|  |
| --- |
| **ТЕХНИЧЕСКИ УНИВЕРСИТЕТ - СОФИЯ**  **Факултет – „Компютърни системи и управление“** |

**ДИПЛОМНА РАБОТА**

**„Наблюдение дейността на студентите по време на изпити провеждани в електронен вариант“**

**Студент: Иван Николаев Колев, фак. № 121213114, гр. 43**

**Специалност: „Компютърно и софтуерно инженерство“**

**Дата: Научен ръководител:**

**/ маг. инж. Д.Андреев/**

**София  
2017**

**СЪДЪРЖАНИЕ**

[Въведение 3](#_Toc493533622)

[1. Постановка на дипломната работа. Цели и задачи. 4](#_Toc493533623)

[1.1. Поток от данни. 4](#_Toc493533624)

[1.2. Вид на събираните данни. 5](#_Toc493533625)

[1.3. Структура на събираните данни. 5](#_Toc493533626)

[1.4. Ресурсоемкост на приложението за събиране на данни. 5](#_Toc493533627)

[1.5. Надеждност и сигурност. 6](#_Toc493533628)

[1.5.1. Балансиране товара на сървърите (Load balancing) 6](#_Toc493533629)

[1.5.2. Менажиране на потребителските сесии 6](#_Toc493533630)

[1.5.3. Защита на вътрешните системи 7](#_Toc493533631)

[1.5.4. JSON Web Token 7](#_Toc493533632)

[1.6. Скалируемост на софтуерната система. 9](#_Toc493533633)

[1.7. Ресурсоемкост на софтуерната система за обработка и съхранение на данните. 9](#_Toc493533634)

[2. Функционално описание на софтуерната система. 10](#_Toc493533635)

[2.1. Програмни езици и библиотеки. 10](#_Toc493533636)

[2.2. Бази за съхранение на данните (Database servers). 11](#_Toc493533637)

[2.3. Поточна линия за пренос на съобщения между системите (Message bus) 12](#_Toc493533638)

[2.3.1. Настройване производителността на Apache Kafka 13](#_Toc493533639)

[2.3.2. Замерване производителността на Apache Kafka 14](#_Toc493533640)

[2.4. Наблюдение на системите 14](#_Toc493533641)

[2.4.1. Наблюдение на логовете 15](#_Toc493533642)

[2.4.2. Наблюдение на различни метрики от системите. 16](#_Toc493533643)

[3. Програмна реализация 17](#_Toc493533644)

[4. Указания за ползване 17](#_Toc493533645)

[5. Експериментални данни 17](#_Toc493533646)

# Въведение

Една от основните цели на преподавателите в училища и университети е да осигуряват подходящи и равни условия за провеждане на изпити, контролни и упражнения. До голяма степен това се изразява в следене на студентите да не преписват от помощни материали, да работят самостоятелно по възложените задачи и да не се отклоняват от целите на упражнението. Това представлява трудна задача, особено при зали с капацитет повече от пет души, защото фокусът на преподавателя не може да бъде върху повече от един човек едновременно и винаги съществува възможността да изпусне моменти, които няма как да върне и види отново. Именно тези моменти студентите ползват, за да използват забранени материали, да препишат или да отворят сайта на любимата си социална мрежа в която да прекарат времето си.

Изследвания проведени по темата за преписване показват стряскащи числа: процентът на студенти/ученици, които са преписвали поне веднъж започва от 17% и стига до 98% в различни части на света. Част от причините за това са, че за да останеш „конкурентоспособен“ си „длъжен“ да преписваш, защото други, по-слаби ученици, го правят и на резултати изглеждат по-способни от теб.

Компютрите и технологиите като цяло стават все по-достъпни и повече училища и университети започват да разчитат на тях за провеждане на упражнения, изпити и контролни. От една страна, работата по следене дейността на цяла група студенти е непосилна за сам преподавател, но от друга страна, е сравнително лесно да се използват наличните технологии за събиране и анализ на данни свързани с дейностите, които дадено лице извърша на компютъра си.

Основният фокус на текущата дипломна работа е наблюдението дейността на студентите по време на изпити, контролни и упражнения, провеждани в електронен вариант с достъп до интернет, по време на които, изпитваното лице трябва да спазва набор от правила и не трябва да използва забранени канали за комуникация, да търси или публикува въпроси във форуми и други вид информационни сайтове, да използва торент тракери и т.н.

Чрез използване на програмно приложение за Windows OS, студентите ще могат да се „регистрират“ в системата и стартират нова „работна сесия“. По време на работната сесия, приложението ще събира информация за процесите стартирани в операционната система, за текущият фокусиран процес, снимки на дисплея, снимки от уеб камерата, натиснати клавиши, изпратени и получени мрежови пакети.

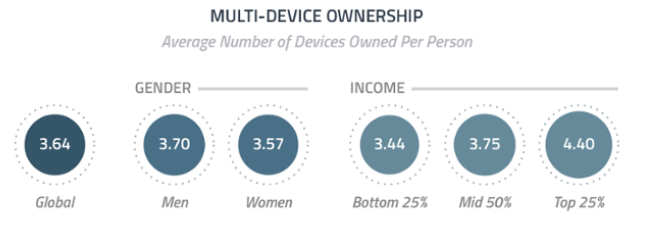
Събраната информация ще може да бъде разгледана и анализирана в реално време или в последващ период.

С предложената имплементация се очаква като резултат, по-ефективно менажиране на учебните и изпитни процеси от страна на преподавателите, по-голяма прозрачност на действията и по-висока ангажираност на студентите.

# 1. Постановка на дипломната работа. Цели и задачи.

## 1.1. Поток от данни.

За последните 20 години, непрестанният технологичен възход промени значително начина по който хората живеят, работят, комуникират и почиват. Масово притежаваме между 3 и 4 мобилни устройства, които ползваме активно през по-голямата част от денонощието. [3] Мобилните устройства могат да бъдат лаптопи, таблети, умни телефони, умни часовници, умни гривни, умни очила, електронни четци за книги, музикални устройства, камери, фотоапарати и т.н. Всяко едно от тези устройства в днешно време има опция за свързване към интернет, към умен телефон или към лаптоп. Това означава постоянен обмен на данни, данни които неусетно могат да станат публични и да бъдат използвани за изграждане на дигитален отпечатък на човека притежаващ устройствата.



