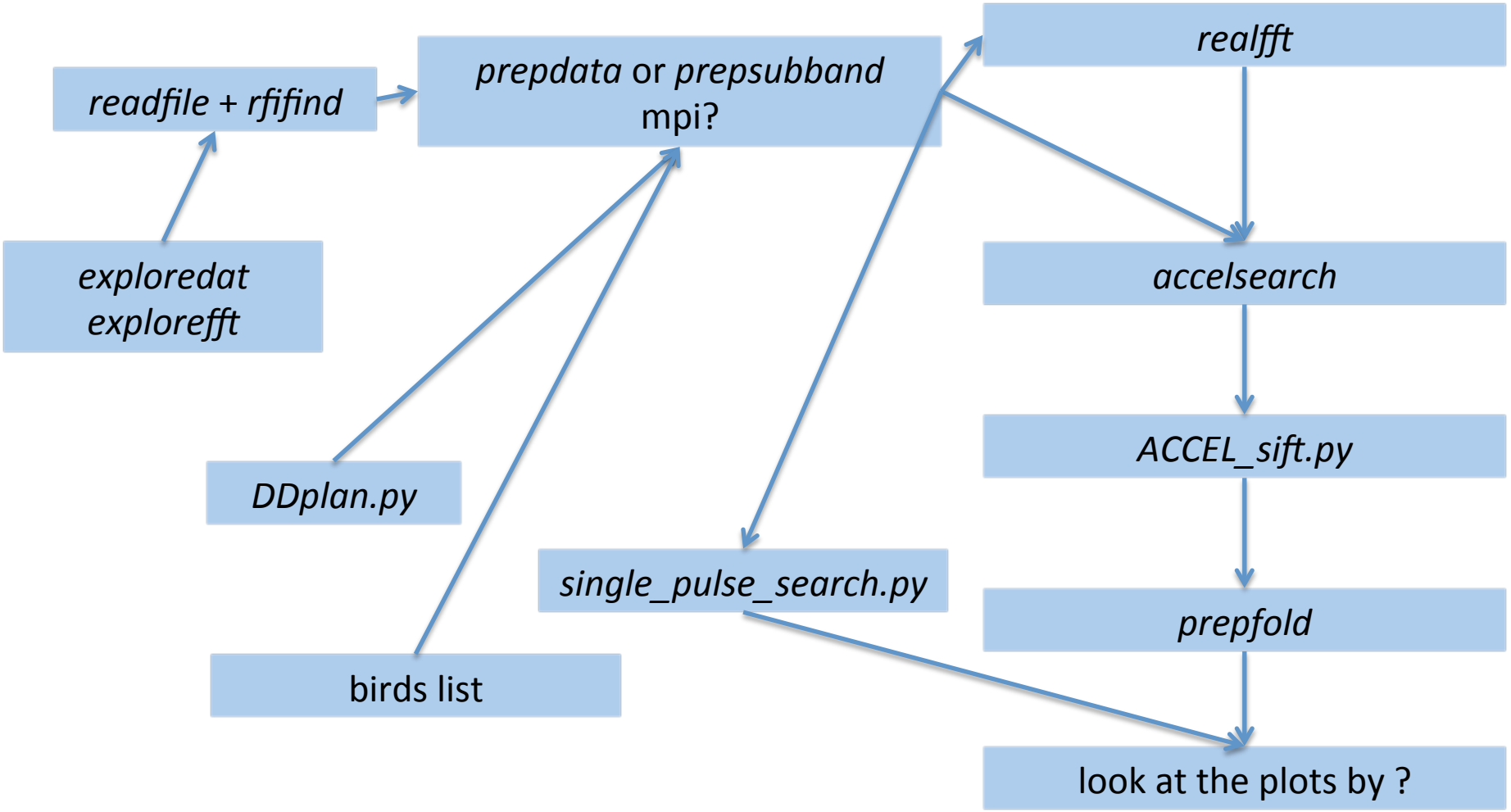
The Test of Running PRESTO Pulsar Search in Parallel

Pan, Zhichen (NAOC), Lei Qian(NAOC), Youling Yue(NAOC), ...

