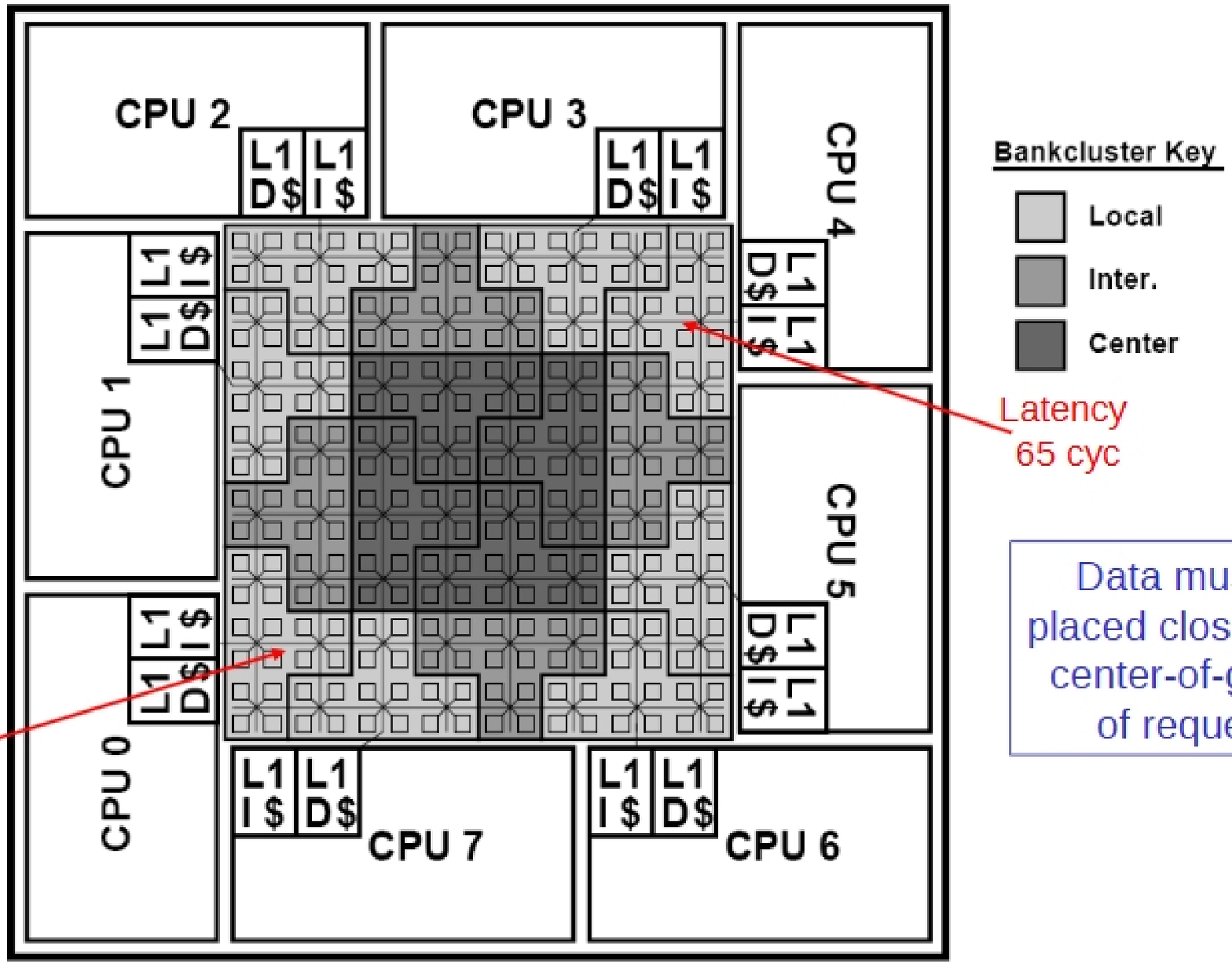
Lecture 12: Large Cache Design

Papers (papers from last class and…):

