

IOR

Escola Supercomputador SDUMONT

IOR

*André Ramos Carneiro*  
*Bruno Alves Fagundes*

# IOR

Roteiro:

- IOR
- Testes
- Resultados

# IOR - Interleaved-Or-Random

- Desenvolvido pelo Lawrence Livermore National Laboratory (LLNL)
- MPI-IO, HDF5, NetCDF e o POSIX
- É possível configurar
  - Tamanho geral de E/S
  - Tamanho das transferências individuais
  - Modo de acessar o arquivo (1->N ou N-N)
  - Os dados serão acessados de forma sequencial ou aleatória
- É possível simular os padrões das operações de E/S de aplicações reais



# IOR - Parâmetros

- `api`: (POSIX, MPI-IO, HDF5 e NetCDF)
- `segmentCount`: Número de datasets
- `blockSize`: Tamanho do subdomínio do dataset (**múltiplo de 64bit**)
- `filePerProc`: Arquivo por processador ou um arquivo compartilhado
- `readFile` e `writeFile`: Realizar teste de Leitura e Escrita
- `transferSize`: Tamanho da requisição de E/S (**múltiplo de 64bit**)
- `numTasks`: Número de processadores participantes
- `collective`: Utiliza operações coletivas para o acesso
- Podem ser passados como parâmetros ou através de um arquivo:
  - `-f /caminho/IOR_SCRIPT`

# IOR - Testes

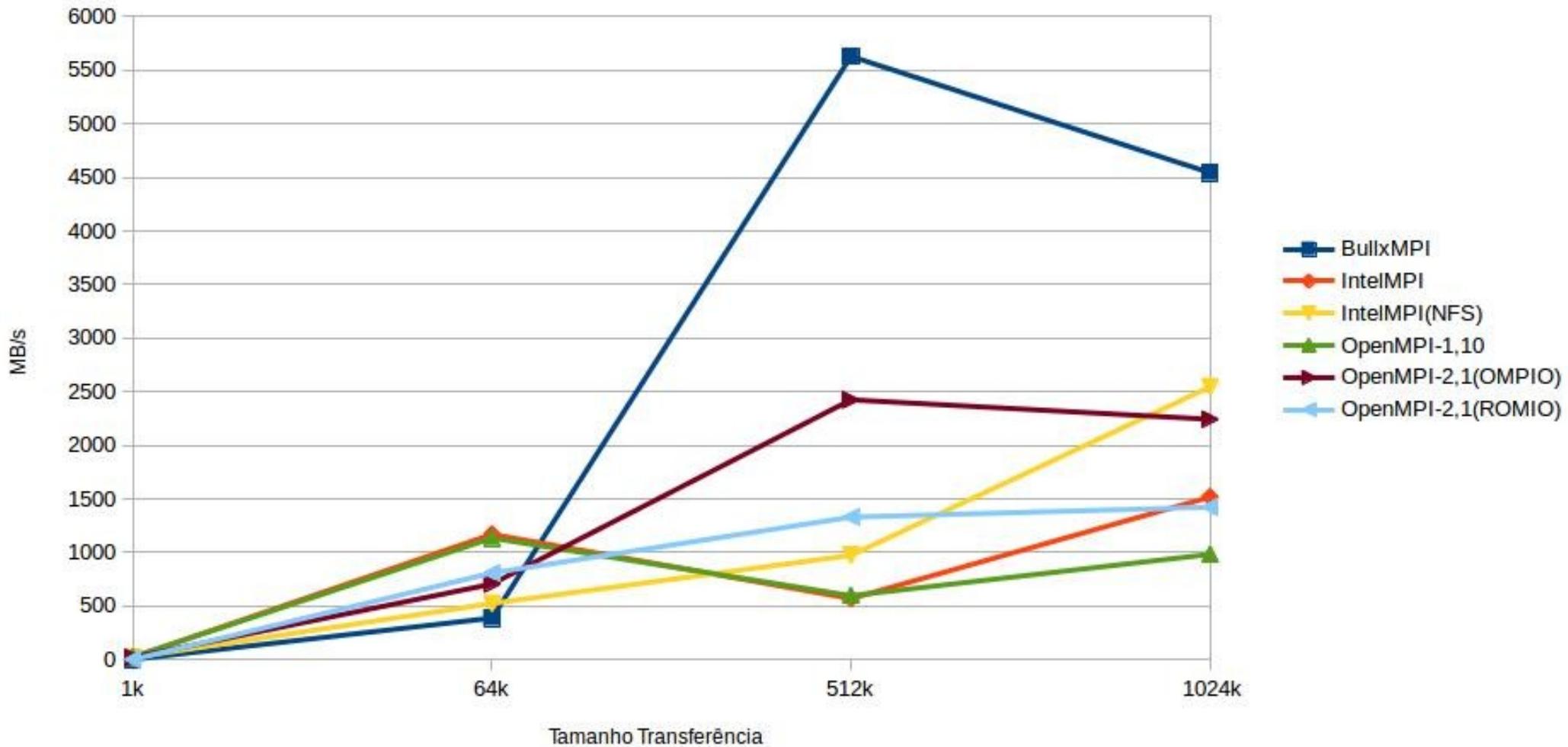
- Testes realizados utilizando as Implementações MPI:
  - BullxMPI, IntelMPI, IntelMPI (vendo o Lustre como NFS), OpenMPI 1.10, OpenMPI 2.1 (ROMIO) e o OpenMPI 2.1 (OMPIO)
- Utilizando 10 nós (240 cores), 10 stripes e Aumentando o tamanho da transferência

