

*7th Workshop on Scalable Shared Memory Multiprocessor
25th Annual International Symposium on Computer Architecture*

Memory System Performance of High End SMPs, PCs and Clusters of PCs

Ch. Kurmann, T. Stricker

Laboratory for Computer Systems
ETHZ - Swiss Institute of Technology
CH-8092 Zurich

Color Slides: www.cs.inf.ethz.ch/CoPs/isca98ws/

Memory Systems

- Low End designs in PCs:
 - ◆ extremely low cost
 - ◆ standard I/O interface
- High End designs in “Killer” Workstations:
 - ◆ well engineered memory systems
 - ◆ support for additional datastreams
 - ◆ better I/O busses
- Are Low End SMPs the universal compute nodes for parallel and distributed systems?

Contribution of this talk

- The answer is probably the memory system performance.
- How significant are the differences in memory system performance?
- Limitations of Low End memory systems
 - ◆ for local computation (e.g. in scientific applications)
 - ◆ for inter-node communication (e.g. in databases)

Extended Copy Transfer Characterization

ECT is a method to characterize the performance of memory systems (ISCA95 and HPCA97):

- ◆ Categories

- ◆ Access pattern, stride (*spatial locality*)
- ◆ Working set (*temporal locality*)

- ◆ Value

- ◆ Transfer bandwidth (large amount of data)

- ◆ Same chart resulting from one microbenchmark

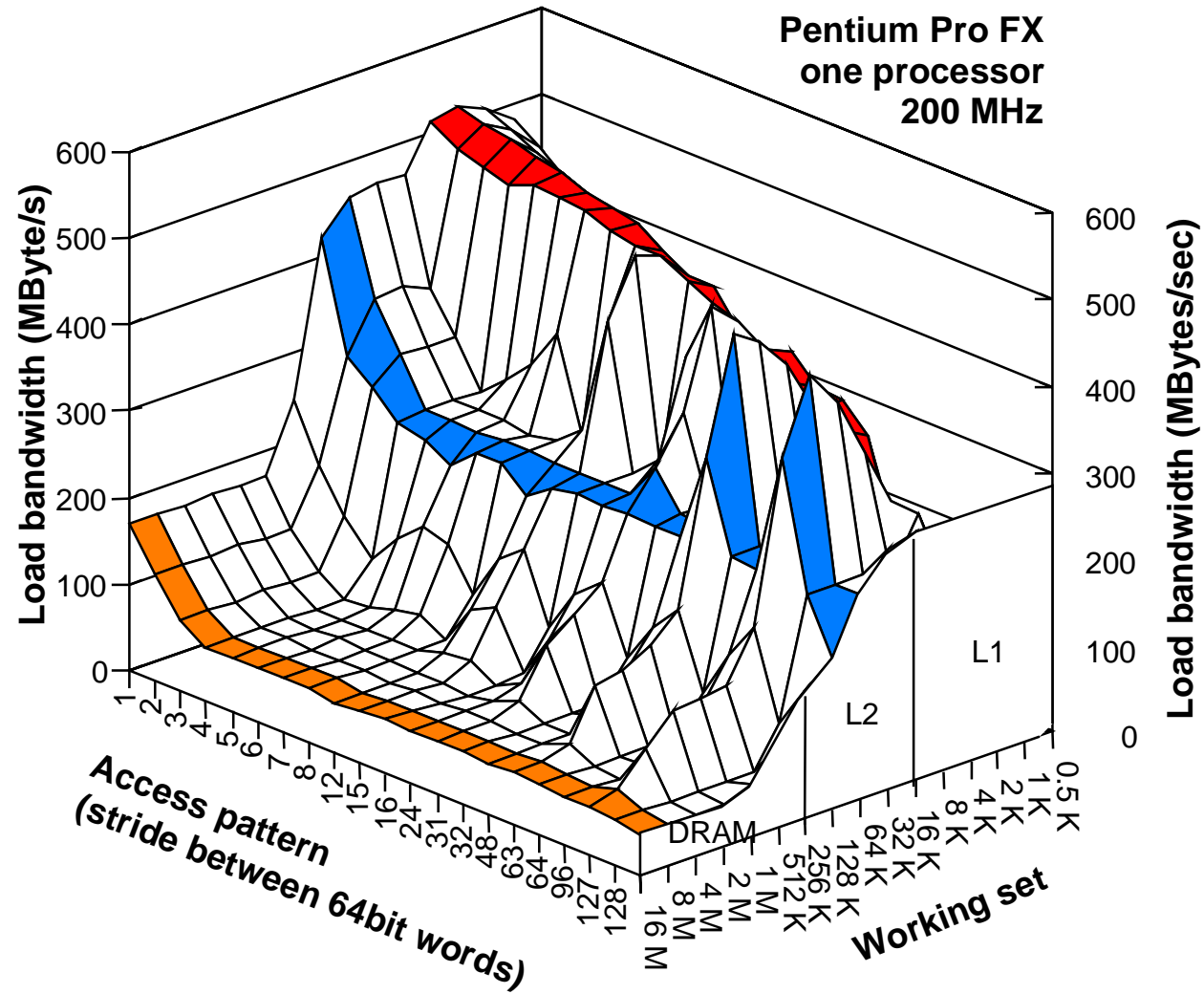
- ◆ *Local* and *Remote* transfers
- ◆ compute and communicate accesses

Measurement Problems

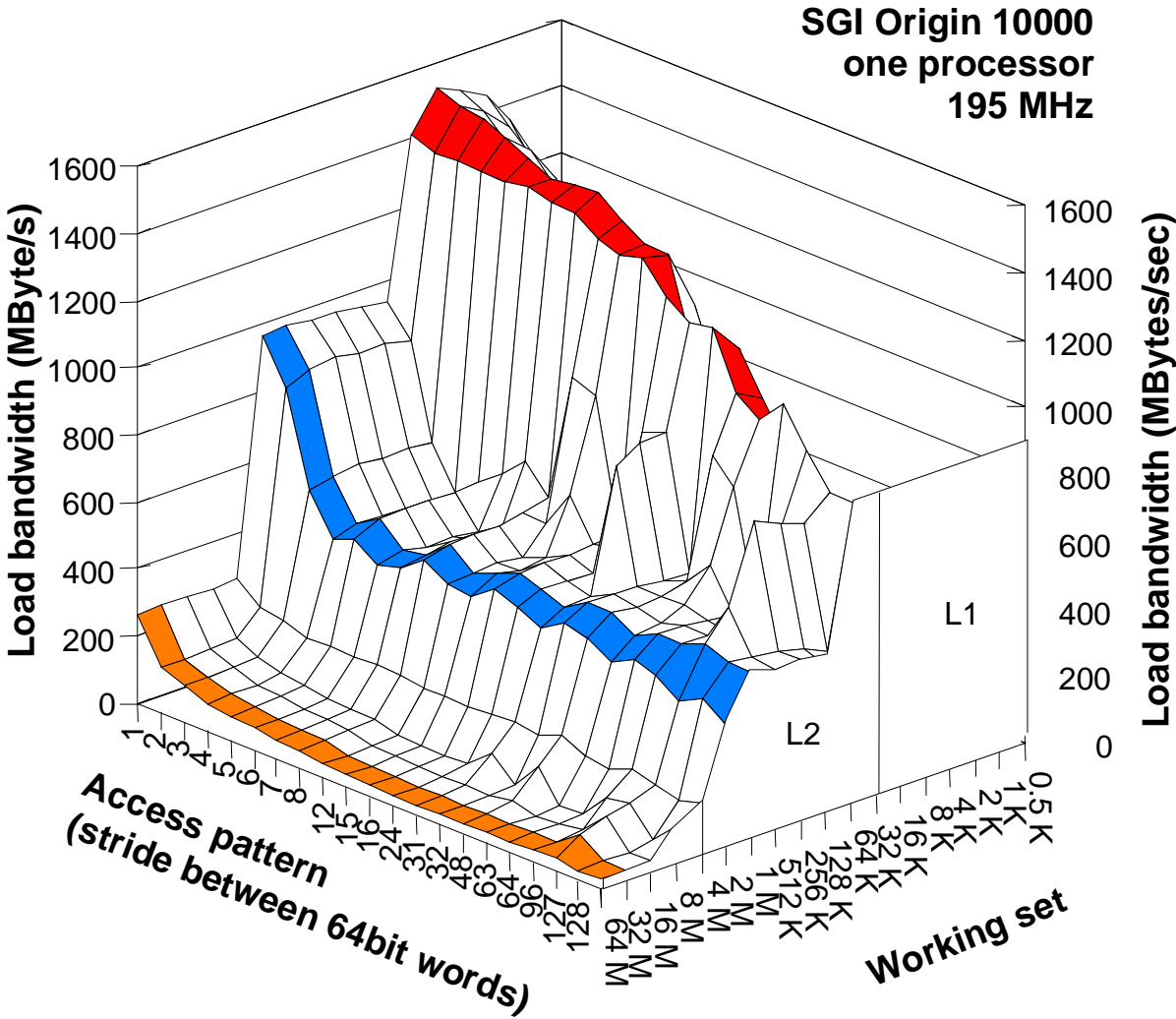
Some parameter combinations are hard to measure, even with carefully tuned C code:

