

зается аналогия Рейнольдса, дает существенное отклонение от аналитического решения, как по средней выходной температуре, так и по критерию Нуссельта. Учет специфики теплопереноса для натриевого теплоносителя с помощью специально разработанной модели LMS позволяет существенно улучшить результаты расчетов для данного класса задач.

### Литература

1. Боришанский В. М. Жидкометаллические теплоносители. М.: Атомиздат, 1967.
2. Исаченко В. П., Осипова В. А., Сукомел А. С. Теплопередача. М.: Энергоиздат, 1981.
3. Осипов С. Л., Рогожкин С. А., Фадеев И. Д. и др. Разработка модели LMS для учета турбулентного теплопереноса в натриевом теплоносителе и ее тестирование // Международный форум «Инженерные системы-2012»: Тез. докл. 2012. С. 21–22.
4. Осипов С. Л., Рогожкин С. А., Соболев В. А. и др. Численное моделирование теплогидравлических процессов в верхней камере быстрого реактора // Атомная энергия. 2013. Т. 115. Вып. 5. С. 295–298.

## БИБЛИОТЕКА ЕФР КАК СРЕДСТВО ЭФФЕКТИВНОГО ДОСТУПА К ФАЙЛОВЫМ ДАННЫМ НА ГИБРИДНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ И СУПЕРКОМПЬЮТЕРАХ

*К. К. Олесницкая, И. А. Антипин, М. А. Петрова*

Российский федеральный ядерный центр –  
Всероссийский НИИ экспериментальной физики, г. Саров

### Введение

Освоение и эффективное использование вычислительных и файловых ресурсов многопроцессорных ЭВМ на базе перспективной архитектуры является одной из актуальных задач в математическом отделении ВНИИЭФ. Поэтому в рамках создания базового программного обеспечения, необходимого для проведения комплексного имитационного моделирования задач основной тематики и инженерных задач различных отраслей промышленности на многопроцессорных ЭВМ, реализована библиотека ЕФР [1]. Проблемно-ориентированная библиотека обеспечивает оперативный и эффективный масштабируемый доступ к файловым данным на многопроцессорных ЭВМ.

Библиотека используется в основных прикладных программных комплексах ВНИИЭФ, является составной частью программ общего сервиса (препроцессор, постпроцессор) и является неотъемлемым компонентом технологии проведения расчетов на вычислительных системах (ВС) ВНИИЭФ.

Библиотека ЕФР предоставляет единый функциональный интерфейс для работы с файловыми данными и единый формат представления расчетных данных, который позволяет сохранять все необходимые данные как для продолжения счета в рамках одного комплекса программ, так и для передачи в другой комплекс программ для полномасштабных последовательных расчетов.

Программы чтения-записи расчетных данных, реализованные средствами библиотеки ЕФР легко портируются между несовместимыми аппаратными платформами. Это достигается благодаря тому, что библиотека ЕФР при решении задач доступа к файловым данным использует реализацию с учетом стратегии оптимального использования аппаратных и файловых ресурсов конкретной ВС.

## Проблематика

Проблема неустойчивой работы ВС не нова и отчетливо проявилась в 2007 году с вводом в эксплуатацию многопроцессорного кластера МП-5. Для решения данной проблемы с момента ее появления постоянно проводятся исследовательские и практические работы по ее устранению на аппаратном, системном и программном уровнях.

Данная проблема значительно усугубилась в настоящее время при использовании более мощных гетерогенных и гибридных вычислительных ресурсов, так как новые архитектуры требуют создания прикладного программного обеспечения, принципиально отличного от существующего.

Поэтому совместно со специалистами, отвечающими за аппаратные и файловые ресурсы ВС, была проведена исследовательская работа. Исследование показало, что корень проблемы в том, что разработка прикладных программных комплексов и программ общего сервиса ведется без учета явных особенностей файловой системы, которые напрямую зависят от аппаратной составляющей вычислительного ресурса. Поэтому решить данную проблему возможно только при выполнении ряда согласованных работ как на аппаратно-системном уровне, так и на уровне прикладных программ.

