



Dynamic Adaptability in Support of Extreme Scale

Dynamic Adaptive Operating Systems -- I/O

Seetharami R. Seelam

Patricia J. Teller

University of Texas at El Paso

El Paso, TX



Goals

Dynamic Adaptability in Support of Extreme Scale

- Present a summary of our ongoing work
- Solicit assistance from the greater community
- Hear comments (+/-)



Project Goal

Dynamic Adaptability in Support of Extreme Scale

Enhanced Performance

Generalized  Customized
resource management

Fixed  Dynamically Adaptable
OS/runtime services



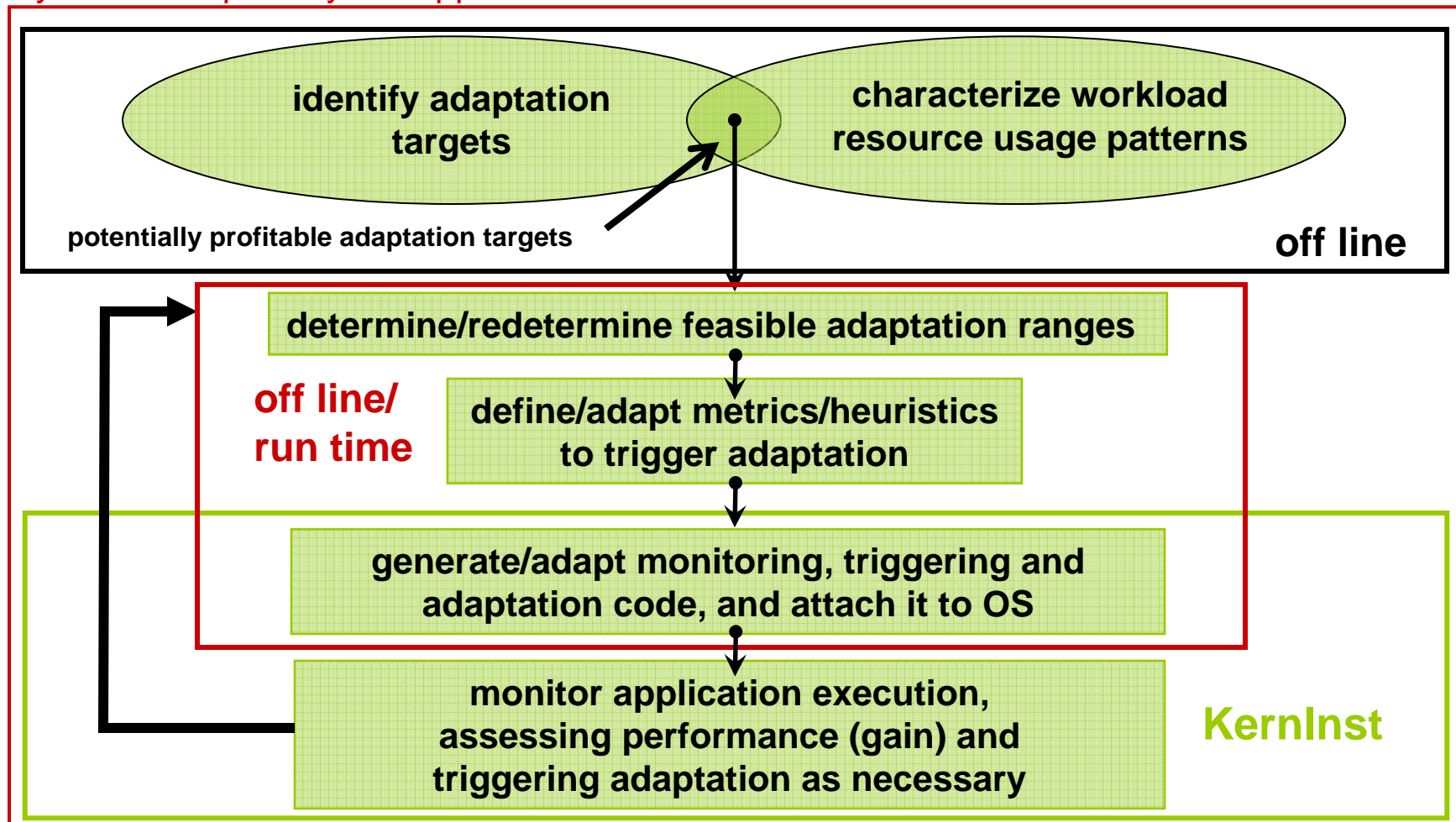
Project Challenges

Dynamic Adaptability in Support of Extreme Scale

Determining

- **What** to adapt
- **When** to adapt
- **How** to adapt
- **How** to measure effects of adaptation

