IC Manufacturing, Cost, Power, and Dependability

COE 501

Computer Architecture

Prof. Muhamed Mudawar

Computer Engineering Department

King Fahd University of Petroleum and Minerals

From Sand to Silicon

Source: Intel corporation

❖ Sand (beach) has a high percentage of Silicon in the form of Silicon dioxide (SiO_2)

❖ Melted Silicon is purified in many steps to reach semiconductor manufacturing quality

❖ Silicon Ingot: diameter = ~30 cm, weight = ~100 Kg, Silicon purity = 99.9999999% (one alien atom per billion silicon atoms)

Ingot Slicing ➜ Blank Wafers

Source: Intel corporation
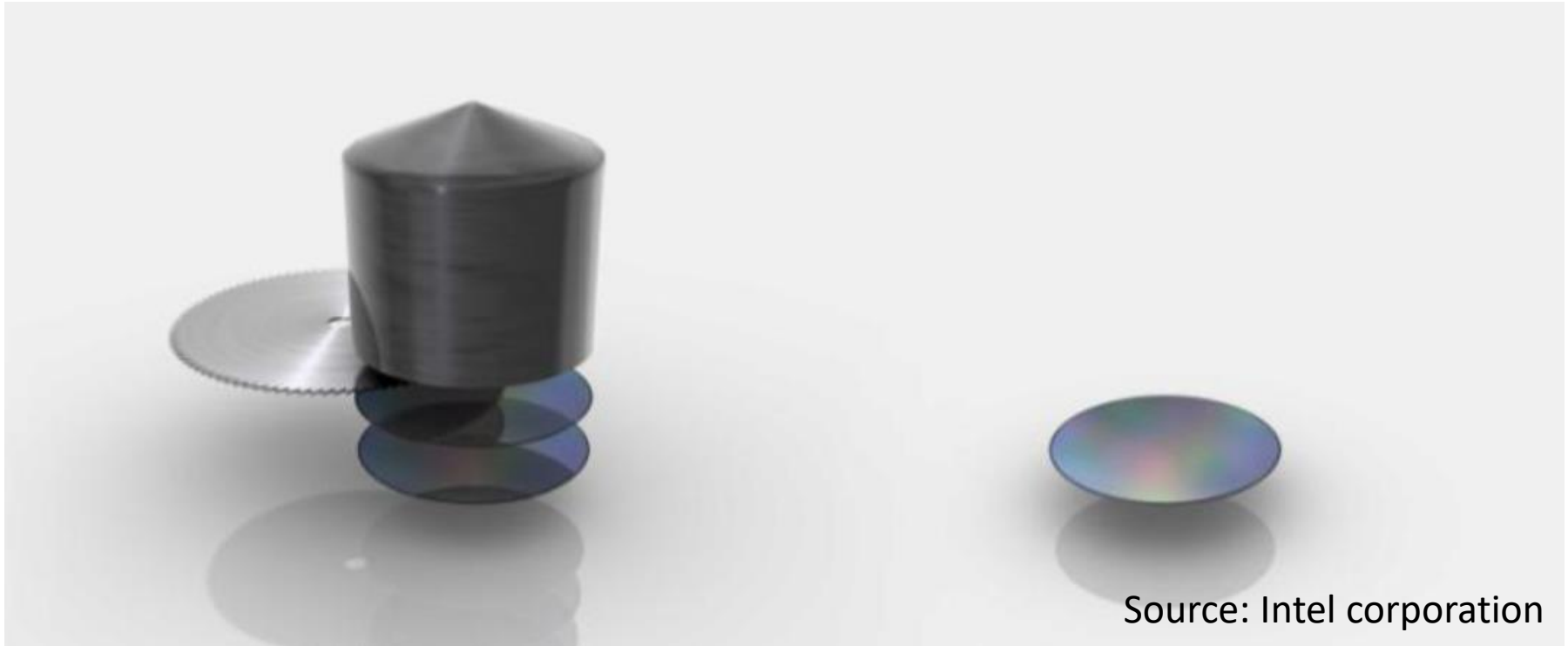
❖ The ingot is cut into individual silicon disks, called wafers

❖ The thickness of a wafer is about 1 mm

❖ The wafers are polished until they have mirror-smooth surfaces

❖ Wafer diameter = 300mm, but future ones can be 450mm

Fabrication of Chips on a Wafer

❖ Hundreds of precisely controlled steps

❖ **Photo Lithography:** uses photo-resist and ultraviolet (UV) light to expose a pattern from a photomask to the surface of a wafer

❖ **Ion Implantation:** the wafer with patterned photoresist is bombarded with a beam of ions that are embedded in selected regions of the wafer. This process is called doping.

❖ **Dielectric Deposition:** using an insulator to reduce leakage

❖ **Etching:** unneeded material is removed to create patterns

❖ **Metal Layers:** multiple metal layers are made to interconnect the various transistors and components

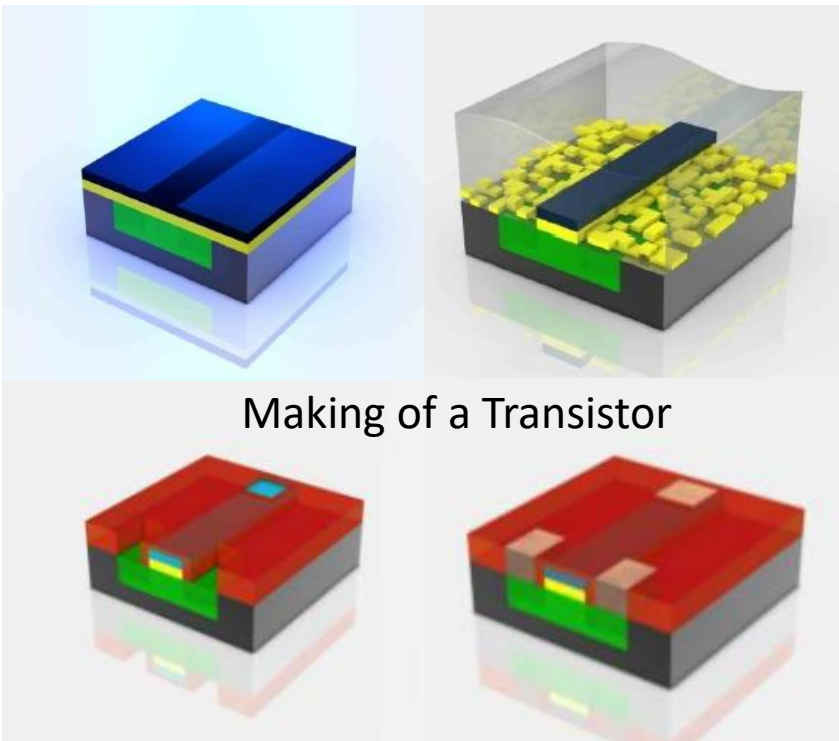# Fabrication of Transistors and Metal Layers

