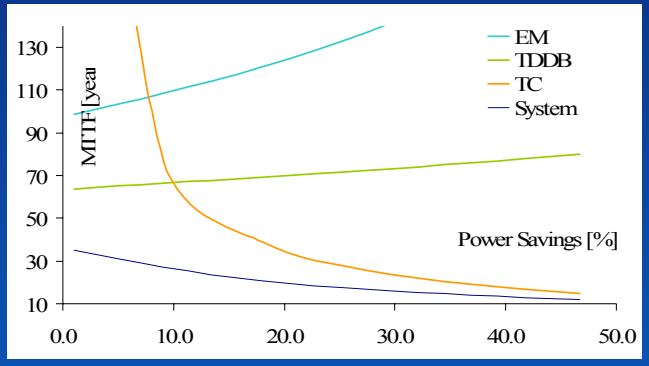
# Optimization of Reliability and Power Consumption in MPSoCs

Tajana Šimunić Rosing

UCSD

# Focus of current work

- Simulate system-level reliability
  - model three sources of hard errors:
    - Electromigration (EM), Time-dependent dielectric breakdown (TDDB), fast (on chip) and slow (package) Temperature Cycling (TC)
  - as a function of a power management policy
    - EM and TDDB typically improve with lower power consumption
    - TC effect can dominate at small feature sizes, thus causing a large drop in reliability with aggressive power management policies




- Design and optimize a system management policy
  - Maximize reliability and minimize energy consumption
  - Time-indexed Markov chain model
    - Combined dynamic reliability (DRM) with dynamic power management (DPM) optimization

# Hard Failure Rate Models

- Electromigration (EM) - result of momentum transfer from electrons to the ions which make interconnect lattice

$$\lambda_{core,s}^{EM} = A_o^{'}(J_s - J_{crit})^n e^{\frac{-Ea}{kT_s}}$$

$$\forall s = active, idle$$

- Time Dependent Dielectric Breakdown (TDDB) - wear out mechanism of dielectric due electric field and temperature

$$\lambda_{core,s}^{TDDB} = A_o^{'} e^{\gamma E_{ox,s}} e^{\frac{-Ea}{kT_s}} \quad ;$$

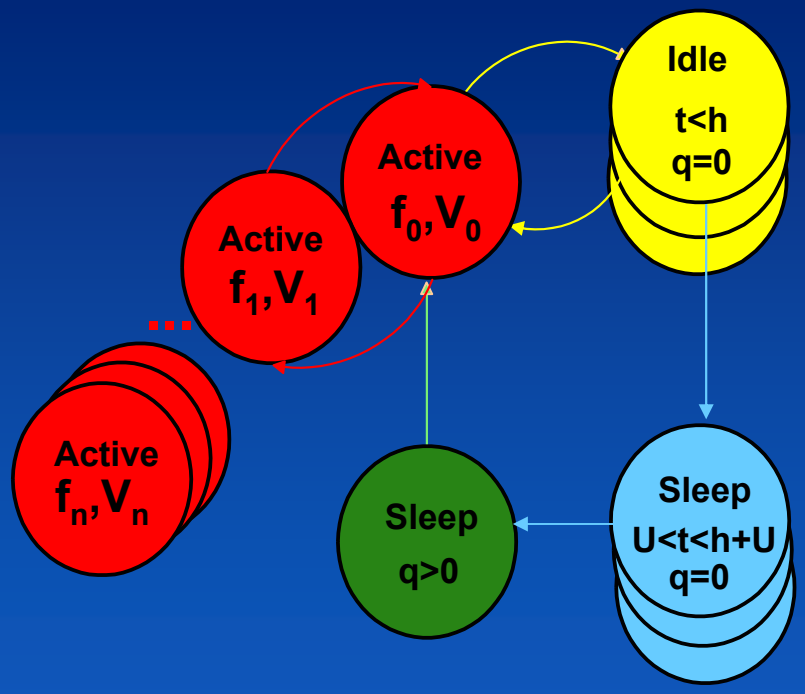
$$\forall s = active, idle, sleep$$

- Thermal Cycles (TC) - occur during power state changes, induce plastic deformations in materials