- ◆ Reduced performance for *large strides* and *small working-sets in L1 caches* is a measurement artifact and not architecture related.
- ◆ Compilers occasionally generate suboptimal instruction schedules for loads / stores.

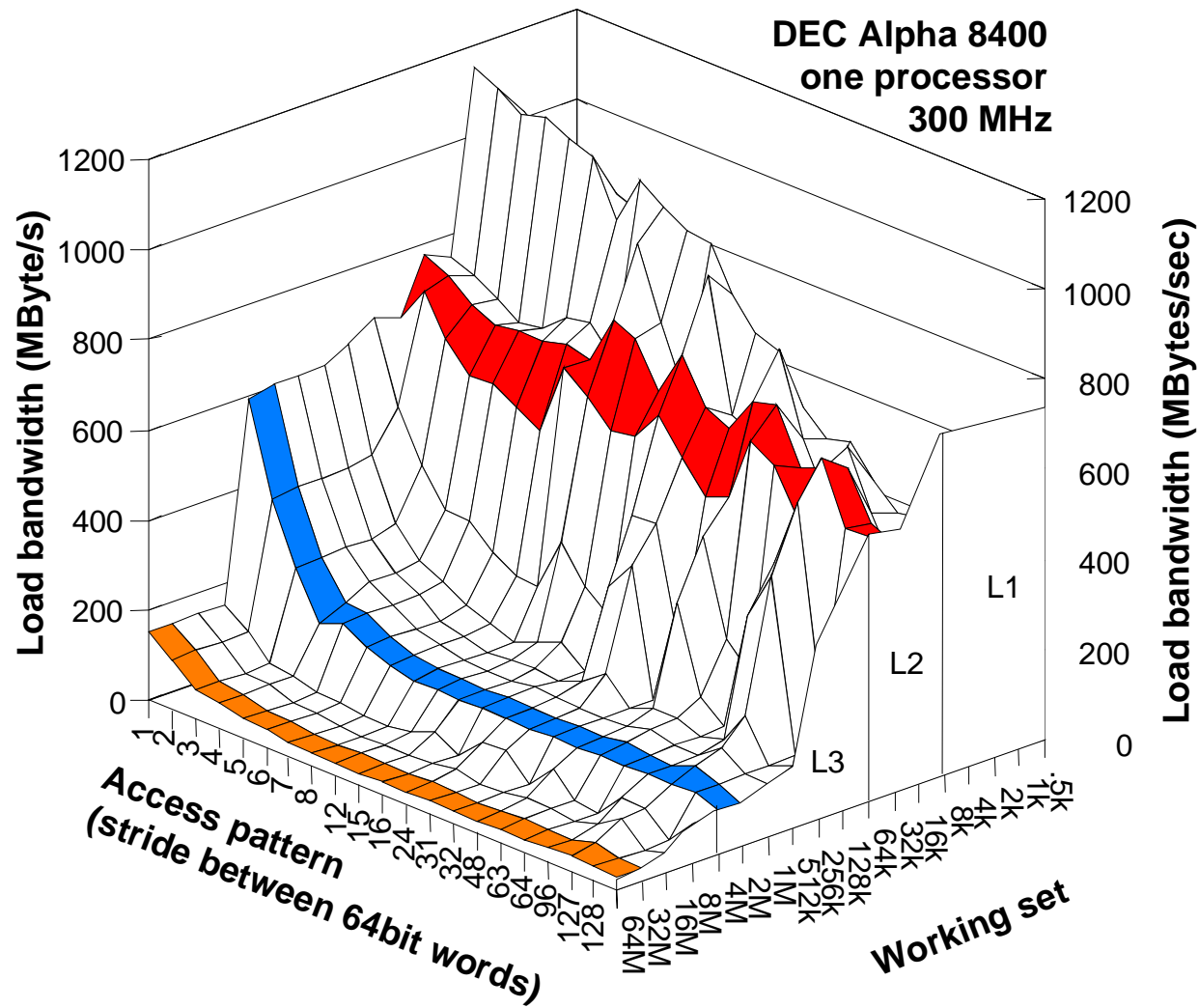
Local Load Access: Pentium Pro PC



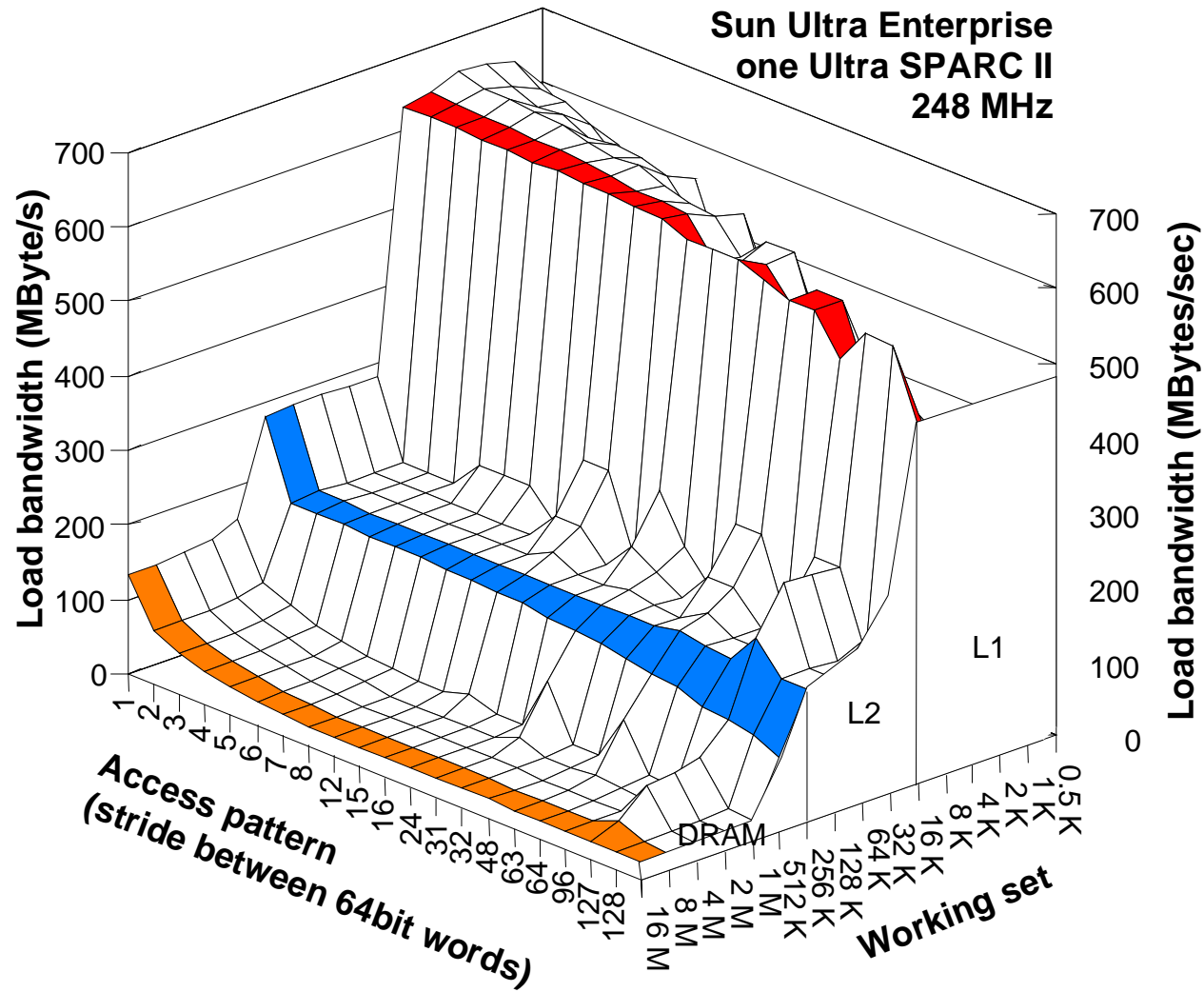