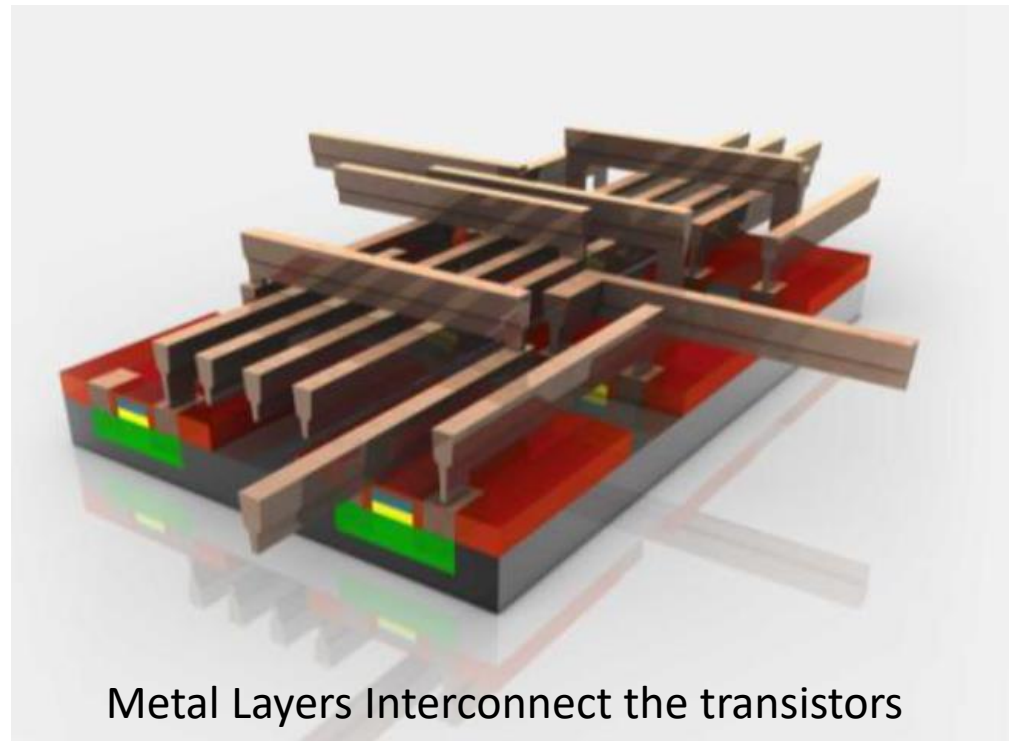

Photo Lithography

Ion Implantation
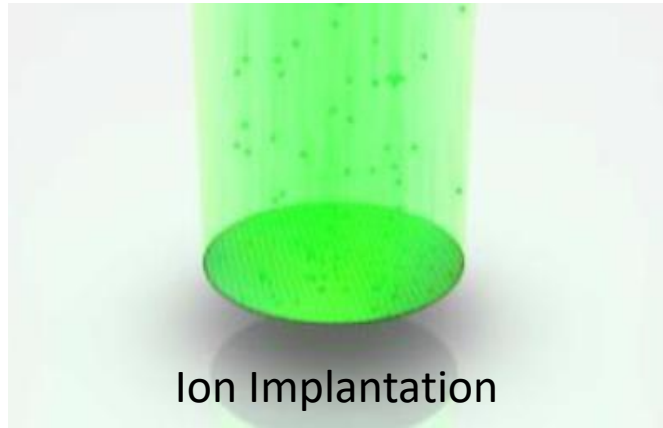
Dielectric Deposition

Making of a Transistor

Metal Layers Interconnect the transistors

Wafer of Intel 8th Generation Core i7 Dies

❖ Wafer diameter = 30cm (300mm)

