
# Future Scaling of Processor-Memory Interfaces

Jung Ho Ahn[†§], Norman P. Jouppi [†], Christos Kozyrakis[‡], Jacob Leverich[‡], Robert S. Schreiber[†]

[†]HP Labs, [‡]Stanford University, [§]Seoul National University

# Executive summary

performance

reliability

## Challenges

energy efficiency

## Mainmemory system

Multicore DIMM

rank subsetting

## Solutions

efficiency/latency/
throughput tradeoffs

system-wide

## Holistic assessments

multithreaded/consolidated

chipkill

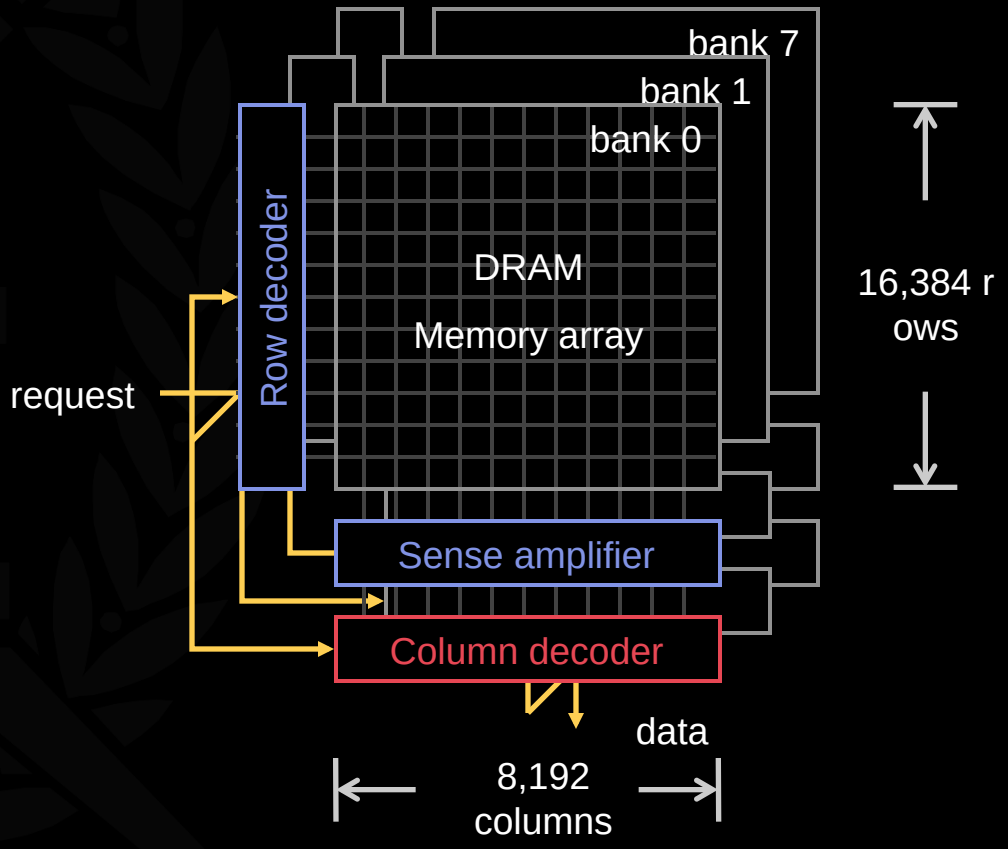
capacity vs. efficiency

# Issues on DRAM based main memory

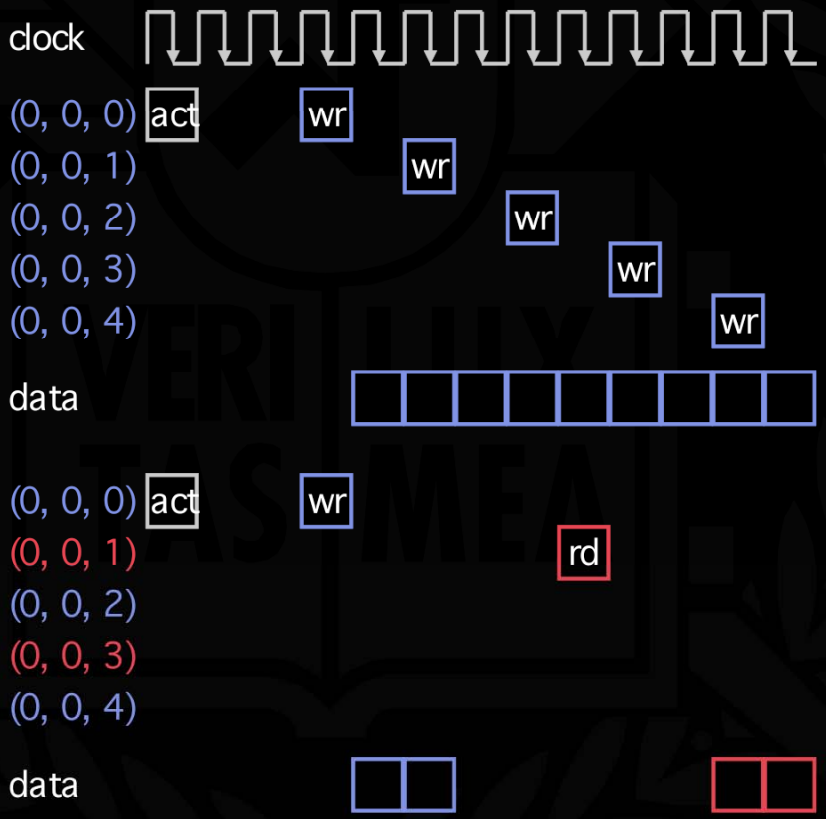
- Chip Multiprocessors (CMPs) demand
  - High capacity
  - High bandwidth
- Global wires improve slowly
  - Energy efficiency challenges = DRAM power matters!
  - Performance/power variation by access patterns
- Hard/soft errors

