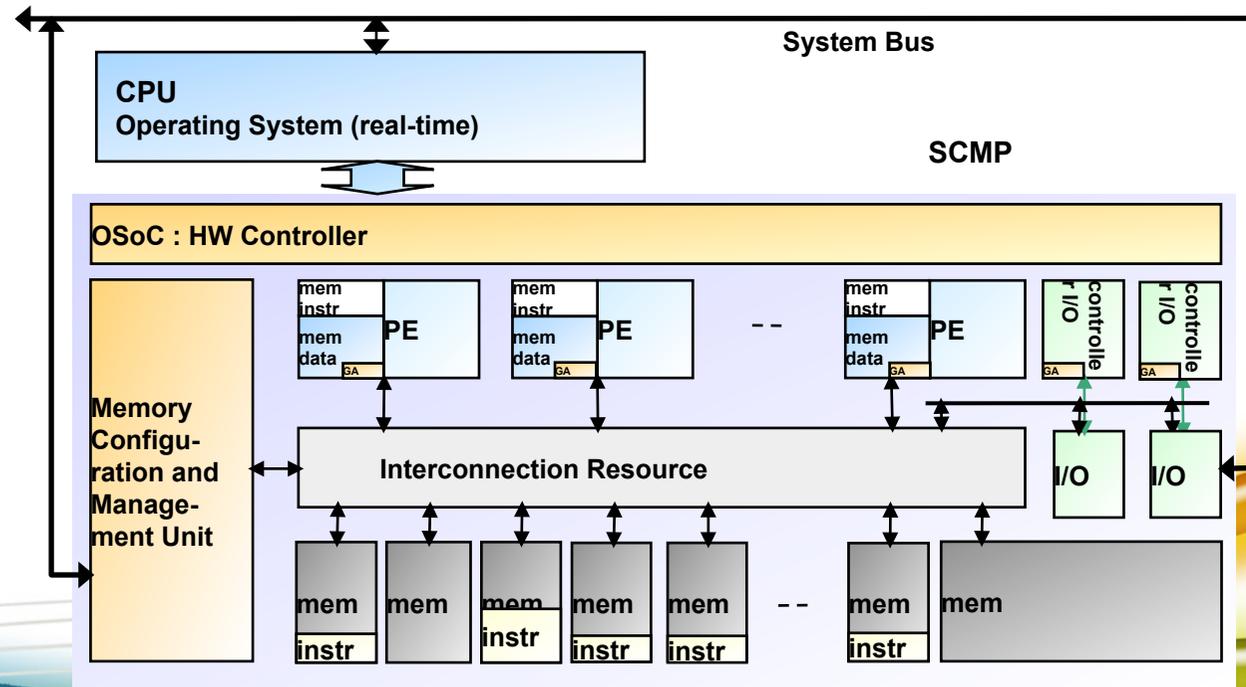
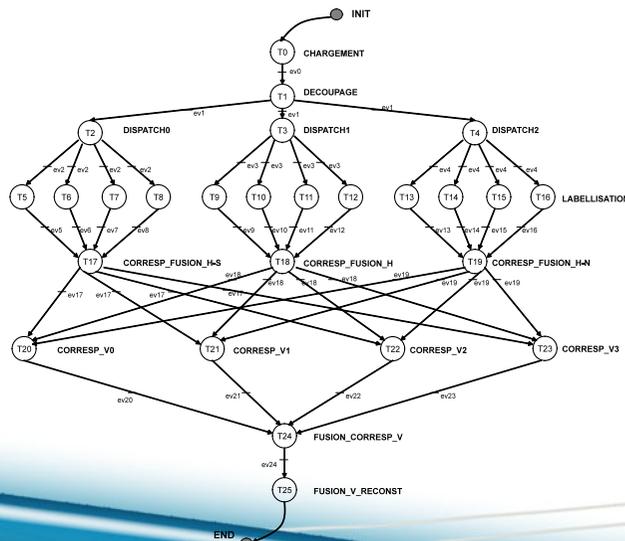
- For multi-purpose asymmetric approaches
- Based on a small RISC processor with optimized coprocessor interface



