CMSC 611

Evaluating Cost

Integrated Circuits: Fueling Innovation

- Chips begins with silicon, found in sand
- Silicon does not conduct electricity well and thus called semiconductor
- A special chemical process can transform tiny areas of silicon to either:
  - Excellent conductors of electricity (like copper)
  - Excellent insulator from electricity (like glass)
  - Areas that can conduct or insulate (a switch)
- A transistor is simply an on/off switch controlled by electricity
- Integrated circuits combines dozens of hundreds of transistors in a chip

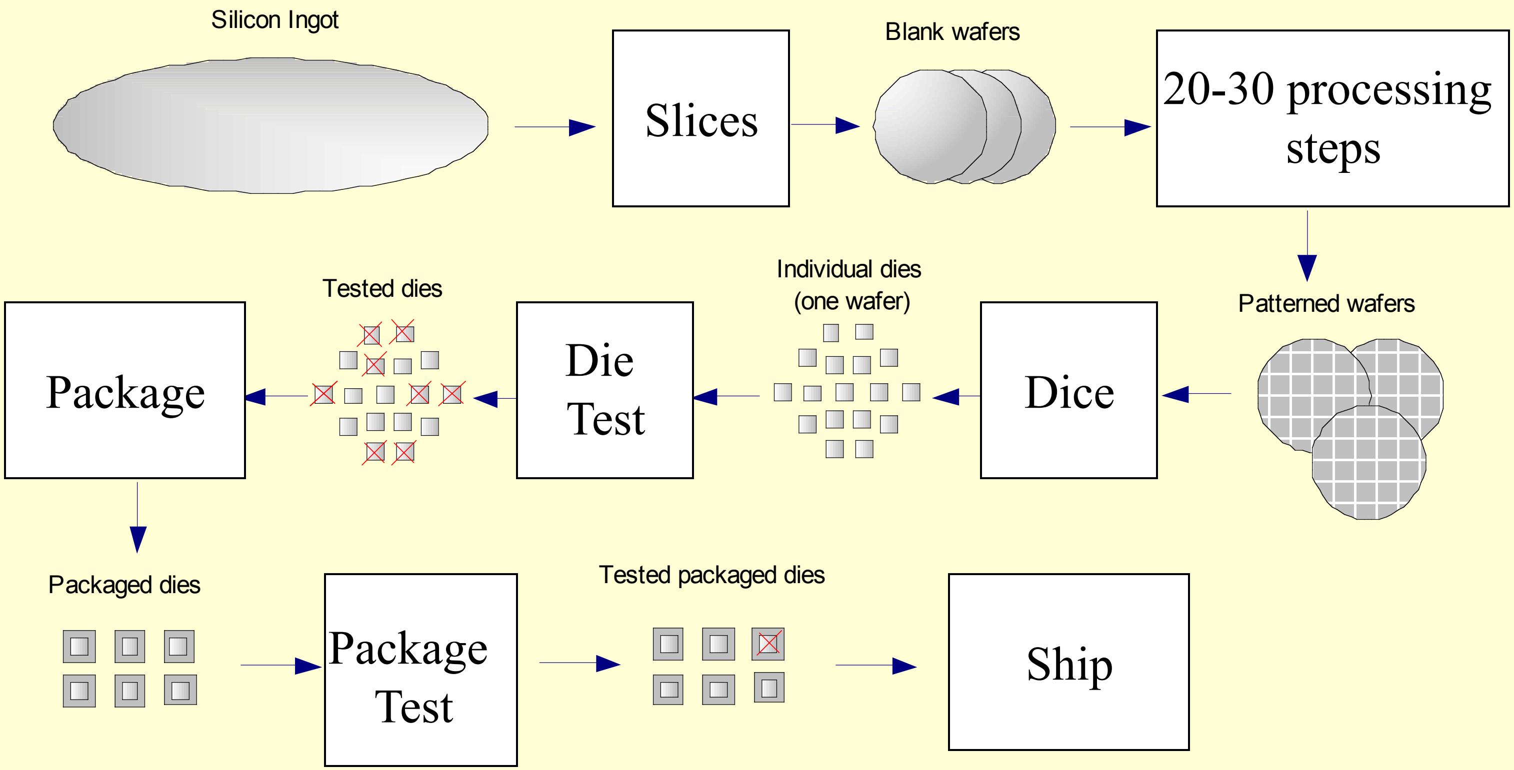
Integrated Circuits: Fueling Innovation

- Technology innovations over time

Year	Technology used in computers	Relative performance/unit cost
1951	Vacuum tube	1
1965	Transistor	35
1975	Integrated circuits	900
1995	Very large-scale integrated circuit	2,400,000