В результате работы был определен ряд возможных причин, которые приводят к снижению работоспособности вычислительных ресурсов.

## Задачи

В качестве решения данной проблемы на программном уровне была разработана новая технология сохранения расчетных данных на ВС и определены задачи, которые должны быть решены в прикладных комплексах, эксплуатирующих вычислительные ресурсы:

- уменьшить количество результирующих файлов;
- уменьшить количество файлов, записываемых одновременно;
- уменьшить количество процессов работающих с одним файлом в режиме записи;
- уменьшить количество файлов в одной директории.

## Библиотека ЕФР

Неотъемлемым компонентом технологии проведения расчетов на ВС является библиотеки ЕФР, которая обеспечивает масштабируемый доступ к файловым данным.

Для большинства прикладных комплексов библиотека ЕФР предоставляет единый функциональный интерфейс для работы с файловыми данными и единый формат представления расчетных данных, который позволяет сохранять все необходимые данные как для продолжения счета в рамках одного пакета программ, так и для передачи в другой пакет программ для полномасштабных последовательных расчетов.

Библиотека ЕФР позволила разработчикам программных комплексов сосредоточиться на предметной области методики и в меньшей степени заботиться о рутинных операциях, связанных с доступом к файловым данным на ВС.

Использование библиотеки ЕФР в подавляющем большинстве прикладных комплексов позволило сосредоточить решение данной проблемы на программном уровне в рамках библиотеки ЕФР – за счет разработки новых функциональных возможностей, которые обеспечили эффективный доступ к файловым данным с учетом стратегии оптимального использования аппаратных и файловых ресурсов ВС ВНИИЭФ.

## Распределенный ЕФР

Одной из основных задач библиотеки ЕФР является обеспечение проблемно-ориентированного масштабируемого доступа к информации на многопроцессорных ЭВМ. В параллельном режиме расчетные данные удобно сохранять в виде распределенного по файлам разреза, где каждый процесс формирует свой локальный файл-фрагмент.

Файл-фрагмент представляет собой обычный файл ЕФР, который содержит локальный фрагмент задачи с процесса.

Получаемый таким образом набор файлов-фрагментов дополняется головным файлом (рис. 1), который содержит глобальную информацию о задаче и информацию о распределении данных по файла-фрагментам.

Созданный таким образом ЕФР-разрез является распределенным, а доступ к данным становится автоматически масштабируемым, т. е. данные можно прочитать на произвольном числе процессов. Для чтения такого распределенного разреза каждому процессу достаточно открыть только головной файл и работать с ним как с нераспределенным. А функции доступа к данным на основе анализа информации о распределении данных будут автоматически открывать необходимые файлы-фрагменты и считывать из них необходимые порции данных.



Рис. 1. Распределенный ЕФР

## Концепция создания распределенного ЕФР

Концепция создания распределенного ЕФР (рис. 2) не накладывает ограничения на результирующее количество файлов-фрагментов. Основное ограничение данной концепции заключается в односвязности и самодостаточности фрагментов области в рамках одного файла-фрагмента.

Несмотря на то, что с самого начала функционирования библиотеки ЕФР существовала возможность топологического объединения нескольких фрагментов области на уровне пользовательской программы, данная возможность не использовалась из-за заметного усложнения программ записи. Поэтому до сих пор большинство комплексов используют подход, при котором каждый счетный процесс формирует отдельный файл-фрагмент. Данный подход успешно применялся на протяжении многих лет, но на текущий момент он не подходит.

