XenFS

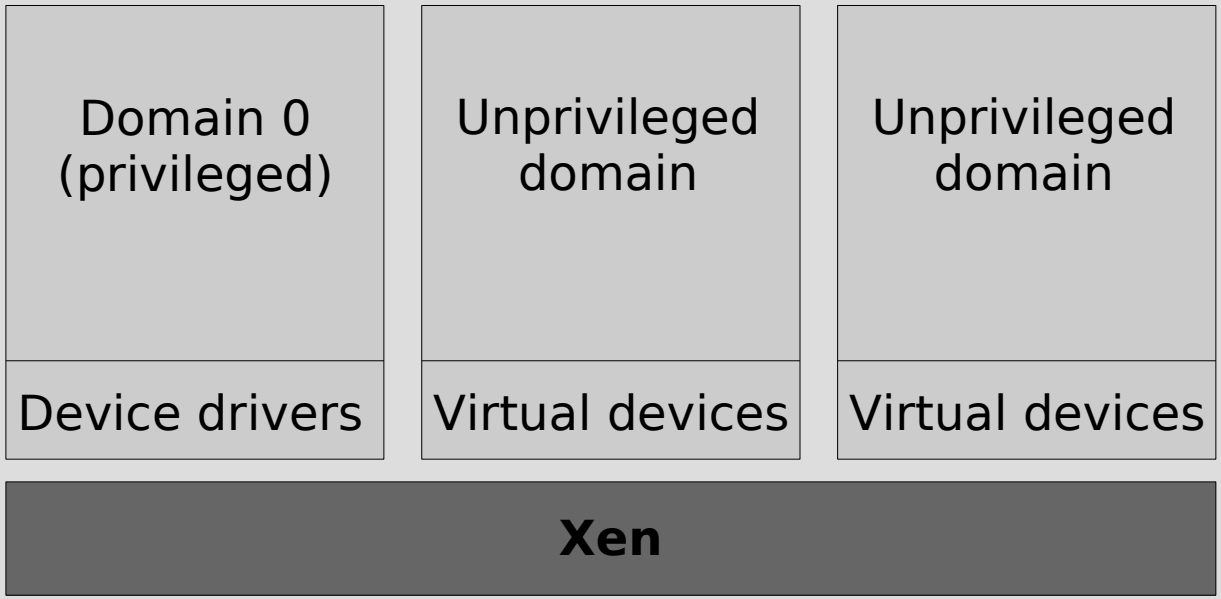
Sharing data in a virtualised environment

