

Aggressive Writeback in the Linux Kernel

DIMACS REU Project

Robert Renaud

About Linux

- Open Source Operating System

About Linux

- Open Source Operating System
- 100s or more contributors

About Linux

- Open Source Operating System
- 100s or more contributors
- 7 contributors with rutgers.edu emails

About Linux

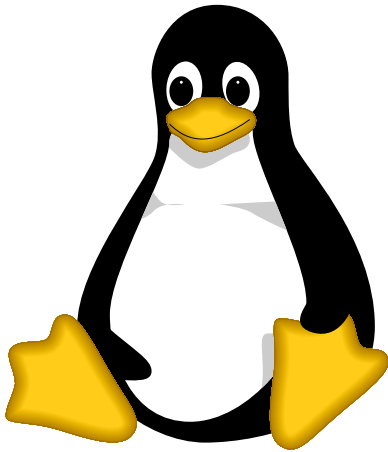
- Open Source Operating System
- 100s or more contributors
- 7 contributors with rutgers.edu emails
- hope to make it 8

About Linux

- Open Source Operating System
- 100s or more contributors
- 7 contributors with rutgers.edu emails
- hope to make it 8
- cute mascot

About Linux

- Open Source Operating System
- 100s or more contributors
- 7 contributors with rutgers.edu emails
- hope to make it 8
- cute mascot



Two Level Memory



RAM

Two Level Memory

- RAM
 - low latency - nanoseconds

Two Level Memory

- RAM
 - low latency - nanoseconds
 - small capacity

Two Level Memory

- RAM
 - low latency - nanoseconds
 - small capacity
 - high bandwidth - GB/s

Two Level Memory

- RAM
 - low latency - nanoseconds
 - small capacity
 - high bandwidth - GB/s
- Disk

Two Level Memory

- RAM
 - low latency - nanoseconds
 - small capacity
 - high bandwidth - GB/s
- Disk
 - enormous latency - milliseconds

Two Level Memory

- RAM
 - low latency - nanoseconds
 - small capacity
 - high bandwidth - GB/s
- Disk
 - enormous latency - milliseconds
 - large capacity

Two Level Memory

● RAM

- low latency - nanoseconds
- small capacity
- high bandwidth - GB/s

● Disk

- enormous latency - milliseconds
- large capacity
- medium bandwidth - 10s of MB/s

Latency/Bandwidth Ratio

● use blocks

Latency/Bandwidth Ratio

- use blocks
- amortized latency = latency / block accesses

Latency/Bandwidth Ratio

- use blocks
- amortized latency = latency / block accesses
- exploits locality of reference

Memory Abstraction

- Virtual Memory

Memory Abstraction

- Virtual Memory
- Hide real location of data

Memory Abstraction

- Virtual Memory
- Hide real location of data
- Evict pages when needed

Eviction

 Clean vs Dirty

Eviction

- Clean vs Dirty
- Dirty write back costly

Eviction

- Clean vs Dirty
- Dirty write back costly
- Eviction Policy
 - Least Recently Used
 - Least Frequently Used

LRU in Linux



active list

LRU in Linux

- active list
- inactive list

LRU in Linux

- active list
- inactive list
- on reference, move up

LRU in Linux

- active list
- inactive list
- on reference, move up
- periodically, refill inactive

LRU in Linux

- active list
- inactive list
- on reference, move up
- periodically, refill inactive
- only approximation to LRU

LRU insufficient

- Databases do own memory management

LRU insufficient

- Databases do own memory management
- Clashes with OS

Properties of Disk

- mechanical movement
 - disk seeks slow
 - small writes costly

Properties of Disk

- mechanical movement
 - disk seeks slow
 - small writes costly
- the big secret
 - page write cost = track write cost

Properties of Disk

- mechanical movement
 - disk seeks slow
 - small writes costly
- the big secret
 - page write cost = track write cost
- about 30 pages per track

Properties of Disk

- mechanical movement
 - disk seeks slow
 - small writes costly
- the big secret
 - page write cost = track write cost
- about 30 pages per track
- try to cluster page writes

Properties of Disk

- mechanical movement
 - disk seeks slow
 - small writes costly
- the big secret
 - page write cost = track write cost
- about 30 pages per track
- try to cluster page writes
- benchmark
 - Sequential page writes - 8 MB/s
 - Random page writes - .35 MB/s

Aggressive Write back

- LRU decides to evict a page

Aggressive Write back

- LRU decides to evict a page
- Aggressively write back nearby pages

Aggressive Write back

- LRU decides to evict a page
- Aggressively write back nearby pages
- Relatively easy to bolt on

Aggressive Write back

- LRU decides to evict a page
- Aggressively write back nearby pages
- Relatively easy to bolt on
- Theoretically wrong

Better Idea

- dirtiness should be bigger factor

Better Idea

- dirtiness should be bigger factor
- prefer dirtier tracks

Better Idea

- dirtiness should be bigger factor
- prefer dirtier tracks
- out of scope of project

Progress

- some source code level understanding

Progress

- some source code level understanding
- printing out data structures in kernel

Progress

- some source code level understanding
- printing out data structures in kernel
- no behavioral modification yet

Acknowledgements

 DIMACS

Acknowledgements

- DIMACS
- Martin Farach-Colton

Acknowledgements

- DIMACS
- Martin Farach-Colton
- All the REU students