Dynamic Adaptability in Support of Extreme Scale





Project Challenges

Dynamic Adaptability in Support of Extreme Scale

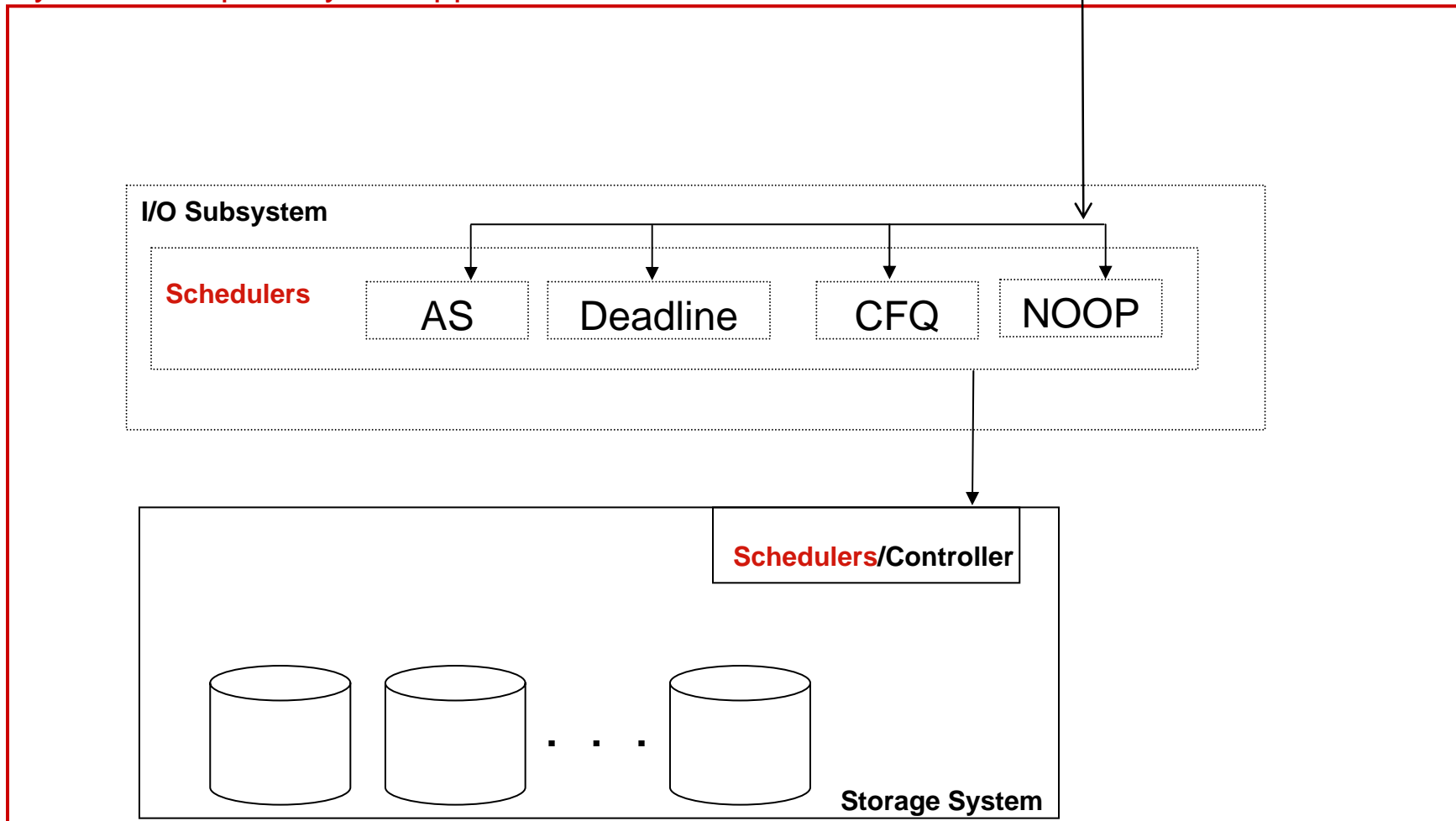
Determining

- **What** to adapt
- **When** to adapt
- **How** to adapt
- **How** to measure effects of adaptation



I/O Schedulers

Dynamic Adaptability in Support of Extreme Scale





Introduction

Dynamic Adaptability in Support of Extreme Scale

- Linux 2.6 provides four I/O schedulers:
 - anticipatory (AS)
 - deadline
 - completely fair queuing (CFQ)
 - noop
- Selection at
 - boot time: one scheduler for all drives
 - runtime: one scheduler per drive
- Default: AS or CFQ



Solution

Dynamic Adaptability in Support of Extreme Scale

Automatic and dynamic selection of scheduler or scheduling policy (within a monolithic scheduler)



Example

Dynamic Adaptability in Support of Extreme Scale

Determining

- **What** to adapt - **I/O Scheduler**
- **When** to adapt - ??
- **How** to adapt - **Explicit/Implicit Selection**
- **How** to measure effects of adaptation - **Execution Time (??)**