Mark Williamson
University of Cambridge
mark.williamson@cl.cam.ac.uk

Overview

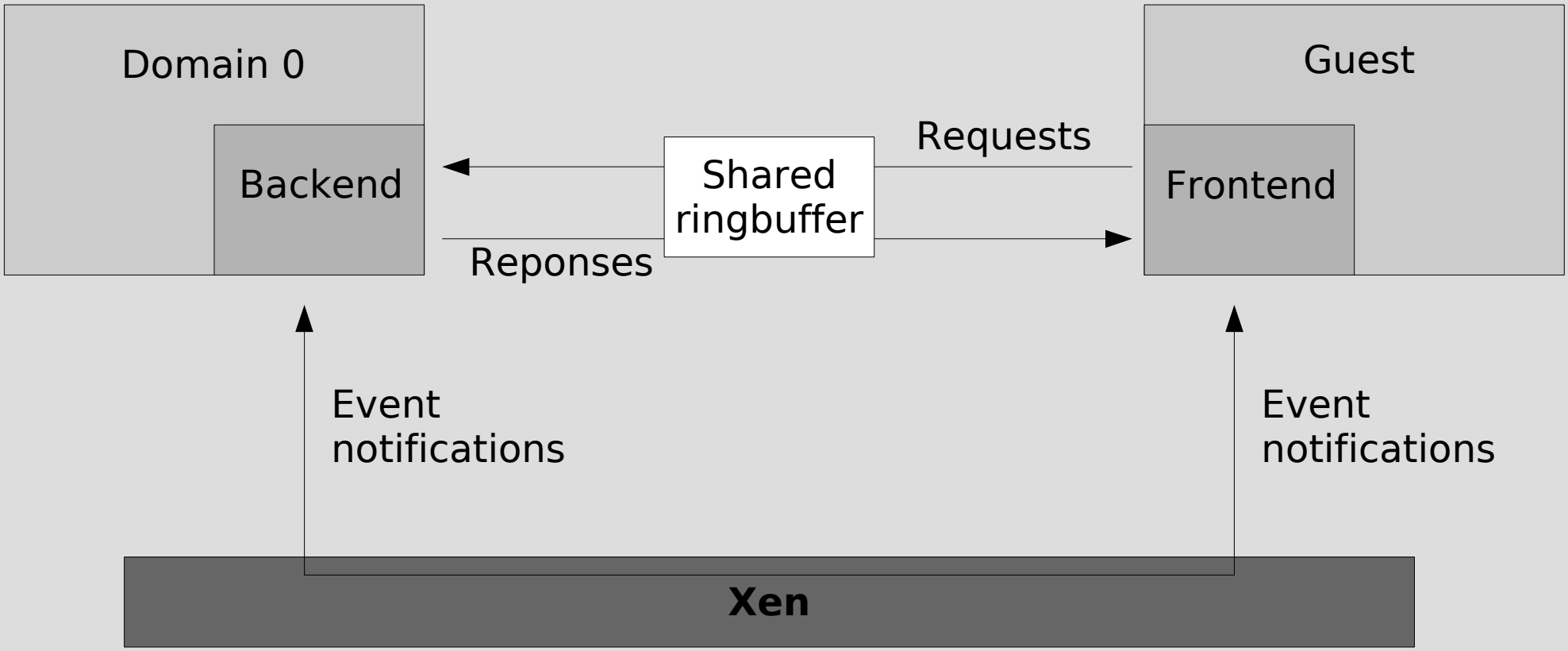
- Introduction to Xen
- Why share filesystems in a VMM?
- Limitations of existing methods
- XenFS will save the world!
- Status and further work

Xen



- Paravirtualised hypervisor for x86
- Run device drivers in privileged virtual machines
- Use inter-VM communication for virtual devices

Virtual devices under Xen



Filesystem sharing under Xen

- Good for administration, efficiency, etc.
- Main options:
  - Share at block level (inflexible but fast)
  - Share via network filesystem (inefficient)

