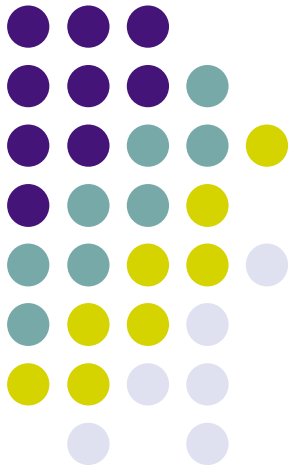
Lost in the Bermuda Triangle:
Energy, Complexity, and Performance

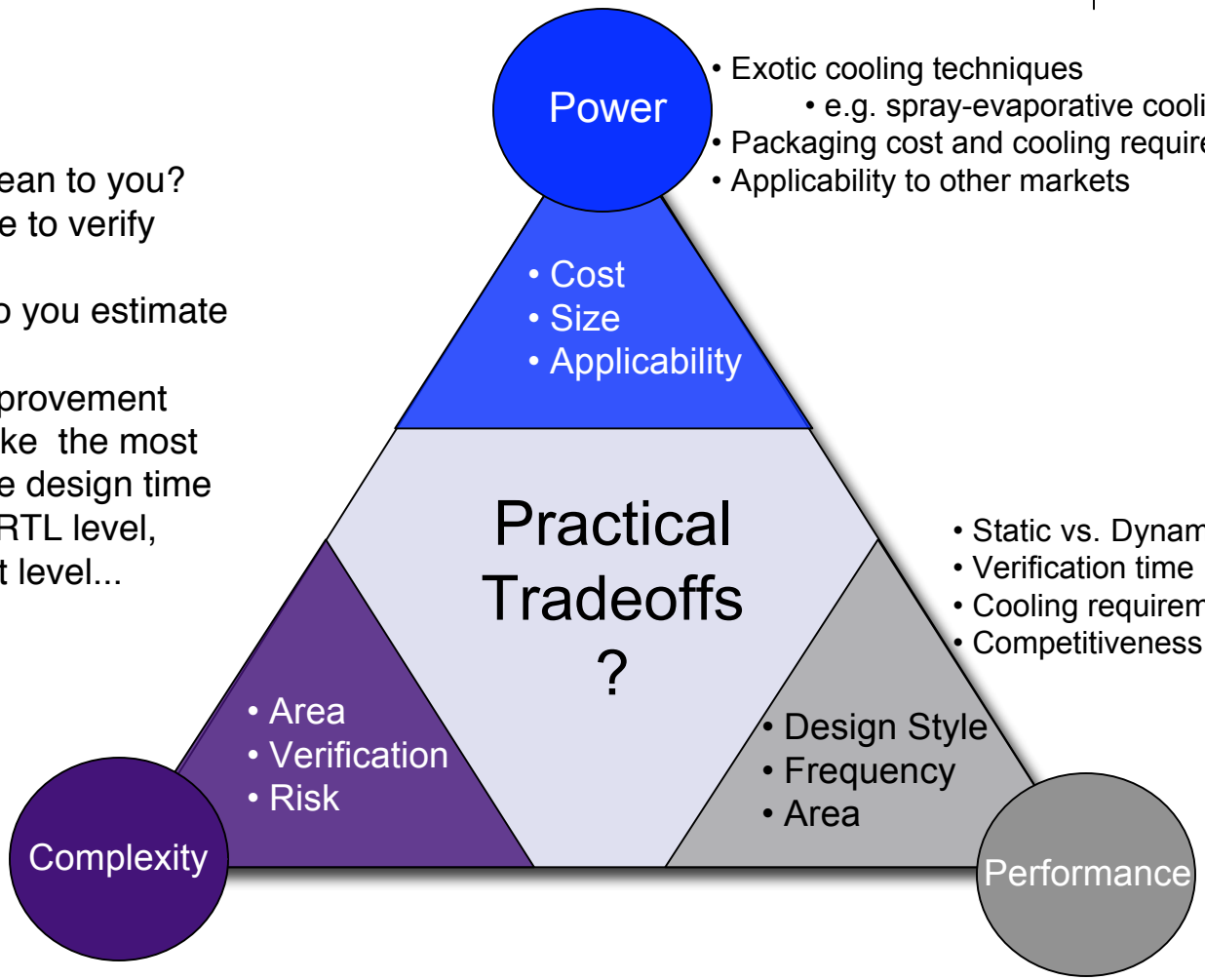
Dennis Abts

Cray Inc.

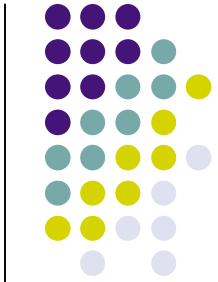


Exploring Uncharted Waters

1. what does complexity mean to you?
2. What takes the most time to verify in your designs?
3. On your projects, how do you estimate design time?
4. In which areas would improvement in the state of the art make the most difference in reducing the design time
5. What can one do at the RTL level, architectural level, layout level... to reduce complexity?

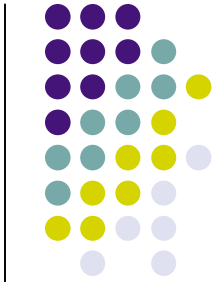


Design Complexity



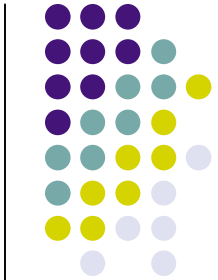
- Aggressive implementation techniques (speculation, O-o-O, etc) complicate pipeline design and verification
  - SIMD architectures (e.g. vectors) provide simpler control logic while still yielding high flop rate
- Multi-core is here to stay
  - Many cores (<8?) are easier to design and verify than a heavyweight processor with lots of aggressive implementation techniques.
    - e.g.Sun's Niagara
  - Not without it's own problems…

