Chapter 10
Class A Results and Analysis

10.1 Class A Simulations

Class A simulations describe situations featuring intermittent workloads. In Simulation2, two users submit equal but complementary workloads. As the number of users increases—to 6 in Simulation6, 11 in Simulation10, and 21 in Simulation14—one user's submission rate stays constant while the workload from the second user (5000 tasks) is distributed among the rest of the users in the system. In Simulation6, users 1–5 submit 1000 tasks each and 5000 total, while user 6 submits 5000 tasks on their own. In Simulation10, users 1–10 submit 500 tasks each, while user 11 submits 5000 tasks on their own. In Simulation14, users 1–20 submit 250 tasks each, while user 21 submits 5000 tasks on their own.

10.2 Class A Results

When the workloads are equal but complementary, as in Simulation2, a difference in performance of less than 2% is observed between fair-share and Rawlsian Fair scheduling methods. As the gap between the largest user and the smallest user increases, however, Rawlsian Fair produces an average performance that is approximately 80% better than that of fair-share (Table 10.1). The average total delay remains constant, confirming the fairness of Rawlsian Fair.

10.2.1 Simulation 2 Analysis and Results

In simulation 2, two users submitted different loads. User 1 submitted 1000 tasks/PC up to a total of 5000 tasks, and user 2 submitted 500 tasks/PC up to a total of 5000

A. Sedighi, M. Smith, *Fair Scheduling in High Performance Computing Environments*, https://doi.org/10.1007/978-3-030-14568-2_10

Table 10.1 Class A Submission type average delay

Class A average delay				
	Simulation 2	Simulation 6	Simulation 10	Simulation 14
User under test	User 1	Users 1–5	Users 1–10	Users 1–20
Fairshare average delay (ms)	62,891	61,824	66,873	73,617
Average Rawlsian Fair delay (ms)	61,455	34,783	34,783	34,783
Rawlsian Fair delay per user (ms)	61,455	User 1: 29,321	User 1: 28,641	User 1:28,641
BH = *inf*		User 2: 32,041	User 2: 30,001	User 2:28,641
		User 3: 34,761	User 3: 38,161	User 3:31,361
		User 4: 37,481	User 4: 39,521	User 4:32,721
		User 5: 40,311	User 5: 41,101	User 5:34,081
			User 6: 34,081	User 6:35,441
			User 7: 35,441	User 7:30,001
			User 8: 36,801	User 8:30,001
			User 9: 32,721	User 9:31,361
			User 10:31,361	User 10:41,101
				User 11:39,521
				User 12:39,521
				User 13:38,161
				User 14:38,161
				User 15:36,801
				User 16:36,801
				User 17:35,441
				User 18:34,081
				User 19:32,721
				User 20:41,101

(continued)

Table 10.1 (continued)

Class A average delay				
	Simulation 2	Simulation 6	Simulation 10	Simulation 14
Average improvement	<2%	86%	95%	114%
Improvement per user	<2%	User 1: 118%	User 1: 133%	User 1: 157%
		User 2: 99%	User 2: 123%	User 2: 157%
		User 3: 84%	User 3: 75%	User 3: 135%
		User 4: 70%	User 4: 69%	User 4: 125%
		User 5: 58%	User 5: 63%	User 5: 116%
			User 6: 96%	User 6: 108%
			User 7: 89%	User 7: 145%
			User 8: 82%	User 8: 145%
			User 9: 104%	User 9: 135%
			User 10: 113%	User 10: 79%
				User 11: 86%
				User 12: 86%
				User 13: 93%
				User 14: 93%
				User 15: 100%
				User 16: 100%
				User 17: 108%
				User 18: 116%
				User 19: 125%
				User 20: 79%

tasks. The number of available CPU's was set at 100 for the duration of the simulation (Fig. 10.1).

Simulation 2 featured 2 users and was run for all BH options and fair-share (Figs. 10.2 and 10.3).

The red line traversing both figures represents the task delay incurred using the fair-share scheduler. As expected, fair-share normalized the delay for both users, irrespective of their different submission profiles. Very little difference was observed at the micro-level between fair-share and Rawlsian Fair, but the results are interesting nonetheless. As Fig. 10.1 shows, the submission is interweaving amongst the two users. As such, each bucket contains tasks from only one user. User 2's tasks came in at a lower rate, and with varying of the BH sizes, user 2's tasks were lowered task delay to start (red circle in Fig. 10.3). This was particularly true for BH = inf, at which user 2, who had a lower task-submission rate, started off with a lower task delay. This, in turn, caused user 1's tasks to fall behind, at which point they were