SCMP: a new architecture for dynamic applications

- **New execution and programming model**

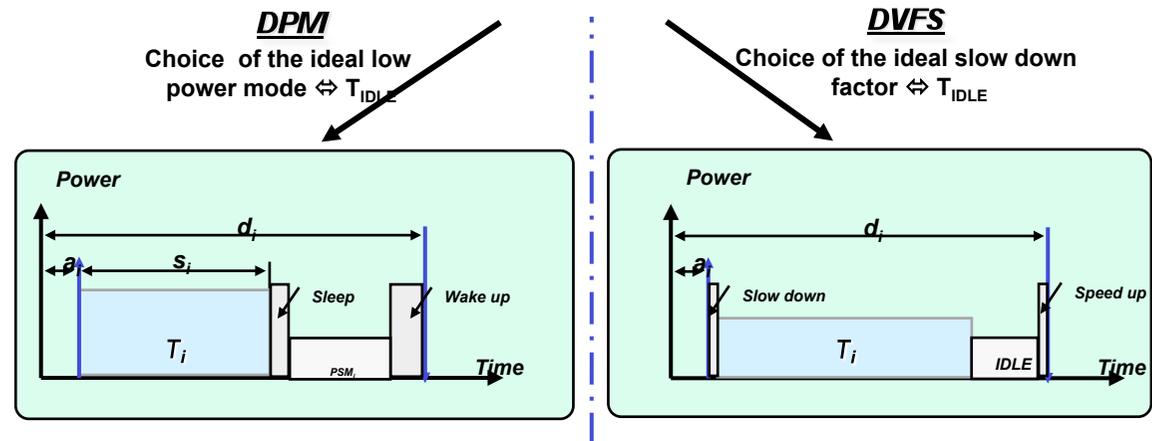
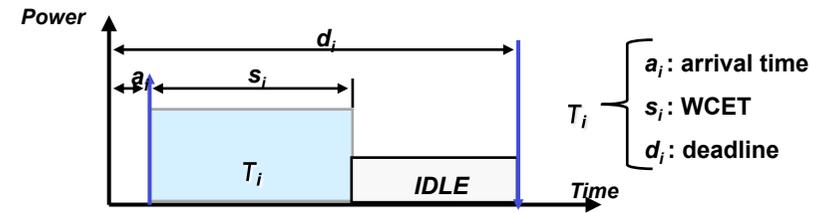
- Simplify the task parallelism management
- Optimize the PE occupancy (load balancing)
- Asymmetric architecture with a global control
- Fast preemption and migration of tasks
- Explicit separated execution between control and computation
 - Low control overhead (HW accelerator)
- **Specific memory management**
 - Physically distributed and logically shared
 - Write exclusive accesses
 - Data and instruction virtualization