The Woes of Multi-core



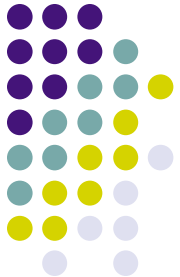
- While each core is simpler, the sharing of inter-core resources (memory controller, network/IO links, etc) is more complicated
- Interconnect among the cores are complicated by conventional crossbars that scale as N^2
  - On-chip networks built from hierarchical crossbars will evolve
- Coherence among the memory hierarchy (private and shared caches) in the cores
- Verification is simplified with the abstraction of many replicated instances of identical logic

Verification Complexity

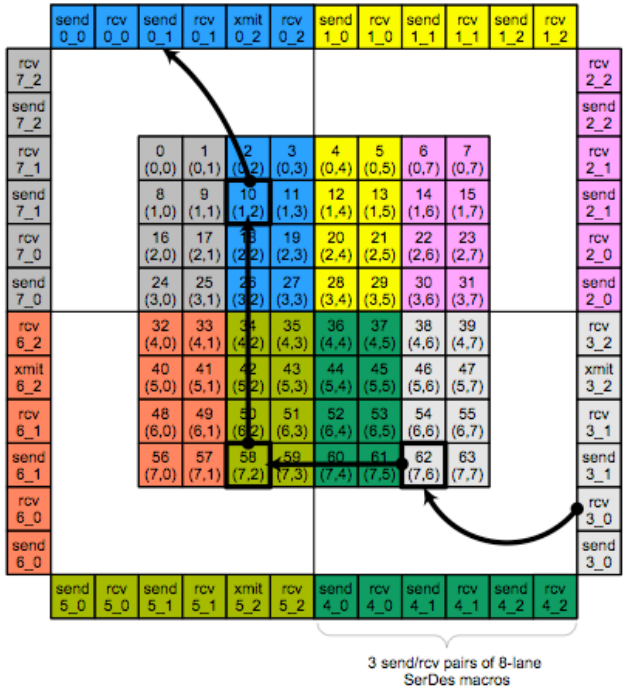


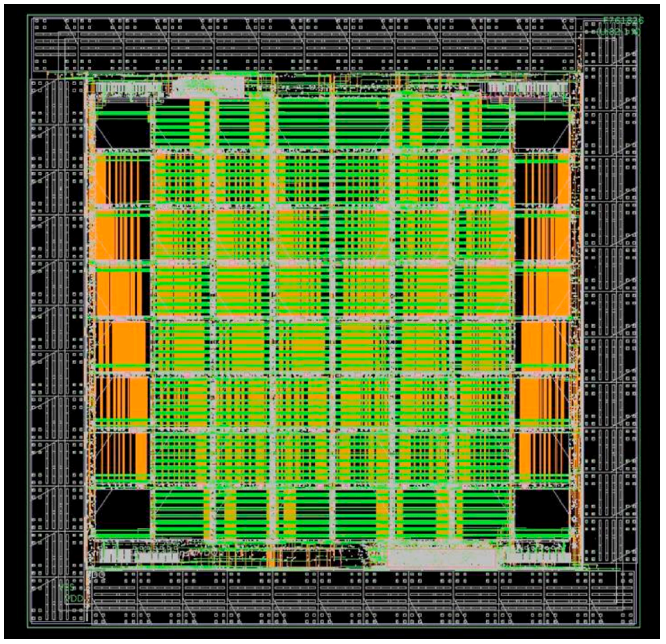
- Protocols, pipeline interaction (instruction permutations), cooperating state machines
- Abstraction and Verification Methods
  - Reference verification methodology (RVM)
  - Transactional verification
  - Assume-Guarantee reasoning to validate behaviors among cooperating logic blocks
  - Hardware Verification Languages
    - Constraint solvers for constrained random verification
    - Formal-informal methods are coming to fruition
- Coverage analysis and metrics for establishing when verification is “done” remains problematic

Innovative Architectures to Mitigate Design Complexity

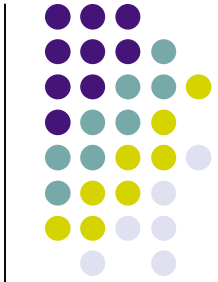


- Tile-based architecture that replicates many simple “tiles” to avoid long global wires.
  - Simplify verification -- since each tile is identical
  - Reduce implementation time
    - Simplified arbitration - easier to close timing





The Wildcard!

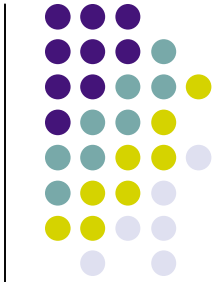


- Error Handling and Verification
- Building reliable systems from unreliable components is becoming increasingly difficult
  - Process variation at feature sizes <90nm
  - Soft errors from natural radiation, and electrical noise
  - Increased cost and complexity of error correcting codes and error handling protocols at the system level

Performance

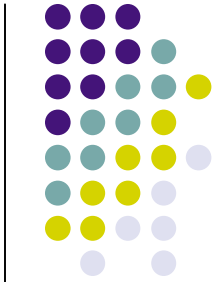
- High-performance microprocessors and systems have very little freedom to make performance tradeoffs for:
  - Ease of implementation and verification
  - Energy efficiency
- Embedded applications have more latitude to make performance vs. complexity tradeoffs
- Worst-case cooling and power dissipation is becoming onerous

Performance…



- Custom vs. ASIC design styles and the implications
  - Std. cell ASIC design is less effort
    - Lower frequency, less control of technology
  - Custom chips can be tuned for technology
    - Domino logic vs. static CMOS
    - Cell geometries are tuned for area/performance tradeoff
  - Mixed? ASIC with Custom logic macros
    - Cray X1 and BW processors take this approach
    - Custom logic used for critical performance areas (func units) and ASIC logic used elsewhere for ease of implementation and verification

