

Приложение М



МИНОБРНАУКИ РОССИИ

ФИЛИАЛ

**ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО УЧРЕЖДЕНИЯ «ФЕДЕРАЛЬНЫЙ
НАУЧНЫЙ ЦЕНТР НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ
ИНСТИТУТ СИСТЕМНЫХ ИССЛЕДОВАНИЙ РОССИЙСКОЙ АКАДЕМИИ НАУК»**

Межведомственный суперкомпьютерный центр

Российской академии наук

(МСЦ РАН – филиал ФГУ ФНЦ НИИСИ РАН)



УТВЕРЖДАЮ

Директор МСЦ РАН – руководитель

ИТ-вычислительными ресурсами

МСЦ РАН – филиала ФГУ ФНЦ

НИИСИ РАН

Б.М. Шабанов

«30» 09 2020 г.

Технический отчет

Разработка системы личных кабинетов

Москва

2020

ИСПОЛНИТЕЛИ:

Ведущий научный сотрудник

 31.08.2020 А. А. Рыбаков

Младший научный сотрудник

 31.08.2020 М. Ю. Воробьев

СОДЕРЖАНИЕ

1	Перечень документов и основание для разработки системы личных кабинетов.	4
2	Цель и перечень выполненных задач по «Разработке системы личных кабинетов».....	4
3	Описание функциональных возможностей пользователей в системе личных кабинетов.....	4
4	Схема взаимодействия системы личных кабинетов	6
5	Модель данных для системы осведомленности о функционировании Центра.....	7
6	Мониторинг использования вычислительных ресурсов Центра.....	11
Приложение. ТЗ №1 от 20.01.2020г. Разработка системы личных кабинетов		

1 Перечень документов и основание для разработки системы личных кабинетов.

- 1.1. Основанием для выполнения работы является Соглашение о предоставлении из федерального бюджета грантов в форме субсидий в соответствии с пунктом 4 статьи 78.1 Бюджетного кодекса Российской Федерации №075-15-2019-1636 от 05 ноября 2019г.
- 1.2. Техническое задание №1 от 20.01.2020г. Разработка системы личных кабинетов.

2 Цель и перечень выполненных задач по «Разработке системы личных кабинетов»

- 2.1. Целью выполнения данной работы является создание системы личных кабинетов для информационной поддержки пользователей и автоматизации обработки информации. Система личных кабинетов позволит упростить предоставление вычислительных ресурсов центром коллективного пользования МСЦ РАН – филиала ФГУ ФНЦ НИИСИ РАН Организациям (далее Центр), проводящим научные исследования по приоритетным направлениям научно-технологического развития Российской Федерации в рамках Проектов в интересах Российской Федерации.
- 2.2. В рамках проведения данной работы были выполнены следующие задачи.
 - 2.2.1. Разработана система подачи заявок на Проект с помощью WEB интерфейса системы личных кабинетов ответственным от Организации/Научным руководителем/Ответственным исполнителем.
 - 2.2.2. Разработана модель данных получения и анализа статистической информации в зависимости от ролей пользователей системы личных кабинетов как по вычислительным ресурсам, так и по публикациям в рамках Проектов.
 - 2.2.3. Оформление отчетных материалов о проведенных научных работах и расчетных исследованиях в рамках Проектов с использованием системы личных кабинетов.

3 Описание функциональных возможностей пользователей в системе личных кабинетов.

Система личных кабинетов создана на основе открытой платформы для WEB - разработки. С помощью разработанного интуитивно понятного пользователям

интерфейса системы личных кабинетов можно выполнять следующие задачи в зависимости от роли пользователя. Система различает следующие роли пользователей:

3.1. **Ответственный от Организации.** При планировании Организацией осуществлять расчетные исследования в рамках Проекта на вычислительных ресурсах Центра между Организацией и МСЦ РАН заключается Договор о присоединении. В Договоре указываются данные Ответственного от Организации и на основании этих данных вводится учетная запись содержащая Логин и первоначальный пароль (все представленные первоначальные пароли в последствие меняются пользователями) для входа в систему личных кабинетов. Ответственный от Организации может осуществлять следующие функции:

- подавать заявки на Проект;
- отслеживать статус ранее поданной заявки;
- аннулировать ранее поданную заявку;
- запрашивать статистику по всем Проектам Организации;
- расторгать Договор присоединения.

Ответственный от Организации имеет возможность в системе личных кабинетов заполнить предварительную упрощенную форму заявки на Проект содержащей данные Руководителя проекта/ответственного исполнителя с целью заведения им учетных(ой) записей в системе и до заполнения заявки на проект по установленной форме с перечнем всех необходимых данных. Частично заполненная заявка на Проект не будет рассматриваться на НТС Центра.

3.2. **Руководитель проекта/ответственный исполнитель** (имеют одинаковые возможности в системе личных кабинетов) и могут осуществлять следующие функции

- подавать заявку на Проект;
- подавать и корректировать состав участников Проекта, которым необходим доступ к вычислительным ресурсам;
- заполнять Форму 1 по правилам работы пользователя Центра;
- подавать запрос на получение статистики использования вычислительных ресурсов

в рамках Проекта;

- формировать отчет о проведенных работах с возможностью продолжения/закрытия Проекта.

3.3. *Участник проекта* имеет следующие функции в системе:

- заполнять Форму 1 по правилам работы пользователя Центра;
- подавать запрос на получение статистики использования вычислительных ресурсов.

3.4. *Аналитик*. Сотрудник ведущий статистику по предоставленным вычислительным ресурсам Организациям и составляющим сводную отчетность по работе Центра перед курирующей организацией. Аналитик имеет следующие функции в системе:

- получение сводной статистики по текущему состоянию всех Проектов Организаций в целом;
- получение данных о задействовании вычислительных ресурсов Центра.

4 Схема взаимодействия системы личных кабинетов

В рамках данной работы разработана схема взаимодействия пользователей системы личных кабинетов в специализированном программном обеспечении. Данная модель описывает все функциональные возможности пользователей относящиеся к различным ролям, что является событиями в системе, а также схема учитывает WEB интерфейс и оборудование. В силу того, что всю схему отобразить в отчете не представляется возможным ввиду её громоздкости, здесь будут приведены отдельные взаимосвязанные с основной схемой выдержки по возможностям пользователей в «критическом» пути предоставления вычислительных ресурсов и проведения Проекта. На рисунке 1 приведена схема критического пути и функциональные возможности с точки зрения ответственного от организации.

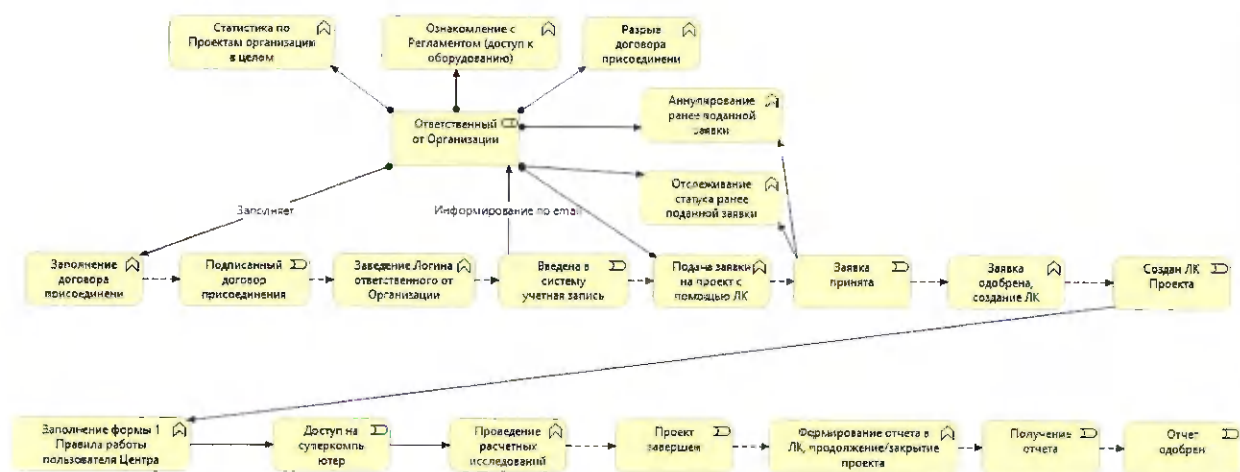


Рисунок 1.

