Lecture 15: Moore's Law and Dennard Scaling

William Gropp
www.cs.illinois.edu/~wgropp

What Do You Think Moore's Law Says?

- Usually cast as X doubles every 18-24 months.
- Is X:
  - Computer performance
  - CPU Clock speed
  - The number of transistors per chip
  - One of the above, at constant cost?

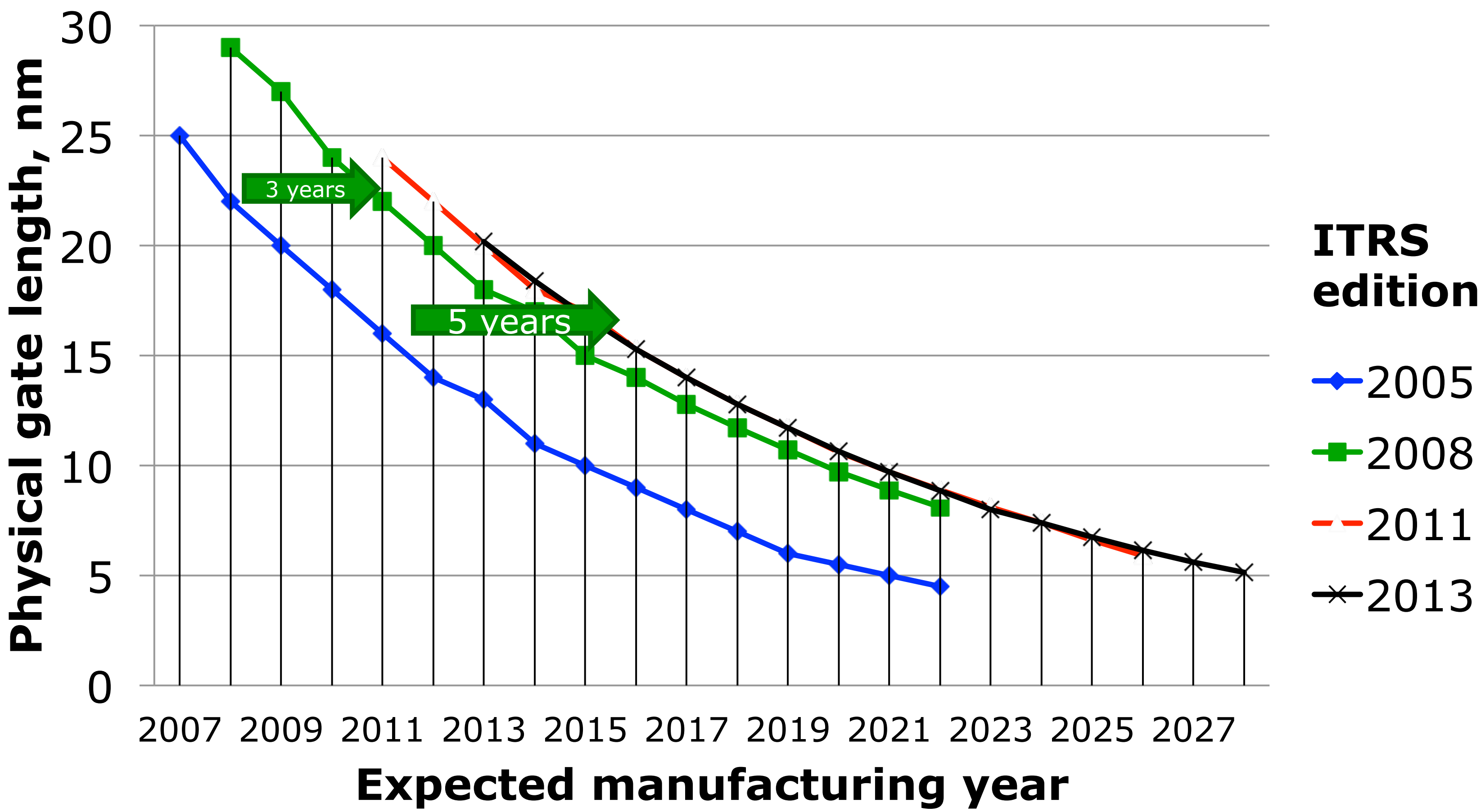
The Original Moore’s Law

- The number of transistors per chip, at constant cost
- This has not been true for years
- The improvement has been remarkable, but it is getting increasingly difficult to maintain this exponential improvement

Road Maps

- The Semiconductor industry has produced a *roadmap* of future trends and requirements
  - ♦ Semiconductor Industry Association (~1977, roadmaps from early '90s)
  - ♦ International Technology Roadmap for Seminconductors (~1998)
- ITRS Home and Current Summary
  - ♦ http://www.itrs.net/
  - ♦ http://www.itrs.net/Links/2013ITRS/2013TableSummaries/2013ORTC_SummaryTable.pdf