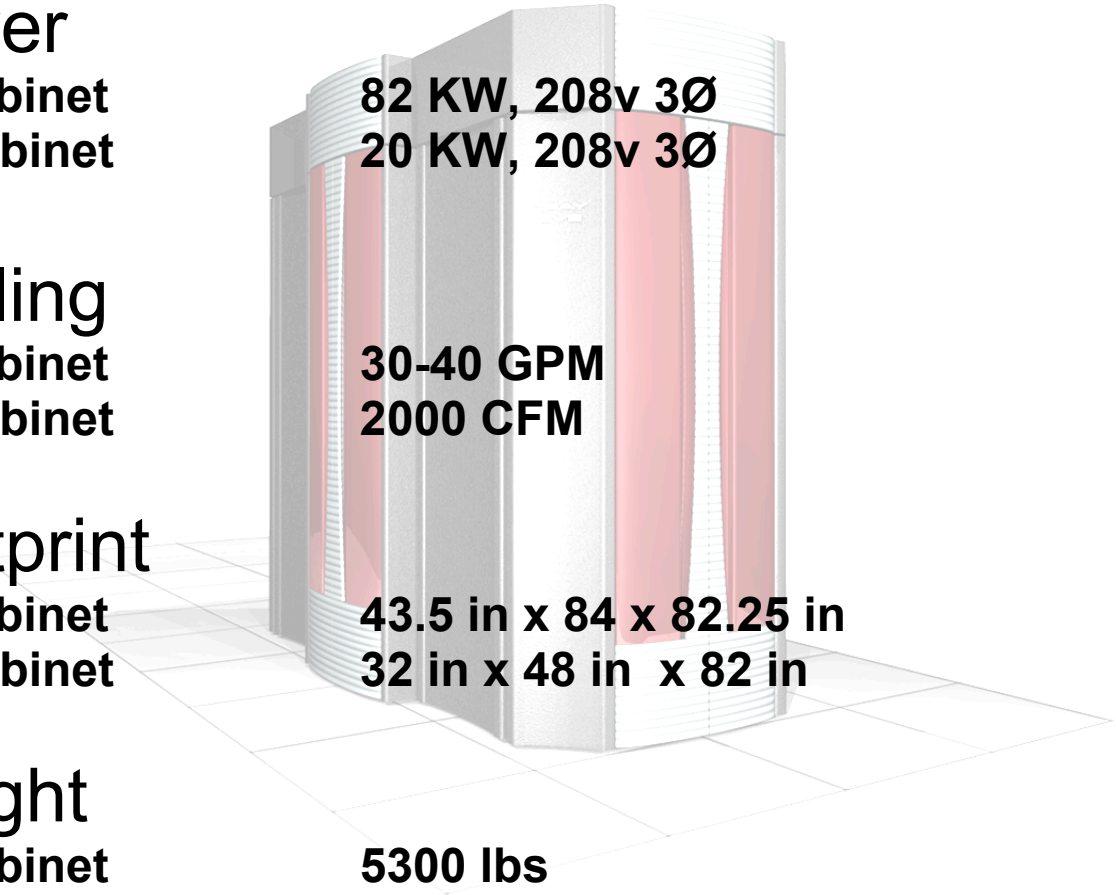
Power and Cooling



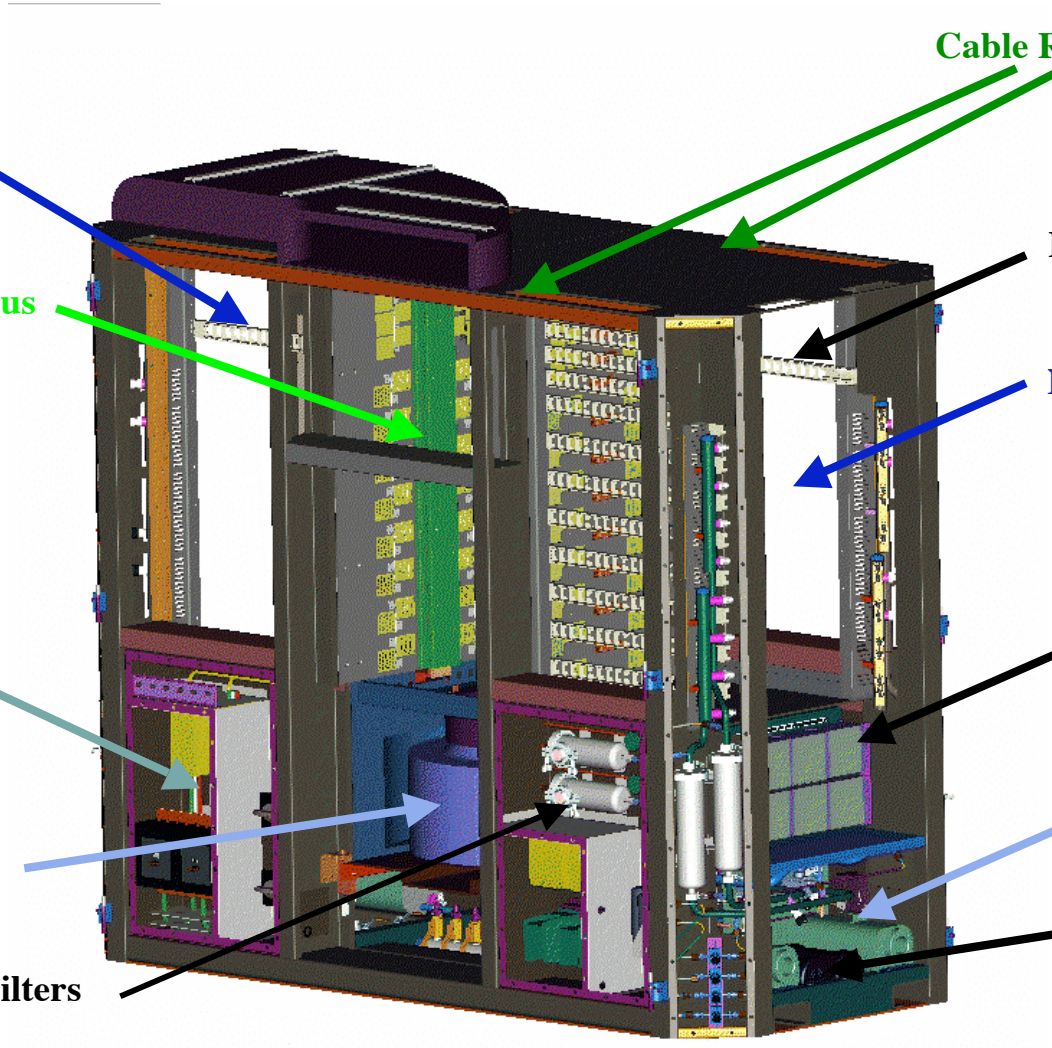
- Scalable multiprocessors must dispense of a LARGE heat load
  - Many KW per cabinet
  - Large systems will have many cabinets
  - See [Pautsch, CoolCon 2005] for details of worst-case cooling…

