

Data Durability and OS Delayed Writes

Kai Shen

Dept. of Computer Science, University of Rochester

Data Durability and OS Delayed Writes

- After a system failure:
 - **Durability/persistence:** ability to return to a consistent state reachable by a series of legal operations that include those whose completions have been indicated to users.
 - Cached asynchronous writes may not be recoverable.
- OS delayed writes:
 - Asynchronous writes are delayed/buffered in memory.
 - Longer delay MEAN more potential loss at failure.
 - Delayed writes with zero delay NOT EQUAL synchronous writes.

2/19/2008

URCS 573 - Spring 2008

2

Delayed Writes

- Benefits:
 - write cancellation, out-of-order scheduling, clustering saves overhead
 - clustering may also allow temporary disk spin-down to save energy [Papathanasiou and Scott, USENIX2004]
- Typical controls:
 - **delay threshold:** limit on how long a given page stays in memory
 - **size threshold:** total amount of memory occupied by delayed writes
- Adaptive per-page delay threshold
 - write working set

2/19/2008

URCS 573 - Spring 2008

3

Rio File Cache [Chen et al., ASPLOS1996]

- Basic idea for surviving OS crashes
 - Non-volatile memory
 - Warm reboot
 - dump all old file cache content after rudimentary warm reboot
 - Analyze and restore the dumped file cache content
- Optimization
 - Keep a concise registry of dirty data in file cache
- Problems:
 - Not overwrite cache content during warm reboot
 - Non fail-stop failure process

2/19/2008

URCS 573 - Spring 2008

4



Virtual Machine Support

- No need for warm reboot
- After a crash, the VM can inspect the existing file cache content and restore delayed writes to disk
- Problems:
 - figure out what pages are in page cache
 - know where to restore the writes to



Rethink the Sync [Nightingale et al., OSDI2006]

- Synchrony/asynchrony is with respect to a user:
 - caller of a system call
 - or a network client
 - or even a real person operating a web browser
- Relaxed user context allows synchronous but delayed writes.
- Challenges/issues?