На рисунке 2 приведена схема функциональных возможностей научного руководителя проекта и ответственного исполнителя.

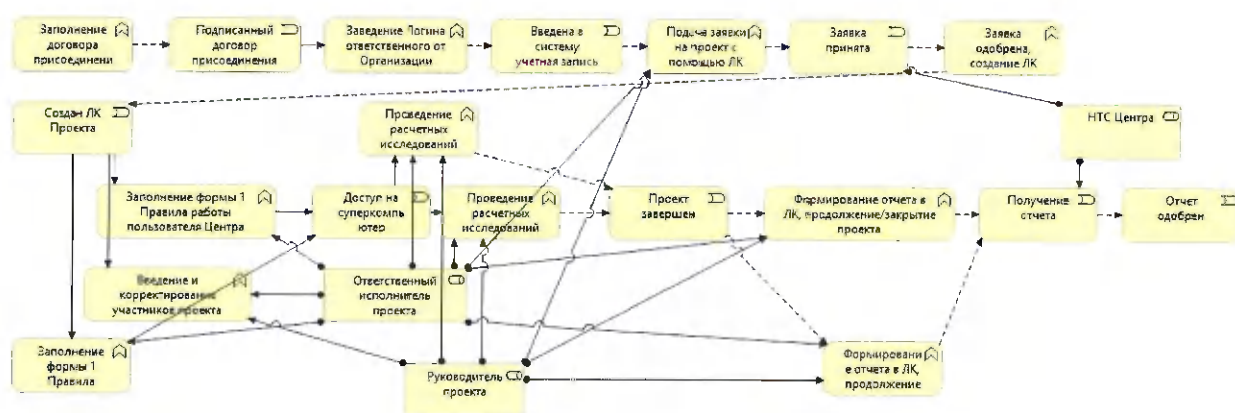


Рисунок 2

Разработанные схемы отвечают требованиям, приведенным в договоре присоединения и нормативным документам Центра по функционалу для пользователей системы личных кабинетов.

5 Модель данных для системы осведомленности о функционировании Центра

Помимо упрощения, с использованием системы личных кабинетов, регистрации пресектов и доступа пользователей к вычислительным ресурсам Центра, также имеется возможность получать и анализировать корректную, целостную, связанную информацию по функционированию Центра. В зависимости от роли пользователя, вошедшего в систему личных кабинетов, система предоставляет различные функции, которые описаны выше, а также доступ к различным уровням статистики. Для пользователя Проекта - статистику по

своим расчетным исследованиям; для научного руководителя и ответственного исполнителя всю информацию по Проекту/Проектам в которых он принимает участие в данном качестве; для ответственного от Организации – всю статистику по Проектам Организации; аналитик всю статистику по всем Проектам организаций.

Основными ресурсами Центра являются его вычислительные мощности (вычислительные сегменты, состоящие из отдельных вычислительных узлов). Наиболее важной задачей, стоящей перед системой осведомленности (или системой мониторинга), является накопление и визуализация информации о состоянии каждого вычислительного ресурса в каждый момент времени. В рамках функционирования Центра минимальной единицей вычислительных ресурсов является узел, который может быть отведен на выполнение расчетной задачи на любой заданный промежуток времени. Таким образом, может быть проведен мониторинг работоспособности каждого вычислительного узла, а также в целом более крупных единиц.

С другой стороны, кроме обеспечения технически эффективного функционирования аппаратуры Центра требуется выполнять отслеживание эффективности использования вычислительных ресурсов пользователями. На сегодняшний день основным показателем эффективности использования суперкомпьютеров является публикационная активность пользователей: количество и качество публикаций, выполненных ими с использованием вычислительных ресурсов Центра. Таким образом, учет и обработка публикаций пользователей является второй важнейшей составляющей системы осведомленности о функционировании Центра.

Итак, будем рассматривать модель данных как совокупность информации о запусках задач пользователей через системы управления заданиями (СУПЗ и SLURM), а также информации о публикациях пользователей, связанных с организациями, проектами, в рамках которых выполняются расчеты, а также с набором программного обеспечения, которое было использовано при проведении расчетов. Основным принципом организации данных является полнота и унификация этих данных, все данные представлены в одном месте и доступ к ним осуществляется по единой схеме. Контроль за доступом к данным для конкретного пользователя осуществляется с помощью ролей данного пользователя. При извлечении статистической информации одновременно извлекается информация о роли запросившего данную информацию пользователя. С помощью ролей пользователя извлеченная статистическая информация фильтруется и подается для визуализации уже не в полном виде, а только в виде того набора, который должен быть доступен запросившему его пользователю (так, например, пользователь конкретной организации не имеет доступа к информации о публикациях пользователей других организаций или к статистике

использованного ими расчетного времени; наоборот, аналитик Центра, имеет доступ к любой информации и в этом случае она не подвергается фильтрации). После фильтрации информация передается на визуализацию и отображается в объекте на веб-форме (все объекты визуализации являются независимыми, на одной странице может быть отображено произвольное количество объектов визуализации) (см. рис. 3). Основными объектами визуализации являются круговые диаграммы, графики, гистограммы, календарные схемы. Возможно комбинирование различных объектов визуализации.

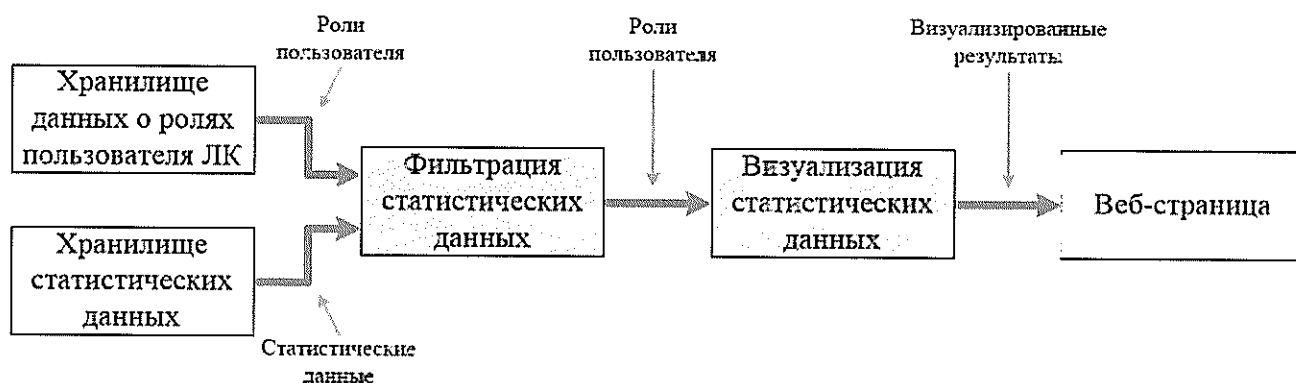


Рисунок 3. Модель обработки статистических данных при выполнении визуализации.

