

# Распределенная система потоковой обработки радарных изображений на базе фреймворк Apache Spark

**АВТОРЫ:** д.т.н. Потапов В.П., к.т.н. Попов С.Е., Костылев М.А.

Разработан прототип распределенного программного комплекса массово-параллельного исполнения заданий на базе фреймворк Apache Spark для потоковой пре- и постобработки радарных снимков. Текущая версия прототипа системы включает следующие расчетные модули с внутренней высокопроизводительной программной алгоритмизацией: расчет интерферометрической фазы для пары снимков в бинарном формате BSQ; расчет значений когерентности и формирования интерферограммы; усовершенствованный высокопроизводительный алгоритм развертки фазы с использованием метода роста регионов. Функциональными особенностями разработанной системы являются:

- распределенная отказоустойчивая файловая система на базе технологии HDFS, позволяющая хранить данные на всех этапах обработки радарных изображений и предоставляющая доступ к различным вычислительным узлам кластера;
- распределенная вычислительная платформа на базе технологии Apache Spark, позволяющая производить параллельную обработку пакета интерферометрических изображений на множестве узлов;
- система планирования потоков работ на базе технологии Apache Oozie, позволяющая организовать процесс обработки пакета радарных данных в виде последовательности действий, каждое из которых выполняется после успешного завершения предыдущего.

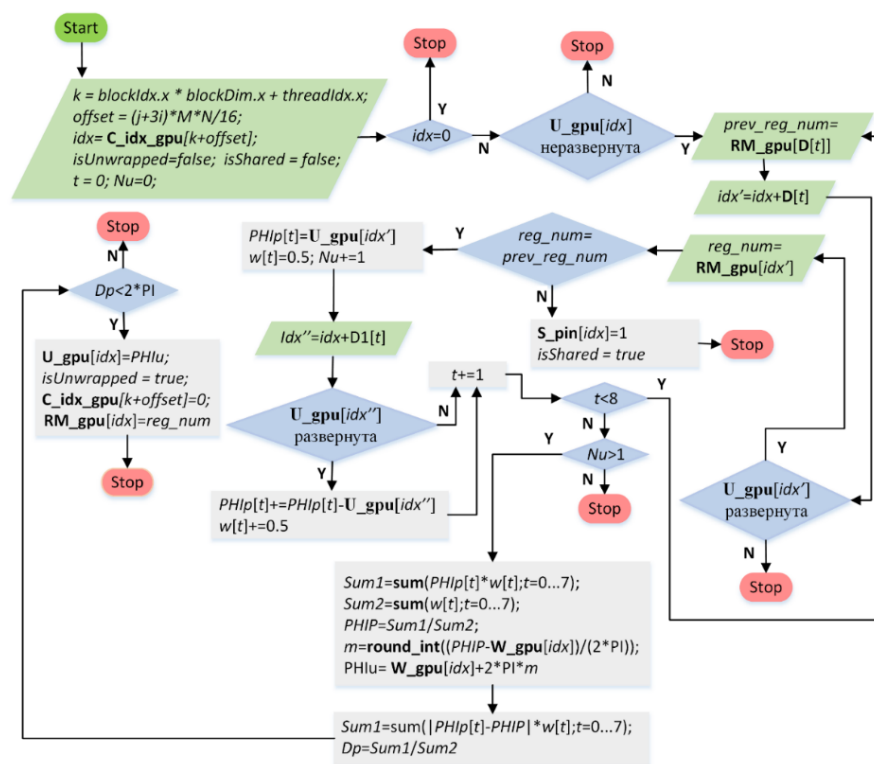


Рисунок 2. Блок-схема алгоритма развертки фазы, запускаемая потоком warp-блока на стороне GPU.

Новизну подхода обеспечивает предложенный способ интерпретации и обработки интерферометрического изображения при загрузке пакета из нескольких (от 10 файлов) интерферограмм и их значений когерентности в распределенную файловую систему HDFS, а также взаимодействие разработанного алгоритма с данными в HDFS во время процедуры подготовки данных.

Оценка полноты и достоверности полученных результатов подтверждается проведенным тестированием программных компонентов усовершенствованного алгоритма (прирост скорости работы в среднем в 10 раз по сравнению с базовым алгоритмом). Интеграция усовершенствованного алгоритма развертки фазы в систему Apache Spark позволила улучшить функциональность и эргономику применения его в пакетной обработке интерферометрических изображений без использования специализированных кластерных систем и программного обеспечения.

Полученные результаты программной реализации этапов обработки радарной информации могут быть использованы для интеграции их в аналогичные системы, разрабатываемые в рамках НИР, как в виде отдельных компонентов и алгоритмов, так и в форме базовых систем для научных исследований.

Результаты оценки научно-технического уровня рассматриваемой исследовательской работы показали высокие характеристики в плане производительности разработанной системы с сохранением требуемой точности результатов.

## **ПУБЛИКАЦИИ:**

1. Потапов В.П., Попов С.Е. Высокопроизводительный алгоритм роста регионов для развертки интерферометрической фазы на базе технологии CUDA // Программная инженерия. 2016. № 2. С. 61-74.
2. Потапов В.П., Попов С.Е. Высокопроизводительный алгоритм Band Interleave Conversion для данных гиперспектральных снимков сенсора EO-1 Hyperion // Информационные технологии и вычислительные системы. 2016. № 1. С. 76-83.
3. Шокин Ю.И., Потапов В.П., Попов С.Е., Гиниятуллина О.Л. Спутниковая радарная интерферометрия: информационно-вычислительные аспекты // Вычислительные технологии. 2016. Т. 21. № 1. С. 141-151.