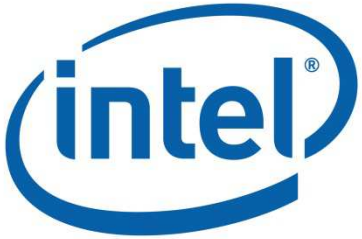
Leveraging Heterogeneity in DRAM Main Memories to Accelerate Critical Word Access



Niladrish Chatterjee

Manjunath Shevgoor

Rajeev Balasubramonian

Al Davis

Zhen Fang[‡†]

Ramesh Illikkal*

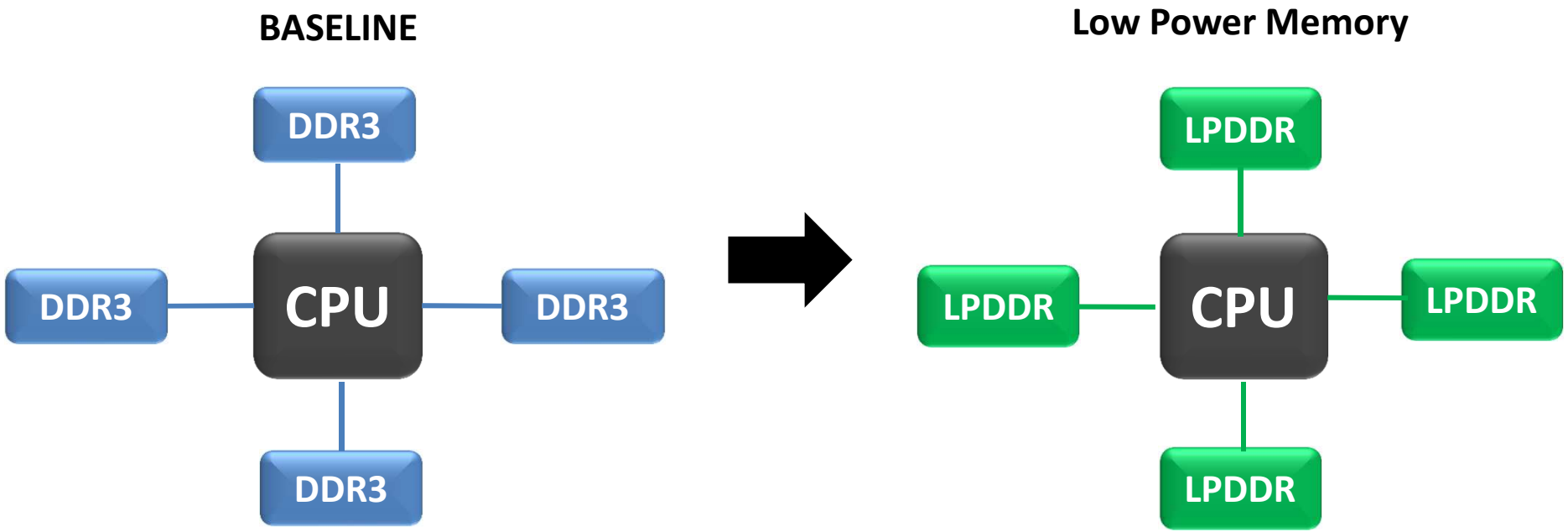
Ravi Iyer*

University of Utah , NVidia[‡] and Intel Labs*

[†]Work done while at Intel

Memory Bottleneck

- High DRAM power consumption
- Recent proposals for Low-Power DRAM instead of DDR3