Local Load Access: SGI Origin



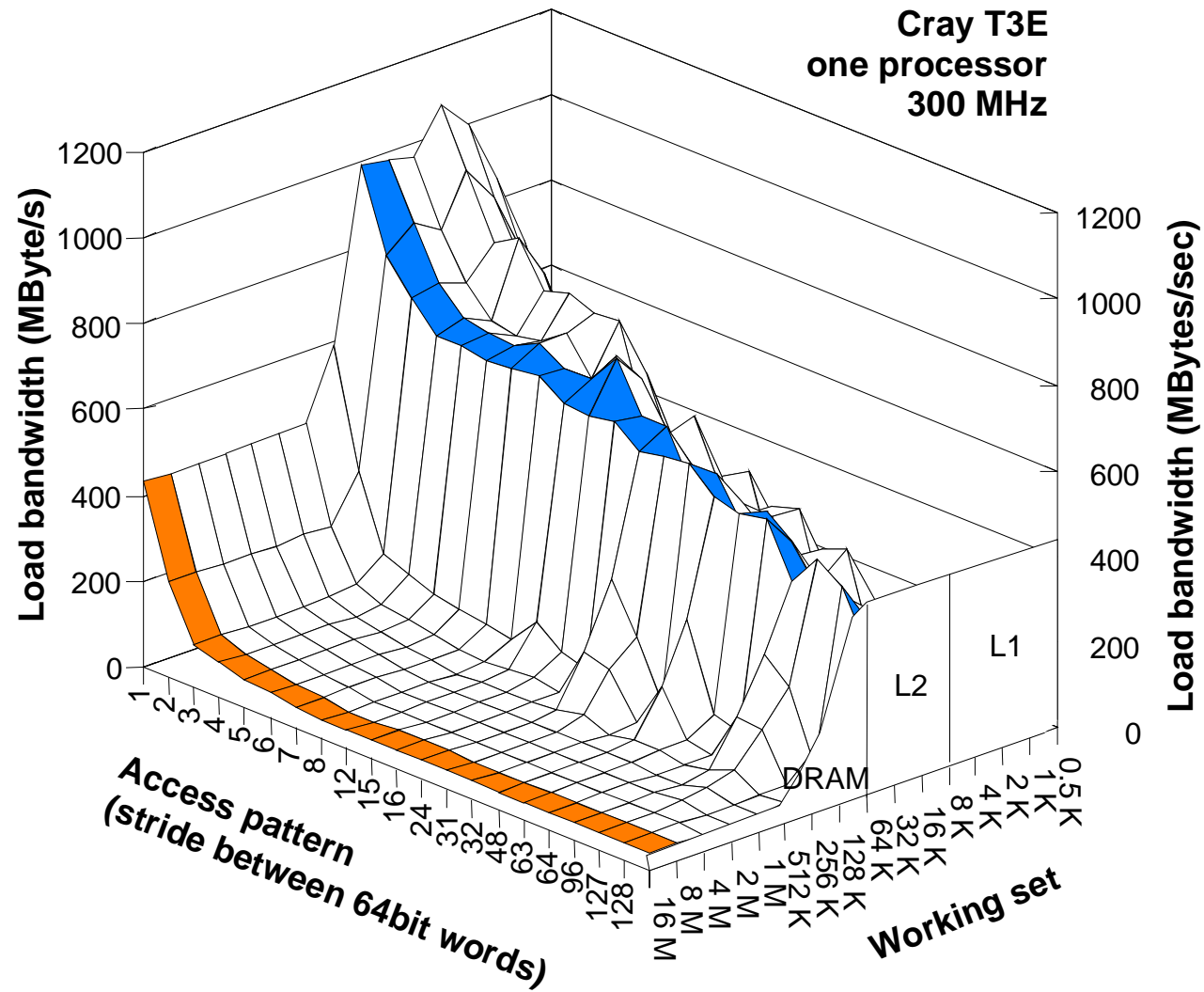
Local Load Access: DEC 8400



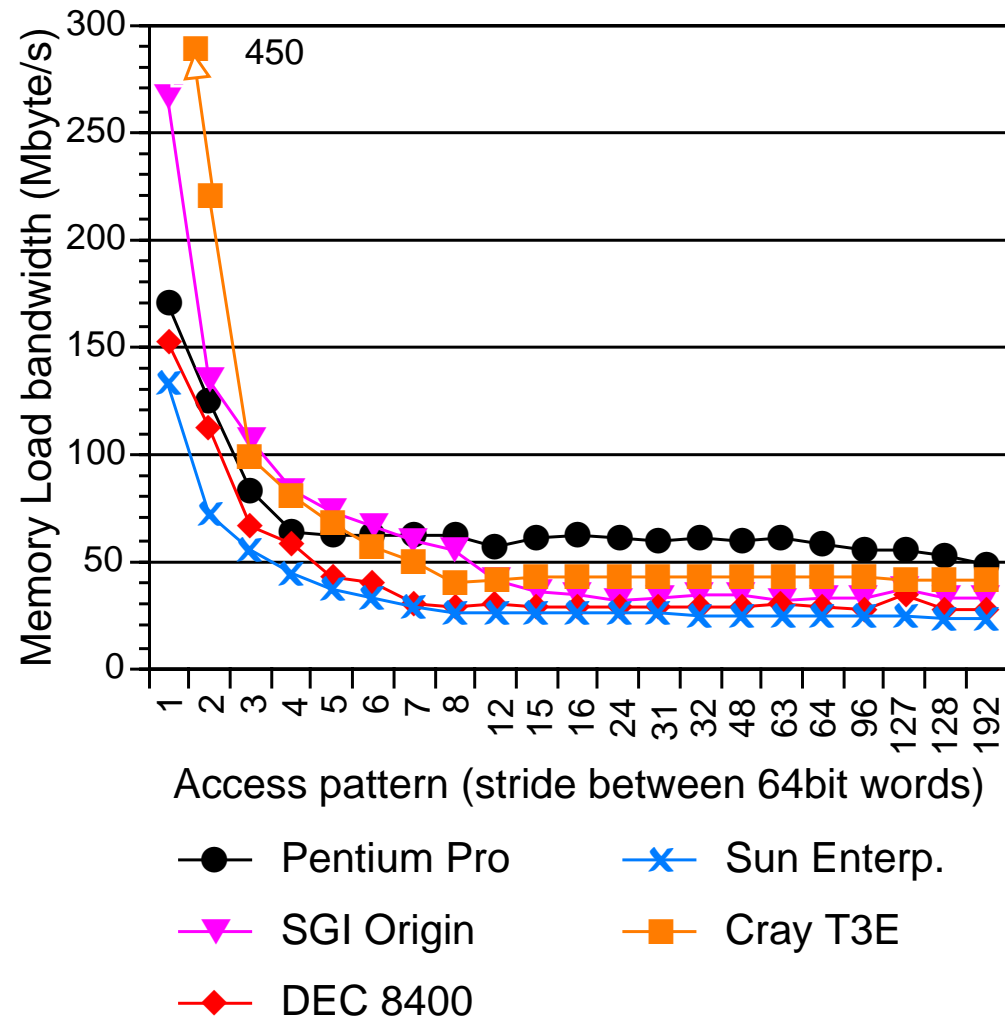
Local Load Access: Sun Enterprise



Local Load Access: SGI Cray T3E



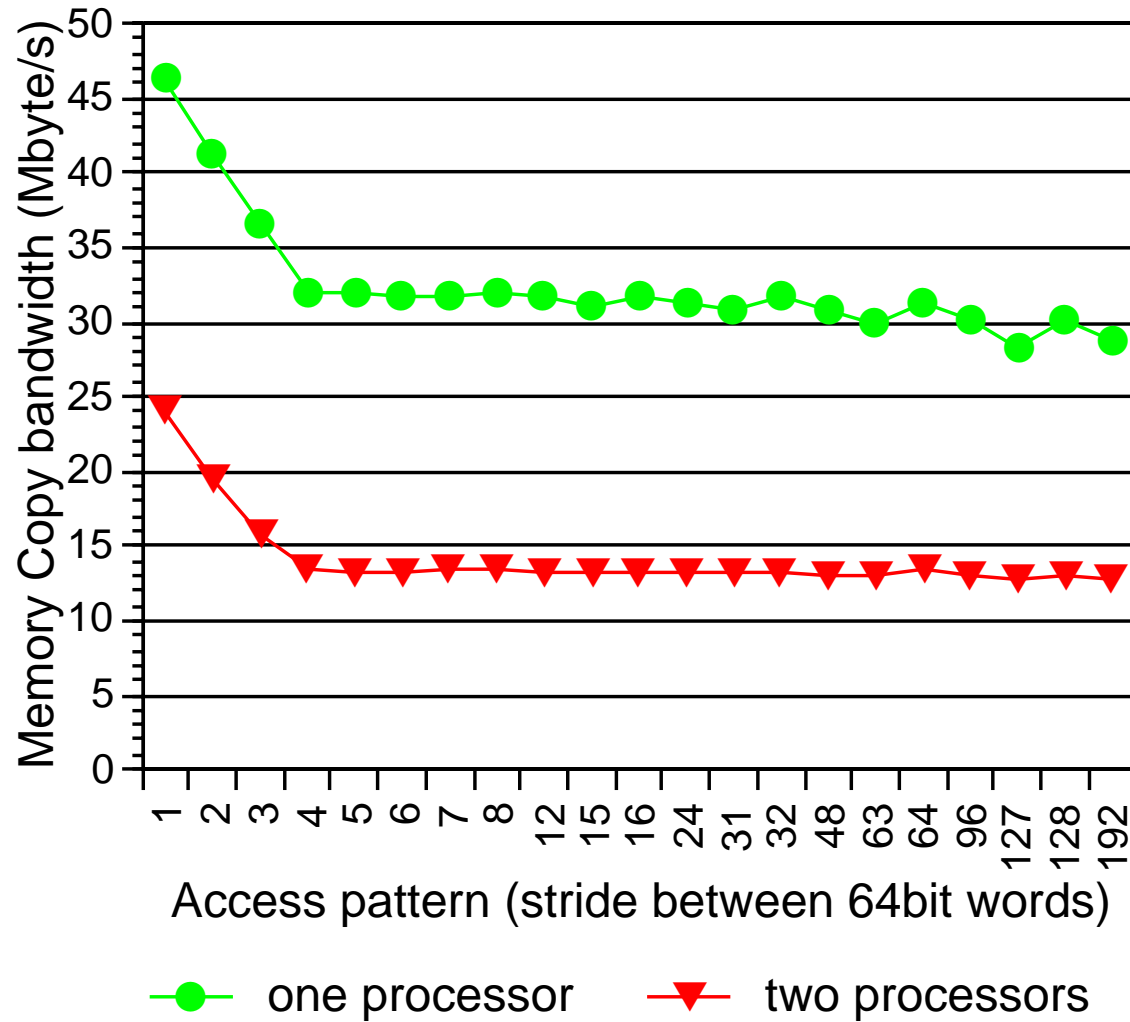
Comparison - Local Access



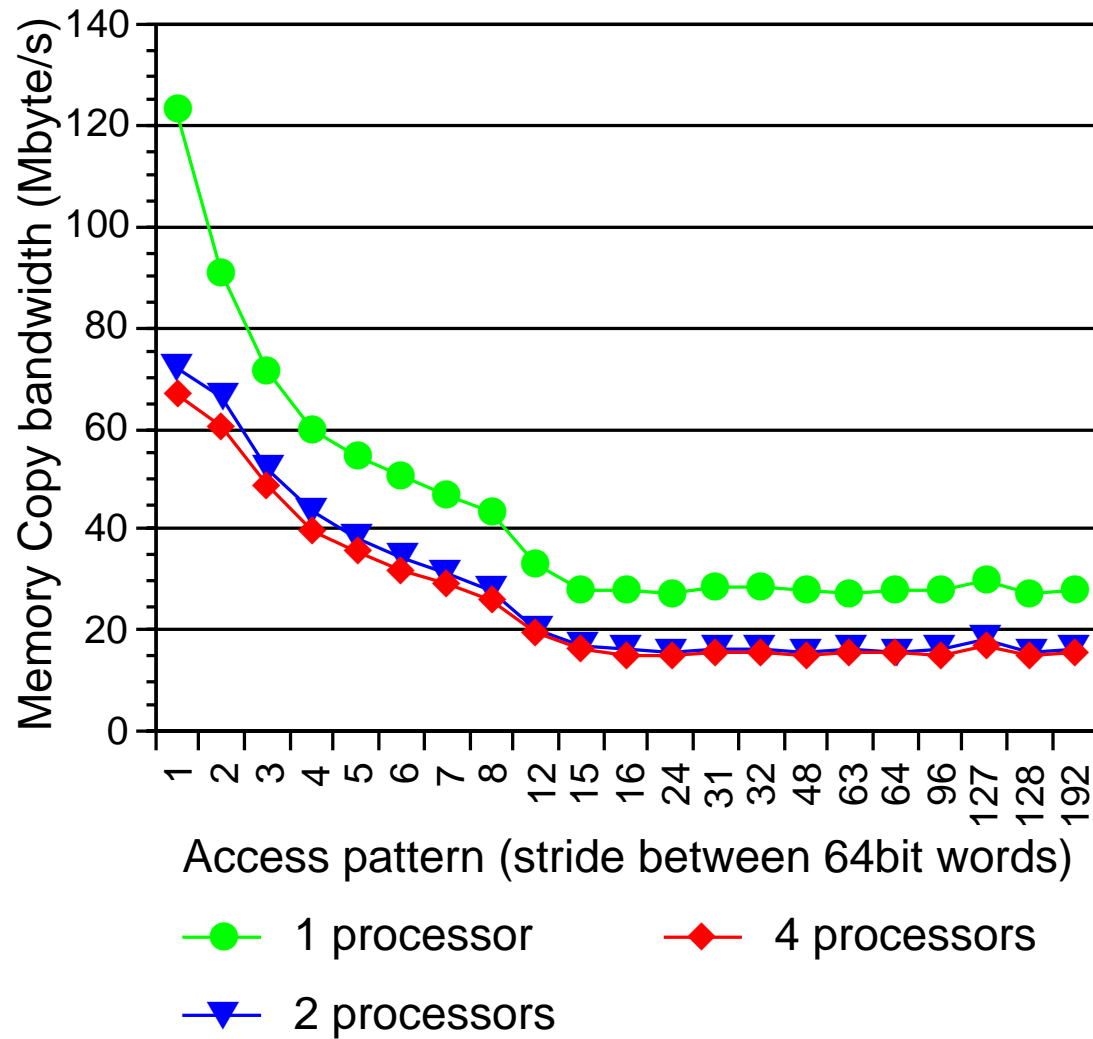
Performance in an SMP setting

- Copy bandwidth decreases for simultaneous access with 1, 2, 4 and 8 processors
- Topics of interest:
 - ◆ small working sets in caches: performance remains same
 - ◆ large working sets in memory: interesting differences
 - ◆ behavior for even/uneven strides
- “Gather copy stream”
(strided load / contiguous store)

