PERMEDIA and GLINT Delta

New Generation Silicon for 3D Graphics

Neil Trevett
Vice President Marketing
(408) 436 3455
www.3dlabs.com
3Dlabs New Generation Silicon

- **Professional 3D** - Strong competition to Workstations
- **Pervasive 3D** - Making 2D chips obsolete

**GLINT 300SX**

**3D Blaster Chip**

**GLINT Delta**

**PERMEDIA**

The first single chip geometry pipeline processor

1st Generation

- Q195
- Q395

2nd Generation

- Q196
- Q396

Games

Professional
Alternative Approaches...  
... to low-cost 3D silicon

No Compromises

3D Performance?

2D and 3D Performance?

No 2D?

Performance

Pervasive 3D  FreeD  Games 3D  Arcade 3D

Cost

3Dlabs

Matrox

ATI

S3

3Dfx

VideoLogic

3D Performance?

Pervasive 3D FreeD Games 3D Arcade 3D

ATI

S3

3Dlabs

Matrox

VideoLogic

3Dfx

Rendition

Performance

Cost
Pervasive 3D - the Challenge

No compromises!

- 3D performance must be much greater than software only
  - Software = 5 million texture-mapped pixels per second
  - Hardware should deliver >25 million bilinear-filtered texture-mapped pixels per second

3D for games
3D for authoring

Video Acceleration ➔ Pervasive 3D ➔ 3D for browsing

Fast Windows Acceleration
Fast VGA for DOS games
PERMEDIA Design Targets

• Robust 100% pixel functionality of all key 3D APIs
  • Direct 3D, OpenGL, Heidi, QuickDraw 3D, QuickDraw 3D RAVE
• No 2D compromises
  • >30 Million Winmarks
• Fast 3D performance
  • Balanced performance for both textures and polygons
  • 600,000 textured polygons/second
  • 30 Million bilinear filtered texture-mapped pixels /second
• Low cost
  • Selling on boards costing <$200 (2MBytes)
Avoids byte-swapping

Provides setup-fetch overlap

Avoids polling the FIFO

Glueless PCI Interface

Provides fetch-draw overlap

Allows software drawing

Provides upgrade path
PERMEDIA Memory Interface

- SGRAM for next generation graphics
  - Good random access speed - vital for texture mapping
  - Block fills - very important for clearing buffers
  - Write-per-bit mask, needed for per-window double buffering
  - Upgrade path to 100 MHz and beyond

- 2 to 8 MBytes
  - Up to 4 pages open at any time
  - E.g. front color buffer, back color buffer, depth buffer, texture buffer
Consolidated Memory

*All buffers in same physical memory*

- Efficient and Flexible
  - Dynamically allocate color and depth buffers
  - Any spare memory available for textures
  - Trade resolution for depth buffer, color depth for texture space etc.

- All data in same memory, scope for optimization, e.g.
  - Clear depth buffer with framebuffer block fills

- Use texture operations on *any* image
  - Full scene anti-aliasing
  - Video texture-mapping
  - 3D sprite processing
PERMEDIA Pixel Core

- Hyper-pipelined function units
- Message passing protocol between units
Pipeline Principles

- Each unit in the pipeline is independent
  - Can be designed, tested and synthesized separately

- Unit State Machine
  - Wait for message in input FIFO
  - If message is not relevant, pass to next unit
  - Else process message and pass on any messages as required
  - Return to waiting

- Some units know their place
  - Some units are completely self contained
    - E.g. scissor/stipple unit
  - Some units know where they are in the pipeline
    - E.g. YUV can absorb localbuffer data if the chroma test fails
Unit Pipeline Stage

- The pipeline uses a message passing paradigm
- A message is made up of a tag field and a data field
  - The tag identifies the message type

Two stage FIFO

Tag = 9 bits
Data = 32 bits

Core Unit

Input Stage

Processing

Control

Output Stage

Pipelining as required

Two stage FIFO
Message Passing

- Everything that moves through the pipeline is a message
  - Messages are used to program control registers - e.g. enable texture
  - Messages are used to carry transient information e.g. texture color for current pixel
  - Messages are used as commands - e.g. start new primitive
  - Messages are used for synchronization - e.g. Sync message