2017-06-28

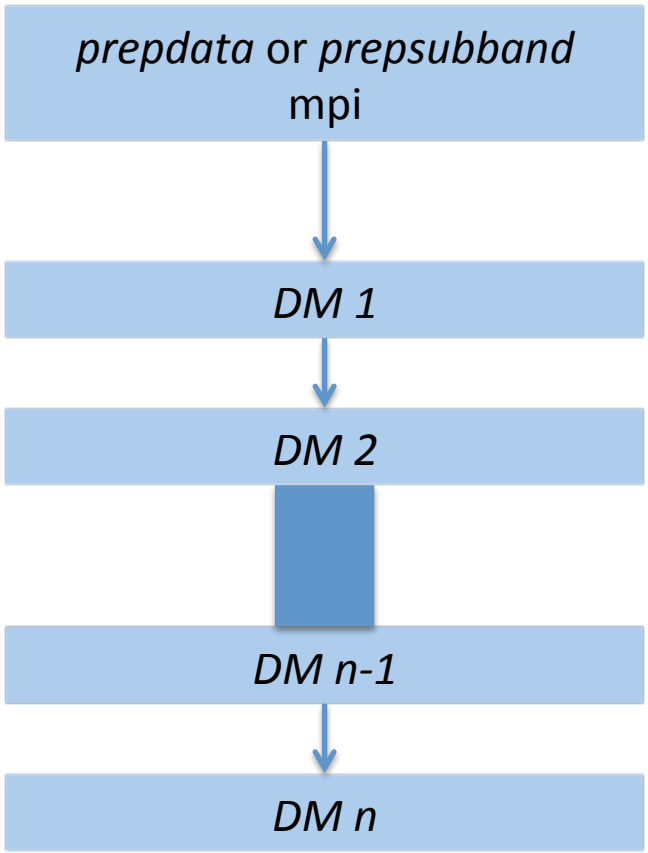
Basic PRESTO Pulsar Search

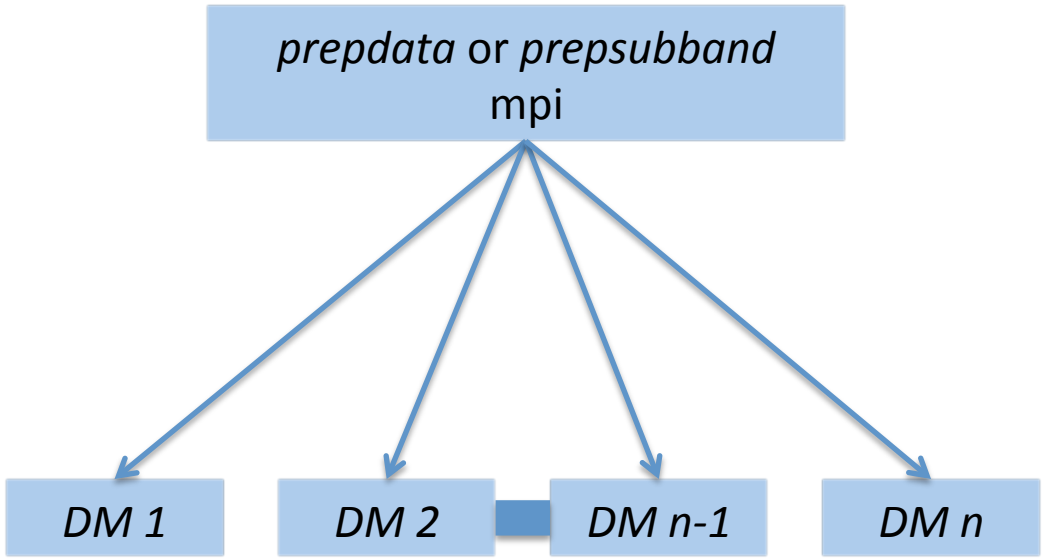
From PRESTO tutorial of Scott Ransom



Our PRESTO Pulsar Search – How to Run Faster

From PRESTO tutorial of Scott Ransom

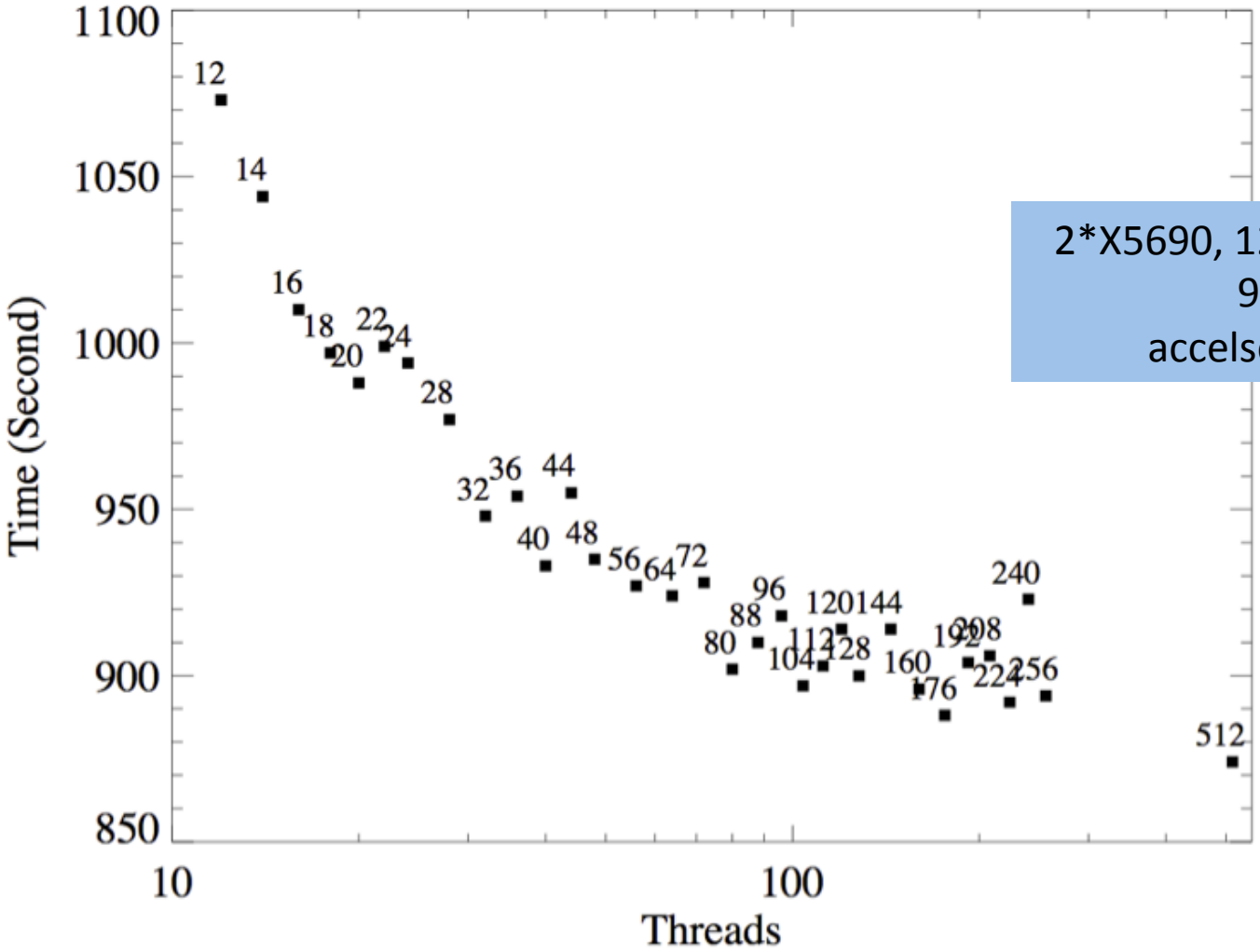




Same with realfft, accelsearch, and prepfold.

Our PRESTO Pulsar Search – Time Costing

