

*Aggressive Prefetching*  
*An Idea Whose Time Has Come*

**Athanasios E. Papathanasiou**

and

**Michael L. Scott**

**University of Rochester**

**HotOS X**

**June 13, 2005**

# *Prefetching, Prepaging or Read-ahead*

---

- **Standard practice for more than 30 years**
- **Complements caching policy (LRU)**
- **Goal: Reduce latency of access to non-cached data**
  - Predict future accesses
  - Read-ahead predicted data
- **A traditional, well-established and well-studied field**
  - **Very rich** literature

# *Far too rich...*

- **Prediction**

- History-based predictors
- Application-disclosed hints
- Compiler-assisted
- Speculative execution

- **Data compression**

Cao et al. – OSDI'94  
Cao et al. – Sigmetrics'95  
Cao et al. – Usenix'94  
Chang et al. – OSDI'99  
Curewitz et al. – Sigmod'93  
Feirtag et al. - SOSP'71  
Griffioen et al. – PDCS'95  
Griffioen et al. - Usenix'94  
Kaplan et al. – ISMM'02  
Kimbrel et al. – OSDI'96

Kistler et al. – TOCS'92  
Kotz et al. – PDIS'91  
Kroeger et al. – Usenix'01  
Kuenning et al. – MSSWIM'02  
Kuenning et al. – SOSP'97  
Kuenning et al. – Usenix'94  
Li et al. – OASIS'04  
McKusick et al. – TOCS'84  
Mowry et al. – OSDI'96  
Palmer et al. – VLDB'91

- **Handling prefetched pages**

- Models to estimate benefit of prefetching and cost of evictions
- Specialized policies for prefetched pages
- Application-controlled

Papathanasiou et al. – Usenix'04  
Papathanasiou et al. – WMCSA'03  
Patterson et al. – SOSP'95  
Satya – IEEE Transactions on Computers 90  
Shen et al. – IPDPS'04  
Steere – SOSP'97  
Tait et al. – ICDCS'91  
Tomkins et al. – Sigmetrics'97  
Yeh et al. – ISPASS'91  
Yeh et al. – WPMRTES'01

# *Common Debate on Prefetching Conservative or Aggressive*

---

- **Aggressiveness aspects**

- Timing: How early should prefetching occur?
- Data coverage: How speculative should it be?

- **Conservative: Reasonable**

- Fetch data incrementally when confidence is high
- Variants of sequential read-ahead widely adopted

- **Aggressive**

- Prefetch deeper and earlier than necessary
- Speculate more to increase data coverage
- Potential to improve I/O performance significantly
  - ✓ Reduce demand-misses (increased data coverage)
  - ✓ Avoid increased delays due to irregular device response times
- **BUT...**

# *Aggressive Prefetching*

## *Traditional Concerns*

---

- **Risk of severe negative impact on performance**
  - Buffer cache pollution
    - ✓ Prefetch *unnecessary* and eject *useful* data
  - Increased physical memory pressure
    - ✓ May even lead to *thrashing*
  - Inefficient use of I/O bandwidth
    - ✓ Speculation may *waste* bandwidth on unnecessary data
  - Increased device congestion
    - ✓ Increased number of *asynchronous* requests in I/O queues
- **Solutions exist, but**
  - Introduce significant implementation complexity
  - Problems more severe on resource-limited systems
  - Forgo potential benefits at high-end to avoid problems at low-end

# *What has changed?*

---

- **Technological and market trends**

- Dramatic improvements in:
  - ✓ Processing power, storage capacity, I/O bandwidth (lesser)
- Modest improvements in I/O latency
- Decrease risks and increase need for aggressive prefetching

- **Emerging usage patterns and design goals**

- Multitasking, larger data sizes, multimedia applications
- Power efficiency, weakly-connected operation
- Increase value of aggressive prefetching

