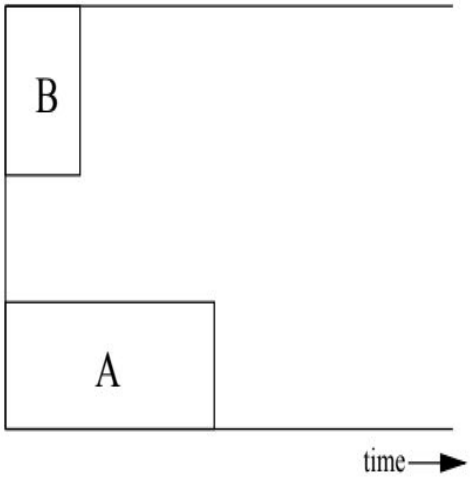
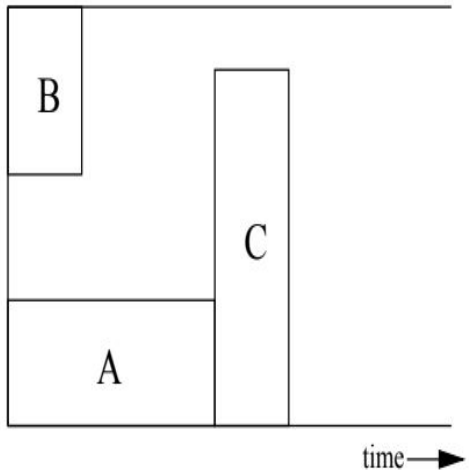
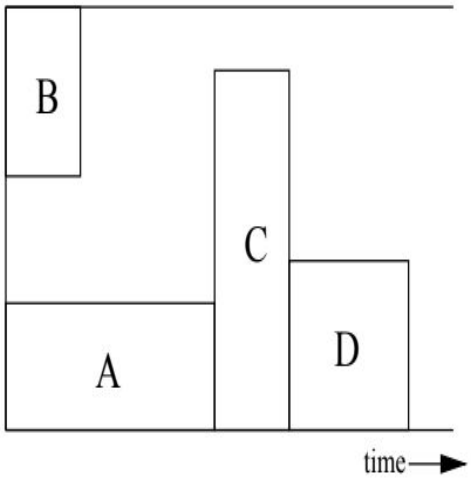
Backfilling with Guarantees Granted Upon Job Submission

Presented by
Lizzy Shakman & Oscar Gardella

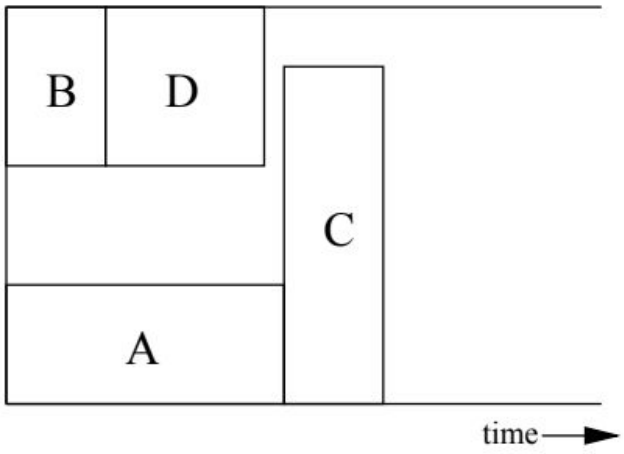
How are jobs scheduled?







What is Backfilling?