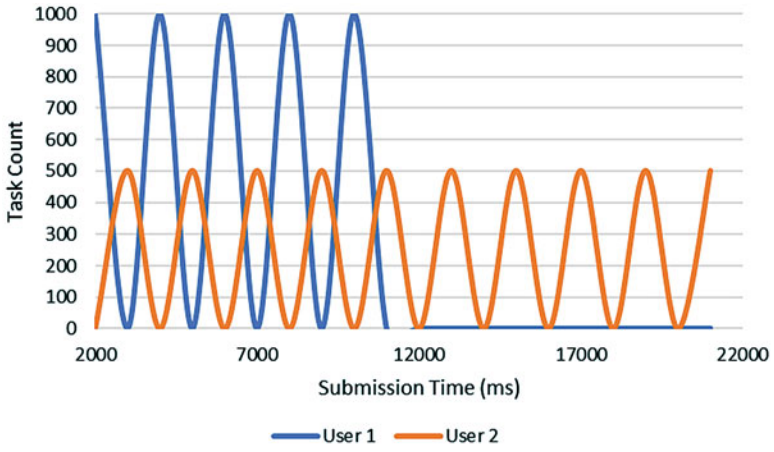

Fig. 10.1 Task submission for simulation 2

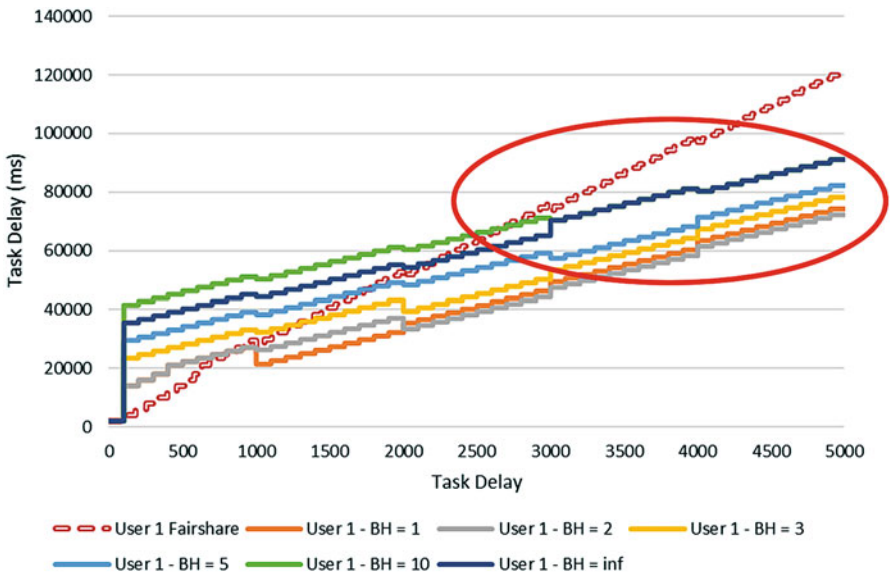

Fig. 10.2 User 1's task delay for all BH sizes in simulation 2

assigned greater seniority and were executed ahead of user 2's tasks (as is indicated by the red circle in Fig. 10.3).

High BH values, such as 10 or *inf*, were the best-case scenarios. In other cases, user 2's tasks lagged in execution behind user 1's because of how they were sorted into buckets. Once each user had submitted all of their 5000 tasks, the FCFS aspect of the scheduler took over. Table 10.2 compares the results obtained with Rawlsian Fair and fair-share in simulation 2.

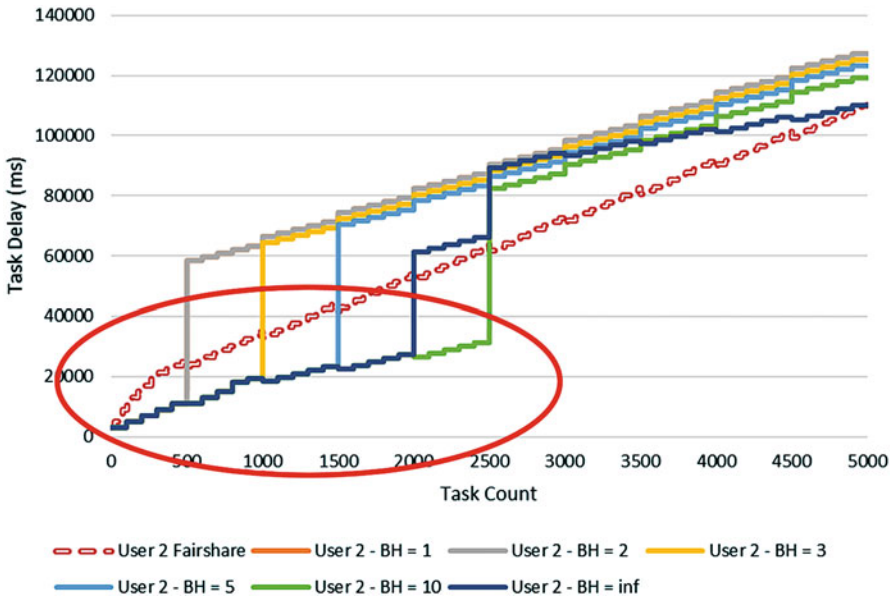

Fig. 10.3 User 2's task delay for all BH sizes in simulation 2

Table 10.2 Comparison of the delays generated by Fair-Share (FS) and Rawlsian Fair (RF) in simulation 2