❖ Wafer has 393 dies

❖ Area of a single die ≈ 149.6 mm^2

   ✧ ~9.19 mm by ~16.28 mm

❖ Six cores + GPU on a single die

   ✧ About 3.6 billion transistors per die

❖ 14 nm manufacturing process

   ✧ 11 metal layers

❖ Incomplete dies on the boundary are useless and discarded. They use the same masks used to pattern the wafer.
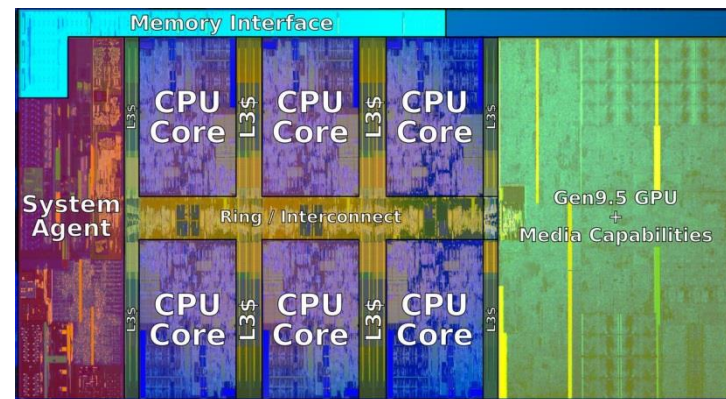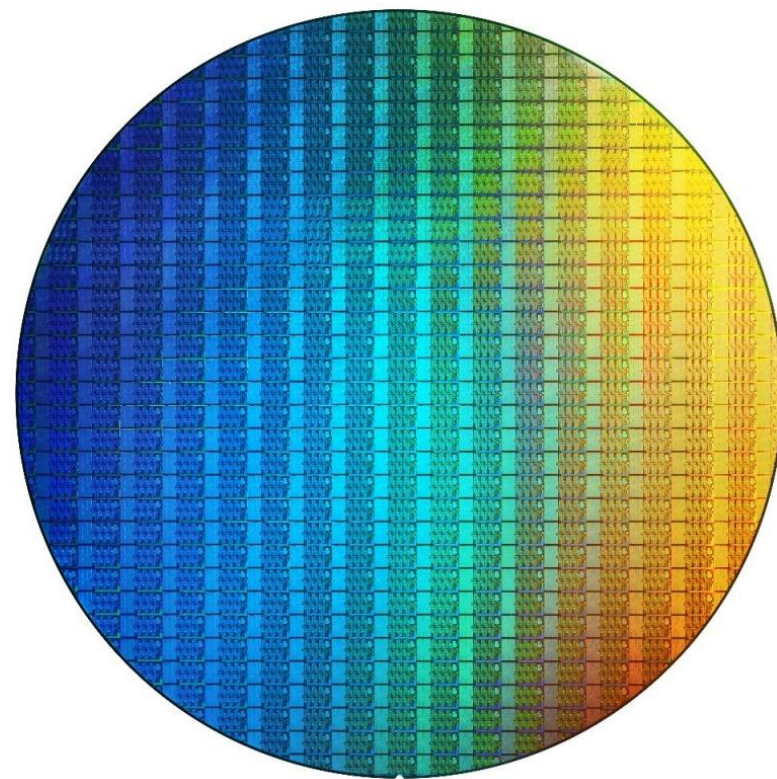
Testing the Dies on the Wafer

Source: Intel corporation
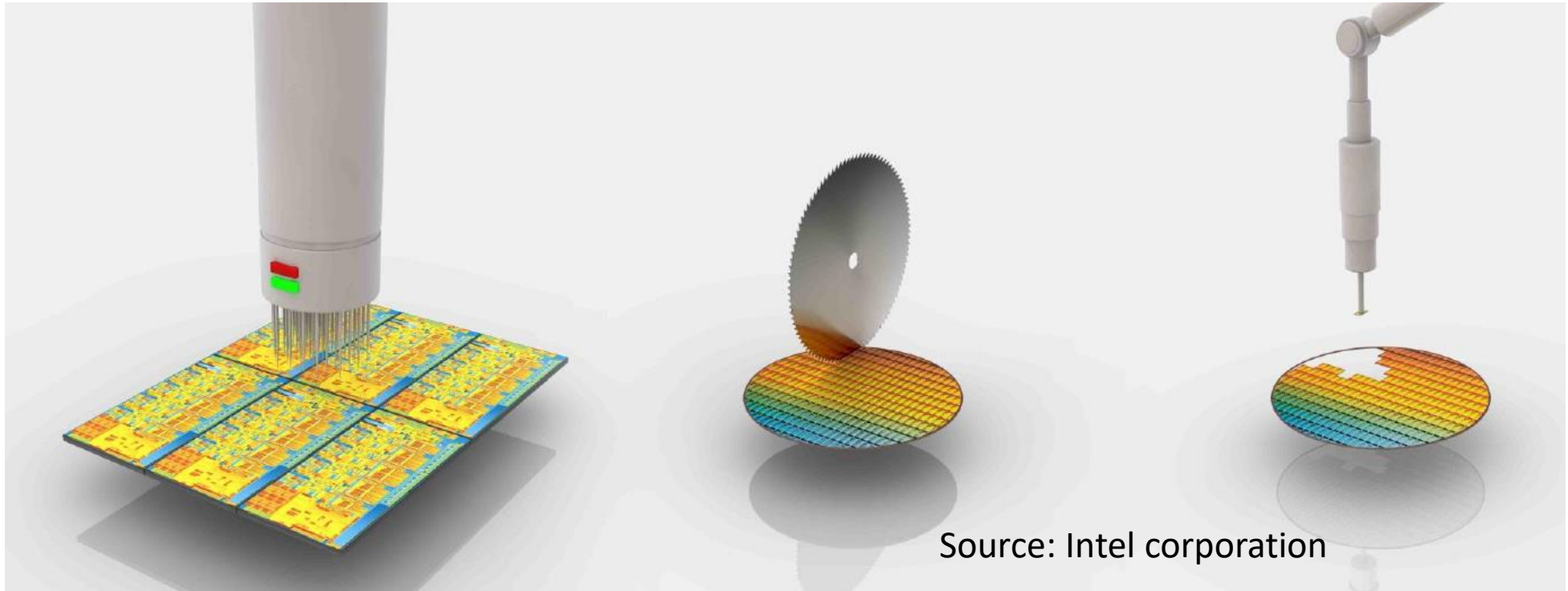
❖ Test patterns are fed into the inputs of each chip on the wafer, and the outputs are compared with the correct values

❖ The wafer is cut into pieces called dies

❖ The dies that passes the test are kept, and the rest is discarded

Bonding Die to Package and Final Test

❖ The substrate, the die, and the heat-spreader are put together to form a complete processor

❖ The green substrate is the interface between the processor and the rest of the PC system (up to several thousand pins)

❖ Finally, the processor are tested for their key characteristics, such as power dissipation and maximum frequency

Die Yield

$$Die\ yield = \frac{Number\ of\ good\ dies\ after\ die\ testing}{Total\ number\ of\ dies\ on\ a\ wafer}$$