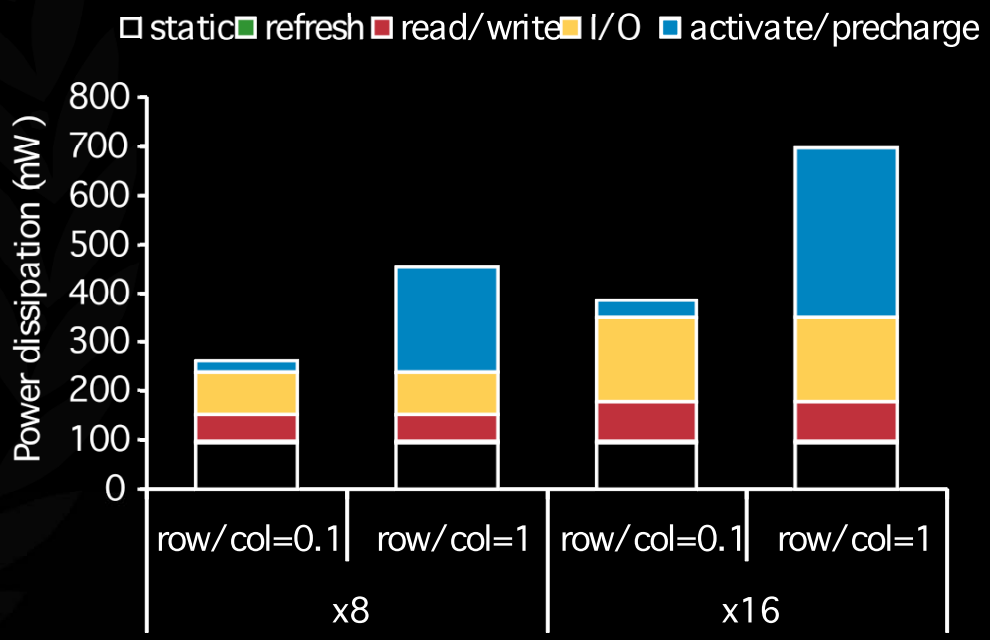
# How DRAM works


DRAM – 1T1C cell

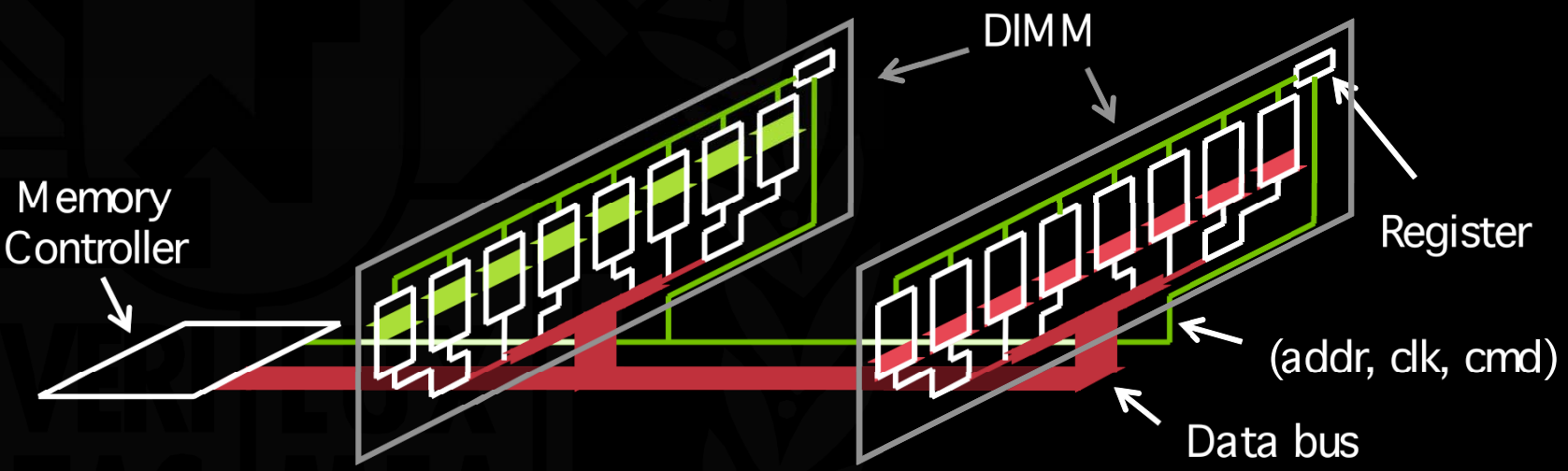


# Performance/power variations



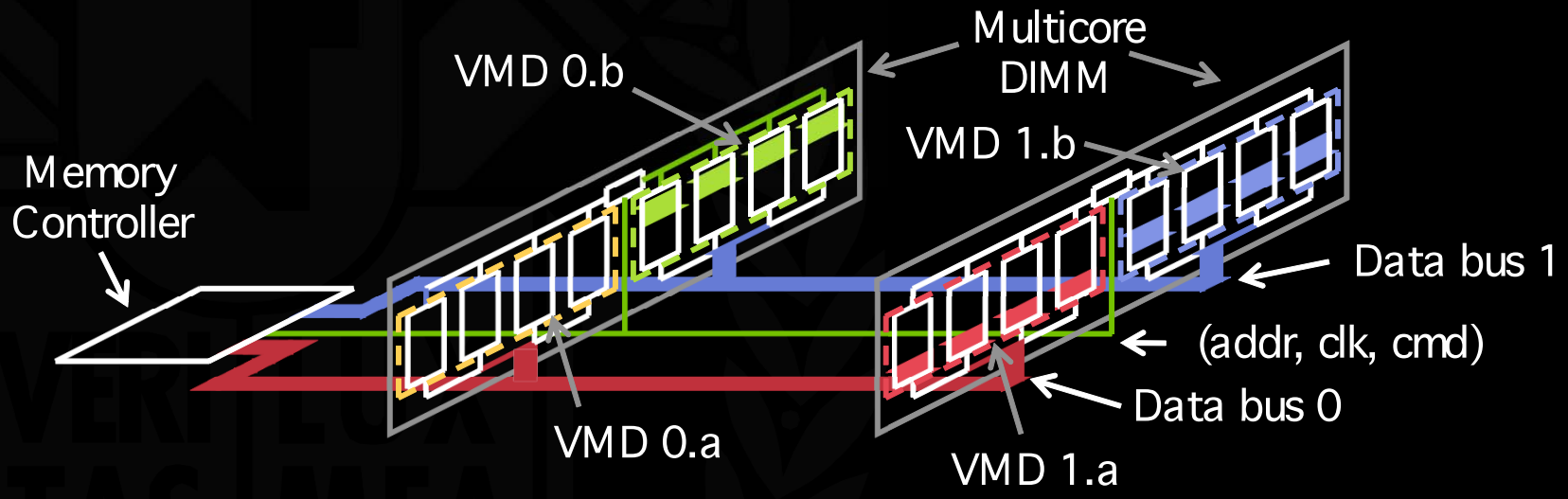
# DIMM = Dual Inline Memory Module




Overfetching problem

- DRAM row size = 8kb, 8 or 16 DRAMs per DIMM