Power Management strategies in SCMP

- Use all hardware support

- Idle modes
- Variable Voltage/frequency



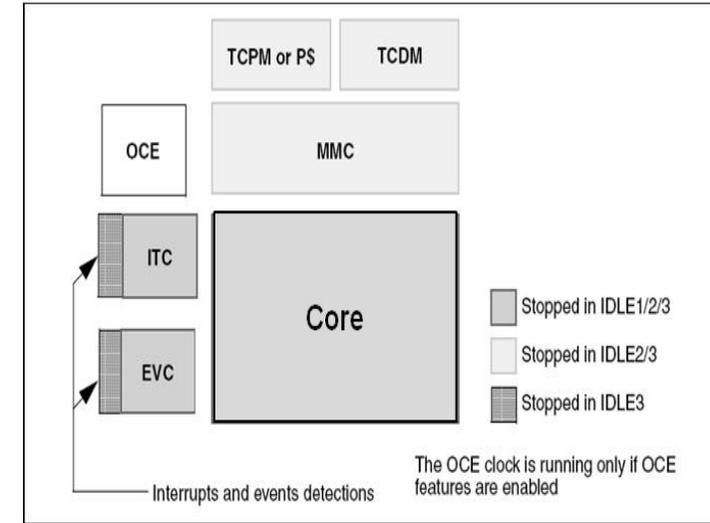
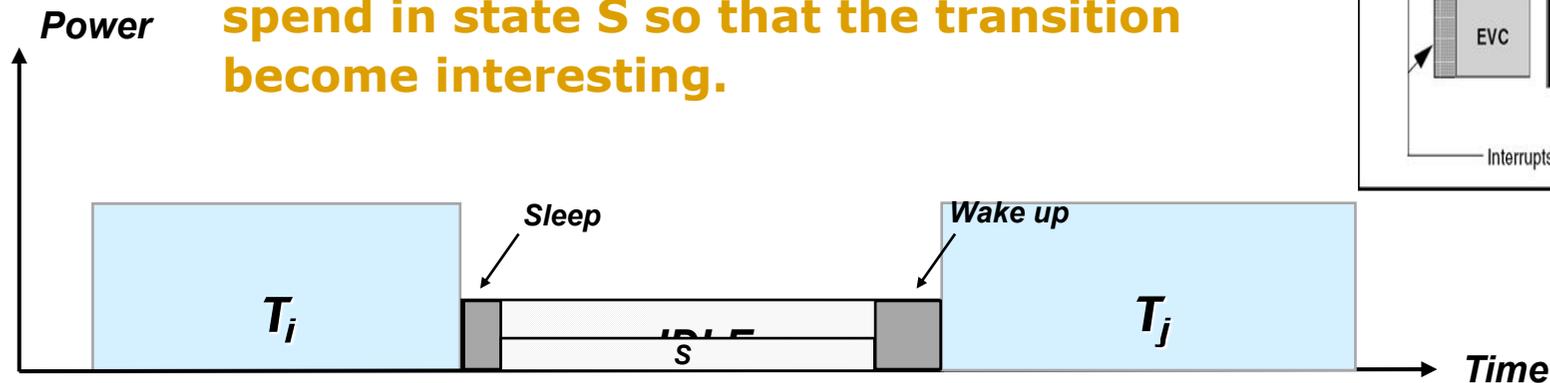
- Adapt power management strategies to application needs

- Real Time
- Best effort in dataflow applications

Idle mode management

– Low Power State Parameters :

- **PS = State Power Consumption.**
- **ETR = Transition energy consumption**
- **TTR = Transition latency**
- **TBE = Break-Event Time: minimum time to spend in state S so that the transition become interesting.**



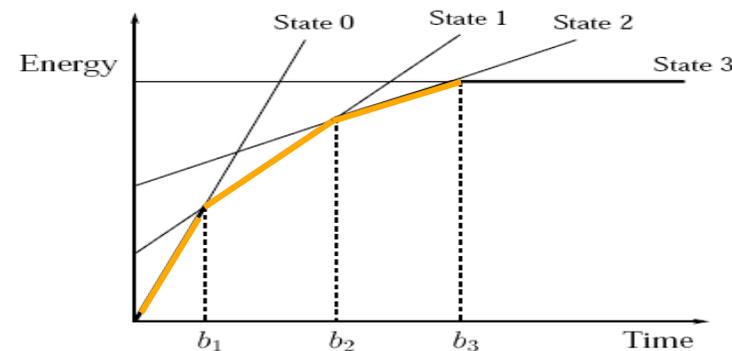
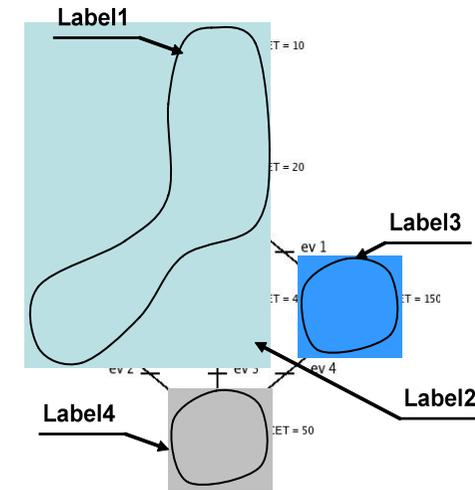
① What is the Idle period on processor P (predict the future !!!)