System Specifications

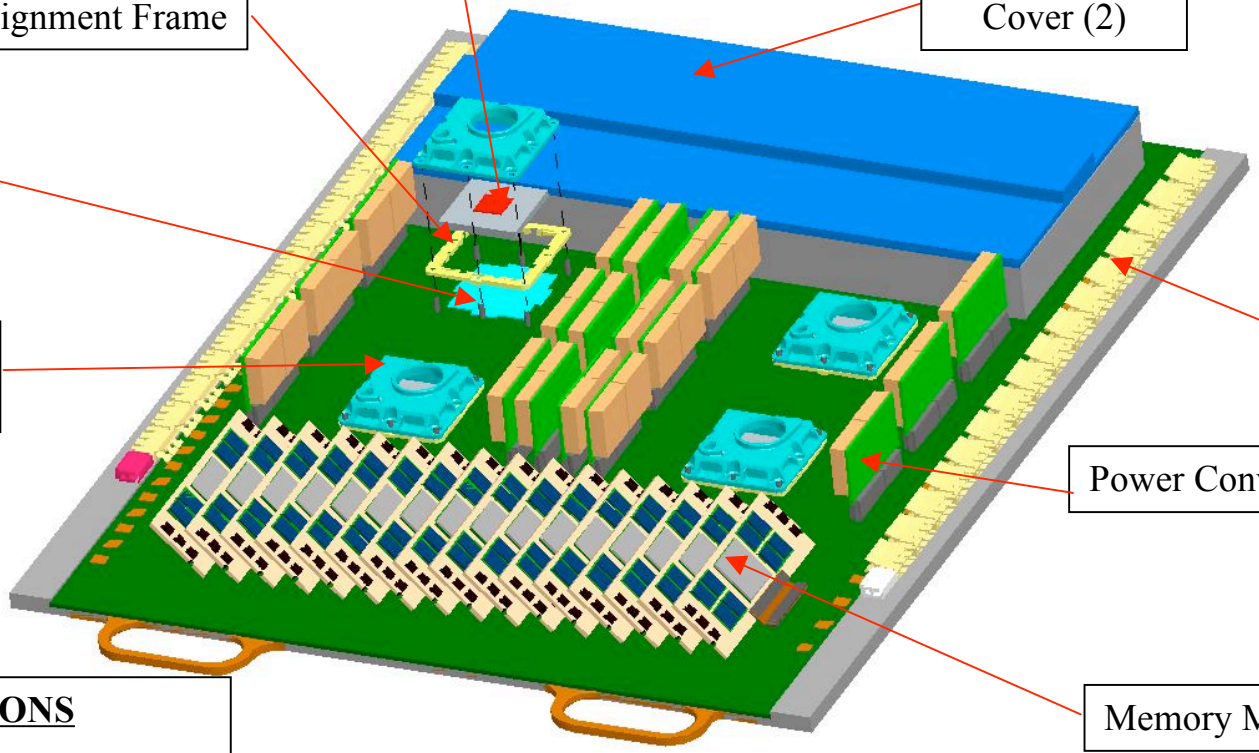
- Power
  - **LC Cabinet** — **82 KW, 208v 3Ø**
  - **AC Cabinet** — **20 KW, 208v 3Ø**
- Cooling
  - **LC Cabinet** — **30-40 GPM**
  - **AC Cabinet** — **2000 CFM**
- Footprint
  - **LC Cabinet** — **43.5 in x 84 x 82.25 in**
  - **AC Cabinet** — **32 in x 48 in x 82 in**
- Weight
  - **LC Cabinet** — **5300 lbs**
  - **AC Cabinet** — **1800 lbs**