# Technological and Market Trends

# Magnetic Disk Performance

- **Annual performance improvement**

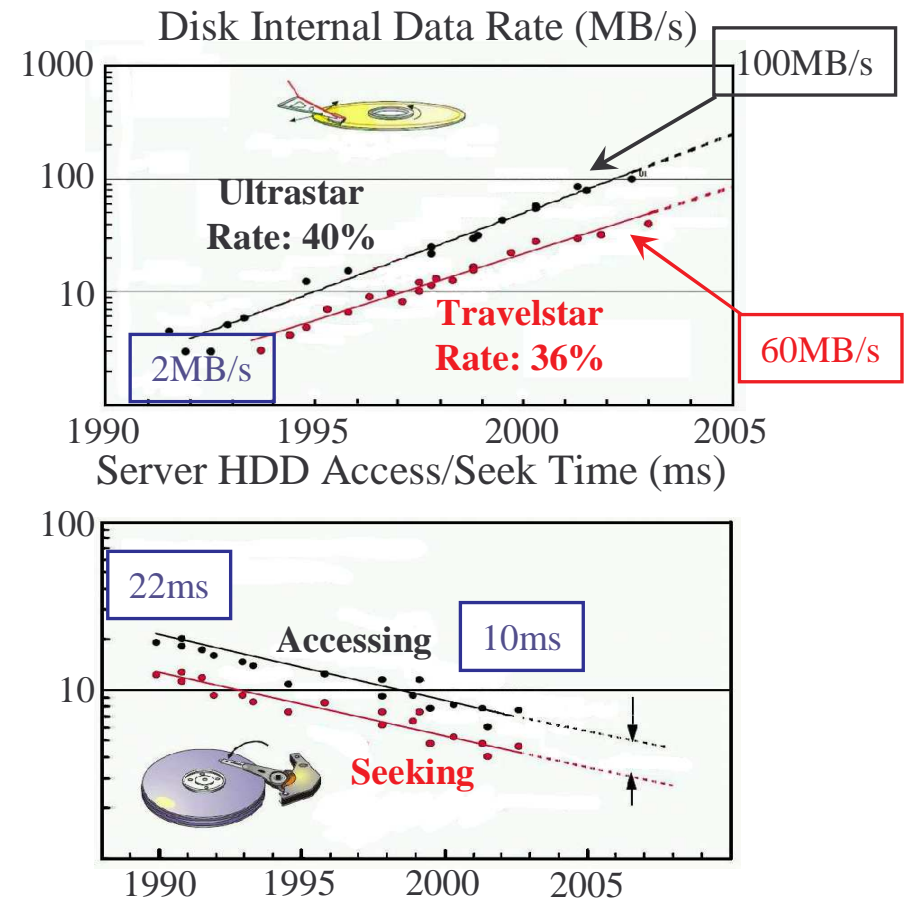
- Bandwidth: 40%
- Latency: **10% only**
- Gap will increase