***Фигура 1.*** *Статистика за средния брой притежавани мобилни устройства, глобално.*

Потокът от данни включва, но не се ограничава до: уеб сайтовете, които посещаваме и времето прекарано в тях; скоростта, с която натискаме последователно клавиши на клавиатурата; софтуерните приложения, които инсталираме и използваме; имейлите, които изпращаме; хората, с които сме приятели и познаваме; стилът музика, който предпочитаме и любимите ни песни; местата, които посещаваме и пътищата, по които минаваме (благодарение на вградени GPS системи и триангулация от страна на мобилните оператори към които сме абонирани) и т.н.

Всеки ден, използвайки преносими или стационарни устройства, средностатистическият човек може да генерира стотици гигабайти информация. Голяма част от тази информация вече се използва от маркетинг компании с цел продуктово позициониране, показване на подходящи реклами и промоции, целящи да подтикнат човека да закупи вещи, които са в сферата на неговите интереси.

Този голям обем от данни може да бъде използван, както за добронамерени, така и за злонамерени цели.

Затова е важно всеки потребител да използва технологиите с предпазливост и с ясното съзнание, че всеки момент може с действията си да стане уязвим.

## 1.2. Вид на събираните данни.

За целта на приложението, ще се събират данни свързани с начина, по който потребителят използва компютъра, а именно:

* Стартирани процеси/приложения (active processes)
* Текущите процеси на фокус (foreground processes)
* Изпратени и получени мрежови пакети (network traffic)
* Моментни снимки на дисплея, който потребителят вижда (display snapshots).
* Моментни снимки през уебкамерата, ако такава съществува (camera snapshots).
* Натиснати клавиши и скоростта на писане (captured keyboard keys).

## 1.3. Структура на събираните данни.

За целта на системата, данните ще следват един и същ шаблон. Всеки прихванат обект, ще принадлежи на определена „сесия“. Сесия, в контекста на системата, означава всяко ново стартиране на процес за наблюдение.

Всеки прихванат обект, ще се състои от следните 5 полета:

* Име на потребителя
* Идентификационен номер на потребителя
* Идентификационен номер на текущата сесия
* Дата на стартиране на сесията
* Обекта, който е прихванат от приложението (натиснат клавиш, стартиран процес, фокусиране на нов процес, снимка на дисплея, снимка на уебкамерата или мрежови пакет)

## 1.4. Ресурсоемкост на приложението за събиране на данни.

С оглед естеството на данните, които ще се събират, основната част от системата за наблюдение, а именно десктоп приложението, трябва да бъде конфигуруемо.

Следните параметри са от основно значение за ресурсоемкостта му:

* Период, през който ще се правят снимки на дисплея.
* Период, през който ще се правят снимки на уеб камерата.
* Период, през който ще се изпращат агрегираните данни на отдалечения сървър за съхранение.

Информация за активните процеси ще се прихваща, само когато нов процес бива стартиран.

Информация за процеса на фокус ще се прихваща, само когато се фокусира процес, различен от текущия.

Информация за натиснатите клавиши, ще се прихваща при всяко натискане на клавиш.

От прихванатите мрежови пакети ще се записват само първите 24 байта информация, за да се избегне претоварване на мрежата или консумиране на прекалено много RAM памет за временното съхранение на пакетите.

С оглед предоставената информация, само първите три изброени параметъра, ще оказват голямо влияние на разхода на ресурси от приложението, затова те трябва да бъдат избрани внимателно, спрямо машината на която ще работи продукта.

Конфигурациите ще бъдат четени от конфигурационен JSON файл, и ще могат да бъдат променяни спрямо нуждите на различните сесии/потребители/машини.

## 1.5. Надеждност и сигурност.

Първото нещо, за което трябва да мислим, когато работим с потребителска информация е „Законът за защита на потребителските данни“. Всичко което е чувствителна информация, или може да навреди по един или друг начин, ако попадне в злонамерени ръце, трябва да бъде съхранявано по сигурен начин.

За целта, всяко едно предаване на потребителски данни, от компютърът домакин, до сървърите, които ще обработват тази информация, трябва да се случва като информацията се криптира, изпраща се на сървъра приемник, а той я декриптира и използва в оригиналният й вид.

За жалост, все още не съществува 100% сигурен начин за предаване на информация между 2 отдалечени точки и методите, които ще използваме, могат единствено да затруднят по-опитните специалисти.

Второто нещо, което искаме да подсигурим е интегритета на данните и информацията, която постъпва в нашата система.

Понеже ще градим решение, което може да скалира достатъчно добре, за голям обем от потребители, първата линия, която обработва всички потребителски заявки е ферма от front-end уеб сървъри, чиято единствена цел е да получат данните, да идентифицират потребителя, който ги е изпратил, и след това да ги запишат в съответното хранилище, където да бъдат поети за обработка по поточната „тръба“.

Ако при ферма от сървъри, за менажиране на потребителските сесии разчитаме на това състояние да бъде пазено на сървъра, ще си създадем потенциални проблеми и неудобства за самите потребители.

### 1.5.1. Балансиране товара на сървърите (Load balancing)

Има няколко техники за рутиране на потребителските заявки при наличие на ферма от сървъри:

* “Sticky session” е стратегия, при която пред фермата стои load balancer, който прихваща потребителска заявка и винаги за дадения потребител, рутира заявката към един и същ уеб сървър. Създава се таблица с връзки „user” – “server”. По този начин се гарантира, че потребителят няма да има проблеми със сесиите си, както ако бъде рутиран към друг сървър. Тази стратегия не е много надеждна, защото при непланирано спиране на съответния сървър, всички потребителски сесии ще бъдат прекъснати, а load-balancer-ът ще пренасочи потребителите към нов активен сървър.
* “Round robin” е стратегия, при който пред фермата отново стои load balancer, който избира следващият сървър, който да обработи заявката, на ротационен принцип. Това е най-семплият начин за балансиране на натоварването между уеб сървърите, но отново води до проблеми при менажиране на потребителските сесии, ако те се съхраняват на ниво сървър.

### 1.5.2. Менажиране на потребителските сесии

Фазата на идентификация е деликатен момент при наличието на ферма от дистрибутирани уеб сървъри. За решаване на този проблем ще използваме RESTful архитектура, чиято основна цел е да не пазим потребителско състояние/сесия на сървъра, а цялата необходима информация да се предава всеки път, криптирана в бисквитка на браузъра (browser cookie) или като част от тялото на заявката, която ще се декриптира с машинен ключ на ниво уеб сървър и ще се проверява за автентичност.