- ‘Step’ messages drive the units
  - A Step message for each pixel to be plotted
  - Passive steps are pixels not to be plotted
  - If a pixel fails a test it is converted from active to passive
  - Passive steps cannot be deleted because they advance DDA units
  - Step messages hold the pixel X,Y coordinate in the data field
Unified 2D/3D Pixel Engine

- 3D is a superset of 2D
  - Don’t separate them
- 2D operations use 3D pipeline and use special features
  - Texture units used for tiled blits
  - Chroma key test used for transparent blits
  - Bilinear filter used for stretch blits
- Using the 3D units is gate efficient
  - No duplication of functions
  - No compromise on performance
**PERMEDIA Performance**

- 30 Mpixels/sec, 600K polygons/sec, textured, bilinear, no Z
  - With full per pixel perspective correction, 16-bit framebuffer, 4-bit palletized textures, 50 displayed pixels per polygon, meshed, 640x480 at 75Hz

- 640x480 full screen bi-linear textured, x2.5 depth complexity = 40Hz frame rate

- 2 GBytes/sec Fill rate using SGRAM block fill
  - 2 GBytes/sec Color expansion

- > 30 Million Winmarks

- Video Playback performance - 30fps
  - 320x200 YUV source zoomed and filtered to 640x480x16-bit RGB
PERMEDIA Physical Characteristics

- Packaging
  - 256 pin BGA
  - Wire-bonded into a plastic BGA package
  - 3W at 3.3V

- Process
  - 0.35μ, 4 layer metal
  - 60 MHz

- Shipping now
Board Design

Low component count

- Single PERMEDIA Chip
  - plus SGRAM, RAMDAC, ROM
- External interfaces
  - Glueless PCI Interface
  - High performance 64-bit SGRAM Interface
  - High speed pixel port to RAMDAC

Typically 2MBytes, with optional upgrade to 4MBytes
But where’s the bottleneck?

Geometry!

- The fastest Pentium Pro cannot keep PERMEDIA saturated if running the geometry in software

1K polygons/MHz on a Pentium Class machine (90K polygons on a P5/90)

100% of Rasterization in PERMEDIA silicon

70% of the CPU cycles spent in setup!
GLINT Delta

*Breaking the Geometry Bottleneck*

- Hardwired 3D Pipeline Processing
  - 1M vertex/sec Vertex Setup Processor
  - Performs all delta calculations and floating point conversions
  - 100 MFlop floating point processor
- Reduces PCI Bandwidth - just passing vertices - no slopes

![Diagram](image)

- 110 Bytes / polygon
- 33 Bytes / polygon
GLINT Delta

Setup Processing in a PCI Bridge

Allows transparent use of VGA and 8514 behind bridge

- Full Bus Master
- Primary PCI Bus
- Provides setup-fetch overlap

176 Pin PQFP
3.3V power
5V I/O

Function 0 Decode with VGA/8514
Function 1 Decode

PCI 0
Master & Slave Interface

DMA Controller
Input FIFO
Delta Setup Engine
Output FIFO

PCI 1
Master Interface and DMA Control

Bypass Path

Slope and Setup Calculations for GLINT and Permedia

DMA to GLINT or PERMEDIA

Primary PCI Bus

Secondary PCI Bus

3DLabs - Hot Chips - Stanford August 1996 - Page 20
GLINT Delta

Setup Engine Functionality

- Follows the message passing architecture of PERMEDIA
  - Delta is just another unit in front of the rasterizer
- API neutral - low-level functionality
- Triangle primitive setup (AA and non-AA)
- Line primitive setup (AA and non-AA)
- Interpolation Parameters - XYZ, RGBA, F, STQ, Ks, Kd
- Accepts floating point (IEEE SP) or fixed point inputs
- Texture coordinate auto normalization
- Optional input value clamping
- High precision sub-pixel correction
GLINT Delta Setup Engine

Hardwired processing

Vertex Store 0

Vertex Store 1

Vertex Store 2

Working Store

Input Vertex Information - 16x32

VHDL Coded, Inferred Routing

Data Routing

FMUL

FADD

FDIV

FDIV

FConvert

Output FIFO

Floating Point Operation Units

Temp Storage 8x32
GLINT Delta Calculations

Floating Point improves robustness and visual quality

- Input parameter score-boarding
- All internal calculations in custom floating point format
  - Less dynamic range, but more precision than IEEE
- RGBA/BAZ triangle set-up involves:
  - 41 floating point adds or subtracts
  - 27 floating point multiplies
  - 5 floating point divides
  - plus.. compares, clamping, fixed point/floating point conversions