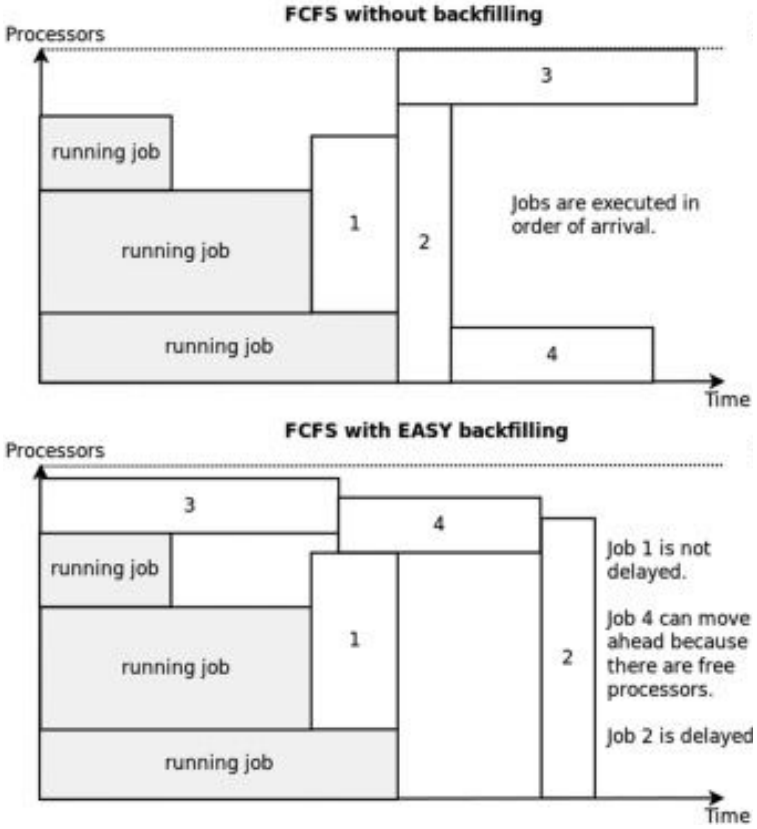
Meaning

- It's what happens when jobs finish early
- Backfilling moves other jobs into the space that is freed up by jobs finishing early
- Attempts to minimize time that processors are idling

Parameters

- Has some number of reservations or jobs with guaranteed start times
- Has an order to the queue of jobs
- Has an amount of lookahead into the queue

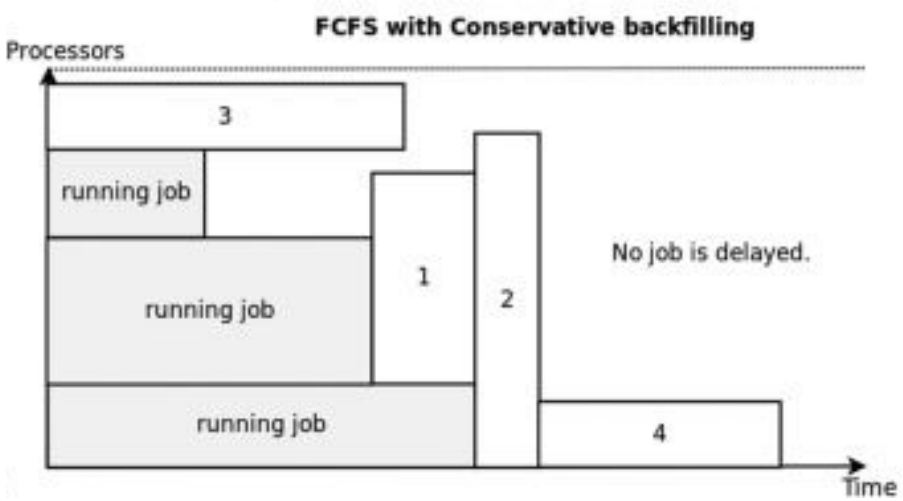
EASY Algorithm



- Jobs backfill if the <u>first</u> job in the queue is not delayed
- Only first job gets a start time reservation
- Sufficient to prevent starvation
- Benefits small/short jobs

Aside: Fattened backfilling is a variation on EASY which doesn't delay jobs more than the average wait time of completed jobs

Conservative Compression



- Jobs backfill only if no other job gets delayed
- Every job gets a start time reservation when it arrives
  - Hence no jobs can be pushed back
- Good for wide/long jobs so they don’t get pushed back by smaller jobs
- Runs fast because it doesn’t do much
- Doesn’t maximize space the best
- First-Come First-Serve

Introduction

Purpose

- Support guaranteed times
- Favor jobs with system-desired traits

Assumptions

- Rigid jobs
- Pure space-sharing
- No interference between jobs

Parameters

- Number of reservations or jobs with guaranteed start times
- Order or queue jobs
- Amount of lookahead into the queue
- **When decisions are made**

Flexibility

- Job selection
- Timing

Conservative with Prioritized Compression (PC)

