ECE 225
High-Speed Digital IC Design
Lecture 10

Nanoscale Power and Thermal Management:

Analysis of Chip Cooling—device, circuit, system considerations, Local vs Global Cooling

Prof. Kaustav Banerjee
Electrical and Computer Engineering
E-mail: kaustav@ece.ucsb.edu

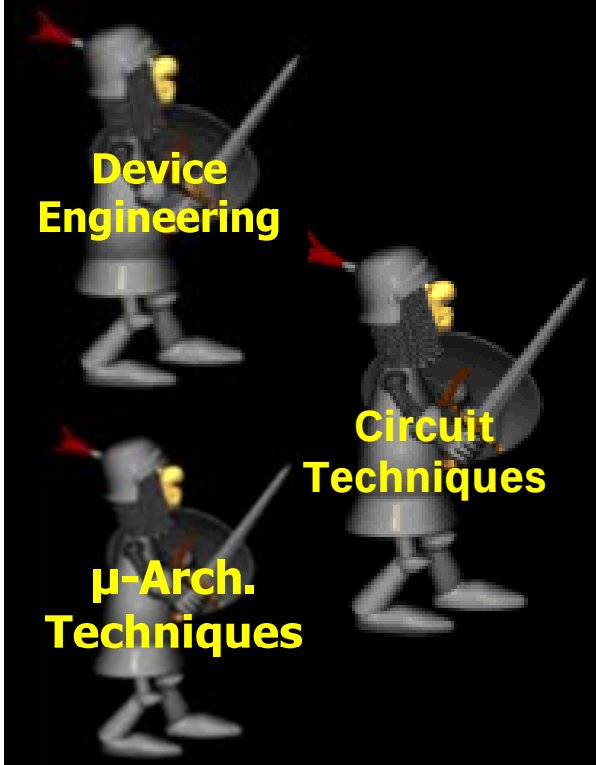
IC Cooling for Power/Thermal Management?

Lin et al., TED Jan 2008

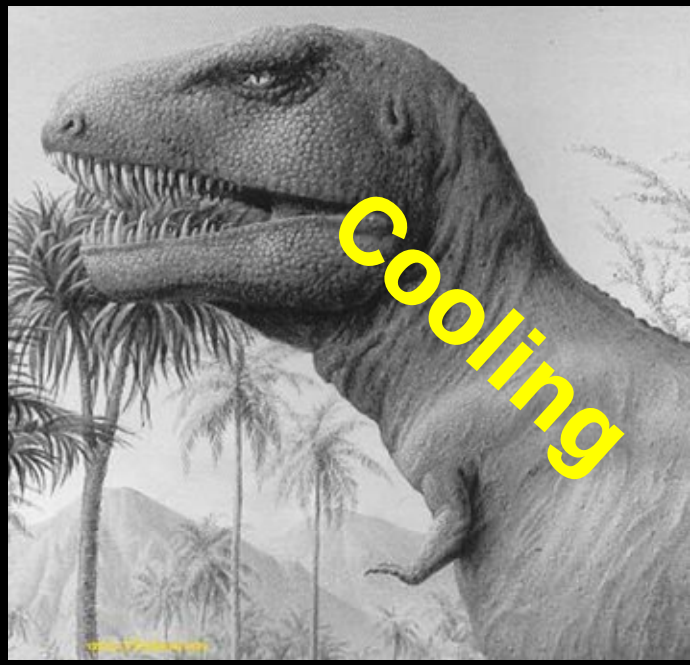
Efforts on Low-Power…..without hurting performance …

- Device Engineering
  - **Enhanced Channel Mobility – Strained Silicon**
  - **Reduced Gate Leakage – High-K Gate**
  - **Reduced Junction Leakage – Nitrogen Doped**
- Circuit Level
  - **Adaptive body-biasing, Dual V_{DD}-CMOS, Dual-V_{th}**
  - **Sleep Transistor, Clock/Power Gating**
- Micro-Architecture Level
  - **Multi-Core**