- Main floating point operators are:
  - One multiplier (one pipeline stage, single cycle).
  - One adder/subtracter (single cycle)
  - Two dividers (5 cycle iterative, autonomous)
  - Four comparators
  - Float to fixed point conversion with clamping
Hard-Wired Processing

Cost-effective floating point performance

• Control is a VHDL state machine.
  • No RAM or ROM for program storage (less gates)
  • No program sequencer or instruction set (less gates)
  • No program fetch (less memory bandwidth)

• Data paths are inferred directly from VHDL
  • No general purpose routing costs

• No software maintenance

• 35 cents / MFlop
GLINT Delta

Physical Characteristics

• Low cost device - 176 pin PQFP
• 0.45µ, 40MHz, 3 layer metal
• Shipping now

Performance
• 1M Meshed Shaded, Z buffered triangles/sec
• 2M 2D polylines/sec
Combined Board Design

*Delta and PERMEDIA*

- Matched Geometry and Rasterization performance
- High performance Arcade machines, VR/simulation engines, entry-level desktop OpenGL acceleration
- Sub $350 street price
# GLINT Delta

## Measured Performance Increases

<table>
<thead>
<tr>
<th>Tspeed3 V3.0 OpenGL</th>
<th>No Delta</th>
<th>With Delta</th>
<th>X Faster With Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meshed Triangles (Z, Shaded) 50 Pixel per second</td>
<td>155,146</td>
<td>238,997</td>
<td>1.54</td>
</tr>
<tr>
<td>Meshed Triangles (Z, flat) 50 Pixel per second</td>
<td>205,870</td>
<td>321,247</td>
<td>1.56</td>
</tr>
<tr>
<td>Meshed Triangles (Z, Shaded) 25 Pixel per second</td>
<td>180,744</td>
<td>427,242</td>
<td>2.36</td>
</tr>
<tr>
<td>Meshed Triangles (Z, flat) 25 Pixel per second</td>
<td>232,398</td>
<td>573,212</td>
<td>2.47</td>
</tr>
<tr>
<td>Meshed Triangles (Z, Shaded) Small Triangles per second</td>
<td>187,454</td>
<td>599,762</td>
<td>3.20</td>
</tr>
<tr>
<td>Meshed Triangles (Z, flat) Small Triangles per second</td>
<td>249,629</td>
<td>586,527</td>
<td>2.35</td>
</tr>
<tr>
<td>Meshed Triangles (Z, Shaded) Single Pixel Triangles per second</td>
<td>187,454</td>
<td>600,476</td>
<td>3.20</td>
</tr>
<tr>
<td>Meshed Triangles (Z, flat) Single Pixel Triangles per second</td>
<td>249,629</td>
<td>586,527</td>
<td>2.35</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, Shaded) 50 Pixel per second</td>
<td>182,048</td>
<td>277,016</td>
<td>1.52</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, flat) 50 Pixel per second</td>
<td>223,155</td>
<td>365,412</td>
<td>1.64</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, Shaded) 25 Pixel per second</td>
<td>199,290</td>
<td>514,781</td>
<td>2.58</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, flat) 25 Pixel per second</td>
<td>272,531</td>
<td>585,847</td>
<td>2.15</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, Shaded) Small Triangles per second</td>
<td>200,159</td>
<td>646,607</td>
<td>3.23</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, flat) Small Triangles per second</td>
<td>271,068</td>
<td>586,527</td>
<td>2.16</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, Shaded) Single Pixel Triangles per second</td>
<td>200,079</td>
<td>646,607</td>
<td>3.23</td>
</tr>
<tr>
<td>Meshed Triangles (No Z, flat) Single Pixel Triangles per second</td>
<td>271,214</td>
<td>586,527</td>
<td>2.16</td>
</tr>
</tbody>
</table>
Future Directions

- More Geometry Pipeline in Hardwired Logic
  - CPUs just aren’t fast enough
  - Hardwired logic is more cost-effective
- Unified Memory
  - Using system memory for texture
  - Intel’s AGP - Accelerated Graphics Port
- 3D Graphics on the Motherboard
  - High integration - RAMDACs and geometry included on-chip
- Aggressive Performance Increases
  - Next generation silicon - single chip million polygon devices
- Major silicon vendors entering graphics chips market