Bose-Einstein Formula:

$$Die\ yield = \frac{1}{(1 + Defects\ per\ unit\ area \times Die\ area)^N}$$

❖ Empirical formula by looking at the yield of IC manufacturers

❖ Defects per unit area ≈ 0.01 to 0.05 defects per cm^2

  ⬦ Defects are due to microscopic particles landing on the wafer surface

  ⬦ A tiny particle with the size of the smallest feature can "kill" the die

❖ *N* is a parameter that measures the manufacturing complexity

Effect of the Die Size on the Yield
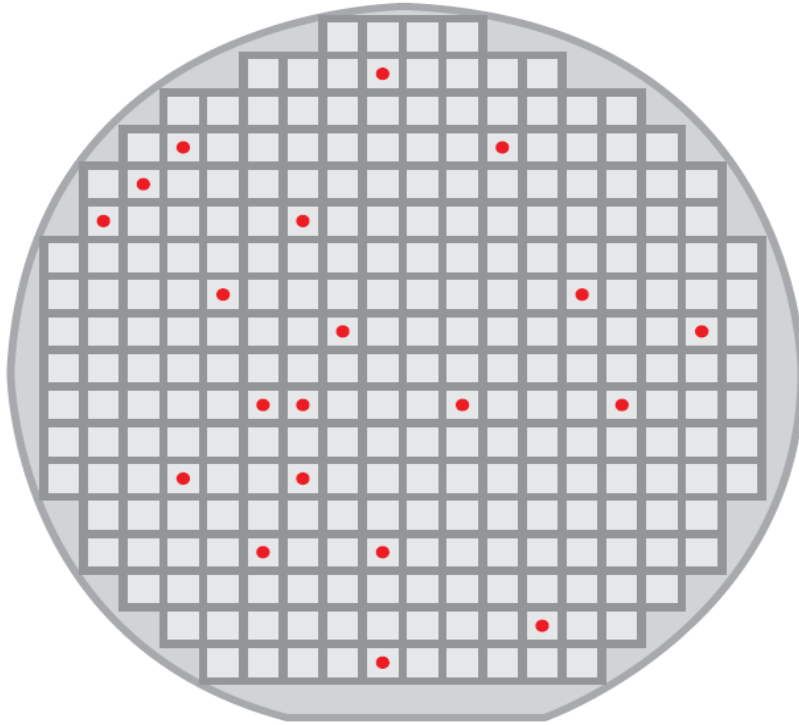
20 Defects on a wafer
20 Bad dies
264 dies on a wafer
Good dies = 264 – 20 = 244
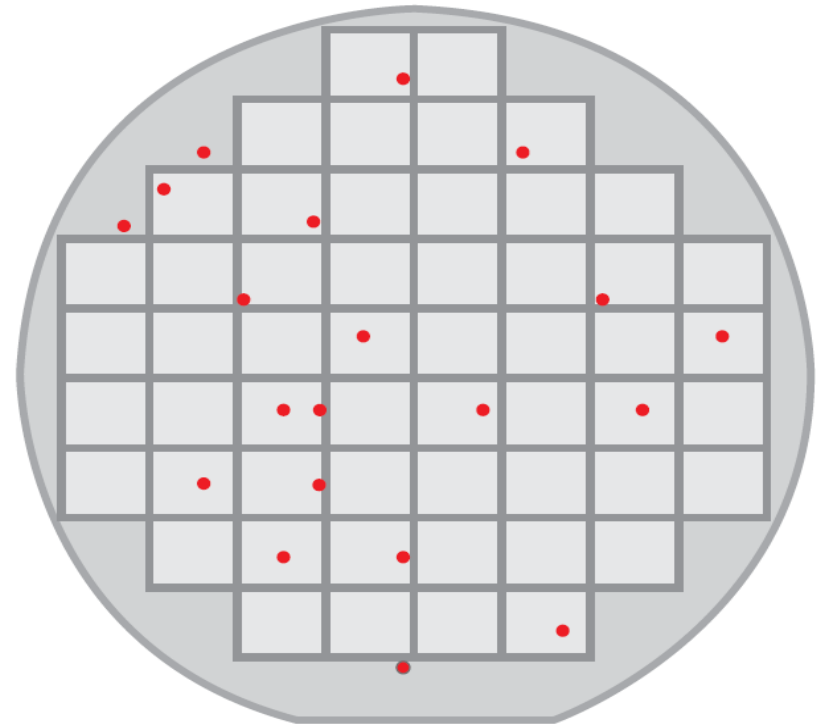**Yield = 244 / 264 = 92.42%**

20 Defects on a wafer
16 Bad dies
54 large dies on a wafer
Good dies = 54 – 16 = 38
Yield = 38 / 54 = 70.37%

Example on the Die Yield

❖ Find the die yield for 60 mm^2, 120 mm^2, and 240 mm^2 dies

The defect density is 0.023 defects per cm^2 and N is 12

Solution

❖ For the small die area = 60 mm^2 = 0.6 cm^2

Die Yield = $1/(1 + 0.023 \times 0.6)^{12}$ = 0.8483 = **84.83%**

❖ For the medium die area = 120 mm^2 = 1.2 cm^2

Die Yield = $1/(1 + 0.023 \times 1.2)^{12}$ = 0.7213 = **72.13%**

❖ For the larger die area = 240 mm^2 = 2.4 cm^2

Die Yield = $1/(1 + 0.023 \times 2.4)^{12}$ = 0.5248 = **52.48%**

❖ The die **yield drops** as the die **area increases**

Cost of an Integrated Circuit

$$Cost\ of\ die = \frac{Cost\ of\ wafer}{Dies\ per\ wafer \times Die\ yield}$$

$$Dies\ per\ wafer = \frac{Wafer\ area}{Die\ area} - Incomplete\ dies\ on\ boundary$$

$$Dies\ per\ wafer \approx \frac{\pi \times (wafer\ radius)^2}{Die\ area} - \frac{\pi \times Wafer\ diameter}{\sqrt{2 \times Die\ area}}$$

$$Cost\ of\ IC = \frac{Cost\ of\ die + testing\ die + package + final\ test}{Final\ test\ yield}$$