Информация о запуске задачи пользователя представлена с помощью объектов, содержащих следующие составляющие: время постановки задачи в очередь, время начала и завершения выполнения задания, описание запрошенных и потребленных во время счета ресурсов (количество узлов, промежуток времени), имя пользователя, название организации, идентификатор проекта, информацию о запускаемой задаче, дополнительные опции и другие параметры. На основании информации о запусках задач пользователей, а также других данных формируется информация о состоянии вычислительных узлов. Для вычислительного узла выделяются его различные состояния, и в числе основных можно перечислить следующие: R (Run) – узел находится в состоянии работы и выполняет некоторую задачу в режиме общей очереди, E (Error) – в работе узла наблюдается некоторая ошибка, в текущий момент он не может быть использован для выполнения заданий; O (Ordered) – состояние вычислительного узла в будущем, на данный момент времени узел зарезервирован для выполнения задачи, но этот момент времени еще не настал; A (Away) – узел отключен от общей очереди и в данный момент выделен в отдельный сегмент, предоставляемый некоторому пользователю в режиме монопольного использования (для статистики Центра более точное состояние узла в этот момент времени интереса не представляет); N (None) – узел не находится ни в одном из перечисленных выше состояний (если не определены какие-то специальные состояния, то этот статус означает, что узел исправен, но простаивает). Таким образом для отслеживания состояния всех узлов в базе данных накапливается множество кортежей со следующей информацией: {NodeName,

TimeStart, TimeEnd, State}, то есть определяются интервалы времени пребывания узла в некотором состоянии. Все интервалы времени, которые не описаны данными кортежами по умолчанию считаются инициализированными состоянием N (None).

Данные о публикации пользователей также представлены набором кортежей, основными полями при описании публикации являются следующие: язык публикации, название публикации, год публикации, список авторов, значение ISSN, название издания или журнала, другие реквизиты публикации, список учетных баз для группирования публикаций (сюда входят и базы цитирования и признак учета публикации в отчетности по ЦКП и могут входить другие категории), DOI, идентификатор проекта, описание проекта, название организации и перечень программных пакетов, которые были использованы при выполнении запусков в рамках работы над публикацией, бинарные данные публикации (возможность привязать полный текст или иллюстрации, набор бинарных данных также представляет собой кортеж с произвольным набором плей). Пример кортежа с описанием данных о публикации имеет следующий вид:

```
Publication = {Language = "английский",
               Name = "Molecular Dynamics Study of Polyethylene: Anomalous
                     Chain Mobility in the Condis Phase",
               Year = 2019,
               AuthorsList = ["Sultanov V.I.", "Atrazhev V.V.",
                             "Dmitriev D.V.", "Erikhman N.S.",
                             "Furrer D.U.", "Burlatsky S.F."],
               ISSN = "0024-9297",
               Pubhouse = "Macromolecules",
               Included = ["CKP", "RINC", "WoS", "Scopus"],
               DOI = "10.1021/acs.macromol.8b02710",
               ProjectId = "ibhf4",
               ProjectDescription = "Теоретическое исследование
                                   электрофизических, электрохимических и
                                   оптических свойств проводящих
                                   полимеров, композиций, сильно
                                   коррелированных электронных систем
                                   пониженной размерности и наноразмерных
                                   систем",
               Organization = "ИБХФ РАН",
               Programs = [],
               ELOB = {...}}.
```

Кроме данных полей в описание публикации могут быть включены другие данные, если существует потребность получения по ним статистики. Построение диаграмм и гистограмм по данным публикаций выполняется в унифицированном виде, как будет показано ниже. Построение визуализации статистической информации может быть выполнено по любым категориям, которые представлены в данных публикаций в виде отдельных значений (пример: поле Pubhouse) или списков значений (пример: поле Included).

Записи о состоянии вычислительных узлов и описание публикаций хранятся в NoSQL

базе данных и выгружаются из нее в объектную модель, реализованную на JavaScript, дальнейшая визуализация выполняется на стороне клиента средствами JavaScript. Основной функционал отображения данных базируется на генерации трех основных визуальных объектов: круговая диаграмма (выполняющая визуализацию по словарю данных {<ключ>, <числовое значение>}), горизонтальная гистограмма (также работающая со словарем данных {<ключ>, <числовое значение>}), прямоугольная диаграмма (также работает со словарем данных {<ключ>, <числовое значение>} и выполняет закраску частей прямоугольника пропорционально соответствующим значениям ключей из словаря), календарная сетка (объект, позволяющий генерировать наполнение для каждого дня месяца, в том числе данное наполнение может быть представлено другими примитивами визуализации, в частности круговыми и прямоугольными диаграммами).

6 Мониторинг использования вычислительных ресурсов Центра

Для мониторинга моментальной загрузки сегментов вычислительных систем Центра предусмотрена прямая выгрузка из систем очередей, снимок экрана которой представлен на рис. 4.

Mon Oct 5 12:32:01 MSK 2020				Queue state at Mon Oct 5 12:32:01 2020			
Node	User	Task	CPU Load	Current schedule:	2	Hold	
n05p052	levpr	clk.vasp_std.1	0.16	Accumulation period:	168 hours.		
n05p053	levpr	clk.vasp_std.1	0.01	---	05.10.20 08:30: R=1 total-1440, debug-0/20, packet-1440, priority scale		
n05p054	orishina	clk.vasp_std.2	0.01	clk.g09e01avx.sh.2 :	khramem 4.12 0* 48 1440/1440 finishing		
n05p055	orishina	clk.vasp_std.2	0.01	clk.g09e01avx.sh.1 :	khramem 4.11 0* 48 1440/1440 finishing		
n05p056	orishina	clk.vasp_std.2	0.01	clk.g09e01avx.sh.1 :	maxast 2.10 0* 48 1440/1440 finishing		
n05p057	orishina	clk.vasp_std.2	0.01	clk.g09e01avx.sh.1 :	galimov 3.9 0* 48 1440/1440 finishing		
n05p058	levpr	clk.vasp_std.1	0.01	clk.vasp_std.3 :	komarov 2.8 0* 288 1440/1440 finishing		
n05p059	levpr	clk.vasp_std.1	0.01	clk.vasp_std.2 :	orishina 3.7 0* 192 1440/1440 finishing		
n05p060			0.01	clk.vasp_gam.2 :	vvp 5.6 0* 48 1200/1200 finishing		
n05p061	omega	clk.nmd1.sh.1	0.01	clk.vasp_gam.1 :	vvp 5.5 0* 48 1200/1200 finishing		
n05p062	komarov	clk.vasp_std.3	0.01	clk.vasp_std.1 :	levpr 4.4 0* 288 1440/1440 finishing		
n05p063	khramem	clk.g09e01avx.sh.2	0.01	clk.vasp_gam.1 :	avmres 3.3 0* 48 1200/1200 finishing		
n05p064	komarov	clk.vasp_std.3	0.01	clk.nmd1.sh.1 :	iredjan 4.2 0* 48 1440/1440 finishing		
n05p065	levpr	clk.vasp_std.1	0.01	clk.nmd1.sh.1 :	omega 3.1 0* 48 1440/1440 finishing		
n05p066	komarov	clk.vasp_std.3	0.01	---	05.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p067	komarov	clk.vasp_std.3	0.01	---	05.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p068	vvp	clk.vasp_gam.1	0.01	---	05.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p069	komarov	clk.vasp_std.3	0.01	---	07.10.20 10:00: R=4 total-1440, debug-0/0, packet-1440, priority scale		
n05p070	komarov	clk.vasp_std.3	0.01	---	07.10.20 11:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p071	levpr	clk.vasp_std.1	0.01	---	08.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p072	iredjan	clk.nmd1.sh.1	0.01	---	08.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p073	vvp	clk.vasp_gam.2	0.01	---	09.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p074	galimov	clk.g09e01avx.sh.1	0.01	---	09.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p075			0.01	---	10.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
n05p076	khramem	clk.g09e01avx.sh.1	0.01	---	10.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 21:00: R=1 total-1440, debug-0/0, packet-1440, priority scale		
				---	11.10.20 08:30: R=1 total-1440, debug-0/0, packet-1440, priority scale		

На данном снимке представлено состояние сегмента Cascade Lake (Xeon 8268) на момент времени 05.10.2020 12:32. В информацию входит отображение состояния всех узлов сегмента (находящихся под управлением системы управления прохождением параллельных заданий Центра) в текущий момент времени. Кроме того, для каждого узла отображается имя работающего в данный момент пользователя, название исполняемого файла и уровень загрузки микропроцессора. Кроме того, предусмотрена возможность просмотра текущего состояния очередей, в том числе представлена следующая информация: количество ресурсов в очереди, количество доступных для использования процессоров. Для каждой задачи, находящейся в очереди отмечено имя исполняемого файла (или скрипта запуска), количество заказанных ресурсов, приоритет задания и состояние задачи.

