

**Резюме проекта, выполняемого  
в рамках ФЦП**  
**«Исследования и разработки по приоритетным направлениям развития научно-  
технологического комплекса России на 2014 – 2020 годы»**

по этапу № 1

Номер Соглашения о предоставлении субсидии: 14.58121.0023

Тема: «Разработка аппаратно-программного комплекса для прогнозирования сбоев в работе системы хранения данных с целью предотвращения критических ситуаций, в том числе деградации производительности, отказа сервиса записи/чтения данных и потери данных»

Приоритетное направление: Информационно-телекоммуникационные системы (ИТ)

Критическая технология: Технологии информационных, управляющих, навигационных систем

Период выполнения: 03.10.2017 - 30.06.2020

Плановое финансирование проекта: 500.00 млн. руб.

Бюджетные средства 250.00 млн. руб.,

Внебюджетные средства 250.00 млн. руб.

Получатель: федеральное государственное автономное образовательное учреждение высшего образования "Санкт-Петербургский политехнический университет Петра Великого"

Индустриальный партнер: Общество с ограниченной ответственностью "КНС Групп"

Ключевые слова: хранилище данных, прогнозирование сбоев, предотвращение критических ситуаций, защита данных, система хранения данных, предсказательные алгоритмы, предотвращение деградации производительности

## **1. Цель проекта**

Разработка комплекса прикладных научно-технических решений в обеспечение создания интеллектуальных автоматических аппаратно-программных систем контроля работоспособности, прогнозирования сбоев и предотвращения критических ситуаций в режиме реального времени для новых видов отказоустойчивых систем хранения данных.

## **2. Основные результаты проекта**

Подготовлен документ «Комплектность» по форме Т1, согласованный с Индустримальным партнером. Выполнен аналитический обзор научно-технической, нормативной, методической литературы, затрагивающей научно-техническую проблему, исследуемую в рамках прикладных научных исследований и экспериментальных разработок, в том числе обзор научных информационных источников. На основании проведенных исследований подготовлен раздел промежуточного отчета о прикладных научных исследованиях и экспериментальных разработок.

Проведены патентные исследования в соответствии с ГОСТ 15.011-96 в направлении разработки аппаратно-программного комплекса прогнозирования сбоев системы хранения данных (СХД). На основе проведенных исследований составлен отчет о патентных исследованиях.

Проведен анализ результатов мониторинга параметров и состояний СХД – на базе ЦКП СКЦ «Политехнический» (СПбПУ) и ЭО СХД на платформе TATLIN (ООО «КНС ГРУПП») в конфигурации 1. На основании проведенного анализа подготовлен раздел промежуточного отчета о прикладных научных исследованиях и экспериментальных разработок.

Разработана конструкторская документация на стенд для тестирования программного комплекса прогнозирования сбоев СХД. Разработаны детерминированные и стохастические имитационно дискретно-событийные модели функционирования аппаратных компонентов СХД.

Разработаны алгоритмы диагностирования СХД на основе имитационно дискретно-событийной модели функционирования аппаратных компонентов СХД. Разработанные алгоритмы описаны в программном документе «Описание алгоритма».

Разработан программный документ «Руководство системного администратора» по установке и настройке стенда для тестирования программного комплекса имитации функционирования СХД.

Индустриальным партнером - ООО «КНС ГРУПП»:  
- изготовлен экспериментальный образец СХД на платформе TATLIN в конфигурации 1. Составлены Акт изготовления и протоколы проверки экспериментального образца СХД на платформе TATLIN в конфигурации 1.

- проведен мониторинг параметров и состояний экспериментального образца СХД на платформе TATLIN в конфигурации 1. Результаты мониторинга зафиксированы в разделах промежуточного отчета о прикладных научных исследованиях и экспериментальных разработок
- разработаны частные технические задания на экспериментальный образец СХД на платформе TATLIN в конфигурациях 2 и 3 в соответствии с ГОСТ 34.602-89
- получены фактические системные журналы по работе системы хранения данных в разных состояниях.