Stronger Knob is needed



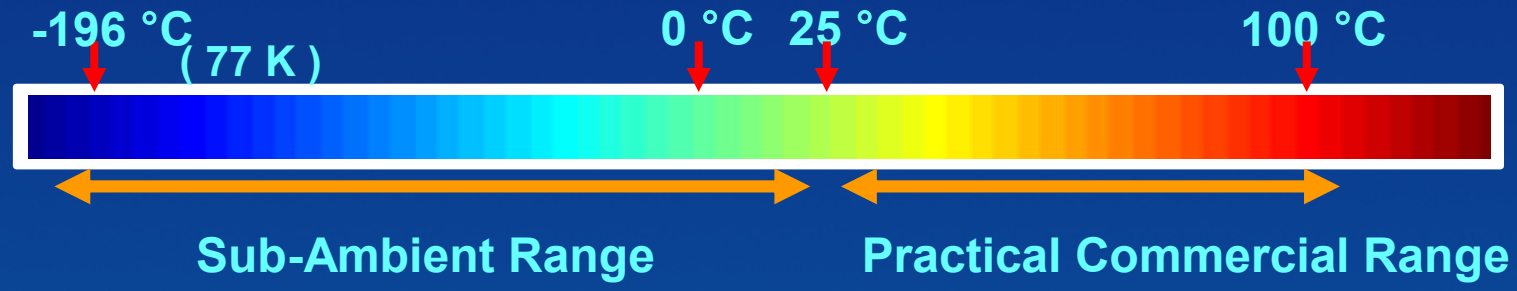
Power Dissipation



Cooling is the Knob !

Does IC Cooling Make Sense?