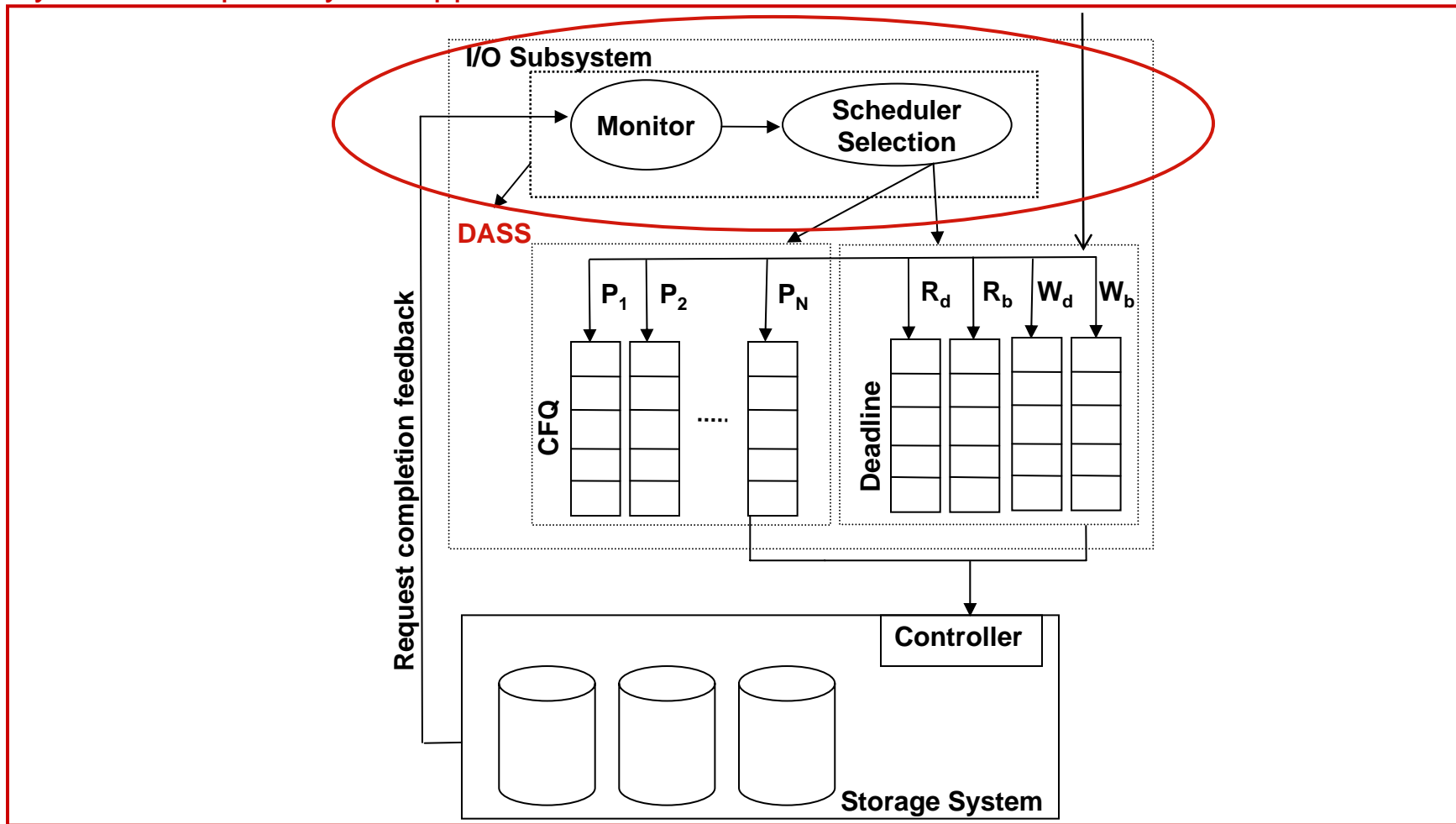


Example Adaptation

Dynamic Adaptability in Support of Extreme Scale

- Automatic and dynamic scheduler selection
- Proof of concept with two schedulers: one providing bounded latency on requests, the other providing fair allocation of bandwidth (presented at OSIHPA Workshop)

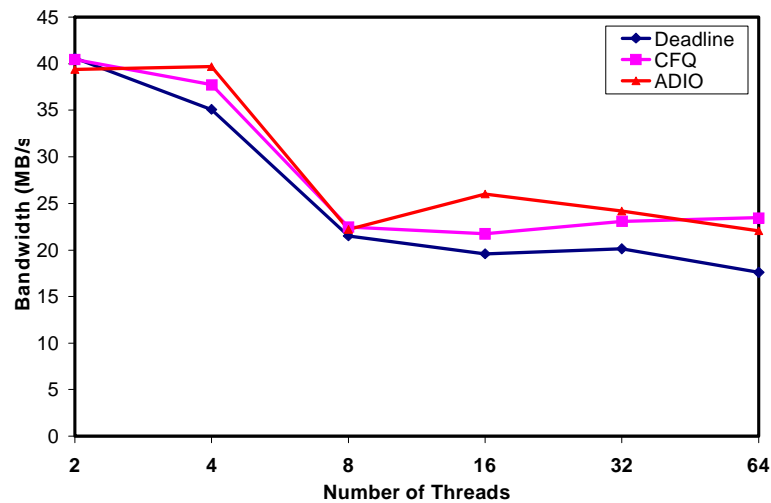
Dynamic Adaptability in Support of Extreme Scale





