Processors in memory
- Chip architecture of DRAM integration -

Takao Watanabe
ULSI Research Department
Central Research Lab., Hitachi, Ltd.
Outline

- Introduction - history and basics -
- LSI architecture in promising applications
- Future perspectives
- Conclusion
Processor and DRAM integration
- example chip -

0.35 um
4 x 2-Mb DRAM
32-b RISC CPU
SRAM, ROM, etc.
Trends in DRAM and MPU
History of embedded DRAMs

- '90
  - 3D RAM [7]
  - PIPRAM [10]
  - Computational RAM [2]
  - 1.5-V Neuro Chip [3-5]

- '00
  - Configurable Macro [18]
  - FPGA+DRAM [16,19]
  - RISC+DRAM [11]
  - PPRAM [14]
  - 3D CG Chip [8]
  - 3D CG Media Chip [6,12,15]
  - IRAM [13]
  - 8-ns Cycle Macro [23]
  - 56.8-GB/s 3D CG [24]
  - 1-GHz Macro [25]
  - 16-MB Cache [26]

R & D chips selected from major conferences
Application of embedded DRAMs

- Game machine
- Digital camera
- Car navigator
- STB
- PC
- WS
- HDD, DVD
- PDA

Consumer electronics
Information technology
PC and WS
Car electronics
Advantages of embedded DRAM

- Higher data-transfer rate
- Lower I/O power dissipation
- Lower chip count
Separate chips vs. embedded DRAM

<table>
<thead>
<tr>
<th></th>
<th>Separate chips</th>
<th>Embedded DRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRAM</td>
<td><img src="image" alt="Separate DRAM" /></td>
<td><img src="image" alt="Embedded DRAM" /></td>
</tr>
<tr>
<td>Logic</td>
<td><img src="image" alt="Separate Logic" /></td>
<td><img src="image" alt="Embedded Logic" /></td>
</tr>
<tr>
<td>External connection</td>
<td>8-32 I/O, 20-50 pF</td>
<td>128-1024 I/O, &lt;1 pF</td>
</tr>
<tr>
<td>Speed</td>
<td>0.1 – 2 Gbyte/s</td>
<td>1 – 100 Gbyte/s</td>
</tr>
</tbody>
</table>
I/O power vs data transfer rate

Data transfer rate (Mbyte/s)

I/O driver power (mW) (without termination power)

LVTTL (50 MHz) x16
SSTL, HSTL (100 MHz) x16

Off-chip I/O

On-chip I/O CMOS (100 MHz)

3.3 V x512
1.5 V x1024

Takao Watanabe - 10th ORAP Forum / 28th SPEEDUP Workshop - Slide 9
Problems with embedded DRAM

- Process cost and side effects in fabrication
- Time consuming design and test
- Mismatch of operating voltages
- Speed penalty due to long row access and cycle
3D structure of cell capacitor

- problems: complexity & side effects -
Trade-off in process optimization

Feature size = 0.18 μm
Memory area = 50 mm²
Occupancy of memory array = 60 %

- **1-T DRAM**
  - Small cell with SAC
  - Poly Si gate transistor

- **6-T SRAM**

- **DRAM-based process**
  - Cell without SAC
  - Siliside gate transistor

**Memory capacity (Mbit)** vs. **Memory cell area (μm²)**
W/poly-Si stacked gate transistor
- small cell and high-speed transistor -
Design and test methodology

Module-based synthesis

Built-in test scheme
Mismatch of operating voltages

Operating voltage: DRAM > Logic

Thick-oxide transistors are used for DRAM
Long row access and cycle time

Clock: Actv., Read, Pre.

Command

Address Row

Column

Data out

Row access time tRAC

Row cycle time tRC
Architectures in promising applications

- Processor with DRAM cache
- Pixel processor with frame buffer for CG
- System on a chip for consumer products
- Special-purpose processors with DRAM
DRAM cache

Configuration example

Examples:
1. I. Naritake, et al., pp. 420-421, ISSCC1999 (ref. [22]).
2. O. Takanashi, et al., pp. 396-397, ISSCC2000 (ref. [25]).
3. M. Nakayama, et al., pp. 398-399, ISSCC2000 (ref. [26]).
Issue: first-access bottleneck

Line size: 128 bits, tRCD: 3 clocks, Column latency: 2 clocks

Cache-fill latency

Column access
First-row access

Number of I/O lines
16 32 64 128
Circuit scheme for fast row access

1T cell

Bit-line pair

Word line

Sense amp.

Common I/O

Direct sensing (I/O separation)

‘Y’ switch can be activated before sense amplifier is turned on.