A Simple Test for PMPS Data



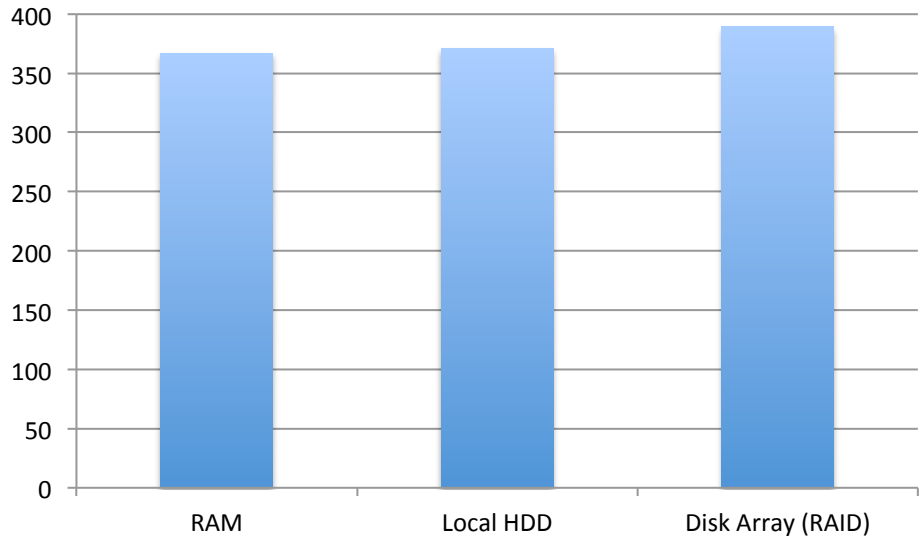
2*X5690, 12 cores, 24 threads
96 G RAM
accelsearch -zmax 0

Our PRESTO Pulsar Search – Time Costing

Q1, do different storages affect?

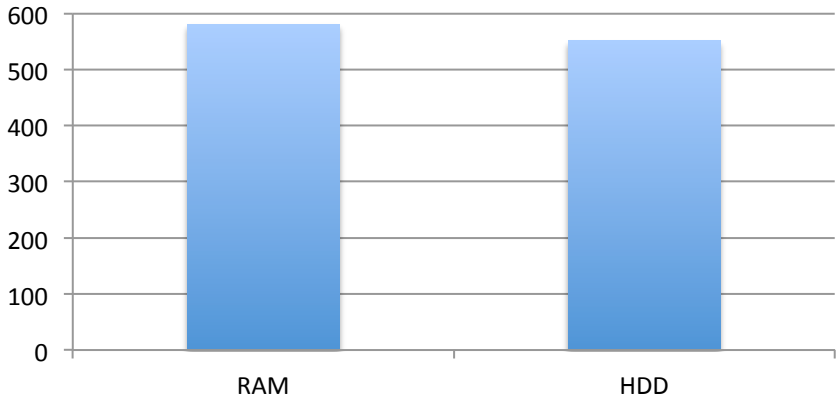
Storage types:

- Local HDD
- Local SSD
- RAM (/dev/shm)
- Disk array (RAID)
- Parallel file system
- Distributed file system
- .....
- .....