- Cache line size = 512b
- Over 99% of bits are unused if row/col = 1

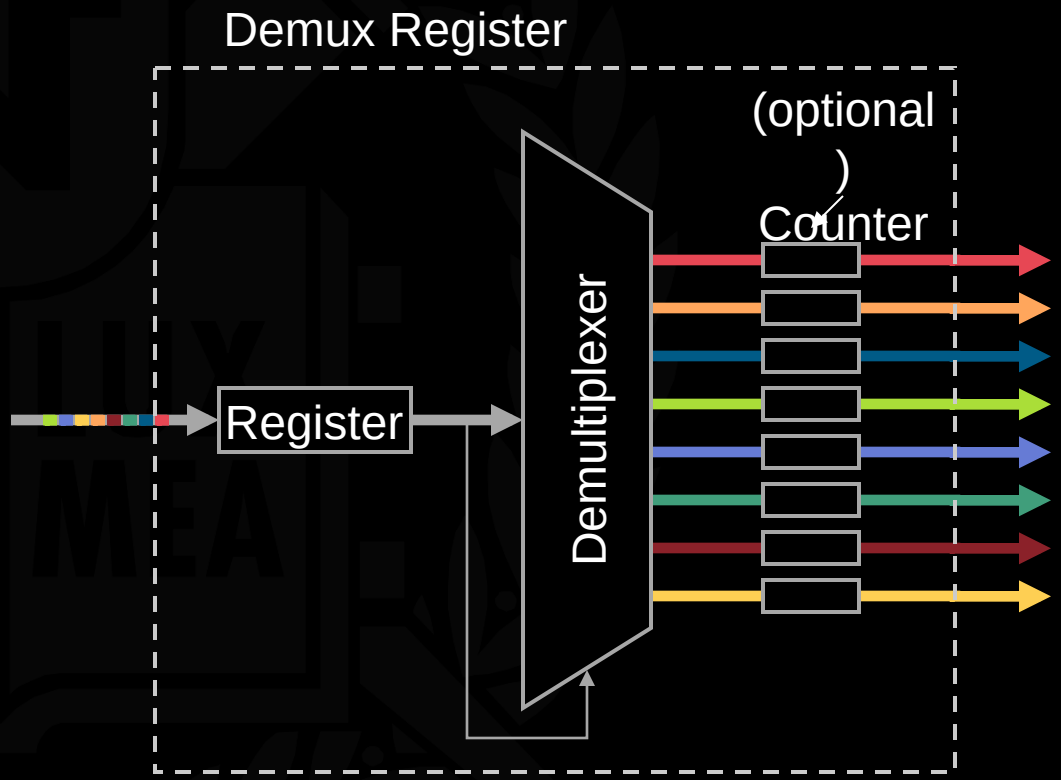
# Solution = Multicore DIMM



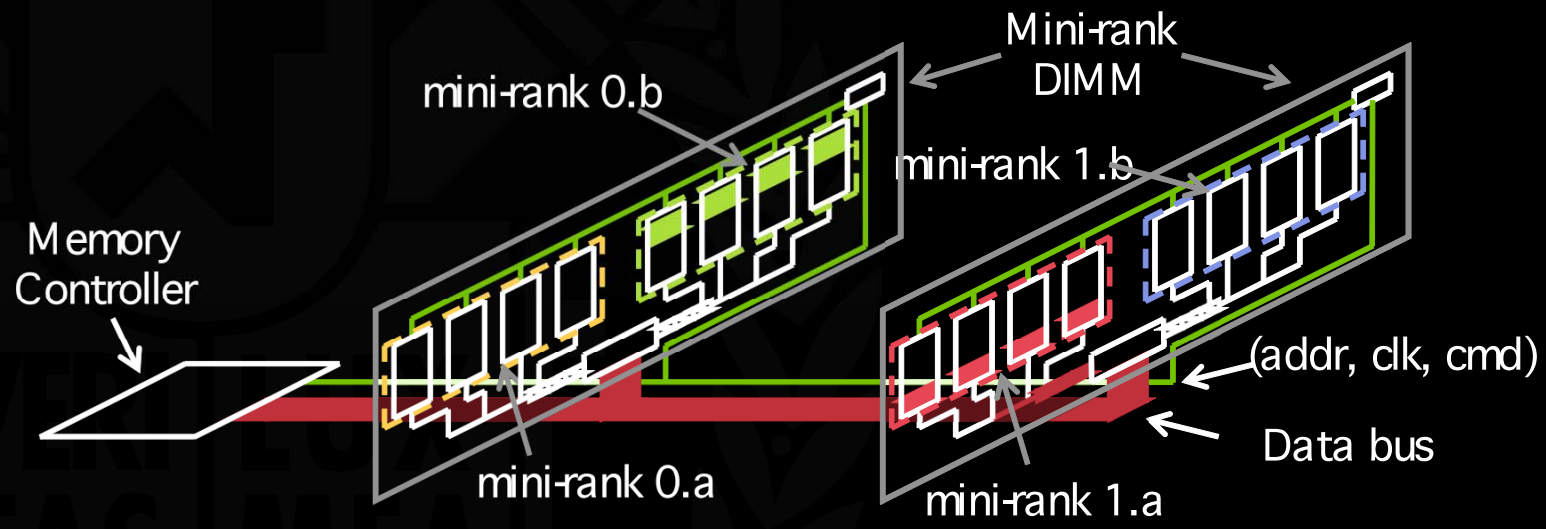

## MCDIMM features

- VMD = Virtual Memory Device : rank subsetting
- Demux register
- Over 99% of bits are unused if row/col = 1

