Modeling the power-reliability tradeoff in on-chip networks

Claas Cornelius, Frank Sill, Dirk Timmermann

8th – 10th October 2007

12. SYMPOSIUM MARITIME ELEKTROTECHNIK, ELEKTRONIK UND INFORMATIONSTECHNIK

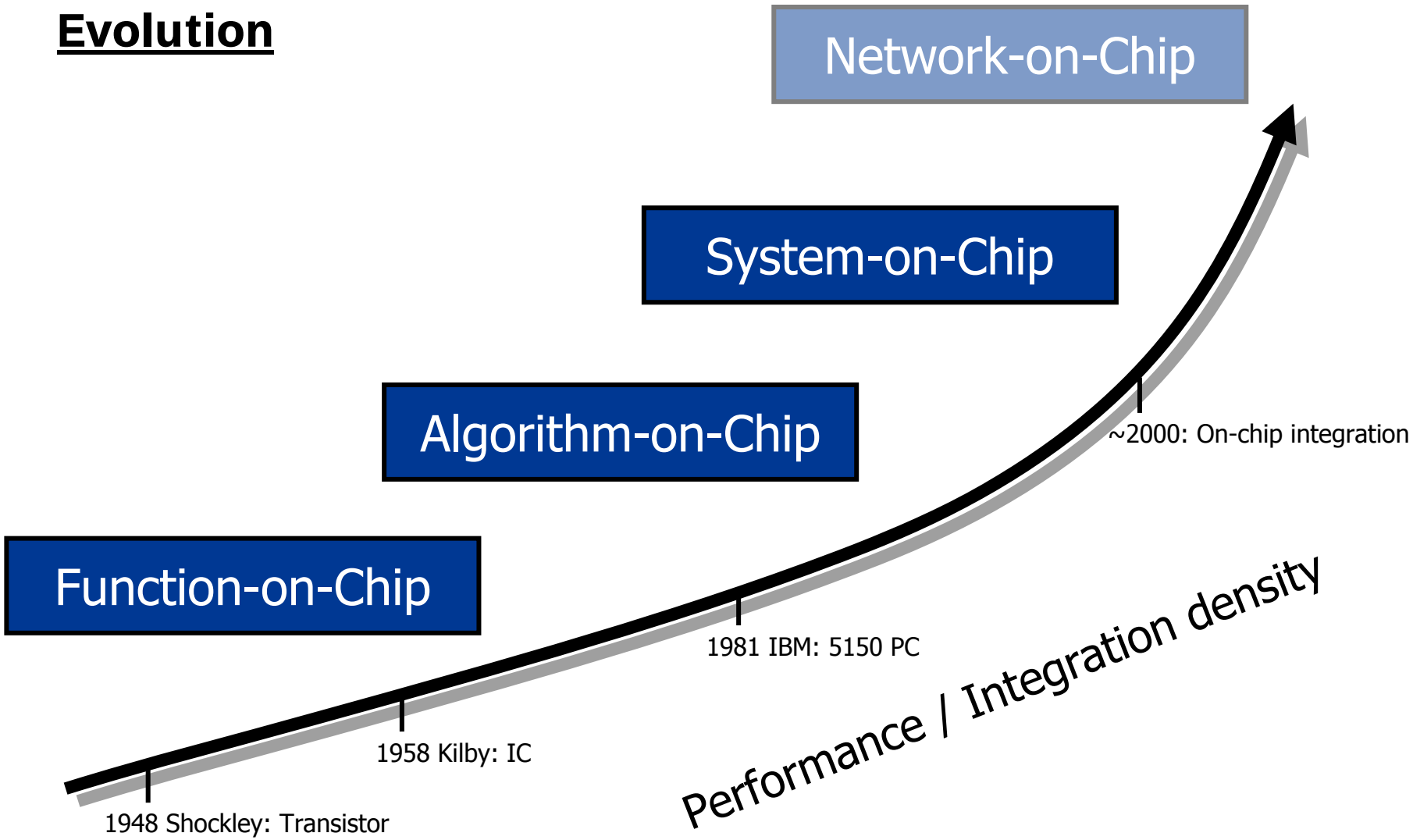
Outline

- Introduction
  - System design
  - Technology issues
- Approach
  - Analytical models
  - Simulation
- Results
- Summary

Claas Cornelius, October 2007
Institute of Applied Microelectronics and Computer Engineering, University of Rostock

System design

Evolution



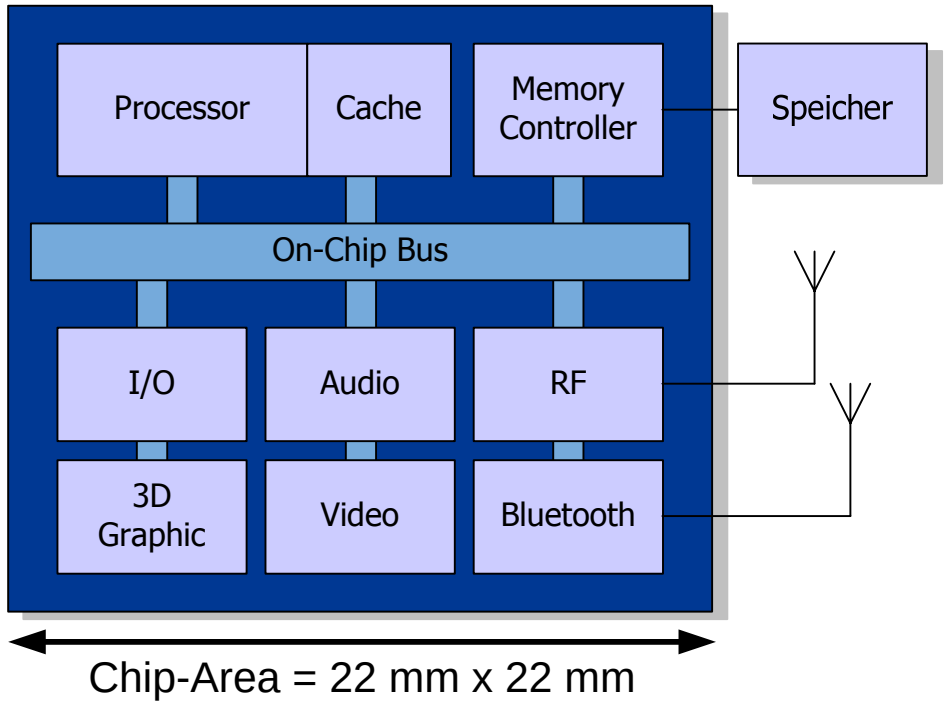
Claas Cornelius, October 2007
Institute of Applied Microelectronics and Computer Engineering, University of Rostock

System design

Bus-based system design

➔ Shared communication medium

- Connected problems:
  - Memory-Bottleneck
  - Design-Productivity-Gap
  - Synchronous design