На основании данных очереди формируются corteжи состояния узлов. Однако не все состояния узлов могут быть собраны из этих данных. В частности, состояние A (Away) должно быть занесено в базу данных инженером Центра при исключении узлов из-под управления системой прохождения параллельных заданий.

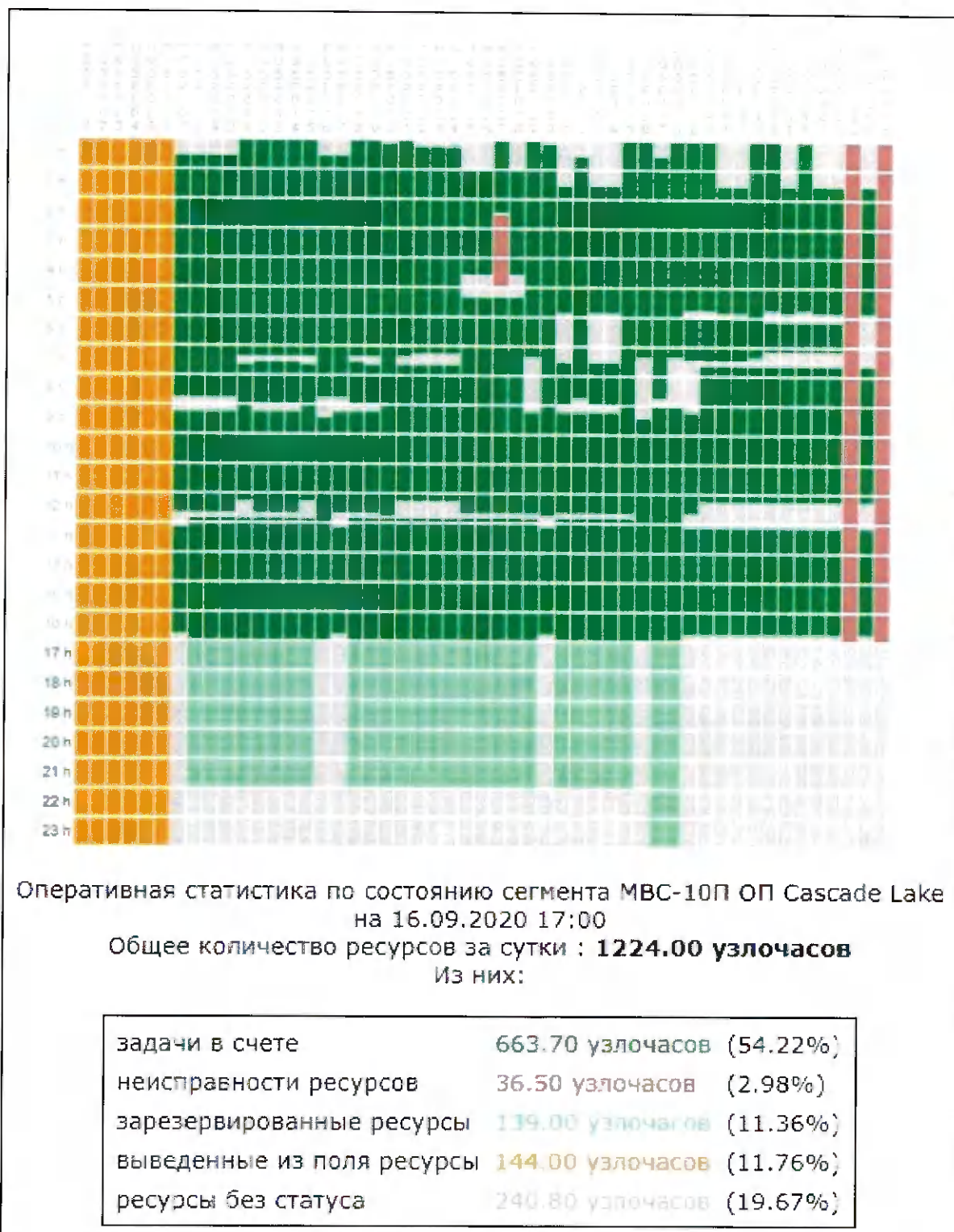


Рисунок 5. Снимок экрана с отображением информации о статусах узлов сегмента вычислительной системы, начиная с точки времени 00:00 до текущего момента.

На основании статистики о состоянии узлов может быть собрана полная картина состояния вычислительного поля, показанная на рис. 5. В данном случае показана суточная загрузка сегмента Cascade Lake, состоящего из узлов n05p001-n05p051. Отображение состояния вычислительного поля представлено матрицей, в которой столбцы соответствуют вычислительным узлам, а строки – отдельному часу в сутках.

