

PFLOCK Report

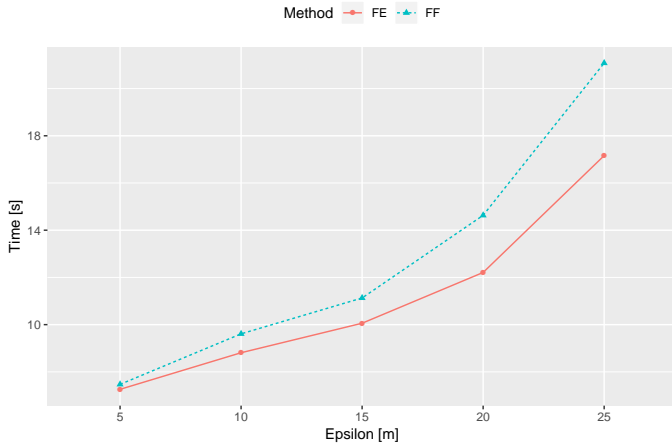
Andres Calderon

University of California, Riverside

January 24, 2020

New experiments...

LA 25K Dataset



Longest tasks in Time-by-time...

| timestamp | executor | duration | recordsRead | bytesRead | recordsWritten | bytesWritten | shuffleRead | shuffleWritten | taskId | index | stageId | jobId | phase | phaseTime | timeInstant | appId |
|-------------------------|--------------|----------|-------------|-----------|----------------|--------------|-------------|----------------|--------|-------|---------|-------|-------------------|-----------|-------------|-------|
| 2020-01-17 09:55:50.789 | 10.10.1.4:3 | 7.015 | 13 | 8216 | 0 | 0 | 0 | 0 | 252451 | 15 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.788 | 10.10.1.4:3 | 7.012 | 15 | 10328 | 0 | 0 | 0 | 0 | 252461 | 15 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.788 | 10.10.1.4:3 | 7.01 | 9 | 5608 | 0 | 0 | 0 | 0 | 252471 | 25 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.784 | 10.10.1.4:3 | 7.006 | 7 | 4752 | 0 | 0 | 0 | 0 | 252481 | 35 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.777 | 10.10.1.7:0 | 7.0 | 9 | 6584 | 0 | 0 | 0 | 0 | 252467 | 13 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.776 | 10.10.1.7:0 | 7.0 | 14 | 9888 | 0 | 0 | 0 | 0 | 252457 | 15 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.774 | 10.10.1.9:0 | 6.999 | 5 | 3384 | 0 | 0 | 0 | 0 | 252468 | 9 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.777 | 10.10.1.7:0 | 6.999 | 12 | 7680 | 0 | 0 | 0 | 0 | 252477 | 23 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.777 | 10.10.1.7:0 | 6.998 | 9 | 6176 | 0 | 0 | 0 | 0 | 252487 | 33 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.771 | 10.10.1.5:7 | 6.996 | 9 | 6432 | 0 | 0 | 0 | 0 | 252458 | 7 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.775 | 10.10.1.9:9 | 6.996 | 10 | 7112 | 0 | 0 | 0 | 0 | 252488 | 29 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.771 | 10.10.1.8:6 | 6.996 | 8 | 4880 | 0 | 0 | 0 | 0 | 252455 | 8 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.774 | 10.10.1.9:9 | 6.996 | 7 | 5600 | 0 | 0 | 0 | 0 | 252478 | 19 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.772 | 10.10.1.5:7 | 6.995 | 11 | 7984 | 0 | 0 | 0 | 0 | 252468 | 17 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.772 | 10.10.1.5:7 | 6.994 | 12 | 8296 | 0 | 0 | 0 | 0 | 252478 | 27 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.769 | 10.10.1.6:4 | 6.993 | 12 | 8176 | 0 | 0 | 0 | 0 | 252462 | 14 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.773 | 10.10.1.9:9 | 6.993 | 9 | 5816 | 0 | 0 | 0 | 0 | 252498 | 39 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.778 | 10.10.1.6:4 | 6.993 | 13 | 8960 | 0 | 0 | 0 | 0 | 252472 | 24 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.769 | 10.10.1.1:2 | 6.992 | 12 | 7448 | 0 | 0 | 0 | 0 | 252469 | 11 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.771 | 10.10.1.5:7 | 6.992 | 12 | 7952 | 0 | 0 | 0 | 0 | 252488 | 37 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.767 | 10.10.1.8:6 | 6.991 | 11 | 8120 | 0 | 0 | 0 | 0 | 252465 | 18 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.770 | 10.10.1.8:6 | 6.991 | 12 | 8328 | 0 | 0 | 0 | 0 | 252485 | 38 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.767 | 10.10.1.10:5 | 6.991 | 13 | 8672 | 0 | 0 | 0 | 0 | 252463 | 16 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.765 | 10.10.1.1:2 | 6.99 | 15 | 10368 | 0 | 0 | 0 | 0 | 252459 | 1 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.770 | 10.10.1.6:4 | 6.99 | 10 | 6272 | 0 | 0 | 0 | 0 | 252482 | 34 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.764 | 10.10.1.6:4 | 6.99 | 11 | 6848 | 0 | 0 | 0 | 0 | 252452 | 4 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.768 | 10.10.1.8:6 | 6.99 | 11 | 7304 | 0 | 0 | 0 | 0 | 252475 | 28 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.768 | 10.10.1.10:5 | 6.99 | 8 | 6216 | 0 | 0 | 0 | 0 | 252473 | 26 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.769 | 10.10.1.1:2 | 6.988 | 14 | 9488 | 0 | 0 | 0 | 0 | 252469 | 31 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |
| 2020-01-17 09:55:50.766 | 10.10.1.2:8 | 6.988 | 9 | 6256 | 0 | 0 | 0 | 0 | 252466 | 12 | 473235 | 0 | 4. Flocks updated | 10.68 | 49 | 0252 |