На основании построенных моделей был предложен и описан алгоритм диагностики сбоев в работе СХД, основанный на сложной комбинации подходов анализа временных рядов, нелинейной фильтрации, деревьев выбора, корреляционного и статистического анализа. Настройка параметров таких алгоритмов на реальные данные СХД – трудоемкий и долгий процесс. Для ускорения и упрощения этого процесса была начата разработка стенда тестирования алгоритмов, который позволит настраивать параметры алгоритмов на имеющиеся статистические данные в автоматическом режиме.

Предварительный анализ результатов показал качественное согласие моделей с общими наблюдаемыми на практике характеристиками работы типовых СХД. Так, при исправности только одного сервера из четырех в ходе имитационного прогона наблюдалась средняя нагрузка носителя информации (дискового шасси HDD) на уровнях 0.15 и 0.14 МБ/с для первой и второй полок соответственно. При исправности двух серверов в ходе имитационного прогона средняя нагрузка носителя информации (дискового шасси HDD) повысилась до уровней 0.202 и 0.201 МБ/с для первой и второй полок соответственно. Добавление третьего исправного сервера слабо изменило среднюю нагрузку носителя информации, которая составила 0.204 и 0.202 МБ/с соответственно. Полностью исправная СХД с четырьмя серверами выдала среднюю нагрузку носителя информации на уровнях 0.248 и 0.243 МБ/с соответственно. Кроме того, появилась нагрузка на полку с SSD дисками, составляющая 0.001 МБ/с.

В целом, подобный рост загрузки качественно соответствует реальным показателям производительности, зависящим от количества задействованных компонент СХД.

Выполненный комплексный анализ результатов мониторинга параметров и состояний СХД – на базе ЦКП СКЦ «Политехнический» (СПбПУ) и ЭО СХД на платформе TATLIN (ООО «КНС ГРУПП») в конфигурации 1 позволил сформировать базу статистических данных о параметрах и состояниях аппаратных и программных компонентов СХД для выполнения имитационного и системно-динамического моделирования функционирования СХД.

Разработанный комплекс имитационных моделей функционирования СХД на платформе TATLIN позволит Индустриальному партнеру:

- осуществлять предсказательную аналитику состояний СХД различных конфигураций с учетом вариабельности параметров системы и режимов функционирования;
- обеспечить повышение надежности СХД, в том числе - уменьшение времени на восстановление в случае отказа дисков и возможность предсказания критической ситуации.

Разработанный новый алгоритм диагностики СХД, основанный на дискретно-событийной и стохастической моделях СХД позволит осуществлять контроль работоспособности СХД и ее компонентов, прогнозирование сбоев и предотвращение критических ситуаций в режиме реального времени для новых видов отказоустойчивых систем хранения данных на платформе TATLIN, выпуск которых будет осуществлен в 2018 году.

Экспериментальный образец СХД на базе платформы TATLIN, изготовленный Индустриальным партнером на 1 этапе, включает высокоскоростную отказоустойчивую фабрику PCI-Express, которая в совокупности с технологией виртуализации SR-IOV обеспечивает рекордную производительность Zero-Copy операций ввода-вывода.

Полученные результаты полностью соответствуют требованиям технического задания к выполняемому проекту и плана-графика к работам и отчетным документам, которые подтверждаются актом изготовления экспериментального образца и промежуточным отчетом.

По результатам аналитического обзора и патентных исследований установлено, что предлагаемые технические решения позволяют обеспечить предсказательную аналитику для сбоев и отказов и предотвращение критических ситуаций в СХД, соответствующие мировому уровню отказоустойчивости для данного класса сложных вычислительных систем. СХД на платформе TATLIN не имеют прямых аналогов на российском рынке и предназначены для ускоренной обработки больших объемов данных.