Takao Watanabe - 10th ORAP Forum / 28th SPEEDUP Workshop - Slide 20
Example of a DRAM cache

- Separate chip example -

0.2-um, 6-metal process
14.52 x 19.3 mm
128-Mb DRAM
tRAS = 9.0 ns, tCAS = 7.2 ns
800 kG (SOG)
VDD = 1.8 V (VDDQ = 1.5 V)
560 MHz, 35.8 Gbyte/s

M. Nakayama, et al., ISSCC2000 (ref. [26]).
Pixel processor with DRAM for CG

One-chip integration

Geometry processor → Pixel processor → Frame memory → Display

Floating-point calculation → Drawing of pixels

Examples:
1. S. Miyano, et al., ISSCC95 (ref. [7]).
2. K. Inoue, et al., ISSCC95 (ref. [8]).
3. T. Watanabe, et al., 1996 Symp. on VLSI circuits (ref. [12]).
Issue: row-address changes

DRAM array

Pixel configuration (simplified)
Architecture for high data transfer rate

- pipelined multi-bank access -
Example: 3-D CG ‘Media Chip’

- 0.4-um process
- 8.35 x 14.6 mm
- 8-Mb DRAM
- 4-pixel processors
- 100 MHz, 512 I/O
- 6.4 Gbyte/s, 800 Mpixel/s
- 3.3 V, 1.5 W

Problems facing system LSIs

Access conflict between CPU and other DRAM masters degrades performance.
Access control for system on chip

- smart access control using sense-amplifier caches -

System-LSI chip

Access optimizer

CPU
Cache

DSP

IP
User's logic

DRAM macro

DRAM macro

Multi-port interface

Bank0
SA cache

Bank1
SA cache

Bank n
SA cache

Main amp.

I/O lines
Inter-bank, non-blocking access

- to avoid access conflicts between DRAM masters -
Effects of ‘Access optimizer’

![Graph showing the effects of access optimizer on CPU cycles. The graph plots the total CPU cycles against the access ratio (DRAM masters/CPU). The graph compares the performance of common I/O, DRAM macro, multi-I/O port, and access optimizer. The y-axis represents the total CPU cycles in arbitrary units, ranging from 0.0 to 2.0. The x-axis represents the access ratio, ranging from 0 to 10. The graph shows a degradation in performance as the access ratio increases, with a performance degradation of 1.39 for Linpack and 1.00 for access optimizer.]
Test chip and waveforms

Access optimizer
1.495 mm²

8-Mbit DRAM Macro

0.18-um process

Clock

Read command

Hit flag

Bank active

Column select

Vdd = 1.8 V, 100 MHz

T. Watanabe, et al., ISSCC 1999 (ref. [20]).
Special purpose processor

- DRAM-based massive parallel processing -

Examples:
1. D. G. Elliott, et al., CICC 1992 (ref. [2]).
2. T. Watanabe, et al., IJCNN 1992 Baltimore (ref. [3]).
3. Y. Aimoto, et al., ISSCC96 (ref. [10]).
Low-voltage parallel operation

Approximately,

\[ P_w: N \times f \times C \times V_{dd}^2 \]
\[ S: N \times f \]
\[ f: V_{dd} \]

\[ P_w: S \times C \times V_{dd}^2 \]
\[ N: S / V_{dd} \]

Total processing speed is constant.

Power dissipation (a.u.)

Supply voltage (V)

No. of processing elements (a.u.)
Pitch-matched configuration

One processing node

- Y decoder
- DRAM cell array
  - Word line
  - Bit line
- DTC
- Processing elements

64 Bit-line pitch

DRAM cell array

- Word line
- Cell
- Bit line
- SA

Data transfer circuit

Buffer

Takao Watanabe - 10th ORAP Forum / 28th SPEEDUP Workshop - Slide 33
Example: 1.5-V neuro chip

Full-scale chip performance

- 0.5-um process
- 8-Mb DRAM
- 256 processing elements
- \( V_{dd} = 1.5 \) V
- 2.74 GOPS (1.37 GCPS)
- 75 mW
- 1.05 x 10E6 synapses
- 15.4 x 18.6 mm

T. Watanabe, et al., 1992 IJCNN, Baltimore (ref. [3]).
Future perspectives (1)

Specialization in LSI architecture will further progress in each application.

1. Cache DRAM: higher random access speed
2. Pixel processor: higher throughput in pixel drawing
3. System on a chip: efficient use with on-chip masters
4. Special purpose processor: highly parallel processing
Future perspectives (2)

Issues to be attacked:

1. Sub-1V memory, logic, and on-chip I/O circuits.
   - 1-T cell has inherent problems in low-voltage region.

2. Easy-to-use and high-performance memory-core architecture.
   - Interesting example - D²RAM (2-Tr, 1-C cell, ref. [21]).

3. Combination of embedded and MCP technologies.

4. Low-cost embedded-memory process
Conclusion

- Embedded DRAM has strong potential, but there are inherent problems.

- DRAM cache, pixel processors, systems on chip, and special-purpose parallel processors are promising applications.

- Many challenging issues still remain.
References (1)

References (2)


References (4)

