

### Literatura

1. Zhukov A.O. *i dr.* Vysokotochnaya sistema navigatsionno-ballisticheskogo obespecheniya sverkhmalykh kosmicheskikh apparatov // Mekhanika, upravlenie i informatika. 2015. T. 7, № 2 (55). S. 332–339.
2. Zhukov A.O. *i dr.* Metod provedeniya obzora i variant postroeniya kosmicheskoy sistemy dlya monitoringa kosmicheskikh ob"ektov na okolozemnykh orbi-takh // Ekologicheskij vestnik nauchnykh tsentrov Chernomorskogo ekonomicheskogo sotrudnichestva (CHES). 2013. T. 3, № 4. S. 71–75.
3. Bondarenko A.P. *i dr.* O nekotorykh ogranicheniyakh korrekcirovki modeli sredy rasprostraneniya na baze informatsii GPS/GLONASS-priemnikov // Oboronnij kompleks – nauchno-tehnicheskomu progressu Rossii. 2018. № 1 (137). S. 57–63.
4. Bondarenko A.P. *i dr.* K voprosu vybora minimal'nogo sostava spetsializirovannogo programmno-apparatnogo kompleksa i vremennykh intervalov ego raboty // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2017. № 2. S. 73–82.

DOI: 10.25586/RNU.V9187.19.03.P.024

УДК 004.042

А.С. Марковский, Н.И. Свеколкин, С.А. Шаров

РЕШЕНИЕ ЗАДАЧ ОБРАБОТКИ ДАННЫХ  
РАДИОЛОКАЦИОННОГО НАБЛЮДЕНИЯ  
С ПОМОЩЬЮ МАСШТАБИРОВАНИЯ  
РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ

Представлен анализ применения масштабирования распределенных вычислительных технологий при обработке радиолокационных данных. На основании полученных результатов исследований сформулированы предложения по применению параллельных вычислений для групповой обработки радиолокационных изображений.

*Ключевые слова:* распределенные вычислительные технологии, радиолокационные данные, облачная инфраструктура.

A.S. Markovskiy, N.I. Svekolkina, S.A. Sharov

THE TASK OF DATA PROCESSING  
OF RADAR OBSERVATIONS  
USING SCALE DISTRIBUTED COMPUTING

The article presents an analysis of the application of scaling of distributed computing technologies in the processing of radar data. On the basis of the results of the research, proposals for the use of parallel calculations for group processing of radar images are formulated.

*Keywords:* distributed computing, radar data, cloud infrastructure.

### Введение

Данные радиолокационного зондирования Земли широко применяются во многих отраслях промышленности и науки, служат основой для решения самых различных задач. Дистанционное зондирование осуществляется тысячами сенсоров, установленных на спутниках, самолетах и беспилотных летательных аппаратах. Активность в данной деятельности постоянно нарастает, обеспечивая непрерывное поступление радиолокационных изображений (РЛИ) для их обработки. Особое место занимает информация, получаемая с космических аппаратов (КА) радиолокационной съемки.

Отдельной задачей является создание и эксплуатация распределенных систем обработки радиолокационных данных (РЛД). При этом важной проблемой построения таких систем является масштабируемость, т.е. способность системы или процесса обработки РЛД справляться с обработкой при увеличении рабочей нагрузки и добавлении ресурсов. Внедрение высокоскоростных и устойчивых протоколов связи в компьютерных сетях, распространение стандартов широкополосного Интернета и стабильность отмечаемой тенденции по наращиванию вычислительной мощности привели к развитию способов получения вычислительных и информационных услуг.

В общем понимании система называется масштабируемой, если она способна увеличивать производительность пропорционально добавляемым дополнительным вычислительным ресурсам.

Рассмотрим процесс обработки РЛД. В классической технологии обработки РЛД действует наземная станция обработки данных зондирования, на которой производятся операции фильтрации, синтеза РЛИ, постобработки и географической привязки полученных снимков. При этом представление РЛД потребителям производится с использованием существующих наземных каналов передачи данных.

В перспективных космических средствах задачи синтеза изображения могут частично решаться на борту КА, а представление данных осуществляться по спутниковым каналам связи.

