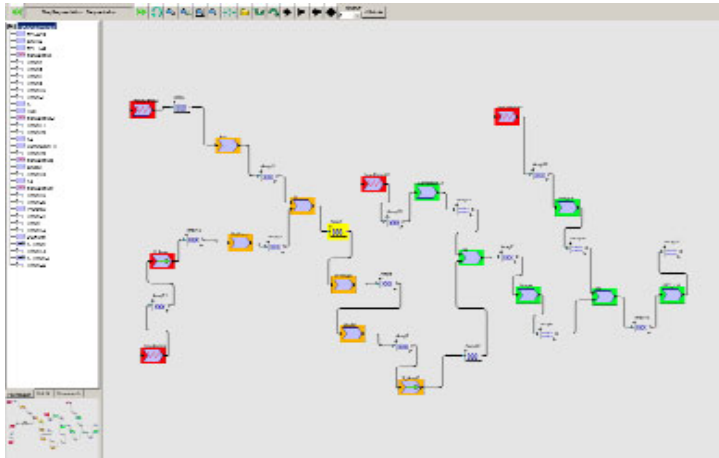




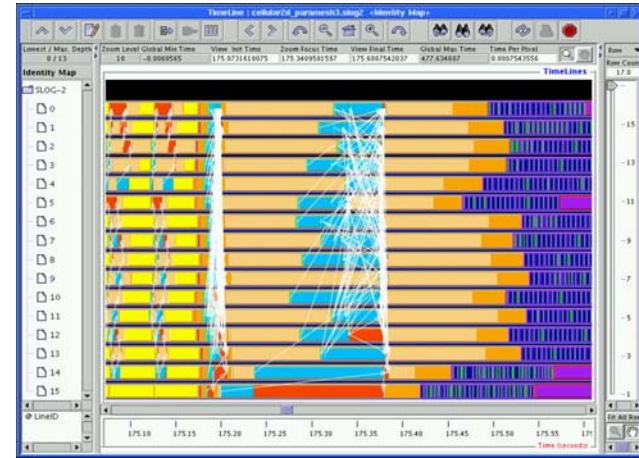
**Open HPEC Systems:  
Design and Profiling tools for  
multiprocessor signal processing applications using MPI**

HPEC 2005

# Typical HPEC application development cycle



Design



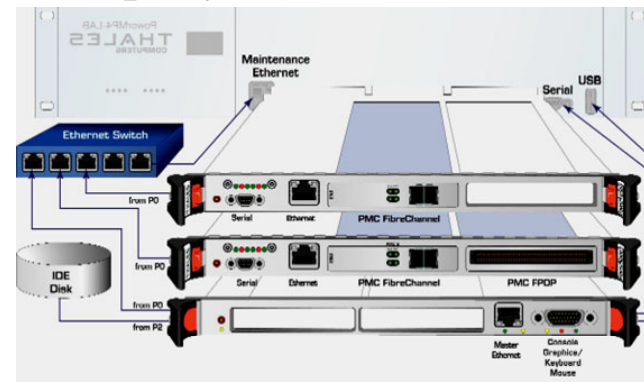
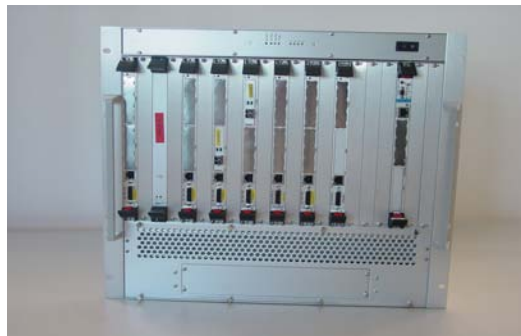
Validation