Advances of the IC technology affect H/W and S/W design philosophy

Microelectronics Process

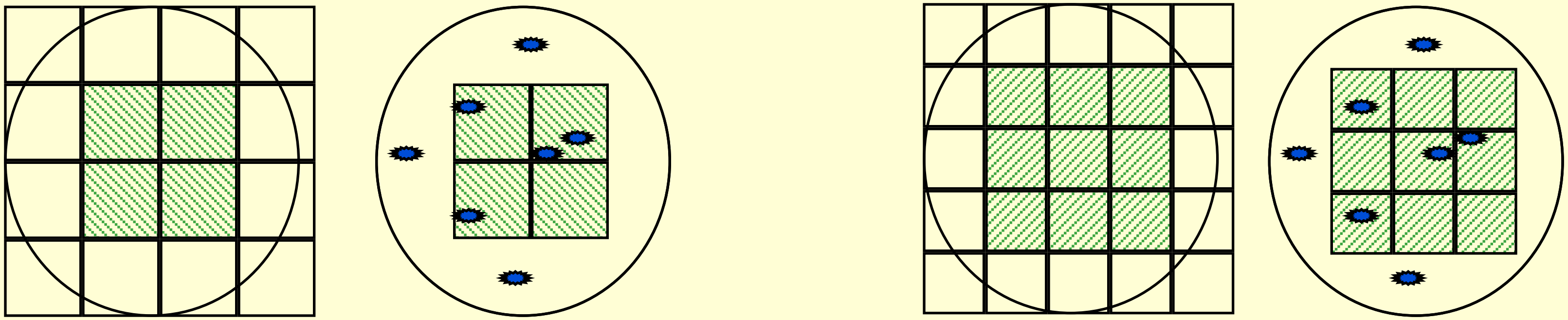


- Silicon ingots:
  - 6-12 inches in diameter and about 12-24 inches long
- Impurities in the wafer can lead to defective devices and reduces the yield

Integrated Circuits Costs

$$\text{Dies\_per\_Wafer} = \frac{\pi \times (\text{Wafer\_diameter}/2)^2}{\text{Die\_Area}} - \frac{\pi \times \text{Wafer\_Diameter}}{\sqrt{2 \times \text{Die\_Area}}}$$

$$\text{Die\_Yield} = \text{Wafer\_Yield} \times \left[1 + \frac{\text{Defects\_per\_Unit\_Area} * \text{Die\_Area}}{\alpha}\right]^{-\alpha}$$



$$\text{Die\_Cost} = \frac{\text{Wafer\_Cost}}{\text{Dies\_per\_Wafer} \times \text{Die\_Yield}}$$

Die cost roughly goes with die area^4

$$\text{IC\_Cost} = \frac{\text{Die\_Cost} + \text{Testing\_Cost} + \text{Packing\_Cost}}{\text{Final\_Test\_Yield}}$$

What Affects Cost?