В настоящее время инфраструктура облачных вычислений (ОВ) является одной из перспективных и востребованных направлений развития распределенной обработки больших массивов РЛД. Например, ее используют в целях повышения оперативности и эффективности методов обработки интерферометрических данных радиолокационной съемки. Следует отметить, что оптимизации и повышению производительности процедур обработки РЛД, в том числе и за счет распараллеливания алгоритмов, посвящено большое количество научных работ [1; 2; 3; 4].

Облачную инфраструктуру необходимо рассматривать как интегратор распределенного исполнения программного кода для поступающих в потоковом режиме данных.

Основными достоинствами распределенной обработки на основе ОВ являются:

- возможность хранения и работы с большими объемами РЛД за счет применения распределенной файловой системы;
- высокий уровень масштабируемости, достигаемый за счет применения дополнительных узлов вычислительного кластера без вынужденного внесения трудоемких изменений в алгоритмы обработки;
- возможность работы с данными в потоковом режиме.

В качестве открытого решения для разработки распределенных приложений в современных проектах систем потоковой обработки активно внедряют платформу разработки распределенных приложений обработки данных Apache Spark [5]. Она имеет встроенные технологии машинного обучения (Spark ML), потоковой обработки (Spark Streaming), работы со структурированными данными (Spark SQL) и вычислений на графах (Spark GraphX). Ключевым понятием в Spark является указатель на распределенную коллекцию данных (RDD). Большинство операций над RDD не приводит к каким-либо вычислениям, а только создает «обертку», обещая выполнить операции только тогда, когда они понадобятся.

Основные преимущества:

- производительность, удобный программный интерфейс;
- параллельная обработка и анализ слабоструктурированных данных в оперативной памяти (ОП);
- поддержка четырех языков программирования: Scala, Java, Python, R;
- возможность кэширования данных и хранения промежуточных результатов вычислений в ОП;
- поддержка потоковой обработки данных.

Целью статьи является представление практических расчетов применения масштабирования распределенной обработки РЛД, формирование рекомендаций по их использованию в автоматизированных системах.

#### Постановка задачи масштабирования обработки радиолокационных данных

Рассмотрим концептуальную модель распределенной системы обработки РЛД как многоканальную систему массового обслуживания (СМО) с неограниченной очередью, где одновременно может быть открыто два и более каналов (рис. 1).

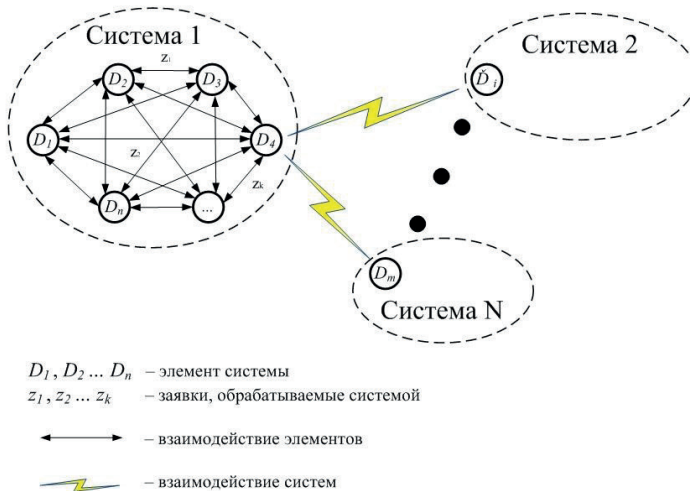


Рис. 1. Концептуальная модель распределенной системы обработки РЛД

Предполагается, что заявки ожидают в общей очереди и обращаются в первый освободившийся канал обслуживания.

Марковский А.С., Свеколкин Н.И., Шаров С.А. Решение задач обработки данных...

Система обработки РЛД имеет следующие состояния (пронумеруем их по числу заявок, находящихся в системе):  $S_0, S_1, S_2, \dots, S_k, \dots, S_n$ , где  $n$  – индекс состояния системы,  $S_k$  – состояние системы, когда в ней находится  $k$  заявок  $\{z_0, z_1, z_2, \dots, z_k\}$ .