	FS delay (ms)	RF delay (ms) BH = 1	RF delay (ms) BH = 2	RF delay (ms) BH = 3	RF delay (ms) BH = 5	RF delay (ms) BH = 10	RF delay (ms) BH = inf
User 1	62,891	42,133	41,913	47,955	53,615	64,935	61,455
vs FS	N/A	49%	50%	31%	17%	−3%	2%
User 2	61,824	84,221	84,221	78,069	72,079	60,099	63,029
vs FS	N/A	−27%	−27%	−21%	−14%	3%	−2%

10.2.2 Simulation 6 Analysis and Results

In simulation 6, one user submitted 5000 tasks and five other users submitted 1000 tasks each (Fig. 10.4). The difference between simulation 6 and simulation 2 is that in simulation 6, users 1–5 submitted simultaneously at a rate of 200 tasks for every two PC ticks. In Simulation 6, the performance level that users 1–5 experienced under Rawlsian Fair with a minimum BH value of 1 (Table 10.3) exceeded the performance level they experienced under fair-share by as much as 170%. This was mainly because with BH = 1, only the size of the immediate bucket was taken into account. The remainder of this section compares the performance levels of Rawlsian

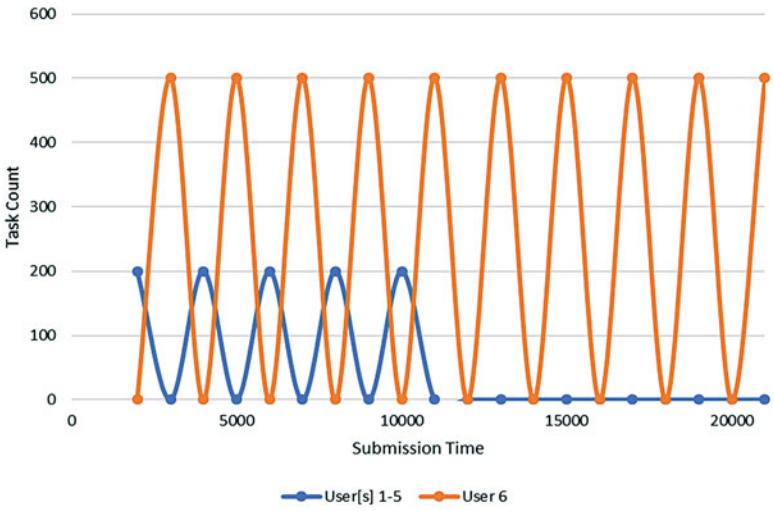

Fig. 10.4 The task submissions of all 6 users in simulation 6

Table 10.3 Comparison of the delays generated by Fair-Share (FS) and Rawlsian Fair (RF) in simulation 6

	FS delay (ms)	RF delay (ms) BH = 1	RF delay (ms) BH = 2	RF delay (ms) BH = 3	RF delay (ms) BH = 5	RF delay (ms) BH = 10	RF delay (ms) BH = inf
User 1	63,883	23,481	27,821	31,421	35,041	29,321	29,417
vs Fairshare	N/A	172%	130%	103%	82%	118%	121%
User 6	62,635	85,891	87,091	83,381	83,271	90,251	89,701
vs Fairshare	N/A	−27%	−28%	−25%	−25%	−31%	−30%

Fair and fair-share for users 1–5, and again for user 6. The other comparison is amongst users 1–5 for a given simulation parameter.

Figure 10.6 shows the task delays that user 1 experienced with the different BH values. User 1 submitted their tasks first, followed by user 2, and so on. This is important because with time, the FCFS aspect of the Rawlsian Fair scheduler negatively affects tasks with the same seniority. Every user was still better off than they would have been under fair-share, but user 1's tasks were completed first (Fig. 10.5). Rawlsian Fair pushed out user 6 in favor of the smaller users (users 1–5). This feature delayed user 6, which is apparent in Fig. 10.7. In every case, the dotted red line represents the task delay generated by the Fair-Share scheduler.

Table 10.3 shows the different performance levels that user 1 experienced as the BH value varied. User 6, the largest user, experienced a level of performance that

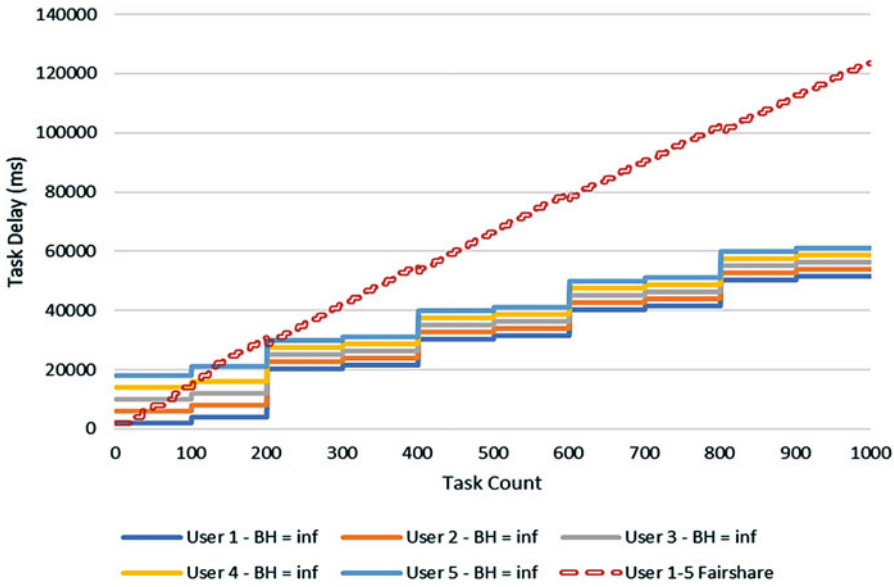

Fig. 10.5 The task delays experienced by all users with BH = *inf* and Fair-Share in simulation 6