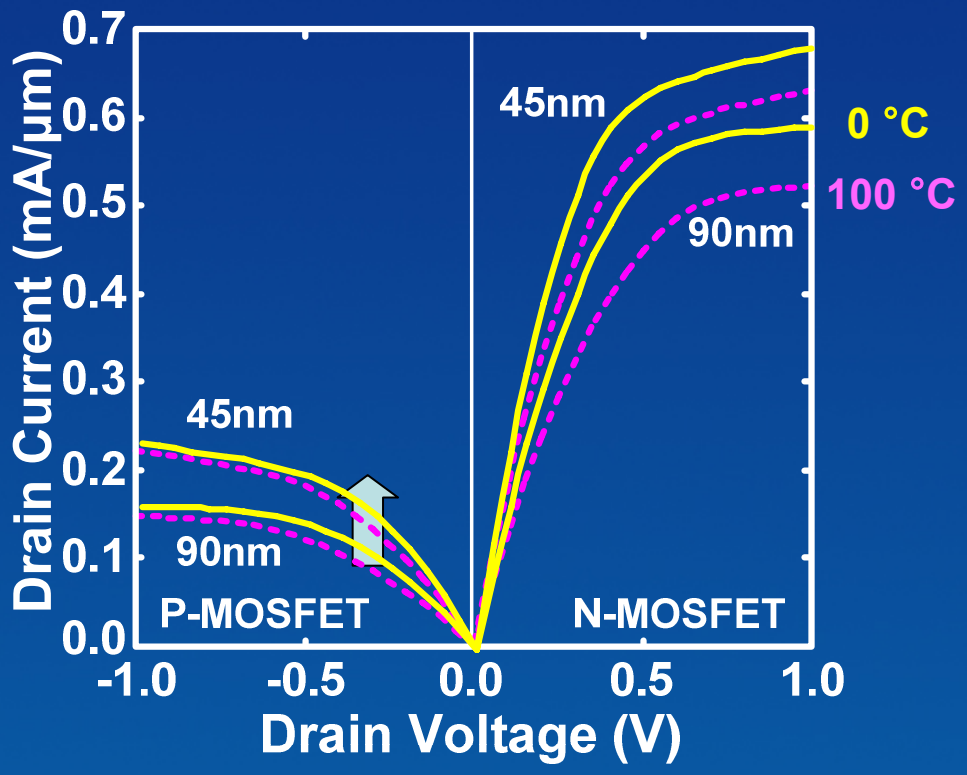
- Range of IC Cooling

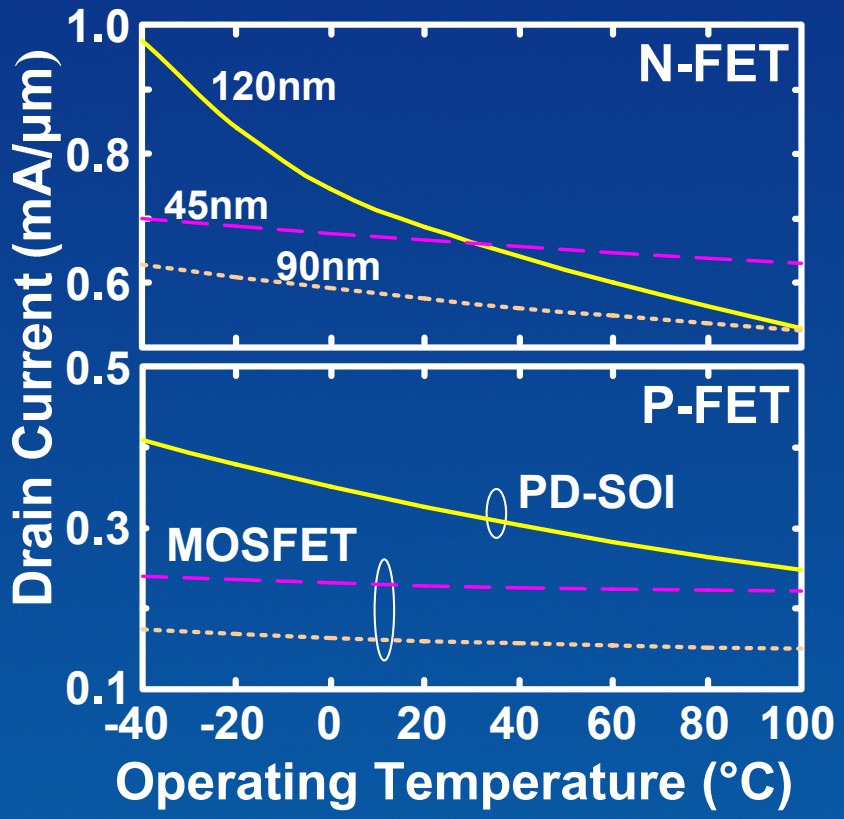


- Prior Research on Cooling
  - Analysis under Cryogenic Condition
    - Operating near liquid nitrogen temperature (77 K) requires complex cooling facilities
    - NOT for commercial purpose
  - Analysis of Performance under Cooling
    - Does NOT consider electrothermal couplings
    - Does NOT consider the system level power dissipation