Counting the Dies on a Wafer

❖ Find the number of dies on a wafer with a diameter = 30cm

When the die area = 60 mm^2,120 mm^2, and 240 mm^2

Answer:

❖ For the small die area = 60 mm^2 = 0.6 cm^2

Dies per wafer $\cong$ (π × 15^2)/0.6 $-$ (π × 30)/$\sqrt{}$(2×0.6) $\cong$ **1092**

❖ For the medium die area = 120 mm^2 = 1.2 cm^2

Dies per wafer $\cong$ (π × 15^2)/1.2 $-$ (π × 30)/$\sqrt{}$(2×1.2) $\cong$ **528**

❖ For the large die area = 240 mm^2 = 2.4 cm^2

Dies per wafer $\cong$ (π × 15^2)/2.4 $-$ (π × 30)/$\sqrt{}$(2×2.4) $\cong$ **251**

❖ If the die area is **doubled**, the total number of dies on a wafer is **less than half** because we loose dies on the boundary

Cost of a Die

❖ Given three different die areas = 60mm^2, 120mm^2, and 240mm^2, the total count of dies on a 30cm-diameter wafer is calculated as 1092, 528, and 251, respectively. The die yield is also calculated as: 84.83%, 72.13%, and 52.48%, respectively.

If the cost of a wafer is $7000, calculate the cost per die

Answer (does not include cost of testing and packaging)

❖ For 60mm^2, number of good dies = 1092 × 0.8483 ≅ **962**

Cost of a 60mm^2 die = $7000 / 962 = **$7.28**

❖ For 120mm^2, number of good dies = 528 × 0.7213 ≅ **381** (rounded)

Cost of 120mm^2 die = $7000 / 381 = **$18.37** (increased by **2.52X**)

❖ For 240mm^2, number of good dies = 251 × 0.5248 ≅ **132** (rounded)

Cost of 240mm^2 die = $7000 / 132 = **$53.03** (increased by **7.28X**)

Things to Remember about Cost

❖ Bottom line is the number of good dies per wafer

❖ Good dies per wafer = Number of dies × Die yield

❖ The manufacturing process dictates the wafer cost, wafer yield, and the defects per unit area

❖ Sole control of the designer is die area and hence the cost

❖ Die should be tested, packaged, then tested again

  ◇ These steps add more costs, which can be significant

❖ Most microprocessor dies fall between 100 and 300 mm^2

❖ Low-end embedded processors are below 10 mm^2

❖ Designers also add redundancy to raise the yield

Power in Integrated Circuits

❖ Power is the biggest challenge facing computer design

  ✧ Power should be brought in and distributed around the chip

  ✧ Hundreds of pins and multiple layers just for power and ground

  ✧ Power is dissipated as heat and must be removed

❖ Thermal Design Power (TDP)

  ✧ Characterizes sustained power consumption

  ✧ Used as target for power supply and cooling system

  ✧ Lower than peak (1.5X higher), higher than average power consumption

❖ Clock rate can be reduced dynamically to limit power

Power versus Energy

❖ Which metric is the right one: Power or Energy?

❖ Power is Energy per Unit Time: 1 Watt = 1 Joule / Second

❖ Energy for a given task is a better measurement (Joules)

❖ Energy for a workload = Average Power × Execution Time

❖ **Energy efficiency** is important for battery-operated devices and for large servers or cloud

❖ Example: which processor is more energy efficient?

    ◇ Processor A consumes 20% more power than B on a given task

    ◇ However, A requires only 70% of the execution time needed by B

❖ **Answer:** Energy consumption of A = 1.2 × 0.7 = 0.84 of B

    ◇ Processor A consumes less energy than B (more energy-efficient)

Dynamic Energy and Power

❖ For CMOS technology, primary energy consumption has been in switching transistors, called **dynamic energy**

❖ **Dynamic Energy ∝ Capacitive Load × Voltage^2**

   ♦ Capacitive Load = Capacitance of output transistors & wires

   ♦ Voltage has dropped from 5V to just under 1V in 20 years

❖ **Dynamic Power ∝ Capacitive Load × Voltage^2 × Frequency**

❖ Reducing Clock Frequency reduces Power but not energy

❖ Reducing Clock Frequency increases execution time

Example of Dynamic Energy & Power

❖ Microprocessors today have adjustable voltage and clock frequency. Assume 10% reduction in voltage and 15% reduction in frequency, what is the impact on dynamic energy and dynamic power?

❖ **Answer:**

10% reduction in Voltage ➔ Voltage$_{new}$ = 0.90 × Voltage$_{old}$

15% reduction in Frequency ➔ Frequency$_{new}$ = 0.85 × Frequency$_{old}$

$$\frac{Energy_{new}}{Energy_{old}} = \frac{Voltage_{new}^2}{Voltage_{old}^2} = (0.90)^2 = 0.81$$

$$\frac{Power_{new}}{Power_{old}} = \frac{Voltage_{new}^2}{Voltage_{old}^2} \times \frac{Frequency_{new}}{Frequency_{old}} = 0.81 \times 0.85 = 0.6885$$