② What is the ideal low power mode S corresponding to the predicted idle time ???

Idle mode management

- **In real Time Systems**

- Characterize offline (using WCET) the variation of the parallelism rate of each application
- Detect online these variations and deduce the Idle periods
- Activate the ideal low power mode (modified LEA) and predict the corresponding awakening time



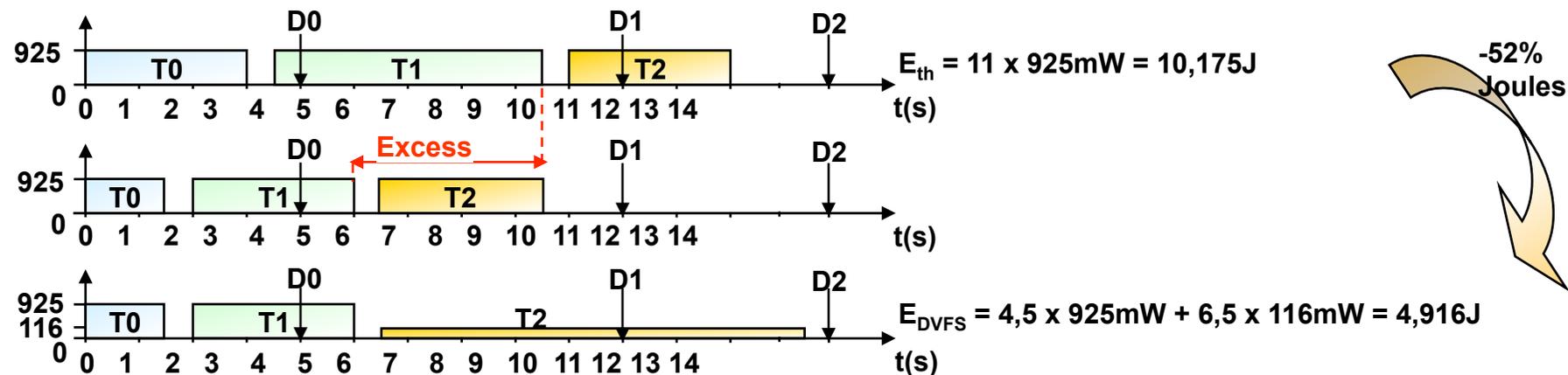
- **In dataflow mode**

- Switch to idle mode as soon as data buffer reaches a critically low filling rate threshold
- Awake whenever data buffer has been filled enough

Voltage and frequency scaling

- **In real time applications**

- Take benefits of slack times
 - Limit the amount of Voltage/frequency modification
 - Accumulate slacks until being able to activate DVFS for a long period of time
- Assign the slack to the next task on the same processor (% resource dependency) to avoid wasting slack time when reaching joining points in applications