TransferSize	BlockSize	SegmentSize	SegmentCount
1024	1024	245760	6554
65536	65536	15728640	103
524288	524288	125829120	13
1048576	1048576	251658240	7

- Um arquivo compartilhado de 1,5 GB
  - `file size = blockSize * numTasks * segmentCount`

# IOR – Resultados MPIIO

Desempenho Leitura  
10 nós x 10 stripes

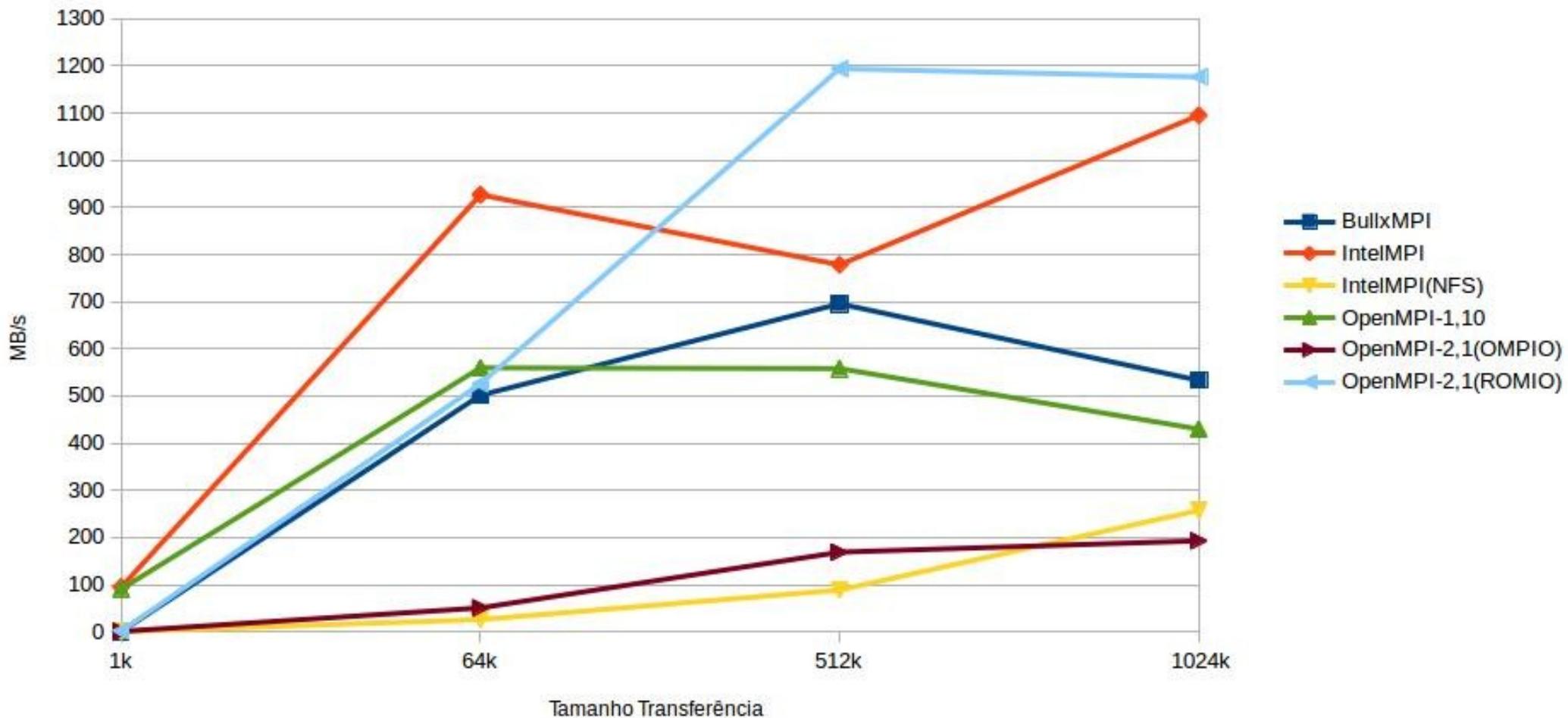


# IOR – Resultados MPIIO



## Desempenho Escrita

10 nós x 10 stripes





