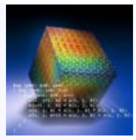
PARO – A Design Tool for the Automatic Generation of Hardware Accelerators

Frank Hannig, Hritam Dutta, and Jürgen Teich
Hardware/Software Co-Design, Department of Computer Science
University of Erlangen-Nuremberg, Germany
{hannig, dutta, teich}@cs.fau.de
URL: http://www12.cs.fau.de/research/paro

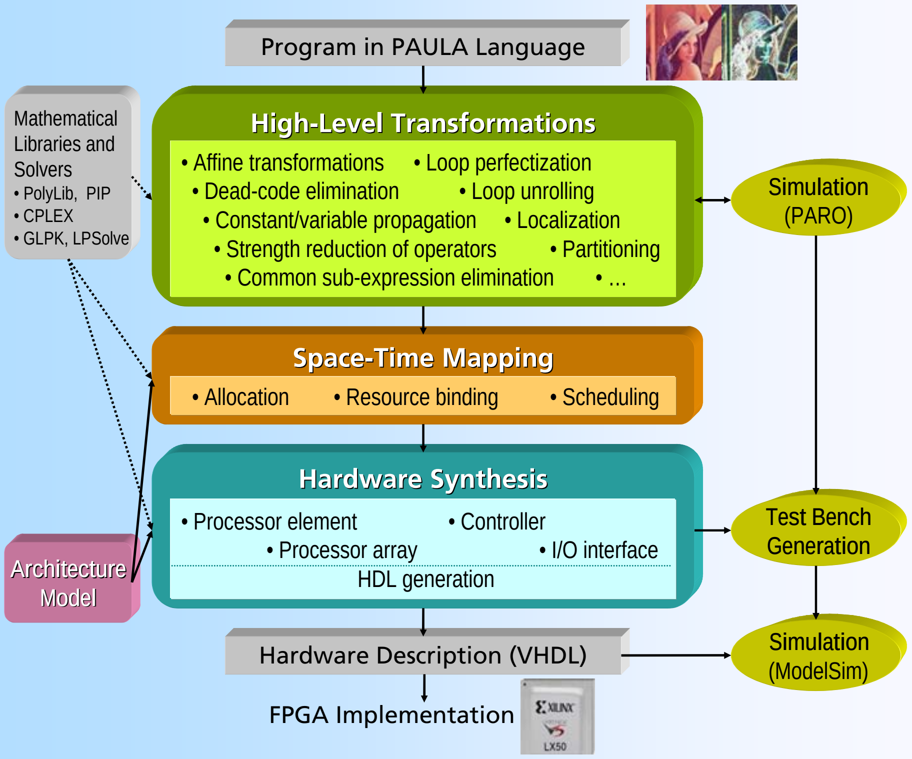
Friedrich-Alexander-Universität Erlangen-Nürnberg

Overview

- Tool for the automated hardware synthesis of massively parallel embedded architectures
- Application domains: Video, audio, image, and other digital signal processing, scientific computing, …
- Design entry of dataflow-intensive algorithms in form of a compact and intuitive language
- Advanced partitioning and scheduling techniques in order to balance trade-offs in cost and performance



PARO Design Flow



PAULA Language

- Functional programming language
- Based on class of DPLA/DPRA
- Compact description of multi-dimensional loop programs with polyhedral iteration spaces
- Reductions: Σ, Π, max, min

  Example, 2-D Gaussian window filter:

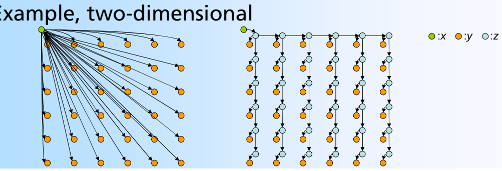
```
par (x>=0 and x <=1021 and y>=0 and y<=1021)
{ w[0,0]  = 1; w[0,1] = 2; w[0,2] = 1;
  w[1,0]  = 2; w[1,1] = 4; w[1,2] = 2;
  w[2,0]  = 1; w[2,1] = 2; w[2,2] = 1;
  h[x,y]  = SUM[i>=0 and i<=2 and j>=0 and j<=2]
             (pi[x+i,y+j] * w[i,j]);
  po[x,y] = h[x,y] >> 4;  // divided by 16
}
```