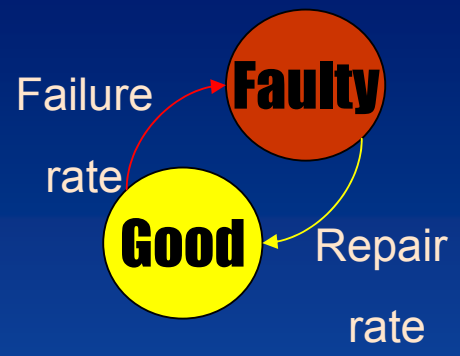
$$\lambda_{core,s}^{TC} = C_o^{'}\left[\left(T_{active} - T_s\right) - \left(T_{avg,s} - T_{mold}\right)\right]^q t^{-1} \quad \forall s = sleep$$

# System Modeling

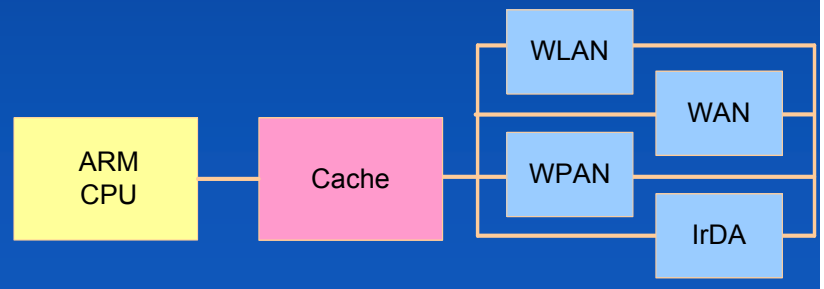
## Core DVS & DPM




## Core reliability

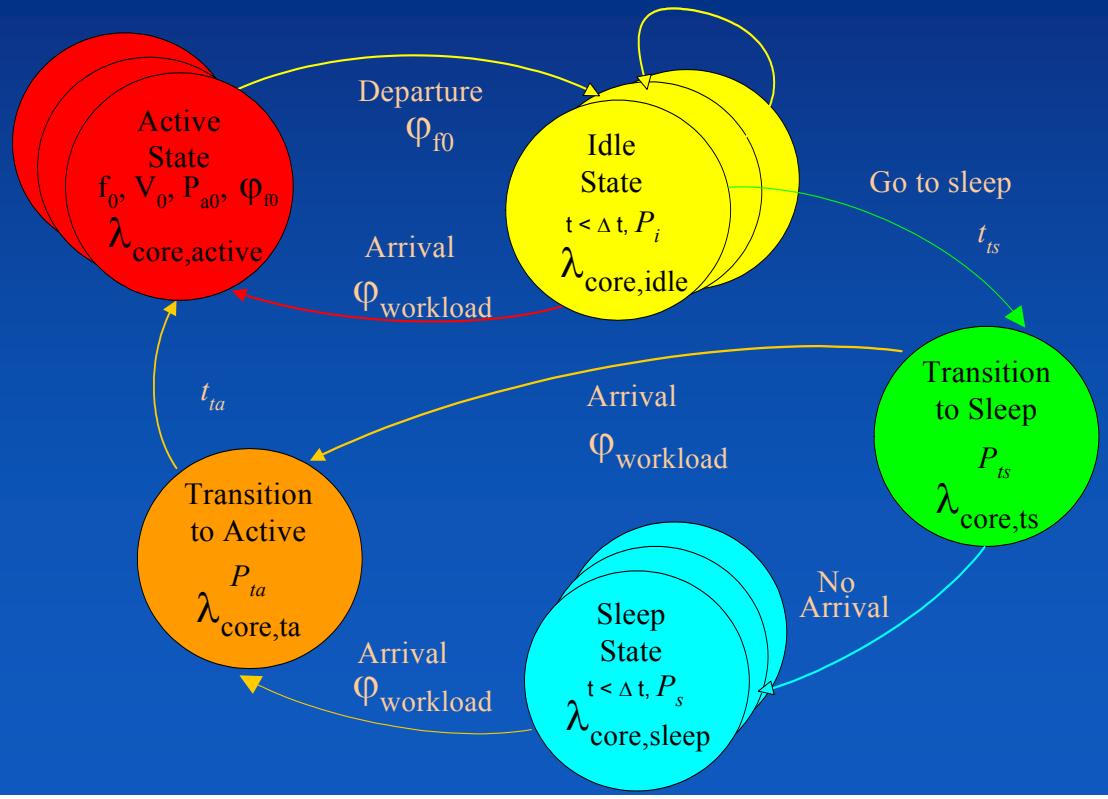



## Reliability topology



# Optimization: DPM&DRM System Model

- Combine:
  - Power-state machine model - TISMDP
  - Reliability model - Markov process
- Represent the overall system as combination of components' power-state machines where failure rates depend on the system state
- System control aims to increase energy efficiency and enhance reliability



# DPM&DRM Policy Optimization

Minimize average energy consumed under reliability and performance constraints – get randomized policy

$$\min \quad \sum_{c=1}^{N} \text{cost}_{energy,\ c}$$

$$s.t. \quad \sum_{a\in A} f(s,a) - \sum_{a\in A}\sum_{s'\in S} M(s'|s,a) f(s',a) = 0; \quad \forall s, \forall c_s$$

$$\sum_{a\in A}\sum_{s\in S} T(s,a) f(s,a) = 1; \quad \forall c_s$$

$$\sum_{c=1}^{N} \text{cost}_{perf,\ c} < Perf_{const}; \quad \forall c$$

$$Tpl\,(\lambda_c) \le \text{Rel}_{const}; \quad \forall c_s$$

$$\lambda_c = \sum_{i\in F}\sum_{a\in A}\sum_{s\in S} \lambda^i_{core}\,(s,a) y(s,a) f(s,a)$$

Variable definitions:

| | |
|---|---|
| cost (s,a) | average cost incurred while in state s given action a |
| f( s,a ) | frequency of executing action a while in state s |
| M( s'\| t,s,a ) | probability of arriving to state s' given action a taken in state s |
| T( s,a ) | expected time spent in state s given action a |
| Tpl($\lambda_c$) | reliability constraint as a function of network topology Tpl |