- **To balance disk activity**

- Read more on disk context switch
- Disks speculatively read whole tracks

- **Aggressive prefetching**

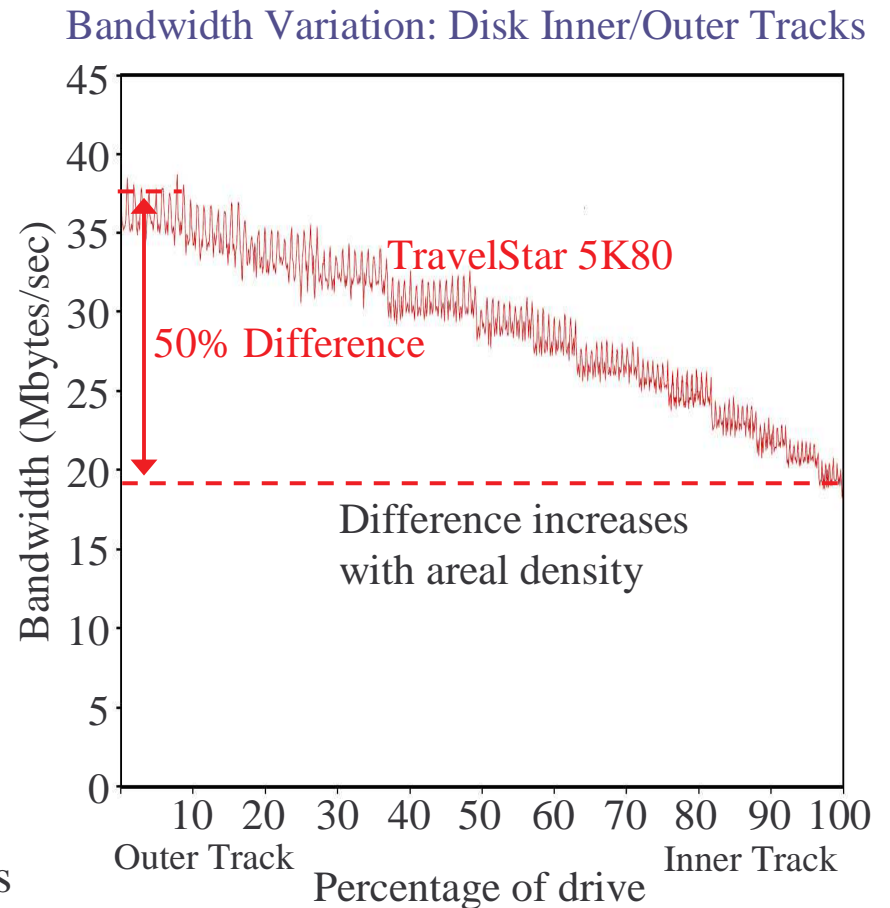
- Exploits bandwidth and caches
- Improves disk utilization



Source: Hitachi Global Storage Tech: HDD Technology'03

# I/O Performance Variability

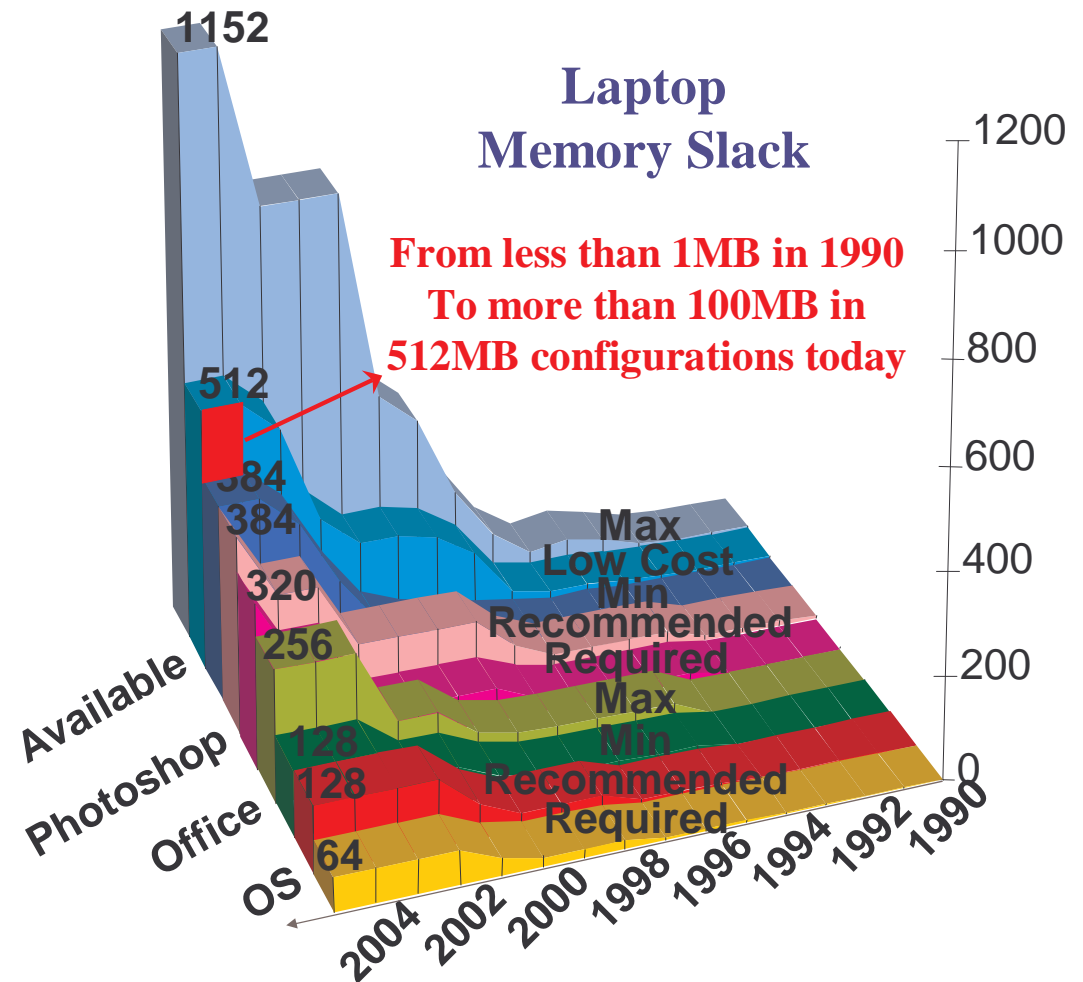
- **Constant bandwidths and latencies**
  - Assumption no longer holds
- **Significant variability**
  - Several I/O devices
  - Variation within a single device
    - ✓ Up to 47% improvement for server-class applications
- **Aggressive prefetching**
  - Exploit periods of higher bandwidth
  - Hide periods of higher latency
  - Smooth out performance fluctuations