След като е решен проблема с идентификацията на потребителите, трябва да подсигурим, че всички „входни“ точки в нашите уеб сървъри изискват съответното ниво на автентикациям преди да допуснат каквито и да било данни, напред по поточната линия.

Входните точки още се наричат “endpoint”-и [mark] на уеб сървъра. Това са адресите, към които изпращаме заявки за извличане или създаване на информация.

### 1.5.3. Защита на вътрешните системи

Друга съществена част при защитаване на софтуерната система е подсигуряване комуникацията между отделните модули и валидиране на автентичността на всяка заявка вътрешно в системата.

За решаване на този проблем стандартно се ползват 2 стратегии:

* Конфигуриране на защитна стена (firewall), която да допуска заявки само от определени IP адреси.
* Изпращане на идентификационен токен, който да съдържа идентичността на системата, която прави заявката.

За идентификацията между системите ще използваме много семпла и надеждна концепция, а именно “JWT” (JSON Web Token) [mark].

### 1.5.4. JSON Web Token

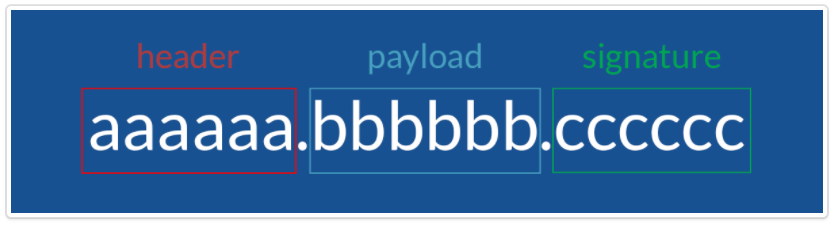
Накратко JWT представлява отворен стандарт (**RFC 7519**), който дефинира компактен начин за сигурно предаване на информация (payload) между различни страни, под формата на JSON обект. Предадената информация може да бъде проверена посредством дигитален подпис. JWT токените могат да бъдат подписани използвайки „тайна” ключова дума (за **HMAC** алгоритъм) или двойка публичен/частен ключ (за **RSA** алгоритъм)

**JWT е подходящ за следните случаи**:

* **Автентикация** – това е най-честия случай за използване на JWT. Веднъж щом потребителят се идентифицира в системата, всяка последваща заявка ще включва токена, позволявайки на потребителят да достъпва пътища, услуги и ресурси, които изискват съответното ниво за достъп. Single Sign On функционалността широко използва JWT в днешно време, заради малкия му размер, бързодействието което осигурява и възможността да се споделя лесно между различни домейни.
* **Предаване на съобщения** – JSON Web Token-ите са добър начин за сигурен пренос на информация между две или повече страни. Поради факта, че тези токени са подписани – можем да сме почти сигурни, че страната изпращаща информацията е тази за която се представя. В допълнение, подписът се генерира използвайки хедърите и самото съобщение, съответно можем да валидираме, че съдържанието не е фалшифицирано.

За надеждна и сигурна работа използвайки този тип токени, ключът, който се използва за подписването им трябва да бъде ръчно сложен в конфигурационния файл на системата, на скрито, непублично място. Добра практика е през определен период от време да се подменя ключа за подписване.

#### 1.5.4.1 Примерна структура на JWT токен:



***Фигура 2.*** *Структура на base64-encoded JWT.*



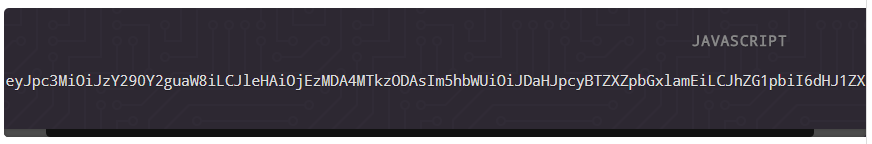
***Фигура 3****. Примерен Header на JWT.*



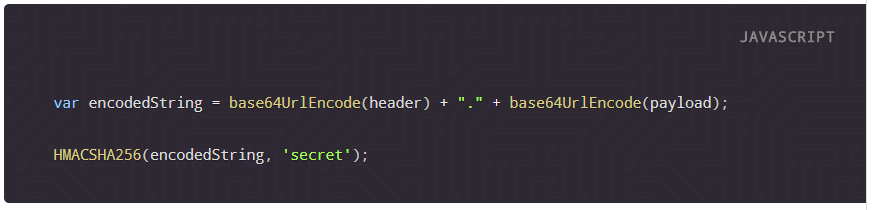
***Фигура 4.*** *Примерен резултат от кодиране на Header в base64 формат.*



***Фигура 5.*** *Примерен Payload на JWT.*



***Фигура 6.*** *Примерен резултат от кодиране на Payload в base64 формат.*



***Фигура 7.*** *Примерно подписване (signature) на JWT посредством JavaScript.*



***Фигура 8.*** *Примерен Signature на JWT.*

## 1.6. Скалируемост на софтуерната система.

Тъй като трябва да създадем надеждна система, която да продължи да работи дори при внезапно спиране на някой от съставните й модули, товa означава, че трябва да мислим за разпределение и мултипликиране на отделните сървъри.

Не искаме всичките ни ресурси да бъдат на едно централизирано място, защото това означава, че при най-малък проблем, като спиране на захранването или дефект в компютъра домакин, цялата ни система ще преустанови работата си.

За целта, трябва да изберем подходящи технологии и стратегии за изграждане на системата по начин, който ще я направи независима от местоположение, устойчива и скалируема.

Под скалируемост се има предвид системата да може да се разширява хоризонтално (вдигане на инстанции върху нови физически или виртуални машини), за да може да поеме по-голям работен товар, когато това е наложително. Например, ако имаме университет с 30 000 студенти, които всеки ден провеждат поне 1 упражнение включващо употребата на компютри, това означава, че нашата система ще трябва да може да поеме голям обем входни данни, които съответно трябва да обработи в реално време, без да има осезаемо забавяне или срив.

