Virtualizing Disk Performance with Fahrrad

Anna Povzner, Scott Brandt, Richard Golding†, Theodore M. Wong†, Carlos Maltzahn

Computer Science Department, University of California, Santa Cruz
†IBM Almaden Research Center

FAST WIP Session
February 27, 2008
Storage performance virtualization

- Guaranteed I/O performance in shared storage systems
  - Virtual disk: Ensure desired throughput and latency for clients
  - LUNs virtualize capacity
  - We also want to virtualize performance

- Goal: Throughput equivalent to standalone throughput
  - Amount of data transferred, given configured time interval \( p \):
  \[
  \forall i, \quad D_i(x\%, n \cdot p) = D_i(100\%, x \cdot n \cdot p)
  \]
  - virtual disk with share \( x\% \) during time \( t = n \cdot p \)
  - using disk alone during time \( x \cdot t \)
Performance isolation

- Primary challenge: **performance isolation**
  - Seeks introduced by competing workloads

- Existing approaches provide soft guarantees
  - Façade[lumb:fast03], Argon[wachs:fast07]
The basis for virtual disk abstraction

- Fahrrad real-time disk I/O scheduler
  - Guarantees **disk time utilization** = time spent servicing I/O requests
  - Clients reserve a portion of disk time
  - Reservation granularity bounds latency
  - Minimizes interference between streams

![Pie chart diagram showing disk time allocation](chart)

- Client 1: 30% of disk time every second
- Client 2: 20% of disk time every hour
- Client 3: 50% of disk time every 250 ms
Guaranteeing performance isolation

- Some seeks between streams are unavoidable

**Approach**: Account for inter-stream seeks

- Account for inter-stream seeks caused by competing workloads
- Reserve overhead utilization for time to perform these seeks
- Charge streams responsible for inter-stream seeking
- So I/O performance depends only upon workload behavior
Fahrrad's virtual disk performance

- Semi-sequential stream does not affect sequential stream

![Graph showing Fahrrad's virtual disk performance](image-url)
Fahrrad's virtual disk performance

- Semi-sequential stream does not affect sequential stream
- Virtual disk performance is within 2% of standalone performance