ServerSet III and MXT Technology

Sectored Memory Management

Control Bit Field	Pointer 0	Pointer 1	Pointer 2	Pointer 3

Physical Address Minimum

Sector Translation Table

CTRL PTR0 ... PTRn

CTRL PTR0 ... PTRn

256B Sector

Sectored Memory Region

256B Sector

Physical Address Maximum

16-Byte
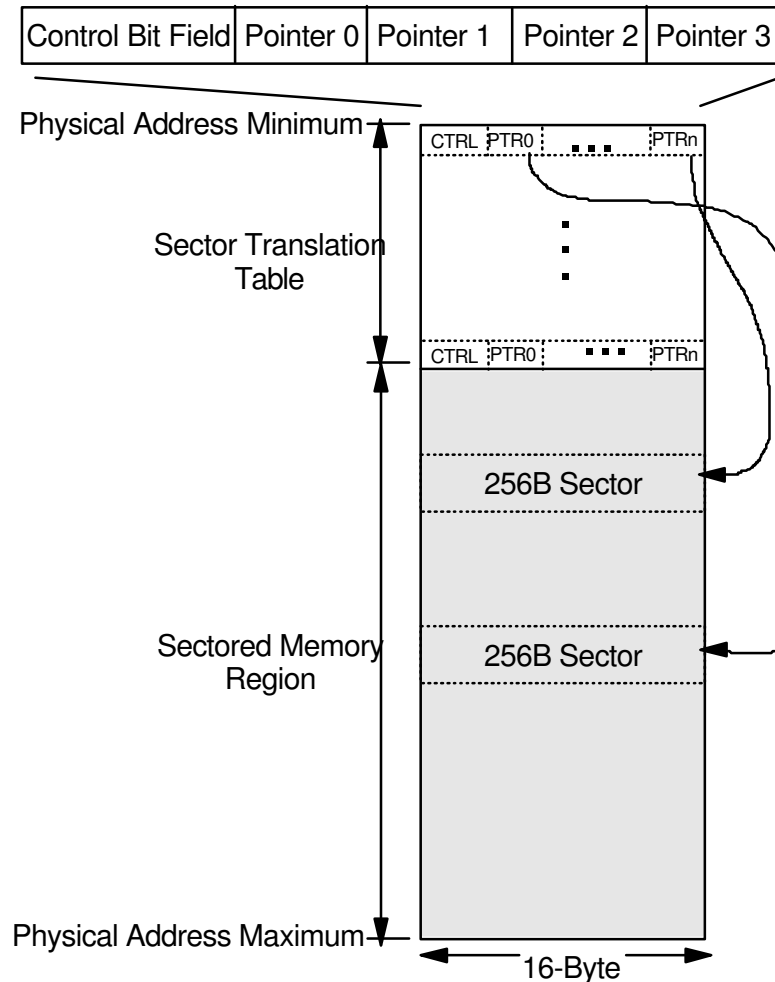
- Indirect "virtual" mapping of main memory
  - Multi-step access to reference data:
    1. Access Translation Table Entry, decode and parse sector addresses
    2. Access data from sector
    3. Update Translation Table Entry if changes
- Sector Translation Table
  - Direct mapped16B entry, per 1KB real memory block (size = real memory / 1KB * 16)
  - Uncompressed entries encoded with flags and up to 4 address pointers to 256B sectors for storing data, associated with a 1KB block, but less than 121bits
  - Compressed entries encoded to contain actual compressed data when less than 121 bits (64:1 compression)
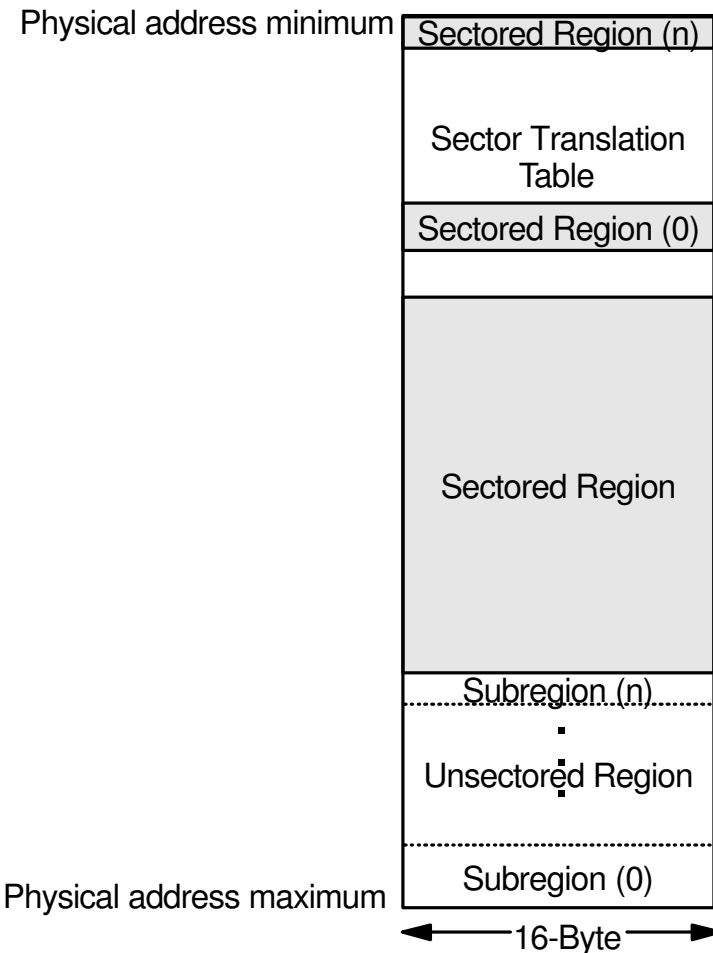- Sectored Region
  - Memory contains a "sea" of 256B sector stored as a linked list.
  - List address pointers exist in unused sectors (i.e., as sectors are consumed, the linked list shrinks)
  - Sectors may be shared by adjacent block data to reduce fragmentation.