Trends in Power & Clock Frequency
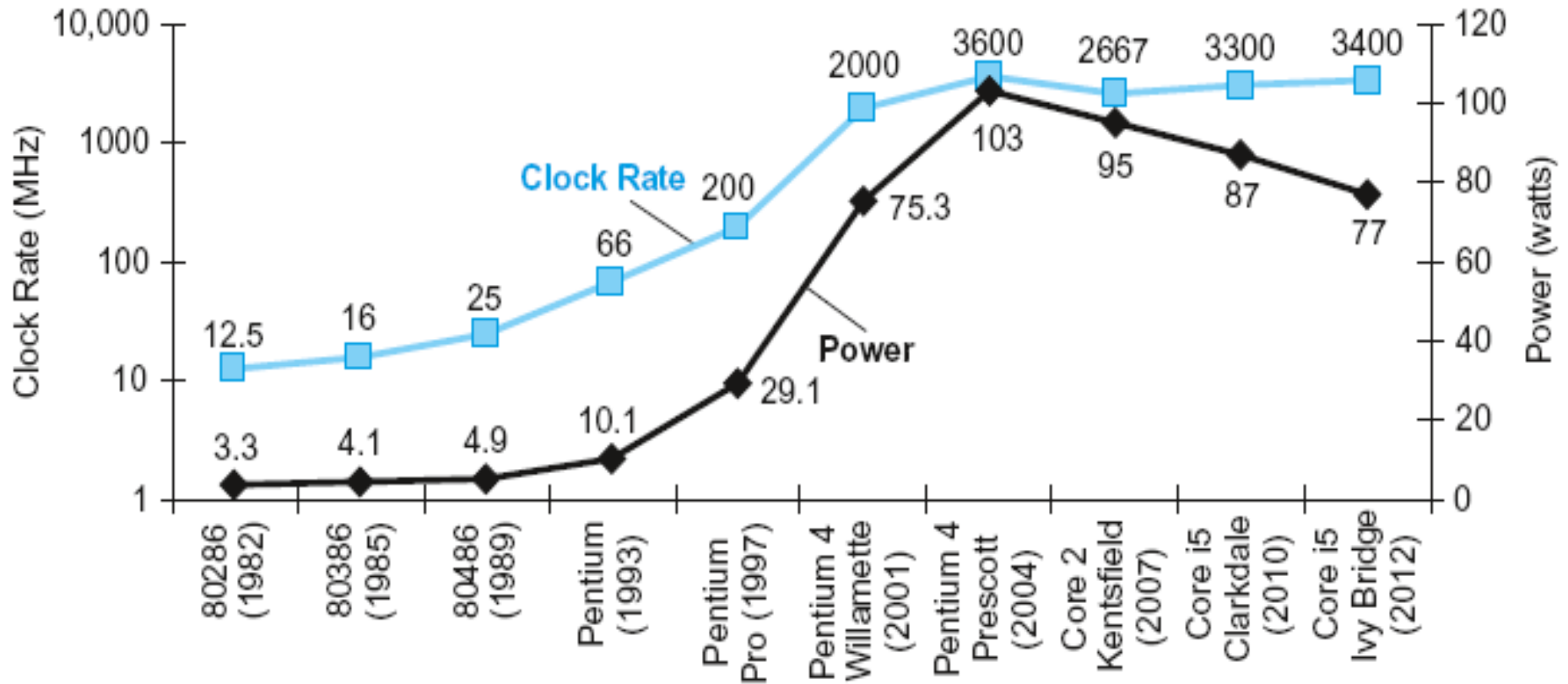
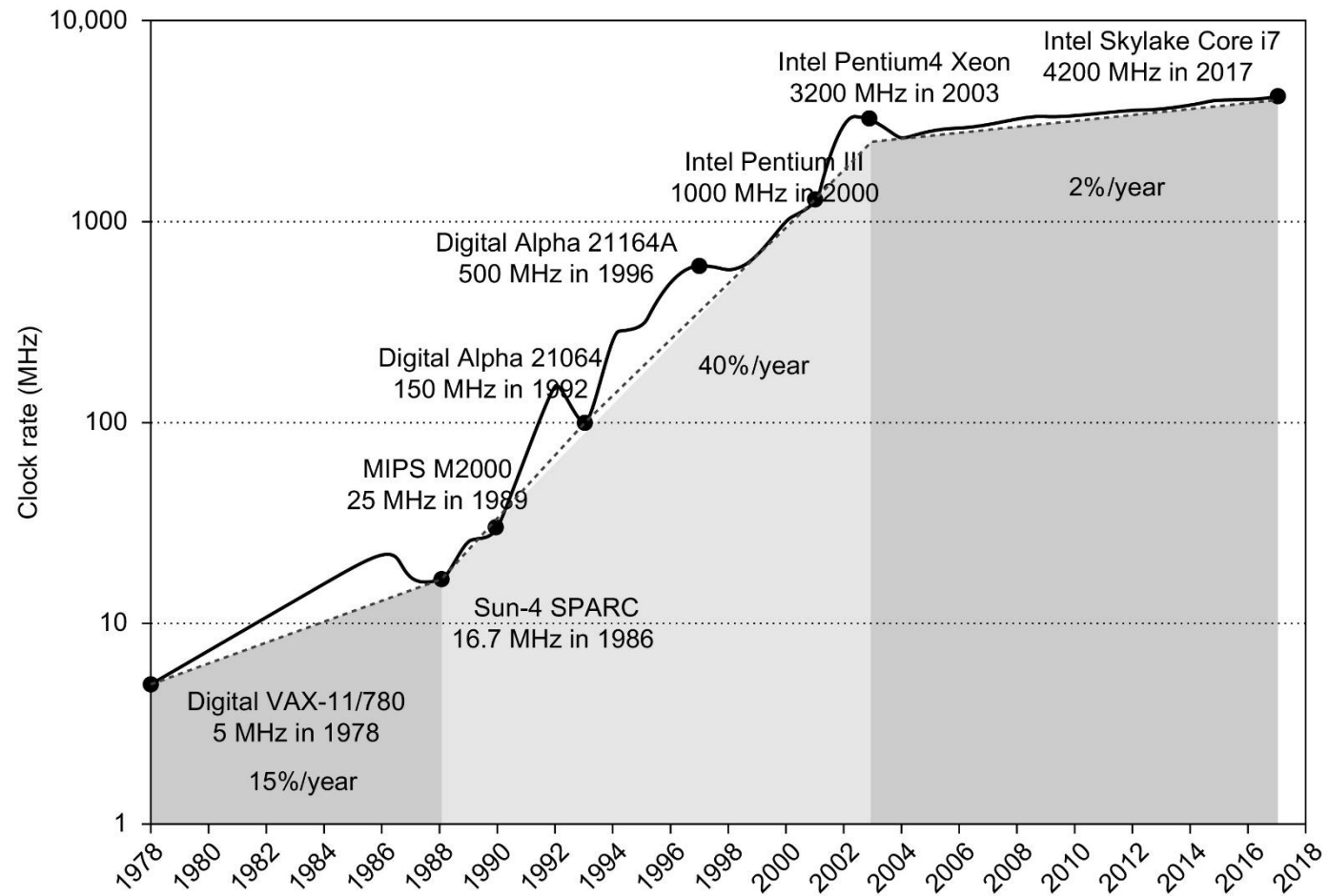Power ∝ Capacitive Load × Voltage^2 × Frequency

| 24X | 2X | 5V→1V | 300X |

Clock Frequency is Slowing Down

❖ Intel 80386 consumed 4 W

❖ Some high-end server processors consume 130 W

❖ Heat must be dissipated from ~200 mm^2 chip

❖ This is the limit of what can be cooled by air

Copyright © 2019, Elsevier Inc. All rights reserved.