Liquid Cooled Cabinet



Liquid Cooled Module



SPECIFICATIONS

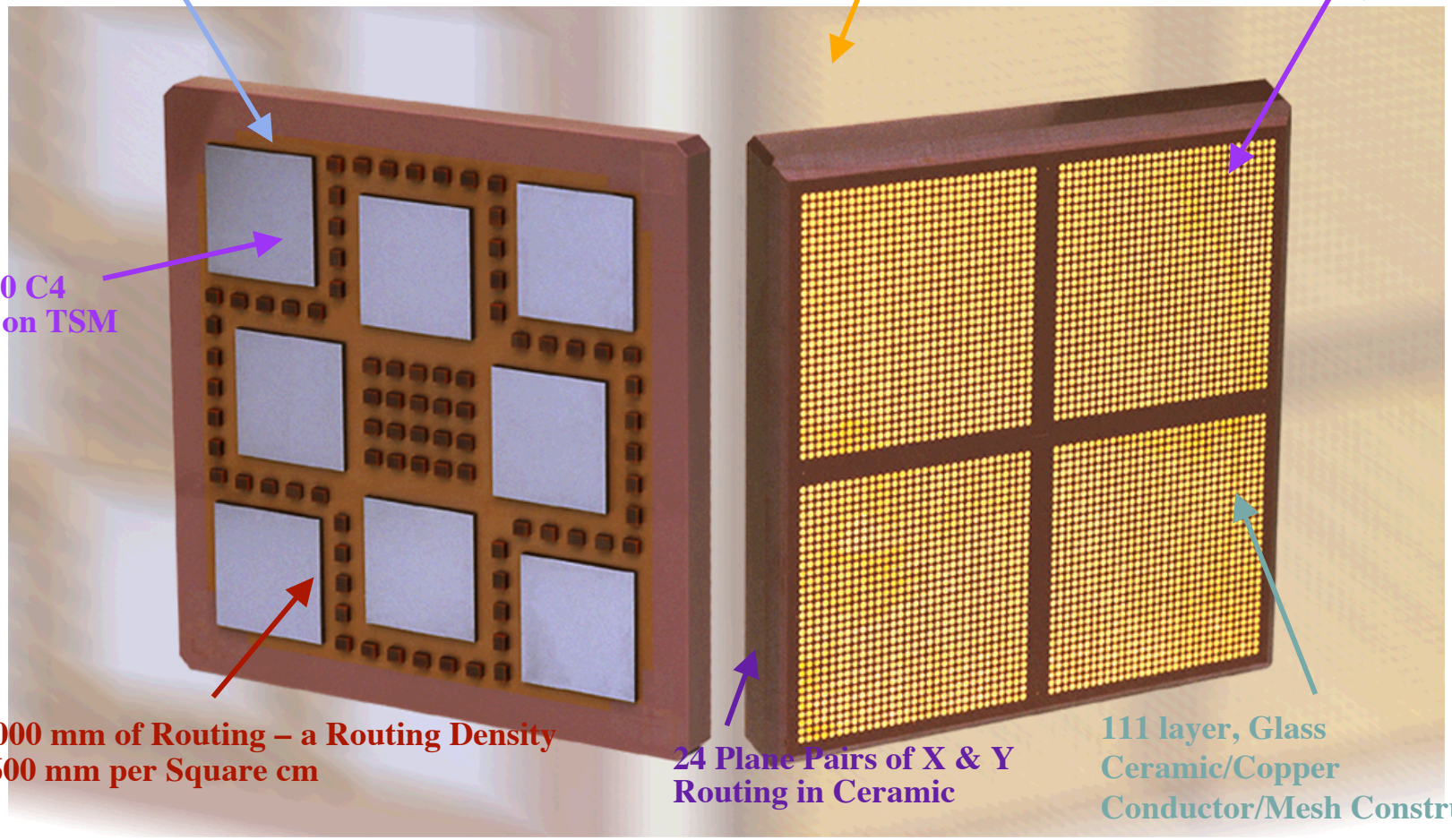
Module Power - 7000 watts

Liquid Flow Req. - 25-35 lpm

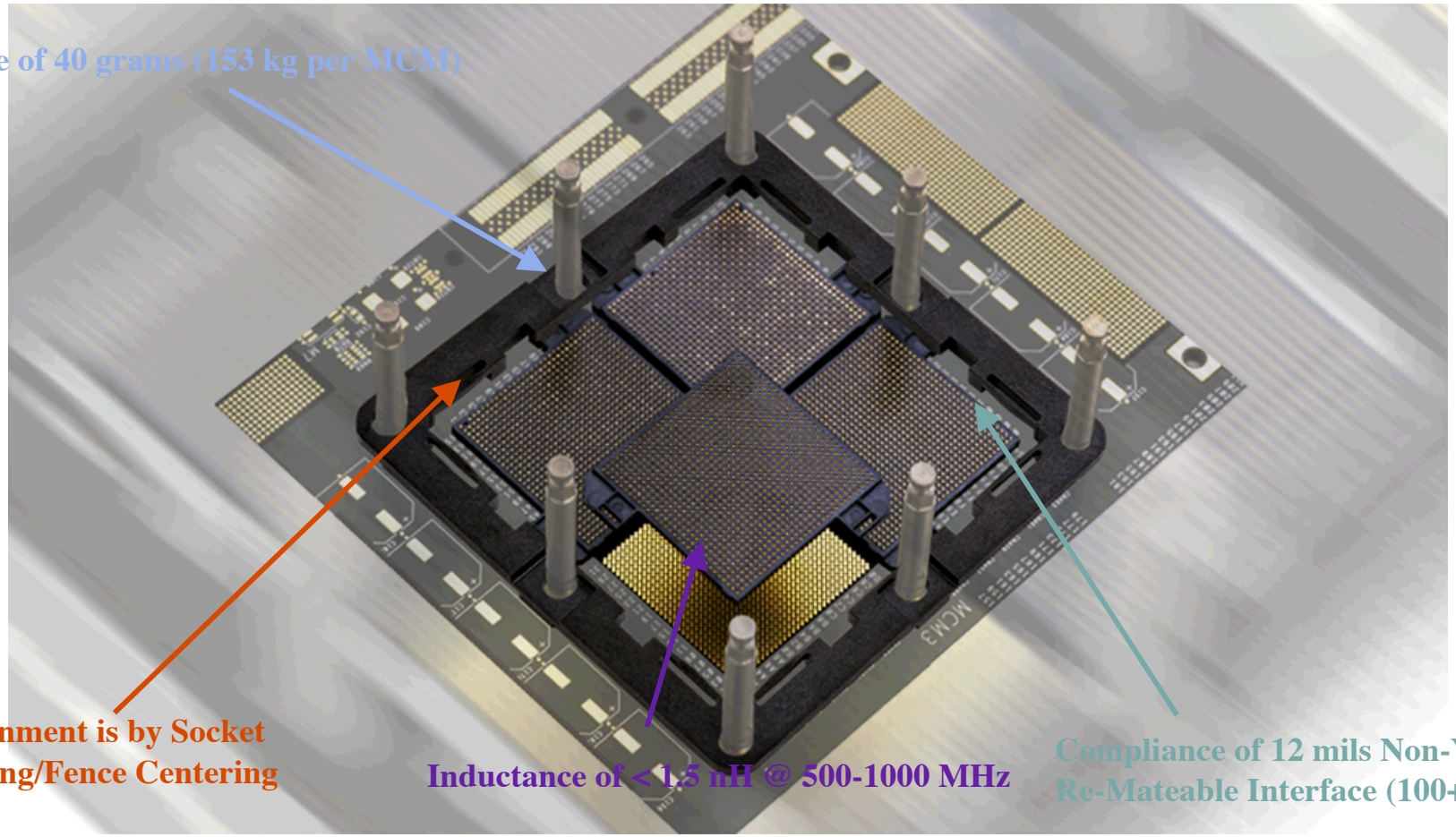