## 1.7. Ресурсоемкост на софтуерната система за обработка и съхранение на данните.

Със скалирането на сървърите, разходите за ползване на хардуерни ресурси могат да скочат драстично, ако избора на технологии и дизайна на системата не са добре обмислени.

За да ограничим разходите до разумен минимум, ще целим да използваме основно Linux-базирани операционни системи и системи с отворен код, които са безплатни за ползване с комерсиални цели, и които се характеризират с ниска цена за притежание (low-cost of ownership).

Набора от системи ще включва:

* Apache Zookeeper (Linux)
* Apache Kafka (Linux)
* ElasticSearch (Linux)
* Kibana (Linux)
* Grafana (Linux)
* Logstash (Linux)
* MongoDB (Linux)
* ASP.NET Web Api (Windows, но може да бъде пренесен на Linux операционна система, ако приложението е разработено върху .NET Core платформата)
* SQL Server (Windows)

Аргументация за избора на технологичен стек ще бъде представена в следващата глава.

## 2. Функционално описание на софтуерната система.

## 2.1. Програмни езици и библиотеки.

Основният програмен език който ще използваме е **C#**, той е създаден да се изпълнява от .NET Framework-а [mark]. Езикът е разработен от Microsoft, а общността, която го използва е много широка. Съществуват огромен брой библиотеки, примери с код и решения на често срещани проблеми, които могат лесно да бъдат използвани в други проекти.

Езикът е обектно-ориентиран със силна типизация и мощна поддръжка на Generic класове и функции, което позволява гъвкавост, надеждност и изграждане на добра абстракция. Друг плюс е, че зад него стои корпорация, чиито продукти са базирани на него, което автоматично означава, че езикът ще продължи да се поддържа и развива, докато Microsoft съществува.

Друг плюс е, че с пускането на **.NET Core**, C# става **мултиплатформен** и може да работи под Линукс-базирани операционни системи, като например **Ubuntu, Fedora, Red Hat Enterprise, Debian, CentOS, SUSE/openSUSE**. Това на свой ред води до факта, че системи разработени на C#, ще могат да бъдат инсталирани на безплатен Linux сървър, вместо на платен Windows сървър, а това автоматично редуцира TCO[mark].

От 2 години насам, езикът и целият фреймуърк под него стават с отворен код и всеки който иска може да допринесе за развитието му или за поправянето на открити проблеми като направи заявка в официалното хранилище [mark] където е публикуван кодът.

Технологичният стек, който ще използваме за потребителския интерфейс включва HTML, CSS и JavaScript, като за манипулации различните елементи (бутони, текстови полета, изображения и т.н.), ще ползваме библиотека, която е изградена на база JavaScript езика, а именно JQuery. Тя ни предоставя програмен интерфейс, през който лесно можем да обработваме, трием и създаваме елементи от потребителския интерфейс.

Посредством HTML ще изграждаме само и единствено структурата на нашия документ, а използвайки CSS – ще стилизираме елементите и техния външен вид.

За комуникация със всички останали модули и системи, ще използваме драйвери изградени специално за езика C#.

* Apache Kafka – Official Confluent Kafka .NET Driver [mark]
* MongoDB – Official MongoDB .NET Driver [mark]
* Microsoft SQL Server – Entity Framework

## 2.2. Бази за съхранение на данните (Database servers).

Изключително важен е избора на database server за съхранение на различните видове данни и операции, които ще съхранява и извършва софтуерната система.

В случая, избираме стратегия отново с ясната идея, че трябва да можем да скалираме хоризонтално (добавяне на повече машини и инстанции на дадена услуга), а не вертикално (добавяне на повече хардуерни ресурси в една единствена машина).

Стратегията която ще използваме е следната:

* За всички важни потребителски данни, които трябва да бъдат съхранявани по сигурен и консистентен начин – използваме SQL Server. Причината за това е, че по подразбиране, SQL Server-ите носят със себе си множество функционалности свързани със сигурността на достъп и интегритета на данните, като например User Authentication[mark]/Windows Authentication, различни роли за контролиране нивата на достъп до съответни функционалности, криптиране на данните, транзакции и т.н.
* За всички функционалности свързани с търсене на логове и метрики от различните системи – използваме ElasticSearch engine, който представлява индекс силно оптимизиран за търсене.
* За всички останали оперативни заявки, свързани със записване, четене и обработка на данните свързани с дейността на студентите – използваме MongoDB, заради гъвкавостта и скоростта която предоставя.

За оперативните данни се спряхме на NoSQL база, заради предимствата, които има в случая пред релационните SQL бази.

Релационните бази се сблъскват със следните предизвикателства:

* Не са удачни при работа с големи обеми от данни и широк набор от типове на данните (изображения, видеа, текст и т.н.).
* Не могат да скалират вертикално, защото една машина е лимитирана от обема хардуерни ресурси, които може да управлява ефективно (предимно RAM и CPU).
* Не могат да скалират хоризонтално, защото са ограничени от кеш-зависими операции за четене и писане.
* Не могат да бъдат раздробявани и разпръсквани на различни машини (nodes).
* Сложен релационен модел, който прави трудна промяната на схемите (добавяне на нови колони, промяна или махане на стари и т.н.)

Нерелационните бази (NoSQL) атакуват проблемите, които релационните бази срещат по следния начин:

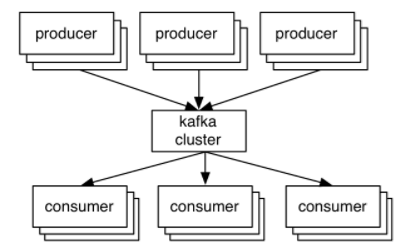
* Създадени с мисъл за хоризонтално скалиране и нулева зависимост между отделните колекции, документи и т.н. Възможност да работят на огромен брой разпределени възли (nodes).
* Незаключващ документите контрол при конкурентни операции, за да може операции свързани с четене да не бъдат блокирани от операции свързани с писане/промяна на документа.
* Скалиране на репликациите и висока дистрибутивност (могат да достигнат хиляди машини)
* Архитектура, която предоставя по-висока производителност на машина отколкото при релационните бази.
* Модел на данните, който не се контролира от схема/шаблон, което позволява голяма гъвкавост при работа с различни данни.

В повечето случаи производителността на NoSQL базите е ограничена от скоростта на мрежата (network-bound) по която минава заявката. При добра инфраструктура, това означава между 2 и 10 милисекунди на заявка.