ITRS projections for gate lengths (nm) for 2005, 2008 and 2011 editions

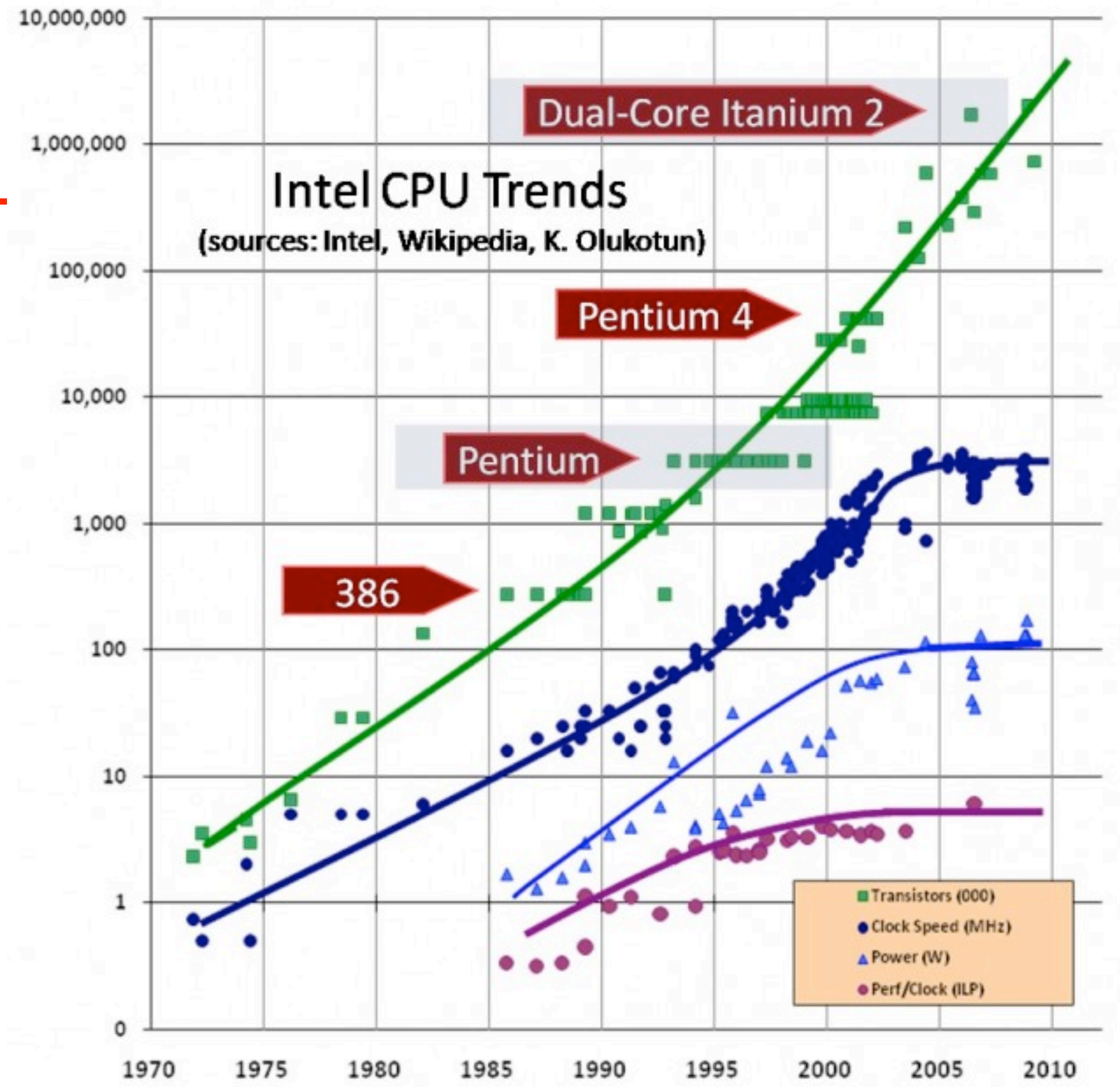


Note the rapid 3- and then 5-year shifts in ITRS projections for physical gate lengths.

PARALLEL@ILLINOIS

Why Is Moore's Law Confused with Performance?