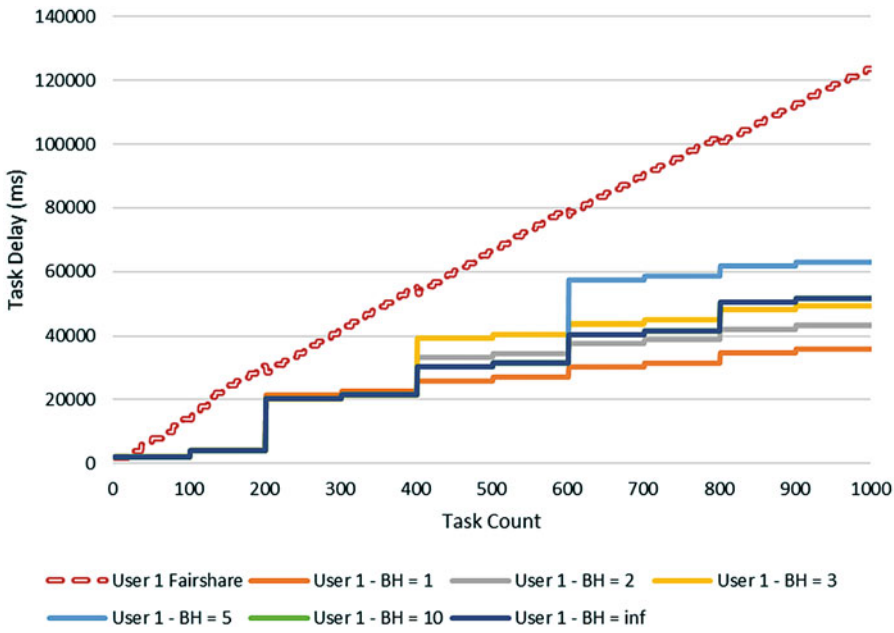

Fig. 10.6 User 1's task delays with Fair-Share and all BH values in simulation 6

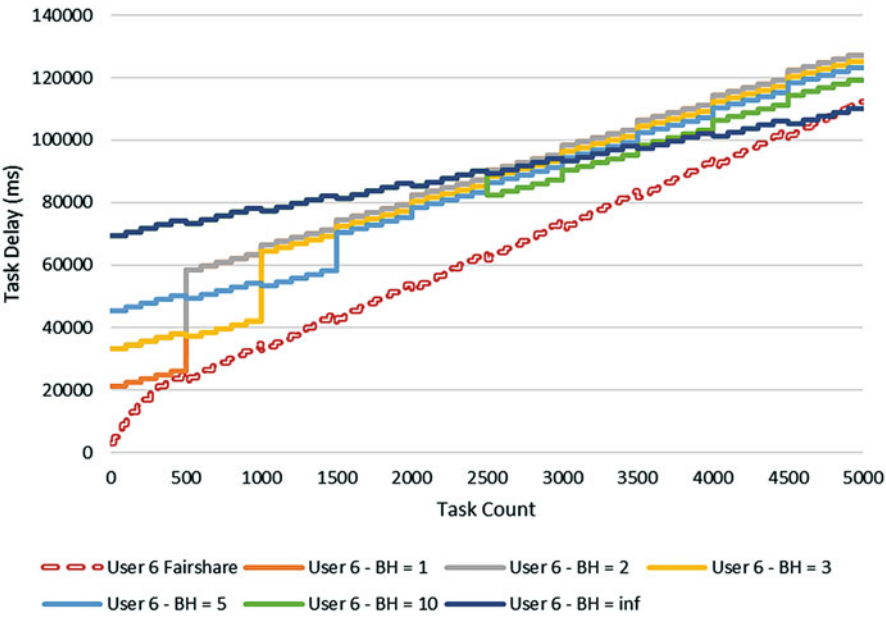

Fig. 10.7 User 6's task delays with Fair-Share and all BH values in simulation 6

was expectedly worse than those of the other users because users 1–5 were given precedence by the Rawlsian Fair scheduler.