Для многоканальной системы с неограниченной очередью, какой является система ОВ, должно выполняться следующее условие [6; 7; 8]:

$$\frac{r}{n} < 1, \tag{1}$$

где  $r$  – параметр загрузки системы (среднее число занятых каналов);

$n$  – минимальное количество каналов, при котором очередь не будет расти до бесконечности.

Зная, что результат исполнения конкретной заявки  $B$  в общем случае зависит от ряда различных факторов и условий, численность которых варьируется в каждом конкретном случае, обозначим его в следующем виде:

$$B = (N_{\text{испол}}, N_{\text{ожид}}, t_{\text{испол}}, I, R, Q_{\text{рес}}, f, M), \tag{2}$$

где  $N_{\text{испол}}$  – количество ожидающих исполнения заявок;

$N_{\text{ожид}}$  – количество исполняемых заявок;

$t_{\text{испол}}$  – время, отводимое на отработку заявки;

$I$  – приоритетность заявки;

$R$  – сложность (трудоемкость) выполняемого задания;

$Q_{\text{рес}}$  – имеющиеся свободные вычислительные ресурсы для исполнения;

$f$  – дестабилизирующие воздействия (форс-мажор, смена заявки и т.п.);

$M$  – рациональная организация исполнения (наличие ранее обработанных данных).

Среднее число занятых каналов рассчитывается по формуле

$$\rho = \frac{\lambda}{\mu}, \tag{3}$$

где  $\lambda$  – интенсивность входящего потока заявок;

$\mu$  – производительность каждого канала.

Состояние незанятости системы находим по формуле

$$\rho_0 = \left( 1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!(n-\rho)} \right)^{-1}, \tag{4}$$

где  $n$  – количество заявок.

Среднее число заявок в очереди рассчитывается следующим образом:

$$L_{\text{оч}} = \frac{\rho^n \rho_0}{n! n \left( 1 - \frac{\rho}{n} \right)^2}. \tag{5}$$

Среднее время нахождения заявки в очереди может быть выражено формулой

$$T_{\text{оч}} = \frac{1}{\lambda} L_{\text{оч}}.$$

Для определения оптимального числа узлов по обработке заявок, помимо показателей эффективности СМО, необходимо ввести критерий оптимизации. В качестве такого критерия, например, может выступить относительная величина затрат, предложенная в работе Н.Ш. Кремера [9].

В статье производится расчет среднего числа заявок в очереди  $L_{оч}$ , среднего времени нахождения заявки в очереди  $T_{оч}$ .

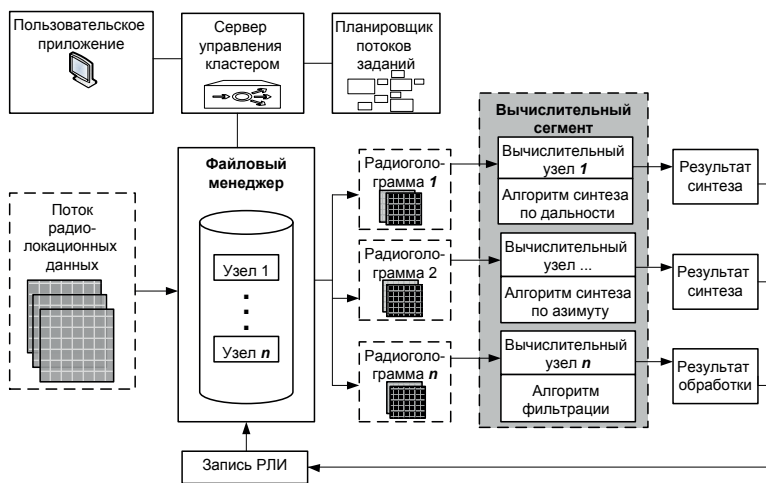
### *Масштабирование распределенной обработки радиолокационных данных*

Решение задачи масштабирования распределенной обработки РЛД в автоматизированной системе необходимо рассматривать на нескольких уровнях выполнения прикладных задач:

- высокопроизводительные серверы потоковой обработки РЛД;
- клиентское (конечное) оборудование обработки РЛД;
- сетевое и каналное оборудование, связывающее удаленные клиенты между собой.