- Its not because of parallelism – for most of its life, Moore's Law and single processor performance correlated well
- The reason is the size and speed are related – the smaller something is, the quicker it can be changed
  - ♦ Thus, smaller transistors can switch at higher speeds


Intel CPU Trends
(sources: Intel, Wikipedia, K. Olukotun)
Dual-Core Itanium 2
Pentium 4
Pentium
386
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
0
1970
1975
1980
1985
1990
1995
2000
2005
2010
Transistors (000)
Clock Speed (MHz)
Power (W)
Perf/Clock (ILP)

ILLINOIS

Dennard Scaling

- Why haven't clock speeds increased, even though transistors have continued to shrink?
- Dennard (1974) observed that voltage and current should be proportional to the linear dimensions of a transistor
  - Thus, as transistors shrank, so did necessary voltage and current; power is proportional to the area of the transistor.

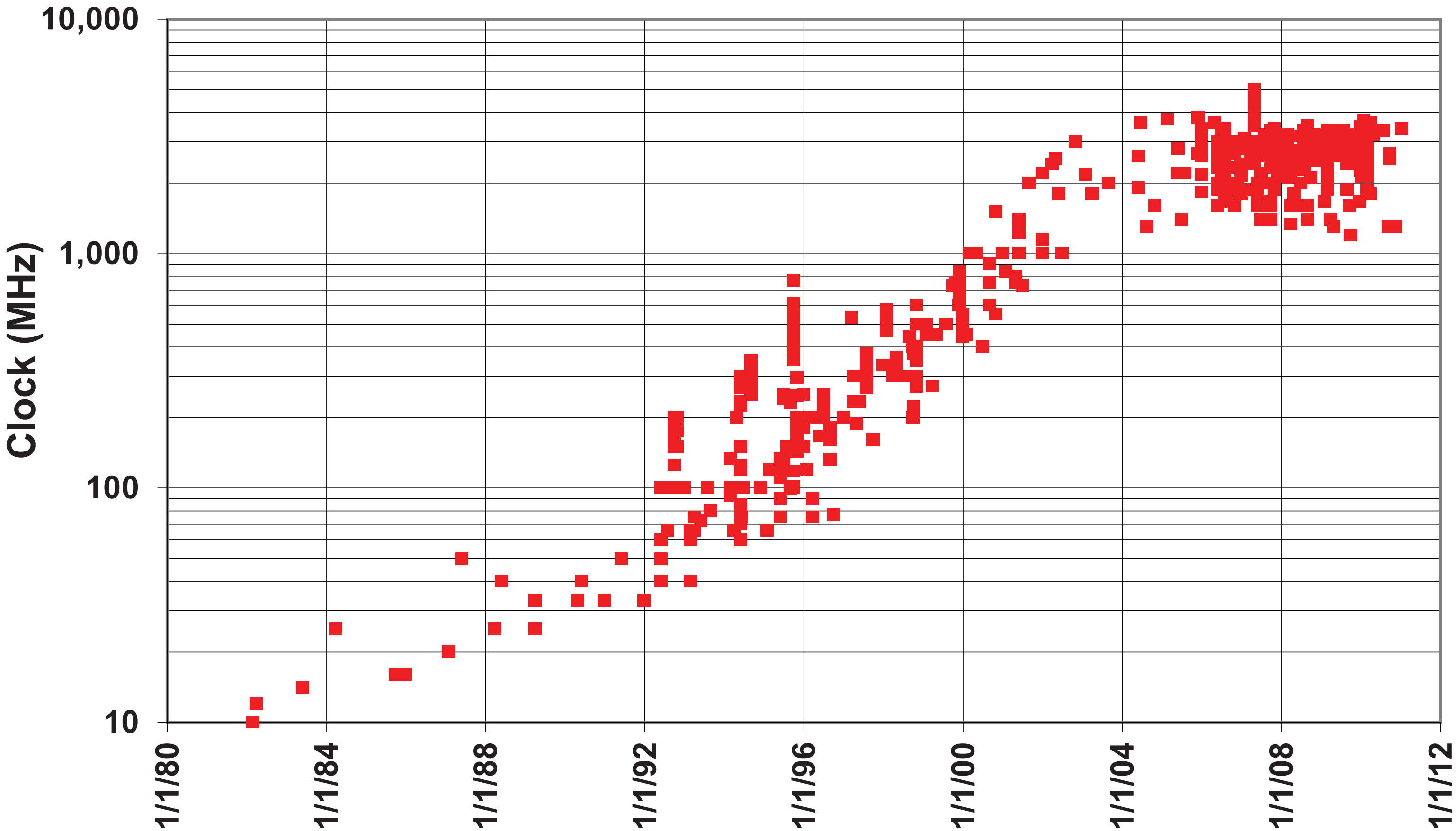
Dennard Scaling

- Power = alpha * CFV^2
  - Alpha – percent time switched
  - C = capacitance
  - F = frequency
  - V = voltage
- Capacitance is related to area
  - So, as the size of the transistors shrunk, and the voltage was reduced, circuits could operate at higher frequencies at the same power