Techniques to Reduce Dynamic Power

1. Turn off the clock of inactive modules or cores

2. Dynamic Voltage-Frequency Scaling (DVFS)

   ✧ Periods of low activity → No need to operate at max power

   ✧ Reduce voltage and frequency to reduce power

3. Design for the Typical Case

   ✧ Battery-operated mobile devices are often idle, DRAM memory and storage offer low power modes to save battery energy

4. Overclocking (Turbo Mode)

   ✧ The chip decides that it is safe to run at higher clock rate for a short time until temperature starts to rise (temperature sensor)

   ✧ Run few cores at higher clock rate, while turning off other cores

Static Power Consumption

❖ **Static power** is dissipated because leakage current flows even when a transistor is off

❖ Leakage current increases with smaller transistor sizes

◇ New transistor technology (better insulator) helps reduce leakage

❖ Static Power ∝ Static Current × Voltage

◇ Static power increases with the number of transistors

❖ Static power can be as high as 50% of total power

◇ Large SRAM caches need static power to maintain their values

❖ Power Gating: Turn Off the Power Supply

◇ To inactive modules to control the loss of leakage current

Energy per Operation

❖ 32-bit FP addition uses 9X more energy than 32-bit INT addition (Area is 31X)

❖ 32-bit SRAM read uses 50X more energy than 32-bit INT addition

❖ 32-bit DRAM read uses 6400X more energy than 32-bit INT addition

Operation:	Energy (pJ)	Relative energy cost	Area (μm^2)	Relative area cost
8b Add	0.03		36	
16b Add	0.05		67	
32b Add	0.1		137	
16b FB Add	0.4		1360	
32b FB Add	0.9		4184	
8b Mult	0.2		282	
32b Mult	3.1		3495	
16b FB Mult	1.1		1640	
32b FB Mult	3.7		7700	
32b SRAM Read (8KB)	5		N/A	
32b DRAM Read	640		N/A	

Energy scale: 1 10 100 1000 10000
Area scale: 1 10 100 1000
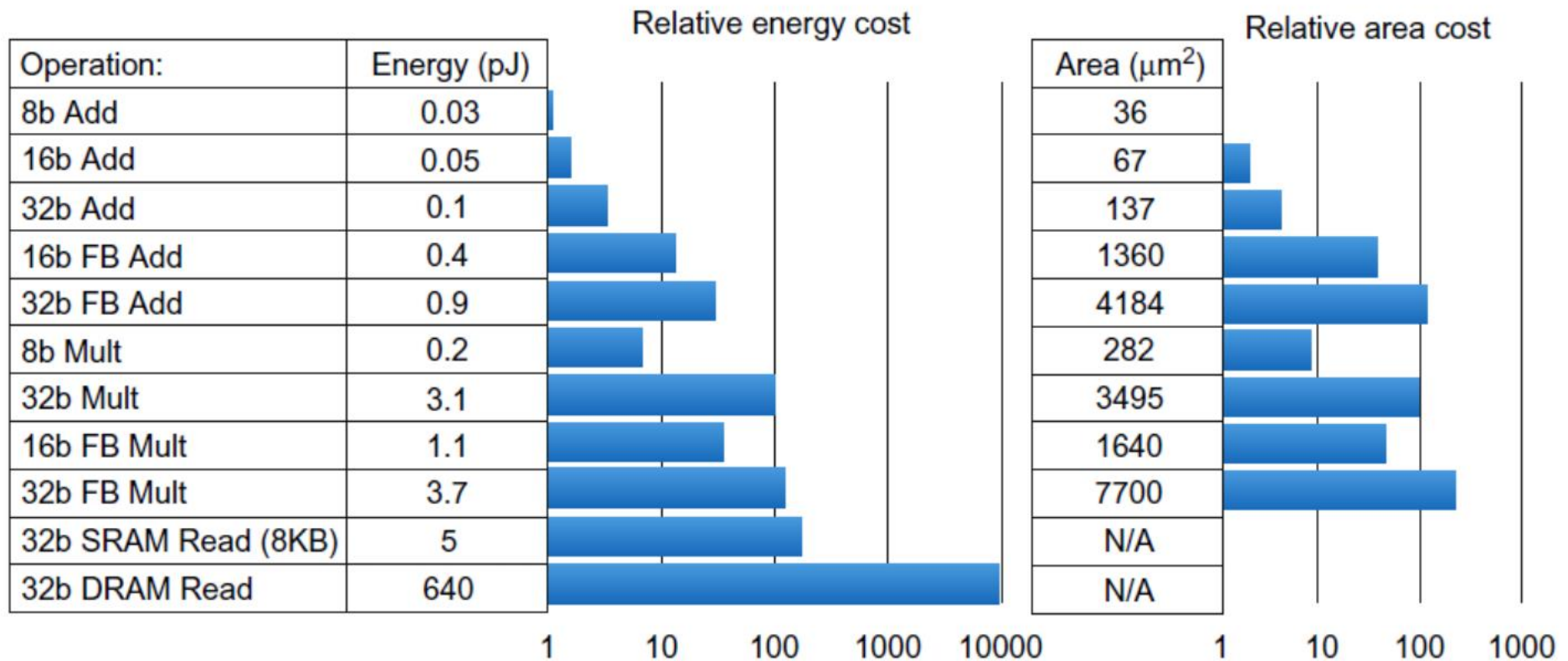
Mark Horowitz, "Computing's Energy Problem and what we can do about it", ISSCC 2014.
Area numbers are from synthesized result using Design Compiler under TSMC 45nm technode.

