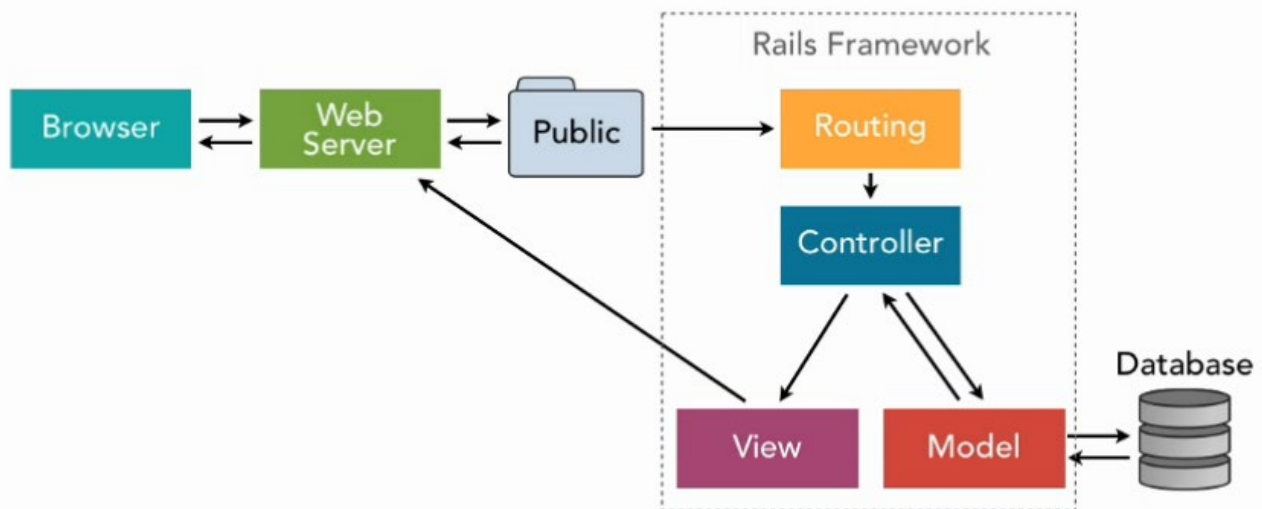


T2A1-A Workbook

Submitted by: Chen Zhang

Q1 Describe the architecture of a typical Rails application

Rails architecture



Ruby on rails is a simple useable framework for full-stack web app development with a database backend. And it follows Model, View and Controller patten which also MVC architecture patten.

- ✧ **Model** – manage the business logic and storage of database. In rails application, model also connect with the database, deal with validation, data storage, data transaction and data table attributions. And it also can add to database, find, update and delete particular data, which we can call it CRUD (create, read, update and destroy).
- ✧ **View** – manage the user interface and representing the information to the customer(end-user). And it is front-end of the application. The view are hybrid files with normal HTML structure, and also use embedded Ruby code.

Assets are normally store and supply visual components such as CSS, JavaScript and other media files (image, wav, movie, etc.)

- ✧ **Controller** – is a connector between View and Model. The requests from browser go through routes, and we use controller to specify the requests (read, create, update or destroy), to call the MODEL for the data then transfer to the VIEW then to the web server.
- ✧ **Routes** – convert URL path into form which more understandable when we develop applications.

To understand the architecture, here s an example, when we input localhost:3000 in browser, it gives the serve (Get) request which go through Routes (config/routes.rb). The it will send the request to the Controller (hello_controller.rb) and execute the action (index). In action, if it has model action (CRUD) it will find the database then jump to the VIEWS (index.html.erb). then back to the web server.

Q2 Identify a database management system (DBMS) commonly used in web applications (including Rails) and discuss the pros and cons of this database

Database management system (DBMS) normally use to add, access and process data in computer database.

MySQL is one of most popular DBMS helps developers with their applications. And there are advantages and disadvantages of MySQL.

Advantages

- Small volume, quick speed, low cost of ownership, open source, support for multiple operating systems.
- Is open-source database, the interface to support multiple language connection operation.
- The core thread it uses is completely multi-threaded and it support multi-processor. And it threads are lightweight processes that can flexibly serve users without excessive system resources.
- a very rapid and stable memory allocation system based on thread, can keep using it without consider about the stability.
- MySQL has a very flexible and secure permission and password system. When a client connects to a MySQL server, all password transfers between them are encrypted, and MySQL supports host authentication.

Disadvantages

- Hot backup not supported.