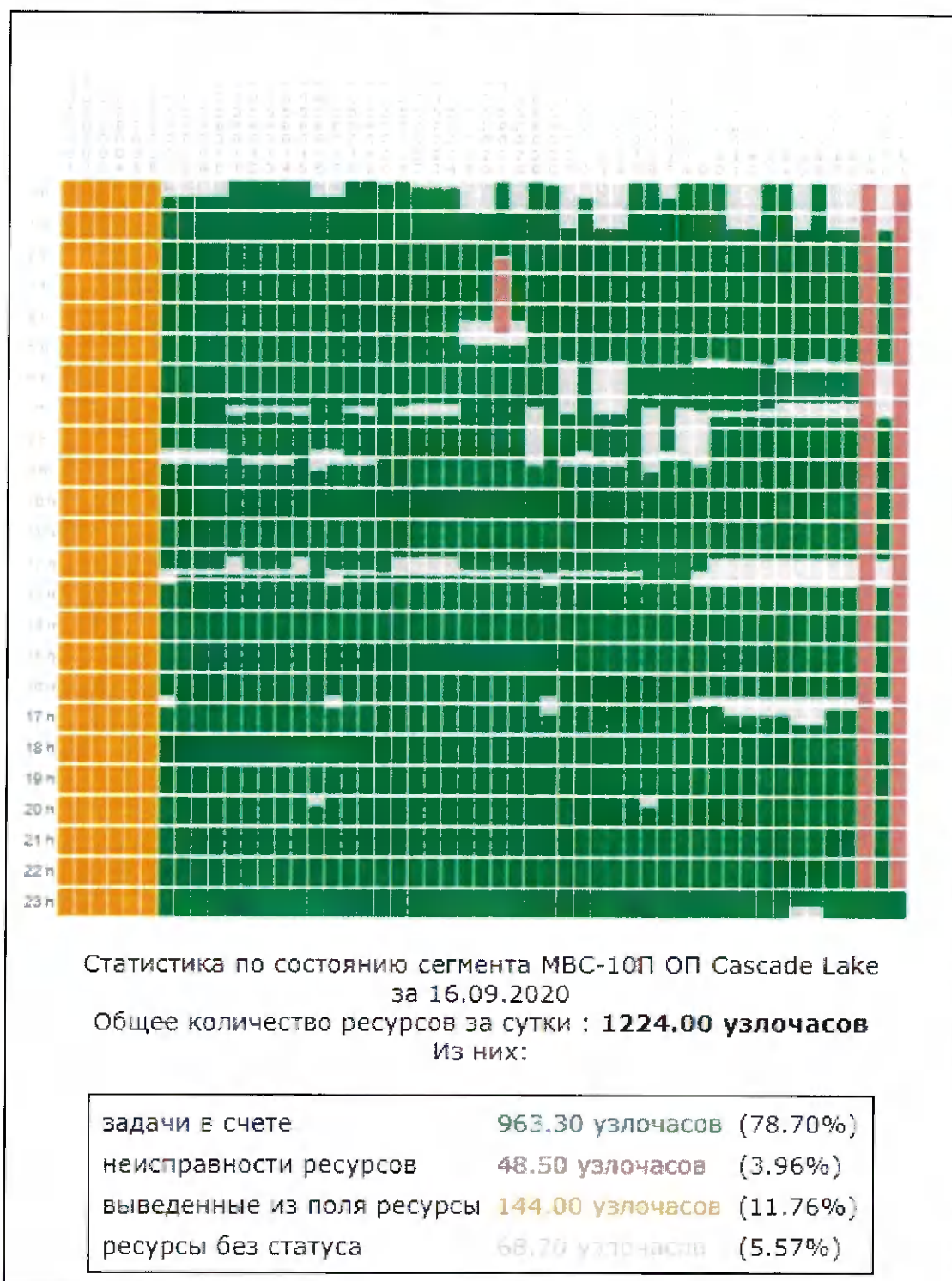


Рисунок 6. Снимок экрана с отображением информации о статусах узлов вычислительной системы за сутки (начиная в точки времени 00:00 и заканчивая точкой 23:59).

В зависимости от статуса вычислительного узла, часть ячейки матрицы окрашивается в соответствующий цвет: R (Run) – зеленый, O (Ordered) – светло-зеленый, E (Error) – красный, A (Away) – золотистый, N (None) – серый. При этом под матрицей состояний выводится сводная информация о статусах сегмента Рис. 6.

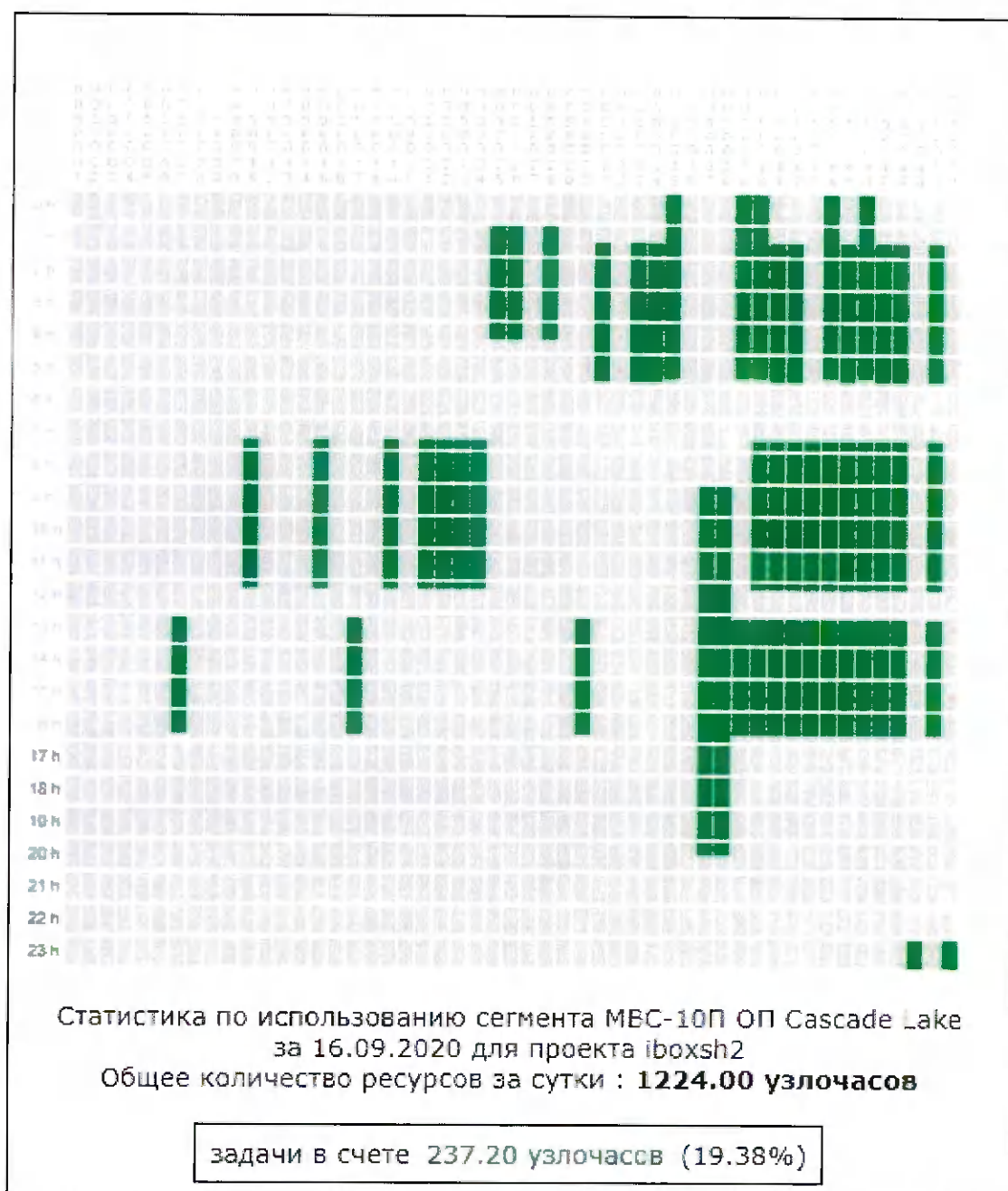


Рисунок 7. Снимок экрана с отображением информации о запусках задач, относящихся к конкретному проекту, на узлах сегмента вычислительной системы за сутки.

Так как на рис. 5 статистика по вычислительному полю собирается до окончания суток (на момент времени 17:00), то единственным возможным статусом после указанного момента времени может быть O (Ordered) (за исключением N (None)). На рис. 4 показана та же статистка, но уже по завершению суток. Отметим, что показанная статистика соответствуют роли аналитика Центра, который имеет доступ ко всей информации. При получении данной информации пользователями конкретного проекта производится фильтрация данных, и пользователь получает информацию, касающуюся только его конкретного проекта, как этого показано на рис. 7 (при других настройках возможен показ информации о нерабочих узлах, однако доступ к информации о вычислениях в рамках

других проектов не предоставляется).