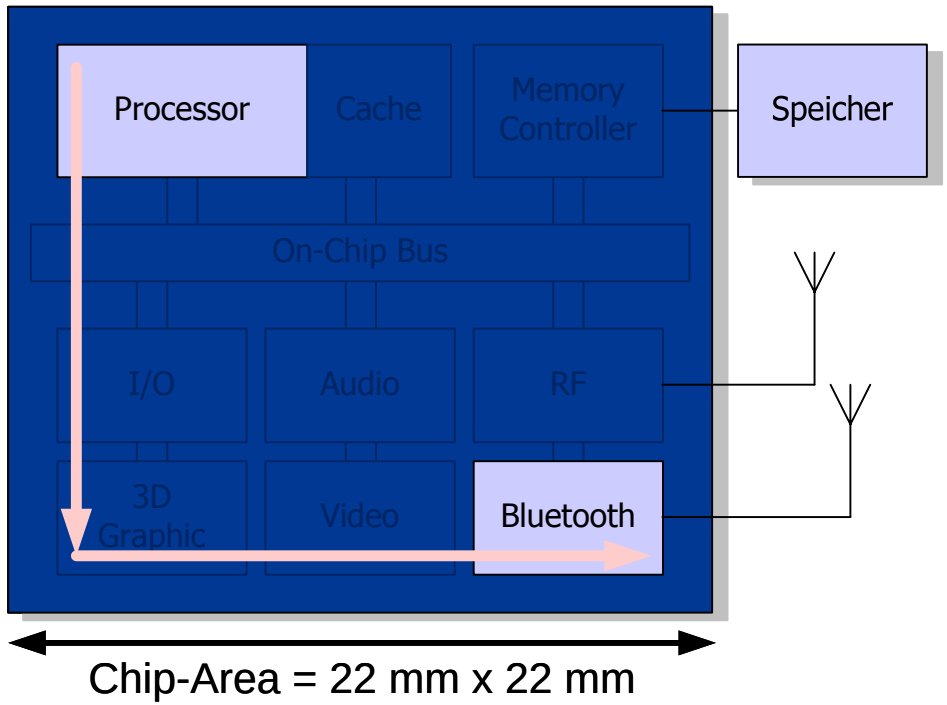
Claas Cornelius, October 2007
Institute of Applied Microelectronics and Computer Engineering, University of Rostock

System design

Bus-based system design

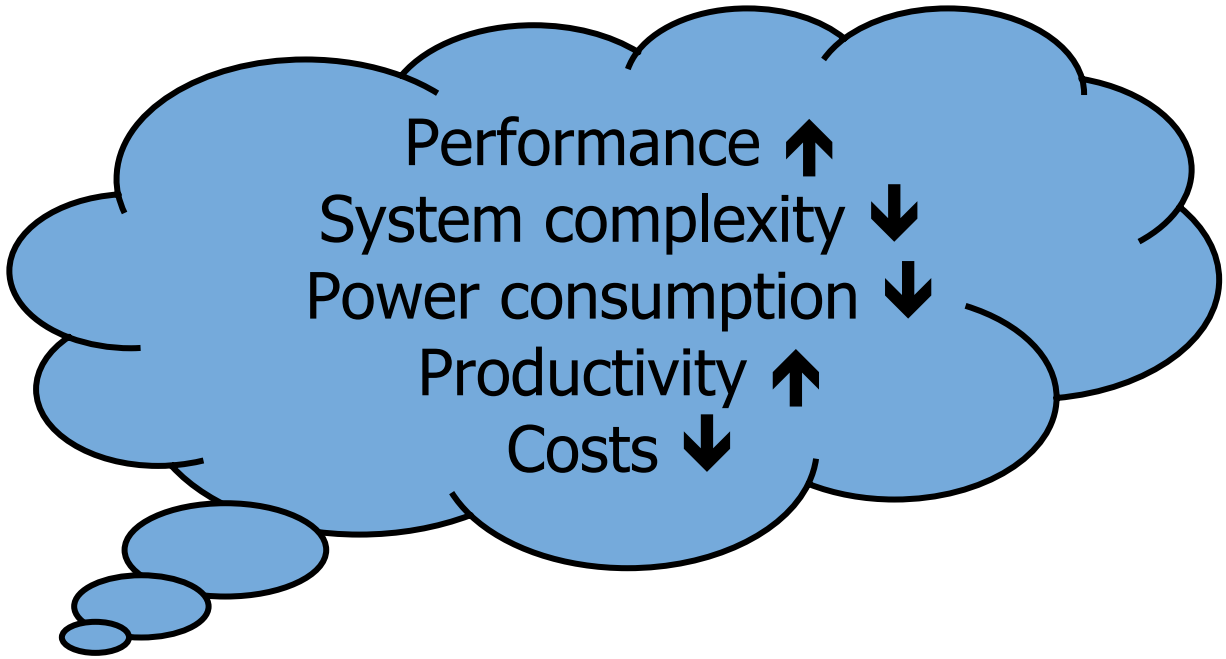
➔ Shared communication medium

- Connected problems:
  - Memory-Bottleneck
  - Design-Productivity-Gap
  - Synchronous design
- Worsening the situation:
  - Chip-Size
  - Interconnects
  - Integration density
  - Parameter variability
  - ... and many more



System design

Wish list



[Intel]

“ … multi-core processors are another […] 10x factor in terms of performance … ”

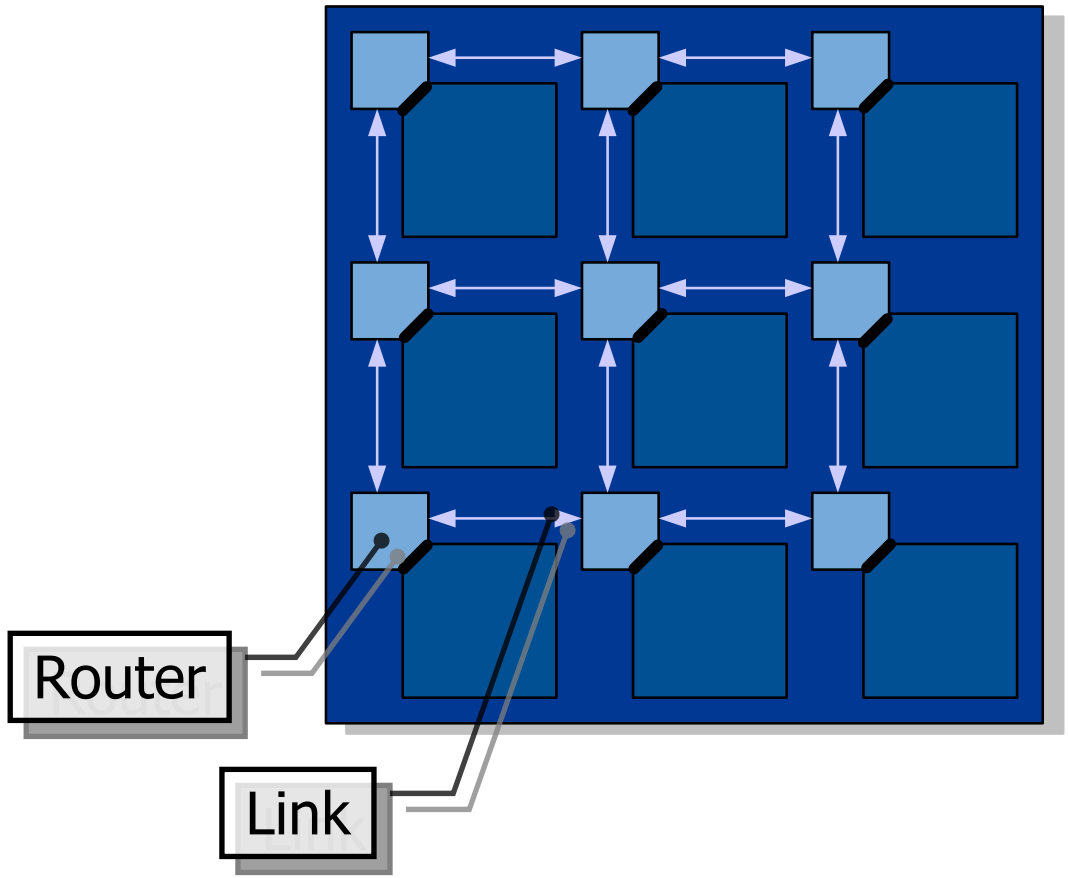
Paul Otellini, Intel President
IDF, September 2004