Като обобщение – изборът на база трябва да зависи изцяло от конкретния случай на ползване и да бъде добре обмислен.

## 2.3. Поточна линия за пренос на съобщения между системите (Message bus)

За предаване на съобщения между отделните системи, ще използваме Apache Kafka, която е известна със способността си за приемане на стотици хиляди съобщения в секунда, което перфектно пасва на нуждите ни от буфер, който да може своевременно да съхранява всички данни, които се събират от студентите.

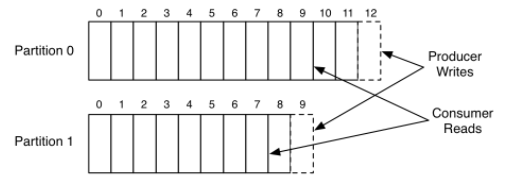


***Фигура 9.*** *Основна схема на работа на Apache Kafka*

Кафка имплементира publish – subscribe модела. Той представлява начин да се абонираме към дадена тема (topic) и в момента, в който някой публикува нещо свързано с тази тема, автоматично да получим известие/съобщение.

Процеса на работа с Кафка е сравнително прост: Производител (Producer[mark]) изпраща съобщение относно дадена тема (Topic[mark]) до Кафка, която на свой ред записва съобщението на твърд диск и след това го препраща към всички абонирани за тази тема потребители (Consumers[mark]).

Ключовата абстракция при Кафка е темата. Тя представлява разделен регистрационен дневник (sharded write-ahead log). Производителите добавят записи към тези днвеници, а потребителите се абонират за промени. Всеки запис е свързана двойка ключ/стойност. Ключът се използва за записване на стойността(съобщението) в съответстващ регистрационен дял (log partition).



Фигура 10. Пример за производител и консуматор, които четат и пишат в двудялова тема.

Всеки запис се асоциира с входящ номер, който наричаме „отстояние“ (offset). Това отстояние се използва от консуматорите и указва позицията, до която са стигнали във всеки един от логовете, които четат.

Регистрационните дялове, могат да бъдат разпръснати по различни машини в рамките на един Кафка клъстер (съвкупност от кафка брокери), което позволява една тема да съхранява повече данни, отколкото могат да се поберат на единична машина.

Важно е да се отбележи, че за разлика от други системи за предаване на съобщения, Кафка логовете винаги са персистентни. Съобщенията се записват на файловата система в момента на получаването им. Съобщенията не се изтриват, когато биват прочетени, а се запазват за определен срок от време (който може да бъде конфигуриран спрямо нуждите на системата, например 2 седмици). Това покрива случаи, когато бихме искали повторно да проиграем(обработим) последователност от съобщения.

Също така, Кафка позволява репликиране на логовете между множество различни сървъри с цел устойчивост на грешки. Например ако 1 от Кафка брокерите спре, консуматорите ще могат да бъдат прикачени към друг жив брокер и ще продължат нормален работен процес.

## 2.3.1. Настройване производителността на Apache Kafka

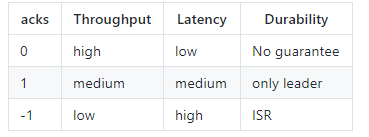
Кафка идва с конфигурация по подразбиране, която работи изключително бързо в повечето от случаите и не се налага допълнително конфигуриране на отделните модули, но е важно да отбележим, че има няколко елемента от значителна важност за бързодействието:

* Размер на твърдия диск (по-голям размер означава по-голям период на задържане на съобщенията).
* Скорост на писане и четене от твърдия диск (по-висока скорост означава по-голямо бързодействие на Производителите).
* Капацитет на мрежата (по-голям капацитет означава по-висока пропускливост на съобщения).
* Размер на енергозависимата памет (повече RAM означава по-висока производителност на Консуматорите)

Друг важен фактор е дали записваните съобщения ще бъдат компресирани. Кафка предоставя няколко различни алгоритми за компресия, като всеки от тях си има плюсове и минуси в зависимост от структурата и размера на съобщенията.

Наличието на компресия ще редуцира значително размера на използваното дисково пространство, но и ще намали малко пропускливостта на клъстера.

Последния фактор, който значително влияе на бързодействието на Кафка производителя е така наречената „надеждност“ (durability) при предаване на съобщенията. Нивото на надеждност се настройва от конфигурация която се казва „request.required.acks”, като спрямо зададеното ниво се променят параметрите: пропускливост (throughput); забавяне (latency); надеждност (durability).



***Фигура 11.*** *Таблица на зависимостите при различни стойности на „request.required.acks”.*

### 2.3.2. Замерване производителността на Apache Kafka

Параметри на физическата машина, на която са проведени тестовете[mark]:

* процесор Intel Xeon 2.5 GHz с 6 ядра.
* шест твърди диска 7200 RPM SATA.
* 32 GB RAM памет
* 1Gb мрежова карта

Размер на записваното съобщение – 100 bytes.

При един единствен Кафка производител, записващ в една тема с 6 дяла и без репликация на съобщенията, резултатът е: **821,557 записа в секунда** (78.3 MB/s).

При един единствен Кафка производител, 3x асинхронна репликация на съобщенията, резултатът е: **786,980 записа в секунда** (75.1 MB/s). Асинхронна репликация означава, че брокерът, който получава съобщението, връща веднага потвърждение, без да чака резултат от останалите 2 брокера, към които е пратил заявката за репликация.

При един единствен Кафка производител, 3x синхронна репликация на съобщенията, резултатът е: **421,823 записа в секунда** (40.2 MB/s). Синхронна репликация означава, че брокерът, който получава съобщението, връща потвърждение, чак след като е получил потвърждение от останалите 2 брокера, към които е пратил заявката за репликация.

При един единствен Кафка консуматор, абониран за една тема, резултатите са: **940,521 прочетени записа в секунда** (89.7 MB/s).

При три Кафка консуматора, абонирани за една и съща тема, резултатите са: **2,615,968 прочетени записа в секунда** (249.5 MB/s).

При един Кафка консуматор и един Кафка производител, които работят едновременно, резултатите са: **795,064 записа в секунда** (75.8 MB/s). Забавянето от момента на записване на съобщението до момента на прочитане на съобщението е: **2ms** за **98%** от записите; **3ms** за **1.9%** от записите; **14ms** за **0.1%** от записите.

