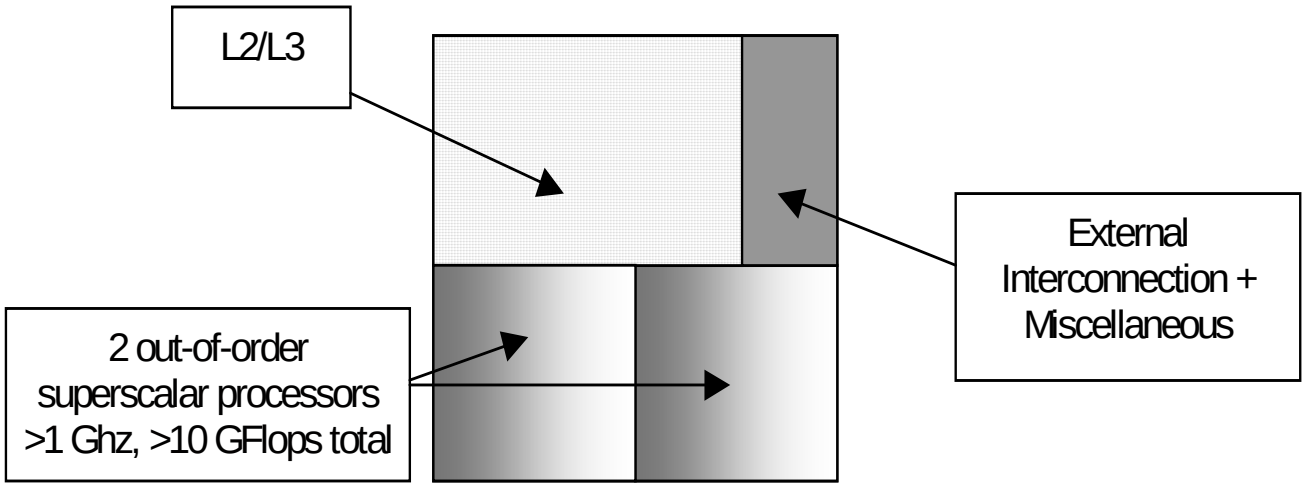
# Evolution of Computer Architectures

Ravi Nair
IBM Thomas J. Watson Research Center
Yorktown Heights, NY
July 10, 2008

*Thanks to Dan Prener*

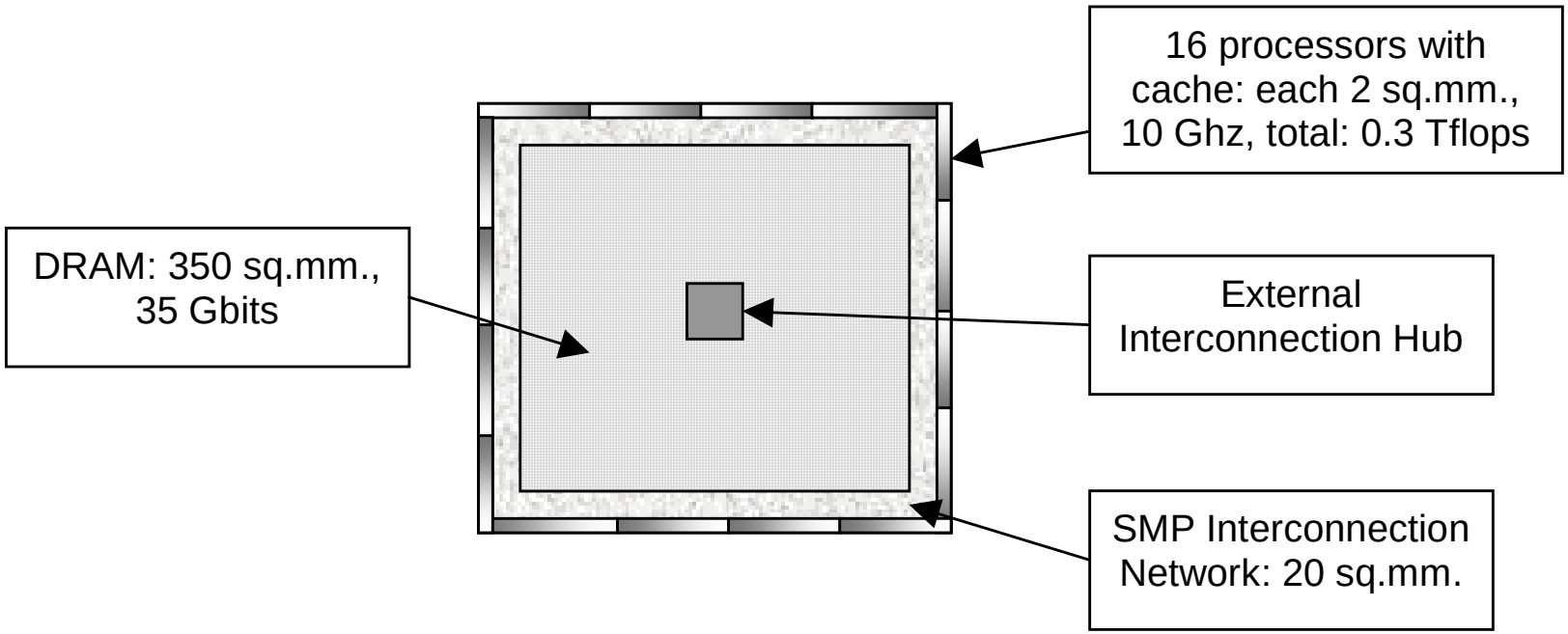
*for collaboration and useful discussions*