*Code Generation*



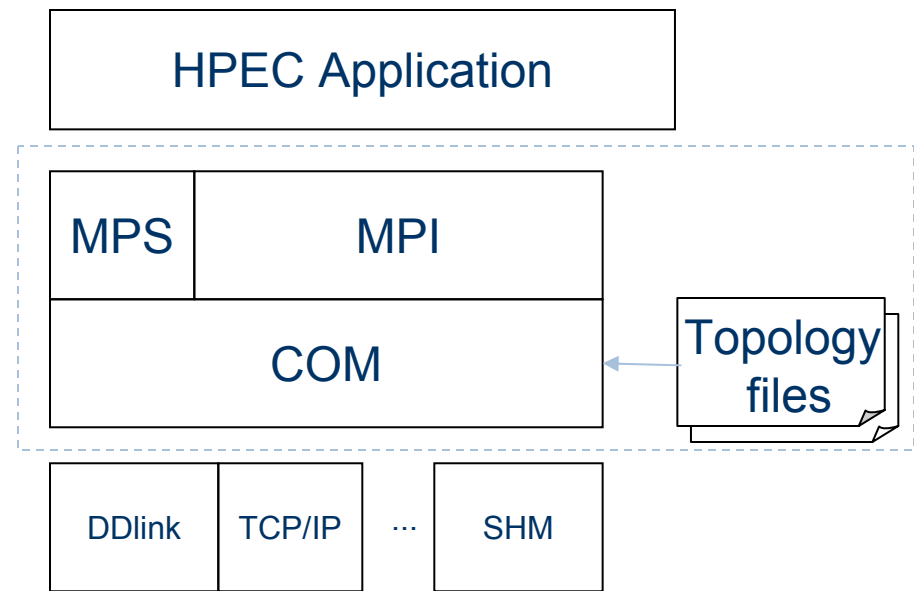
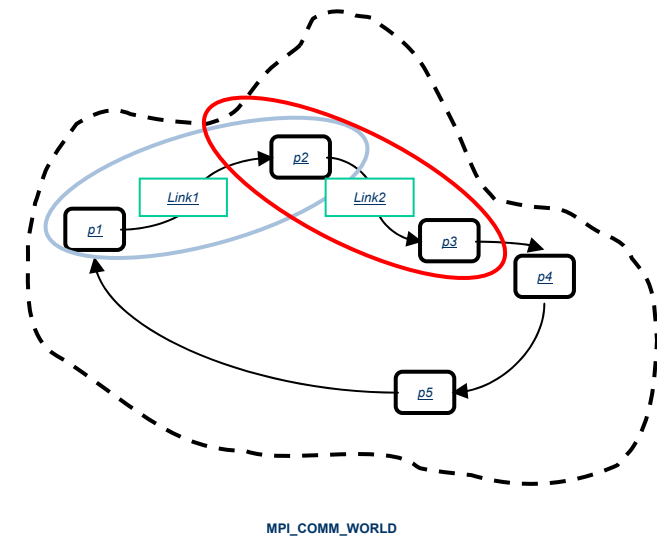
*Profiling*

*MPI based HPEC platform*



# Zero-Copy MPI Implementation

- MPI Services (MPS) side to side with MPI
  - MPI application source portability
  - Links/Connector relationship
    - Links to select communication channels (~ QoS)
    - Requests timeout support
  - Real-Time support
  - Real zero-copy transfer
    - Buffer Management API (MPS)
  - Heterogeneous machine support
    - Topology files outside application