IBM

ServerWorks

ServerSet III and MXT Technology
Unsectored Memory Management

Physical address minimum

| |
|---|
| Sectored Region (n) |
| Sector Translation Table |
| Sectored Region (0) |
| |
| Sectored Region |
| Subregion (n) |
| Unsectored Region |
| Subregion (0) |

Physical address maximum

◄——16-Byte——►

- **Direct mapping of main memory**
  - Requires one-step access to reference data:
    1. Access data from relocated direct mapped data address.
  - Lowest latency access to performance sensitive data structures.

- **Sector Translation Table**
  - Unused "holes" in table are redefined as sectored memory to eliminate wasted space.

- **Un-Sectored Region**
  - Selectable as 0-4 independent regions (32KB - 256MB) begining at 32KB aligned addresses.
  - Contain only uncompressed data.

IBM

ServerWorks

ServerSet III and MXT Technology
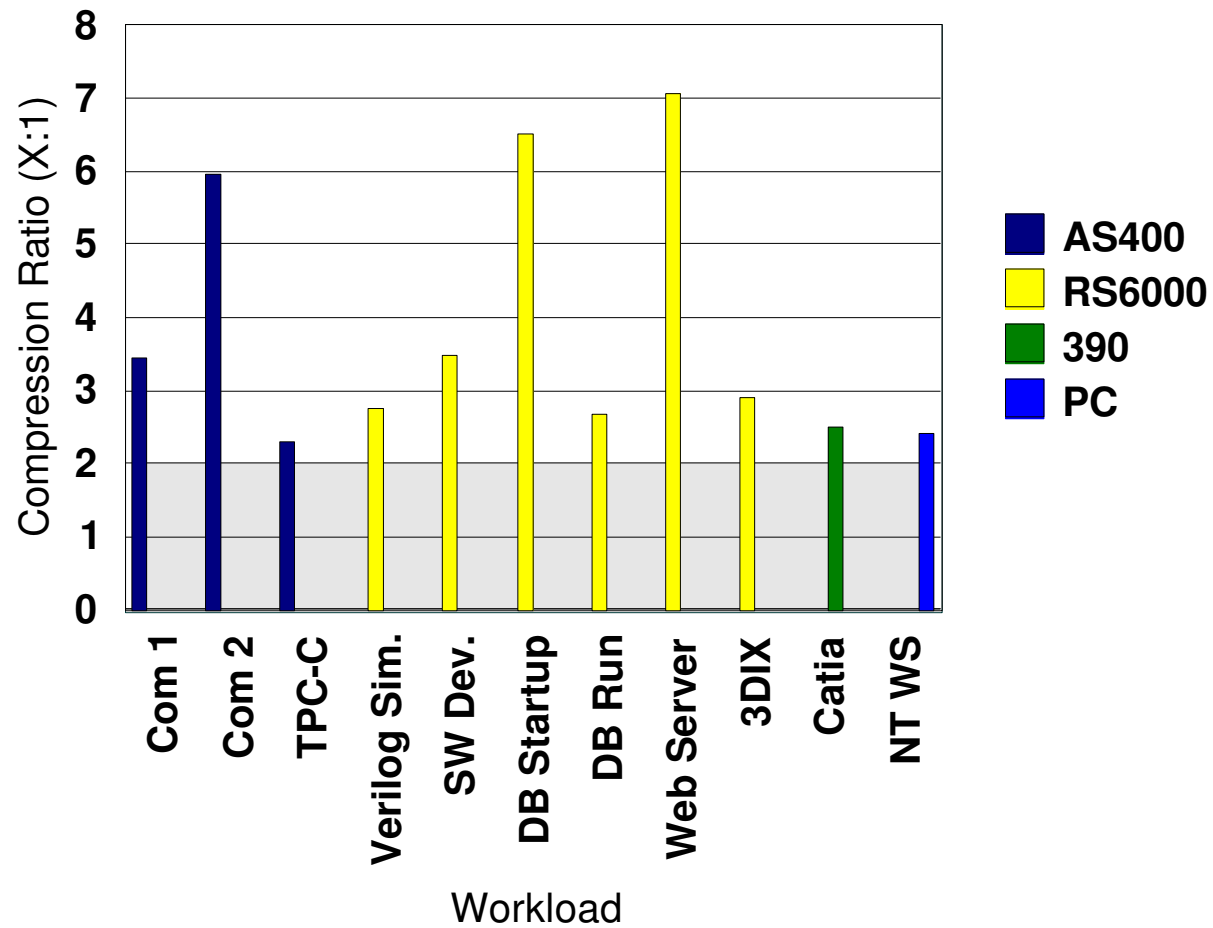
Compression Technology
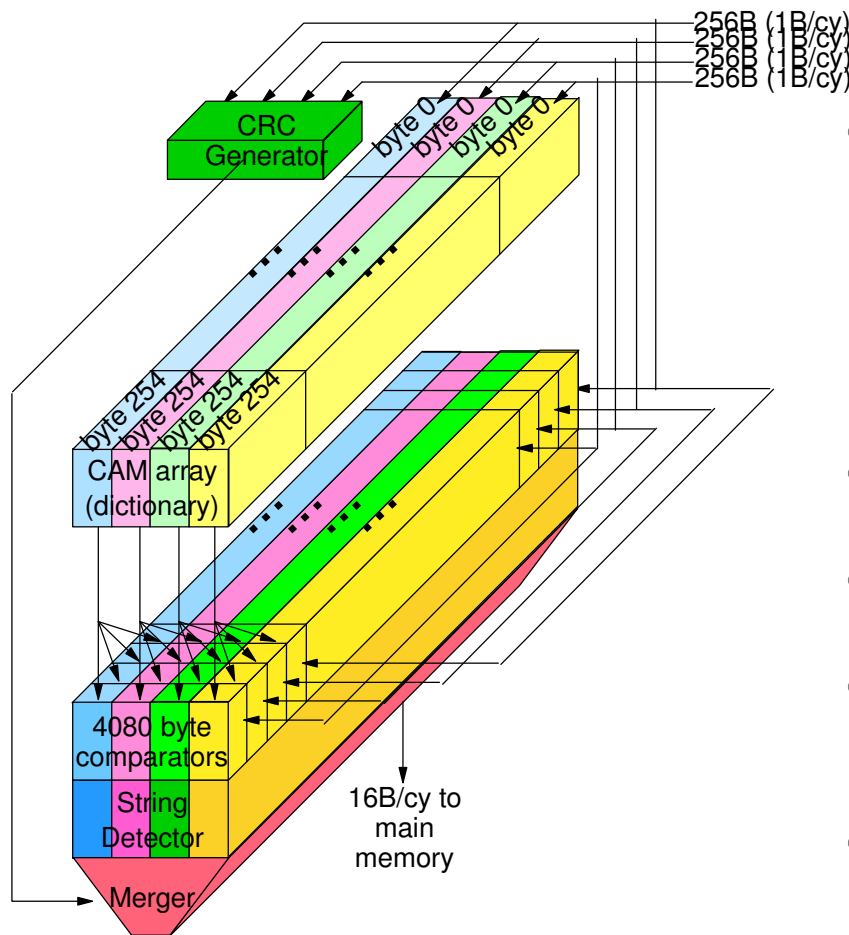
- Variation of Ziv-Lempel algorithm (LZ77 or LZ1) modified for parallel single pass loss-less compression.