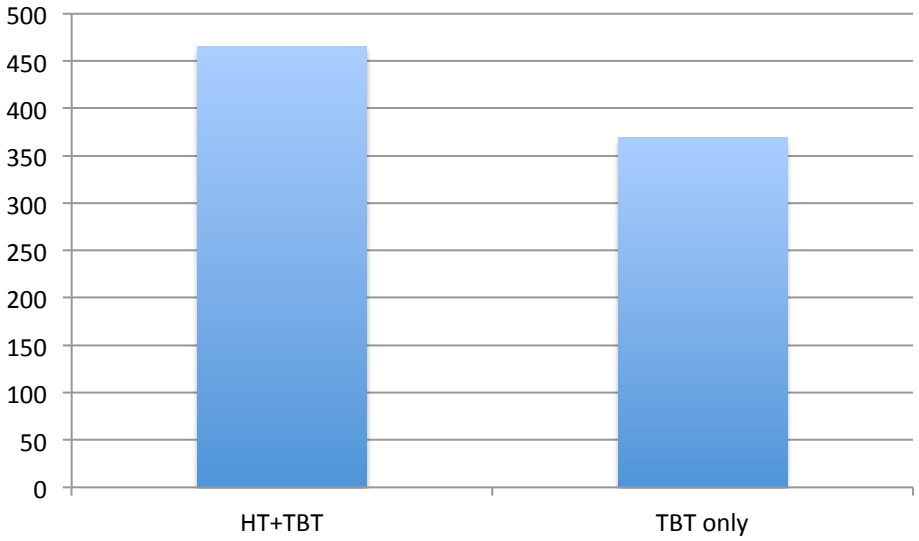
Best in reading, writing, and IOPs: RAM
Worst: local HDD

Test running on E5-2690 v2 *2 for FAST simulation data (upper left) and PMPS data (lower left)

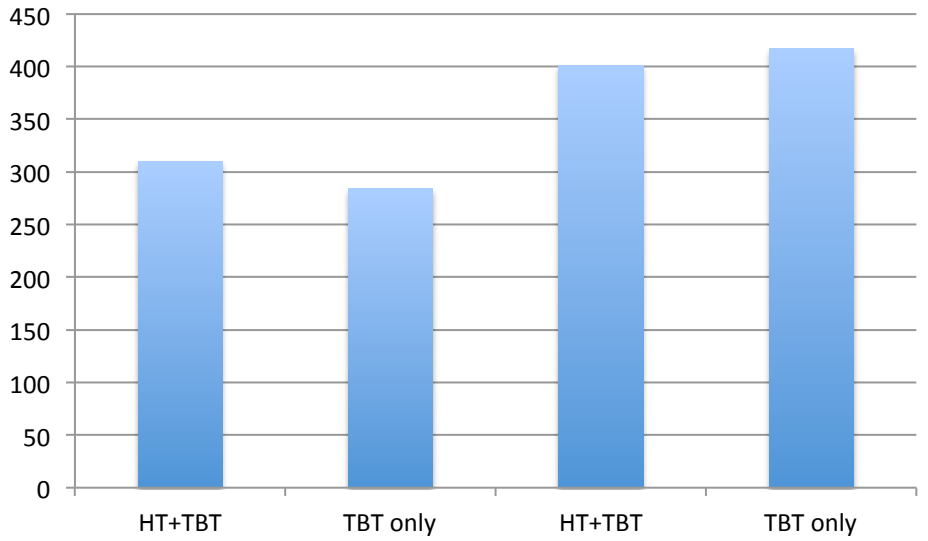


Our PRESTO Pulsar Search – Time Costing

Q2, Hyper Thread (HT) and Turbo Boost (TB)



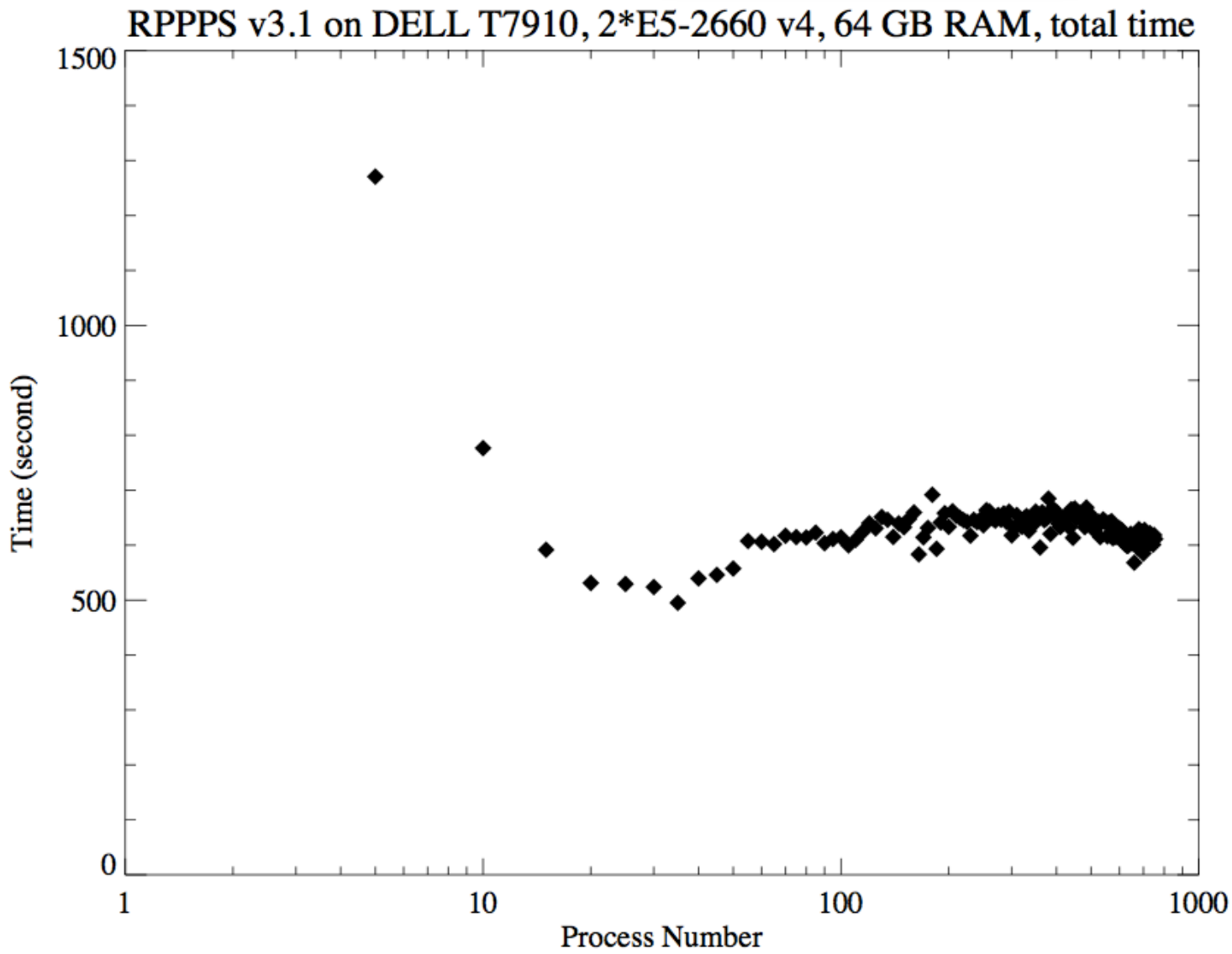
E5-2690 v4 *2, 512 GB DDR4 RAM
for FAST simulation data
non-acceleration search