- Uses a prioritized compression queue
- Reorders the profile when a job arrives or terminates early
- Tries to reschedule each job in the order given by the compression queue
- Returns to the front of the queue when a job reschedules
- Allows high-priority jobs to benefit
- Avoids idle time
- More time consuming

Conservative with Prioritized Compression (PC)

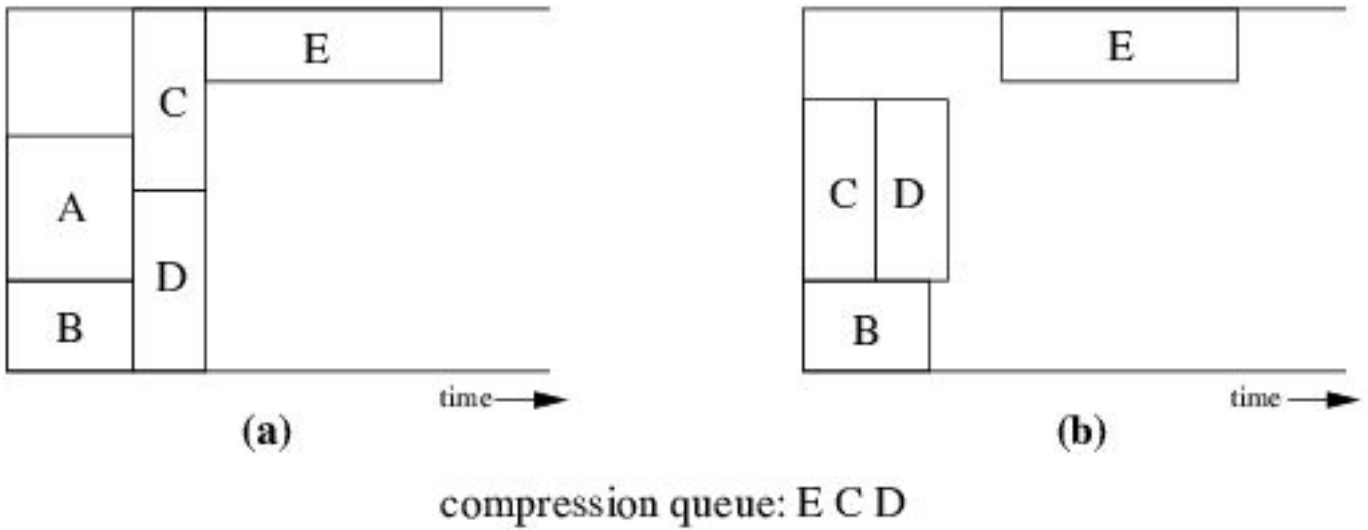


Fig. 1. Profile showing need to return to beginning of the compression queue after each successful rescheduling. (a) Initial profile before job A terminates early. (b) Profile after rescheduling jobs E, C, and D once each in that order.

Conservative with Prioritized Compression (PC)