В статье рассмотрены вопросы масштабирования серверных и клиентских решений обработки РЛД. В общем случае для проведения качественного масштабирования автоматизированной системы специализированного назначения необходимо: определить степень важности информации, оценить производительность всех средств обработки в географическом районе, произвести оценку пропускной способности каналов передачи данных и технического состояния средств обработки, а также провести формирование динамического задания на обработку РЛД по заявкам и запросам потребителей.

На примере структурной схемы серверного программного комплекса распределенной системы обработки (рис. 2), являющейся типовой для современных решений, рассмотрим его работу по обработке РЛД.



**Рис. 2.** Структурная схема серверного программного комплекса распределенной системы обработки РЛД

В зависимости от реализации конкретного алгоритма обработки радиолокационных космических снимков деление входных данных в распределенной серверной среде производится следующими способами:

- отдельные файлы РЛД, предназначенные для обработки, делятся на области, вычисление результата для каждой из них производится на отдельном узле, после чего полученные данные объединяются в готовое изображение;

Марковский А.С., Свеколкин Н.И., Шаров С.А. Решение задач обработки данных...

- отдельное изображение обрабатывается целиком на вычислительном узле без деления на области, параллельность исполнения достигается за счет обработки большого количества изображений на множестве узлов.

На вычислительных узлах серверного программного комплекса производится синтез РЛИ, а также выполняются операции фильтрации с целью повышения качества РЛИ.

Кроме масштабирования обработки на стороне сервера актуальным является вопрос масштабирования обработки РЛД с разнородных технических средств. Так, в отличие от задач обработки оптических изображений РЛД требуют мощных вычислительных ресурсов, которые не всегда доступны абонентскому оборудованию (сенсору наблюдения).

В качестве программно-аппаратной платформы для расчетов использовалась 64-рядная операционная система (ROSA Linux) на 8-ядерном процессоре Intel Core i7 с тактовой частотой 3,2 МГц.

Результаты практической обработки РЛД обычным способом и с использованием многопоточной реализации приведены в таблице 1.

С целью определения требуемых вычислительных ресурсов были проведены расчеты для различных реализаций алгоритмов по типовым операциям обработки РЛД – свертки РЛИ по дальности и азимуту.

Таблица 1

#### Оценка производительности алгоритмов обработки РЛД

Размеры исходного файла, пиксель	Наименование операции	Время реализации варианта обработки, с	
		Обычный	Многопоточный
20 × 23 552	Свертка РЛИ (дальность)	15,72	11,04
	Свертка РЛИ (азимут)	5,28	4,48
200 × 23 552	Свертка РЛИ (дальность)	140,05	98,03
	Свертка РЛИ (азимут)	76,35	64,89

Анализ оценок производительности алгоритмов синтеза РЛИ показывает наибольший прирост при многопоточной реализации для свертки по дальности.

Кроме указанных операций синтеза на производительность обработки РЛИ большое влияние оказывает выбор вычислительной платформы и способа реализации алгоритмов, что особенно актуально для сенсоров, установленных на авиационных носителях. При наличии стандартизированной вычислительной платформы задачи оптимизации кода являются для данных средств приоритетными (в том числе низкоуровневое исполнение на языках С и ассемблере, использование специализированных плат обработки).

В таблице 2 приведены результаты практической обработки РЛД для различных программных платформ.

Как видно из полученных результатов, программная платформа накладывает существенные ограничения на производительность обработки РЛД. Так, реализация алгоритма свертки по дальности на языке С дает 75%-й прирост по сравнению с интерпретируемым языком среды Matlab.

Таблица 2

## Оценка производительности программных платформ

Параметр	Свертка РЛИ	
	по дальности	по азимуту
Входные данные	+	+
Радиолокационное изображение	+	+
Размер исходного файла, Гб	2,19	12
Размер выходных данные, Гб	12	5
Время работы алгоритма (Matlab), с	3 919,25	74,49
Время работы алгоритма (C, многопоточное исполнение), с	548,67	40,08

На практике любой элемент системы (первоисточник заявки на обработку или ее непосредственный исполнитель) обладает ограниченными возможностями по пропускной способности каналов приема-передачи, вычислительным ресурсам, количеству одновременно обрабатываемых заявок, что в целом может отразиться на сроках решения задачи.