Cooling---Benefit at the Device Level

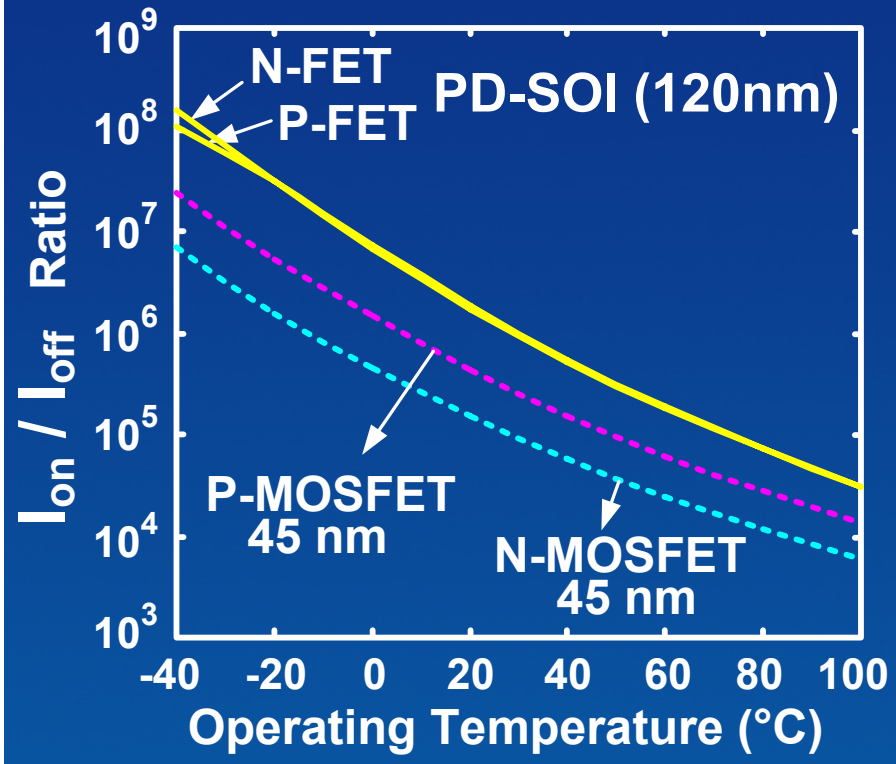
Transistor Drive Current Increases

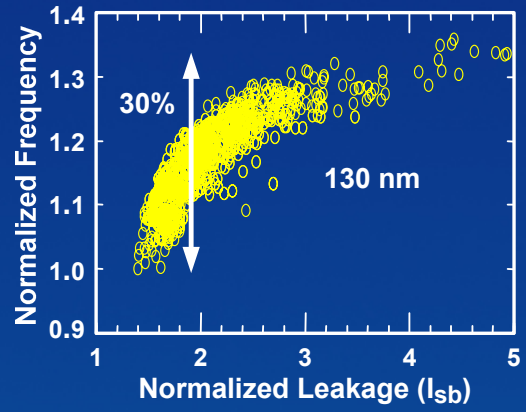




Higher drive current can be achieved by cooled operation across all technology nodes due to higher carrier mobility