System design

Network-on-Chip

Promising properties:

- Parallelism
- Modularity



Claas Cornelius, October 2007
Institute of Applied Microelectronics and Computer Engineering, University of Rostock

System design

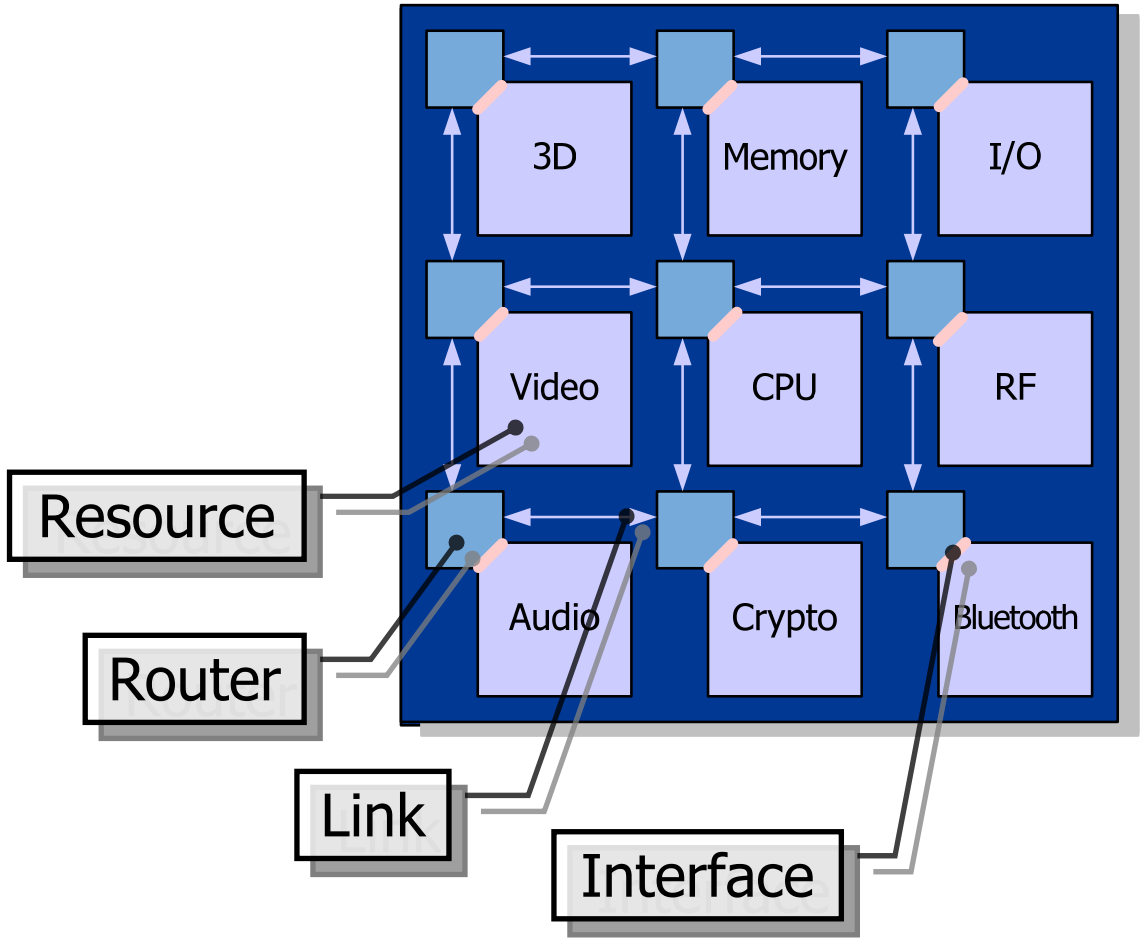
Network-on-Chip

Promising properties:

- Parallelism
- Modularity

Level of abstraction:

Modules
Logic-Gates
Transistors
Polygons



Change of Paradigms:

Computation → Communication

Claas Cornelius, October 2007
Institute of Applied Microelectronics and Computer Engineering, University of Rostock

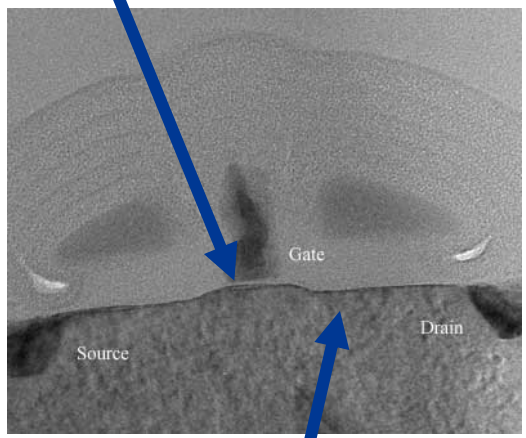
Technology issues

Technology scenario 2010

- Technology node 45 nm
- Chip size 620 mm²
- Transistor count 4424 Mio.
- Physical gate length ~20 nm
- Metal layers 12-16
- Frequency (global-local) 2-12 GHz
- Total wire length ~2.2 km/cm²
  - For M1-M6, 33% usage

[ITRS, 2003 and 2005]

<10 Ångström gate oxide thickness



[Intel, 2001]