10.2.3 Simulations 10 and 14 Results and Analysis

Simulation 10 featured 11 users, and simulation 14 featured 21 users. The total number of tasks submitted was the same in both simulations, but the number of tasks each user submitted decreased as the number of users increased. In simulation 10, users 1–10 submitted 100 tasks every two PCs up to a total of 5000 tasks, while user 11 submitted 500 tasks every two PCs up to a total of 5000 tasks (Fig. 10.8). In Simulation14, users 1–20 each submitted 50 tasks every two PCs, while user 21 submitted 500 tasks at a time up to a total of 5000 tasks (Fig. 10.10). In simulation 10, users 1–10 submitted their tasks to the same bucket simultaneously, starting with bucket 2. As a result, bucket 2 held 100 tasks from each of ten users, for a total of 1000 tasks. Users 1–20 in simulation 14 submitted their tasks in the same way: each submitted 50 tasks at a time, so bucket 2 held 1000 tasks. In both simulations, bucket 3 held 500 tasks submitted by the larger user.

In both simulations, the larger user—user 11 in simulation 10 and user 21 in simulations 14—experienced delays with Rawlsian Fair that exceeded the delays they experienced with traditional fair-share. The same pattern was observed in simulation 6. The delays the larger users incurred in both cases were almost identical

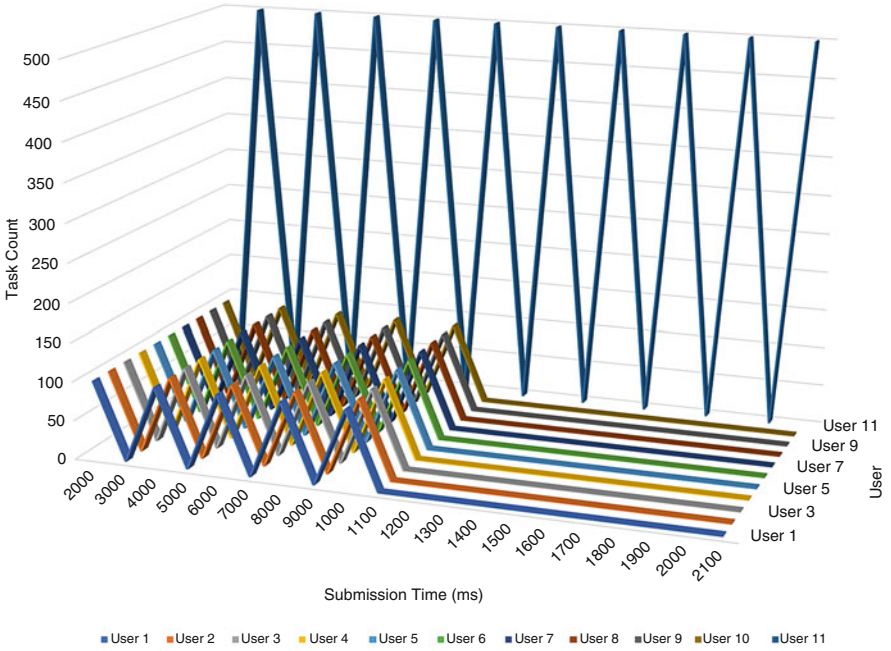

Fig. 10.8 The task submissions of all 11 users in simulation 10

(Figs. 10.13 and 10.14) because in both cases, the scheduler did not send the larger user's task for execution until all of the smaller users' tasks had been executed. Figure 10.12 shows a close-up of the of one of the smaller users in simulation 10 (user 1), and Fig. 10.11 shows a close-up of the of one of the smaller users in simulation 14. Rawlsian Fair scheduling was able to pick this smaller user from the others in the system and execute their tasks earlier than Fair-Share would have. The different Bucket-History (BH) sizes affected the task delay only after each user's submission number had exceeded the minimum number of buckets required for BH to be valid—with BH = 3, there must be at least three buckets, and so on. The red circle in Fig. 10.12 shows this effect in simulation 10, and the red circle in Fig. 10.11 shows this effect in simulation 14. The delays incurred by the smaller users were due largely to the number of available resources (100 CPU's), which determined the number of tasks that could be executed simultaneously. This is apparent in Fig. 10.9 (for Simulation10) and Fig. 10.15 (for Simulation14), which show the task delays experienced by all of the users. The task delays experienced by the smaller users are similar because the scheduler picked the largest seniority in the order of arrival before moving on to the next user on a first-come-first-served basis.

Table 10.4 compares the performance characteristics of one of the smaller users (user 1) to those of the larger user (user 11). With every BH value in simulation