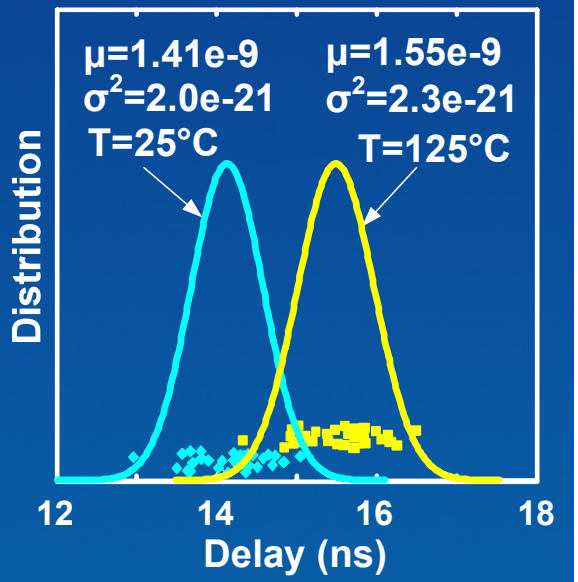
Cooling---Benefit at the Circuit Level





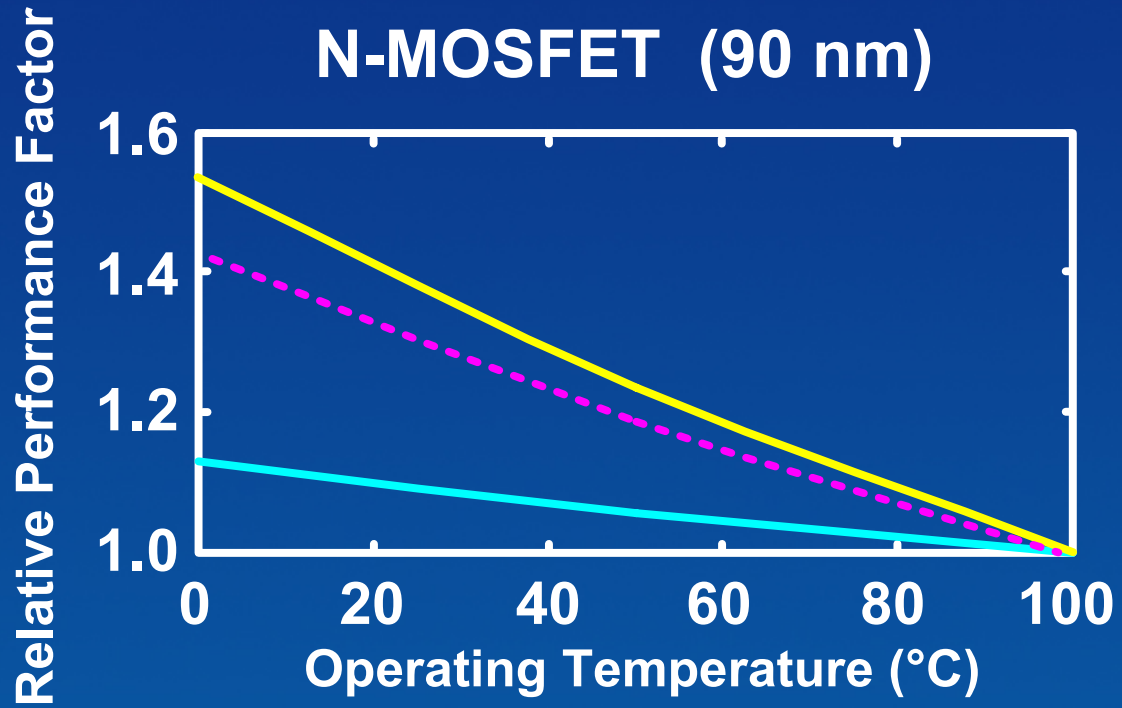
S. Borkar, Intel

9-stage Inverter Chain



- Enhance I_{on} to I_{off} ratio
- Reduce propagation delay and variance