| $\lambda_c$ | core reliability |

**Obtain globally optimal policy using linear programming**

- Policy is obtained from state-action frequencies f(s,a) in form of a table of probabilities of issuing command a when system is in state s

$$p(s,a) = \frac{f(s,a)}{\sum_{a'\neq a} f(s,a')}$$

# Reliability Constraint Formulation

- Failure rate of each state is a sum of the failure rates due to all mechanisms (EM, TDDB, TC) acting in that state
  - Expected temperature in a state needs to be calculated

$$T_{state} = (T_{active} - T_{state,ss})e^{-\frac{y(s,a)}{\tau}} + T_{state,ss}$$

$$T_{active} \propto P_{active}(R_{th\,die} + R_{th\,package})$$

- Total failure rate of a core is a weighted sum of state failure rates, for example:
  - core has three power states: active, idle and sleep
  - two actions: "go to sleep" (S) and "continue" (C)

$$\lambda_A y(A,C) f(A,C) + \lambda_I y(I,C) f(I,C) + \lambda_I y(I,S) f(I,S) + \lambda_S y(S,C) f(S,C) \le \mathrm{Rel}_{const}$$

- System failure rate is calculated based on core topology as a function of series and parallel combinations

# DPM Constraint Formulation

- Energy and performance cost:
  - $k(s_i, a_i)$ - lump sum cost
  - $c(s_{i+1}, s_i, a_i)$ - cost rate (e.g. power or performance penalty)
  - $F(t_i / s_i, a_i)$ - probability distribution of next event occurrence
  - $p(s_{i+1} | t_i, s_i, a_i)$ – probability of transition into next state $s_{i+1}$

$$Cost(s_i, a_i) = \begin{cases} k(s_i, a_i) + \int_0^\infty \left[ F(du \mid s_i, a_i) \sum_{s_{i+1} \in S_{i+1}} \int_0^u c(s_{i+1}, s_i, a_i) p(s_{i+1} \mid t_i, s_i, a_i) dt \right] & \forall dt \\ k(s_i, a_i) + \sum_{s_{i+1} \in S_{i+1}} c(s_{i+1}, s_i, a_i) T(s_i, a_i) & \forall \Delta t \end{cases}$$