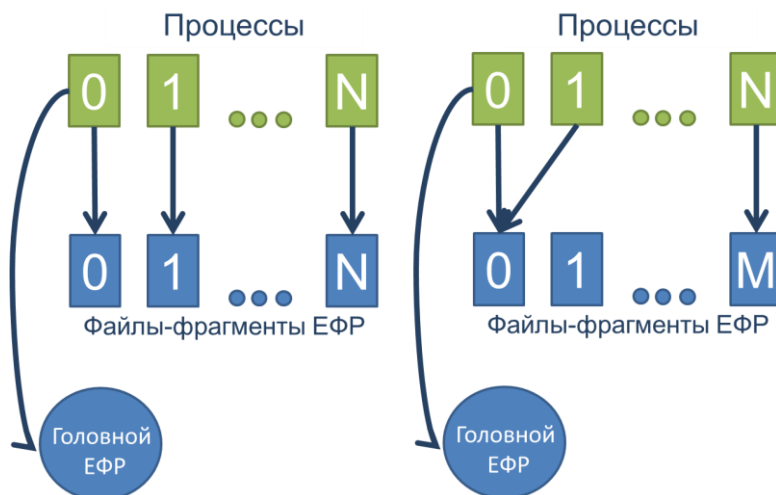


Рис. 2. Создание распределенного ЕФР

### Технология сохранения данных в параллельном режиме

Технология заключается в объединении данных с нескольких процессов в один файл-фрагмент, а алгоритм сохранения представляет собой следующую последовательность действий:

- Процессы разделяются на группы, для работы с одним файлом-фрагментом (группы процессов первого уровня (рис. 3)). Количество групп определяется количеством файлов-фрагментов.
- Группы процессов для работы с одним файлом разделяются на группы по файловым серверам. Количество групп определяется количеством файловых серверов ВС.

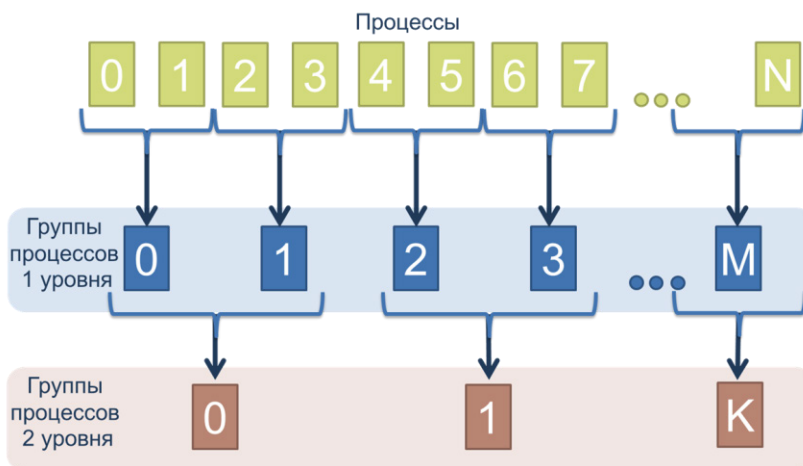


Рис. 3. Группировка процессов

Процессы в группах выстраиваются в очереди и в соответствии с ней выполняют доступ к файлам-фрагментам. В результате достигается следующее: ограничивается количество одновременно записываемых файлов-фрагментов по количеству файловых серверов и ограничивается количество процессов, одновременно работающих с одним файлом-фрагментом в режиме записи.

Объединение данных с нескольких процессов можно осуществлять разными способами: топологическое объединение, объектное объединение.

По ряду причин пришлось отказаться от топологического объединения, а для поддержания объектного объединения в библиотеке ЕФР был реализован новый объект данных – подобласть.

## Область (подобласть)

Структура представления данных в ЕФР ориентирована на схему раздельного счета по математическим (счетным) областям, то есть данные должны относиться к какой-либо области. Область может быть распределенной, то есть состоять из любого количества подобластей. Подобласть представляет собой самодостаточный фрагмент области, сохраненный в файл-фрагмент. Один файл-фрагмент может содержать несколько подобластей одной области.