От тук нататък няма причина да продължаваме с тестовете, защото тези метрики покриват нуждите ни предостатъчно.

## 2.4. Наблюдение на системите

При система, състояща се от множество отделни модули, които трябва да работят в синхрон е от огромно значение да имаме начин за наблюдение на процесите, които се случват в реално време. Това означава да знаем кога нова заявка е достигнала до сървъра ни, кога е била обработена, какъв е резултата от обработката на тази заявка, от кой потребител е дошла, за колко време е минала през поточната линия и т.н.

Също така е полезно да имаме достъп до всяка една грешка или изключение, което е възникнало по време на работата на нашата система, например: връзка между уеб сървър и база от данни е прекъсната по време на трансфер на данни или просто не може да бъде установена; изпратена заявка до друг уеб сървър е неуспешна и е върнат код за грешка; входни данни не са в очаквания формат и не могат да бъдат трансформирани и т.н.

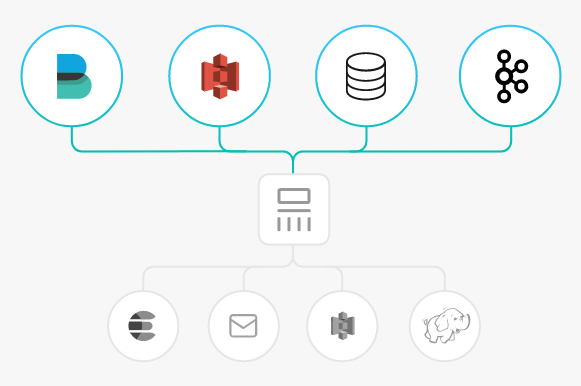
Възможността за записване и проследяване на всяко едно подобно събитие е ключът към ефективната поддръжка на една софтуерна система.

### 2.4.1. Наблюдение на логовете

За наблюдение и на логовете и лесно търсене измежду милионите записи, ще използваме стандартен ELK стек (ElasticSearch, Logstash & Kibana). Комбинацията от тези три системи ни дава невероятноа гъвкавост и лекота при работата с логове.

**Logstash** [mark] представлява поточна линия с вградени функционалности за четене, обработка и записване на входни данни от и към множество източници. Системата е реализирана на Ruby и използва Ruby Gems за разширяване на съществуващите функционалности. Общността, която подпомага развитието на Logstash е разработила десетки плъгини, които можем да ползваме наготово за нашите цели.

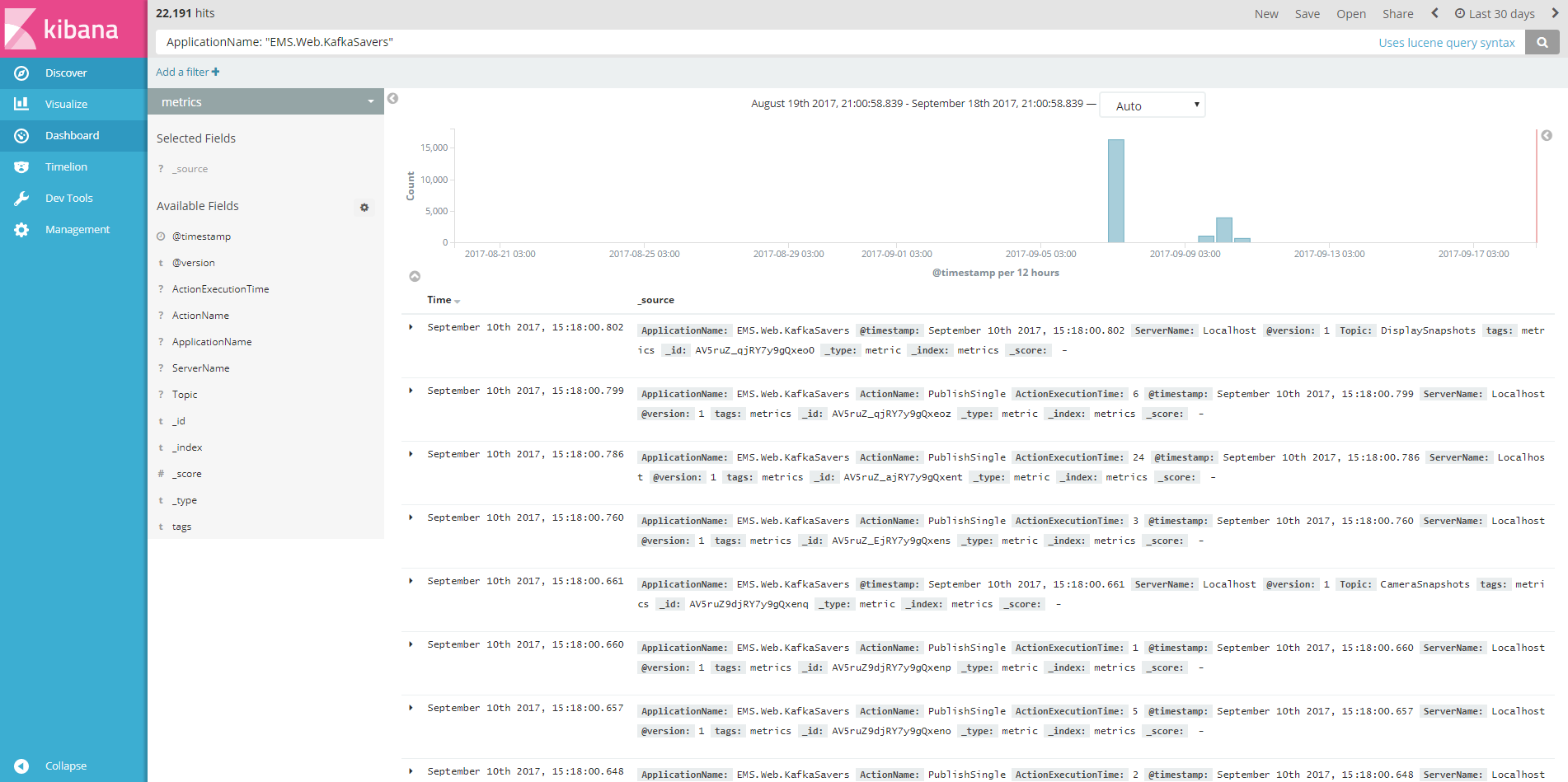
Също така, авторите са отворили кода публично в Github [mark] и той е достъпен за преглед и модификация от всеки, който има интерес.