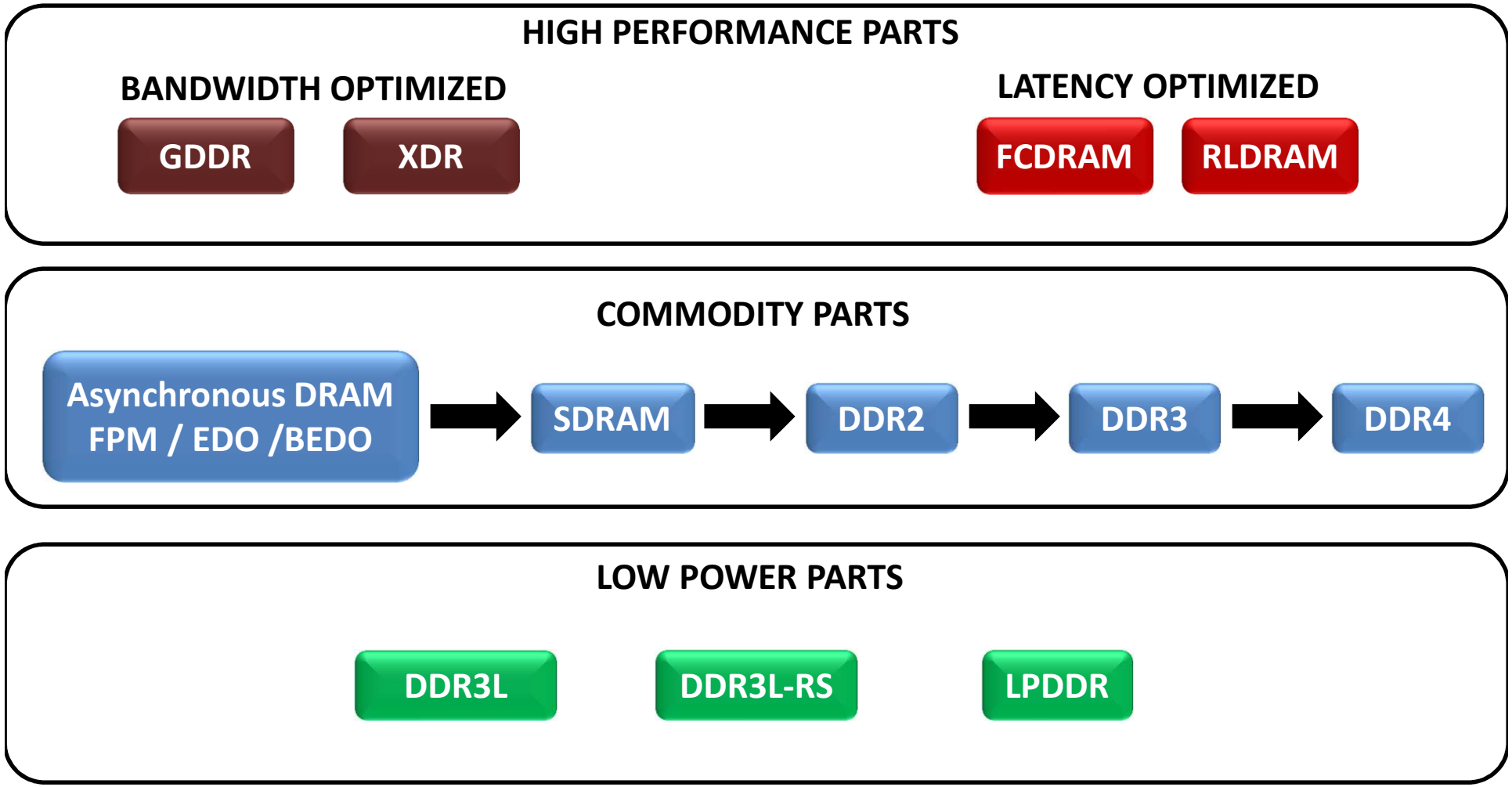


- The use of LPDRAM does not address the latency wall problem that continues to grow