В качестве примера рассмотрим файл-фрагмент, структура которого представлена на рис. 4. Файл-фрагмент содержит девять подобластей трех распределенных областей с идентификаторами: 1, 2 и 4. Область 1 в файле-фрагменте представлена пятью подобластями с линейными индексами: 1, 2, 4, 5 и 6. Область 2 в файле-фрагменте представлена одной подобластью с линейным индексом: 3. Область 4 в файле-фрагменте представлена тремя подобластями с линейными индексами: 7, 8 и 9.

## Коллективные функции доступа библиотеки ЕФР

На текущий момент функции доступа к файловым данным в библиотеке ЕФР имеют два режима работы: асинхронный и коллективный. Режим работы может меняться и переопределяться по ходу выполнения программы.

*Установка распределения MPI-процессов для коллективных функций доступа.* Для установки коллективного режима работы используется функция `EFR_Mpi_Set_Group`. Функция выполняет распределение процессов по файлам-фрагментам и файловым серверам для заданной ПФС. Информация о распределении сохраняется в глобальной области видимости (в настройках библиотеки) текущего процесса. Функция поддерживает два метода распределения процессов по группам: один файл-фрагмент с узла и по количеству заданных файлов-фрагментов.

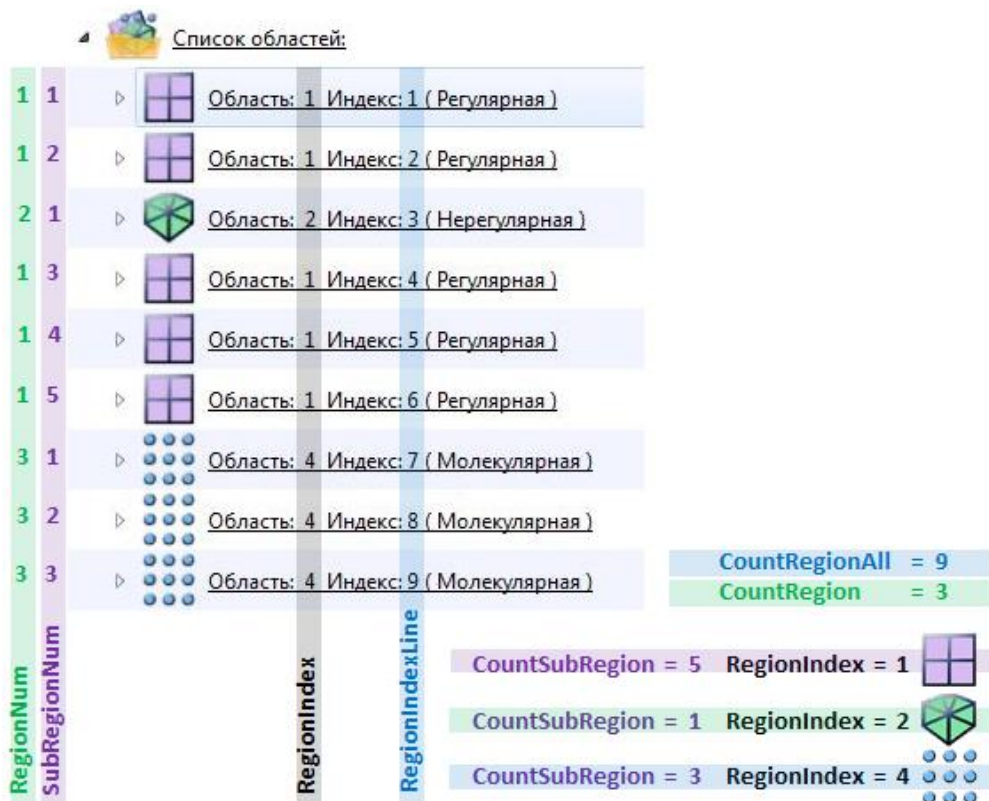


Рис. 4. Структура файла-фрагмента с подобластями

Функция работает по схеме (рис. 5):

- определяется количество файлов-фрагментов (если задан режим – один файл-фрагмент с узла);
- определяется количество OST-серверов;
- MPI-процессы распределяются равномерно по файлам-фрагментам. Для каждого MPI-процесса определяется файл-фрагмент и коммуникатор группы MPI-процессов для работы с ним;
- файлы-фрагменты распределяются равномерно по OST-серверам. Для каждого MPI-процесса определяется OST-сервер и коммуникатор группы MPI-процессов для работы с ним.