Source: Hitachi Global Storage Tech: Adaptive Formatting

# Large Memory Size at Low Cost

- **Production**
  - 70% annual increase
- **Price**
  - 32% annual decrease
  - Today: 12 cents per MB
- **Memory configurations**
  - 512 MB common for laptops
  - Several GBs for desktops
- **Increased memory slack**
  - Prefetch aggressively
  - Lower risk of
    - ✓ Pollution
    - ✓ Memory Pressure



# Usage Patterns and Design Goals

# *Larger Data Sets, Less Reuse*

---

- **Significant digital content increase**
  - 114% increase of hard disk content production from 1999 to 2003
- **Proliferation of multimedia content**
  - Each datum touched only once
  - Multimedia application exhibit predictable patterns
- **Large data sizes diminish benefit of larger memories**
- **Importance of prefetching will increase**

# *Processor I/O Gap and Multitasking*

---

- **Processing power far ahead of I/O latencies**
  - Prefetching has to work multiple blocks ahead
- **Applications**
  - Alternating compute and I/O intensive phases
  - Prefetch to hide latency of **last** I/O access
    - ✓ Prefetch just-in-time for next access is not enough
    - ✓ Prefetch earlier and deeper
- **End-users work concurrently on multiple tasks**
  - Switching among several data hungry applications
  - Reduces effectiveness of LRU-style caching
- **Aggressive prefetching on task switch**