В таблице 3 представлены результаты моделирования параллельной обработки РЛД (при одинаковых начальных условиях) для вычислительного кластера из трех, пяти и семи узлов распределенной системы с неограниченной очередью.

Таблица 3

## Оценка модели распределенной системы обработки РЛД

Характеристика	Количество задействованных узлов			
	Три	Пять	Семь	
Вероятность простоя	0,02	0,06	0,06	
Число заявок в очереди, шт.	10,07	0,22	0,02	
Время ожидания в очереди, ч	7,30	0,16	0,01	
Относительная величина затрат, у. е.	24,07	4,10	5,10	
Длительность обработки данных объемом $V$ , с	$V = 200$ Мб	35,29	24,54	19,43
	$V = 2$ Гб	387,15	274,33	231,07

При выполнении численных расчетов было установлено, что система, составленная из пяти узлов, работает эффективнее по сравнению с системой из трех узлов, так как существенно снижается вероятность возникновения очереди, ее длина и среднее время пребывания в очереди.

## Заключение

Предложенный подход масштабирования ориентирован на оптимизацию обработки большого количества потоковых данных на сервере и клиентском оборудовании. Результаты анализа по распределению обработки показали, что различные факторы способны



Марковский А.С., Свеколкин Н.И., Шаров С.А. Решение задач обработки данных...

повлиять на совместное функционирование высокопроизводительных алгоритмов различных технологических этапов обработки РЛД в облачной инфраструктуре. К ним относятся: способ реализации потоковой обработки РЛД на серверном программном обеспечении, оптимизация программного кода, а также использование специализированных аппаратных средств для низкопроизводительных клиентских платформ [10; 11].

А решением по снижению возможности массового появления заявок на обработку данных несмежных участков земной поверхности является предоставление благоприятных условий на использование ранее полученных результатов интересующих территорий (их повторное использование).

### Литература

1. Mistry P, Braganza S, Kaeli D, Leeser M. Accelerating Phase Unwrapping and Affine Transformations for Optical Quadrature Microscopy Using CUDA // Proc. of 2nd Workshop on General Purpose Proc. on Graphics Proc. Units. Washington: ACM, 2009. P. 28–37.
2. Sheng G, Qi-Ming Z., Jian J, Cun-Ren L., Qing-xi T. Parallel Processing of InSAR Interferogram Filtering with CUDA Programming // Science of Surveying and Mapping Engineering. 2015. № 1. P. 54–68.
3. Потанов В.П., Попов С.Е., Костылев М.А. Метод обработки радарных данных на базе системы массово-параллельного исполнения заданий Apache Spark // Вычислительные технологии. 2017. Т. 22, спец. вып. 1. 2017. С. 60–74.
4. Zinno I, Mossucca L., Elefante S., De Luca C., Casola V, Terzo O., Casu F, Lanari R. Cloud Computing for Earth Surface Deformation Analysis Via Spaceborne Radar Imaging: A Case Study // IEEE Trans. Cloud Computing. 2016. № 4. P. 104–118.
5. Apache Spark Unified Analytics Engine for Large-Scale Data Processing. URL: <http://spark.apache.org> (date of the application: 11.05.2018).
6. Русинов Ю.А., Титов К.И., Шаров С.А. Использование алгоритмов многопоточной реализации для решения задач синтеза радиолокационного изображения // Радиолокационное исследование природных сред: материалы XXIX Всероссийского симпозиума (Санкт-Петербург, 25–26 марта 2015 г.). СПб., 2015.
7. Чистяков В.П. Курс теории вероятностей. М.: Наука, 1987. 240 с.
8. Кауштанов В.А., Ивченко Г.И., Коваленко И.Н. Теория массового обслуживания. М.: Либроком, 2012. 306 с.
9. Кремер Н.Ш., Путко Б.А. Эконометрика. М.: ЮНИТИ-ДАНА, 2010. 328 с.
10. Нечай А.А., Борисов А.А., Борисова Ю.И. Точечный анализ данных дистанционного зондирования земли средствами языка программирования Python // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2019. № 1. С. 49–55.
11. Шаймарданов А.М., Нечай А.А., Лепехин С.В. Математические модели систем автоматического управления с широтно-импульсной модуляцией // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2019. № 2. С. 27–39.