# Demux register



# Alternative solution = mini-rank




MCDIMM vs. mini-rank

- Register for data path vs. control path
- Timing constraint due to access interference
- Load balancing between rank subsets

# Governing equations

Total main memory power =

$$D \cdot S \cdot R \cdot SP + E_{RW} \cdot BW_{RW} + D \cdot E_{AP} \cdot f_{AP}$$

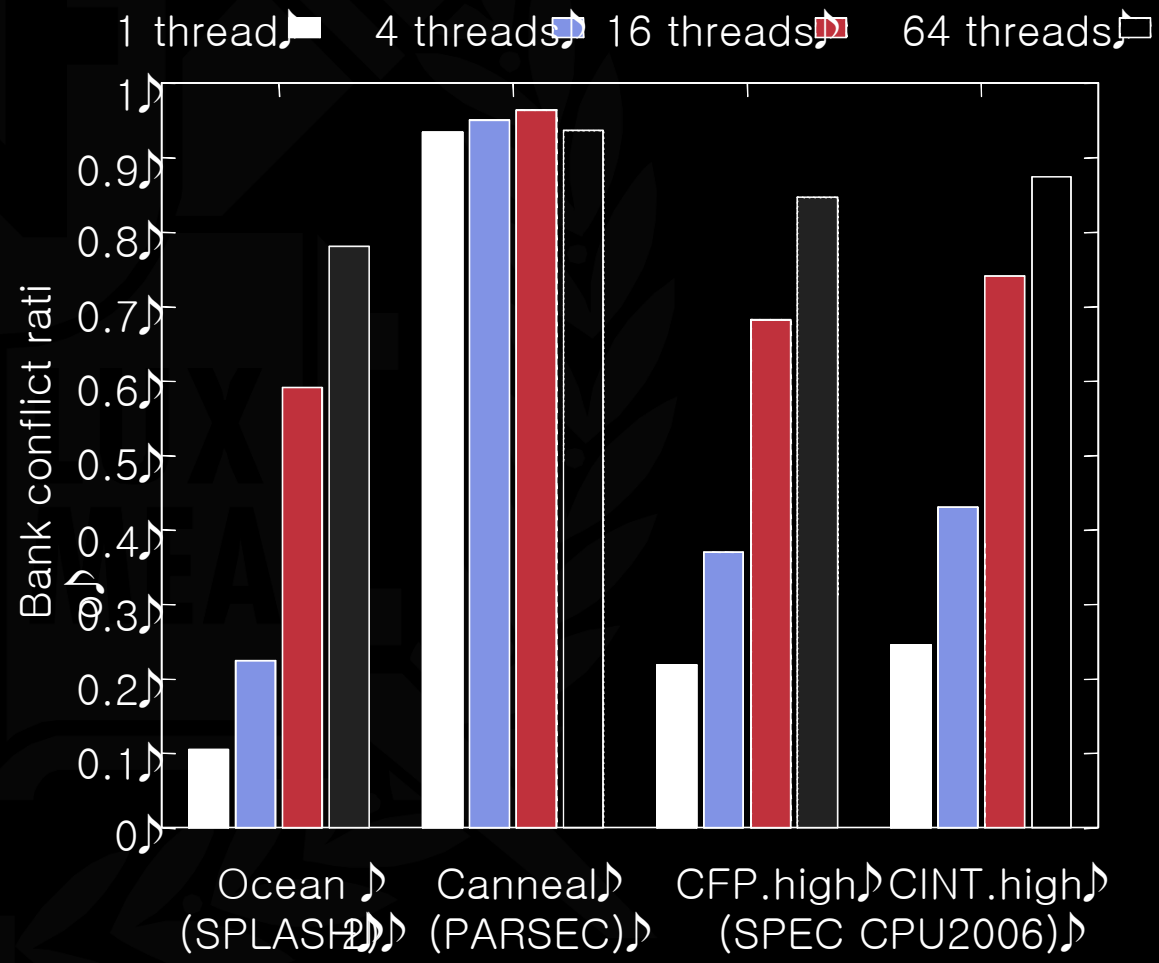
- D : # of DRAM chips per subset
- S : # of subsets per rank
- R : # of ranks per channel
- SP : static power of a DRAM chip
- $E_{RW}$ : energy needed to read/write a bit
- $BW_{RW}$ : read/write bandwidth per memory channel
- $E_{AP}$ : energy to activate/precharge a row
- $f_{AP}$ : frequency of activate/precharge per memory channel

# Governing equations

$$f_{AP} = \frac{f_{AP}}{f_{CM}} \cdot f_{CM} = \frac{f_{AP}}{f_{CM}} \cdot \frac{BW_{RW}}{CL} = \beta \cdot \frac{BW_{RW}}{CL}$$

- $BW_{RW}$ : read/write bandwidth per memory channel
- $f_{AP}$ : frequency of activate/precharge per memory channel
- $f_{CM}$ : frequency of cache miss
- CL : line size of last-level cache
- β : row/col (bank conflict ratio)