- Co-Operative Caching for Chip Multiprocessors, Chang and Sohi, ISCA'06
- Victim Replication, Zhang and Asanovic, ISCA'05
- Interconnect Design Considerations for Large NUCA Caches, Muralimanohar and Balasubramonian, ISCA'07
- Design and Management of 3D Chip Multiprocessors using Network-in-Memory, Li et al., ISCA'06
- A Domain-Specific On-Chip Network Design for Large Scale Cache Systems, Jin et al., HPCA'07
- Nahalal: Cache Organization for Chip Multiprocessors, Guz et al., Comp. Arch. Letters, 2007

Beckmann and Wood, MICRO'04



Data must be placed close to the center-of-gravity of requests

Examples: Frequency of Accesses

Dark → more accesses

←OLTP (on-line transaction processing)

Ocean → (scientific code)

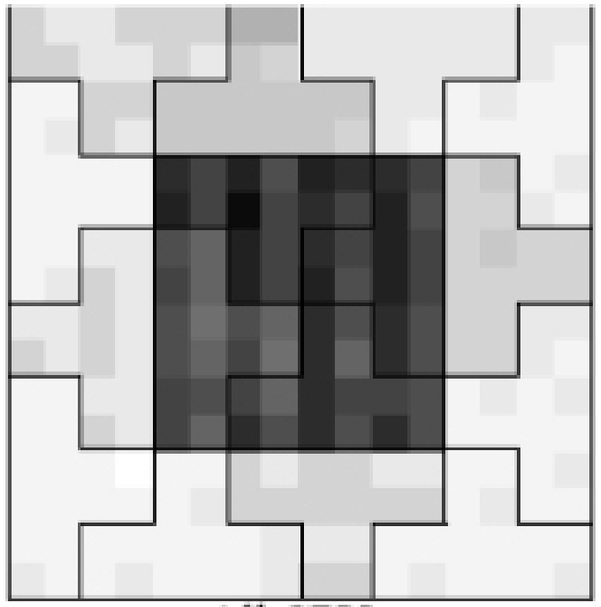


Figure 10. oltp L2 Hit Distribution

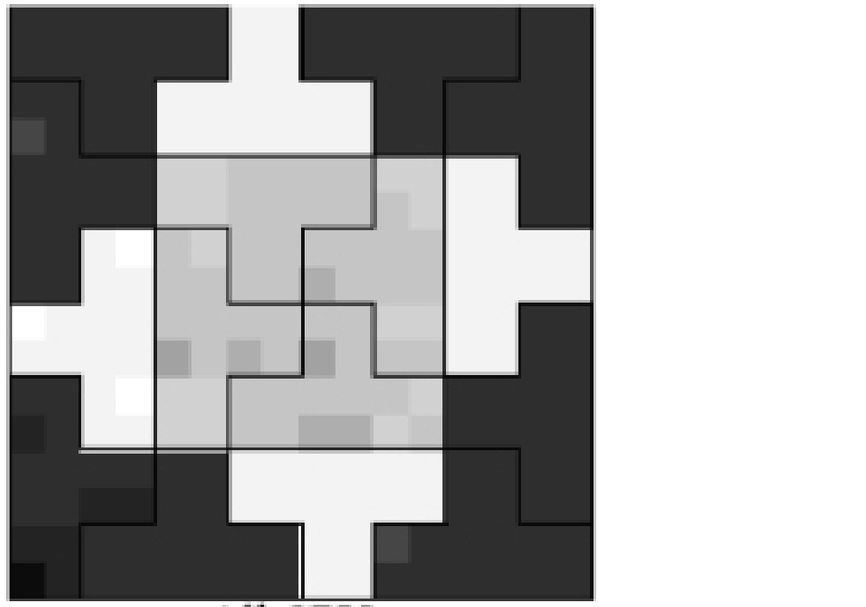


Figure 11. ocean L2 Hit Distribution