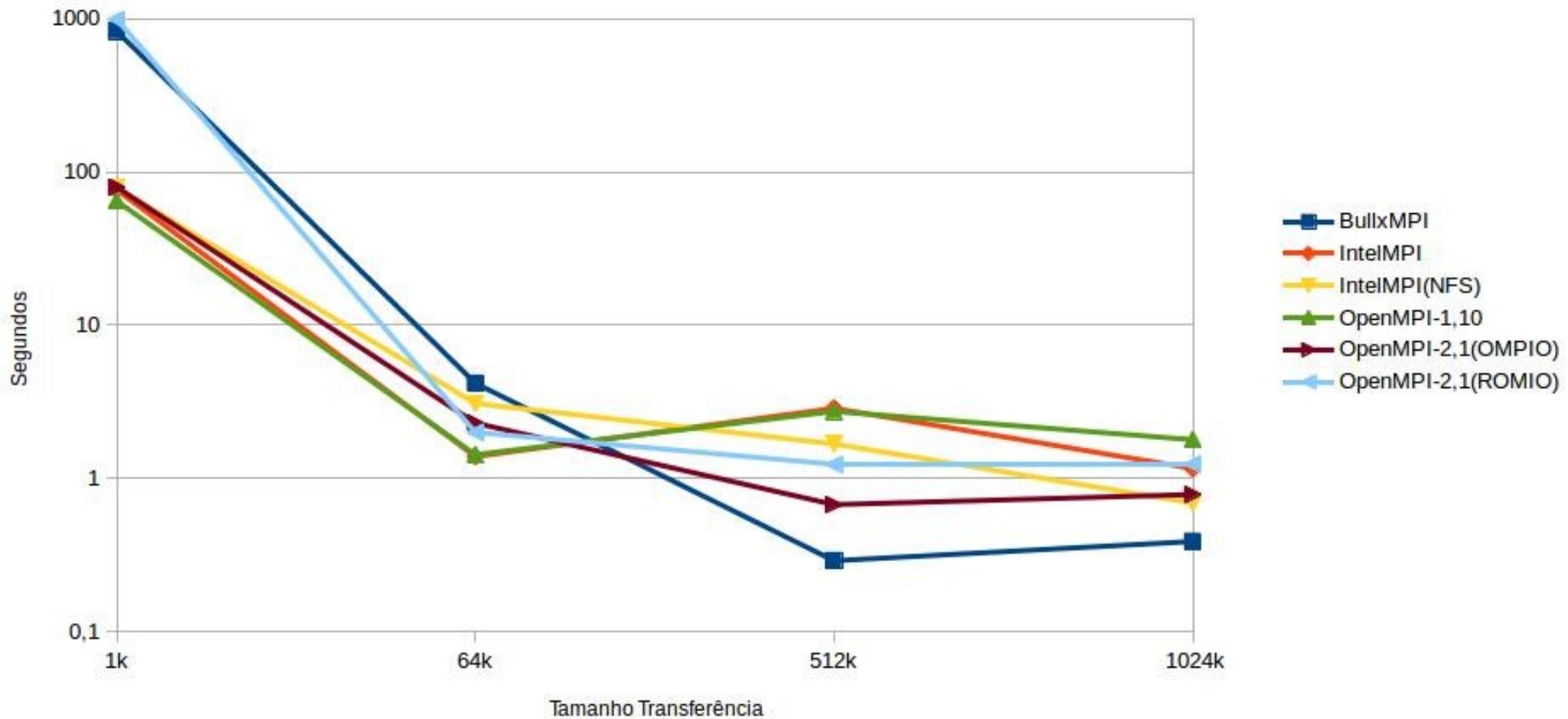
- Maior desempenho leitura:
  - BullxMPI = 5628,29 MB/s | transferSize 512k
    - \* Partiu de 1,94 MB/s (1k)
    - Resultado na escrita: 3º Lugar – 695,45 MB/s | transferSize 512k
- Maior desempenho escrita:
  - OpenMPI 2.1 (ROMIO) = 1193,96 MB/s | transferSize 512k
    - \* Partiu de 1,91 MB/s (1k)
    - Resultado na leitura: 3º Lugar – 2424,05 MB/s | transferSize 512k