# IBM Power4 chip (circa 2001)

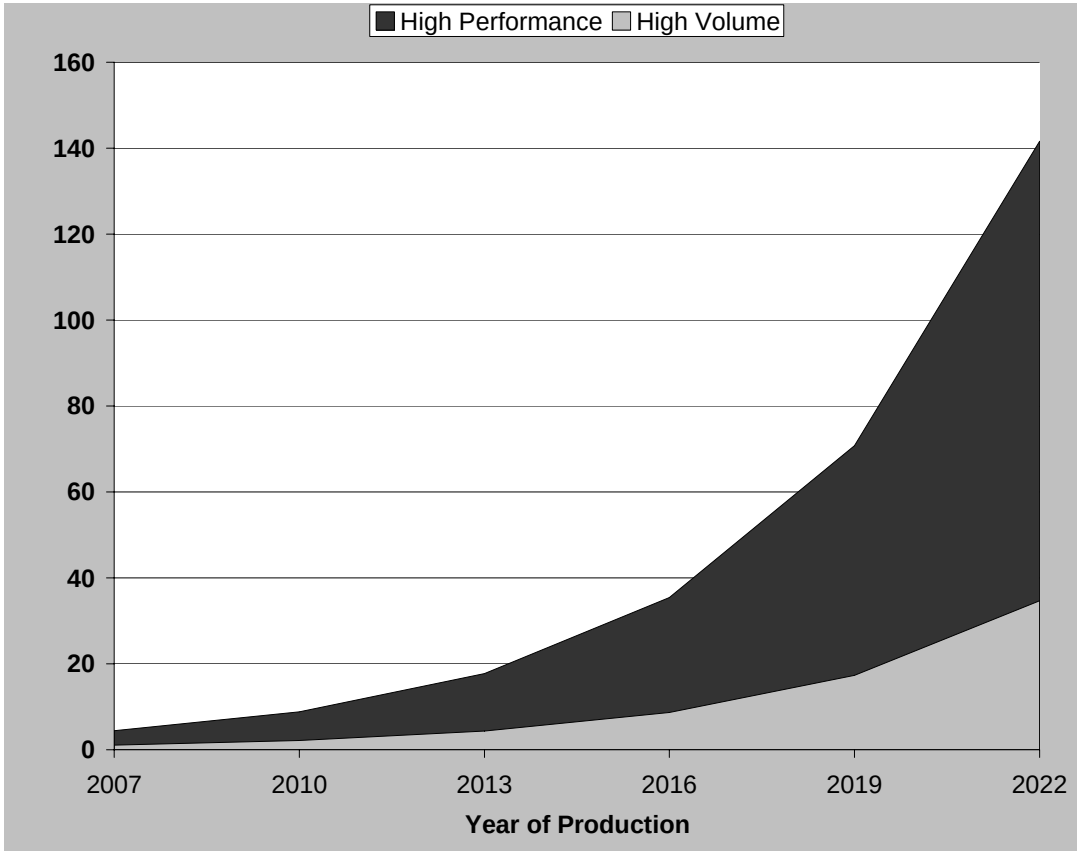


# Projection in 2001 for a 2011 chip

(Based on the SIA Roadmap of 2001)