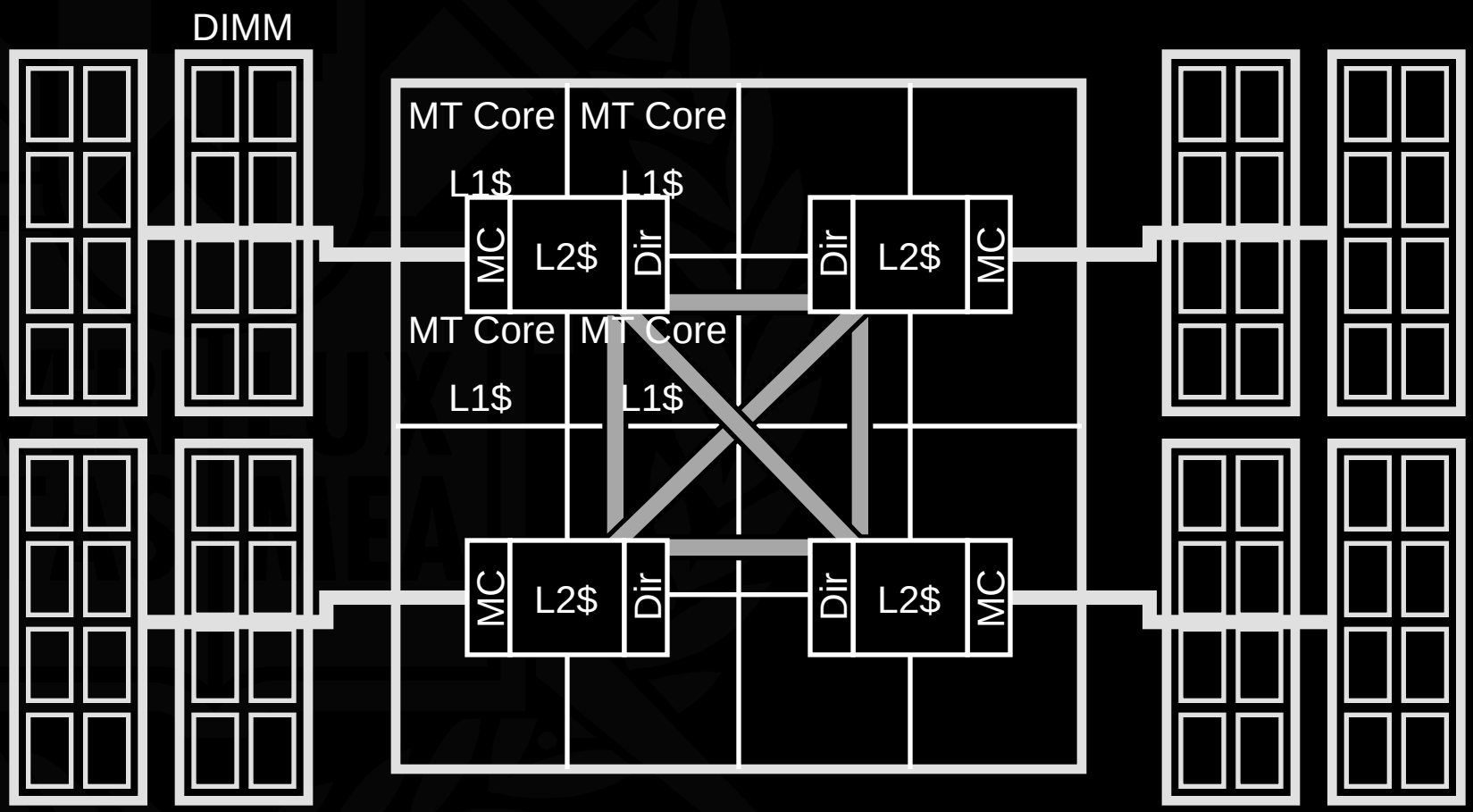
# β on multicore applications



# MCDIMM reliability issues

- SECDED
  - Single error correction, double error detection
  - Typically, (64 + 8) ECC solution is enough.
- SCCDCD
  - Single chip-error correction, double chip-error detection
  - Chipkill
  - Implementations
    - Interleaving SECDED over multiple ranks
    - Employing stronger error correcting code
      - 2b + l additional bits to correct b bits of bursty error + to detect l bits of bursty error