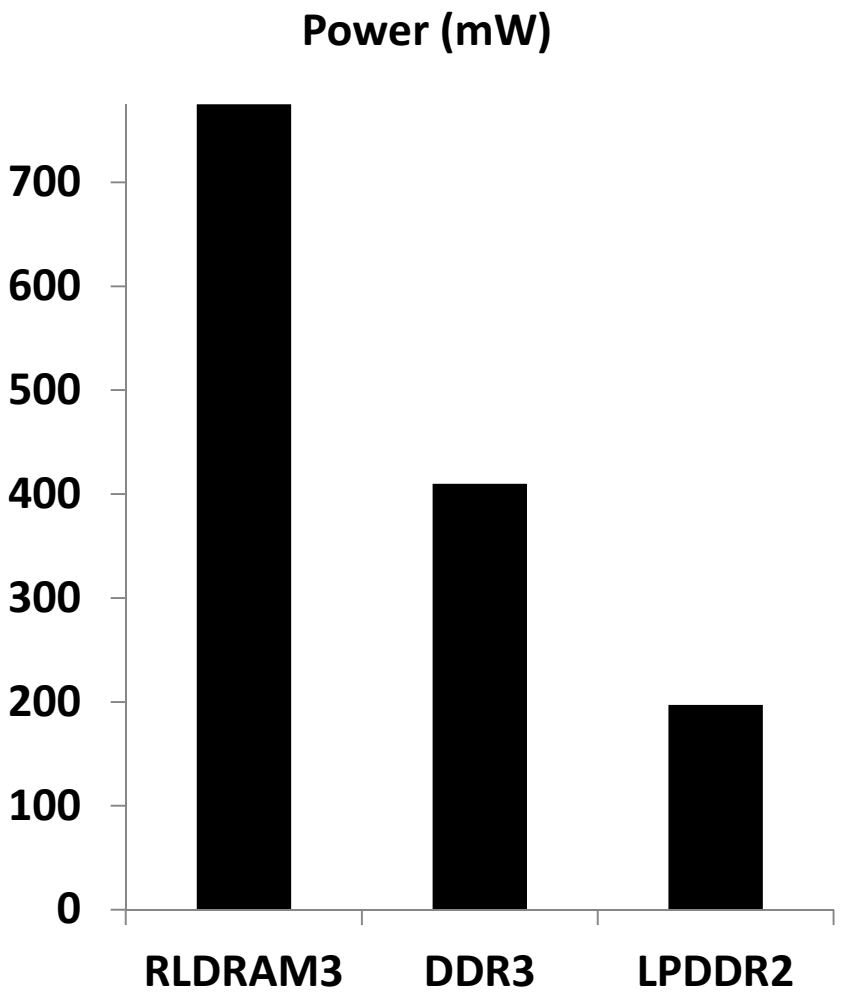
UTAH ARCH

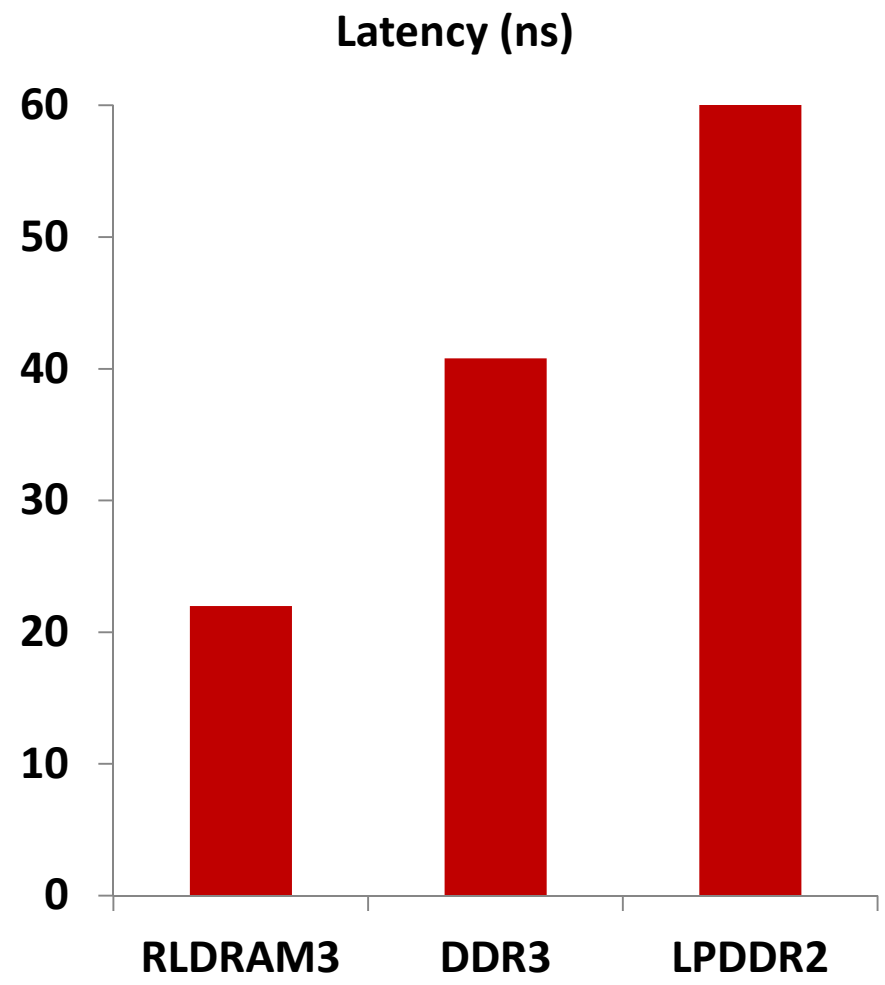
intel

DRAM Variants



No one memory works best

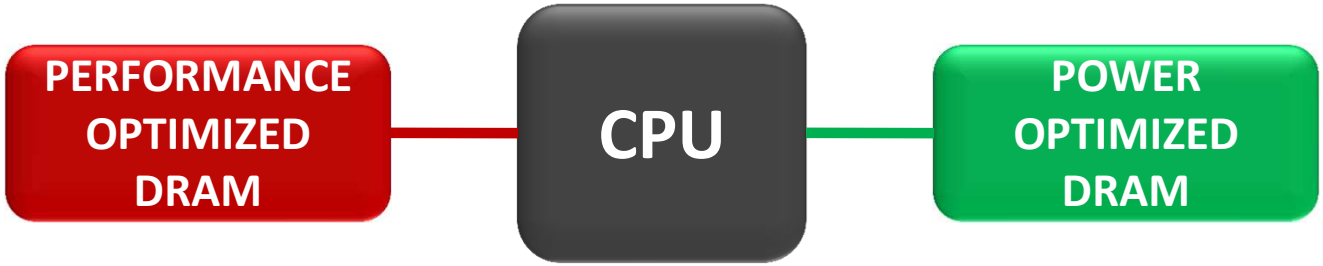




UTAH ARCH

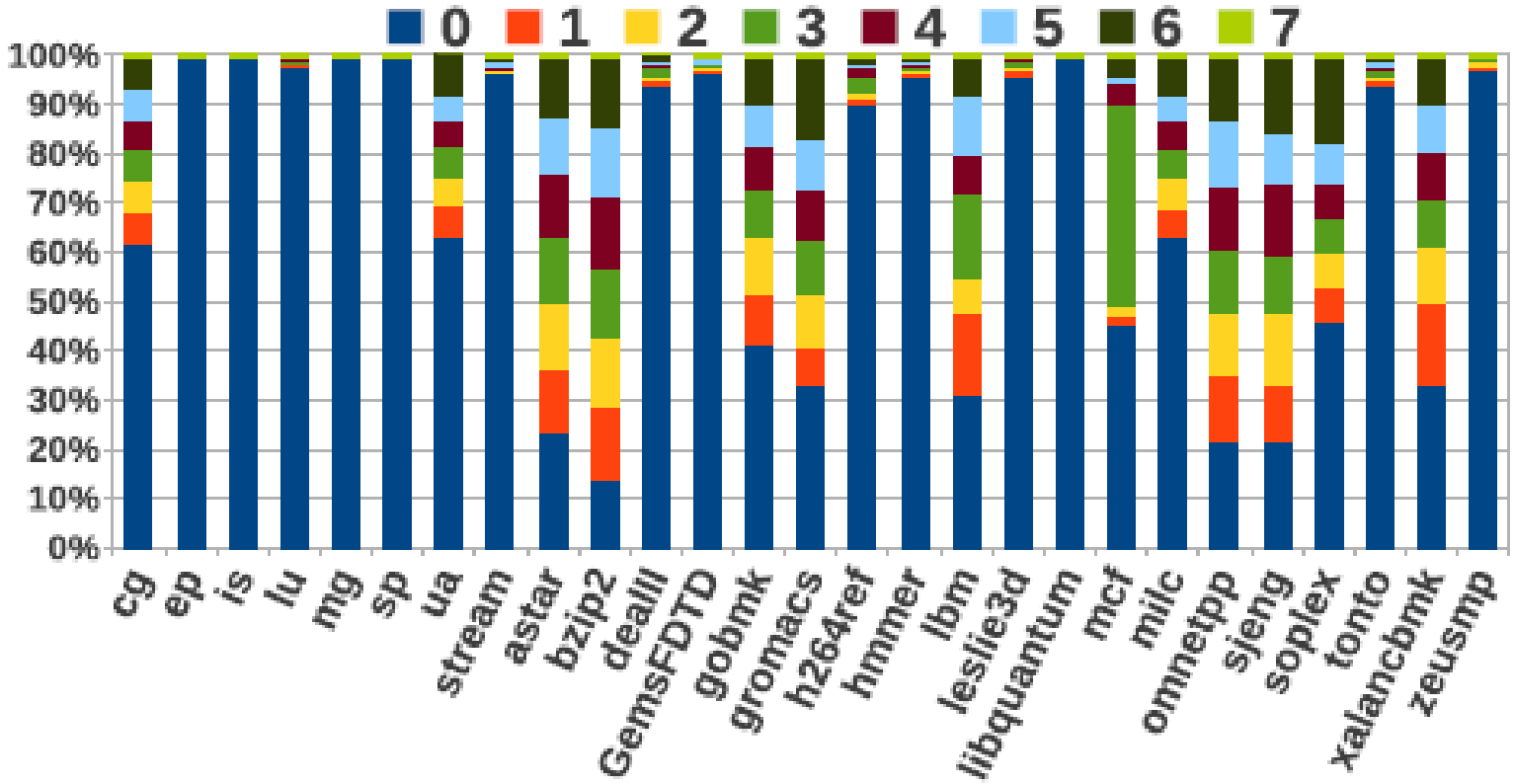