E5-2680 v3 *2, 512 GB DDR3 RAM
for PMPS data
acceleration search (left, zmax=50)
and non-acceleration search (right, zmax=0)

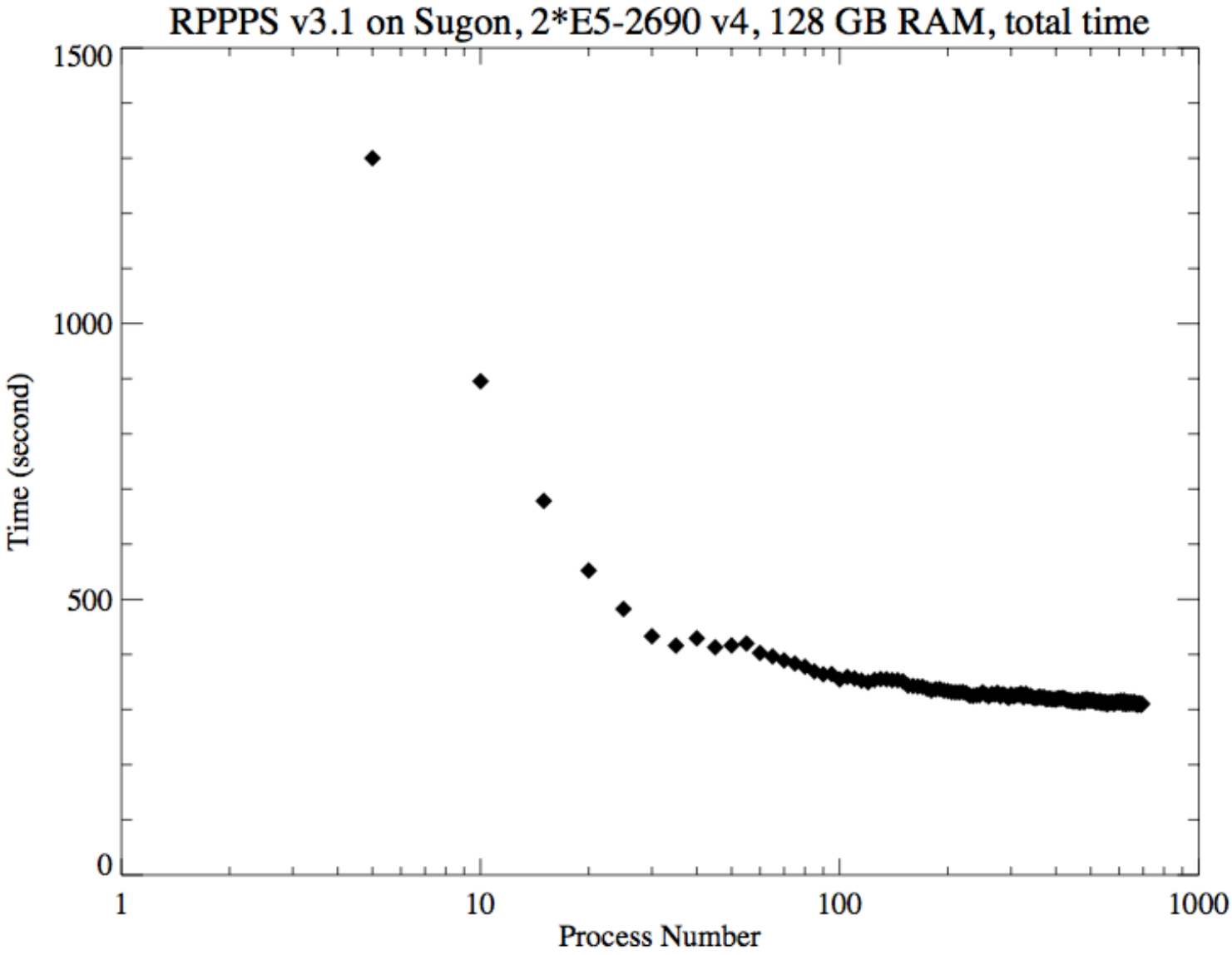
Our PRESTO Pulsar Search – Time Costing

Q3, Process Number to Time



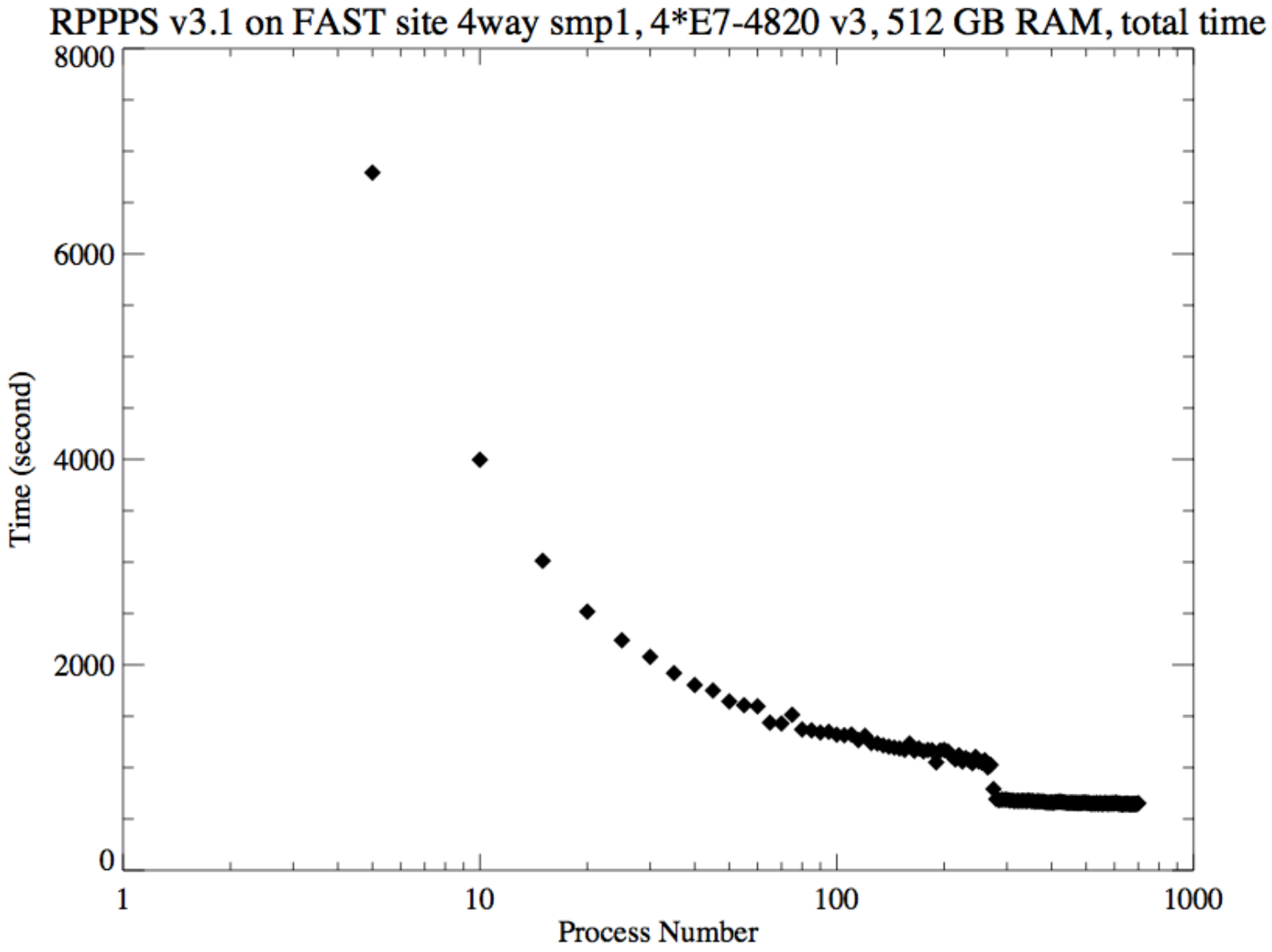
Our PRESTO Pulsar Search – Time Costing