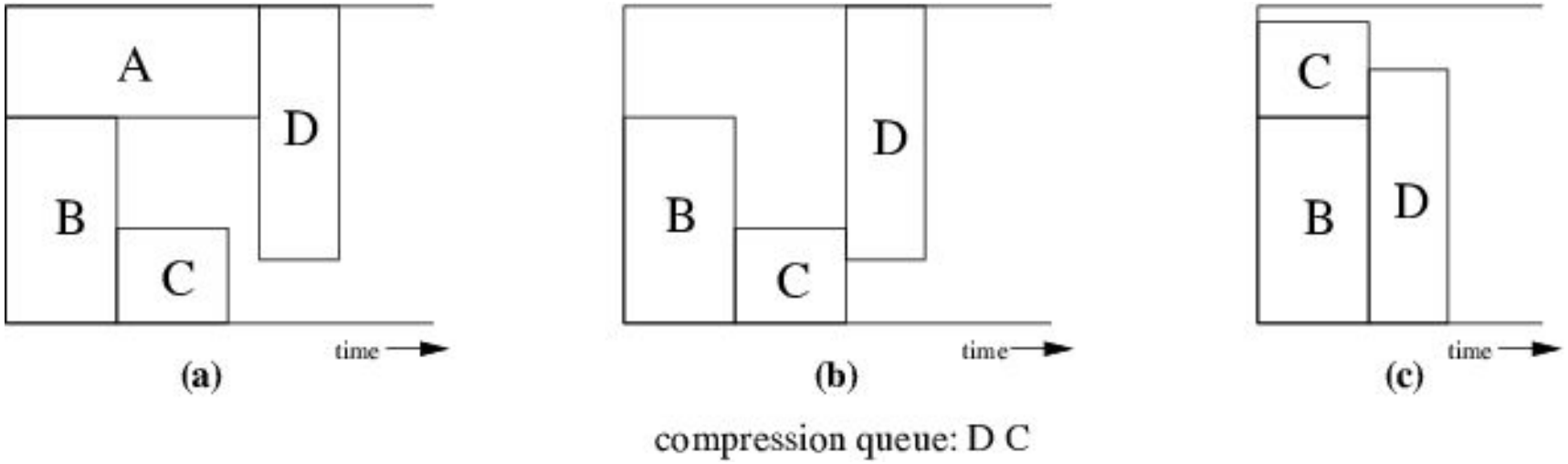


Fig. 2. Example where PC compression moves the same job twice. (a) Initial profile before job A terminates early. (b) Profile after first compression of job D. (c) Profile after compressing job C and then job D again.

Conservative with Delayed Compression (DC)

- Uses a prioritized compression queue like PC, but also defers some rescheduling operations
- Only reschedules jobs that can begin immediately
- Reorders the profile only when a job arrives or can run immediately
- Requires a check after a job finishes and some processors are idle
- Favors short jobs because they can fill the gaps
- Downside: Jobs can be moved more than once

Conservative with Delayed Compression (DC)

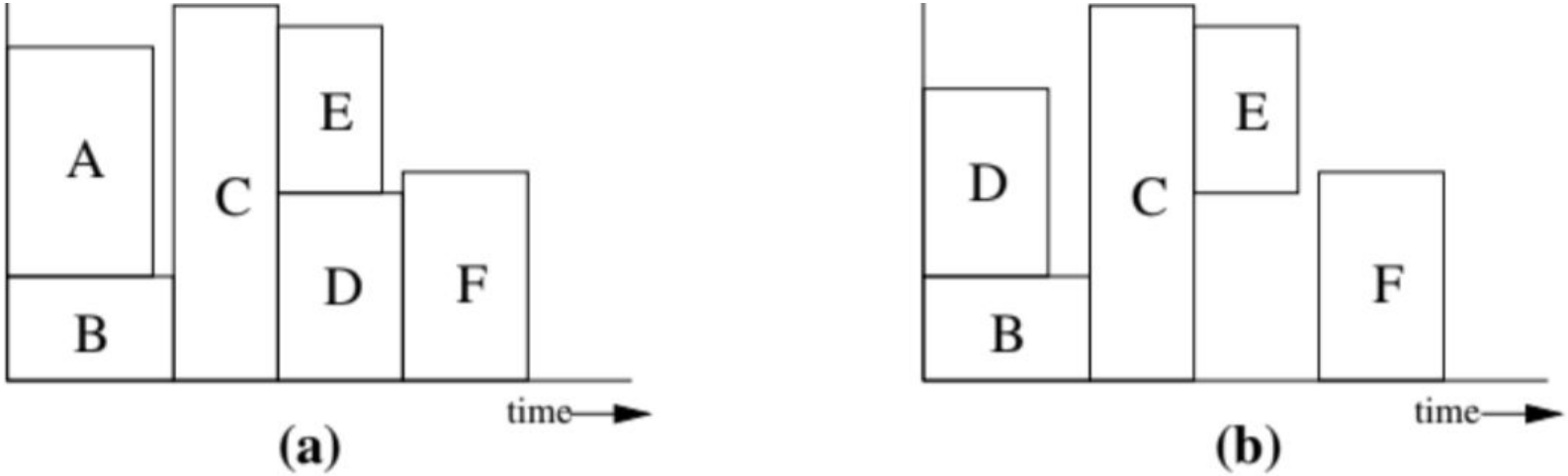


Fig. 3. Example where the DC algorithm deliberately leaves a hole in the profile. (a) Initial profile before job A terminates early. (b) Profile after compression.