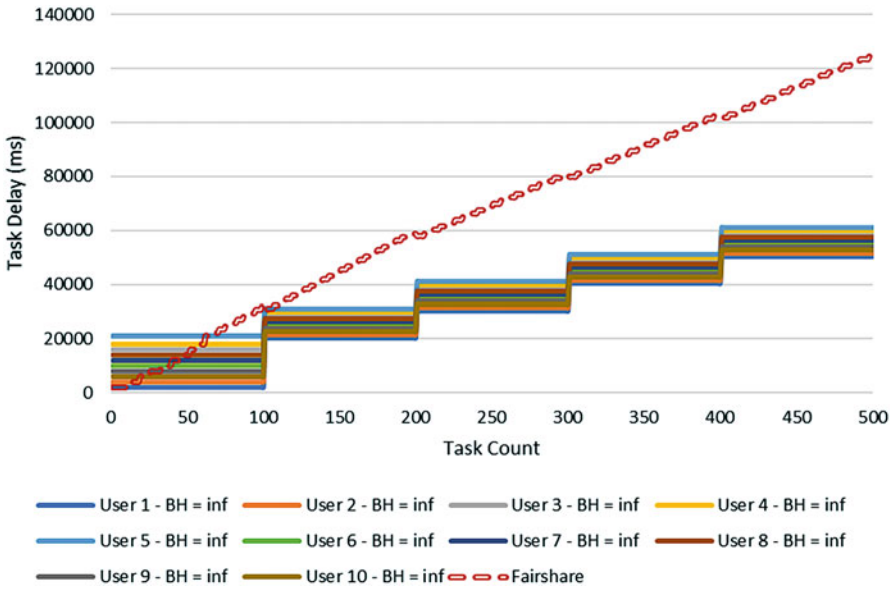

Fig. 10.9 The task delays experienced by users 1–10 with BH = *inf* in simulation 10