Figure: Top 30 longest tasks

Longest tasks in Time-window...

| timestamp | executor | duration | recordsRead | bytesRead | recordsWritten | bytesWritten | shuffleRead | shuffleWritten | taskId | index | stageId | jobId | phase | phaseTime | timeInstant | appId |
|-------------------------|--------------|----------|-------------|-----------|----------------|--------------|-------------|----------------|---------|-------|---------|-------|-------------------------|-----------|-------------|--------|
| 2020-01-17 12:09:51,570 | 10.10.1.10:5 | 4.675 | 4465 | 12537872 | 0 | 0 | 0 | 0 | 1316194 | 48 | 13296 | 0 | 1.Getting maximal disks | 19.11 | 46 | [0261] |
| 2020-01-17 12:03:53,765 | 10.10.1.7:8 | 5.14 | 2969 | 1641880 | 0 | 0 | 0 | 0 | 199668 | 44 | 8315 | 0 | 1.Getting maximal disks | 16.67 | 29 | [0261] |
| 2020-01-17 12:02:52,003 | 10.10.1.4:3 | 5.064 | 2862 | 1587632 | 0 | 0 | 0 | 0 | 179183 | 47 | 7436 | 0 | 1.Getting maximal disks | 15.69 | 26 | [0261] |
| 2020-01-17 12:09:50,381 | 10.10.1.10:5 | 5.028 | 3690 | 2071712 | 0 | 0 | 0 | 0 | 316190 | 44 | 13296 | 0 | 1.Getting maximal disks | 19.11 | 46 | [0261] |
| 2020-01-17 12:10:36,093 | 10.10.1.4:3 | 4.93 | 3562 | 2003208 | 0 | 0 | 0 | 0 | 329777 | 50 | 13882 | 0 | 1.Getting maximal disks | 16.92 | 48 | [0261] |
| 2020-01-17 12:10:14,089 | 10.10.1.6:4 | 4.869 | 3077 | 2075112 | 0 | 0 | 0 | 0 | 322952 | 47 | 13589 | 0 | 1.Getting maximal disks | 17.44 | 47 | [0261] |
| 2020-01-17 12:11:22,697 | 10.10.1.10:5 | 4.688 | 4296 | 2396616 | 0 | 0 | 0 | 0 | 343436 | 48 | 14468 | 0 | 1.Getting maximal disks | 17.92 | 58 | [0261] |
| 2020-01-17 12:11:21,993 | 10.10.1.10:5 | 4.669 | 3717 | 2156968 | 0 | 0 | 0 | 0 | 343435 | 47 | 14468 | 0 | 1.Getting maximal disks | 17.92 | 58 | [0261] |
| 2020-01-17 12:03:32,811 | 10.10.1.9:9 | 4.652 | 2854 | 1577992 | 0 | 0 | 0 | 0 | 1927905 | 44 | 8022 | 0 | 1.Getting maximal disks | 15.92 | 28 | [0261] |
| 2020-01-17 12:10:35,756 | 10.10.1.4:3 | 4.572 | 3922 | 2216120 | 0 | 0 | 0 | 0 | 329778 | 51 | 13882 | 0 | 1.Getting maximal disks | 16.92 | 48 | [0261] |
| 2020-01-17 12:03:12,499 | 10.10.1.7:8 | 4.428 | 2813 | 1601864 | 0 | 0 | 0 | 0 | 186868 | 41 | 7729 | 0 | 1.Getting maximal disks | 16.15 | 27 | [0261] |
| 2020-01-17 12:10:58,096 | 10.10.1.5:7 | 4.42 | 4265 | 2388672 | 0 | 0 | 0 | 0 | 336610 | 39 | 14175 | 0 | 1.Getting maximal disks | 17.52 | 49 | [0261] |
| 2020-01-17 12:02:31,302 | 10.10.1.2:1 | 4.308 | 2037 | 1606964 | 0 | 0 | 0 | 0 | 172207 | 50 | 7143 | 0 | 1.Getting maximal disks | 14.75 | 25 | [0261] |
| 2020-01-17 12:01:16,166 | 10.10.1.1:2 | 4.312 | 2665 | 1469992 | 0 | 0 | 0 | 0 | 144040 | 41 | 5971 | 0 | 1.Getting maximal disks | 15.3 | 21 | [0261] |
