Technische Universität Kaiserslautern

TU/e Technische Universiteit Eindhoven University of Technology

DRAM Selection and Configuration for Real-Time Mobile Systems

Manil Dev Gomony, Christian Weis, Benny Akesson, Norbert Wehn, and Kees Goossens

Problem statements

1. ***Analyze the worst-case bandwidth, average-case execution time, and power consumption of mobile DRAMs within and across generations.***
2. ***Devise a methodology to select memory configurations in real-time mobile systems running applications with hard and soft real-time requirements.***

DRAM overview

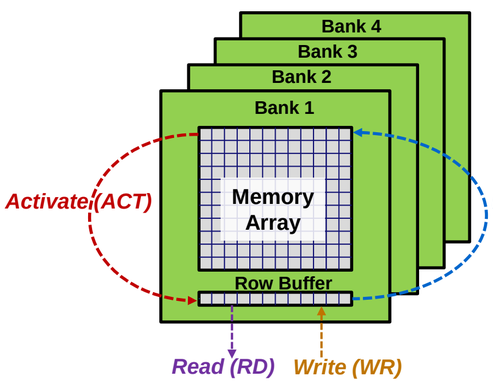


Figure 1: SDRAM architecture

→ Read/Write operations are performed by issuing *memory commands*

→ Timing constraints between memory commands impacts memory efficiency

$$Memory\ efficiency = \frac{Clock\ cycles\ containing\ useful\ data}{Total\ clock\ cycles}$$

→ Memory efficiency depends on *operating frequency, interface width, data-rate and memory map (BL, BI and BC)*

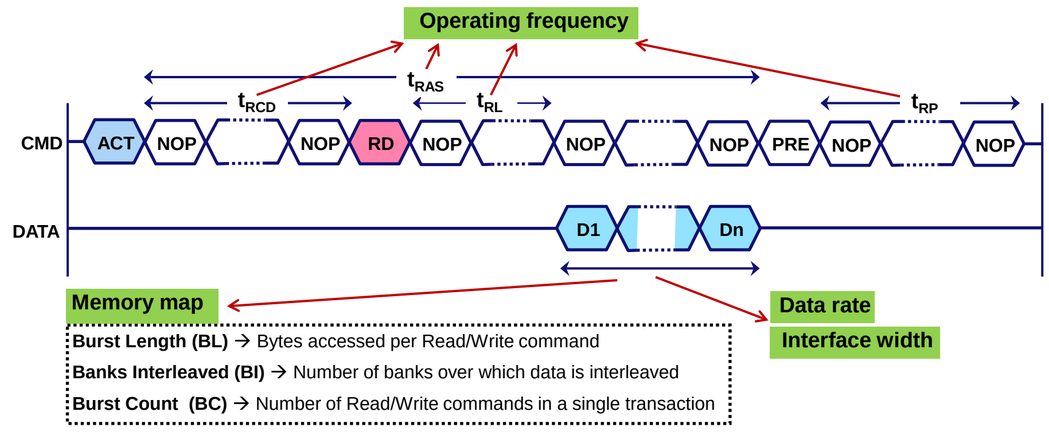
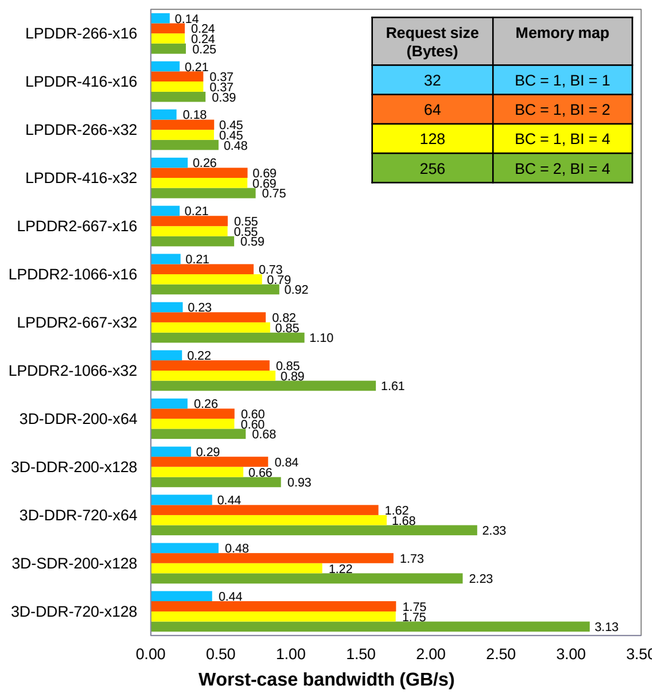


Figure 2: Memory read operation showing the impact of memory configurations on memory efficiency

Worst-case analysis



Request size (Bytes)	Memory map
32	BC = 1, BI = 1
64	BC = 1, BI = 2
128	BC = 1, BI = 4
256	BC = 2, BI = 4

Figure 3: Worst-case bandwidth of mobile memories across and within generations for different request sizes with their optimal memory map configurations.