- Compresses replicate strings of 2 or more bytes or "raw" characters within 1KB block yielding 1:1 - 64:1 compression ratio.
  - Raw character = (0, databyte)
  - String = (1, primary len, position, secondary len)
  - Encoded length (2-12 bit) and position (2-10-bit)
  - CRC-32 protected compressed block

- Each of four engines compress one byte/cycle of a 1/4 block (256B).

- Four dictionaries are incrementally filled and compared to new byte every cycle.

- Emerging strings may occur within any dictionary and are tracked for the longest occurance.

- Strings are encoded and packed into an output buffer with CRC-32 appended to the end of the buffer when complete.

256B (1B/cy)
256B (1B/cy)
256B (1B/cy)
256B (1B/cy)

CRC Generator

byte 0
byte 0
byte 0
byte 0

byte 254
byte 254
byte 254
byte 254

CAM array (dictionary)

4080 byte comparators

String Detector

Merger

16B/cy to main memory

IBM

ServerWorks
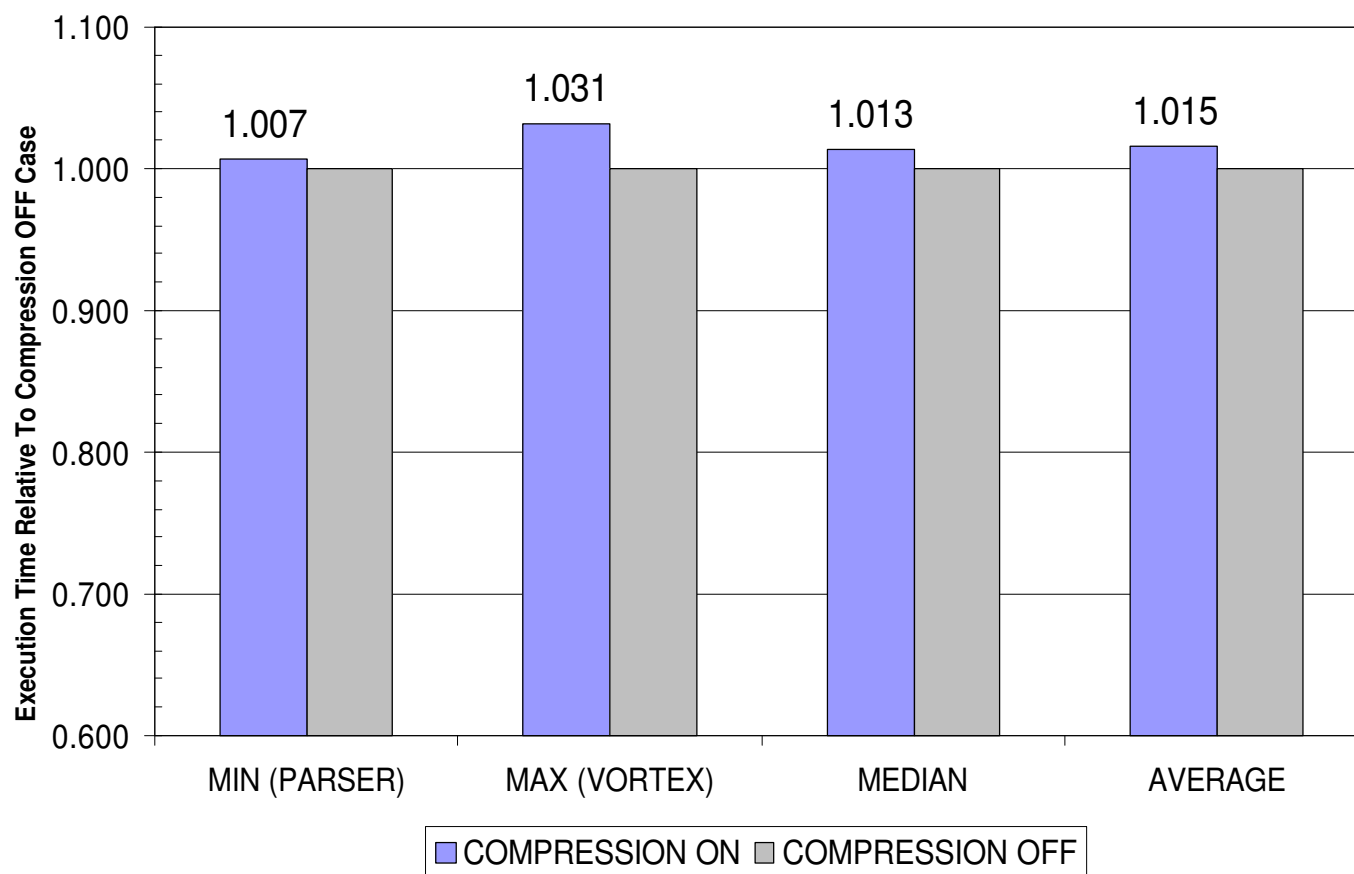
3

ServerSet III and MXT Technology
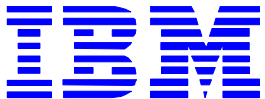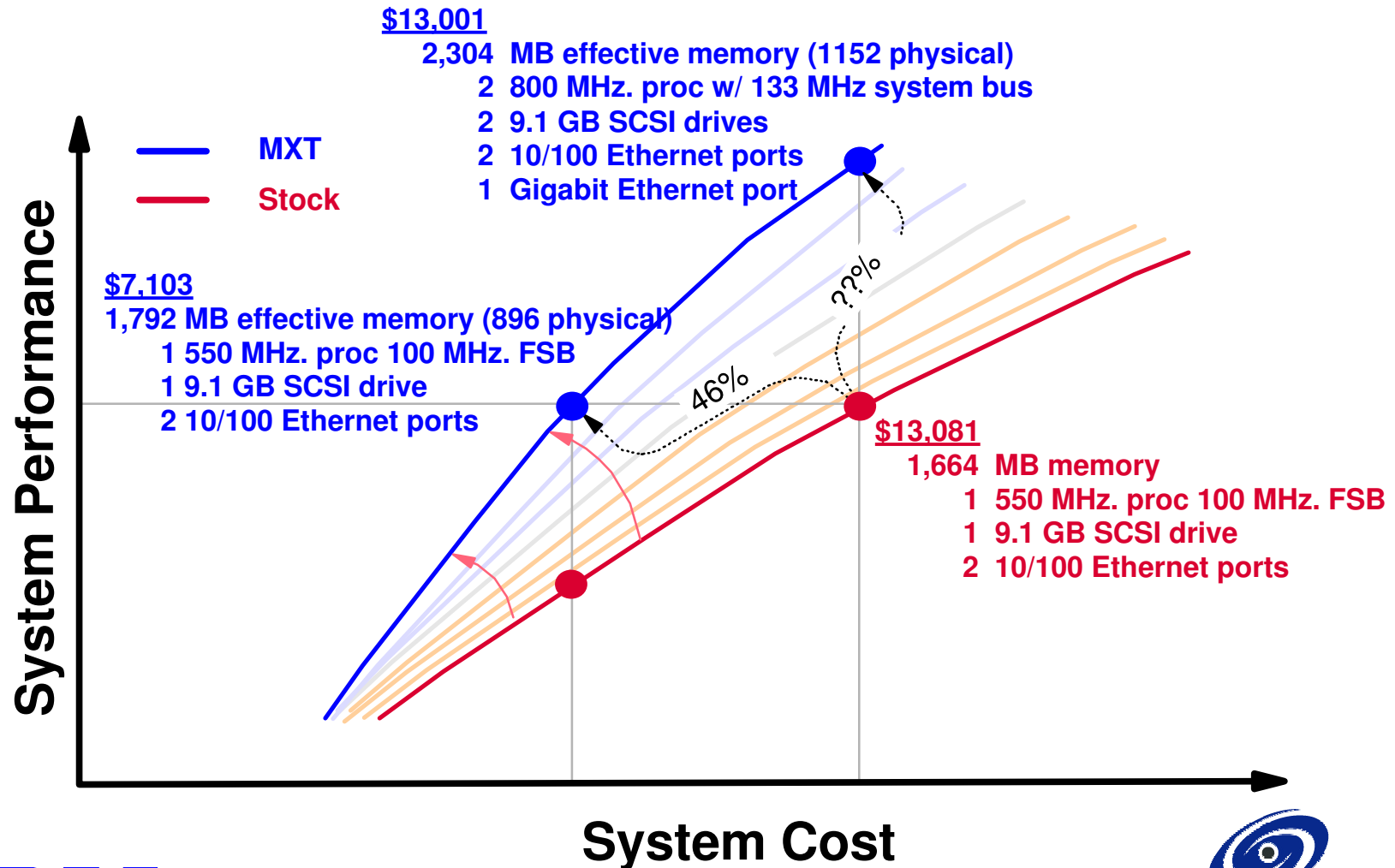## Real-World Compression