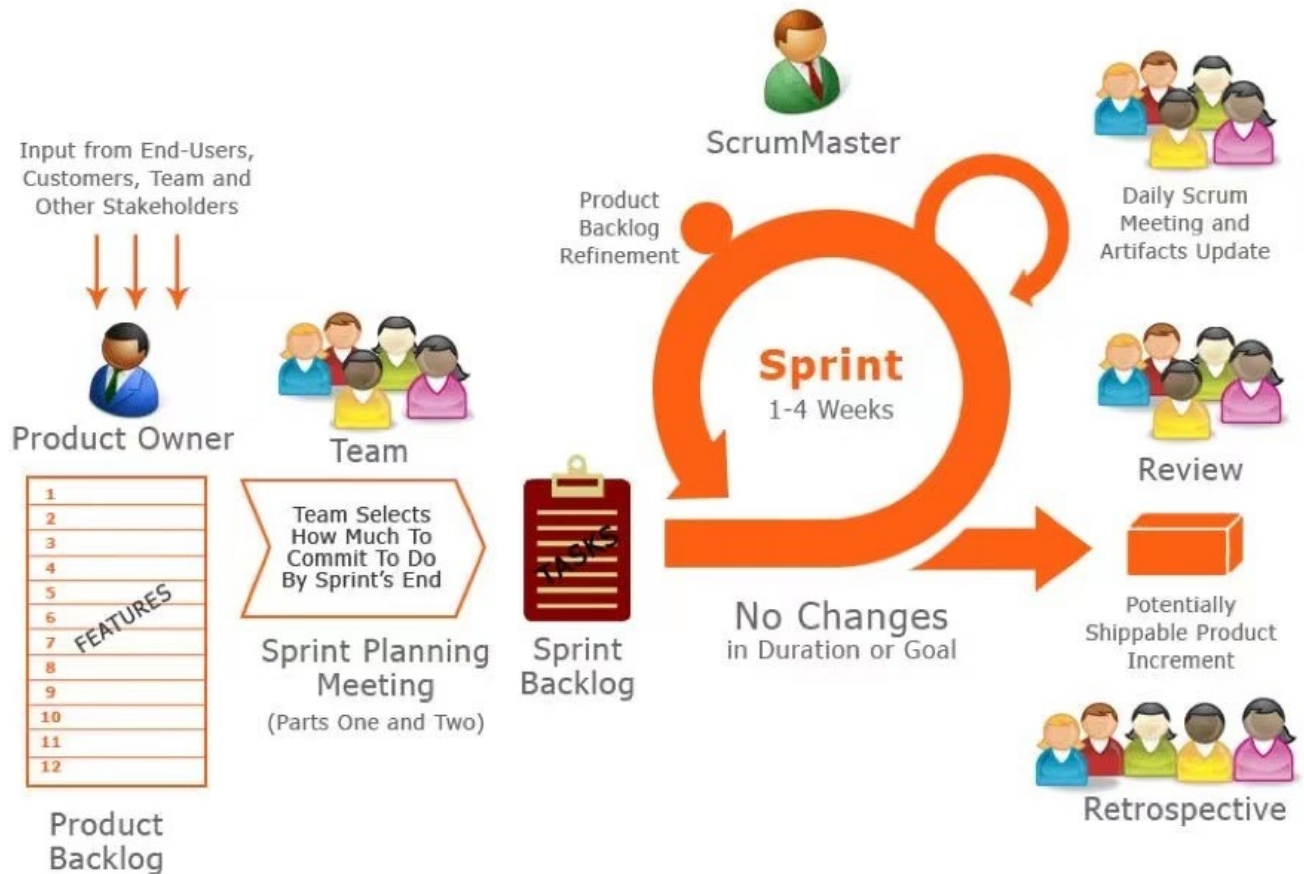
- The biggest drawback of MySQL is its security system, the main reason is it complex and non-standard, and only can be changed when you call mysql admin to re-read the user rights.
- If you have a large number of stored procedures, the memory that used for each connection to these stored procedures will increase significantly. In addition, if you overuse a large number of logical operations in a stored procedure, CPU utilization increases. This is why it is not efficient for very large database.
- The price of MySQL may vary with different platform and installation. And compare to other paid database management systems, it lacks a good developing and debugging tool.

Q3 Discuss the implementation of Agile project management methodology

Agile project management is an iterative and incremental solution for development which aims to deliver value by seeking feedbacks and changes by customers and continuing improving the project.

Scrum is the most popular Agile project management according to Agile Report.

The Scrum framework is composed of three characters, five events and three artifacts and describes how these entities to connect with each other and implementation framework.



Three characters

Scrum masters - plays a critical role in the Scrum team by ensuring that the team uses the right processes, holds the right meetings, and helps everyone understand Scrum theory, practices, rules, and values.

The Product Owner - is the core of authorized Product leadership and is one of the three roles of the Scrum team. Po plays the role as product manager.

Team - all of the people are referred as Engineers and five to seven people is a proper amount, and includes with multiple roles for product, design, front-end, back-end, testing, and so on. Team is the actual value producer;

Three artifacts

Product Backlog – is priority list for products.

Sprint Backlog – current project, only in one sprint cycle.