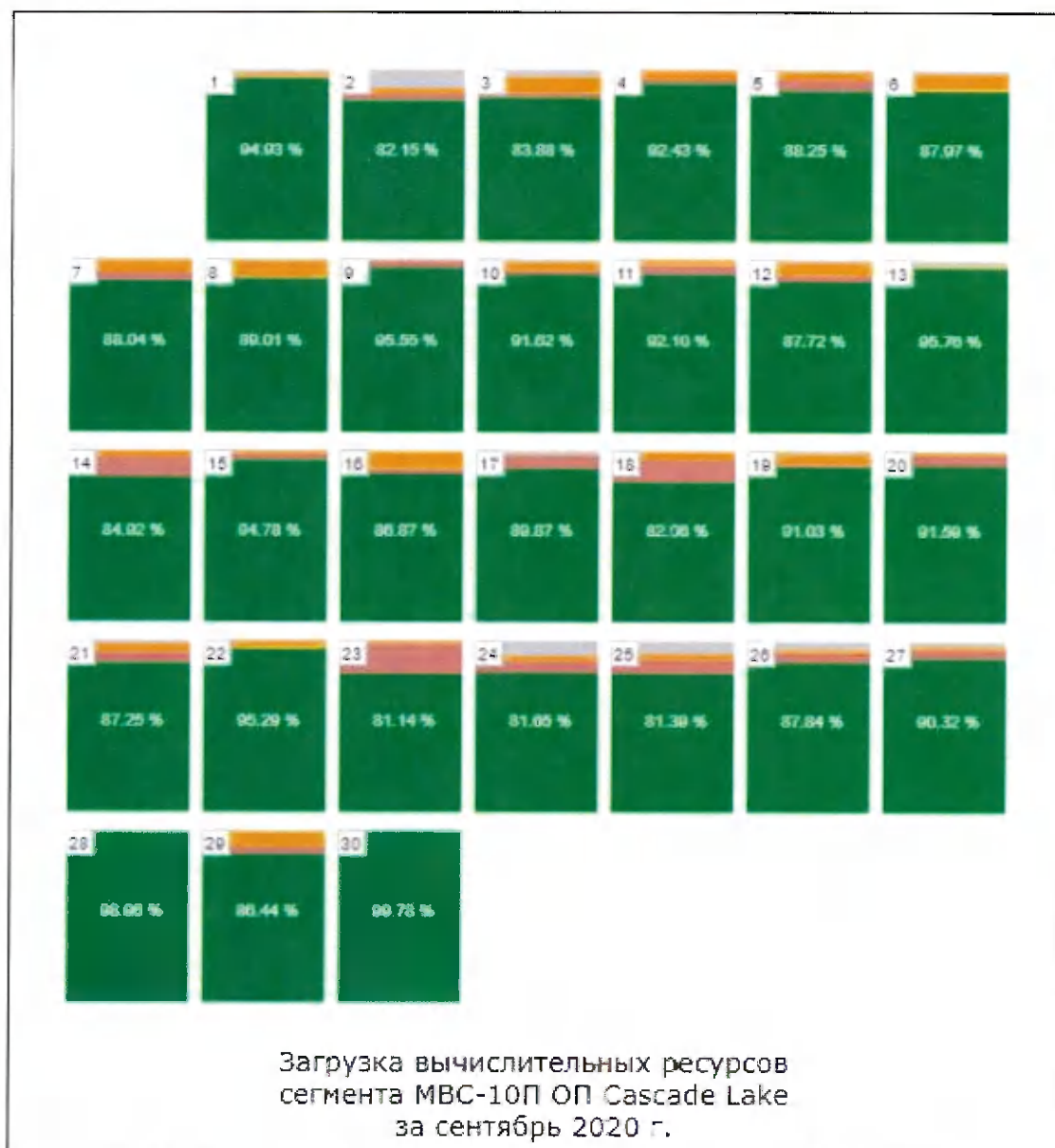


Рисунок 8. Снимок экрана с отображением картины загрузки узлов вычислительной системы за календарный месяц (отображение статусов с помощью прямоугольников).

Другим важным моментом является сбор и визуализация статистики работы сегмента за более длительный период. На рис. 8 продемонстрирована визуализация месячной загрузки сегмента Cascade Lake с помощью комбинированного визуального объекта. Данный объект представляет собой календарное расписание, каждая ячейка которого содержит информацию о одном отдельном дне. Внутри каждой ячейки календарного расписания генерируется отдельный визуальный объект типа "прямоугольная диаграмма". Таким образом, с помощью данной визуализации можно наблюдать динамику загрузки, сбоев и выведения из вычислительного поля узлов на протяжении календарного месяца. По схожему принципу строятся календарные расписания в рамках года с разбивкой на месяцы,

недели или отдельные дни. Годовое календарное расписание также может комбинировать внутри себя отдельные расписания по месяцам.

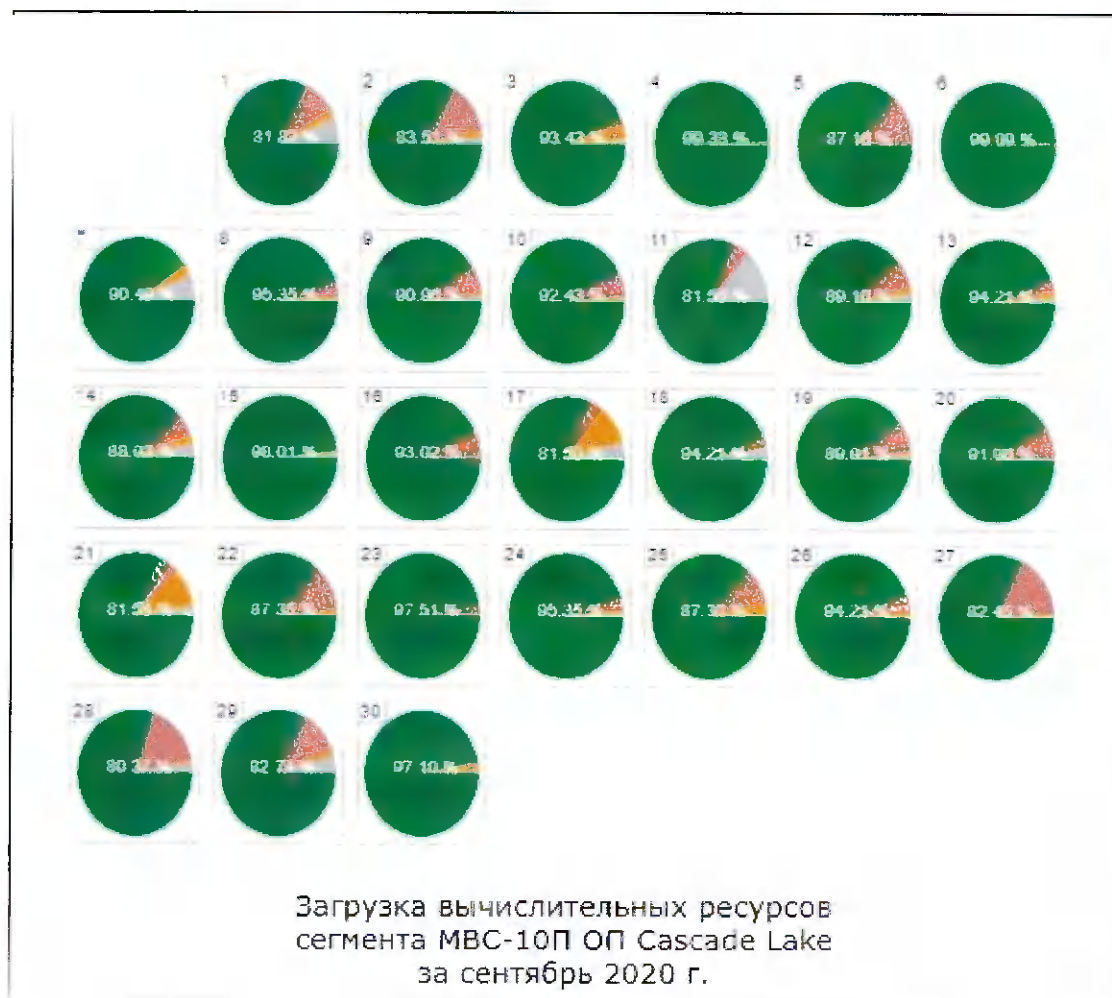


Рисунок 9. Снимок экрана с отображением картины загрузки узлов вычислительной системы за календарный месяц (отображение статусов с помощью круговых диаграмм).