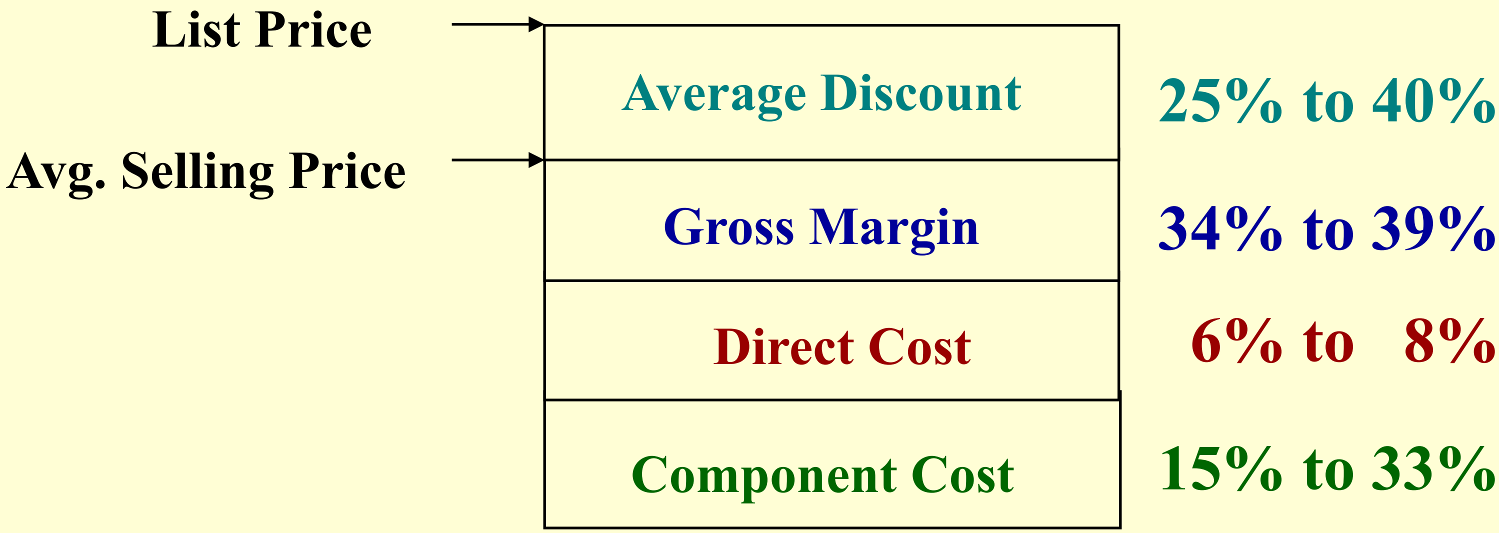
1. Learning curve:
   - The more experience in manufacturing a component, the better the yield
   - In general, a chip, board or system with twice the yield will have half the cost.
   - The learning curve is different for different components, complicating design decisions
2. Volume
   - Larger volume increases rate of learning curve
   - Doubling the volume typically reduce cost by 10%
3. Commodities
   - Are essentially identical products sold by multiple vendors in large volumes
   - Foil the competition and drive the efficiency higher and thus the cost down

Real World Examples

Chip	Layers	Wafer cost	Defect/cm^2	Area (mm^2)	Dies/Wafer	Yield	Die Cost
386DX	2	$900	1.0	43	360	71%	$4
486DX2	3	$1200	1.0	81	181	54%	$12
PowerPC 601	4	$1700	1.3	121	115	28%	$53
HP PA 7100	3	$1300	1.0	196	66	27%	$73
DEC Alpha	3	$1500	1.2	234	53	19%	$149
SuperSPARC	3	$1700	1.6	256	48	13%	$272
Pentium	3	$1500	1.5	296	40	9%	$417