Increment – is the sum of all Product Backlog items completed during Sprint and the incremental value of all previous sprints.

The Scrum implementation process

1. **Before Scrum** – make sure the project suit for Scrum.
2. **Identify the basic requirements of the iteration** – iteration need a fix time period, no more than 6 weeks
3. **Prepare Sprint plan** – backlog must exist; for one project or product you can only have one backlog and one product owner; all important backlog entries have been scored based on importance.
4. **Create Sprint plan** – Sprint planning meeting is the key and most important events in Scrum, this meeting will achieve:

Target of Sprint; Team members; Sprint backlog; Determine the Sprint presentation time; Daily Scrum time and target; decide which features to complete during the sprint.

5. **Sprint planning meeting** - Set sprint goals and a backlog for the given product
6. **Daily Scrum meeting** – focus on Sprint target, normally with three questions:

What have I done to help achieving the target yesterday?

What will I do for today?

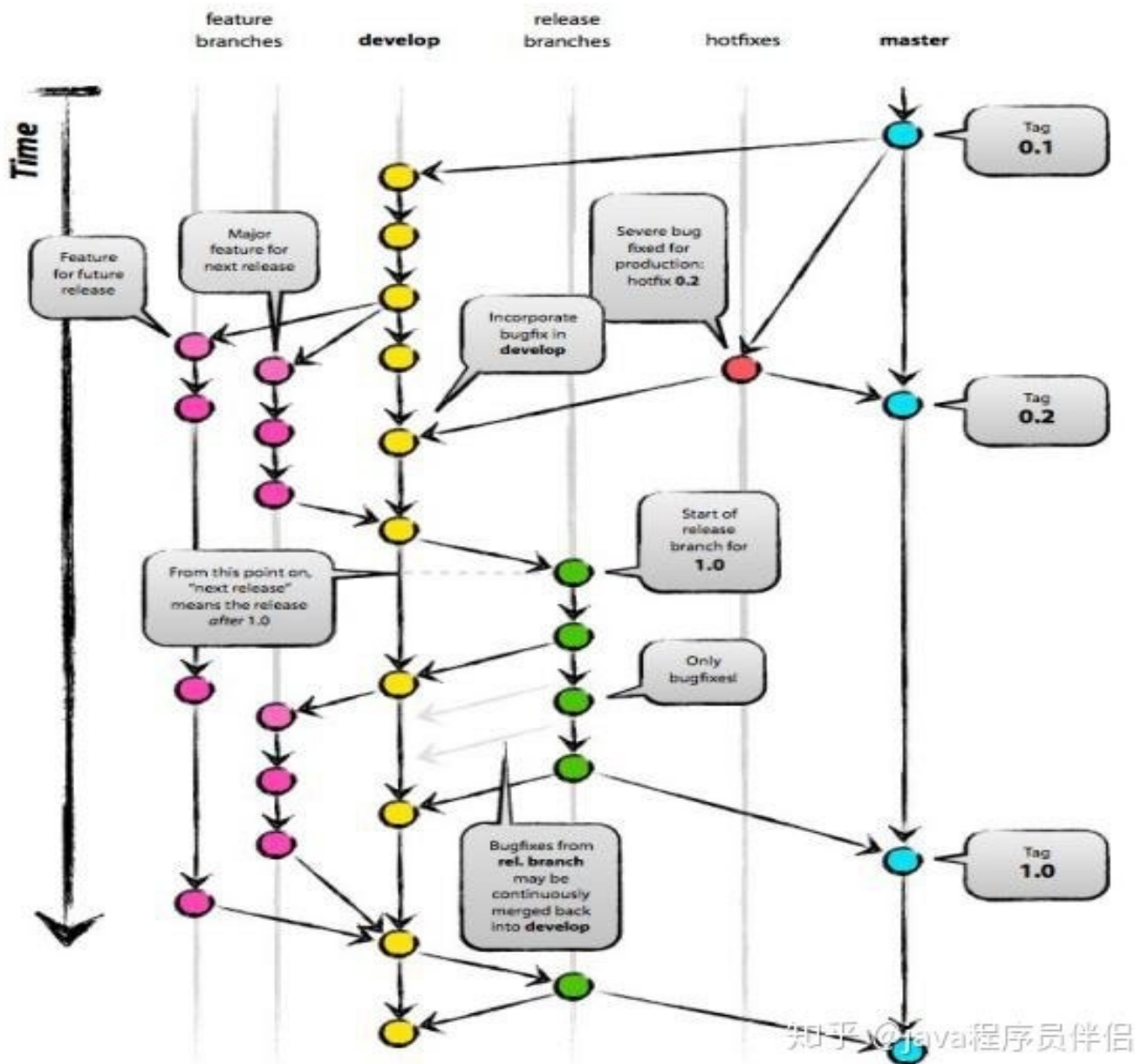
Anything slow me or team down to achieving the Spring target?