- Benefit back-end performance and reliability

Further Exploiting the Benefit of Cooling



I. Aller et al., ISSCC 2000

Same Device

Same V_{th}

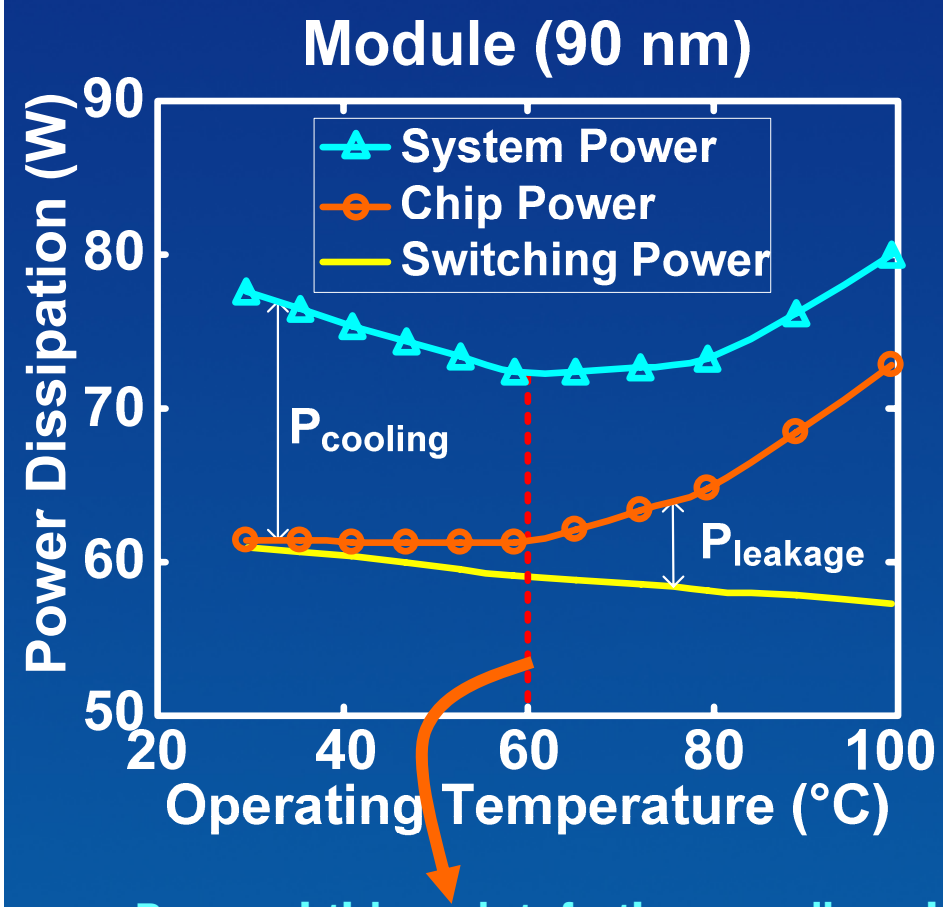
Applying Forward Body Bias (FBB)