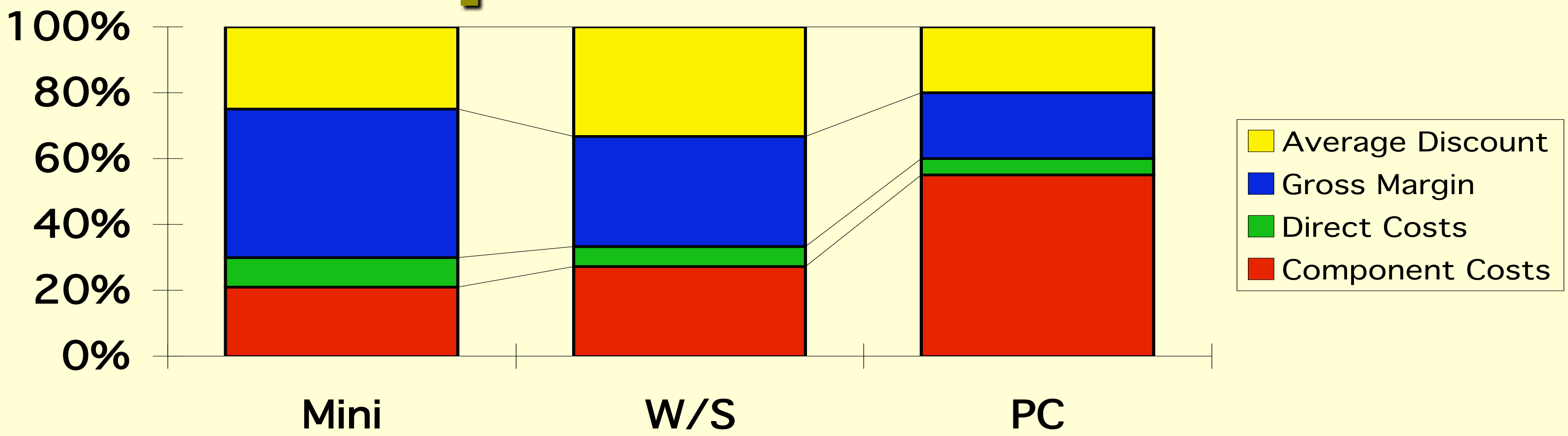
From "Estimating IC Manufacturing Costs," by Linley Gwennap,
***Microprocessor Report*, August 2, 1993, p. 15**

Cost vs. Price



- Component Costs: raw material cost for the system's building blocks
- Direct Costs (add 25% to 40%) recurring costs: labor, purchasing, scrap, warranty
- Gross Margin (add 82% to 186%) nonrecurring costs: R&D, marketing, sales, equipment maintenance, rental, financing cost, pretax profits, taxes
- Average Discount to get List Price (add 33% to 66%): volume discounts and/or retailer markup

Example: Price vs. Cost



Chip Prices (August 1993) for a volume of 10,000 units

Chip	Area (mm^2)	Total Cost	Price	Comment
386DX	43	$9	$31	
486DX2	81	$35	$245	No Competition
PowerPC 601	121	$77	$280	
DEC Alpha	234	$202	$1231	Recoup R&D?
Pentium	296	$473	$965	

Defining Performance

- Performance means different things to different people, therefore its assessment is subtle

Analogy from the airlines industry:

- How to measure performance for an airplane?
  - Cruising speed (How fast it gets to the destination)
  - Flight range (How far it can reach)
  - Passenger capacity (How many passengers it can carry)

Airplane	Passenger capacity	Cruising range (miles)	Cruising speed (m.p.h)	Passenger throughput (Passenger × m.p.h)
Boeing 777	375	4630	610	228,750
Boeing 747	470	4150	610	286,700
BAC/Sud Concorde	132	4000	1350	178,200
Douglas DC-8-50	146	8720	544	79,424

Criteria of performance evaluation differs among users and designers

Performance Metrics

- Response (execution) time:
  - The time between the start and the completion of a task
  - Measures user perception of the system speed
  - Common in reactive and time critical systems, single-user computer, etc.
- Throughput:
  - The total number of tasks done in a given time
  - Most relevant to batch processing (billing, credit card processing)
  - Mainly used for input/output systems (disk access, printer, etc.)

Response-time Metric

- Maximizing performance means minimizing response (execution) time

$$\text{Performance} = \frac{1}{\text{Execution time}}$$

Response-time Metric

$$\text{Performance} = \frac{1}{\text{Execution time}}$$

- Performance of Processor P1 is better than P2 if
  - **For a given work load L**
  - P1 takes less time to execute L than P2

$$\text{Performance}(P_1) > \text{Performance}(P_2) \text{ w.r.t } L$$

$$\Rightarrow \text{Execution time}(P_1, L) < \text{Execution time}(P_2, L)$$

Response-time Metric

$$\text{Performance} = \frac{1}{\text{Execution time}}$$

- Relative performance captures the performance ratio
  - For the same work load

$$\frac{\text{CPU Performance }(P_2)}{\text{CPU Performance }(P_1)} = \frac{\text{Total execution time }(P_1)}{\text{Total execution time }(P_2)}$$