“Unlimited” Event Channels

David Vrabel

24 October 2013
What are Event Channels?

- Paravirtual interrupts
- Edge triggered
- Bidirectional
- Directed at a single VCPU
How do Event Channels Work?

1. Notify
2. Set Pending
3. Upcall
4. Call Handlers

Domain A

Domain B

Shared Memory

Xen
Problems

- Too few
- No priorities
- Unfair
Additional Requirements

- Identical ABI between 32-bit and 64-bit guests
- Reasonable memory usage
- Easily extensible
New (FIFO-based) Design

```
| Bit 15 | 1 0 |
| READY  | 1   |
| HEAD   |     |
| Priority 0 | 1 1 |
| 15     |     |

Per-VCPU shared memory
```

```
| Bit 31 | 30 29 16 0 |
| Port 0 (Reserved) |
| 1       | 1 0 1 6 |
| 2       | 1 0 1 0 |
| 3       | 1 0 1 0 |
| 4       |         |
| 5       |         |
| 6       | 1 0 1 3 |
| 7       |         |

Up to 132,072 words
```

Shared memory

```
"Unlimited" Event Channels
```
function raise(q, p)
    E[p].pending = 1
    if not E[p].masked and not E[p].linked
        E[p].linked = 1
        if T[q] != p
            linked = link(T[q], p)
        else
            linked = false
    if not linked
        C.head[q] = p
    T[q] = p

1. Set pending
2. Unmasked and not in list?
   Add to list
3. Set LINK of tail
   or
   Set HEAD in control block
4. Advance tail
function handle_one_event(q)
    p = H[q]
    if p == 0
        p = C.head[q]
    link = unlink(p)
    H[q] = link
    if E[p].pending and not E[p].masked
        handle(p)

1. Get local head
2. List was empty? Get new HEAD from control block
3. Clear LINKED and LINK
4. Advance local head
5. Pending and unmasked? Handle IRQ
Event Channel Fairness

![Graph showing latency across different event channel ports for 2-Level and FIFO-based systems.](image-url)
Current Status

- Xen support in 4.4
- Linux support in 3.14
Further Work

- Investigate how priorities may be used
- Add support for limiting number of event channels to other toolstacks (XAPI, libvirt)
- Add support in other OSes
- Xen event lock scalability
- Linux PIRQ setup/teardown refactoring
Questions?

- Design Document
  http://xenbits.xenproject.org/people/dvrabel/event-channels-F.pdf