| 2020-01-17 12:10:58,651 | 10.10.1.5:7 | 4.289 | 3647 | 2053936 | 0 | 0 | 0 | 0 | 336609 | 38 | 14175 | 0 | 1.Getting maximal disks | 17.52 | 49 | [0261] |
| 2020-01-17 12:04:15,347 | 10.10.1.10:5 | 3.968 | 2860 | 1588424 | 0 | 0 | 0 | 0 | 206510 | 41 | 8608 | 0 | 1.Getting maximal disks | 15.77 | 30 | [0261] |
| 2020-01-17 12:09:06,163 | 10.10.1.4:3 | 3.877 | 3840 | 2107968 | 0 | 0 | 0 | 0 | 302444 | 45 | 12710 | 0 | 1.Getting maximal disks | 15.83 | 44 | [0261] |
| 2020-01-17 12:04:35,954 | 10.10.1.6:4 | 3.872 | 2978 | 1628744 | 0 | 0 | 0 | 0 | 213335 | 35 | 8901 | 0 | 1.Getting maximal disks | 16.05 | 31 | [0261] |
| 2020-01-17 12:06:55,700 | 10.10.1.6:4 | 3.83 | 2522 | 1421640 | 0 | 0 | 0 | 0 | 137974 | 41 | 5678 | 0 | 1.Getting maximal disks | 15.11 | 28 | [0261] |
| 2020-01-17 12:09:49,131 | 10.10.1.10:5 | 3.76 | 5448 | 1928064 | 0 | 0 | 0 | 0 | 316191 | 45 | 13296 | 0 | 1.Getting maximal disks | 19.11 | 46 | [0261] |
| 2020-01-17 12:07:40,113 | 10.10.1.8:6 | 3.662 | 3960 | 2133488 | 0 | 0 | 0 | 0 | 274867 | 33 | 11538 | 0 | 1.Getting maximal disks | 17.28 | 48 | [0261] |
| 2020-01-17 12:09:27,562 | 10.10.1.9:9 | 3.645 | 4002 | 2276672 | 0 | 0 | 0 | 0 | 309411 | 48 | 13903 | 0 | 1.Getting maximal disks | 16.04 | 45 | [0261] |
| 2020-01-17 12:01:35,481 | 10.10.1.7:8 | 3.636 | 2694 | 1462008 | 0 | 0 | 0 | 0 | 151687 | 38 | 6264 | 0 | 1.Getting maximal disks | 14.72 | 22 | [0261] |
| 2020-01-17 12:10:12,891 | 10.10.1.6:4 | 3.613 | 3635 | 2005056 | 0 | 0 | 0 | 0 | 322953 | 48 | 13589 | 0 | 1.Getting maximal disks | 17.44 | 47 | [0261] |
| 2020-01-17 12:06:35,377 | 10.10.1.8:6 | 3.553 | 3375 | 1872168 | 0 | 0 | 0 | 0 | 254283 | 50 | 10659 | 0 | 1.Getting maximal disks | 16.08 | 37 | [0261] |
| 2020-01-17 12:08:23,550 | 10.10.1.10:5 | 3.55 | 3074 | 2123664 | 0 | 0 | 0 | 0 | 280294 | 51 | 12124 | 0 | 1.Getting maximal disks | 17.62 | 42 | [0261] |
| 2020-01-17 12:08:45,130 | 10.10.1.5:7 | 3.422 | 3702 | 2068936 | 0 | 0 | 0 | 0 | 295330 | 45 | 12417 | 0 | 1.Getting maximal disks | 16.01 | 43 | [0261] |
| 2020-01-17 12:08:01,162 | 10.10.1.2:8 | 3.353 | 3011 | 2185690 | 0 | 0 | 0 | 0 | 281531 | 51 | 11831 | 0 | 1.Getting maximal disks | 16.33 | 41 | [0261] |
| 2020-01-17 12:07:39,244 | 10.10.1.8:6 | 3.34 | 3524 | 1991544 | 0 | 0 | 0 | 0 | 274863 | 29 | 11538 | 0 | 1.Getting maximal disks | 17.28 | 48 | [0261] |
| 2020-01-17 12:06:14,594 | 10.10.1.8:6 | 3.301 | 3288 | 1832672 | 0 | 0 | 0 | 0 | 247451 | 47 | 10366 | 0 | 1.Getting maximal disks | 15.36 | 36 | [0261] |

Figure: Top 30 longest tasks

What is next...

- ▶ Currently running an experiment varying the number of partitions in the finer and coarse grids in order to find which are the optimal values...
- ▶ Measure scale up performance in both alternatives...