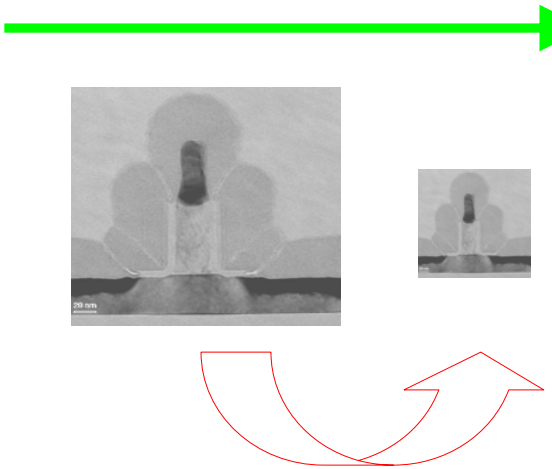


# ITRS 2007 Roadmap



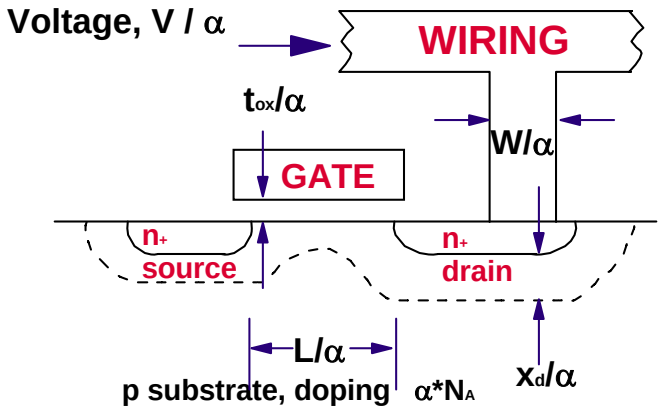
# CMOS Scaling: Dennard's Theory

**Scaled technology generations**



Smaller
Faster
Lower
Power

## Scaled Device




*Dennard, 84*

**SCALING:**

| | |
|---|---|
| Voltage: | V/α |
| Oxide: | tox / α |
| Wire width: | W/α |
| Gate width: | L/α |
| Diffusion: | xd /α |
| Substrate: | α * NA |

**RESULTS:**

| | |
|---|---|
| Higher Density: | ~α2 |
| Higher Speed: | ~α |
| Power/ckt: | ~1/α2 |
| Power Density: | ~Constant |

*Chart - courtesy G. Shahidi*

# Utilization

- Maximal utilization of transistors not as important
- Scarce resources are
  - Power
  - Bandwidth
- Use real estate for
  - Different forms of accelerators
  - Different types of cores
- Turn on only those accelerators or cores that are needed
  - Within chip power budget

# Performance Unpredictability

- Small geometry leads to
  - Process variability
    - Hence performance variability
  - Greater unreliability
    - Need to tolerate failures

# Approximate Computing

- In many applications, approximate results can be tolerated
  - Not of course calculations involving your bank account
- In a sense, Google search results are approximate
- VIA 2020 applications can tolerate imprecision
  - User interfaces
  - Simulation of physical systems
- Do not skirt around unreliability
- Instead, tolerate unreliability through new architectural and software models

# Storage-Class Memory

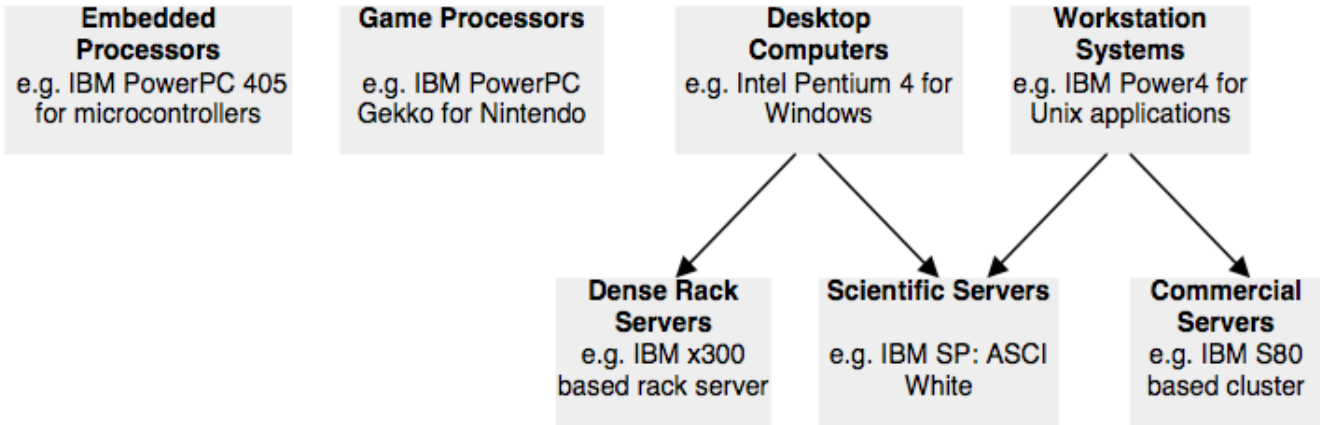
- Exciting developments
- SCM Characteristics
  - Non-volatile
  - Denser than DRAM
  - DRAM-like access times
  - $/bit will approach disk $/bit eventually
  - More power-efficient than DRAM and disk

