MEMORY SHARING REVISITED

Work in Progress

4th EuroSys Conference
Nuremberg, Germany
2 April 2009

Jorrit N. Herder <jnherder@cs.vu.nl>
Dept. of Computer Science
VU University Amsterdam
THE NEED FOR DRIVER ISOLATION

- Memory corruption is major crash cause
- Device drivers need access to memory
  - OS data structures
  - Application memory
EXISTING PROTECTION SCHEMES

- System V IPC and POSIX Shared Memory
- Not suitable for low-level device drivers
  - Coarse-grained, page-based protection
  - Protection based on UID, not on process
  - Access rights cannot be delegated
  - No seamless integration for safe DMA
  - No automatic cleanup after driver crash
MEMORY GRANTS

- **Safe memory access based on least authority**
  - Precise, byte-granularity memory area
  - Fine-grained, per-process access rights

- **Privileged grant operations mediated by kernel**
  - Memory copying
  - Memory mapping
  - Direct memory access

- **Delegation supported via indirect grants**
# GRANT STRUCTURE

## Direct Memory Grant

<table>
<thead>
<tr>
<th>flags</th>
<th>grantees</th>
<th>base address</th>
<th>memory size</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>T</td>
<td>D</td>
<td>R</td>
</tr>
</tbody>
</table>

## Indirect Memory Grant

<table>
<thead>
<tr>
<th>flags</th>
<th>grantees</th>
<th>former grantor</th>
<th>former grant ID</th>
<th>base offset</th>
<th>memory size</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>T</td>
<td>I</td>
<td>R</td>
<td>W</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flags</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG_WRITE</td>
<td>Grantee may write</td>
</tr>
<tr>
<td>MG_READ</td>
<td>Grantee may read</td>
</tr>
<tr>
<td>MG_INDIRECT</td>
<td>Grant from grant</td>
</tr>
<tr>
<td>MG_DIRECT</td>
<td>Grant from process</td>
</tr>
<tr>
<td>MG_TAINTED</td>
<td>Grant used for DMA</td>
</tr>
<tr>
<td>MG_VALID</td>
<td>Grant slot in use</td>
</tr>
</tbody>
</table>
GRANT STRUCTURE

Address Space of Proc A

A allows B to Read+Write

B’s Grant Table

C can Write

B: R+W

192 B

C: W

A: 0x440

A: 0x500

A: 0x600

C can Read+Write

A: 0x4c0

A: 0x5c0

512 B

256 B

A’s Grant Table

B’s Grant Table

Indirect Grants
IDs = 1, 4

Direct Grant
ID = 1

A

B

4th EuroSys Conf., April 2009

Jorrit N. Herder <jnherder@cs.vu.nl>
THANK YOU

- Download WIP paper from EuroSys website
- Visit me during EuroSys poster session

Are you a student, love to hack systems, and have some spare time?

MINIX 3 takes part in GSOC 2009 ... pick up the flyer for more info!