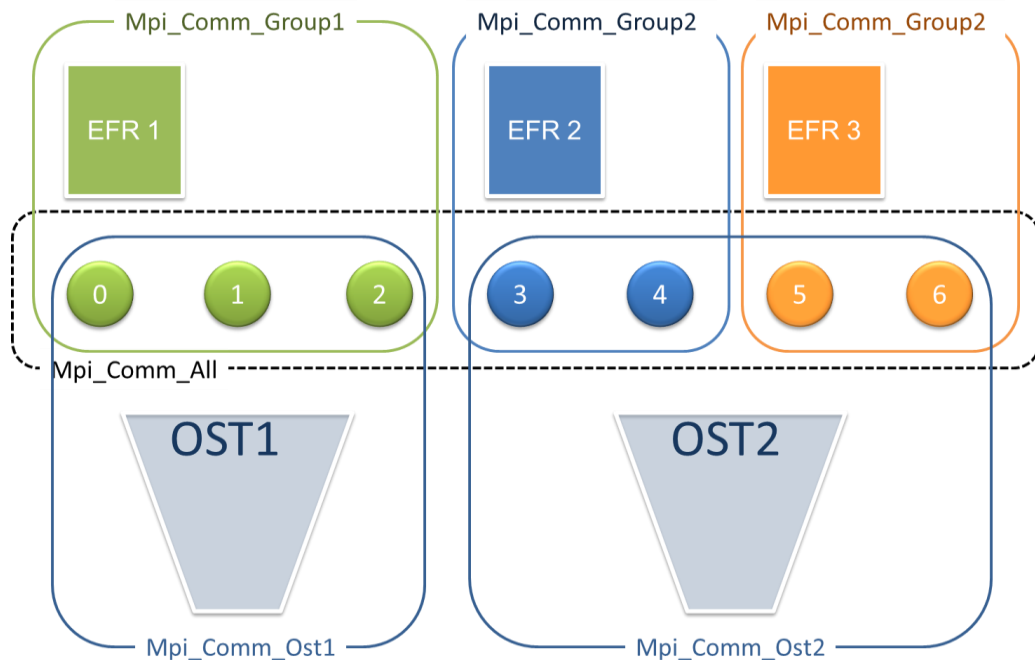


Рис. 5. Распределение MPI-процессов по файлам и OST-серверам

*Создание файлов-фрагментов.* Функции данной группы создают файлы-фрагменты ЕФР для записи расчетных данных. Работа коллективной функции выполняется по схеме: управляющий процесс (из группы Mpi\_Comm\_GroupX) создает файл, а затем остальные – открывают файл на редактирование.

*Установка параметров задачи.* Функции данной группы предназначены для установки общей информации о задаче в файлы-фрагменты. Работа коллективной функции выполняется по схеме: управляющий процесс (из группы Mpi\_Comm\_GroupX) вызывает функцию, остальные ничего не делают.

*Создание областей, установка областных параметров и работа с объектами данных.* Функции данной группы предназначены для установки и записи данных фрагмента области. Работа коллективной функции выполняется по схеме: управляющий процесс (из группы Mpi\_Comm\_GroupX) вызывает функцию, остальные ничего не делают.

*Установка параметров задачи.* Функция EFR\_File\_Create\_Fragments создает файлы-фрагменты ЕФР для записи расчетных данных. Работа коллективной функции выполняется по схеме: все MPI-процессы (из группы Mpi\_Comm\_GroupX) последовательно вызывают функцию.