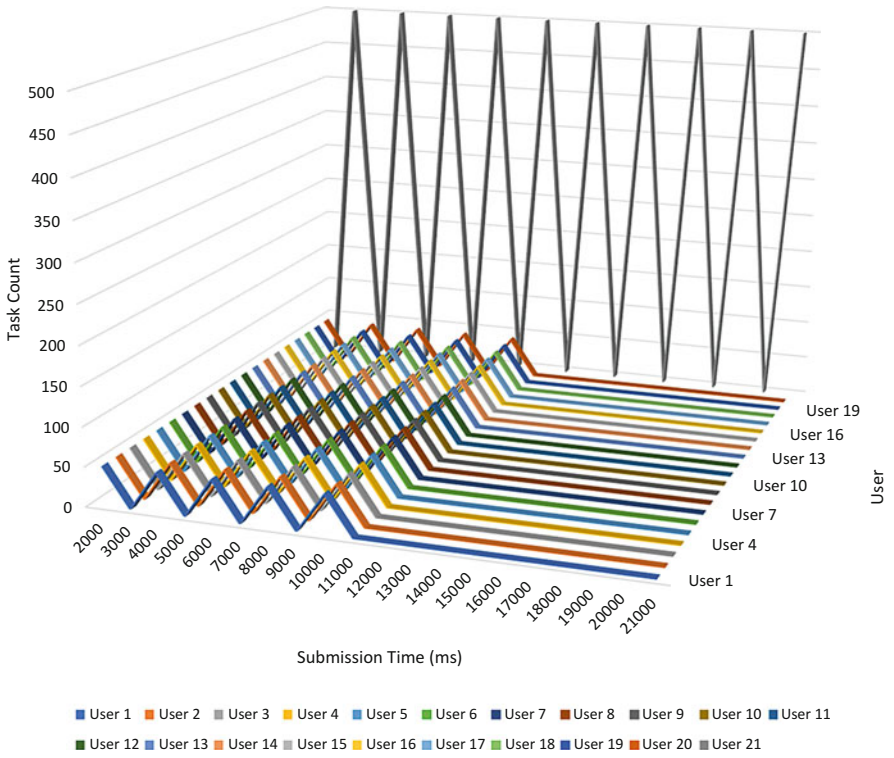

Fig. 10.10 The task submissions of all 21 users in simulation 14

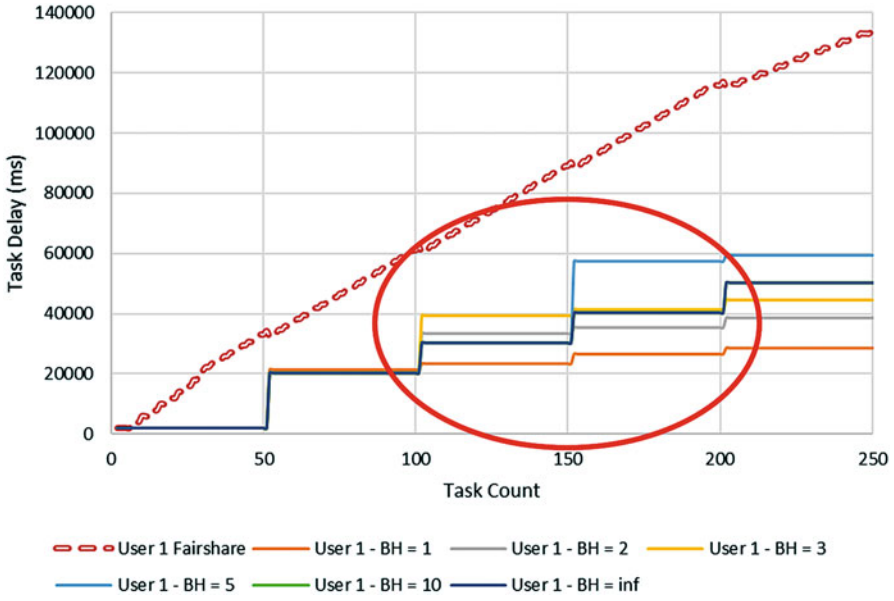

Fig. 10.11 Task delay for user 1 for all BH values for simulation 14

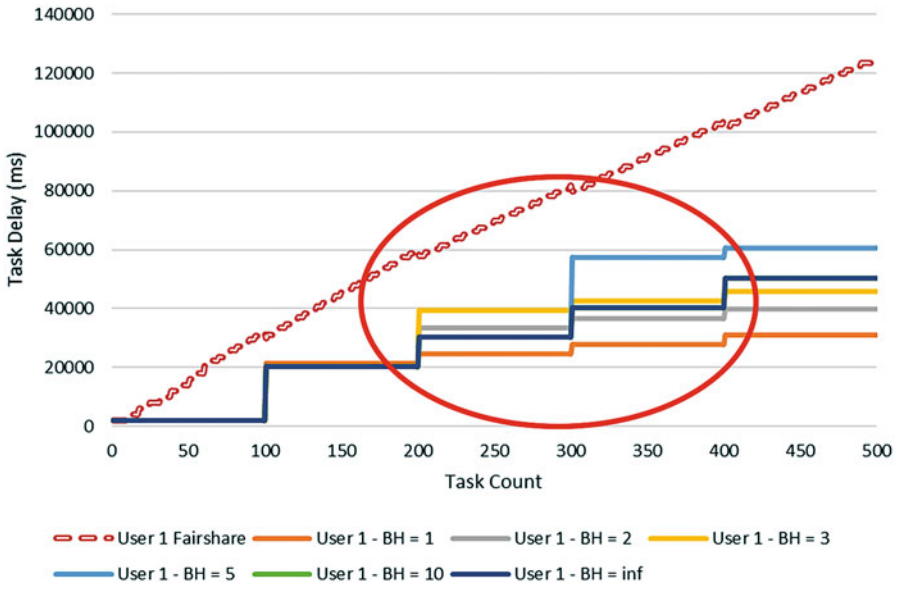

Fig. 10.12 User 1's task delays with all BH values in simulation 10

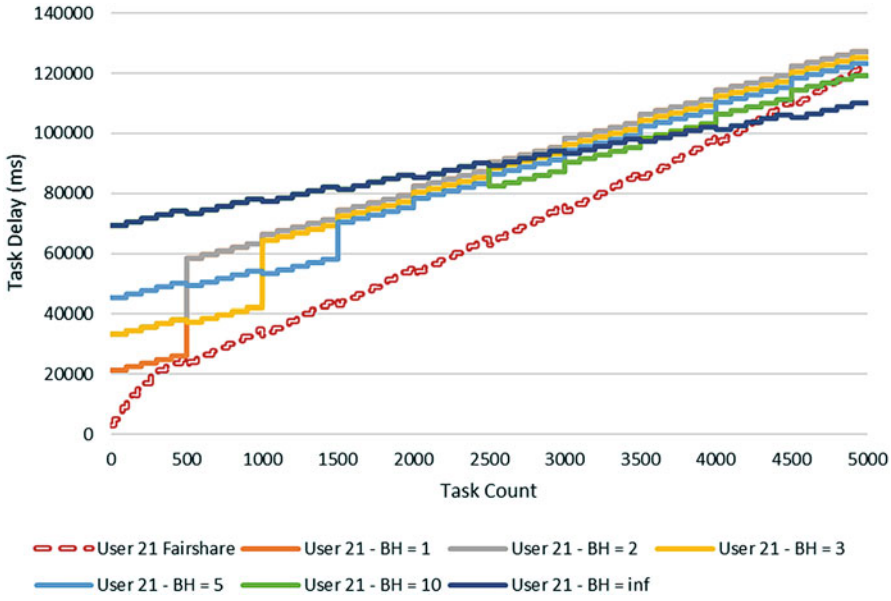

Fig. 10.13 User 21's task delays with all BH values in simulation 14

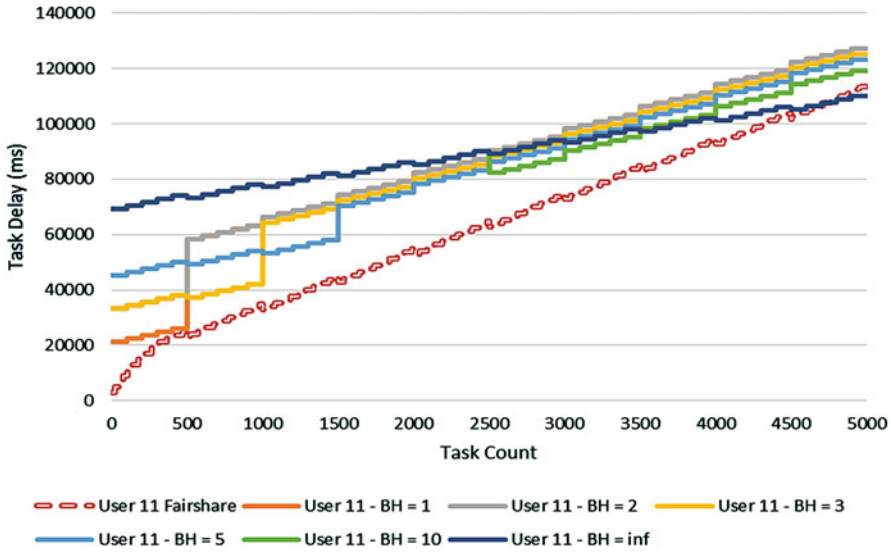

Fig. 10.14 User 11's task delays with all BH values in simulation 10

10, the smaller users (1–10) saw an increase in performance. All of these users saw the most improvement with BH = *inf* (Table 10.1), at which users 1–10 saw an average increase in performance greater than 95%.