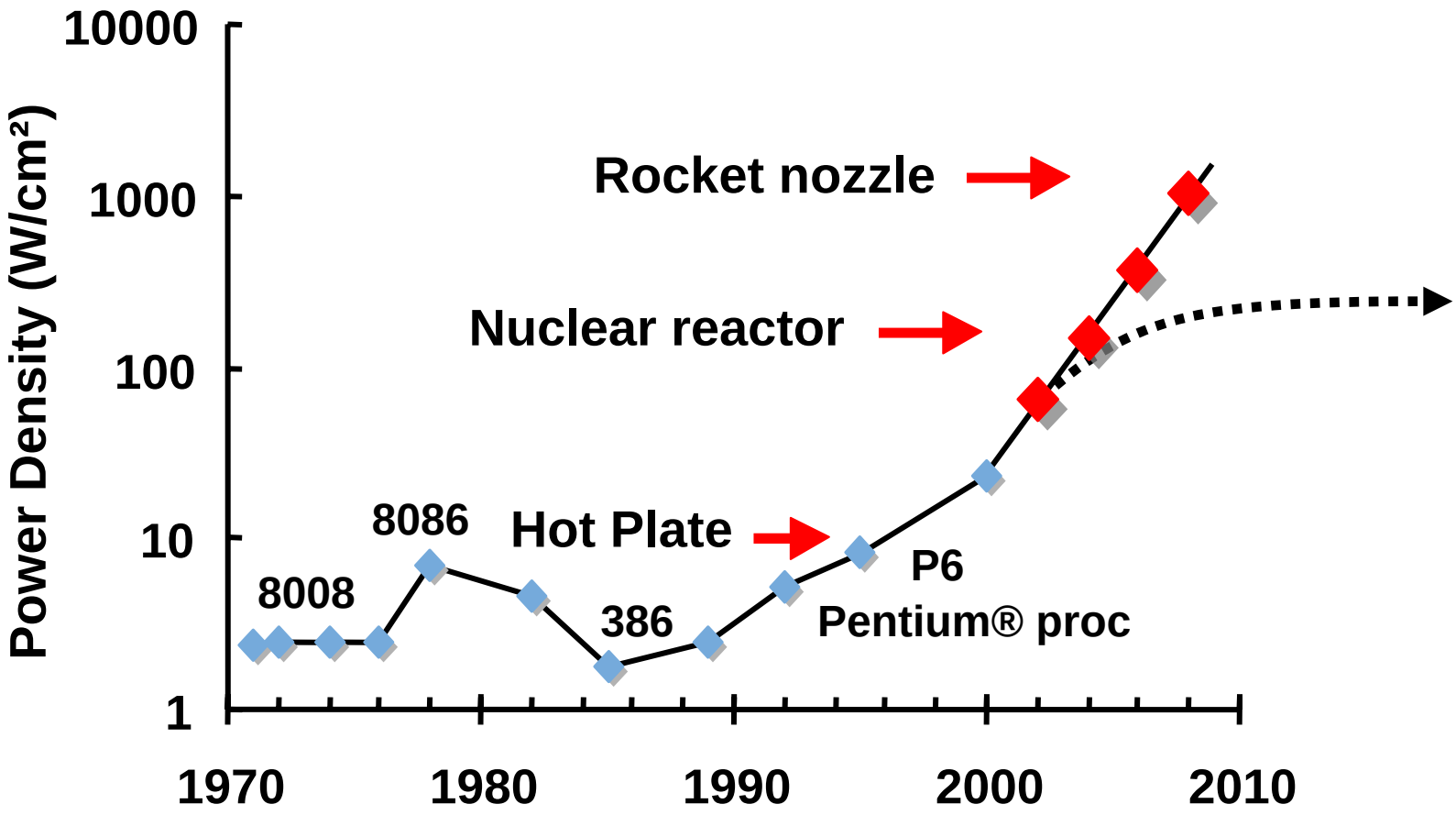
<100 dopant atoms

- Approximate size of a resource
  - Rent's rule
  - Reachable die area with clock signal
  - Acceptable power trade-off

} **~85 resources**

Technology issues

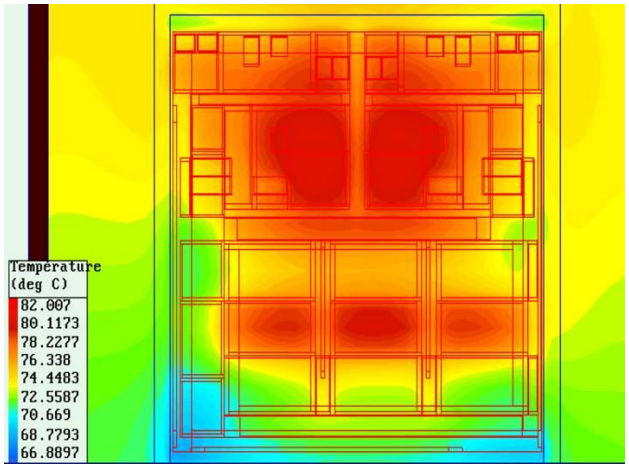
Power density



[Borkar]

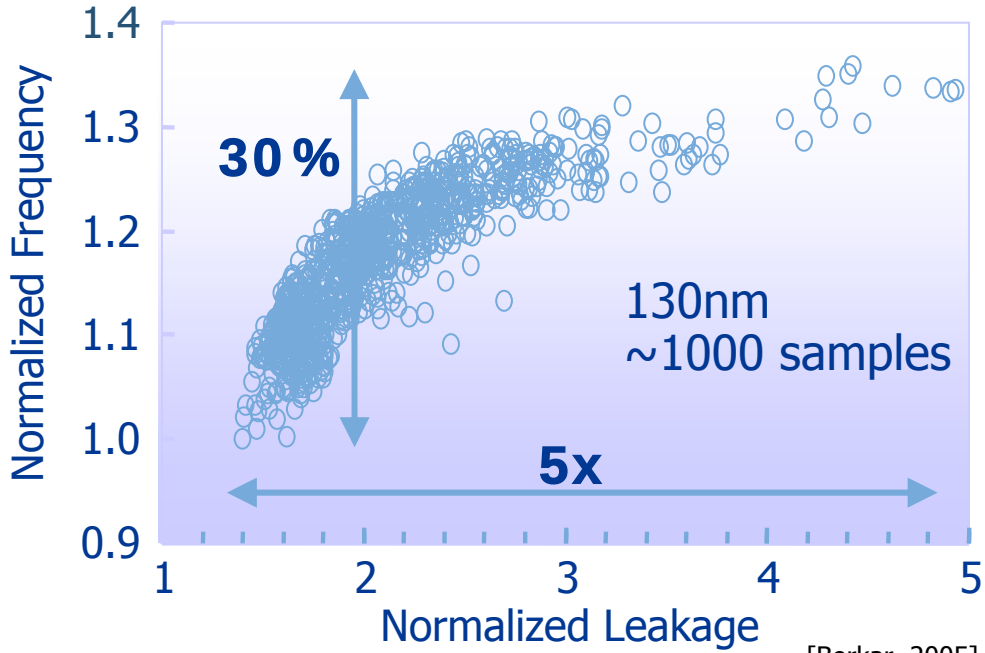
Technology issues

Parameter variability



Power4 Server Chip

[Devgan, 2003]



[Borkar, 2005]

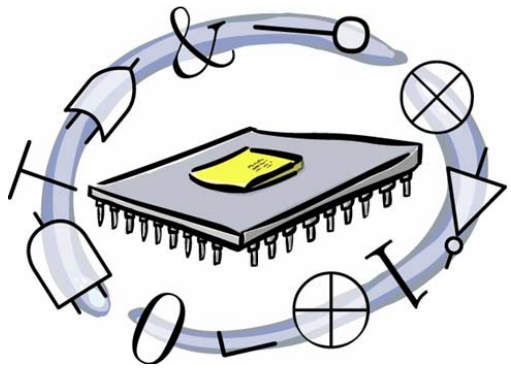
L (nm)	250	180	130	90	65	45
Vt (mV)	450	400	330	300	280	200
σ-Vt (mV)	21	23	27	28	30	32
σ-Vt/Vt	4.7%	5.8%	8.2%	9.3%	10.7%	16%

[ITRS, 2003]