## Примеры программ записи

В помощь разработчикам программных комплексов в освоении и внедрении новой технологии сохранения расчетных данных с помощью коллективных функций доступа библиотеки ЕФР разработан ряд примеров. Примеры моделируют основные алгоритмы расчета и записи рассчитанных данных в параллельном режиме, которые используются в ППК.

В программах создается распределенный ЕФР, в котором сохраняются расчетные данные одной регулярной области. Регулярная область содержит два вещества. Распределение веществ по ячейкам сетки осуществляется следующим образом: первая ячейка - смесь, четные ячейки – вещество № 1, нечетные ячейки – вещество № 2. Количество ячеек регулярной сетки определяется количеством MPI-процессов.

Каждый MPI-процесс формирует данные подобласти, состоящей из одной ячейки, габариты которой –  $10 \times 10 \times 10$  условных единиц. Позиционирование ячеек в сетке выполняется вдоль оси абсцисс.

Основными отличительными особенностями примеров является способ распределения MPI-процессов по файлам-фрагментам и способ описания подобластей в карте фрагментов.

В первой программе используется автоматическое распределение MPI-процессов по файлам (режим: один файл-фрагмент с узла). В данном примере количество файлов-фрагментов вычисляется автоматически и равно количеству узлов на задачу. Подобласти в файлах-фрагментах имеют одинаковый идентификатор, поэтому при формировании карты фрагментов используется линейный индекс подобласти.

На рис. 6 представлена структура распределенного ЕФР, который создан в результате выполнения программы в параллельном режиме на четырех MPI-процессах, принадлежащих одному узлу.

The image shows two windows of the EFR Tools software. The left window displays the 'Parameters' section for a file named 'program1.efr:00000'. It lists four regular regions (Index 1, 2, 3, 4) and other task parameters. The right window displays the 'List: 8099 (Frag\_Map\_Sub)' window, which contains a table of elements.

Элемент 1	
1	0
2	0
3	1
4	1
5	2
6	1
7	2
8	1
9	2
10	0
11	2

Рис. 6. Структура данных распределенного ЕФР

Во второй программе используется полуавтоматическое распределение MPI-процессов по файлам. Количество файлов-фрагментов определяется пользователем и равно половине количества MPI-процессов. Подобласти в файлах-фрагментах имеют одинаковый идентификатор, поэтому при формировании карты фрагментов используется линейный индекс подобласти.

На рис. 7 представлена структура распределенного ЕФР, который создан в результате выполнения программы в параллельном режиме на восьми MPI-процессах, принадлежащих одному узлу.

The screenshot shows the EFR\_Tools application interface. The main window displays a tree view of parameters under 'EFR Файл:'. The tree includes 'Тип файла, версия, размерность:', 'Список областей:', 'Область: 1 Индекс: 1 (Регулярная)', 'Область: 1 Индекс: 2 (Регулярная)', 'Параметры задачи:', 'Информация об исполнителе и заказчике:', 'Комментарий к задаче:', and 'Область: 1 Индекс: 1 (Регулярная)'. A data table is visible in the bottom right corner of the main window, showing the distribution of MPI processes across regions.

	Элемент 1	Элемент 2	Элемент 3	Элемент 4
1	0	1	2	3
2	0	0	0	0
3	1	1	1	1
4	1	3	5	7
5	2	4	6	8
6	1	1	1	1
7	2	2	2	2
8	1	1	1	1

Рис. 7. Структура данных распределенного ЕФР

В третьей программе используется ручное распределение MPI-процессов по файлам. Количество файлов-фрагментов равно количеству MPI-групп, определенных пользователем. В программе создается две группы MPI-процессов, первая половина MPI-процессов работает с файлом-фрагментом 0, вторая – с файлом-фрагментом 1. Для индексации подобластей используется сквозная нумерация. Индекс подобласти формируется как ранг MPI-процесса + 1, а номер области в головном ЕФР равен количеству MPI-процессов + 1. При формировании карты фрагментов используется индекс подобласти, так как в данном случае он является уникальным.