Want (at least) the performance of block-level sharing but NFS-style flexibility

XenFS: the concept

Virtualisation-aware shared filesystem

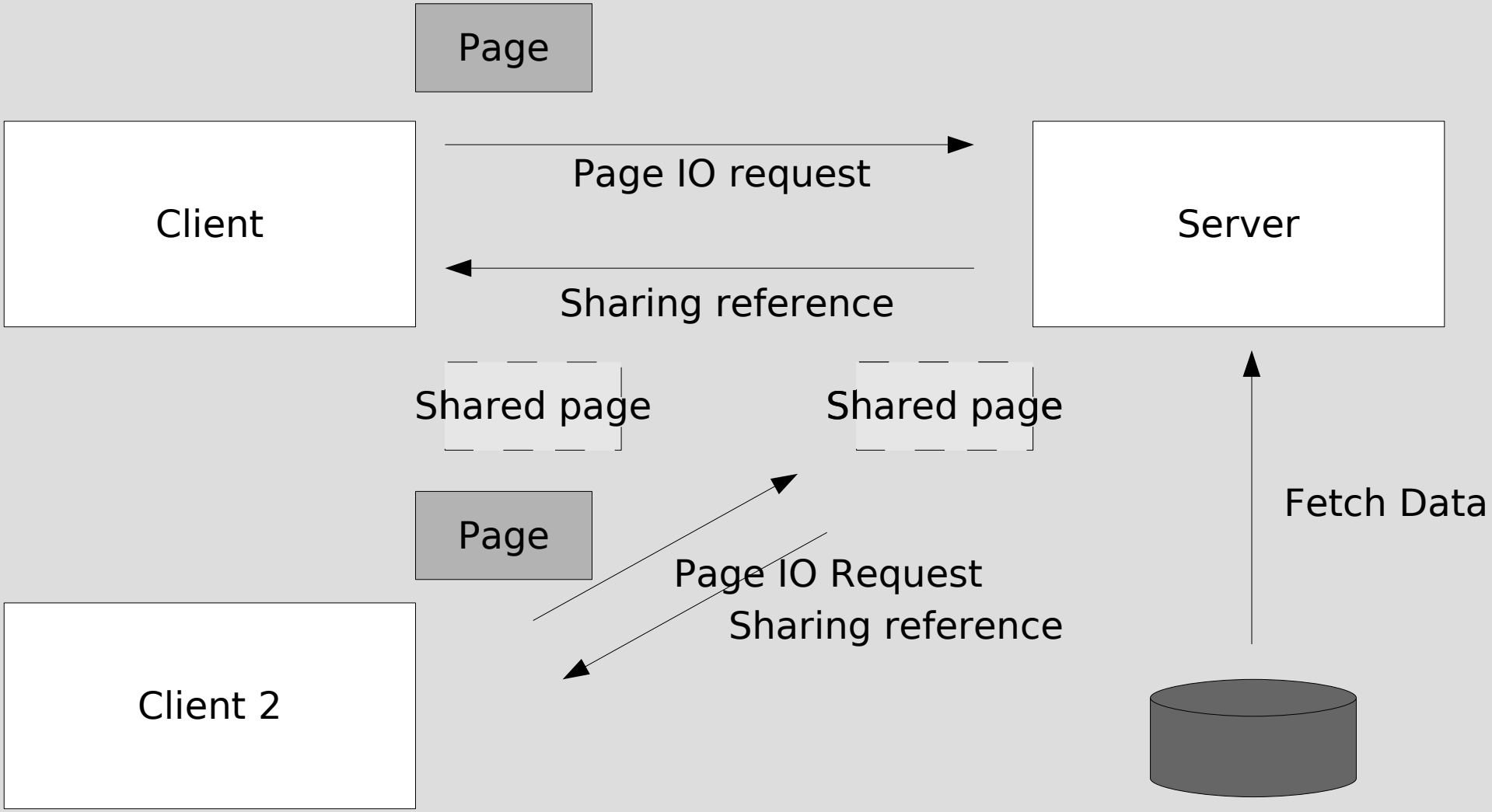
Use shared memory directly as a data transport:

- Filesystem best placed to optimise data transfers
- Maximise performance
- Eliminate duplicated IO and in-memory data
- May perform *better* than existing virtual IO
- **Page sharing**, **page flipping**, **events**

XenFS: planned features

- Export directory hierachies from server
- Multiple modes of operation:
  - FS-level Copy on Write
  - (Reasonably) coherent writable sharing
  - interdomain shared memory (mmap shared files)
- Unify IO caching across the whole host

XenFS: Read mechanism



XenFS: status

- Read-only prototype with page transfer / sharing
- Rewriting for this week's Xen control interface :-)

- Page frame reclamation: designed, no code yet
- Coherent writing mechanisms: but be efficient
- Live migration: soft state in the server
- Benchmarking: XenFS vs NFS vs Block device

Summary

- Sharing data is important in a virtualised system
- Filesystem sharing is an opportunity to improve on block IO
- Virtualising the filesystem gives flexibility & performance

Questions?