# Packaging Technology

- Silicon carrier
- 3-D stacking

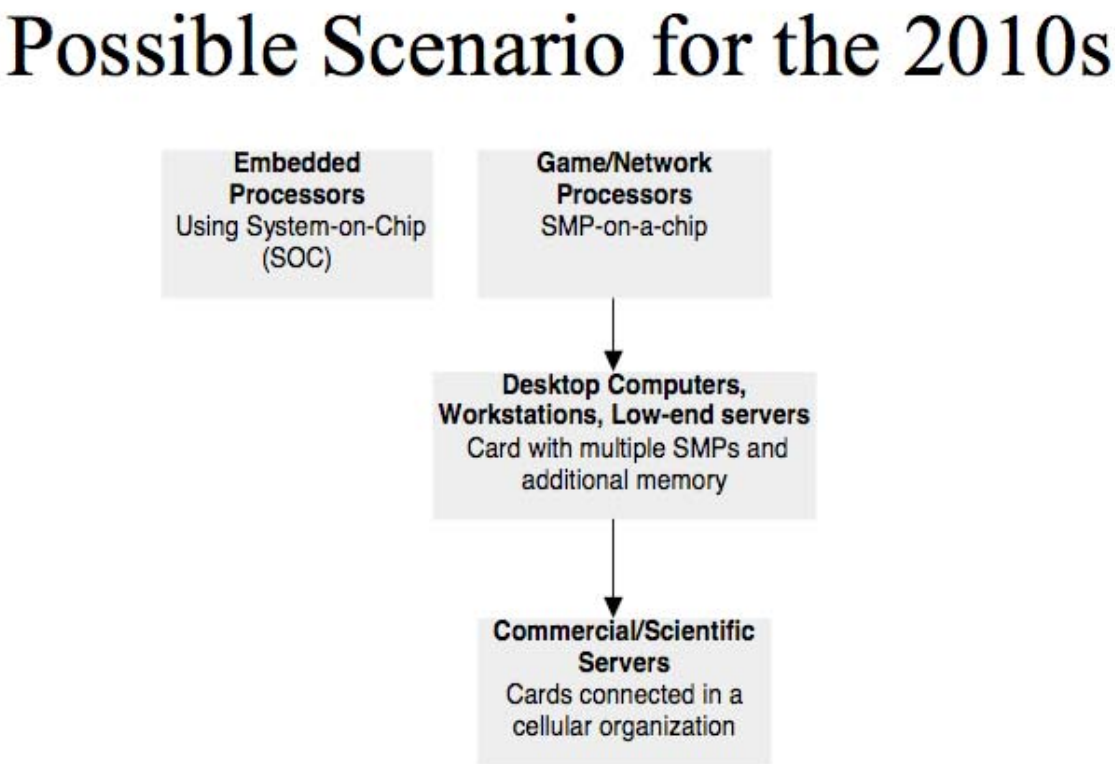
# Prediction in 2001

## Growing Reuse of Microprocessor Cores




*With increasing cost of designing, implementing, and testing microprocessors, there will be a swifter move towards common cores*

# Prediction in 2001



# IBM Roadrunner for Los Alamos National Lab

- 1 Petaflops
- 6562 Dual-core AMD Opteron chips
- 12240 Cell chips (used in Sony Playstation 3)
- 98 Terabytes of memory
- 278 refrigerator-sized racks
- 2.35 MW of power