→ LPDDR, LPDDR2, 3D-SDR (JEDEC) and 3D-DDR guarantee up to 0.75 GB/s, 1.6 GB/s, 2.2GB/s and 3.1 GB/s

→ Selection criteria of memory map (BI, BC and BL):

1. *Access granularity ≤ request size*
   (Access granularity = BI × BC × BL × IO Width)
   → Data fetched from memory is not discarded
2. *Interleave data to the maximum number of banks (BI) to exploit bank-level parallelism*
   → Bank-level parallelism amortizes overhead
3. *After satisfying 1 and 2, increase BC*
   → Maximum efficiency in a single transaction

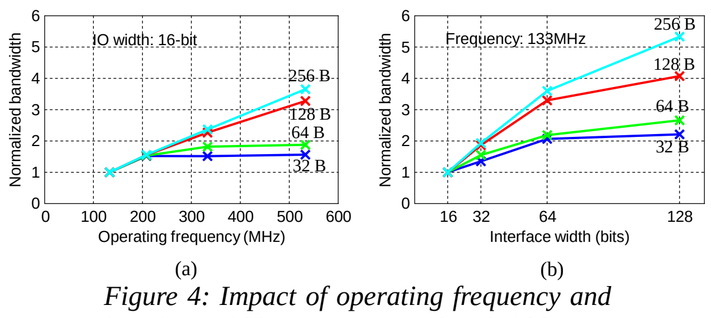


Figure 4: Impact of operating frequency and interface width on worst-case bandwidth.

→ Selection criteria of IO width and operating frequency:

1. *Select the widest interface as long as the access granularity is less than or equal to request size*
2. *Select a higher operating frequency*
   → Overhead increases with increase in operating frequency but not when interface goes wider

Average-case analysis

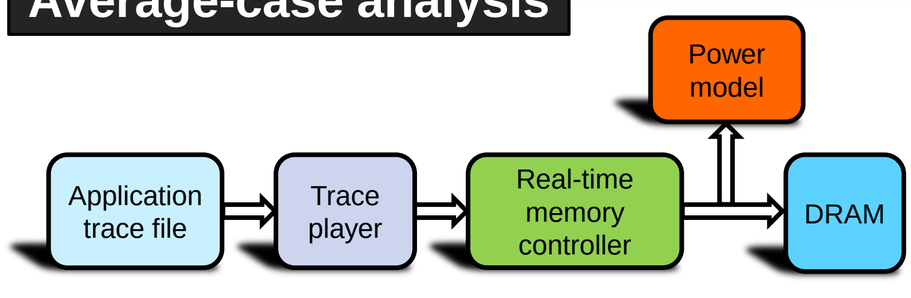


Figure 5: Experimental setup.

→ Application trace: memory requests by running a H.263 video decoder application in *SimpleScalar*

→ Real-time memory controller: *Predator* [2]