ServerSet III and MXT Technology
Performance Overhead for SPEC CINT2000 Benchmarks
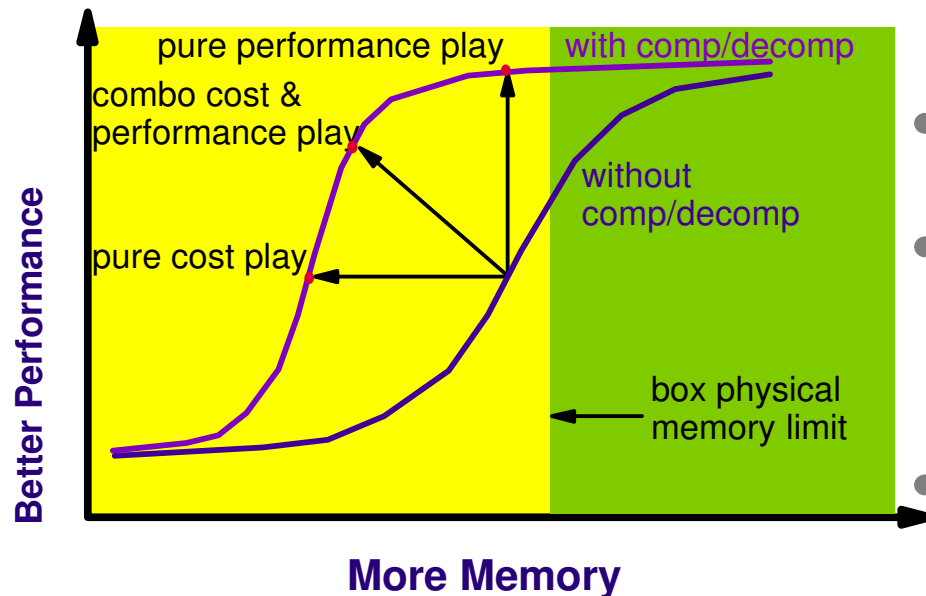### Intel XEON (256K L2), 512MB System

ServerSet III and MXT Technology
System Cost Leverage

$13,001
- 2,304 MB effective memory (1152 physical)
- 2 800 MHz. proc w/ 133 MHz system bus
- 2 9.1 GB SCSI drives
- 2 10/100 Ethernet ports
- 1 Gigabit Ethernet port

MXT

Stock

System Performance

$7,103
- 1,792 MB effective memory (896 physical)
- 1 550 MHz. proc 100 MHz. FSB
- 1 9.1 GB SCSI drive
- 2 10/100 Ethernet ports

??%

46%

$13,081
- 1,664 MB memory
- 1 550 MHz. proc 100 MHz. FSB
- 1 9.1 GB SCSI drive
- 2 10/100 Ethernet ports

System Cost

6

ServerSet III and MXT Technology
Conclusion

- **Better Performance** (y-axis)
- **More Memory** (x-axis)
- pure performance play
- combo cost & performance play
- pure cost play
- with comp/decomp
- without comp/decomp
- box physical memory limit

- Proven technology that empowers consumers to efficiently utililize their memory investment.
- Logical step in the pervasion of ubiquitous compression techniques.
- IT professionals routinely experience $1000's on systems ranging from High density servers to large memory enterprise servers
- Technology well suited to other memory intensive applications too (disk controllers, laptops, etc.)

IBM

ServerWorks