**Gecko: Contention-Oblivious Disk Arrays for Cloud Storage**

Ji-Yong Shin¹, Mahesh Balakrishnan², Tudor Marian³, and Hakim Weatherwax¹

¹Cornell University, ²Microsoft Research, ³Google

**Motivation**

- Cloud/Virtualization accelerates consolidation of servers
  - Numbers of CPU cores and VMs increase per server
  - Storage is typically poorly virtualized

**Challenges to I/O contention**

- **I/O contention to overcome**
  - Write-write, write-read, read-read, write-GC, and read-GC
  - RAID cannot preserve high throughput
  - IO performance varies depending on coexisting VMs
  - Vulnerable to I/O Contention
  - LFS only solves write-write contention
  - GC (Garbage Collection) interferes with logging
  - First class reads interfere with logging

**Caching Gecko Chains**

- Caching tail drive
  - Reduces write-read contention
  - 86% of reads absorbed from mix of MS I/O traces
  - RAM + SSD cache
  - RAM: Small amount of hot data
  - SSD: Large amount of warm data
  - RAM extends SSD lifetime by 8X

**Running Enterprise Workloads**

- Mix of 8 MSR Cambridge and MS enterprise workloads
- 6 Disk Configurations
- **Gecko performs**
  - 2-3X better than Log+RAID10
  - Order of magnitude better than RAID 0

**Metadata Management**

- In memory-logical-to-physical map
  - 4-byte entries per page
  - 8GB for 8TB storage

- In flash physical-logical map
  - Maintains persistence
  - Flushed to flash every 1024 page writes
  - Written in sequential order
  - High flash performance
  - Good for flash lifetime

**Varying Chain Length**

- Logs to one disk at a time to reduce I/O contention
- GC and read occurs at the rest of the drives
- Logging solves
  - Write-write contention
  - Write-GC-write contention
  - Chaining solves/reduces
  - Write-GC-read contention
  - Write-read contention
- Mirroring/Striping chains enables
  - Power saving with consistency concerns
  - High performance
  - High fault tolerance

**Conclusion**

- Gecko maintains high I/O performance
  - Securing single uncontended disk
  - Separating reads and writes
- Log-structured designs
  - Oblivious to write-write contention
  - Sensitive to GC-write and read-write contention
- Gecko fixes GC-write and read-write contention
  - Log tail and head separation using chain logging
  - Use of RAM+SSD tail disk cache