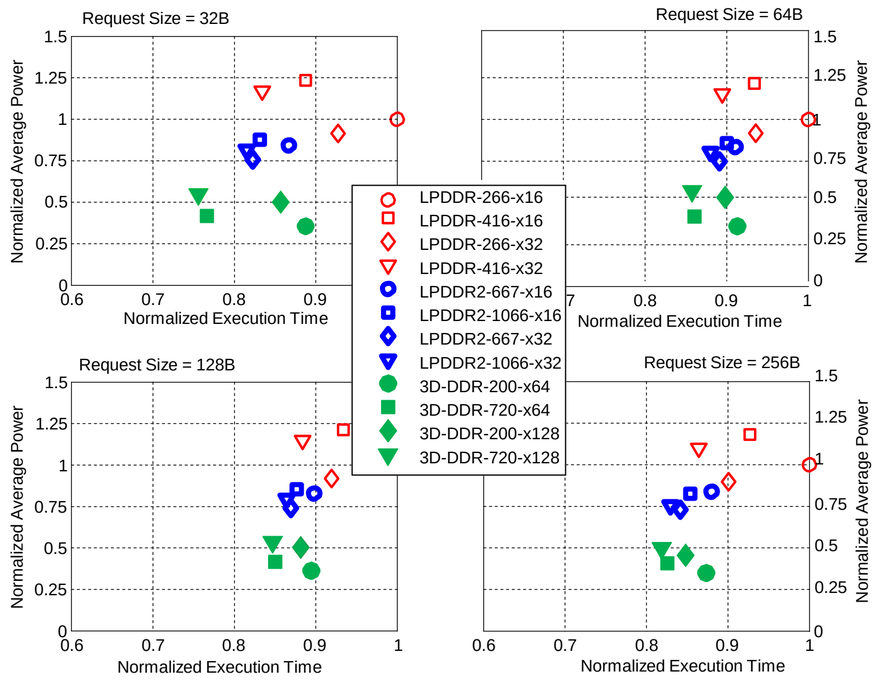
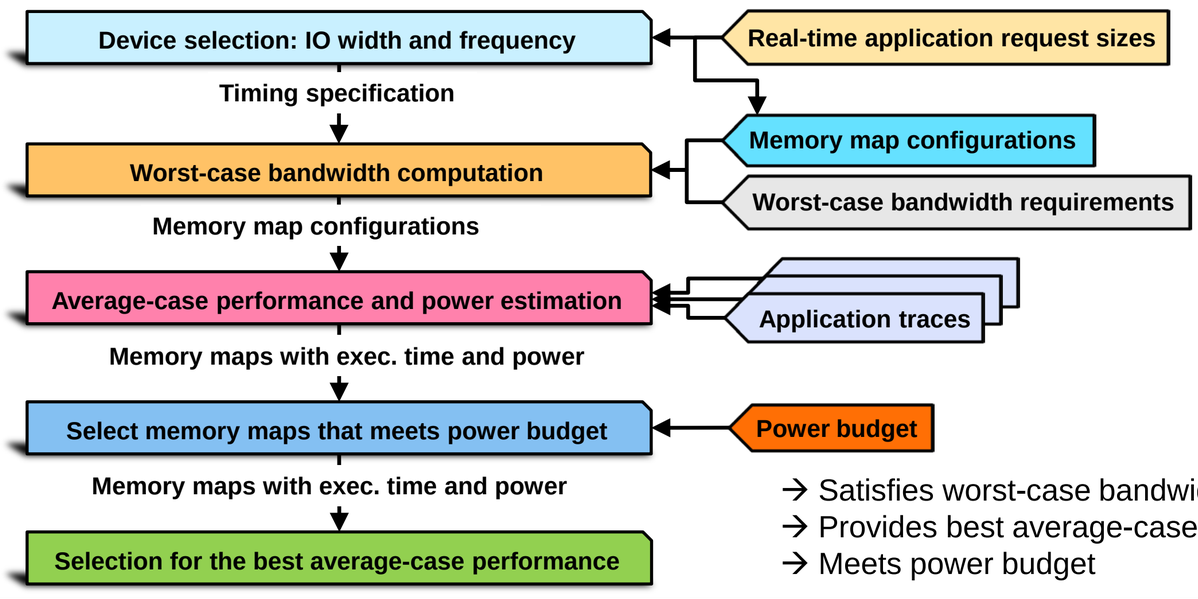


Figure 6: Execution time vs Power of mobile memories normalized to LPDDR-266-x16 for different request sizes.

DRAM selection and configuration methodology



→ Satisfies worst-case bandwidth requirements
→ Provides best average-case performance
→ Meets power budget

Conclusions

Memory Device	Worst-case bandwidth	Power savings w.r.t LPDDR-266-x16	Performance gain w.r.t LPDDR-266-x16
LPDDR-416-x32	*0.75 GB/s*	*-15%*	*14%*
LPDDR2-1066-x32	*1.6 GB/s*	*25%*	*18%*
3D-DDR-720-x128	*3.1 GB/s*	*67%*	*25%*

→ Wider interface and lower operating speed gives better performance at a lower power consumption

References

[1] M.D.Gomony et al., "DRAM Selection and Configuration for Real-Time Mobile Systems." in Proc. DATE, 2012.
[2] B. Akesson et al., "Predator: A predictable SDRAM memory controller," in Proc. CODES+ISSS, 2007.

Manil Dev Gomony
m.d.gomony@tue.nl

Acknowledgements
This work was partially funded by projects EU FP7 288008 T-CREST and 288248 Flextiles, Catrene CA104 Cobra, and NL STW 10346 NEST