Q3, Process Number to Time



Our PRESTO Pulsar Search – Time Costing

Q3, Process Number to Time



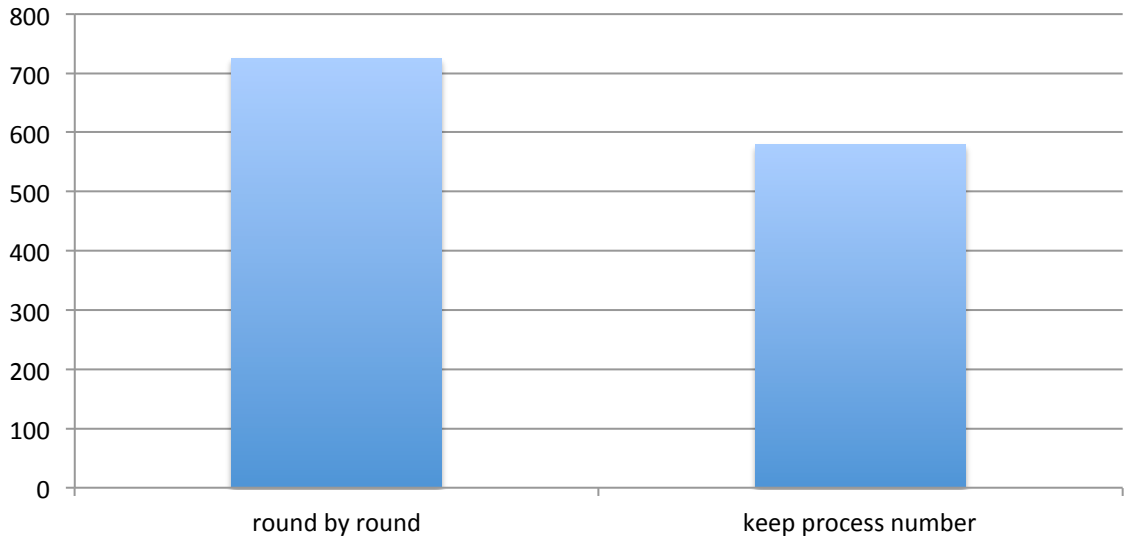
Our PRESTO Pulsar Search – Time Costing

Q4, Two Methods of Parallel

for **N** DM trials, **Round by round**:
1, start **n** process;
2, wait for these processes finish;