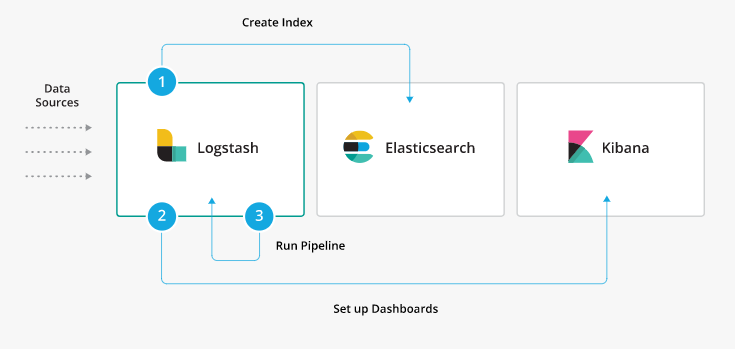
***Фигура 12.*** *Архитектура на Logstash – Схема за консумиране на данни от множество входни точки, обработване и записване в множество изходни точки.*

**ElasticSearch** [mark] ще бъде основната ни база за търсене на логове и метрики. Той представлява документ-базирана система за търсене, която е изградена върху Lucene engine[mark]. Към ElasticSearch можем да изпращаме заявки за търсене по име на поле, стойност на което и да е поле от документа, стойност на точно определено поле, и т.н. Също така е създаден с идеята за скалируемост и устойчивост на грешки, а това е един от основните приоритети на нашата система. Основната абстракция в Elastic са индексите. Индекс представлява колекция от еднородни документи, измежду които ще търсим данни.

**Kibana** [mark] представлява потребителски интерфейс върху ElasticSearch. Тя предоставя удобен модел за визуализиране на записаните документи. Носи със себе си възможност за рендериране на различни графики, топлинни карти и т.н. Основно се използва за бързо навигиране в големи обеми от данни.



***Фигура 13****. Примерна страница в Kibanа, която визуализиране данните от даден индекс в ElasticSearch.*



***Фигура 14.*** *Поток на данните между отделните системи в ELK стека.*

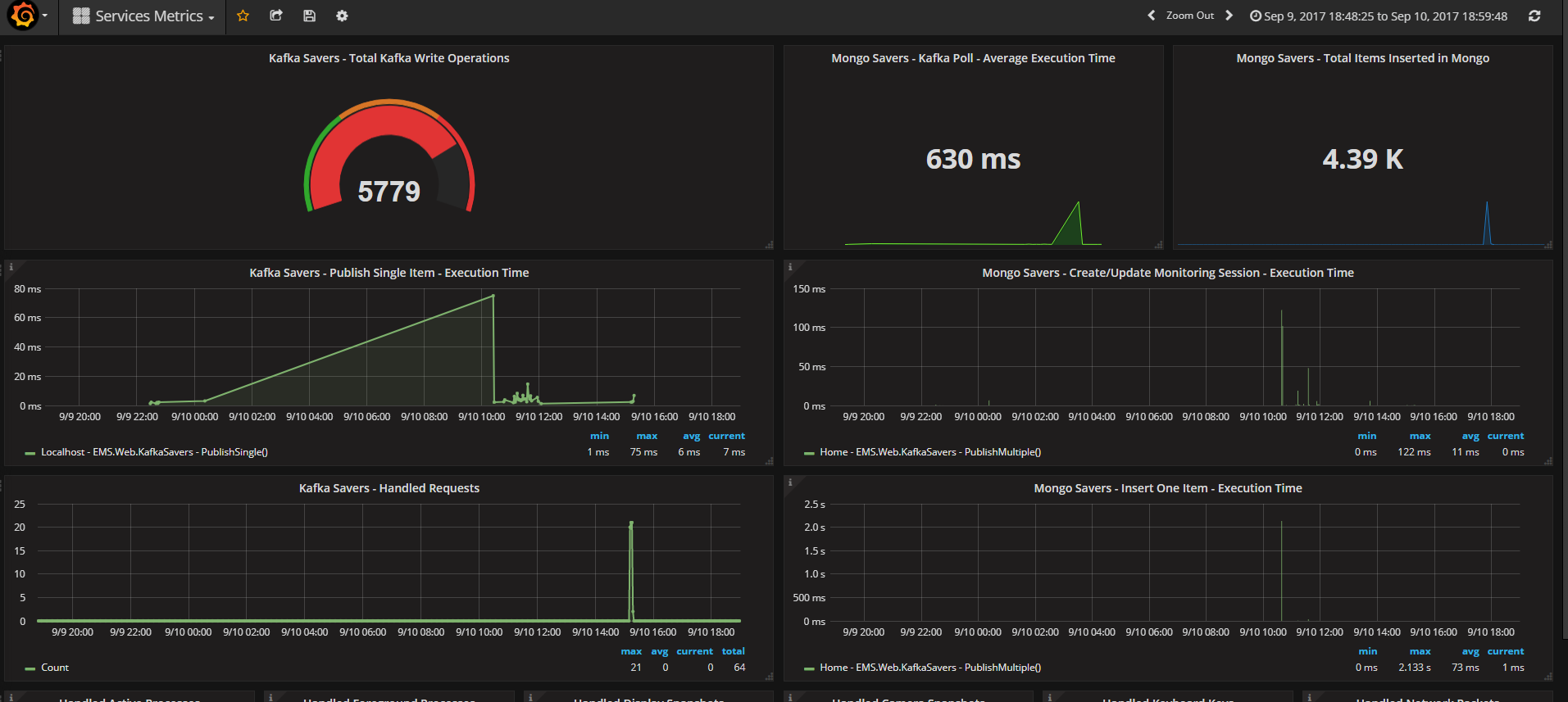
### 2.4.2. Наблюдение на метрики от системите.

За следене на метрики свързани с работата на системите и комуникацията между тях можем да използваме вече инсталираната Kibana, но факт е, че нейният основен фокус е визуализиране и търсене измежду големи обеми данни, а не рендериране на графики.