End of Dennard Scaling

- Dennard scaling ignored the "leakage current" and "threshold voltage", which establish a baseline of power per transistor.
  - As transistors get smaller, power density increases because these don't scale with size
  - These created a "Power Wall" that has limited practical processor frequency to around 4 GHz since 2006

Historical Clock Rates



Prediction Trends

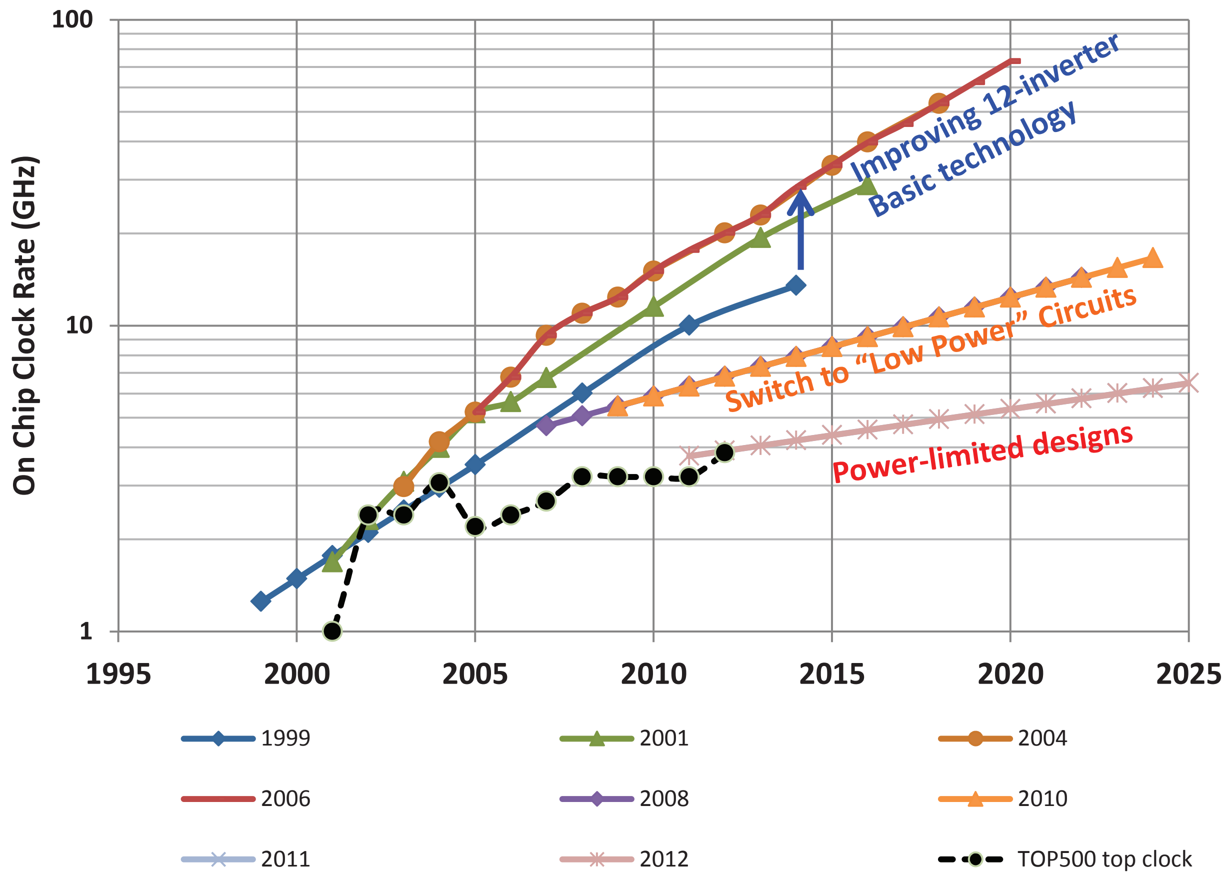


Figure 3.14. On-chip Clock from ITRS.

Questions

- True or False: Moore's law says that computer performance doubles every 18-24 months
- True or False: Moore’s law allows us to predict future properties in the same way that the law of gravity allows us to predict the path of a planet around the sun.

Questions

- If the number of transistors per unit area (i.e., on the same chip size) doubles every 24 months:
  - By what factor does the length of the side of a transistor, assuming it is square, decrease every year?
  - Starting with a length of 18.4nm (2014), what would be the size of a transistor in 2025? How does this compare to the ITRS value of 6.75um?