Processing in Memory

• Remove memory bottleneck
  – lower latency
  – Much higher bandwidth

• Things to think about
  – underlying technology
  – architecture
  – programming model
Integrating Logic and DRAM

• Logic wants to be fast
  – want to minimize capacitance
  – tends to be high power and therefore HOT

• DRAM wants to be stable
  – want to maximize capacitance
  – want low temperatures

• What do we lose by integrating the two?
  – Logic speed
  – DRAM density
Papers

• Architecture papers

Terasys
1 bit processors
Simple PIM

Active Pages
small FPGAs

IRAM
Monolithic Processor
Complex PIM

• Technology handout