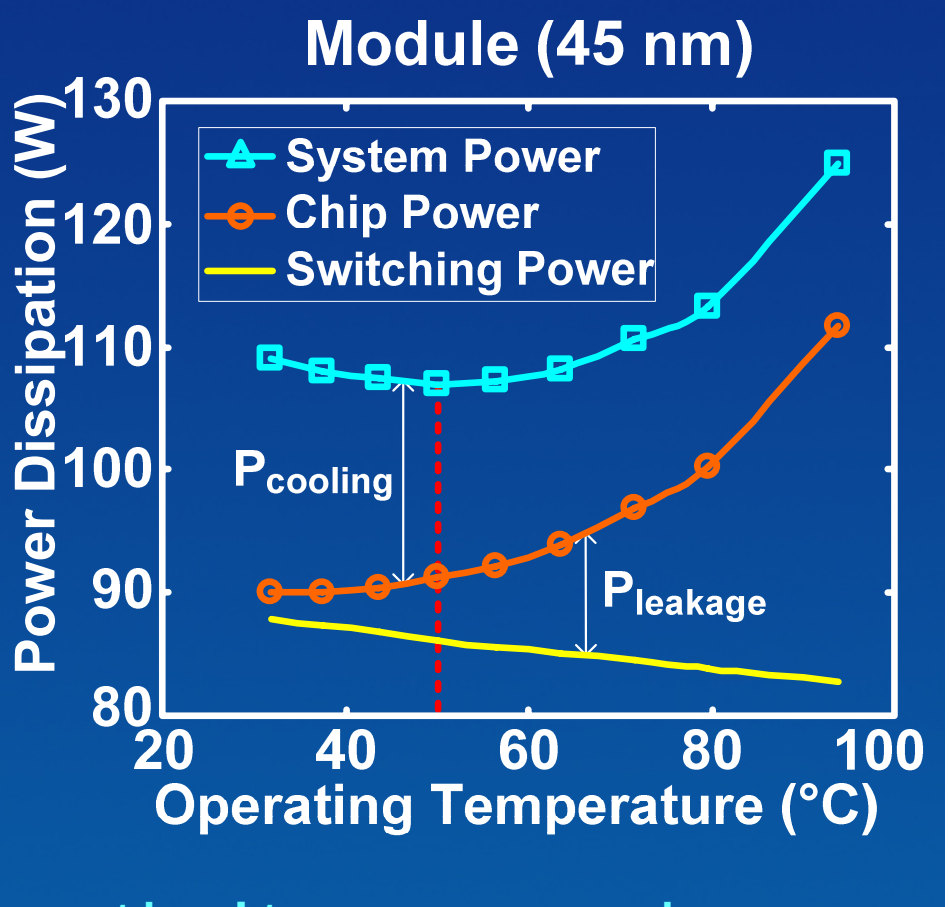
Same Leakage

Increase FBB Further

Performance can be further improved by different design strategies

Cooling---System Level Considerations

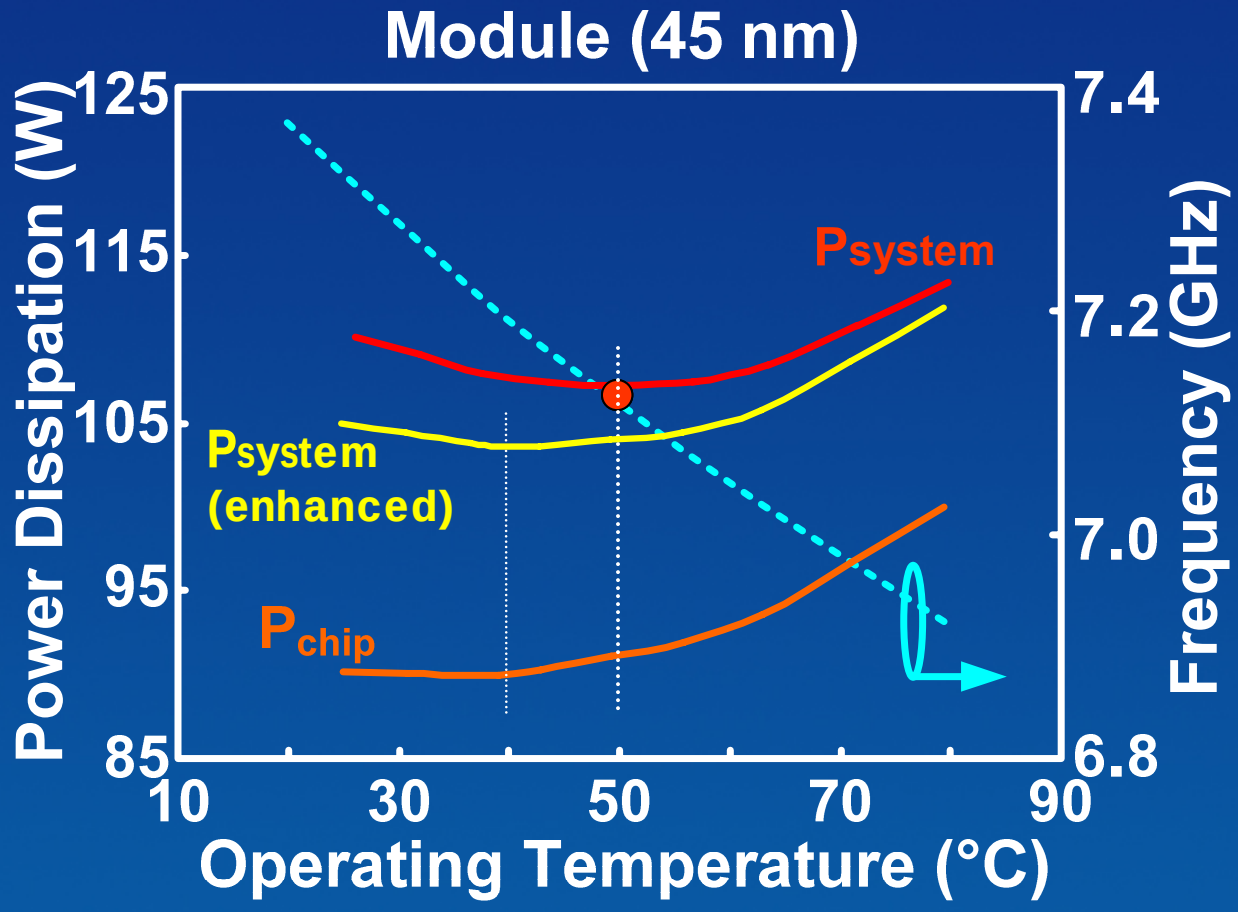




Beyond this point, further cooling does not lead to any power saving.

