PTAR

Table of contents

1 Introduction

Creating tar files from thousands of files is very slow on Lustre. `ptar` addresses this issue by creating tar files using several threads.

2 Comparison with regular tar

Our test data is a ODB database which contains about 1500 files. Total data size is 2.5 Gb.

The regular `tar` command is used with the `--blocking-factor=2000` option which was recommended by Bull. `ptar` is run with 24 threads.

Archive extraction and creation is performed on regular compute nodes in exclusive mode. We take care not to re-use nodes, as they may hold file-system data in cache.

	Create archive	Extract archive
`tar`	100s	17s
`ptar`	9s	6s

3 `ptar` scalability

Threads	Create archive	Extract archive
1	106s	47s
2	45s	28s
3	35s	20s
6	16s	10s
12	10s	8s
24	9s	6s

4 Where is it ?

`/home/gmap/mrpm/marguina/SAVE/ptar/ptar`