Parameter variability dramatically increasing

Approach

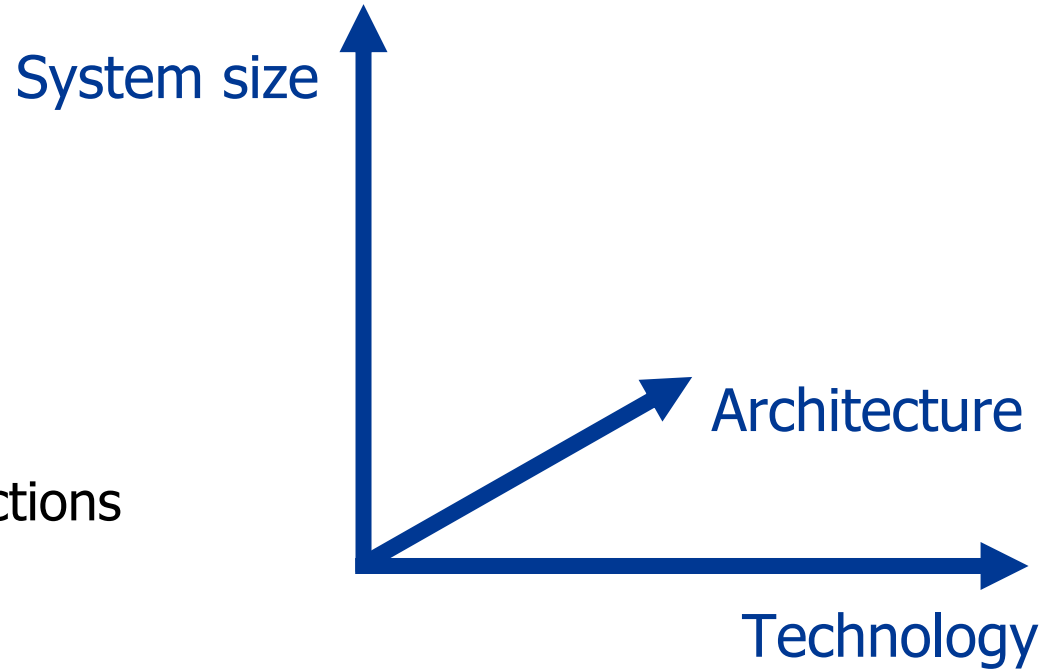
Related work



- Prototyping, Test-Chips
  - Star topology for multimedia applications [Lee, 2003]
  - 4x4 mesh network with traffic generators [Mullins, 2006]
- Parametrizable VHDL-model ported to FPGA [Zeferino, 2004]
- Emulation framework on an FPGA [Genko, 2005]
- High-level VHDL [Sigüenza, 2002]
- SystemC approach and design flow [Jalabert, 2004]
- Event-based C++ Simulator [Wiklund, 2004]

Approach

Design space exploration



1) Analytical models
   - Fast evaluation
   - Inaccurate Predictions

2) Simulation
   - Computation and time intensive
   - Accurate results

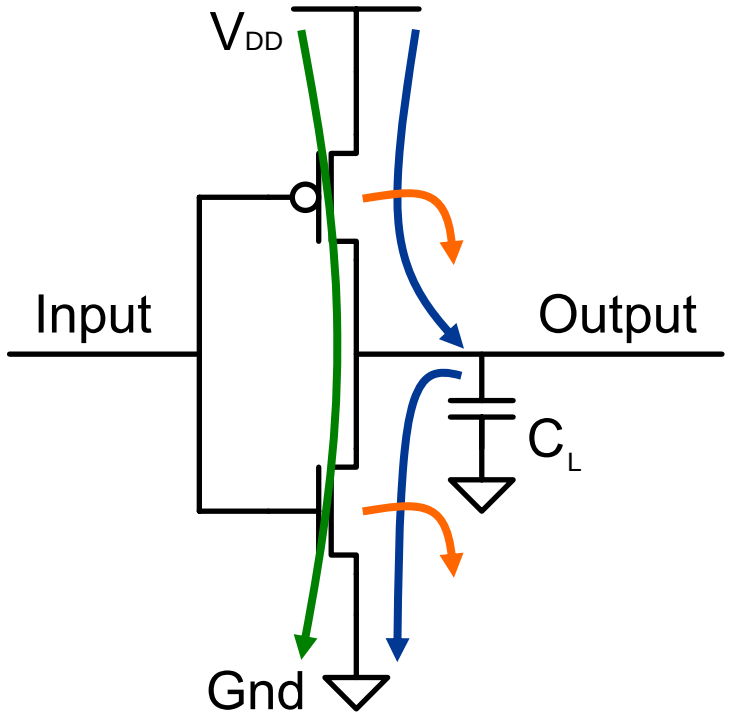
Analytical models

Power equation

$$P_{total} = \underbrace{\alpha\, C_L\, V_{DD}^{\,2}\, f}_{\text{Dynamic}} + \underbrace{t_{SC}\, V_{DD}\, I_{peak}\, f}_{\text{Short-Circuit}} + \underbrace{V_{DD}\, I_{leakage}}_{\text{Leakage}}$$