Size - 23 in x 28.5 in

- 583 mm x 724 mm

Multi-chip Module



Compliant Interconnect

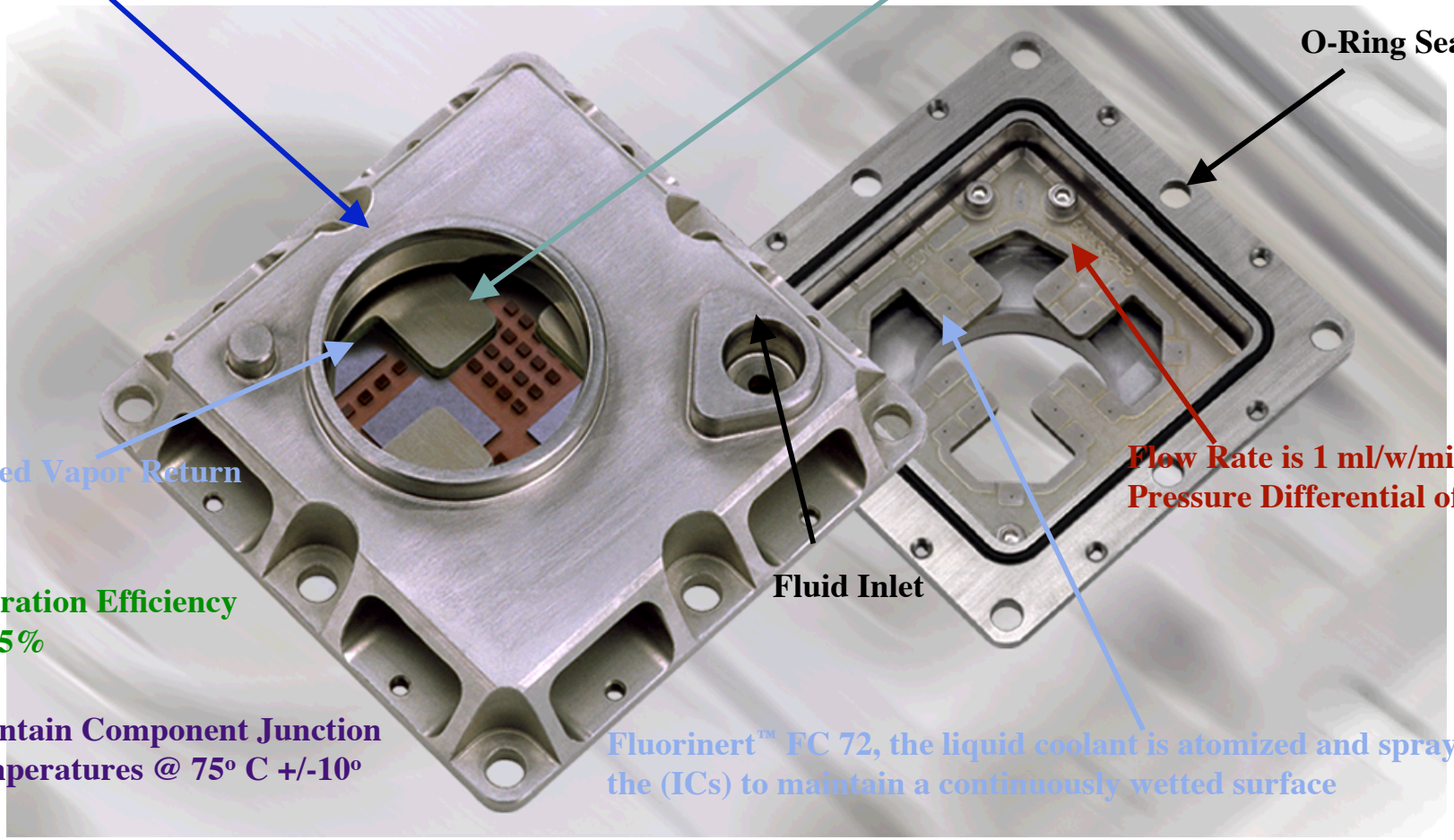


Spary Cap Assembly

Heat Flux of IC's on the MCM are:

- **P+ Chip Heat Flux - 45 W/cm²**
- **E+ Chip Heat Flux - 20 W/cm²**

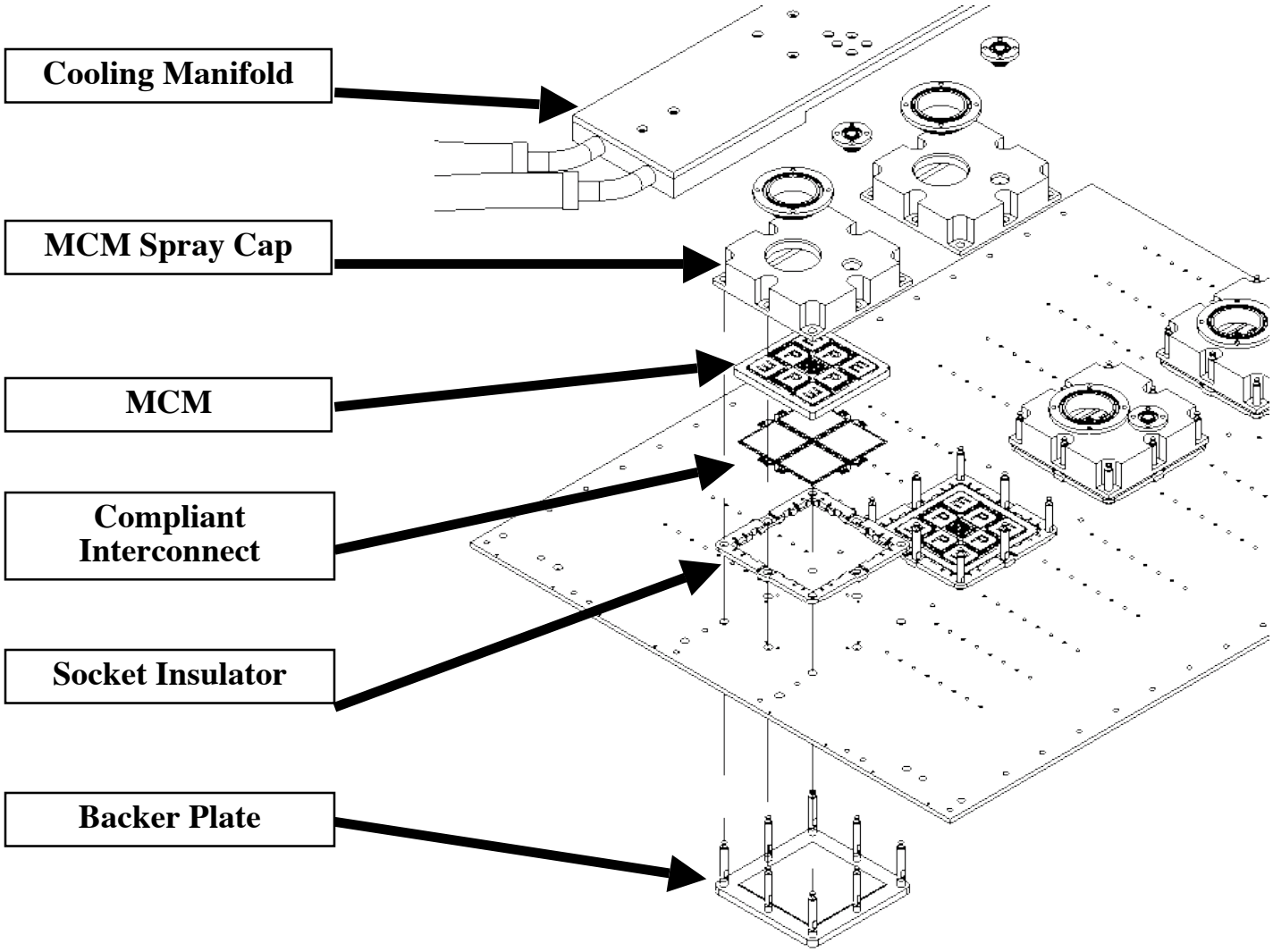
IC Junction Temperature of 85° C with Heat Flux Density up to 70 W/cm²