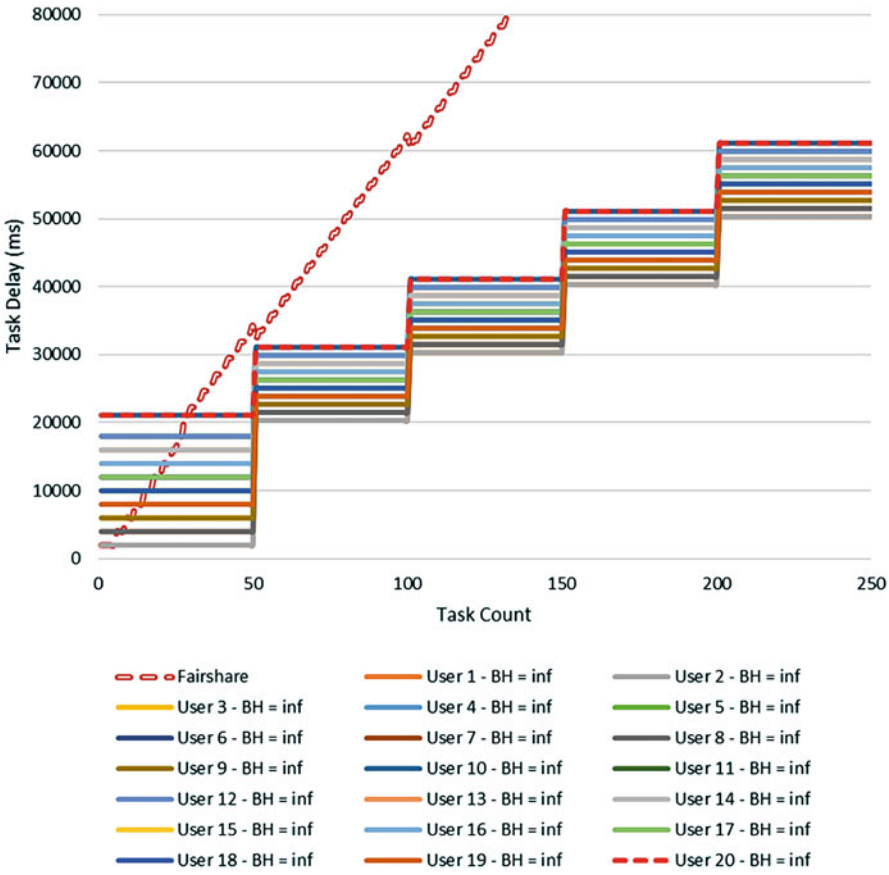

Fig. 10.15 The task delays of users 1–20 with BH = *inf* in simulation 14

Table 10.4 Comparison of the delays experienced by user 1 and user 11 under Fair-Share (FS) and Rawlsian Fair (RF) in simulation 10

	FS delay (ms)	RF delay (ms) BH = 1	RF delay (ms) BH = 2	RF delay (ms) BH = 3	RF delay (ms) BH = 5	RF delay (ms) BH = 10	RF delay (ms) BH = inf
User 1	67,133	21,361	26,421	30,021	34,121	28,641	28,641
vs Fairshare	N/A	214%	154%	124%	97%	133%	133%
User 11	63,302	85,891	87,091	83,381	83,271	90,251	89,701
vs Fairshare	N/A	−26%	−27%	−24%	−24%	−30%	−29%

Table 10.5 Comparison of the delays experienced by user 1 and user 21 under Fair-Share (FS) and Rawlsian Fair (RF) in simulation 14

	FS delay (ms)	RF delay (ms) BH = 1	RF delay (ms) BH = 2	RF delay (ms) BH = 3	RF delay (ms) BH = 5	RF delay (ms) BH = 10	RF delay (ms) BH = inf
User 1	73,617	20,401	25,941	29,541	33,881	28,641	28,641
vs Fairshare	N/A	261%	184%	149%	117%	157%	157%
User 21	65,245	85,891	87,091	83,381	83,271	90,251	89,701
vs Fairshare	N/A	−24%	−25%	−22%	−22%	−28%	−27%

Table 10.5 compares the delay of a smaller user (user 1) in simulation 14 to that of the larger user (user 21). As expected, user 1 saw an increase in performance with every BH value, and all users saw performance increases with BH = *inf* (Table 10.1). The average improvement in performance was over 110%.