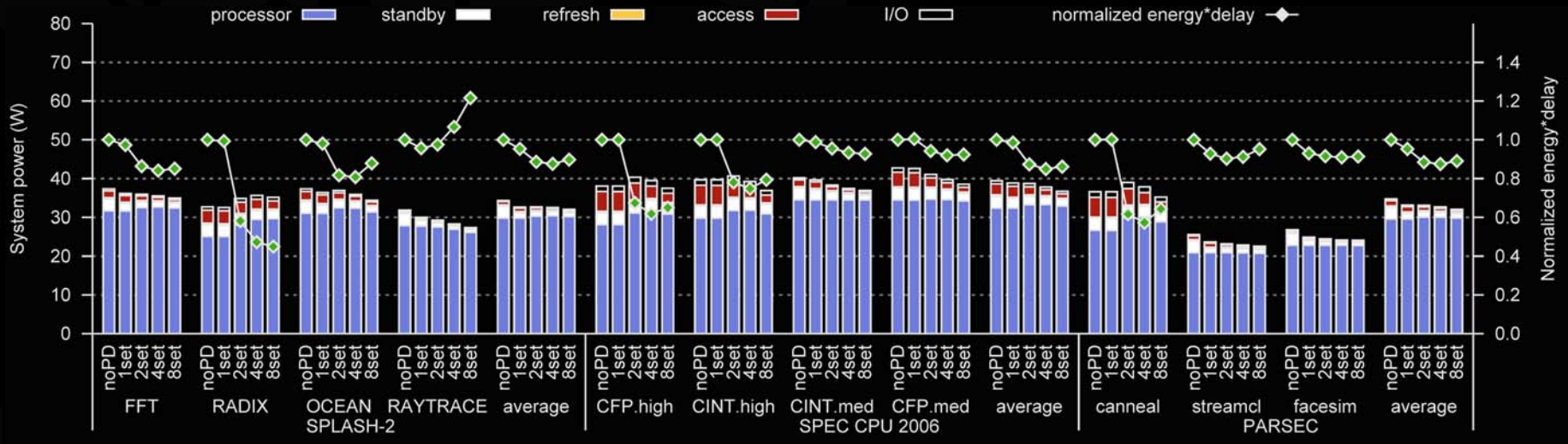
# Multicore configuration



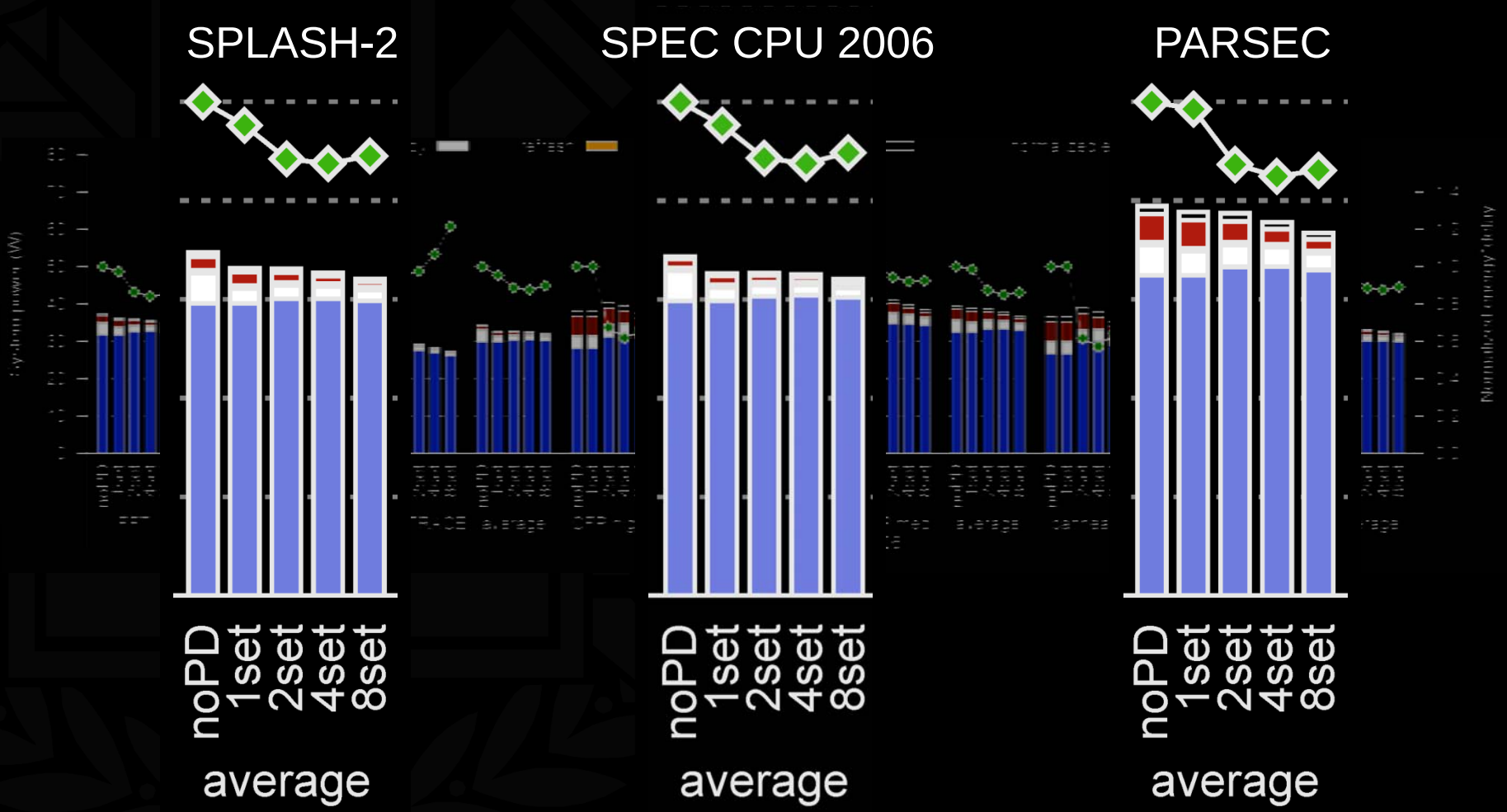
# Experimental setup

- System architecture
  - 32nm, 2GHz in-order CMT, max IPC = 16, 64 threads
  - 64B $ line, 4 1MB L2 $
  - hierarchical MESI, reverse directory
- Simulator/modeling
  - Intel Pin based in-house simulator
  - CACTI
- Applications
  - SPLASH-2/PARSEC/SPEC2006

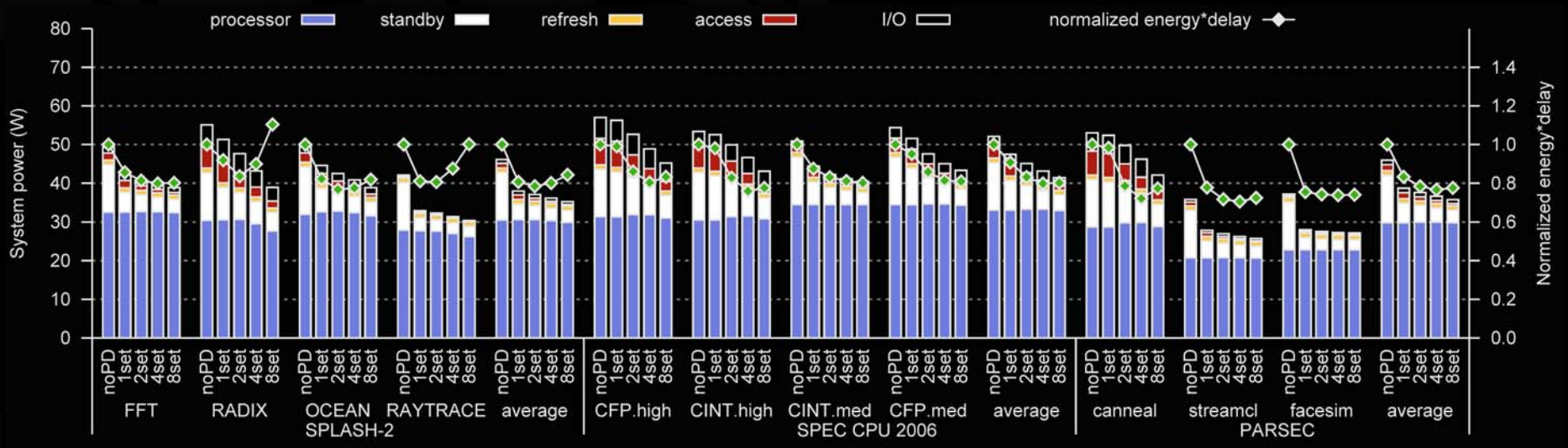
# Energy-delay product & system power with 1 rank per memory channel