intel

Heterogeneous Memory



- Combine high-performance and low-power dram to outperform DDR3 at a lower energy cost
- Many design options

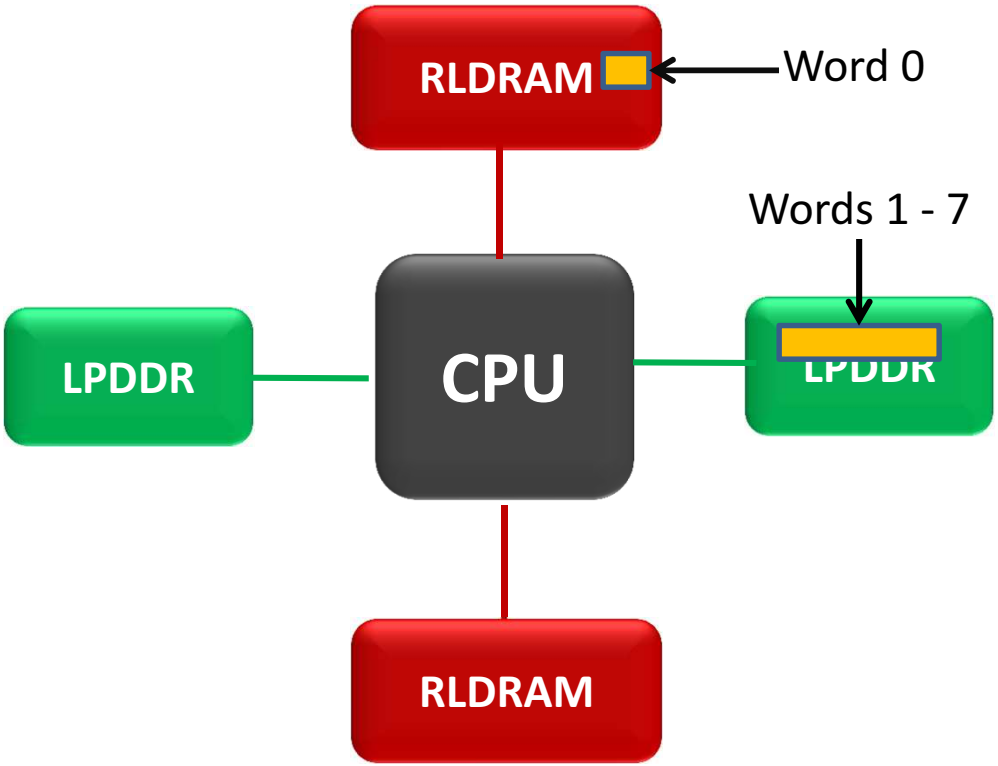
Critical Word Regularity



Frequency of accesses to individual words of a cache-line

- Most DRAM requests are for word-0 of the cache-line

Critical Word Acceleration



- Critical Word from RLDRAM
- Rest of cache-line from LPDRAM
- Improves both system throughput (13%) and energy (6%)

UTAH ARCH

intel