- Expected time spent in each state:

$$T(s_i, a_i) = \begin{cases} \int_0^\infty t \sum_{s_{i+1} \in S_{i+1}} p(s_{i+1} \mid t_i, s_i, a_i) F(dt \mid s_i, a_i) & \forall dt \\ \int_{t_i}^{t_i + \Delta t} \frac{(1 - F(t)) dt}{1 - F(t_i)} & \forall \Delta t \end{cases}$$

- Probability of arrival into each state:

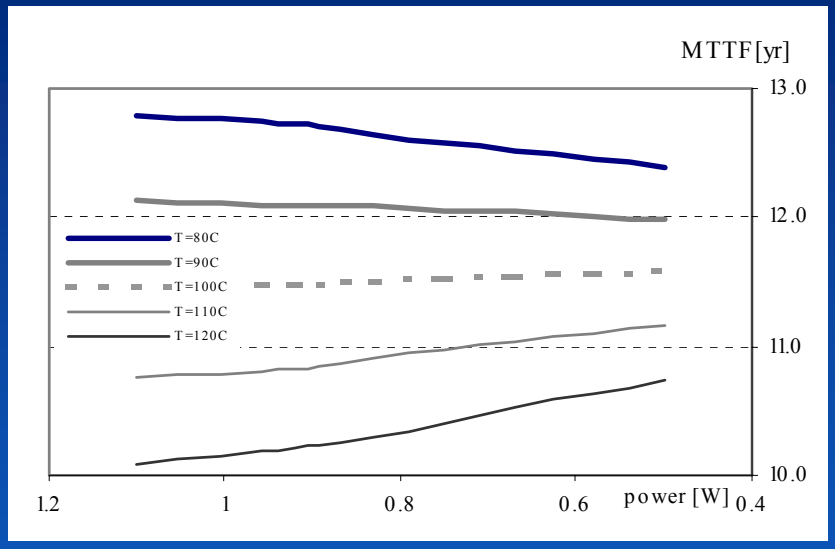
$$M(s_{i+1} \mid s_i, a_i) = \begin{cases} \int_0^\infty p(s_{i+1} \mid t_i, s_i, a_i) F(dt \mid s_i, a_i) & dt \\ p(s_{i+1} \mid t_i, s_i, a_i) & \Delta t \end{cases}$$

# Simulator design

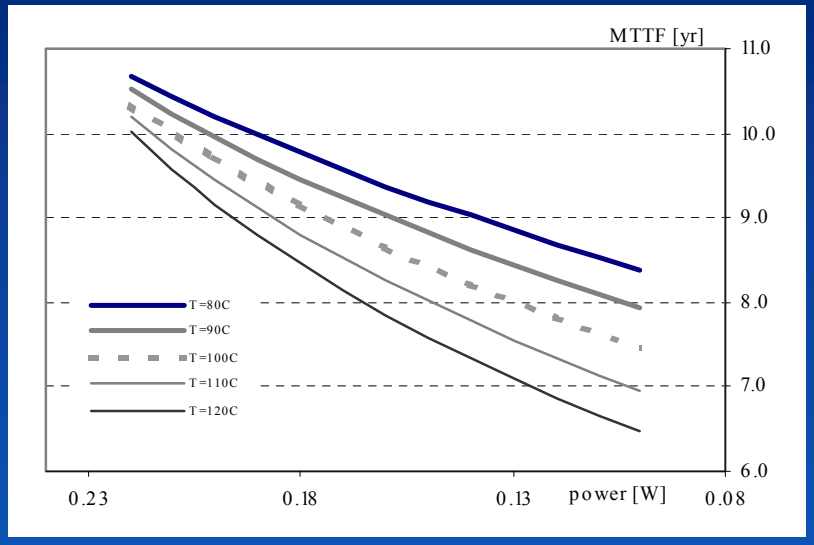
- Event-driven stochastic simulator
- Accepts any workload distribution, including raw data
  - distributions do not need to be stationary
- Input:
  - Reliability topology and failure mechanism characteristics (e.g. EM)
  - Power state specification
  - Workload
  - Time horizon
- Output:
  - MTTF, system reliability, energy consumption, performance
  - For a system of 10 cores runs in a few seconds

