

MPI Fault Tolerance Library

User-level checkpointing

Поддерживаемый и реализованный функционал

- **C/C++**
- **MPI 3.0**
- **User-level checkpoint library**
- **ULFM**

База тестовых примеров

- **head_2d** - Laplace equation solver by Jacobi iteration method
- **n_body** - an n-body simulation approximates the motion of particles, often specifically particles that interact with one another through some type of physical forces

Функциональные возможности User-level checkpoint library

- **Rollback recovery** - checkpoint/restart based
- **Failure detection** - ULFM based
- **Snapshot creation** - hard drive based (in place/via NFS)
- **Incremental checkpointing** - delta encoding based (XOR operation)
- **Additional compress procedure** - zlib based

В процессе разработки

- **Реализация альтернативных отказоустойчивых методов**
- **Расширение количества тестовых примеров**
- **Снижение накладных расходов**
- **Улучшение степени сжатия**
- **Улучшение скорости вычисления дельт**

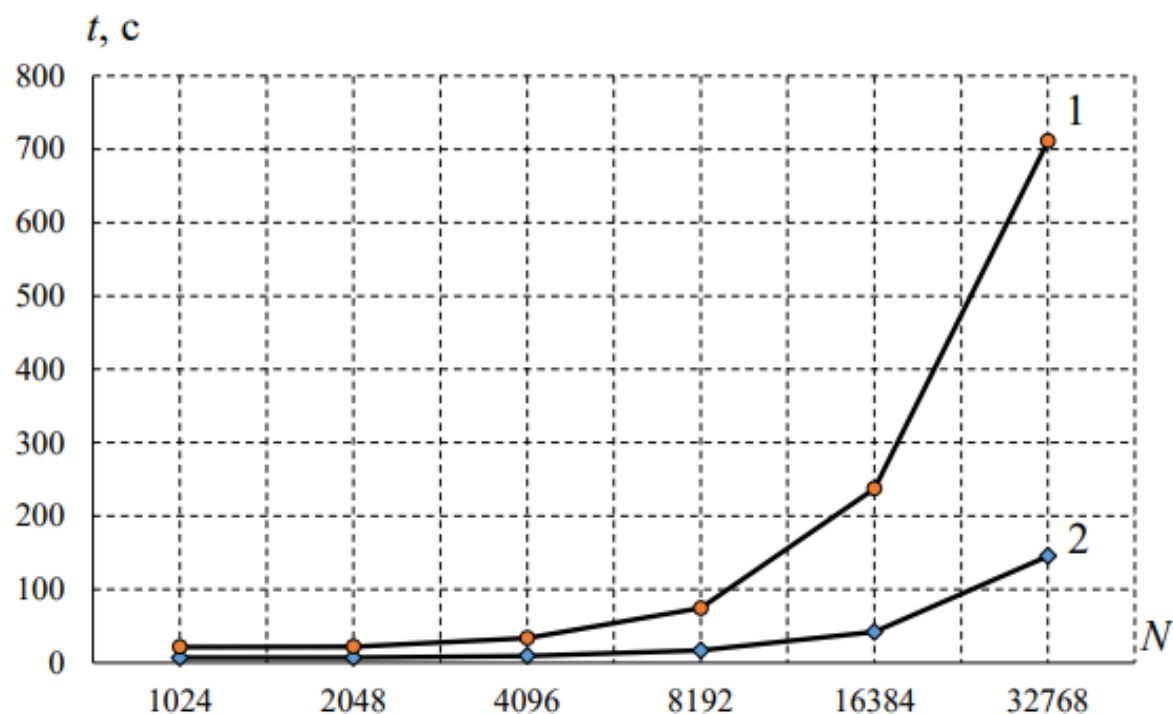
Размер матрицы 8192*8192, количество узлов 2, количество процессоров 8

Пример	Затраченное время (сек)	Ускорение	Количество контрольных точек	Суммарный размер контрольных точек (байт)	Ускорение
heat_2d (w/o checkpoint)	49	-	-	-	-
heat_2d (checkpoint + zlib)	69,3	-44%	5	178,304	-
heat_2d (checkpoint + xor + zlib)	66	-34%/ +5%	5	174,528	+2%
heat_2d (checkpoint + xor + zfp)	61,2	-24%/ +13%	5	11,136	+1600%