3, loop 1 and 2 until **N** DM trials finish;

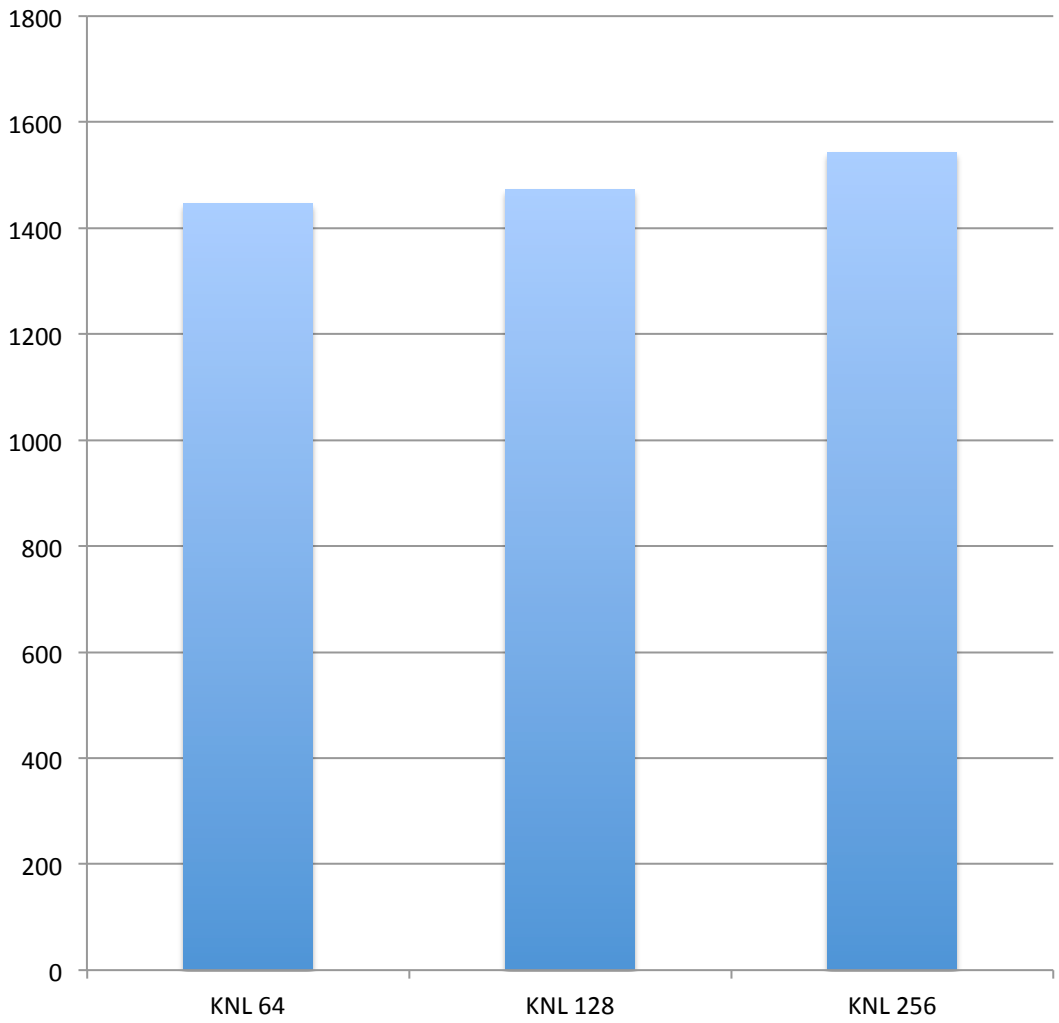
for **N** DM trials, **keep process nuber**:
1, start **n** process;
2, count process number;
3, if process number is less than **n**, start several processes;
4, loop 1 to 3 until **N** DM trials finish;