# IOR – Resultados - TEMPO



## Tempo da Operação de Leitura

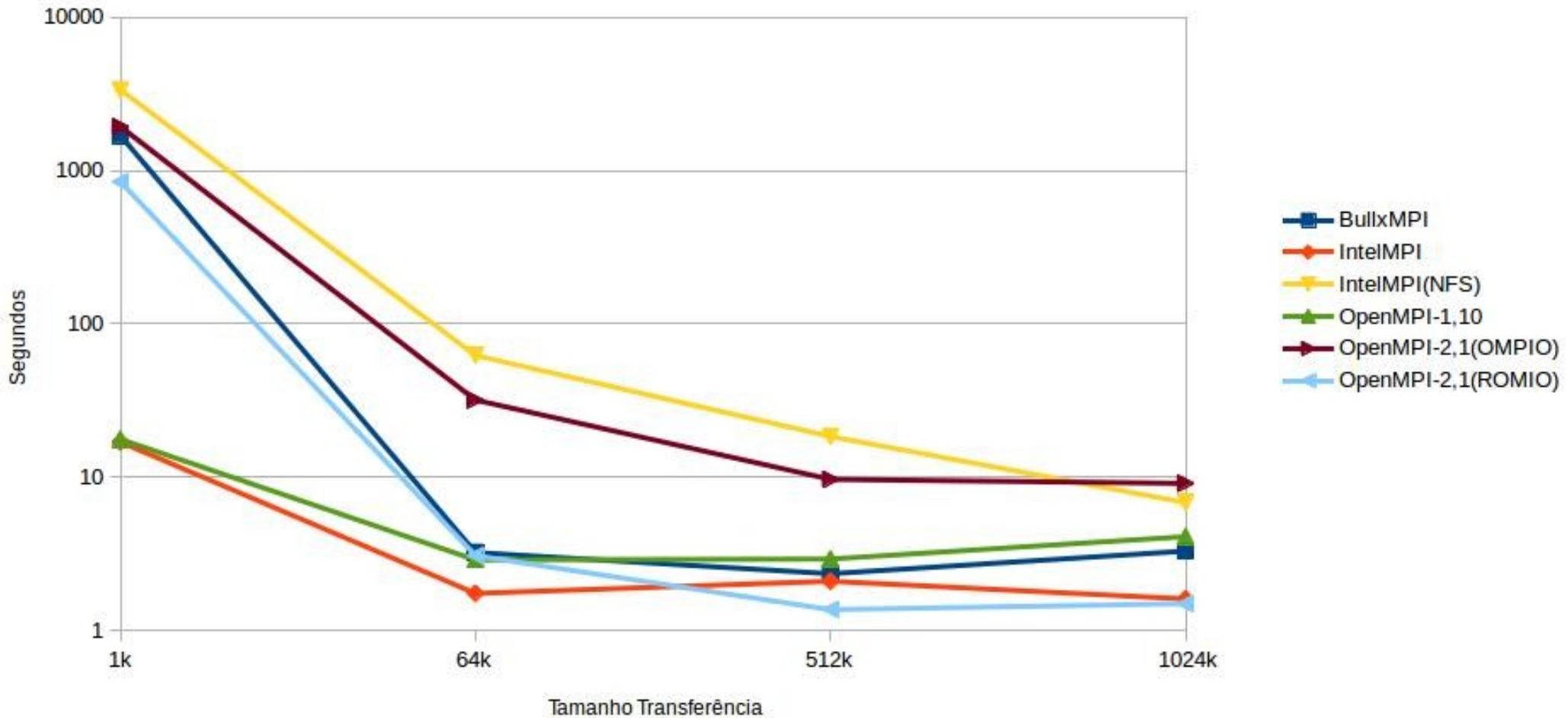
10 nós x 10 stripes



# IOR – Resultados - TEMPO

## Tempo de Operação de Escrita

10 nós x 10 stripes





- Menor tempo de leitura:
  - BullxMPI = 0,29 s, | transferSize 512k
    - \* Partiu de 830,79 s ( $\cong$  13:50 minutos) | transferSize 1k
    - *Resultado na escrita: 3º Lugar – 2,35 s (ts 512k)*
- Menor tempo de escrita:
  - OpenMPI 2.1 (ROMIO) = 1,37 s, | transferSize 512k
    - \* Partiu de 845,39 s ( $\cong$  14:05 minutos) | transferSize 1k
    - *Resultado na leitura: 3º Lugar – 1,23 s (ts 512k)*