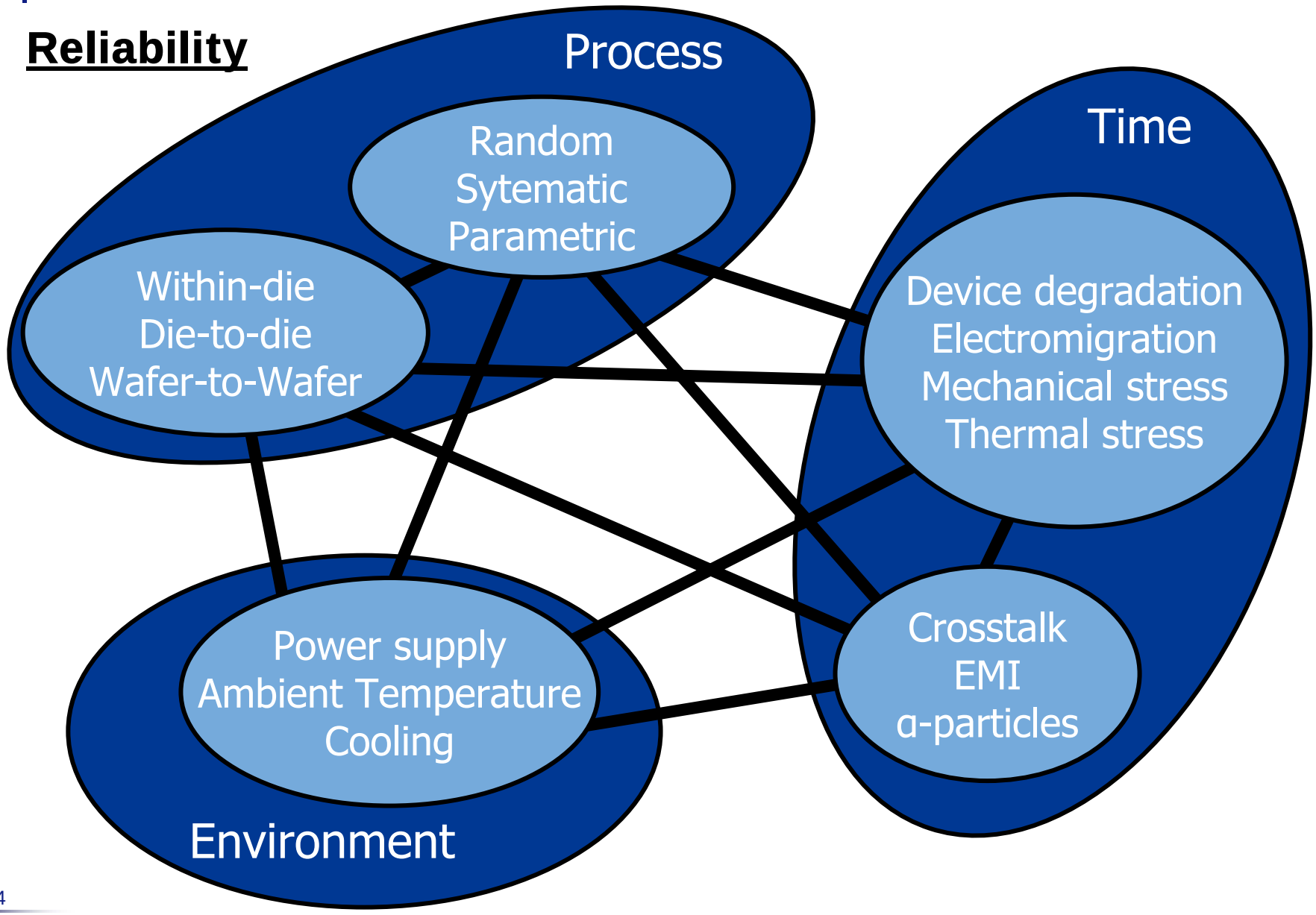
E.g. Inverter

Output = $\overline{\text{Input}}$



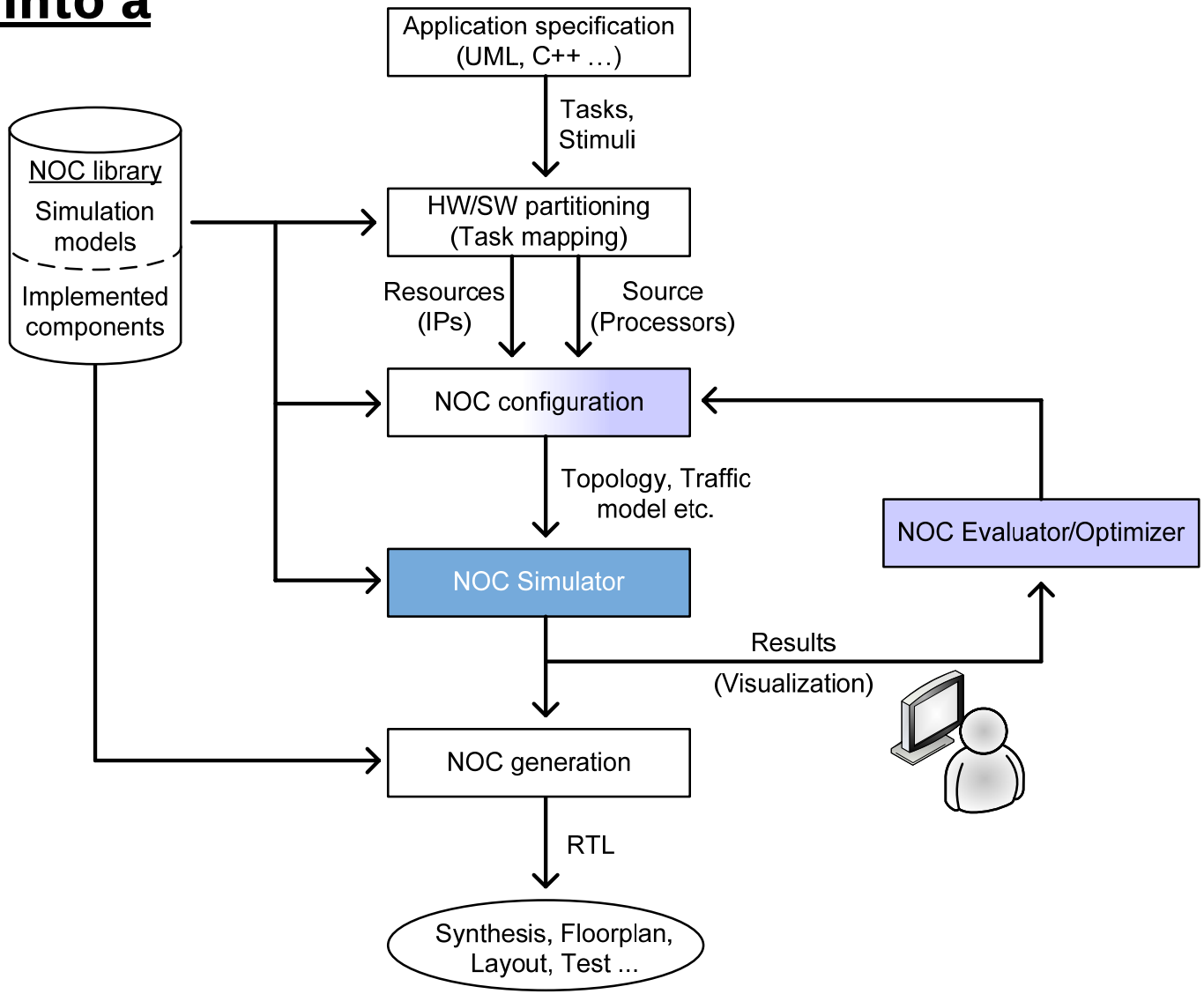
- Glitches, Spikes, Hazards:
  - Due to race conditions of signals
  - Hard to predict
  - Mostly empirical assessment
- Static power due to design style