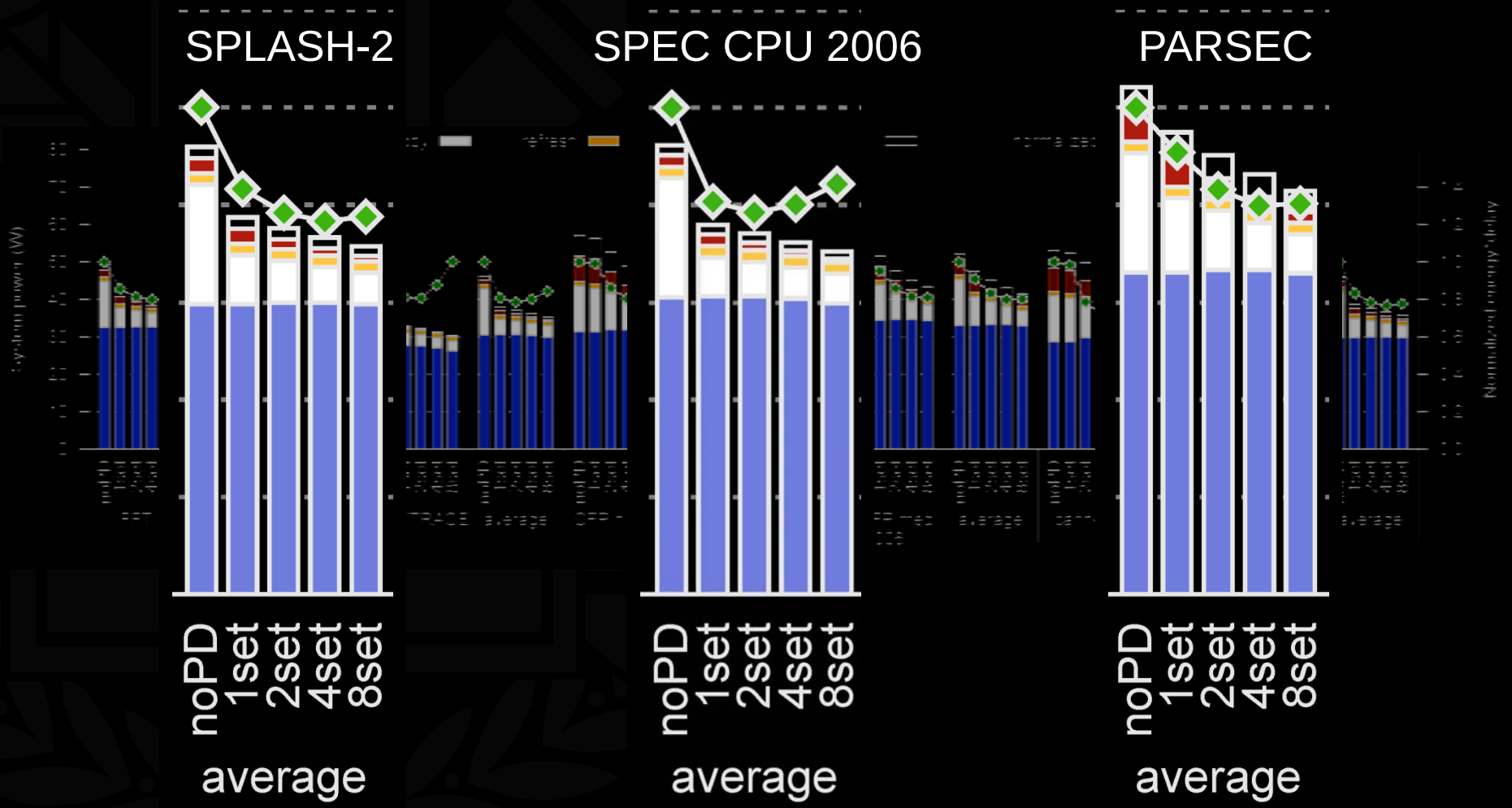
# Energy-delay product & system power with 1 rank per memory channel