# DPM & Reliability

- 95nm technology
- Core A results show a tradeoff between reliability, active core temperature and power management policy due to
- Core B reliability consistently falls with power consumption as TC mechanism dominates at all temperatures




Core A




Core B

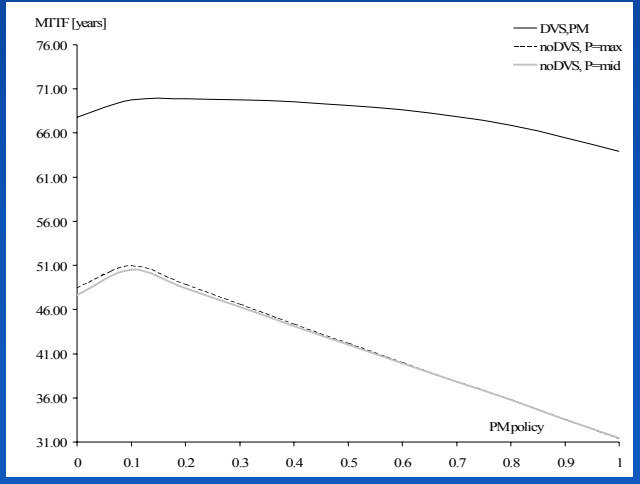
| | $P_{active}$ [W] | $P_{idle}$ [W] | $P_{sleep}$ [W] | $t_{ts}$ [ms] | $t_{ta}$ [ms] | $\lambda_{core}$ [$s^{-1}$] | $\lambda_{workload}$ [$s^{-1}$] |
|---|---|---|---|---|---|---|---|
| Core A | 1.5 | 1 | 0.65 | 40 | 40 | 100 | 1 |
| Core B | 0.7 | 0.2 | 3.00E-04 | 40 | 40 | 10 | 0.1 |

# DVS, DPM and Reliability

- Simulate using a "typical day" workload, consisting of video, audio, www and telnet traffic interspersed throughout the day
- 95nm technology, power/performance properties of Xscale PXA270
- Aggressive DPM: large power savings, but reliability loss due to TC
- DVS only: less power savings, but longer MTTF due to EM/TDDB
- Both DVS/DPM give best tradeoff

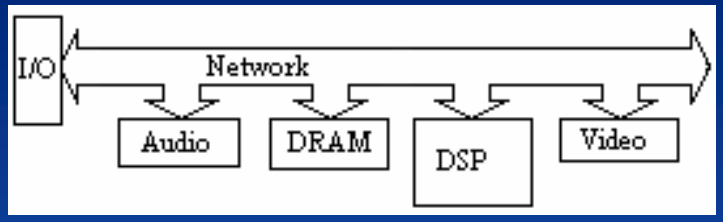
| State | Active (mW) | Idle (mW) | Freq (MHz) |
|---|---|---|---|
| P1 | 925 | 260 | 624 |
| P2 | 747 | 222 | 520 |
| P3 | 279 | 129 | 208 |
| P4 | 116 | 64 | 104 |
| Psleep | 0.163 | 0.163 | 0 |

| Policy | Power | MTTF |
|---|---|---|
| None | 0% | 0% |
| DVS | 35% | 42% |
| DPM (Rmax) | 16% | 6% |
| DPM (ave) | 47% | -12% |
| DPM (Pmax) | 99% | -34% |
| both (Rmax) | 46% | 47% |
| both (ave) | 61% | 45% |
| both (Pmax) | 99% | 34% |