Experimental Evaluation Enforcing Bandwidth

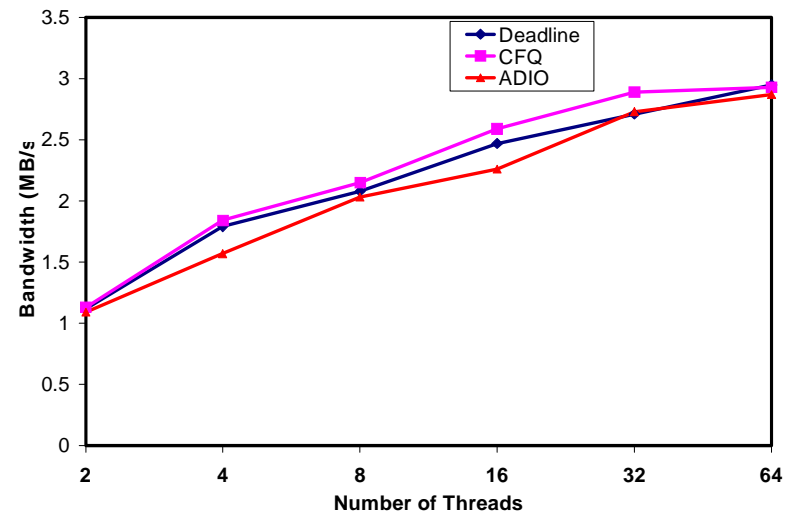
Dynamic Adaptability in Support of Extreme Scale



Sequential utilization

- Disk utilization
- tiobench: 2 GBs data in 4 KB blocks with 2 to 64 concurrent threads
- Several scheduler swaps