### Literatura

1. Mistry P, Braganza S, Kaeli D, Leeser M. Accelerating Phase Unwrapping and Affine Transformations for Optical Quadrature Microscopy Using CUDA // Proc. of 2nd Workshop on General Purpose Proc. on Graphics Proc. Units. Washington: ACM, 2009. P. 28–37.



2. Sheng G., Qi-Ming Z., Jian J., Cun-Ren L., Qing-xi T. Parallel Processing of InSAR Interferogram Filtering with CUDA Programming // Science of Surveying and Mapping Engineering. 2015. № 1. P. 54–68.
3. Potapov V.P., Popov S.E., Kostylev M.A. Metod obrabotki radarnykh dannykh na baze sistemy massovo-parallel'nogo ispolneniya zadaniy Apache Spark // Vychislitel'nye tekhnologii. 2017. T. 22, spets. vyp. 1. 2017. S. 60–74.
4. Zinno I., Mossucca L., Elefante S., De Luca C., Casola V., Terzo O., Casu F., Lanari R. Cloud Computing for Earth Surface Deformation Analysis Via Spaceborne Radar Imaging: A Case Study // IEEE Trans. Cloud Computing. 2016. № 4. P. 104–118.
5. Apache Spark Unified Analytics Engine for Large-Scale Data Processing. URL: <http://spark.apache.org> (date of the application: 11.05.2018).
6. Rusinov Yu.L., Titov K.I., Sharov S.A. Ispol'zovanie algoritmov mnogopotочноj realizatsii dlya resheniya zadach sinteza radiolokatsionnogo izobrazheniya // Radiolokatsionnoe issledovanie prirodnykh sred: materialy XXIX Vserossijskogo simpoziuma (Sankt-Peterburg, 25–26 marta 2015 g.). SPb., 2015.
7. Chistyakov V.P. Kurs teorii veroyatnostej. M.: Nauka, 1987. 240 s.
8. Kashtanov V.A., Ivchenko G.I., Kovalenko I.N. Teoriya massovogo obsluzhivaniya. M.: Librokom, 2012. 306 s.
9. Kremer N.Sh., Putko B.A. Ekonometrika. M.: YuNITI-DANA, 2010. 328 s.
10. Nechaj A.A., Borisov A.A., Borisova Yu.I. Tochechnyj analiz dannykh distantsionnogo zondirovaniya zemli sredstvami yazyka programmirovaniya Python // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2019. № 1. S. 49–55.
11. Shajmardanov A.M., Nechaj A.A., Lepekhin S.V. Matematicheskie modeli sistem avtomaticheskogo upravleniya s shirotno-impul'snoj modulyatsiej // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2019. № 2. S. 27–39.

DOI: 10.25586/RNUV9187.19.03.P.032

УДК 004.94

**В.А. Минаев, К.М. Бондарь, Е.В. Вайц, И.А. Беляков**

---

## ДИСКРЕТНО-СОБЫТИЙНОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССОВ МОНИТОРИНГА И УПРАВЛЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТЬЮ

---

Применен метод дискретно-событийного моделирования для описания процессов мониторинга и управления информационной безопасностью. В ходе имитационных экспериментов показано, что именно эти модели позволяют наиболее адекватно, в режиме реального времени представить процесс реагирования на компьютерные атаки, рассчитать резко меняющуюся при их реализации нагрузку на информационную систему, персонал, обеспечивающий ее защиту, наглядно отобразить функционирование подсистемы мониторинга и управления информационной безопасностью. Дискретно-событийная модель построена в среде AnyLogic и отражает сообщения об атаках из двух источников информации (внутренних и внешних), приходящие от рабочих станций, сетевых устройств, веб-ресурсов и средств защиты информации. Дается описание SIEM (Центр информационной безопасности и управления событиями), где последовательно производится сбор сообщений, их фильтрация, агрегация и корреляция. SIEM дает возможность распределить сообщения по степени риска: высокий, средний, низкий. Далее моделируется SOC (Центр операций по обеспечению безопасности) с тремя линиями обслуживания, отличающимися уровнями