Analytical models



Simulator

Integration into a design flow



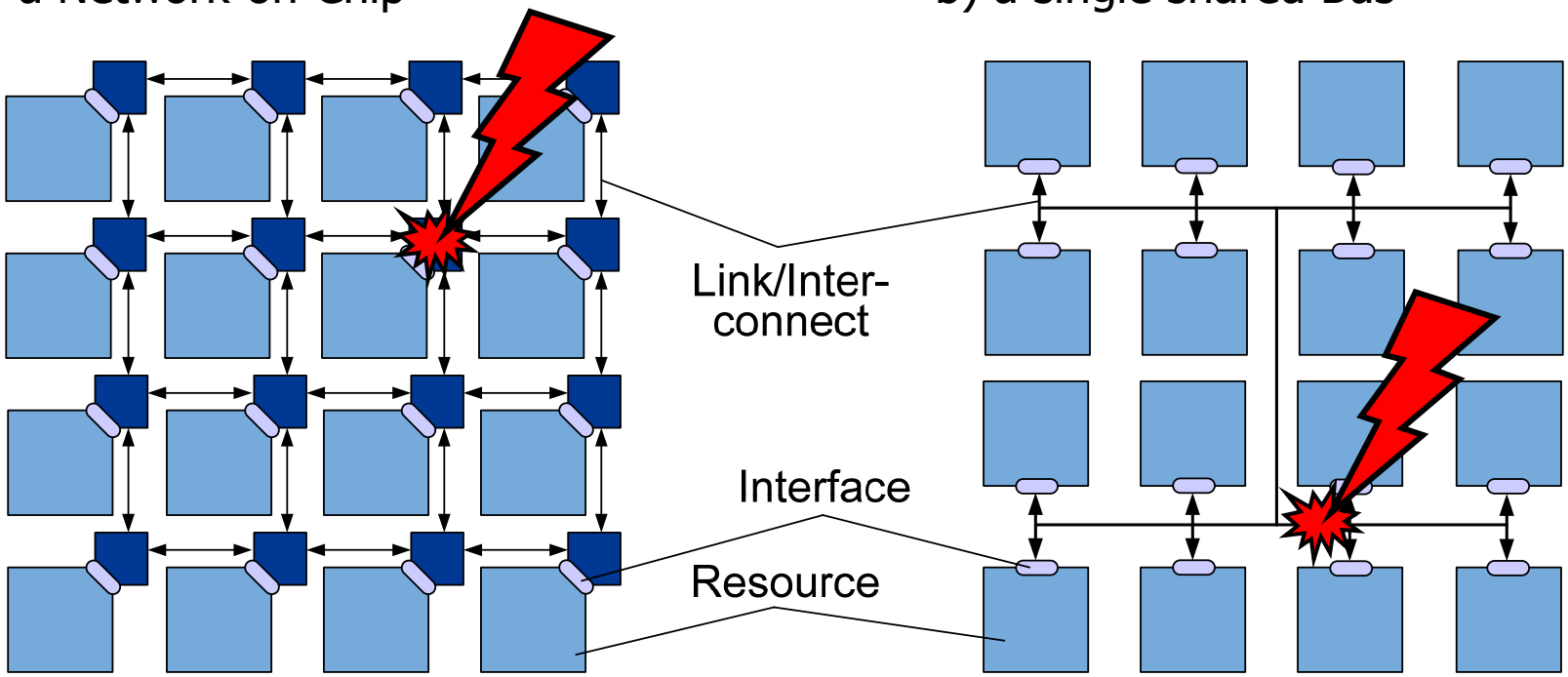
Results

Scenario

Simple example of 16 resources/IPs in a regular 4x4 mesh interconnected with

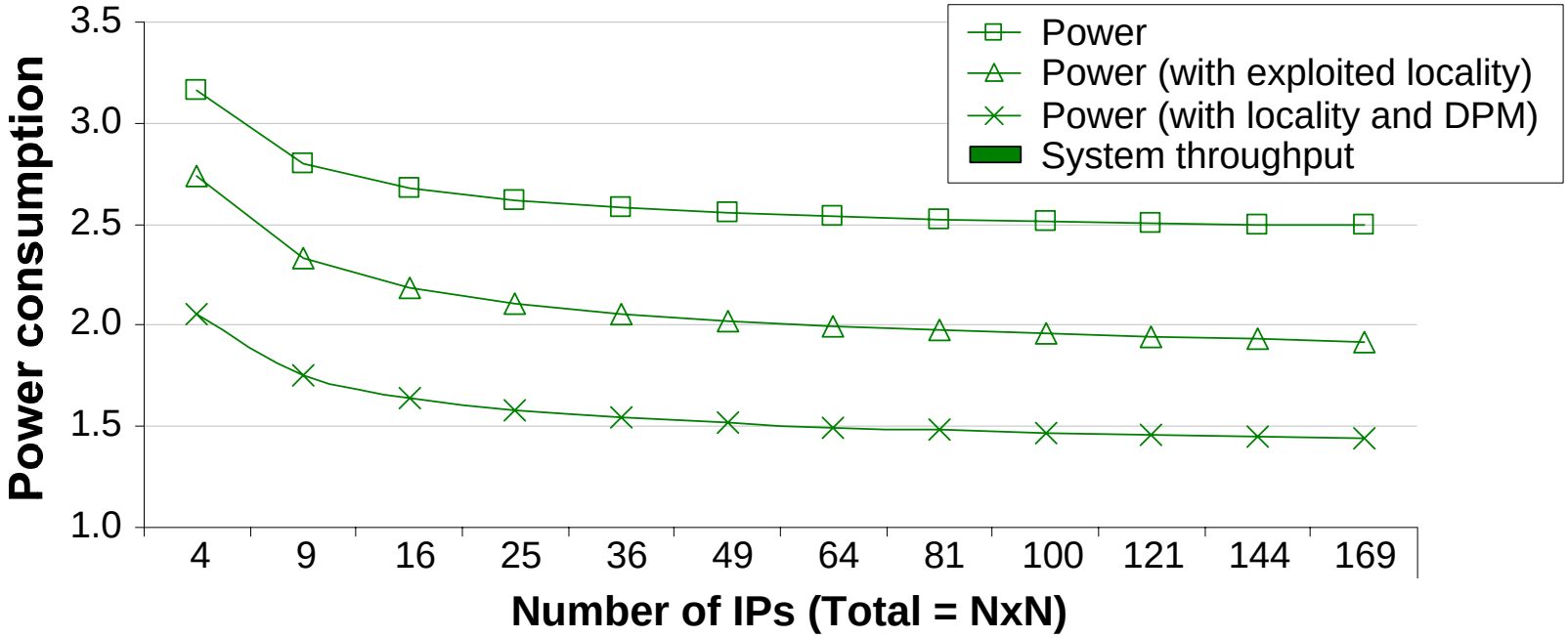
a) a Network-on-Chip

b) a single shared Bus



Results

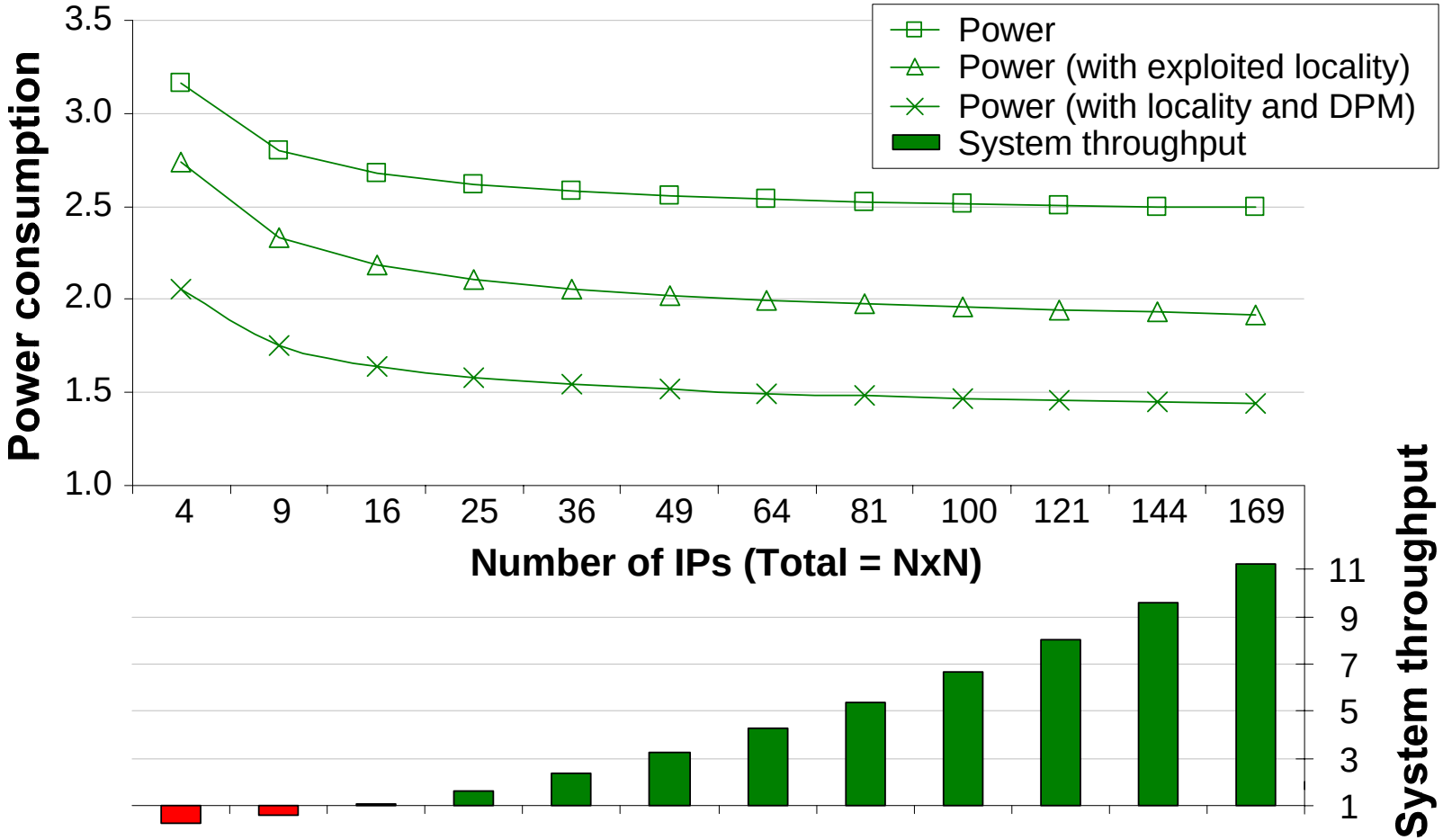
Power and throughput



Claas Cornelius, October 2007
Institute of Applied Microelectronics and Computer Engineering, University of Rostock