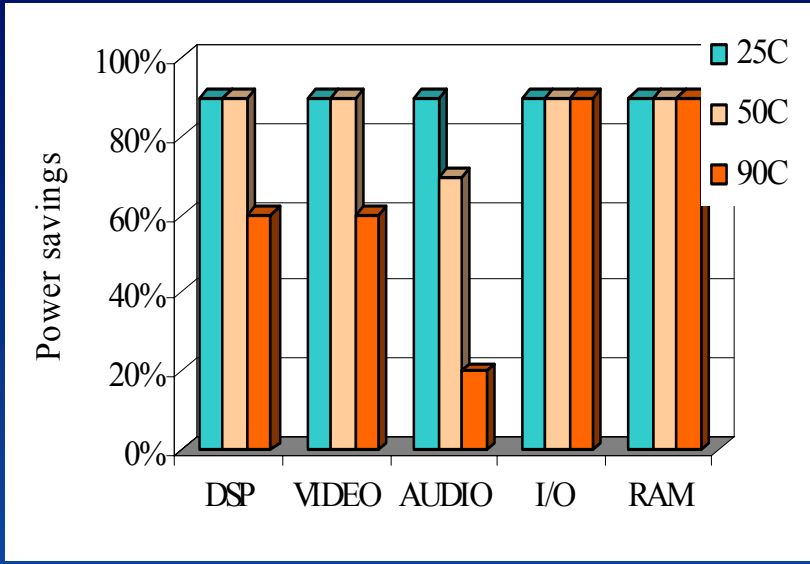
# System design

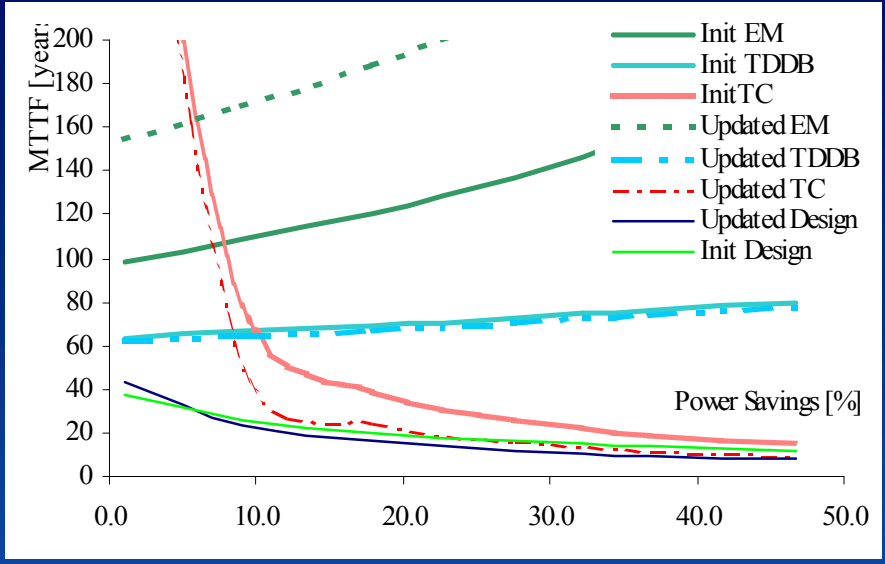
- 95nm technology
- Five cores; standard workloads (audio, video, www etc.)
- MTTF constraint set to 10 years; minimized power consumption




| IP block | $P_{active}$ [W] | $P_{idle}$ [W] | $P_{sleep}$ [W] | $t_{ts}$ [s] | $t_{ta}$ [s] |
|---|---|---|---|---|---|
| DSP (TMS6211) [22] | 1.1 | 0.5 | 0.01 | 250u | 100n |
| Video (SAF7113H) [23] | 0.44 | N/A | 0.07 | 110m | 0.9 |
| Audio (SST-Melody-DAA) [24] | 0.11 | 0.03 | 3.00E-03 | 6u | 0.13 |
| I/O (MSP43011x2) [25] | 1.00E-03 | N/A | 6.00E-06 | 100n | 6u |
| DRAM (Rambus 512M) [26] | 1.58 | 0.37 | 1.00E-02 | 16n | 16n |