Localization

- Example, one-dimensional

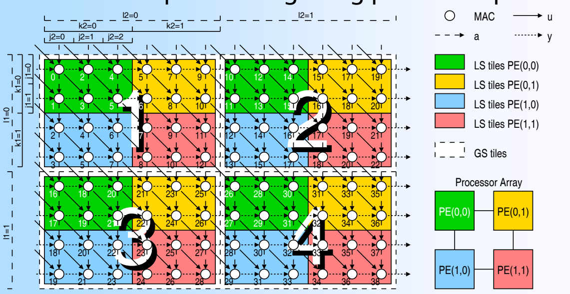
```
for (i=0 to N)          for (i=0 to N)
{ b[i] = a[0];          { a[i] = a[i-1]   if (i > 0);
}                         b[i] = a[i];
                        }
```

- Example, two-dimensional



Partitioning

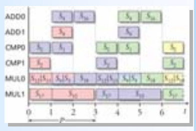
- Transformation which covers the iteration space of computation using congruent *tiles*
- Matches an algorithms to given arch. constraints (functional resources, memory, I/O bandwidth)
- Hierarchical partitioning using parallelotopes



Space-Time Mapping

- Assignment of iteration points to processors (allocation) and start times (scheduling)

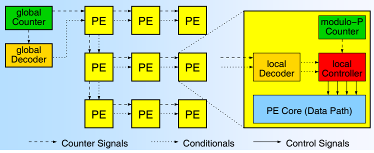
$$\begin{pmatrix} p \\ t \end{pmatrix} = \begin{pmatrix} Q \\ \lambda \end{pmatrix} \cdot I + \begin{pmatrix} q \\ \gamma \end{pmatrix}$$



- Scheduling
  - Resource constraints
  - Module selection
  - Functional and software pipelining
  - Run-time dependent conditionals and hierarchical partitioned algorithms
  - Mixed integer programming (MIP)

Hardware Synthesis

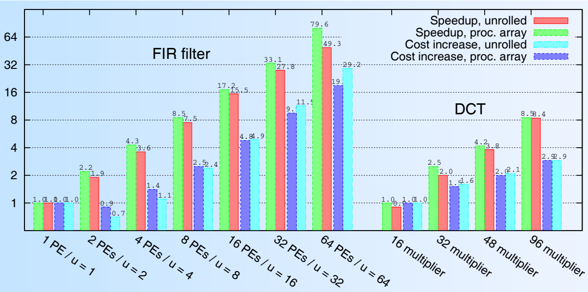
- Generation of platform and language independent RTL description
- Synthesis steps:
  1. Processor elements
  2. Array interconnection structure
  3. Controller
- Backend for VHDL



Case Studies

Algorithm	No. of PEs	No. of LUTs	No. of FFs	No. of MULTs	No. of BRAMs	max. Clock (MHz)	Exec. Time (cycles)	avg. Output Interval (cycles)
Edge Detection								
– 100×100 image, partitioned	1×4	2962	1455	0	4	143	$7.7 \cdot 10^3$	3
– 1000×1000 image, partitioned	1×4	2997	1913	0	44	120	$7.5 \cdot 10^5$	3
Gaussian Filtering								
– 100×100 image, 3x3 mask	3×3	655	1439	9	2	171	$1.0 \cdot 10^4$	1
– 1000×1000 image, 3x3 mask	3×3	683	1463	9	2	171	$1.0 \cdot 10^6$	1
– 2000×2000 image, 3x3 mask	3×3	696	1472	9	4	169	$4.0 \cdot 10^6$	1
– 1000×1000 image, 5x5 mask	5×5	1538	3909	25	4	171	$1.0 \cdot 10^6$	1
FIR Filter								
– 64 Taps, partitioned	1×4	773	834	4	0	125	71	16
– 64 Taps, partitioned	1×8	1915	1014	8	0	132	71	8
– 64 Taps, projected	1×64	5782	9089	64	0	167	68	1
Matrix Multiplication								
– 6x6 matrix size, sequential	1	204	157	1	0	131	250	6
– 6x6 matrix size, partitioned	2×2	829	795	4	0	115	72	1.5
– 6x6 matrix size, projected	6×6	1888	4067	36	0	166	20	0.28
Discrete Cosine Transformation	2	1754	1152	8	1	130	94	0.65
Elliptical Wave Digital Filter	1	1169	624	1	0	94	$2.5 \cdot 10^5$	1
Partial Differential Equation Solver	1	619	502	1	0	128	$1.2 \cdot 10^5$	(1 result)
MPEG2 Quantisizer	1	637	1190	1	0	141	222	2.95
JPEG Loop 1	1	82	79	0	0	224	63	1
JPEG Loop 2	1	570	1139	0	0	158	126	1

Setup: Xilinx ISE 6.3i, Virtex-II 8000



Conclusions

- Good scalability of the PARO methodology
- Processor array approach outperforms conventional loop unrolling (10-61% higher throughput)
- Generation of highly parallelized hardware accelerators