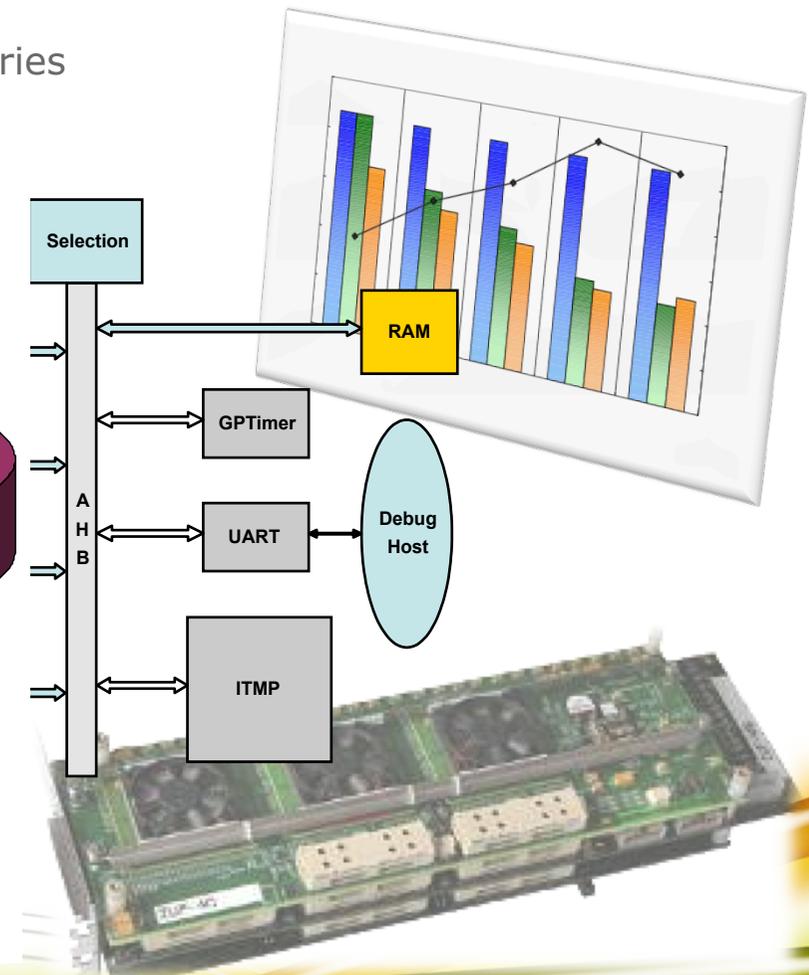
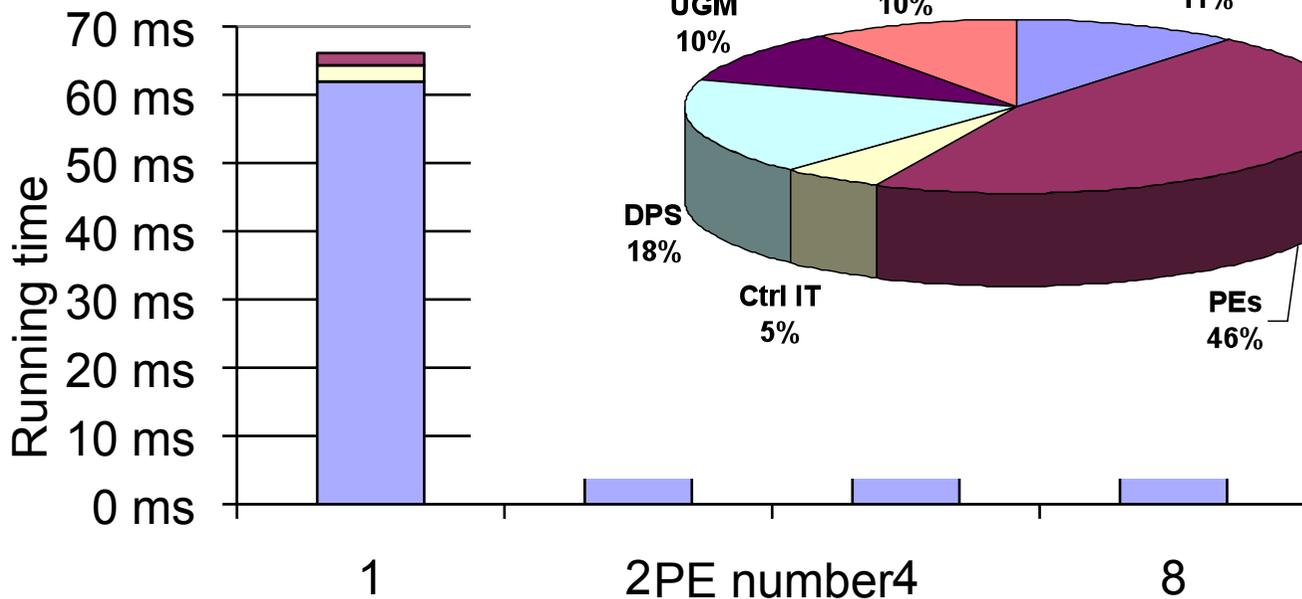