Relative Performance

❖ Absolute performance is about maximizing the performance of a given application or workload, regardless of the cost of the system or the energy consumed

❖ Relative performance can be cost-aware or energy-aware

❖ Cost-aware performance = Performance per Cost of a system

    ◇ Designing systems that deliver higher performance per dollar

❖ Energy-aware performance = Tasks executed per Joule

    ◇ Designing systems that executes more tasks per energy consumed

❖ Performance per Watt = Tasks per Joule

    ◇ Performance / Watt = (Tasks / sec) / (Joule / sec) = Tasks / Joule

    ◇ Designing systems that deliver higher performance per power utilized

Example on Relative Performance

❖ Suppose we run a database application on three different servers: A, B, and C. Performance is measured as the number of database transactions per second. The data is recorded as follows:

Transactions Per Sec (TPS): A = 910,978, B = 976,812, and C=1,840,450

System Cost: A = $9352, B = $9576, and C = $21,658

Power Utilization: A = 570 W, B = 650 W, and C = 1090 W

❖ Which server has the highest absolute performance, performance per cost, and performance per watt?

❖ **Answer:**

Server C has the highest absolute performance = 1,850,450 TPS

Server B has the highest performance per cost = 976,812/$9576 = 102 TPS/$

Server C has the highest performance per watt = 1,840,450/1090 = 1688 TPS/W

Things to Remember about Energy & Power

❖ The bottom line is Energy for a given Workload (Joules)

  ♢ The less the Energy used per task, the higher is the energy efficiency

  ♢ For both battery-operated mobile devices and large servers

❖ Power should be used as a constraint

  ♢ A processor for a mobile device might be limited to 15 watts

  ♢ A server processor might be limited to 120 watts (cooling + power supply)

❖ Power Consumed = Dynamic Power + Static Power

  ♢ Dynamic power is consumed when transistors switch on and off

  ♢ Static power is consumed because of static leakage current

❖ Clock rate can be reduced dynamically to limit power

❖ Reducing voltage reduces both energy and power

Dependability

❖ Dependability is a measure of system's reliability, availability, and maintainability

❖ **Reliability:** continuity of correct service

❖ **Availability:** readiness for correct service

❖ **Maintainability:** ability to undergo modifications and repair

❖ Systems alternate between two states of service:

    1. Service accomplishment: service is delivered as specified

    2. Service interruption: failure in service

❖ **Failure:** transition from state 1 to state 2

❖ **Restoration (or Repair):** transition from state 2 back to state 1

Measuring Dependability

❖ Module Reliability: measure of continuous service

❖ **MTTF:** Mean Time To Failure

  ◈ Measures Reliability

  ◈ Continuous service accomplishment from a reference time

  ◈ Reported in hours

❖ **FIT:** Failures In Time = 10^9 / MTTF

  ◈ Number of failures per billion hours of operation

❖ **MTTR:** Mean Time To Repair

  ◈ Measure service interruption time

❖ **MTBF:** Mean Time Between Failures = MTTF + MTTR

❖ **Module Availability = MTTF / (MTTF + MTTR)**

Example of Calculating MTTF

❖ A disk subsystem has the following components

   ✧ 10 disks, each rated at MTTF = 1,000,000 hours

   ✧ 1 ATA controller, rated at MTTF = 500,000 hours

   ✧ 1 power supply, rated at MTTF = 200,000 hours

   ✧ 1 fan, rated at MTTF = 200,000 hours

   ✧ 1 ATA cable, rated at MTTF = 1,000,000 hours

❖ Failure of a component is a failure of the disk subsystem

❖ Failures are **independent** and lifetimes of modules are exponentially distributed

❖ **Overall failure rate = sum of failure rates of modules**

❖ Compute the MTTF of the disk subsystem

Calculating MTTF (Solution)

❖ Failure Rate (1disk) = 1 / 1,000,000 = 1000 / 10^9 = 1000 FIT

❖ Failure Rate (10 disks) = 10 × 1 / 1,000,000 = 10,000 / 10^9

❖ Failure Rate (ATA controller) = 1 / 500,000 = 2000 / 10^9

❖ Failure Rate (Power supply) = 1 / 200,000 = 5000 / 10^9

❖ Failure Rate (Fan) = 1 / 200,000 = 5000 / 10^9

❖ Failure Rate (ATA cable) = 1 / 1,000,000 = 1000 / 10^9

❖ Failure Rate (disk system) = Sum of failure rates

   = (10,000 + 2000 + 5000 + 5000 + 1000) / 10^9

❖ Failure Rate (disk system) = 23,000 / 10^9 = 23,000 FIT

❖ MTTF (disk system) = 10^9 / 23,000 $\cong$ 43,478 hours $\cong$ 1812 days, which is just under 5 years

Redundancy Improves Reliability

❖ Example: Using two power supplies for a disk subsystem

◇ One power supply is sufficient to run the disk subsystem, but we are adding a second power supply to tolerate the failure of one power supply

◇ If each power supply is rated at MTTF = 200,000 & MTTR = 50 hours, what is the MTTF of the power supply pair?

❖ **Answer:**

◇ Assuming independent failure of power supplies

◇ Failure rate (1st or 2nd power supply) = 2 / MTTF

◇ Probability of 2nd failure while repairing first = MTTR / MTTF

◇ Failure rate (both power supplies) = (2 / MTTF) × (MTTR / MTTF)

◇ MTTF (power supply pair) = MTTF^2 / (2 × MTTR) = 200,000^2 / (2 × 50)

◇ MTTF (power supply pair) = 400,000,000 hours (2000X more reliable)