A test on PMPS data, acceleration search (zmax=50)



Our PRESTO Pulsar Search – Time Costing

Q5, How about KNL (Intel Many Core)



On a KNL node
64 cores @ 1.3 GHz, 265 threads
For FAST simulation data

On a 2-way E5-2690 v4 node
28 cores @ 2.6 GHz
< 420 second

Our PRESTO Pulsar Search – Time Costing

Q6, What is the Possible Shortest Time

1, *rfifind*, single process (single core);
2, *prepdata*, *realfft*, and *accelsearch*, running in parallel;

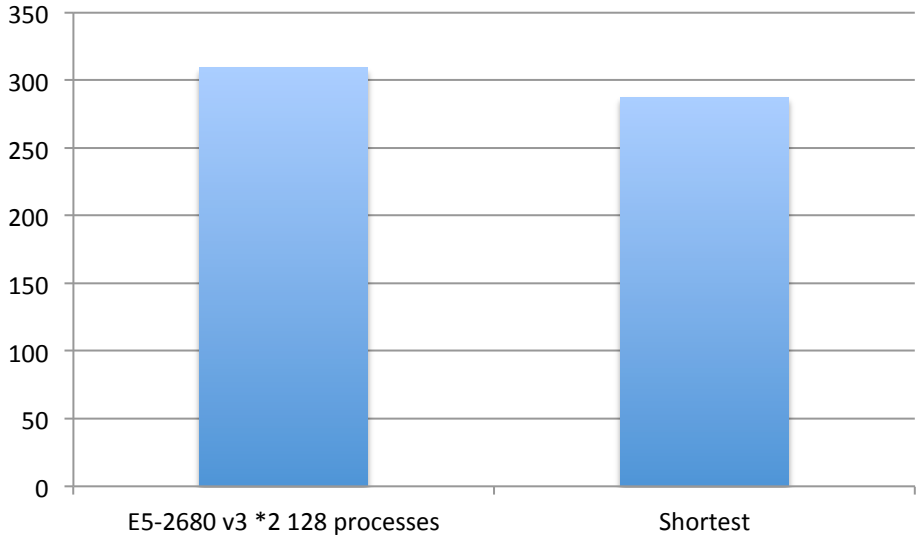
$$\boldsymbol{shortest\_time}$$

$$= \boldsymbol{rfifind}$$

$$+ \frac{(\boldsymbol{prepdata} + \boldsymbol{realfft} + \boldsymbol{accelsearch}) \times (\boldsymbol{number\ of\ DM\ trials})}{\boldsymbol{number\ of\ cores}}$$

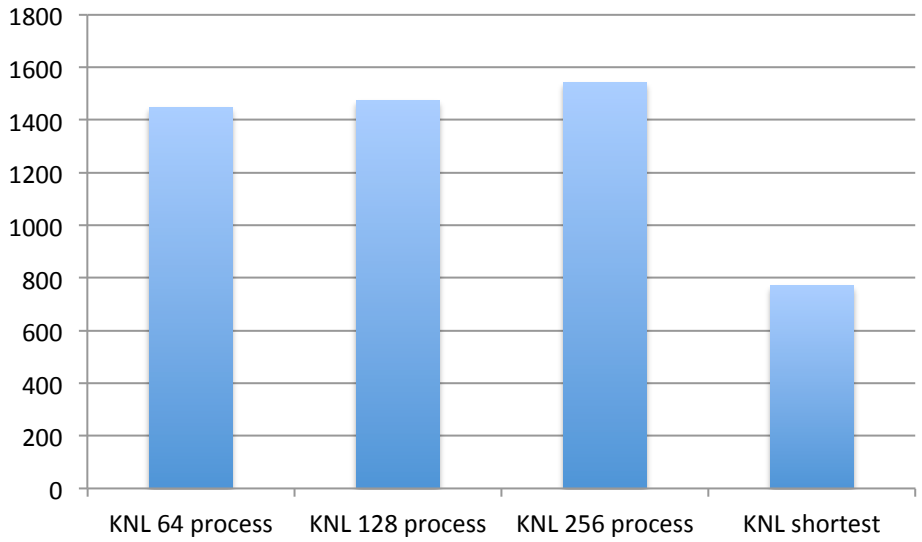
Our PRESTO Pulsar Search – Time Costing

Q6, What is the Possible Shortest Time



On 2-way servers
close to the shortest time

On KNL servers
much longer than the shortest time



Our PRESTO Pulsar Search – Parallel Results

1, to be an RPPPS (Reprocess Pipeline for Parkes Pulsar Survey) package, v3.2 now;

2, storages do not affect a lot;
==>>SSD is not necessary for single computer.

3, HT disable, TBT enable for modern Intel CPUs (AMD?);

4, much higher process numbers than the CPU cores;

5, two parallel method for different search parameters;

6, wish to be close to the possible shortest time;

Thank you!

Please send emails to panzc@bao.ac.cn if you are interested in the parallel scripts.

Pan, Zhichen (NAOC), Lei Qian(NAOC), Youling Yue(NAOC), ...
2017-06-28