- **In dataflow applications**

- Adapt production and consumption rates to avoid stalls when pipeline is unbalanced

- **Complete FPGA prototype**

- 75 MHz prototypes on FCM4 boards from Scaleo Chip (StratixIII based)
- Including OSoC and additional power management accelerators
 - For real time DPM and DVFS management
- 4 Processing Elements
 - Based on sparc processors
 - With Data Prefetch Engines + standard cache memories
- Multibanked memory system
 - With HW support for dynamic allocation



- **Application and architectural trends**
- **The SCMP architecture**
- **Former LIST's power management solutions**
- **New challenges**

Application domain specific systems

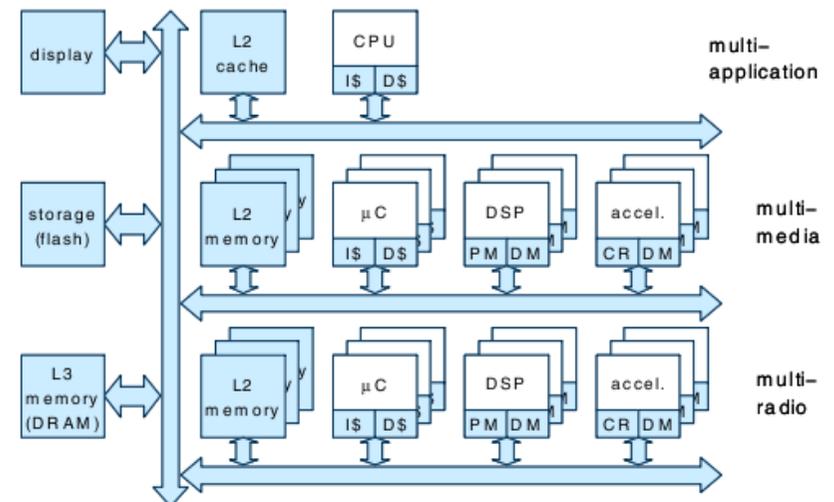
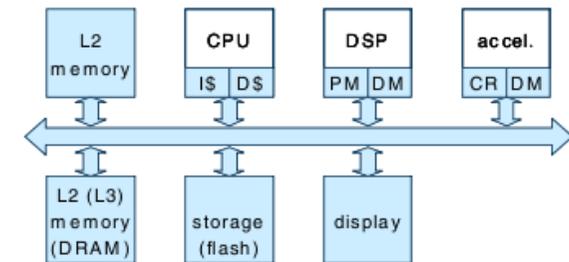
- **Functional Heterogeneity**

- **How to support the load balancing in this context ?**

- Data access modes are IP specific
- Binaries are IP specific
- Performance predictions are IP specific
- ...

- **How to manage the power in this context ?**

- Very simple and drastic (ON-OFF)
- Selective clock and power gating at application sub-system level