На рис. 9 приведена другая визуализация того же самого календарного расписания, которое представлено на рис. 8. Однако, в данном случае вместо прямоугольных диаграмм использованы круговые. При реализации прямоугольных и круговых диаграмм предусмотрены три режима передачи цветов: массив фиксированных цветов с доступом по порядку, массив фиксированных цветов с круговым доступом и использование случайных цветов. Для использования диаграмм в составе комбинированного объекта по умолчанию используется массив фиксированных цветов, чтобы диаграммы для соседних дней календарного расписания использовали одинаковые цветовые раскладки. Кроме того, поддерживаются режимы построения диаграмм с сортировкой данных и без сортировки (опять, же в составе комбинированного объекта по умолчанию сортировка данных не применяется для поддержания однообразия отображения). Для избегания шумовых эффектов при построении диаграмм включена опция отсечения статистически незначимых данных (данная опция важна, когда словарь значений не является фиксированным). В качестве

[illegible]

На рис. 11 приведен другой пример применения горизонтальной гистограммы, но с полностью аналогичнее предыдущему, однако в данном случае анализируется не активность авторов публикаций, а востребованность программных пакетов. Для поддержки данного вида анализа предусмотрена возможность развития функционала по слиянию отдельных категорий. Например, в гистограмме можно наблюдать пакет Quantum Espresso со значением 8 и quantum espresso со значением 12. Очевидно, что это один и тот же пакет.

значения для которого должны быть объединены. Игнорирование регистра написания тут не поможет, так как пакеты Gaussian 09 и Gaussian 16 являются разными версиями одного и того же пакета, и должны быть объединены в одну запись. К тому же объединение записей должно происходить до отсеечения шума, так как иначе необъединенные названия могут быть отфильтрованы несмотря на то, что суммарно они могут иметь шанс пройти в финальную статистику.

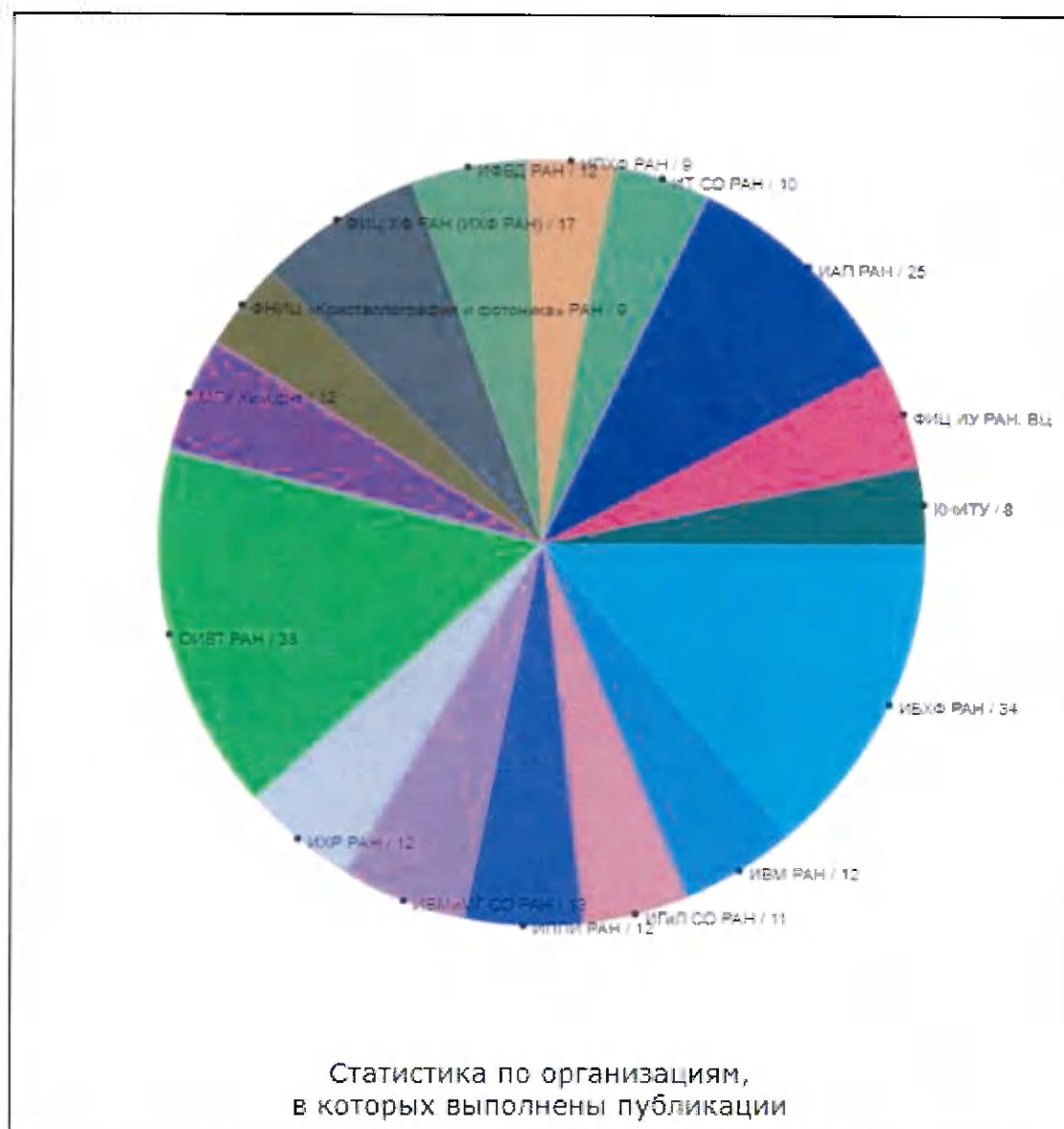


Рисунок 12. Снимок экрана с отображением статистической информации по распределению публикаций пользователей среди организаций.

Для анализа статистики кроме горизонтальных гистограмм могут быть использованы круговые диаграммы. На рис. 12 показан пример такой диаграммы, на которой выявляются институты – лидеры по количеству предоставленных публикаций. Так как количество институтов не фиксировано и может динамически меняться, то при построении данного объекта визуализации используется случайная раскраска секторов диаграммы,

выполняемая по формуле $RGB(28 + 200 * rand(), 28 + 200 * rand(), 288 + 200 * rand())$ для избегания чрезмерного затемнения и осветления диаграммы. При построении данной диаграммы использовалось отсечение шумов (организаций со слишком низким показателем количества публикаций), а также был включен режим отображения меток (для каждого сектора обозначена точка, рядом с которой приводится названия института и количество его публикаций). Данная диаграмма поддерживает режим инвертирования для обнаружения организаций с наименьшим количеством публикаций.