Results

Power and throughput



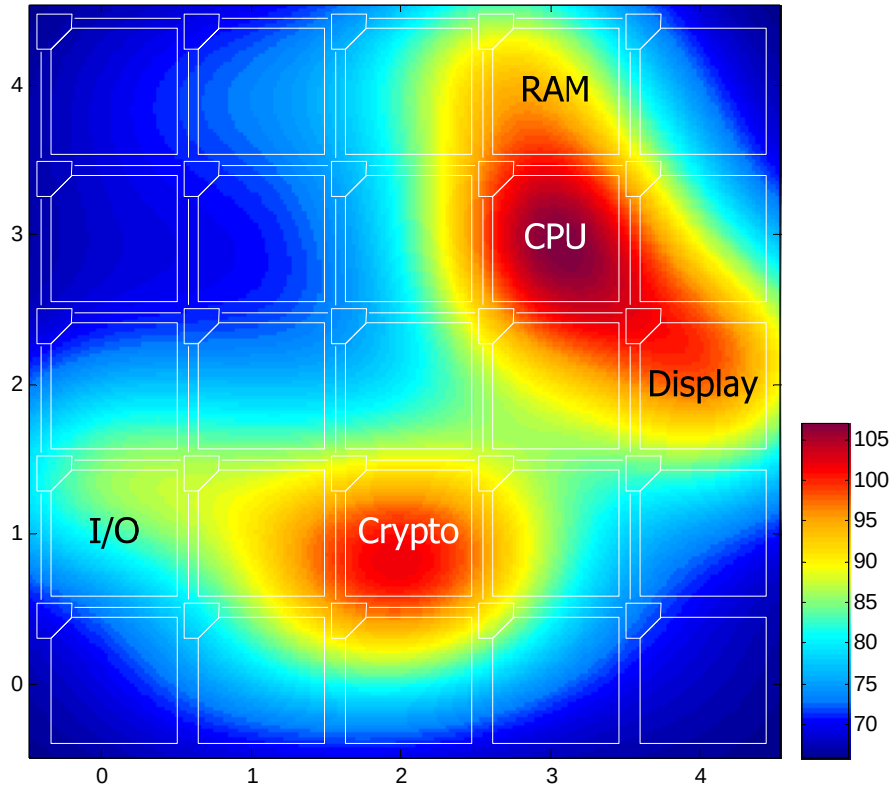
Claas Cornelius, October 2007
Institute of Applied Microelectronics and Computer Engineering, University of Rostock

System control

Results

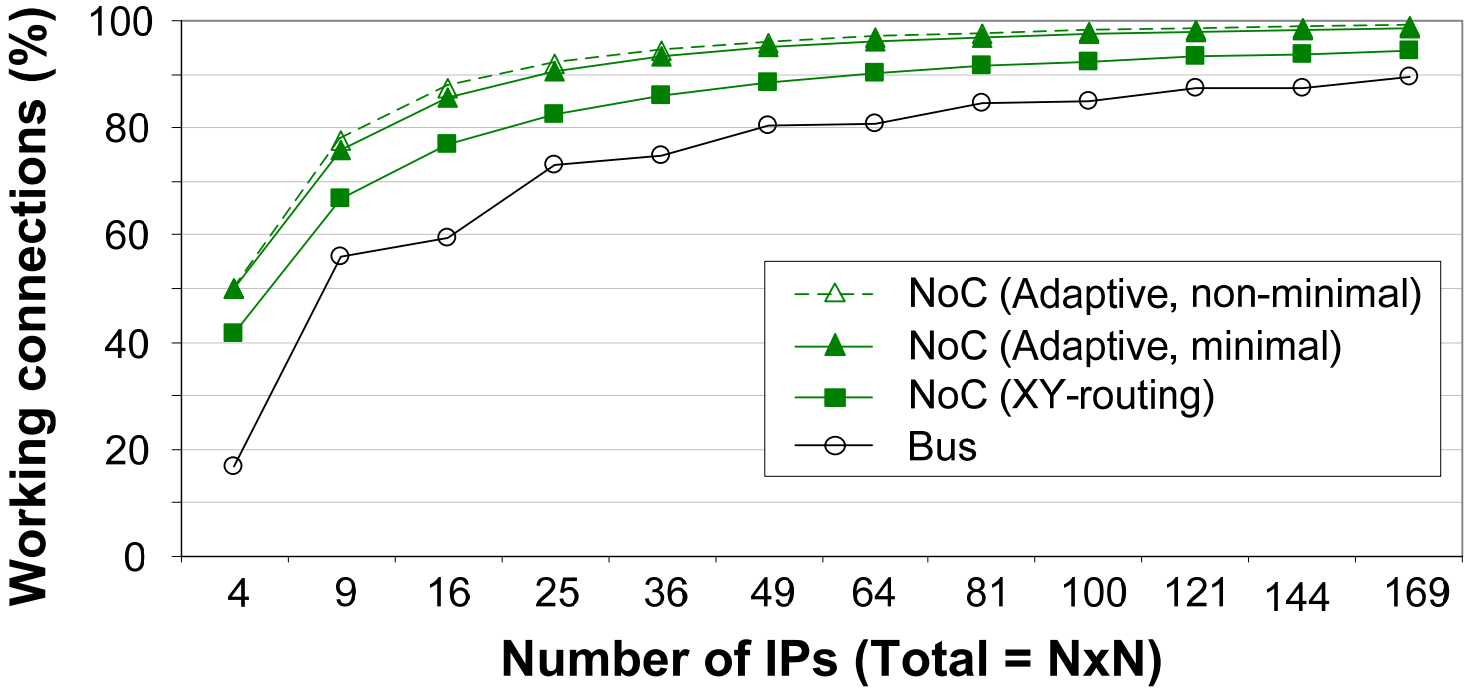
- Temperatur distribution influences
  - ➔ Performance
  - ➔ Power consumption
  - ➔ Reliability
- Dynamic power management
  - ➔ Avoid hot spots
  - ➔ Task mapping
  - ➔ Packet congestions
  - ➔ DV/FS, Clock-gating …

Application example for the temperature distribution of a 5x5 network



Results

Reliability



- Permanent, single error
- Average number of working connections
- Cohesive connected system

Summary

- Motivation for investigating
  - Networks-on-Chip
  - Power and reliability
- Combined approach for design space exploration
- Comparison of results to a reference bus
- Consideration of dynamic behavior and control