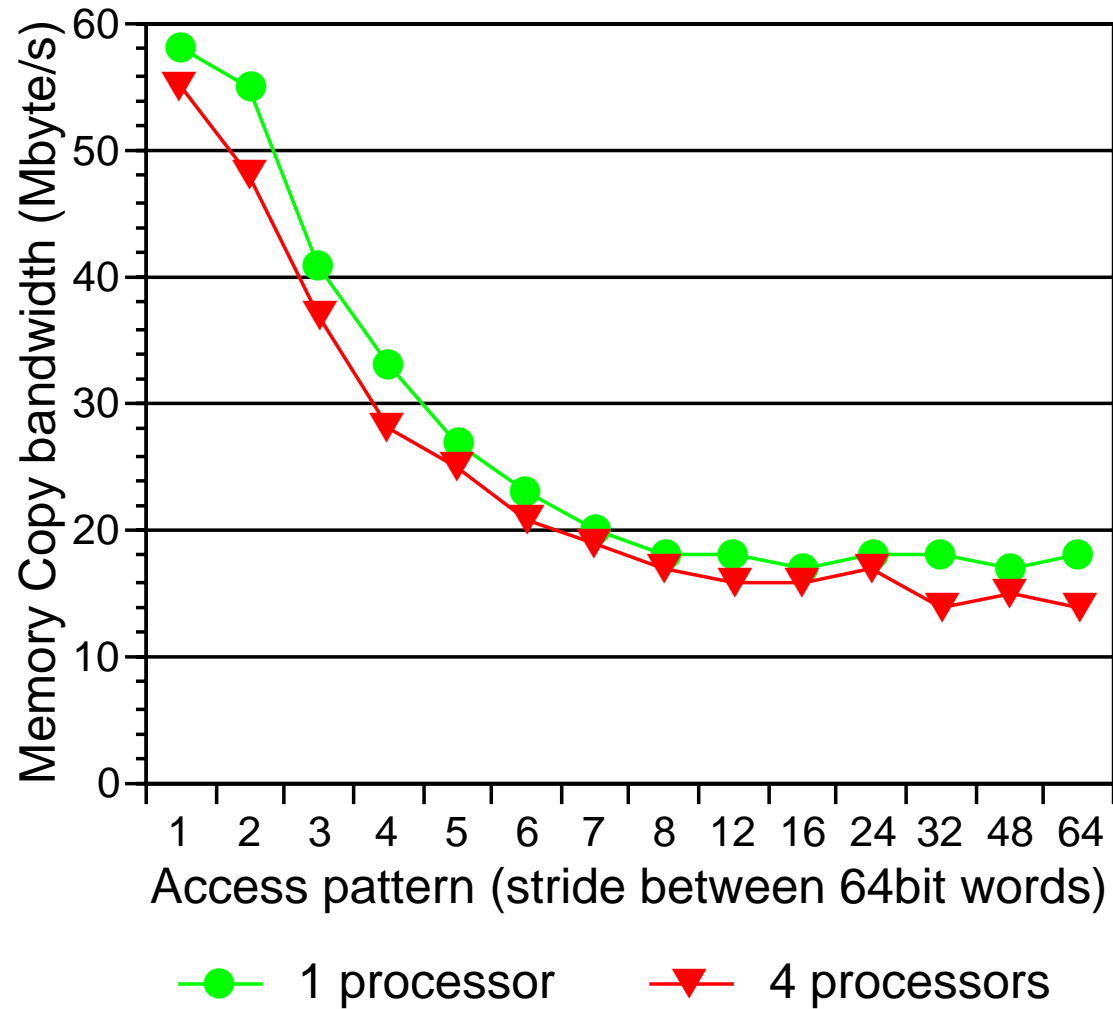
Local Copy: Pentium Pro SMP



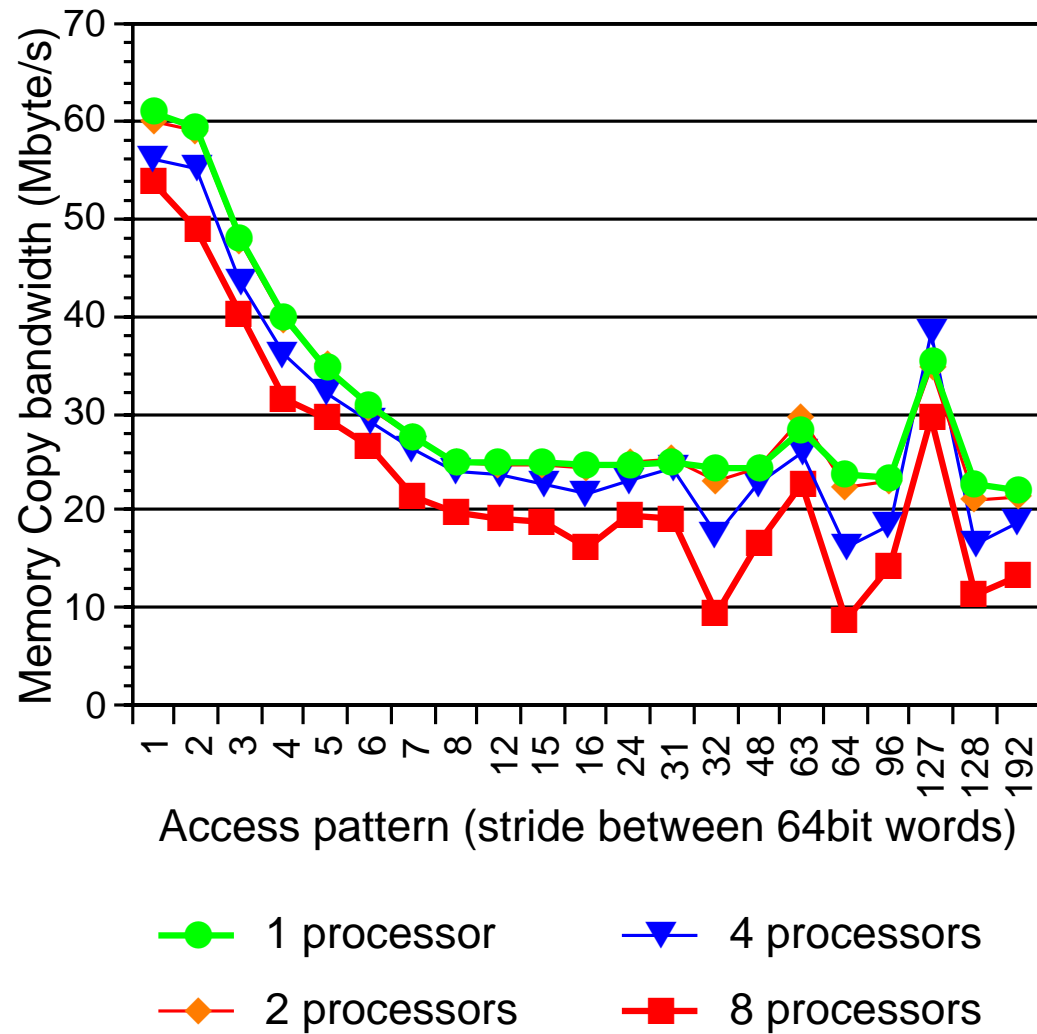
Local Copy: SGI Origin CC-NUMA



Local Copy: DEC 8400 SMP

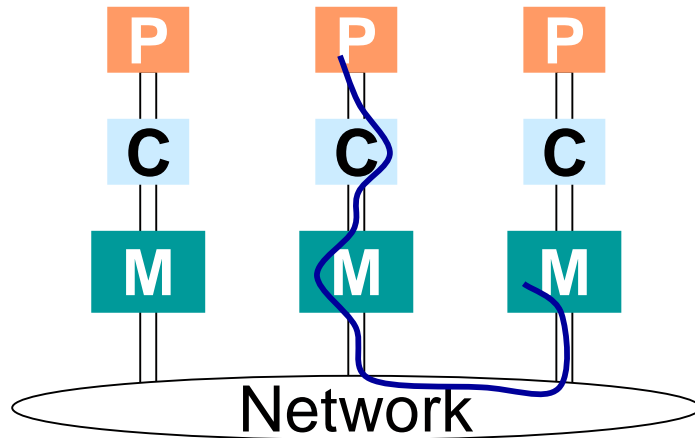


Local Copy: Sun Enterprise SMP



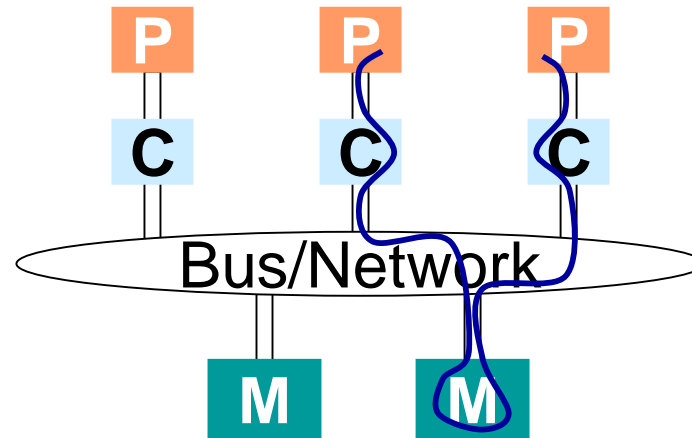
Remote in Parallel Computers

Parallel & Network Computers



SGI Cray T3E, SGI Origin
Clusters of PCs (CoPs)