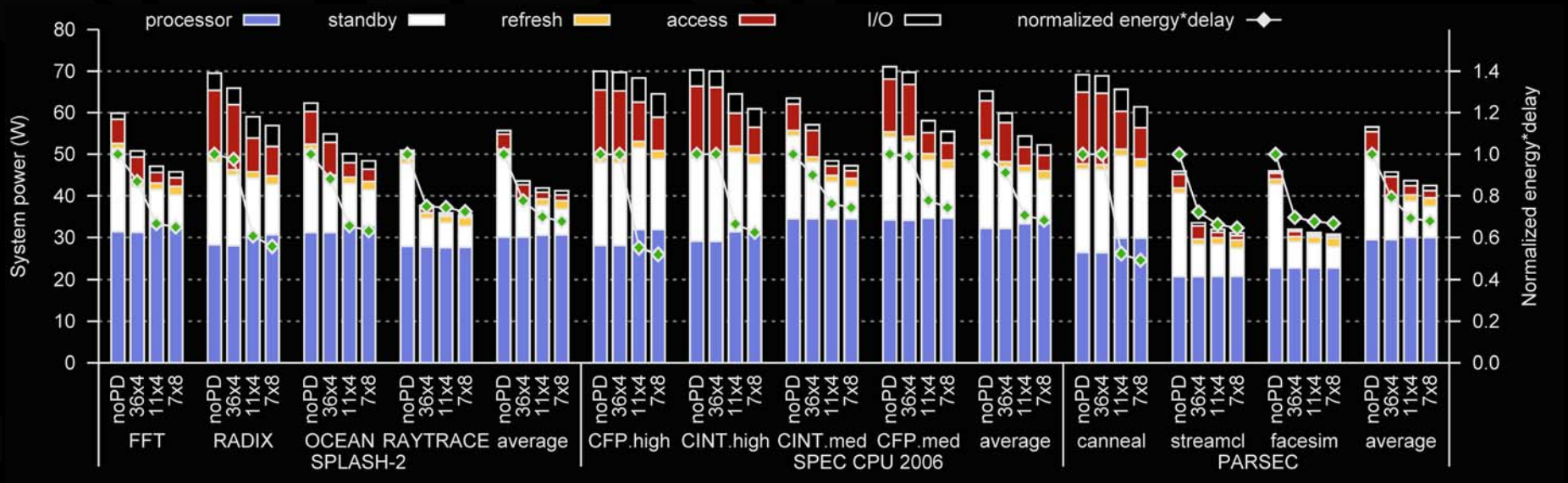
# Energy-delay product & system power with 4 ranks per memory channel



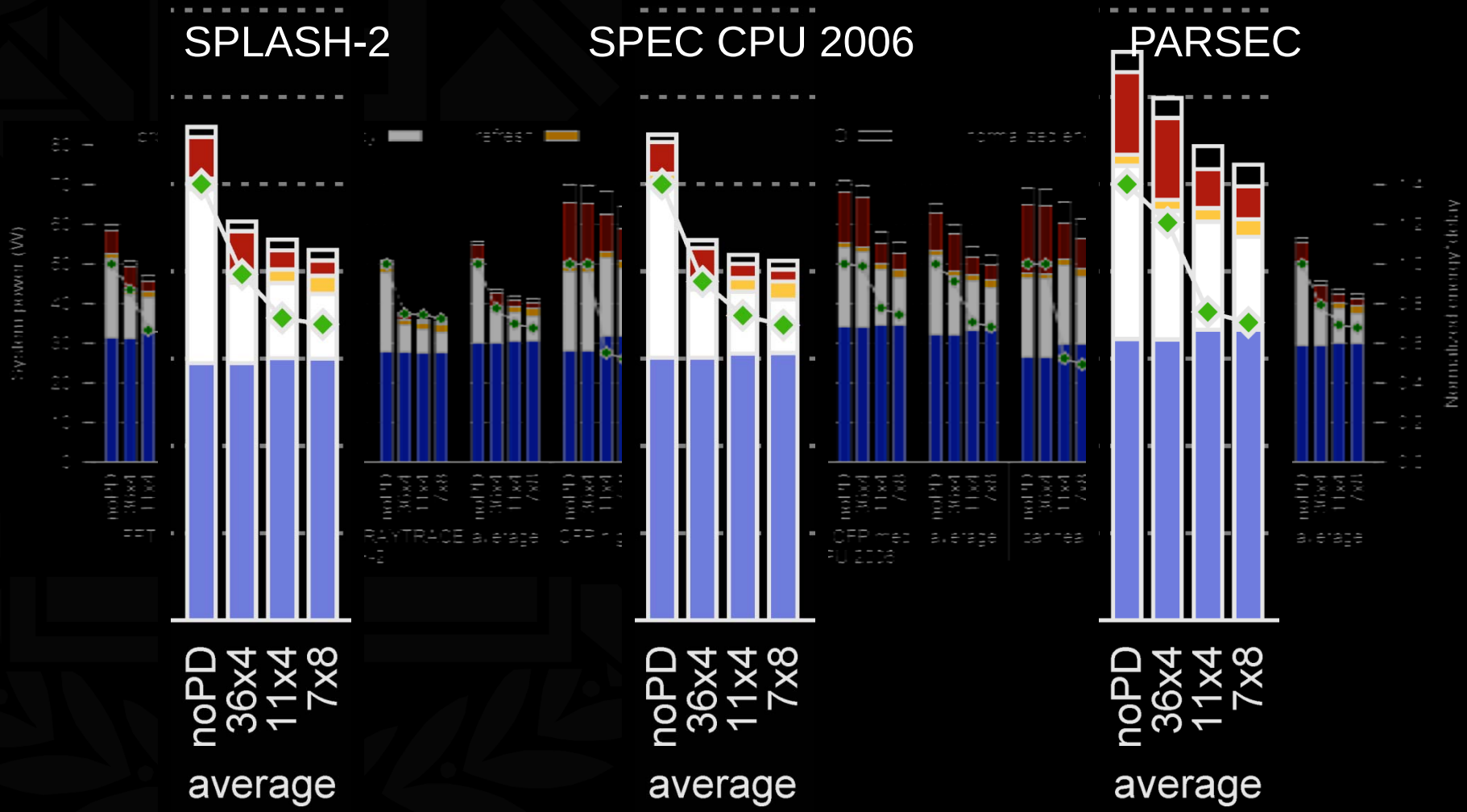
# Energy-delay product & system power with 4 ranks per memory channel



# Energy-delay product & system power with Chipkill enabled



# Energy-delay product & system power with Chipkill enabled



# Conclusion

## Challenges on main memory systems

- Performance/capacity demands
- Energy-efficiency goals
- Reliability constraints

## Multicore DIMM

- Instantiation of rank subsetting
- Gain energy efficiency & concurrency
- Sacrifice serialization latency
- Advantage in EDP (energy-delay product) with proper subsetting
- Energy-efficient, capacity-inefficient reliability solution