T30527-b-en

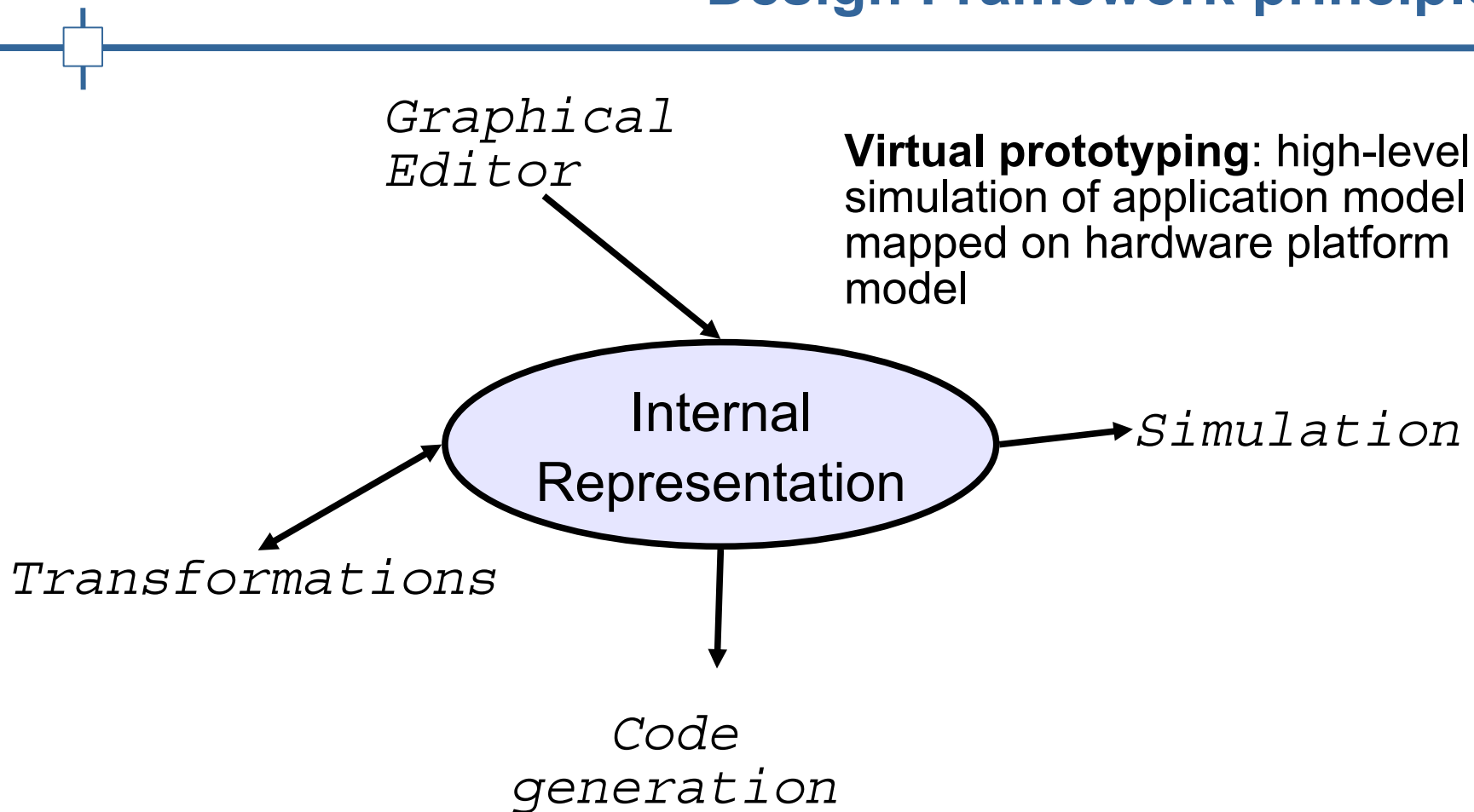


- Performance and determinism
  - Zero copy (e.g., RDMA on Fibre Channel) in addition to tcp/ip
  - Buffering control
- Deployment control
  - Topology file
    - Mapping of processes on processors
    - Mapping of logical channels (MPI communicators) on physical link
- Execution control
  - extended “mpirun” facility (includes “ps”, “kill” features over the application as a whole, or selected processes)

“User friendly” interfaces:

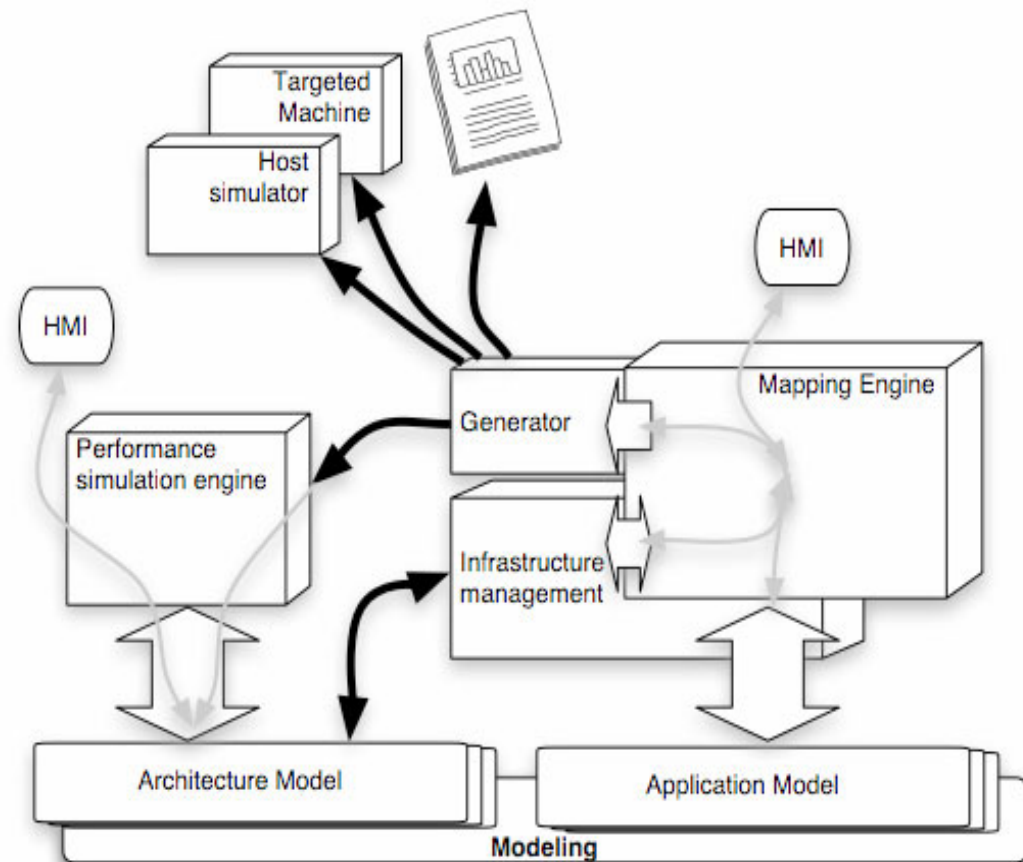
- standard MPI API
- simple XML configuration files
- command line interface

MPI API and XML files provide a for code generation by high-level tools



## 5 main modules

- Modelling (application, architecture)
- Management (profiles, projects, libraries)
- Mapping
- Performance simulation
- Generator (code, stimuli, doc)