Symmetric Multiprocessors

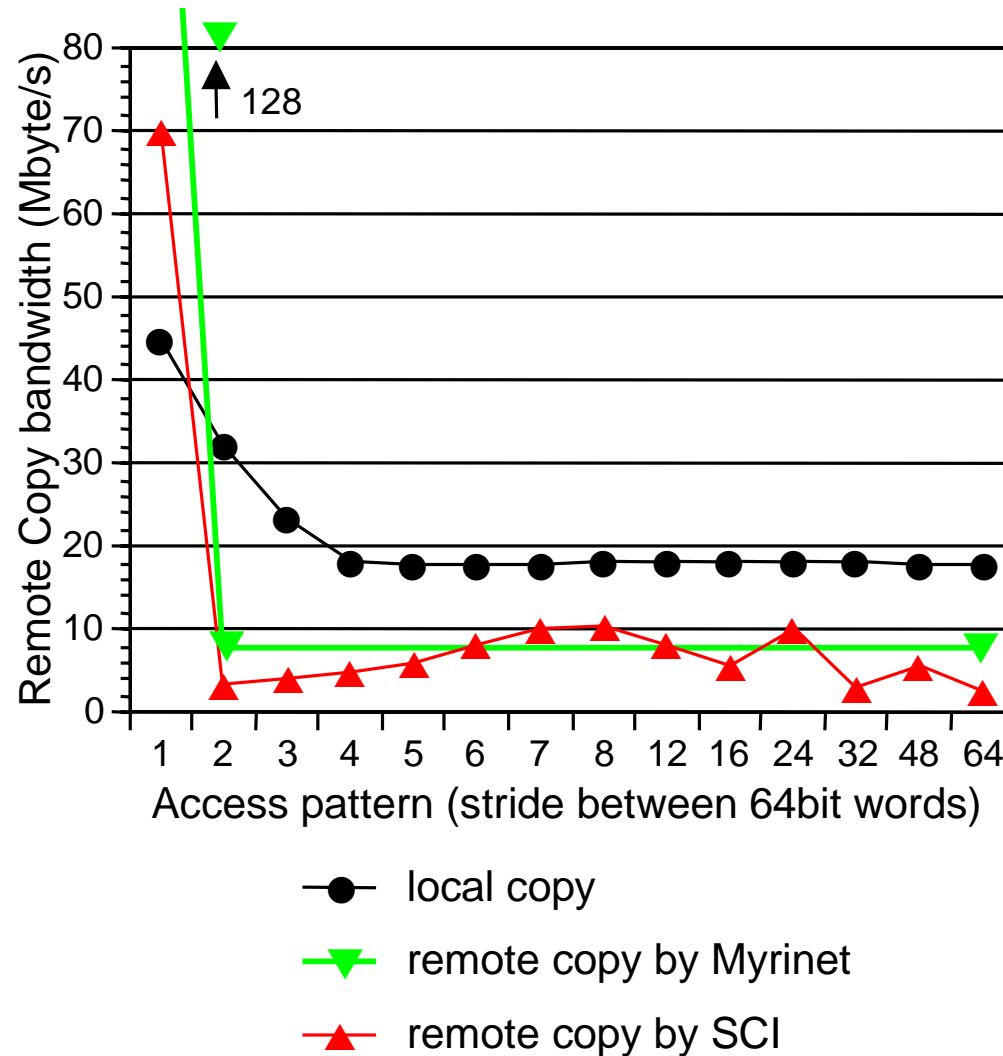


DEC 8400, Sun Enterprise,
Pentium Pro SMPs

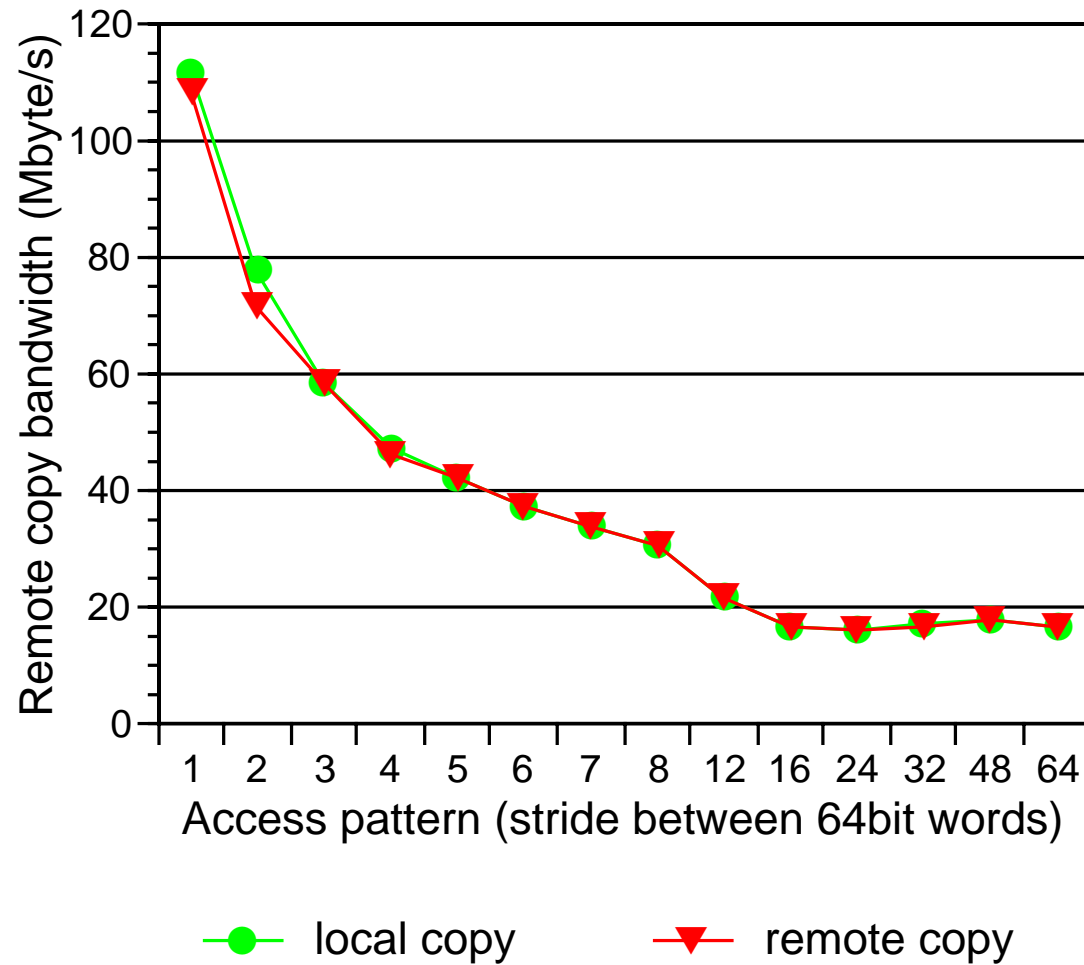
P Processor **C** Caches **M** Memory

Remote Transfers: CoPs

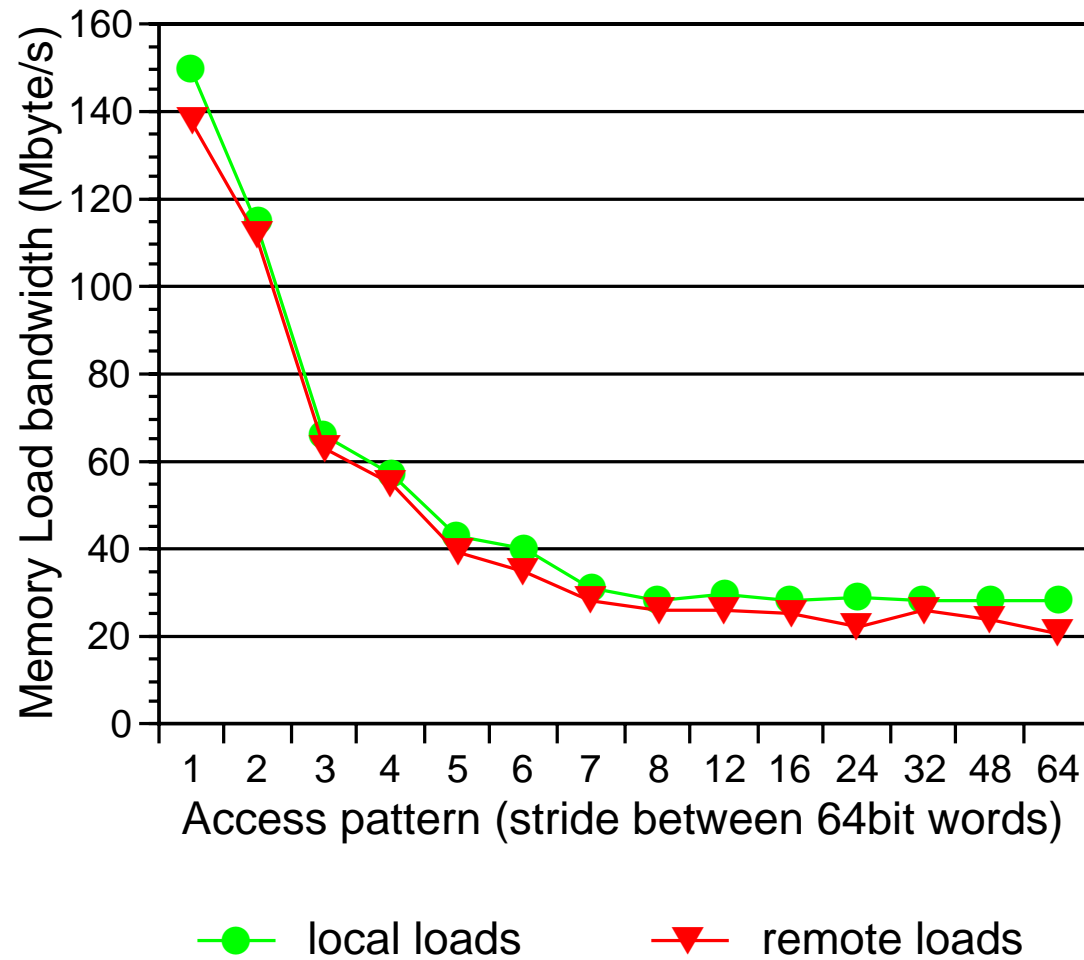
Pentium Pro with SCI / Myrinet



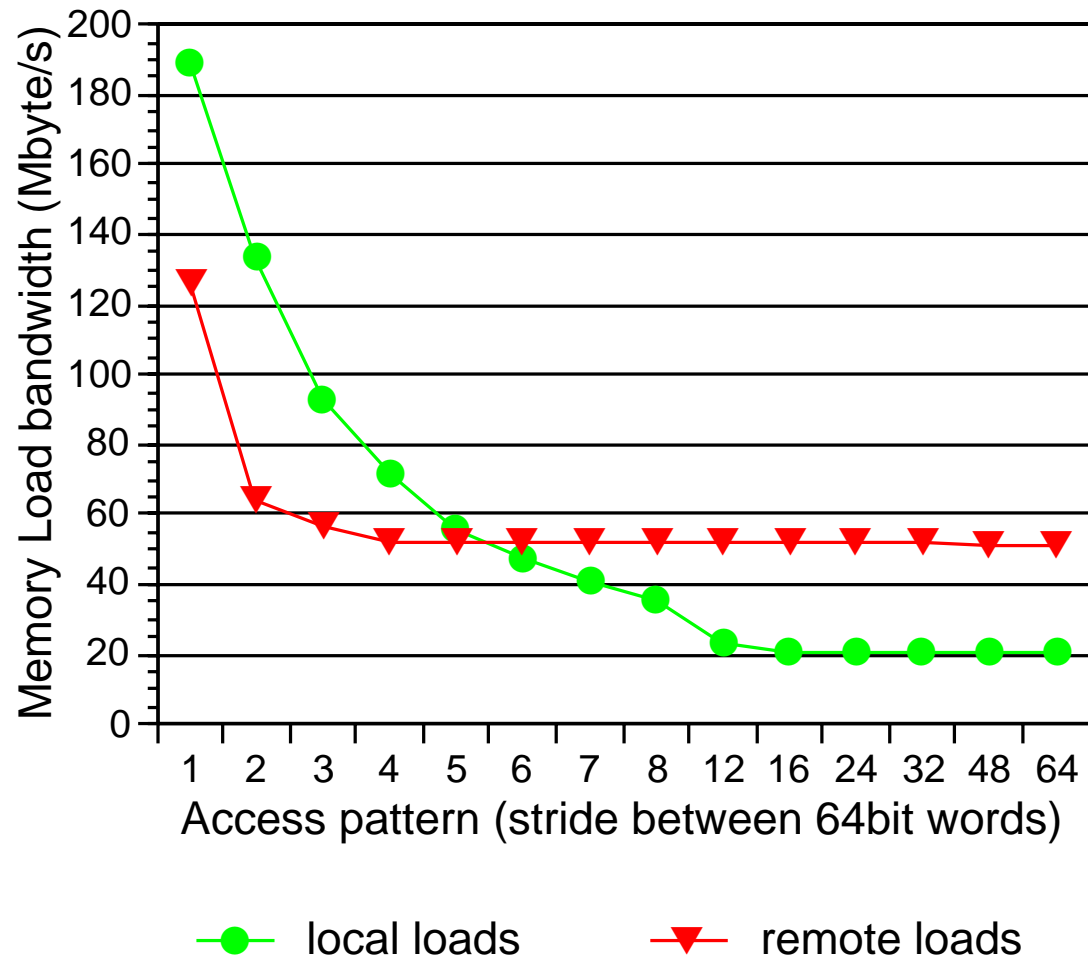
Remote Transfers: SGI Origin



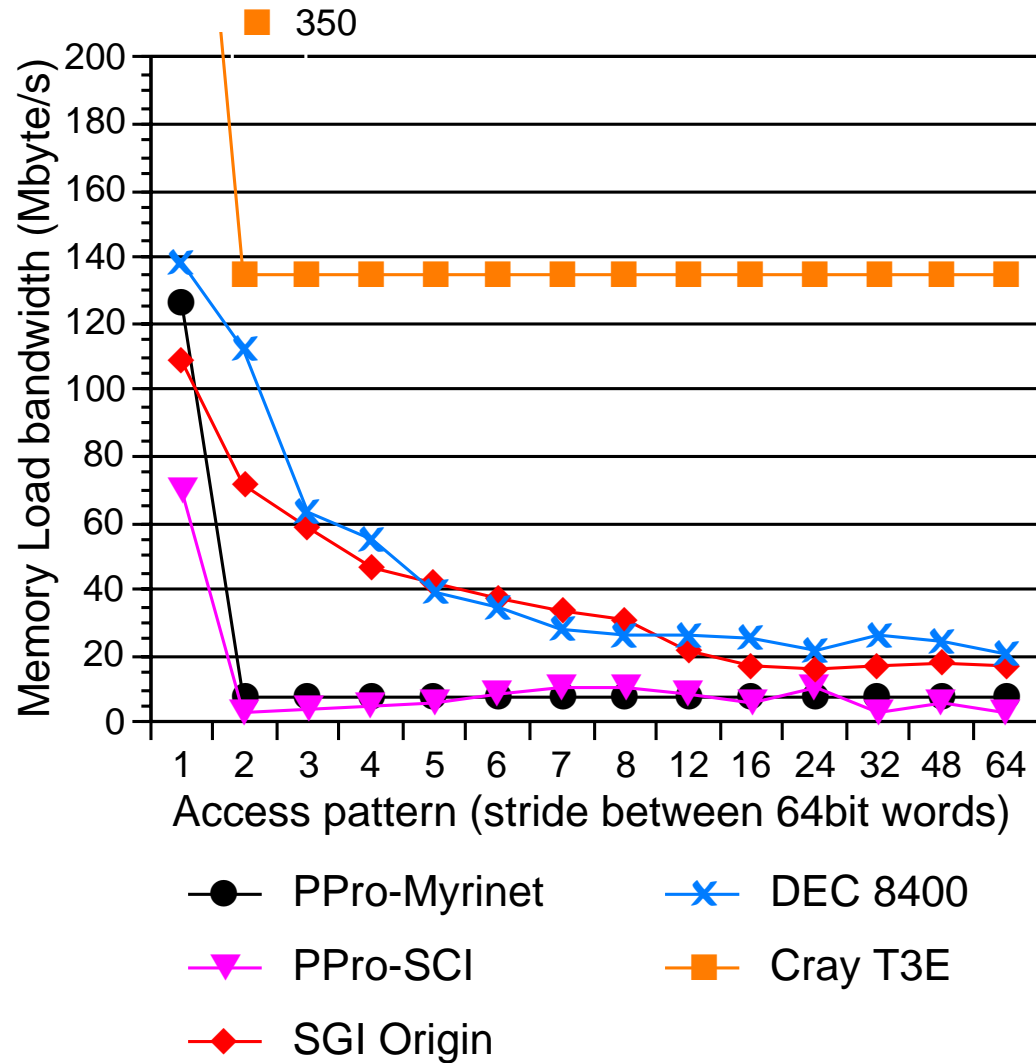