XenFS: Undoing sharing

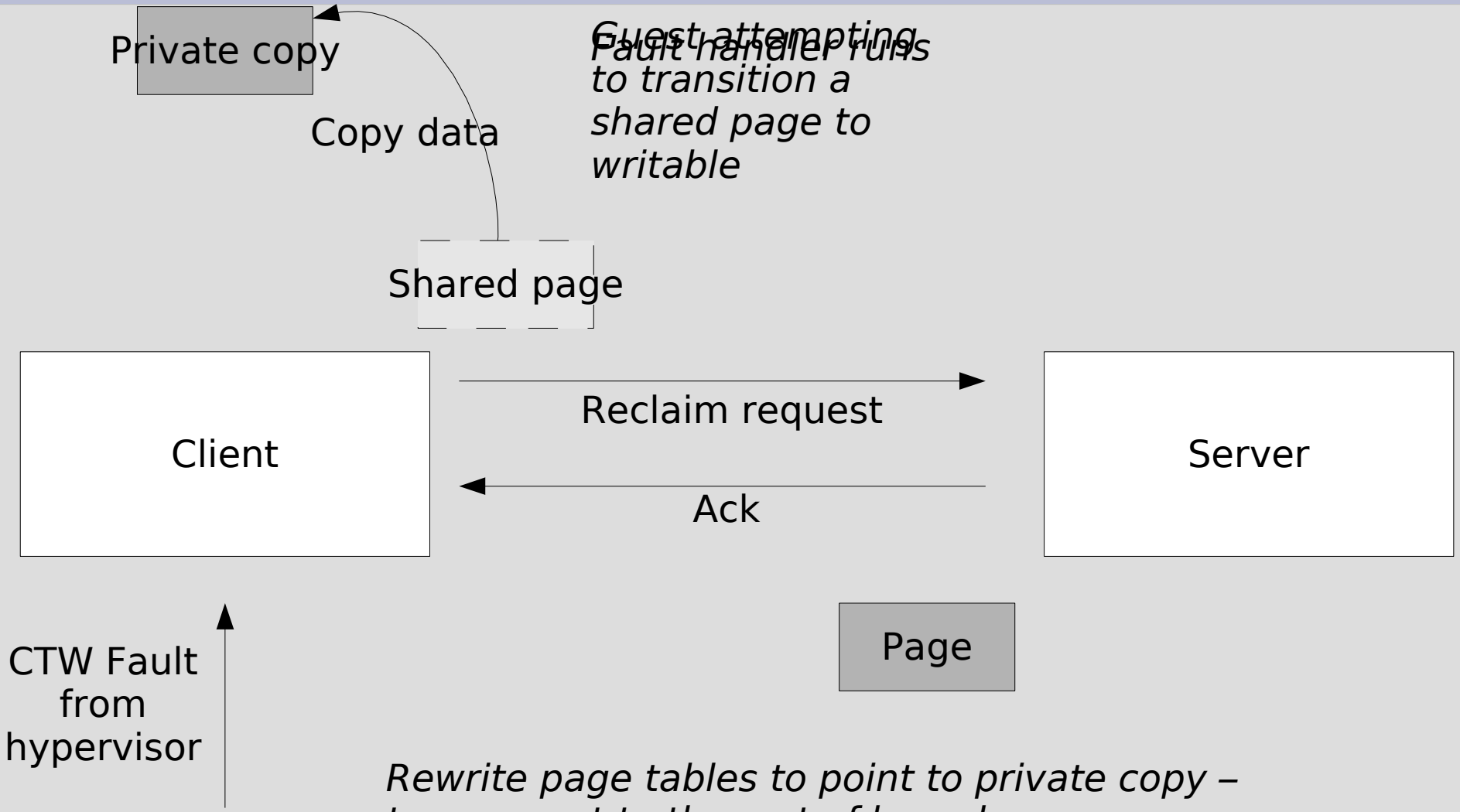
Problem:

- Guests may try to write shared frames:
  - Modify private copy of data
  - Reallocated page cache frames

Solution:

- Mechanism to reclaim memory from server
- New kind of fault: “Copy To Write”

XenFS: “Copy To Write” Faults



Rewrite page tables to point to private copy – transparent to the rest of kernel