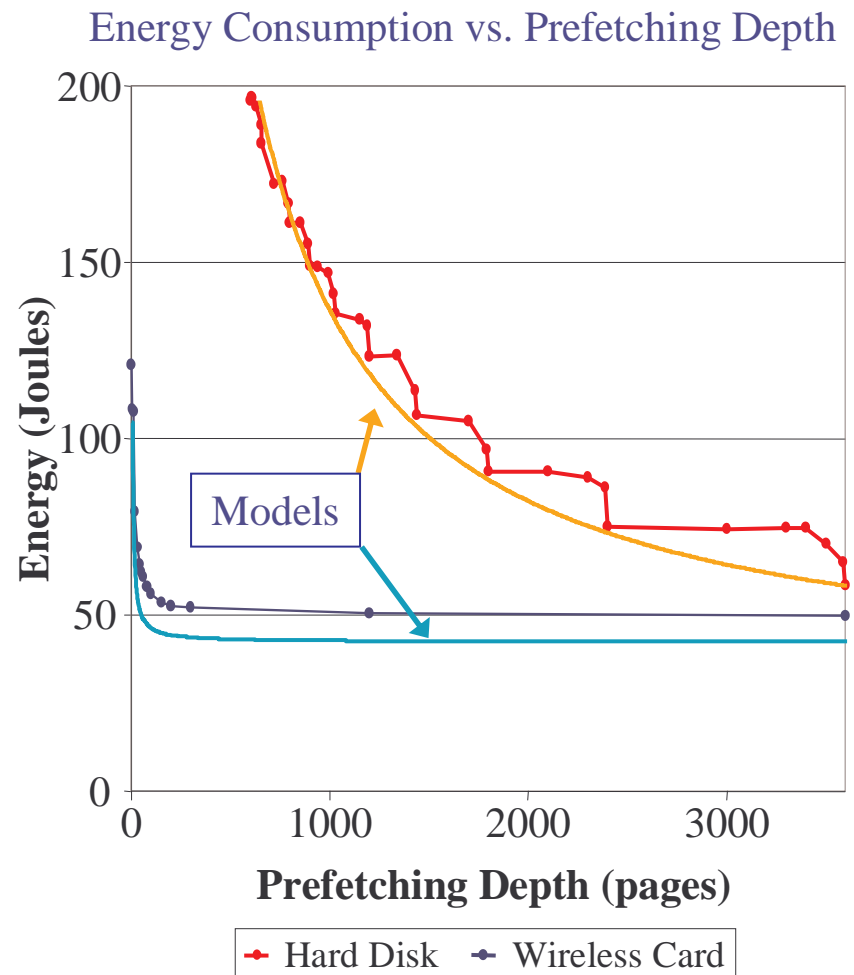
# Energy and Power Efficiency

- **Prefetching impact on energy consumption**

- Bursty, speculative prefetching
  - ✓ Benefits non-operational modes
    - 72% of disk energy savings with just 20% prefetching accuracy
- Smooth prefetching
  - ✓ Benefits operational modes (multi-speed disks)
- Tradeoff?

- **Benefit varies with device**

- Modeling prefetching benefit
  - ✓ Depth, accuracy, speculation
- For informed decisions