- Draining time is a factor in reduced bandwidth

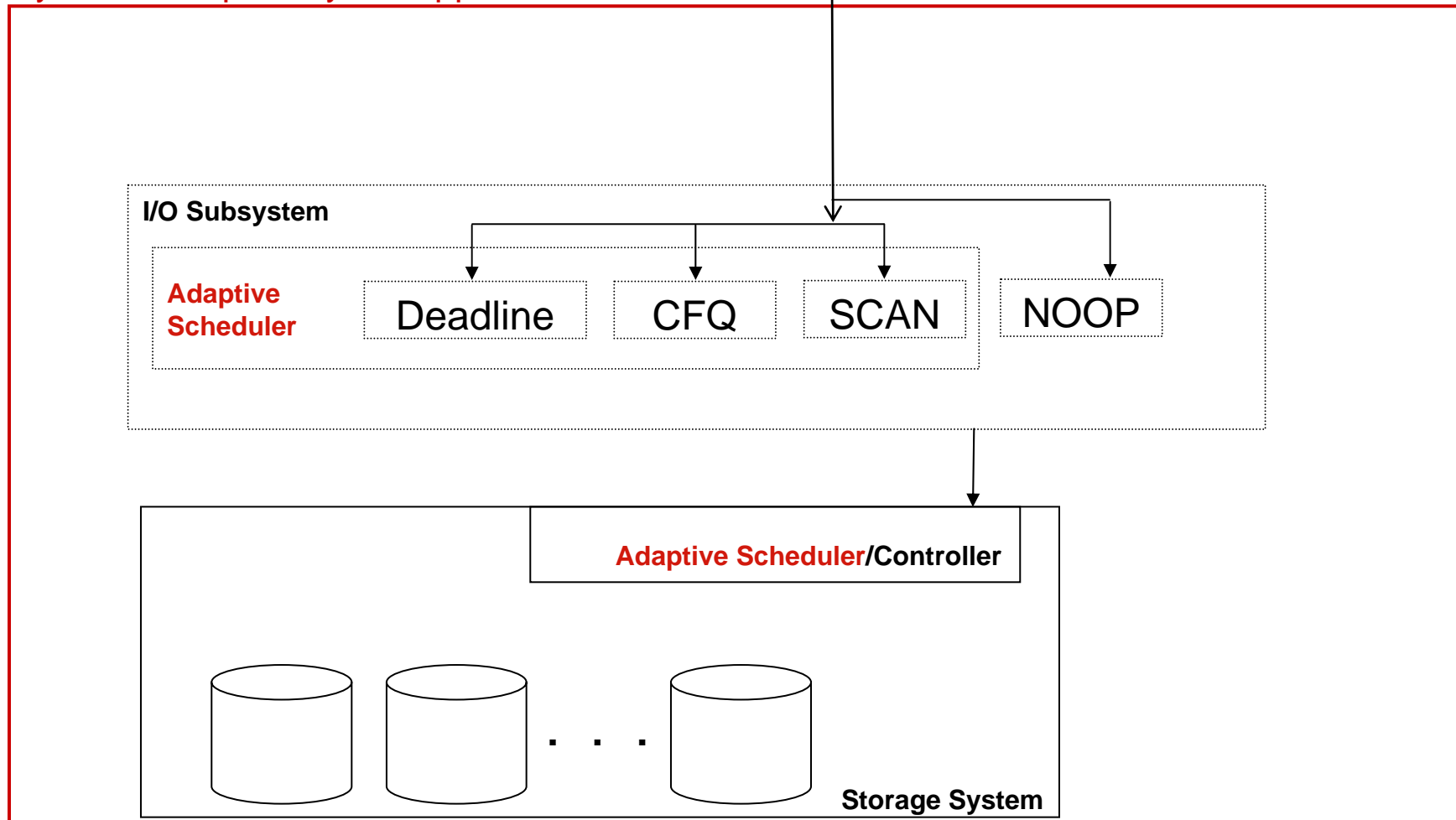


Random utilization



Implicit Selection

Dynamic Adaptability in Support of Extreme Scale





Goals -- Revisited

Dynamic Adaptability in Support of Extreme Scale

- Present a summary of our ongoing work
- Solicit assistance from the greater community
 - I/O scheduling -- Interesting Applications
 - VM -- Memory Hogs
 - Other areas
- Hear comments (+/-)



Acknowledgements

Dynamic Adaptability in Support of Extreme Scale

- We thank:
 - DOE (Grant # DE-FG02-04ER25622) and UTEP for financial support
 - IBM Corporation for IBM Shared University Research Grant (SUR)
 - SC|05 and Harvey Wasserman for creating and conducting this session
 - You for your interest



References

Dynamic Adaptability in Support of Extreme Scale

- Seetharami R. Seelam, Jayaraman Suresh Babu, Patricia J. Teller, "Automatic I/O Scheduler Selection for Latency and Bandwidth Optimization," *Proceedings of the Operating System Interference in High Performance Applications (OSIHPA)*, St. Louis, MO, 17 September 2006.
- J. Axobe, "Time Sliced CFQ I/O Scheduler," <http://kerneltrap.org/node/4406>
- J. Axobe, "Linux Block IO – Present and Future," *Proceedings of the Ottawa Linux Symposium 2004*, Ottawa, Ontario, Canada, pp. 51-61, July 21-24, 2004.
- tiobench: Thread I/O bench for Linux, <http://tiobench.sourceforge.net>
- Choosing an I/O Scheduler for Red Hat Enterprise Linux 4 and the 2.6 kernel, <http://www.redhat.com/magazine/008jun05/features/schedulers/>
- Borril, J., J. Carter, L. Oliker, D. Skinner, and R. Biswas, "Integrated Performance Monitoring of a Cosmology Application on Leading HEC Platforms," *Proceedings of the 2005 International Conference on Parallel Processing (ICPP-05)*, June 2005.
- Pratt, S., and D. Heger, IBM-Austin, "Workload Dependent Performance Evaluation of Linux 2.6 I/O Schedulers," *Linux Symposium*, 2, July 2004, pp. 425-448.



Thank You

Dynamic Adaptability in Support of Extreme Scale

Questions / Thoughts?
seelam@cs.utep.edu
pteller@utep.edu