### **3. Охраноспособные результаты интеллектуальной деятельности (РИД), полученные в рамках прикладного научного исследования и экспериментальной разработки**

В рамках выполнения работ Этапа 1 создание охраноспособных РИД не предусмотрено.

### **4. Назначение и область применения результатов проекта**

Результаты проекта могут применяться в различных областях, например:

- Информационно-телекоммуникационные услуги;
- Хранение персональных данных;
- Хранение данных российских производственных предприятий;
- Хранение данных российских интернет-компаний, мессенджеров, соцсетей, компаний, обеспечивающих функционирование почтовых и поисковых сервисов;
- Хранение данных правоохранительных организаций;
- Хранение данных для тяжелой и легкой промышленности.

Полученные результаты патентного исследования будут использованы для подготовки РИД по теме проекта. Разработанные имитационные модели будут использованы при разработке алгоритма и программного комплекса прогнозирования сбоев для предотвращения критических ситуаций, в том числе деградации производительности и потери данных при выполнении операций чтения/записи в СХД.

Выполненные разработки будут:

- способствовать повышению эффективности мониторинга параметров СХД;
- обеспечивать определение предотказного состояния СХД и ее компонентов для предотвращения критических ситуаций;
- обеспечивать сокращение времени на принятие решения при наличии отказов в процессе функционирования СХД в условиях различных режимов эксплуатации и влияния внешних факторов – температуры воздуха, относительной влажности, давления и вибрации.

## 5. Эффекты от внедрения результатов проекта

Снижение риска потери данных, снижение издержек на хранение единицы объема данных, снижение трудоемкости сбора и обработки данных мониторинга, снижение скорости реагирования и повышение уровня эффективности принятия решения о предотвращении критических ситуаций за счет наличия статистических данных о вероятностном распределении их наступления в зависимости от набора параметров.

## 6. Формы и объемы коммерциализации результатов проекта

Разрабатываемая в рамках проекта система хранения данных, может быть использована, в решении задач корпоративных и государственных заказчиков, индустрии ИТ, а именно:

- операторами связи;
- правоохранительными органами;
- российскими интернет-компаниями;
- компаниями, обеспечивающими функционирование почтовых и поисковых сервисов;
- игровой индустрии;
- промышленностью и т.д.

Система разрабатываемая в рамках данного проекта имеет обширные возможности по коммерциализации, так как сбор и хранение данных, является неотъемлемой частью любой индустрии. Также, опыт в разработке систем мониторинга сложных вычислительных систем будет использован в образовательном процессе и при выполнении научно-исследовательских и опытно-конструкторских работ по заказам российских и зарубежных компаний. Разрабатываемые экспериментальные образы аппаратно-программных комплексов прогнозирования сбоев войдут в состав продуктов на базе платформы TATLIN в парадигме «функция определяет форму», что предполагает, что функциональность продуктов будет в наибольшей степени востребована потребителями:

- гибридная система хранения данных с унифицированным протоколом доступа (Hybrid Unified Storage) (2018);
- система хранения данных на основе флеш-памяти (All-Flash система) (2018);
- системы архивного хранения больших объемов с оригинальной технологией холодного хранения (2019).

На данный момент, предполагается использовать разрабатываемую систему хранения данных на мировом и внутреннем рынках. Рынок хранения данных занимает второе место среди расходов на ИТ и составляет примерно 23% от всех расходов.

В последующие 5 лет, рынок сбыта систем хранения данных на мировом рынке может вырасти на 250% (50% в год) и на 178,5% (37,5% в год) для внутреннего рынка.

## 7. Наличие соисполнителей

2017 год - федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет «Высшая школа экономики».

федеральное государственное автономное образовательное учреждение высшего образования "Санкт-Петербургский политехнический университет Петра Великого"

проректор по научной работе

(должность)

Сергеев В.В.

(фамилия, имя, отчество)

(подпись)

руководитель работ по проекту

(должность)

М.Л.



М.Л.

(подпись)

Болсуновская М.В.

(фамилия, имя, отчество)