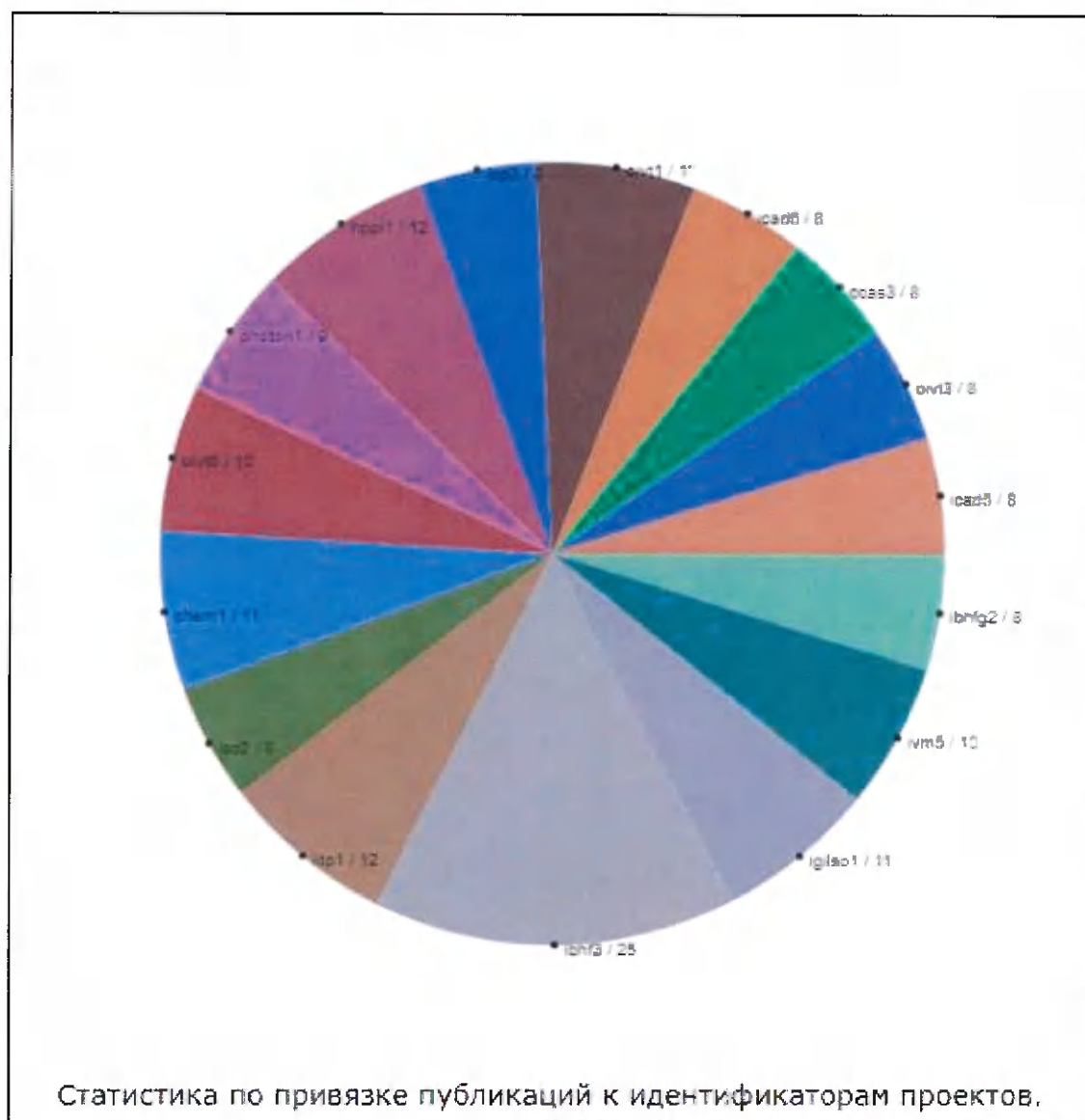


Рисунок 13. Снимок экрана с отображением статистической информации по распределению публикаций пользователей среди проектов.

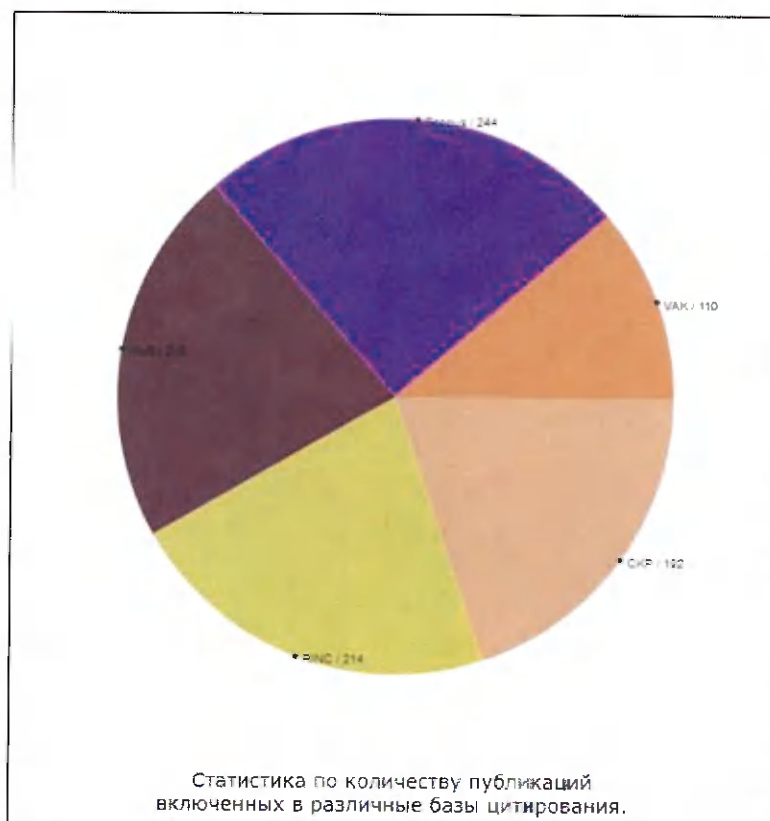


Рисунок 14. Снимок экрана с отображением статистической информации по распределению публикаций пользователей среди баз цитирования и привязки к отчетности ЦКП.

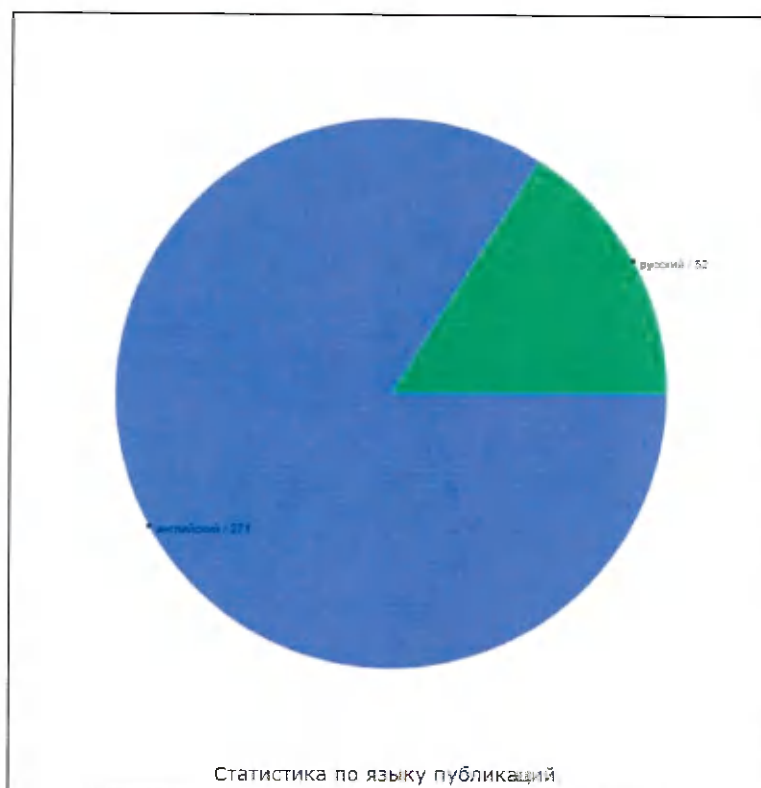


Рисунок 15. Снимок экрана с отображением статистической информации по распределению публикаций пользователей среди языков написания публикации.

На рис. 13, рис. 14 и рис. 15 приведены другие примеры построения круговых диаграмм, направленные на выявление наиболее активных проектов (в результате которых генерируется наибольшее число публикаций), получается распределение публикаций по различным базам цитирования и учета, а также строится распределение публикаций по языкам написания статей. Распределение статей по проектам также поддерживает инвертированное построение для обнаружения проектов, по которым не достигнуто заявленное количество публикаций.

Описанный функционал анализа публикационной активности доступен только для аналитиков Центра, так как пользователь, выполняющий расчеты в любом случае, не имеет доступа к данным о публикациях, выполненных за пределами его проекта.