Затова ще включим още един инструмент в нашата система, а именно Grafana [mark]. Тя също е приложение с отворен код, чиито основен фокус е създаване на красиви графики с изключително мощни функционалности за агрегиране и визуализиране на различни видове данни. Grafana може да се използва и за визуализиране на данни от индустриални сензори, данни свързани с прогнози за времето, контрол на процеси и автоматизиране на домове. Широката й употреба се основава и на това че поддържа различни източници на данни като ElasticSearch, InfluxDB, Graphite, Prometheus, CloudWatch и други.

Двете основни абстракции в Grafana се наричачат „табло“ (Dashboard) и „панел“ (Panel). На едно табло, можем да прикачим множество панели. Можем да ги подреждаме по един или няколко на ред. Можем да скриваме (collapse) даден ред от таблото, за да не товарим сървъра на Grafana излишно, когато не се интересуваме от съответните панели на таблото, и искаме да изолираме само конкретни метрики.

Също така, ако имаме нужда от функционалност, която все още не съществува, можем да разработим собствен плъгин и да го инсталираме върху вече съществуваща инстанция на Grafana.



Фигура 15. Примерно табло с множество различни панели в Grafana.

За началния етап на системата ни, ще следим метрики свързани с времето за обработка на дадена заявка от различните сървъри, времето за записване на документ в MongoDB, времето за записване на документ в Kafka, броя записани документи в MongoDB и Kafka, като можем да разширим графиките да показват и ресурсите използвани от приложенията, като например процесорно време (CPU) и памет (RAM), за да можем да реагираме, ако някоя машина е критично натоварена и има нужда от ръчна намеса.

Като обобщение, имаме за цел да композираме надеждна система за следене поведението на студентите по време на различни занятия. Искаме да имаме добра видимост върху работата на отделните модули в системата и да можем да скалираме лесно при необходимост.

## 3. Програмна реализация

Всеки достатъчно опитен софтуерен инженер знае, че колкото повече редове код съдържа проекта, толкова по-труден става за поддръжка в бъдеще. Писането на гъвкав, лесно разширим и тестваем код е трудно постижима задача, а при постоянно гонене на срокове, става почти невъзможна.

За да минимизираме техническият дълг, който ще се натрупа в проекта ни с течение на времето, ще се фокусираме върху използване на възможно най-голям брой готови модули, библиотеки и системи, които имат силна подкрепа от обществото зад тях и изглеждат устойчиви в бъдеще.

Това ни дава възможността да се възползваме от човешки ресурси, външни за нашата организация, които да се грижат за еволюцията и поддръжката на кода, който ползваме.

За нас остава само частта с интеграцията между отделните системи и разработка на основните функционалности, които са строго индивидуални за нашия продукт и няма от къде да си набавим готови.

Още преди създаване на проект в средата ни за разработка, трябва много добре да сме проучили сферата (domain) с която е свързан софтуера, който ще създаваме. Това ще ни позволи да проектираме системата с много повече внимание към детайлите и спецификите на конкретната сфера и ще намали до минимум нуждата от промяна или нагласяне на модулите спрямо изискванията, които не са били ясни в началото.

Като основни правила при писане на код, ще следваме принципа за „обръщане на контрола“ (Inversion-of-control), който гласи, че даден клас/обект/функция не трябва да се грижи за инициализиране на модулите от които зависи (dependencies). Това означава, че всички елементи, които се явяват зависимости на друг елемент, трябва да бъдат създадени и инициализирани извън неговите рамки. Той не трябва да се грижи за менажиране на живота на зависимостите, които има.

Друго основно правило, което ще следваме е, всяка зависимост да се консумира през интерфейс, а не през конкретен тип. Нивото на абстракция трябва да е такова, че да позволява лесна подмяна на използваните модули, лесно автоматизирано тестване и лесна поддръжка. Всички повтарящи се функционалности трябва да бъдат изнесени в базов клас, който да бъде наследен. Така ще избегнем повтаряем код и всеки път когато правим промяна, тя ще бъде само на едно място.

Това са много семпли и мощни правила, върху които ще се стремим да изградим качествена код-база.

3.1. Структура на софтуерната система

## 4. Указания за ползване

## 5. Експериментални данни

. JWT/Cookies/Retry policies/Global exception handling/Firewalls

Кафка компресира данните посредством gzip за намаляне размера на записваните съобщения.

# Речник

Framework – работна рамка; съвкупност от базови библиотеки и функции, които са често използвани от различни приложения и улесняват работата на софтуерния разработчик, като му предоставят вече разработени и тествани модули или шаблони.

TCO (Total-cost of ownership) – тотална цена за притежание; сумата, която ще ни коства поддръжката и работата на системата.

# Източници

[1] Academic Integrity, Number of university students who admitted cheating.

<http://www.academicintegrity.org/icai/integrity-3.php>

[2] Kessler International, Percentage of students cheating using mobile devices.

<http://www.cleveland.com/metro/index.ssf/2017/02/cheating_in_college_has_become.html>

[3] Global Web Index, Average count of owned mobile devices.

<http://blog.globalwebindex.net/chart-of-the-day/digital-consumers-own-3-64-connected-devices/>

[4] JSON Web Token, Introduction and Anatomy.

<https://jwt.io/introduction/>

[5] MongoDB Licensing.

<https://www.mongodb.com/community/licensing>

[6] Apache Licensing (used by Apache Kafka, Grafana, ElasticSearch, Kibana, Logstash etc)

<https://www.apache.org/licenses/LICENSE-2.0>

[7] Microsoft, Official .NET open-source repository

<https://github.com/dotnet>

[8] Confluent, Official .NET driver for Apache Kafka

<https://github.com/confluentinc/confluent-kafka-dotnet>

[9] MongoDB, Official .NET driver for MongoDB

<https://github.com/mongodb/mongo-csharp-driver>

[10] LinkedIn, Benchmarking Apache Kafka

<https://engineering.linkedin.com/kafka/benchmarking-apache-kafka-2-million-writes-second-three-cheap-machines>

[11] Logstash, Official Website

<https://www.elastic.co/products/logstash>

[12] Logstash, Official Documentation

<https://www.elastic.co/guide/en/logstash/current/index.html>

[13] ElasticSearch, Official open-source repository

<https://github.com/elastic/logstash>

[14] ElasticSearch, Official Website

<https://www.elastic.co/products/elasticsearch>

[15] Grafana, Offical documentation

<http://docs.grafana.org/>

[16] Grafana, Official open-source repository

<https://github.com/grafana/grafana>