Размер матрицы 16384*16384, количество узлов 2, количество процессоров 8

Пример	Затраченное время (сек)	Ускорение	Количество контрольных точек	Суммарный размер контрольных точек (байт)	Ускорение
heat_2d (w/o checkpoint)	177	-	-	-	-
heat_2d (checkpoint + zlib)	209	-18%	5	358,592	-
heat_2d (checkpoint + xor + zlib)	211	-19%/ -0,1%	5	348,288	+%2

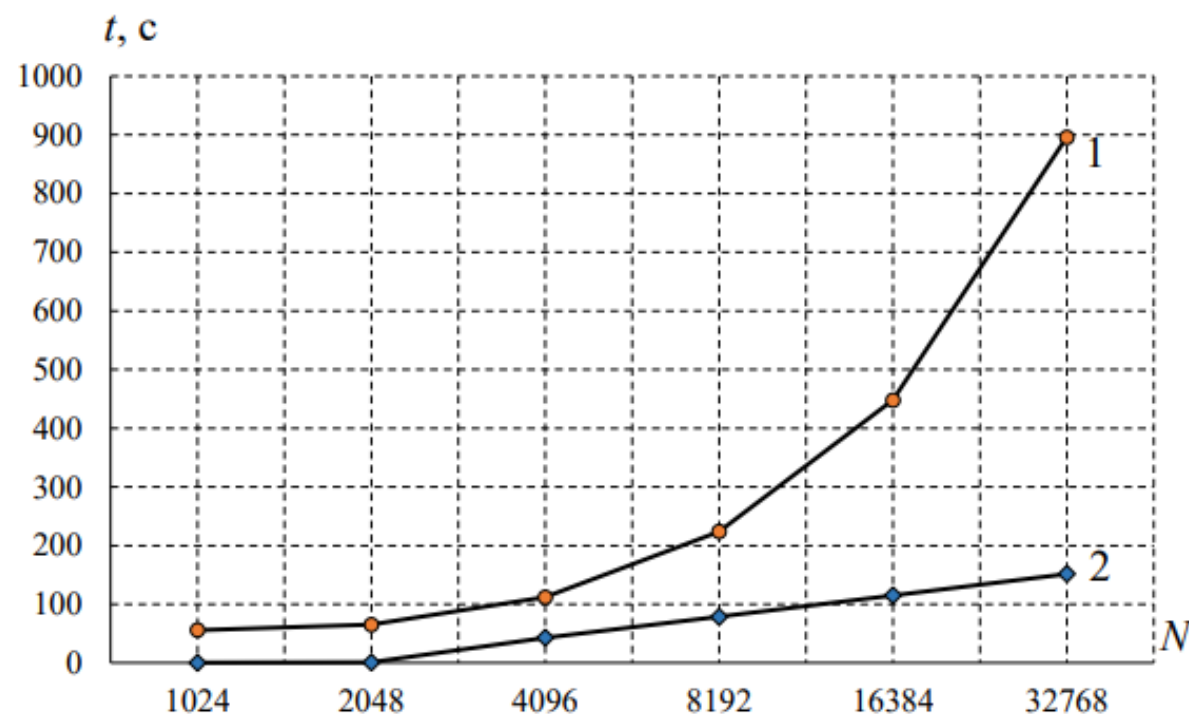
Эксперименты (запись контрольной точки по протоколу NFS)



Зависимость времени t выполнения параллельной программы от размера N расчетной области на кластере Oak (32 процесса – 4 вычислительных узла)

1 – время выполнения параллельной программы с формированием контрольной точки

2 – время выполнения исходной параллельной программы



Зависимость времени t выполнения параллельной программы от размера N расчетной области на кластере Jet (128 процессов – 16 вычислительных узлов)

1 – время выполнения параллельной программы с формированием контрольной точки

2 – время выполнения исходной параллельной программы