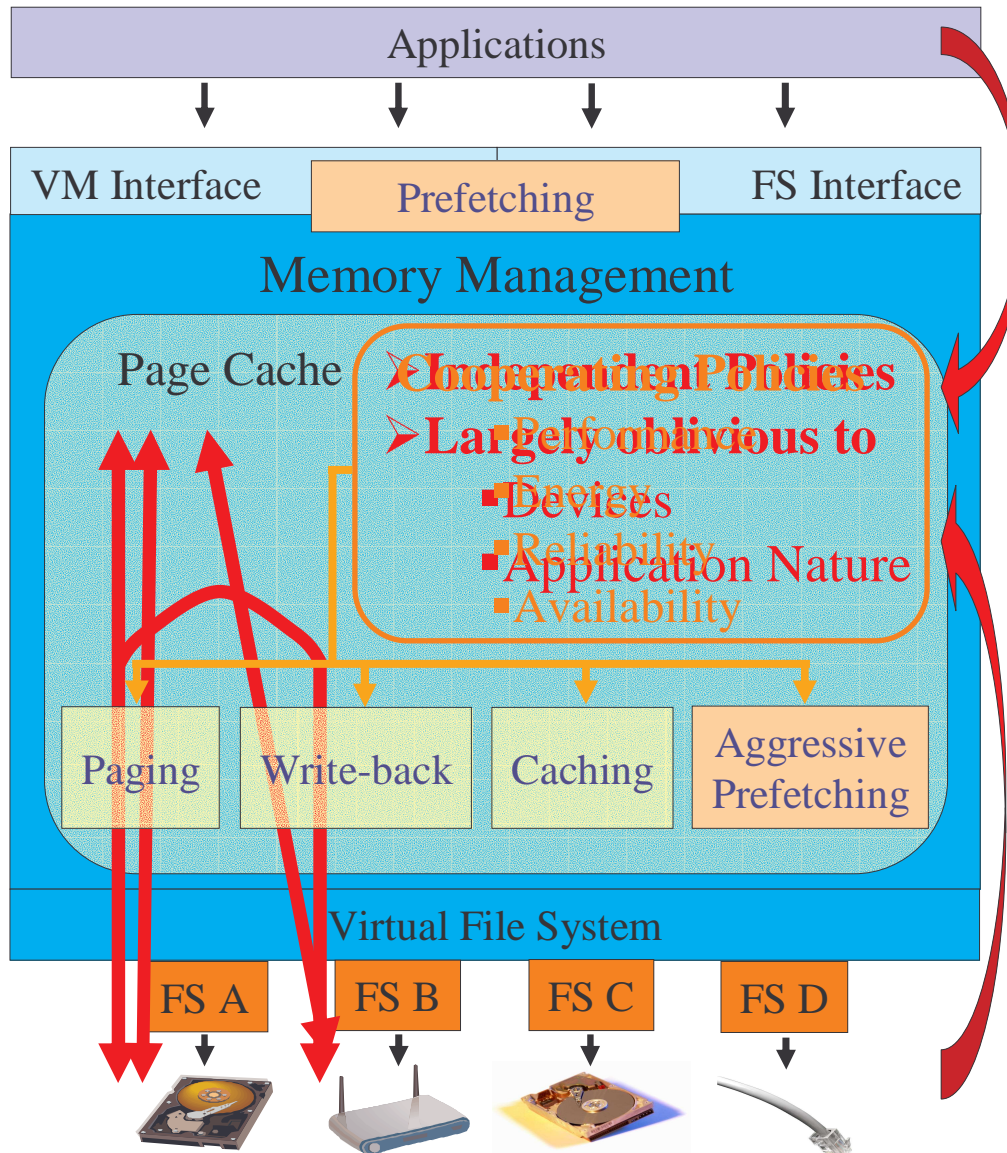
# *Reliability and Availability of Data*

---

- **Mobile systems**
  - Disconnected, weakly-connected operation
  - Portable storage devices
    - ✓ New form of weak connection to data
- **Traditional topics of file system research**
  - **Problem:** Multiple (legacy) file systems
    - ✓ Do not handle disconnection well
- **Hoarding based on aggressive prefetching**
  - Implement at memory management/virtual FS layers
  - Benefits *arbitrary* underlying file systems

# Design Challenges

# Design Challenges for Prefetching



- **Device-centric**
- **Application-sensitive**
  - Access patterns
  - Latency resilience
  - Bandwidth demands
  - Availability
- **Coordination**
  - Across devices and applications
- **Models and Metrics**
  - Miss ratios?
- **Speculative predictors**
  - Patterns and
  - User tasks
- **Fallback mechanisms**
  - Monitoring
  - Canceling pending I/O

# Conclusions

---

- **Memory management of today**
  - Caching secondary storage
  - Oblivious to
    - ✓ Nature of applications
    - ✓ Devices beneath
- **Storage management of the future**
  - Application sensitive
    - ✓ Track and predict behavior
    - ✓ Monitor and coordinate I/O needs
  - Device sensitive
    - ✓ Model latency, bandwidth, availability and power consumption
    - ✓ Move data
      - Between memory and devices, and also
      - Among I/O devices

*Aggressive Prefetching*  
*An Idea Whose Time Has Come*

**Q&A**

<http://www.cs.rochester.edu/~papathan/research/Athena/>

**HotOS X**  
**June 13, 2005**