На рис. 8 представлена структура распределенного ЕФР, который создан в результате выполнения программы в параллельном режиме на восьми MPI-процессах, принадлежащих одному узлу.



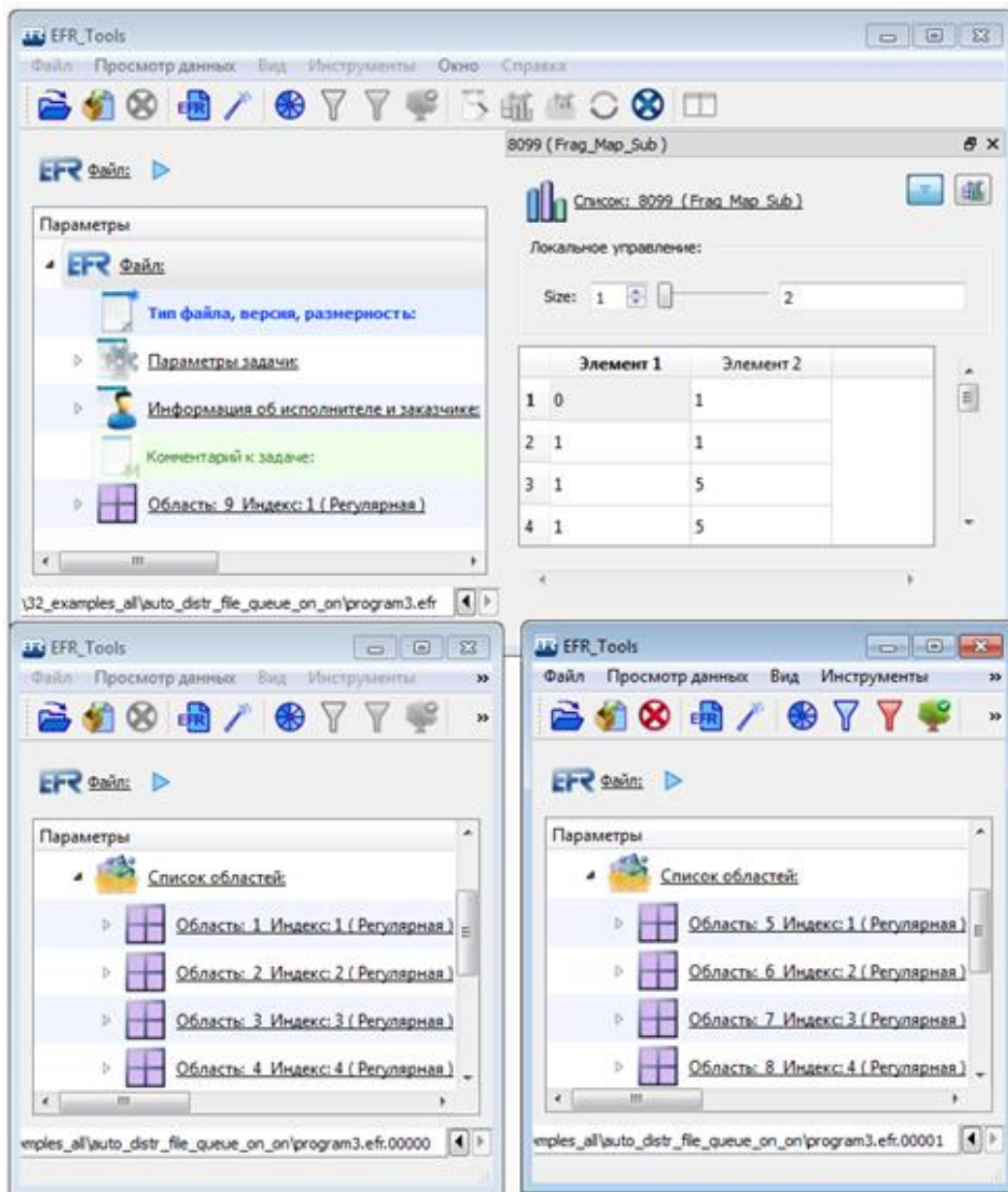


Рис. 8. Структура данных распределенного ЕФР

## Заключение

В отчете описаны возможности библиотеки ЕФР, которые обеспечивают аккуратную работу с параллельной файловой системой LUSTRE и позволяют на программном уровне решить следующие задачи:

- уменьшить количество результирующих файлов;
- уменьшить количество файлов, записываемых одновременно;
- уменьшить количество процессов, работающих с одним файлом в режиме записи;
- уменьшить количество файлов в одной директории.

В настоящее время новая технология сохранения расчетных данных поддерживается в ряде программных комплексах ВНИИЭФ.

## Литература

1. Олесницкая К. К., Антипин И. А., Шубина М. А. Библиотека ЕФР для масштабируемого доступа к файловым данным на многопроцессорных ЭВМ // XI Забабахинские научные чтения: Тез. докл.
2. MPI: A Message-Passing Interface Standard Version 2.2. [Electronic resource]. Mode of access: <http://www.mpi-forum.org/docs/mpi-2.2/mpi22-report.pdf>.

## EFR-TOOLS КАК СРЕДСТВО МОДИФИКАЦИИ, ВЕРИФИКАЦИИ И ВАЛИДАЦИИ РАСЧЕТНЫХ ДАННЫХ ЕДИНОГО ФАЙЛОВОГО РАЗРЕЗА

*К. К. Олесницкая, И. А. Антипин, М. А. Петрова*

Российский федеральный ядерный центр –  
Всероссийский НИИ экспериментальной физики, г. Саров

### Введение

В математическом отделении ИТМФ РФЯЦ-ВНИИЭФ для сохранения расчетных данных в файлы (разрезы) при численном моделировании задач математической физики по прикладным программным комплексам (ППК) в качестве основного инструмента используется библиотека ЕФР [1]. С ее