Experimental Results

Name	Full file name	# jobs	% w/ estimates
CTC-SP2	CTC-SP2-1996-2.1-cln.swf	77,222	99.99
DAS2-fs0	DAS2-fs0-2003-1.swf	219,571	100
DAS2-fs1	DAS2-fs1-2003-1.swf	39,348	100
DAS2-fs2	DAS2-fs2-2003-1.swf	65,380	100
DAS2-fs3	DAS2-fs3-2003-1.swf	66,099	100
DAS2-fs4	DAS2-fs4-2003-1.swf	32,952	100
HPC2N	HPC2N-2002-1.1-cln.swf	202,876	100
KTH-SP2	KTH-SP2-1996-2.swf	28,489	100
LANL-CM5	LANL-CM5-1994-3.1-cln.swf	122,057	90.75
LLNL-Atlas	LLNL-Atlas-2006-1.1-cln.swf	38,143	84.85
LLNL-Thunder	LLNL-Thunder-2007-1.1-cln.swf	118,754	32.47
LPC-EGEE	LPC-EGEE-2004-1.2-cln.swf	220,679	100
SDSC-BLUE	SDSC-BLUE-2000-3.1-cln.swf	223,669	100
SDSC-DS	SDSC-DS-2004-1.swf	85,006	100
SDSC-SP2	SDSC-SP2-1998-3.1-cln.swf	54,041	99.94

Fig. 4. Traces used in simulations

Evaluated using an event-based simulator running traces from the Parallel Workloads Archive

Increasing Responsiveness - Shortest Jobs First

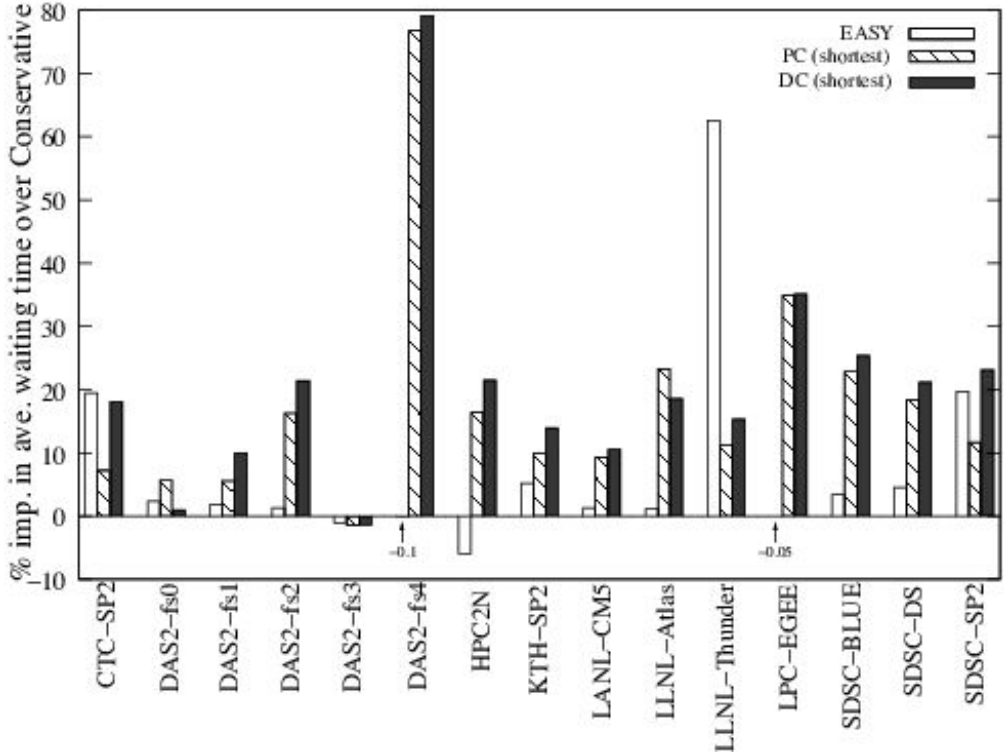


Fig. 5. Average waiting time relative to Conservative

Future research

- Understand why the algorithms perform better on some traces than others and distinguish between the algorithms
- Further explore the flexibility in job selection
- Further explore the flexibility in timing
- Ways to estimate job length quickly before adding it to the queue