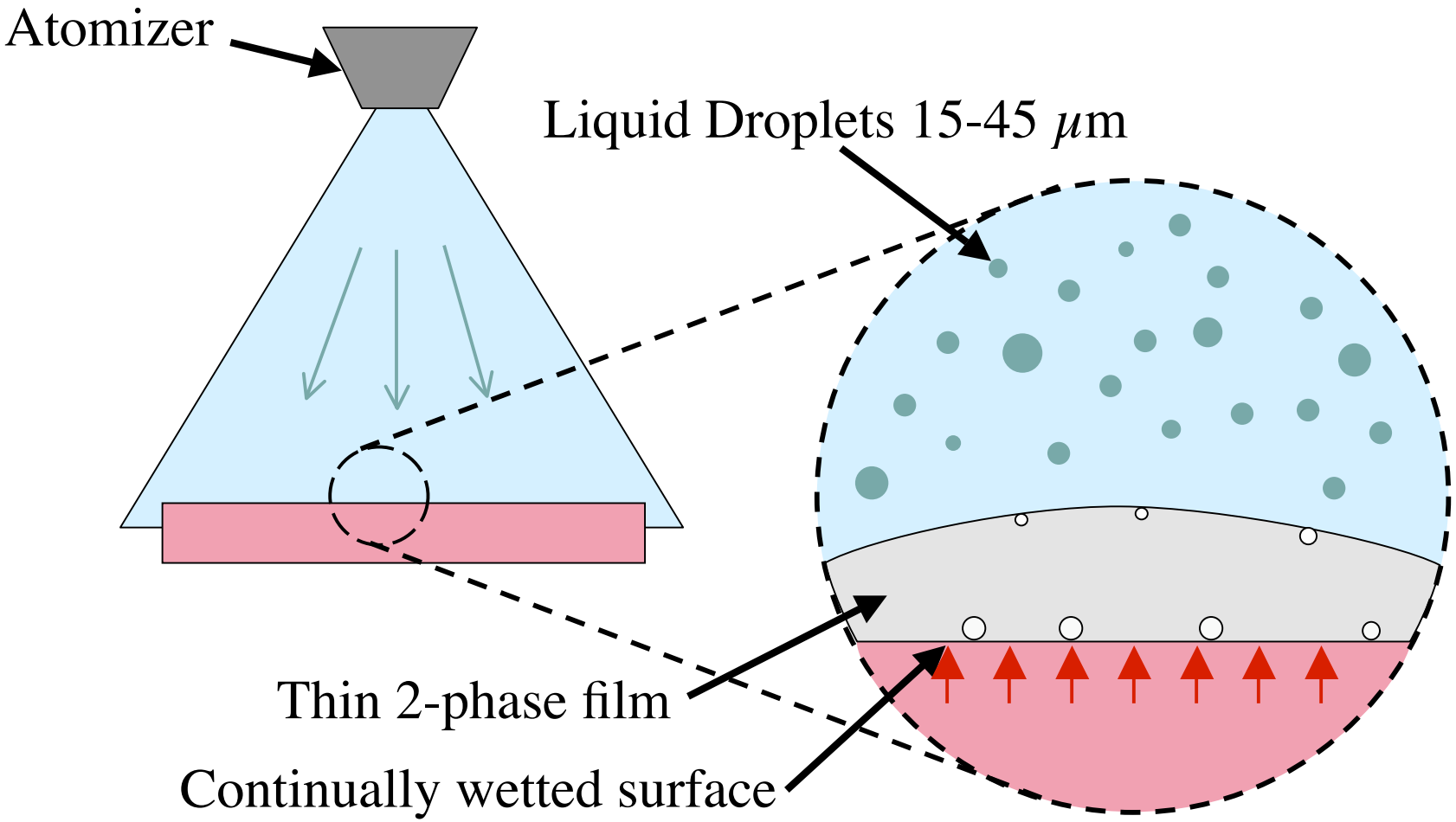
Evaporation Efficiency Of ~ 25%

Maintain Component Junction Temperatures @ 75° C +/-10°

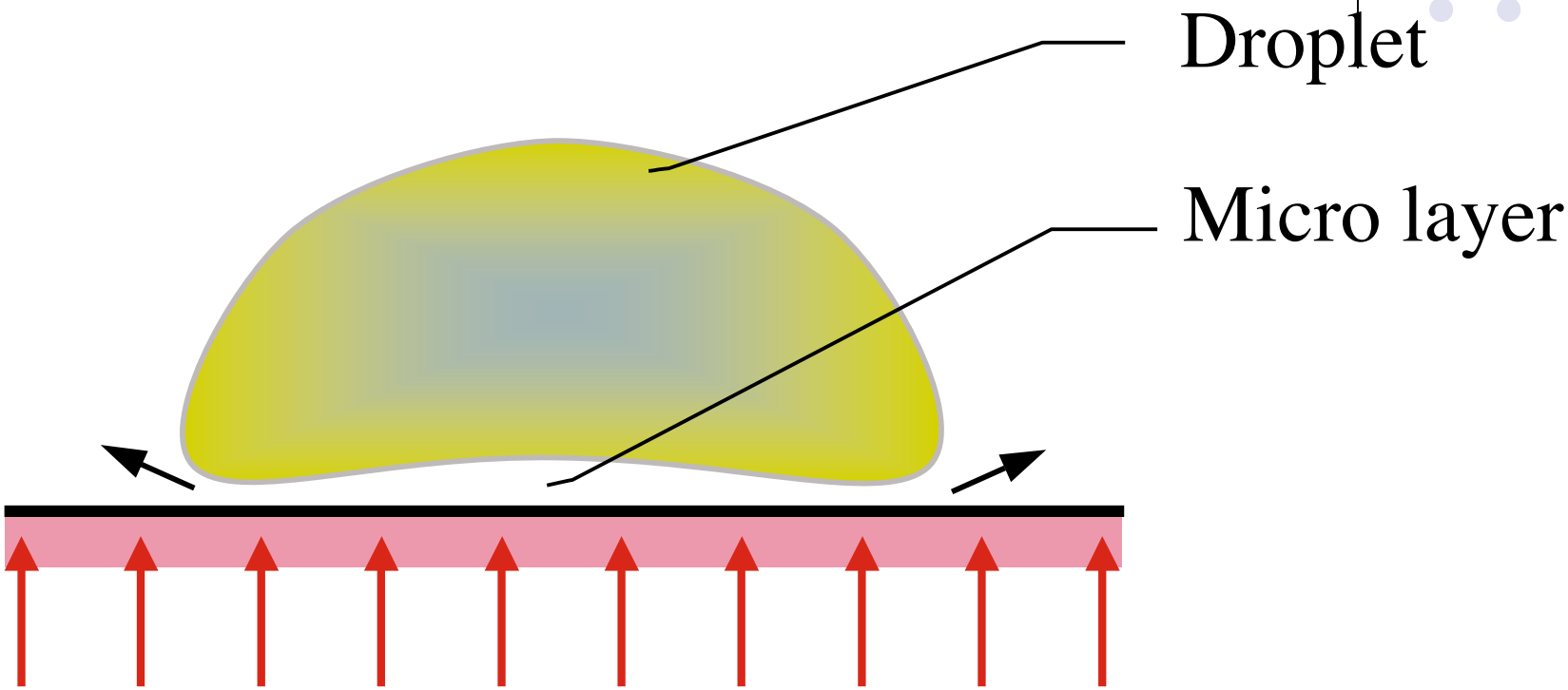
MCM Assembly



How Spray Cooling Works



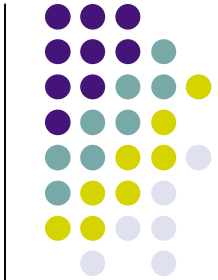
Leidenfrost Effect



Evaporation of the bottom portion of the droplet forms an insulating micro layer of vapor

to 25

Discussion Points



- Large scale multiprocessors (>1K processors) have unique design challenges to balance tradeoffs surrounding
  - Design complexity
  - Verification complexity
  - Performance and Power
- Each node may have multiple custom chips
  - Processor, Interconnect, Memory controller, DRAM parts
- A system with thousands of nodes can easily have >10K components
- Building highly scalable and reliable systems from unreliable components is becoming a daunting task
  - Verification complexity and design complexity of error handling

Thank You

Dennis Abts, Cray Inc