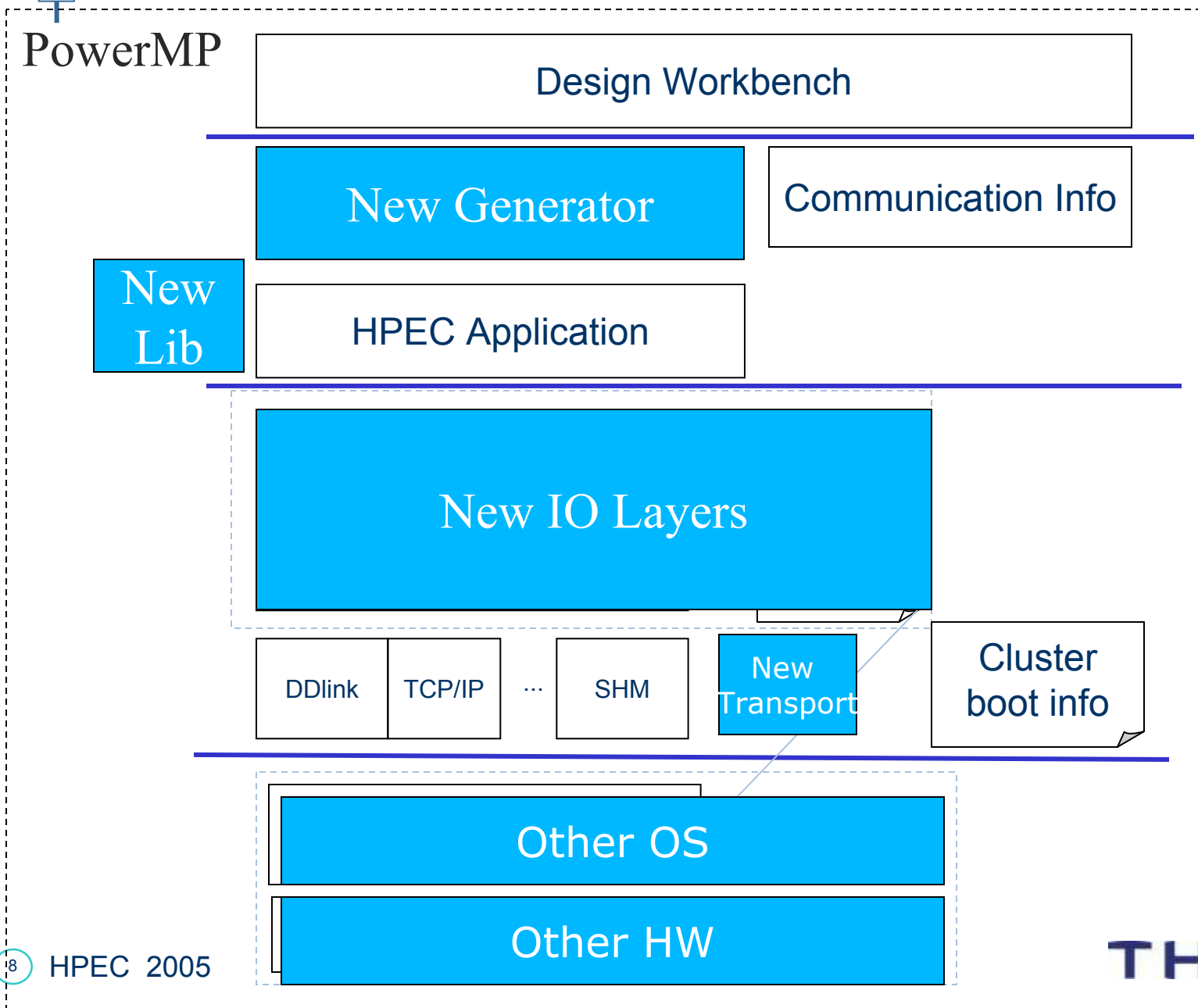
T30527-b-en





- Design support
  - Graphical interface, internal representation of application, and code generation “front end” implemented in Java within the Ptolemy framework (University of California at Berkeley, Pr. Edward Lee)
  - Code generation “back-end” implemented using XML transformations
  - Target may be any MPI library; tested on LAM (University of Indiana) on Pentium M/Linux
- Profiling
  - Optimised (zero copy) MPI implementation compatible with MPE profiling library and “Jumpshot” graphic display and analysis tool (Argonne National Laboratory)
- MPI
  - PowerPC G3, G4, G5 ; Pentium M ; Linux, LynxOS
  - VME 2eSST, Fibre Channel, RapidIO
- Cluster administration and control/monitoring
  - based on “Webmin”

# Designed for Customization







## Design support

- Improvement of data parallelism support – including data reorganisation (signal processing “corner turn”)
- High-level, discrete event, transaction level simulation
- Integrate in Eclipse framework
  - Cooperation's: UCB, Eclipse project

## Profiling and MPI

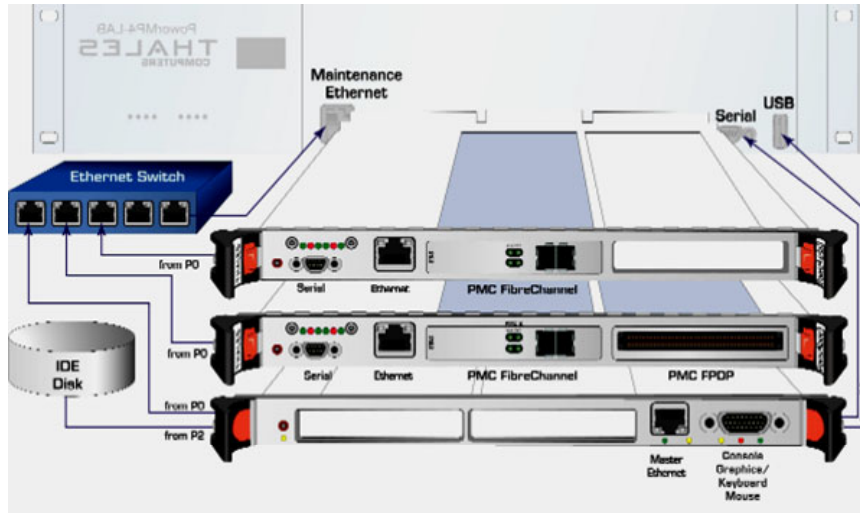
- Provide a trace collection “trigger” facility
- Provide an open “plug in” structure for MPI customisation (e.g., external I/O, specific communication fabric)
  - Co operations: LAM and MPICH teams

## Platform infrastructure

- Adopt existing, open, mature solutions from “non embedded” computing, e.g., HA-OSCAR
  - Cooperation's: Louisiana Tech University, Oak Ridge National Laboratory

Come and see 

## Demo platform outside



See you there for questions

Thank you.

T30527-b-en