помощью данные сохраняются в единый файловый разрез (ЕФР) в виде бинарных файлов унифицированного формата. Неотъемлемым компонентом библиотеки является кроссплатформенная программа EFR-Tools [2], основной задачей которой является ускорение и упрощение процесса внедрения возможностей библиотеки ЕФР в ППК.

Программа EFR-Tools имеет удобный графический интерфейс, предназначенный для просмотра, редактирования, анализа и обработки большого объема данных (более 10 миллиардов счетных узлов) единого файлового разреза.

Для разработчиков ППК EFR-Tools предоставляет ряд вспомогательных средств, для отладки собственных программных комплексов и для решения проблем, связанных с постобработкой данных ЕФР программами общего сервиса. К таким средствам относятся:

- средства создания базовых (тестовых) разрезов;
- средства модификации и коррекции разрезов;
- средства компоновки разрезов;
- средства верификации разрезов;
- средства валидации разрезов.

Разработчики ППК могут самостоятельно расширять функциональность программы необходимыми специфическими возможностями, благодаря тому, что в EFR-Tools поддержана интерактивная работа с плагинами и реализованы средства для их создания.

Для разработчиков библиотеки ЕФР программа предоставляет средства, которые позволяют в кратчайшие сроки восстанавливать «битые» разрезы.

В докладе описываются функциональные возможности программы EFR-Tools версии 1.3.0, которая создана на базе библиотеки ЕФР версии 3.1.2.

### Общие сведения о программе EFR-Tools. Инструменты разработки

Для разработки графического пользовательского интерфейса (GUI) была выбрана кроссплатформенная библиотека Qt [3]. Во-первых, в Qt реализован гибкий подход к визуальному отображе-