# Single Core Design

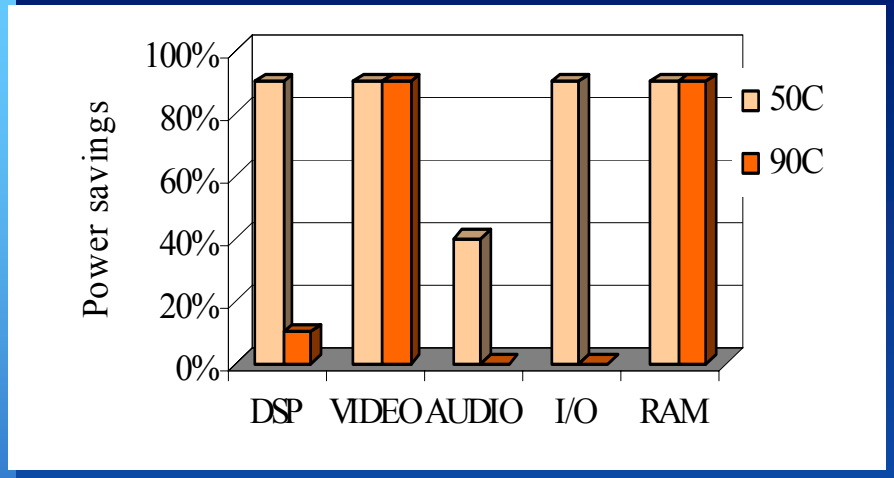


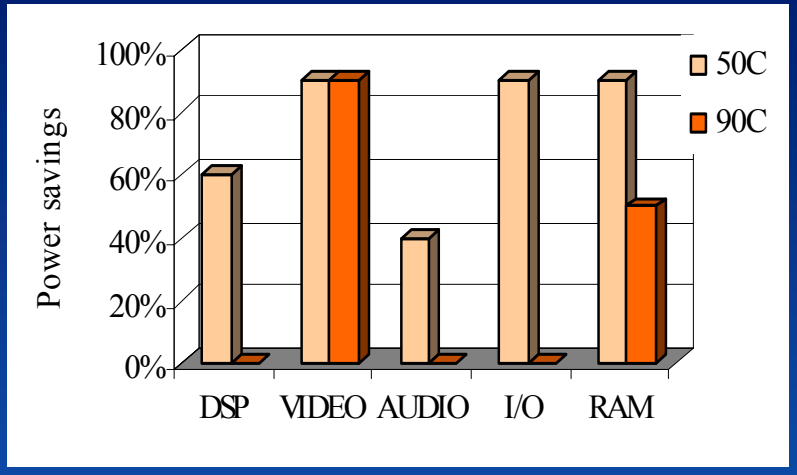




- Maximum power savings achievable given MTTF of 10 years are at 90% for all cores and temperature ranges except for DSP, Video and Audio at 90 C due to TC mechanism
- Design change effect - widening metal lines – Current density down by 20%, core area up by 5%, temperature down by 2%, but TC up by 10%

# System design with redundancy

- Standby off and standby sleep redundancy power savings with MTTF set to 10 years




- System meets MTTF of 10 years when one more redundant core in standby off mode is added to DSP, Audio and I/O; power savings of 40% are achieved

# Summary

- Reliability is strongly affected by both DVS and DPM
- This work presents an integrated methodology for analysis, optimization and management of reliability and power consumption:
  - Simulator gives fast feedback on topology design and system characteristics for a wide range of operating conditions
  - Optimizer provides a policy capable of giving an optimal implementation of reliability and power management control
- Results obtained for a number of integrated systems implemented in 95nm technology show:
  - Large dependence between power management policy and reliability due to tradeoff between EM, TDDB and TC effects
  - 40% power savings on top of meeting MTTF of 10 years for an integrated system consisting of five cores with redundancy

# Relevant Publications

- T. Šimunić , K. Mihic, G. De Micheli: “Optimization of Reliability and Power Consumption in Systems on a Chip,” PATHMOS’05.
- T. Šimunić , K. Mihic, G. De Micheli: “Reliability and power management of integrated systems ,” keynote at DSD'04.
- T. Šimunić , S. Boyd, P. Glynn: “Managing power consumption in Networks on Chips,” IEEE Transactions on VLSI, pp. 96- 107, Jan 2004.
- T. Šimunić , L. Benini, P. Glynn, G. De Micheli: “Event-driven power management,” IEEE Transactions on Computer-Aided Design, July 2001.
- For more see: http://akebono.stanford.edu/users/tajana/Papers.html