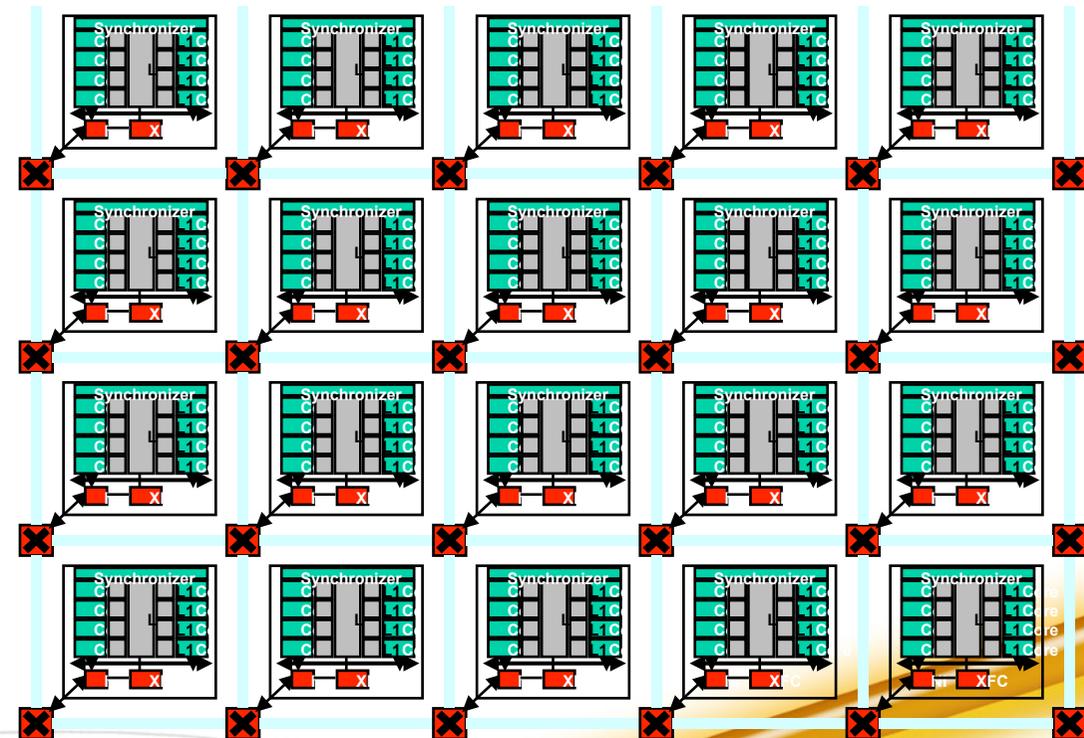


From multi- to many-cores

- **A many-core fabric**
 - Homogeneous or Heterogeneous resources
 - Not clustered by application
- **The fabric dynamic management allows to support:**
 - Faults (remanent or transient, aging...)
 - Complex applications not fully predictable at compile time
 - Load balancing
 - Power Management

- **Possible approaches for complexity management**

- **Globally static locally dynamic management:**
 - Dynamic application deployment,
 - no migration between clusters
 - **Globally dynamic management**
 - Dynamic application deployment,
 - possible inter-cluster migration
- Or
- Dynamic task deployment (at task creation)



- **Advanced process variability**

- From a design based on the worst case to a design based on the average case with online adaptation
 - For each processor, the frequency is evaluated and corrected online
- Implies a GALS functioning model
 - DVFS modes support tends to be generalized in complex systems
 - Need a low footprint HW support for voltage and frequency adaptation (from DC-DC → VDD Hopping)
- High performance heterogeneity
 - Need to reconsider performance homogeneity hypothesis in the large scheduling literature
 - Tighter coupling of Power management policies with the allocation phase:
 - * With the Load balancing policies
 - * With the system monitoring
 - * with the thermal management policies
 - Effect of the temperature on the static power consumption → overall power consumption
 - Effect of the (frequency, voltage) operating point on the temperature

- **The transition from multi-core to many core systems is not straightforward as far as the dynamic system management is concerned :**
 - **HW acceleration of control primitives can lead to high controllability**
 - At the lower hierarchy level for which it can be well dimensioned
- **Power management policies targeting many-core systems must take into account the process variability**
 - **Reactive and proactive policies can be considered**
- **Many core system dynamic power management is more tightly coupled to the Task allocation phase**
 - **With or without Load balancing techniques**
- **The effect of temperature on the power consumption and vice versa is not well estimated for 32nm and beyond ...**
 - **Benefits of coupling the Thermal and power management techniques...**