Remote Transfers: DEC 8400



Remote Transfers: SGI Cray T3E

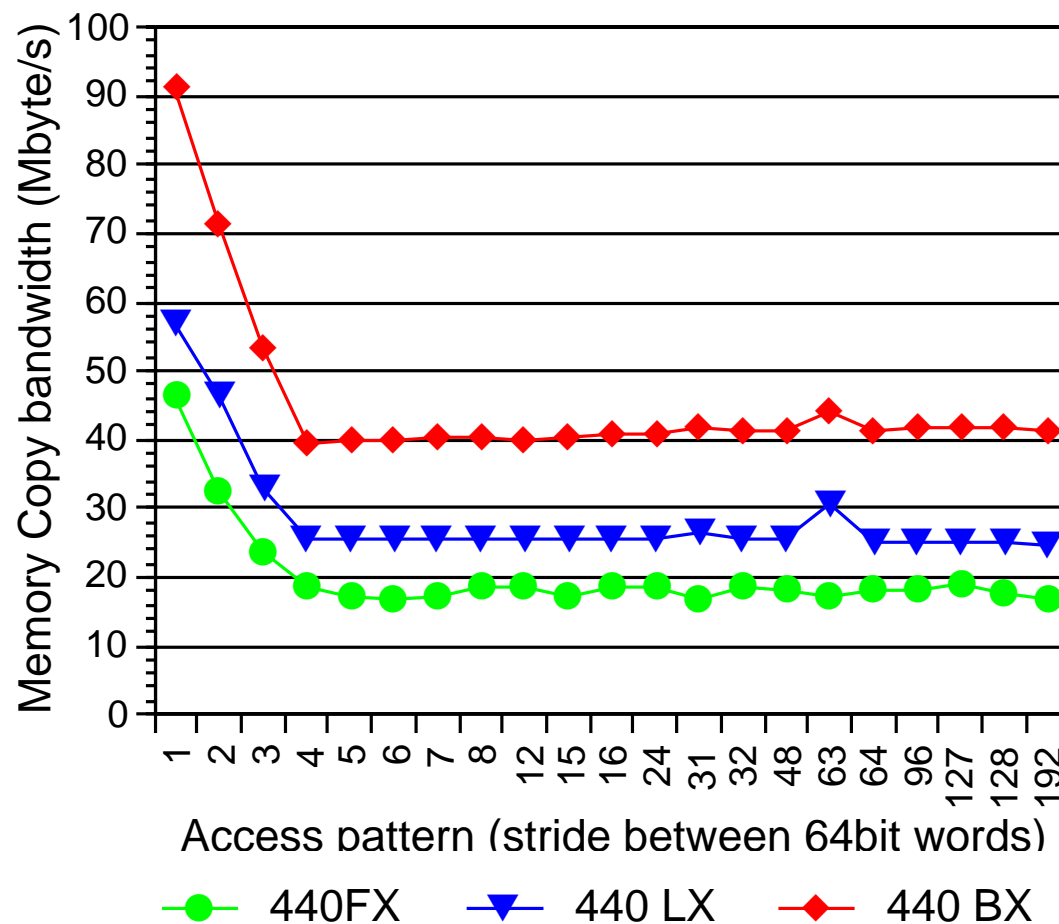


Comparison - Remote Transfers



Improvement of PC Chipsets

- Intel 440 BX
AGP Chip Set
350 MHz / 100 MHz
- Intel 440 LX
AGP Chip Set
233 MHz / 66 MHz
- Intel 440 FX
Natoma Chip Set
200 MHz / 66 MHz



Conclusion

- ECT-Characterizations for different memory systems:
 - ◆ T3E (MMP-Node), Origin (NUMA), DEC8400 (SMP)
 - ◆ CoPs Intel P6 SMPs and Clusters
- High End SMP vs. Low End SMP:
 - ◆ Less than half performance on two processor PCs.
- Fast communication puts high demands on the memory system:
 - ◆ Unlike in traditional SMPs and CC-NUMAs fine grained remote access do not perform at all in PC-SMPs and CoPs
- Adding more commodity microprocessors processors without reinforcing the memory system is therefore questionable.