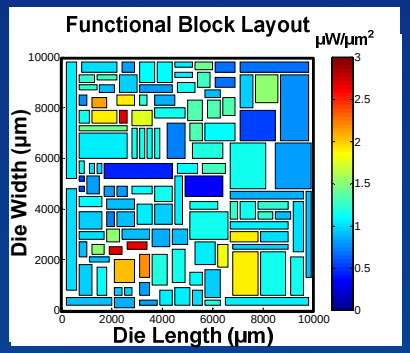
The limit occurs at a lower temperature as technology scales.

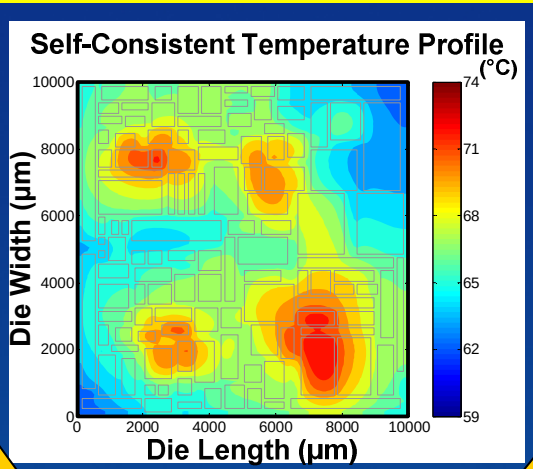
Extending the Practical Limit of Cooling



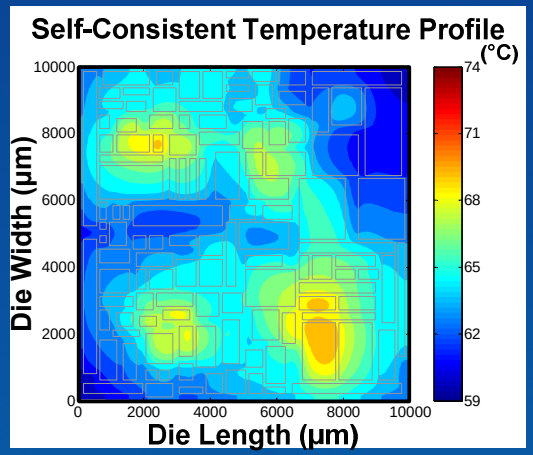
Practical limit can be further extended towards a lower temperature and higher performance is achieved by enhancing cooling efficiency

Hot-Spot Management



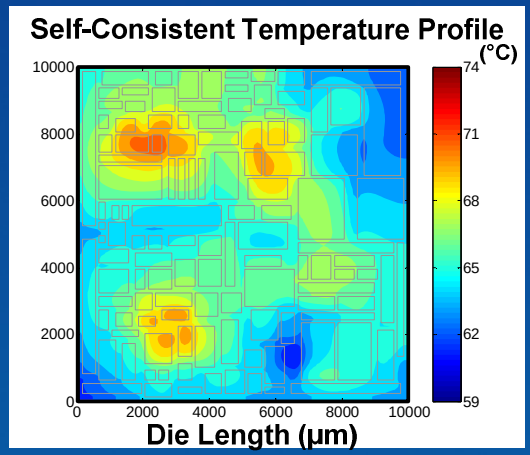


T_{MAX} decreases
but hot-spots remain



Global Cooling

Thermal resistance reduces 20%



Localized Cooling

Localized cooling will be more effective for hot-spot management

Conclusion

- For power/thermal management of nanoscale CMOS ICs, **electrothermal couplings** must be incorporated for trading-off chip-level power, performance, reliability, and cooling cost.
- An accurate leakage-aware **self-consistent** power and silicon-substrate **temperature-profile-estimation** technique has been developed for nanoscale CMOS technologies.
- A systematic methodology for power-performance optimization has been developed to **bring design and packaging closer**.
- **Chip cooling** combining device, circuit, and system level considerations can extend CMOS scaling.

Additional Issues….

- Extend CMOS scaling
  - Time per node can be expanded via cooling
- Application Based
  - High-performance servers
  - Hand-held applications
- Implications for Ultra Low-Voltage Applications
  - Positive Temperature Dependence