7. **Review** – Teams demonstrate the Increment during the Spring.
8. **Spring review meeting** – figure out where did go good and where did go wrong from last spring.

And how to improve it for next Spring.

Q4 Provide an overview and description of a standard source control workflow

Source control helps developer to control and manage changes code or versions in their program and make sure that they are on the right version of the source code.



A standard process for managing source control is as above picture.

Support branches:

The feature branch - is used to develop features and must be pulled from the latest develop branch code. The branch naming is basically feature.

Release branch - When the develop branch already has all the code on the line this time, and when all the tests have passed, a release branch can be created from the develop branch. The release branch is designed for releasing new product versions.

Hotfix branch - When there is a bug online that needs to be fixed urgently, the hotfix branch is derived from the current master branch.

Main branches are Master Branch and Develop Branch

Master branch - only stores online code (final code). Only when it is confirmed to go online, then merge it into the master branch and tag it (version 1.0, 2.0).

Develop branch - is the final branch at the time of development and has all the features that need to go live for all current versions. When develop team have developed different requirements in their own feature branch and need to merge tests, then we need to merger all feature branches into Develop Branch. Then submit the test, and fix the bug on the develop together. Then push the Develop Branch to Release Branches.

Source Control is keeping the codebase clean and tidy and also isolating the development among the members to avoid inefficiency and confusion caused by the mutual influence of the code in development.

Q5 Provide an overview and description of a standard software testing process (e.g., manual testing)

Software testing process is the process of operating or measuring a software system, using manual or automatic ways, to verify that the software system meets specified requirements and to identify deviations from expected results. And also, software testing is the process of judging and evaluating the quality or usability of software using by certain methods like manual testing, black box testing, static and dynamic testing, etc.

The steps in software testing:

1. Refine the test requirements according to project and product requirements.

2. According to the test requirements and the overall plan of the project, formulate the test plans which includes test time arrangement, human resource arraignment, test strategies, and conduct review.
3. According to the test requirements and related design documents, write test cases, specific operation steps of each test point, expected results, and review the use cases.
4. Prepare test environment and test data, including hardware environment and software environment for test system deployment.
5. Execute test cases, submit bugs found during testing, and perform regression testing through version iteration to verify related bugs.
6. After completing the functional test of the internal software system and the system test, submit to the customer for acceptance test
7. Write software test reports
8. Summarize the testing process and file all documentation during the testing process.

Q6 Discuss and analysis requirements related to information system security and how they relate to the project

The basic requirements of information system security mainly include the following four parts:

Confidentiality - means controlling the open scope of information resources, and preventing those who should not access confidential information from being involved the information. Related to the project, classifying the information into different levels and assigning access rights to users (admin and users).

Integrity – refers to ensuring that information system is "remained intact or in an undamaged state". Any interruption, theft, tampering or falsification of the proper characteristics or status of system information is an act that undermines the integrity of system information.

Availability – means that legitimate users can correctly use the required information when they need it without being denied service. Related to the project, the system can take many security measures to control

illegal access, but the system should not prevent legitimate users from using the information in the system. Double authentication password and SMS or email authentication is common way to prevent illegal access.

Controllability – refers to the ability to control the information flow and behavior within the authorized scope of the information system. Related to the project, if user upload or list illegal items the system can detect and control them.

Q7 Discuss common methods of protecting information and data and how you would apply them to the project

For users:

Two-factor authentication - Two-factor authentication requires users to have at least two verification methods before logging in: one is a user name and password, and the other is a mobile phone SMS verification code.

Login notification – Send SMS or email notification when log in the project with new machine or ip address.

For project

Data encryption - refers to the conversion of plaintext into ciphertext through encryption algorithms and encryption keys, and is the most reliable way for computer systems to protect information. It uses cryptographic technology to encrypt information, realize information concealment, and effectively protect the security of core data information. In rails we can use ‘bcrypt’, ‘devise’.

Files encryption - Encrypted files cannot be opened without permission and permission, and operations such as copying, printing, and taking screenshots of files are also not allowed. In rails we can use ‘lockbox’. The issue with files encryption is the files in SQL database is encryption as well. It means checking files in SQL, we need decryption first.

Access controls - define which users can access which files or folders, and what users can do with specific data.

Data backup - It's important to back up your data regularly. Once project is accidentally compromised, having a complete backup software can help you recover any lost data, avoid ransomware, and more.

Q8 Research what your legal obligations are in relation to handling user data and how they can be met for the project

The Privacy Act 1988 (Cth) (Privacy Act), which includes the Australian Privacy Principles (APPs), is the principal data protection legislation which is in order to balance the relationship between data utilization and data protection. Because most of users, cannot fully know what data the platform collects and use and cannot control the usage of the data by the platform. Related to our project:

By using, collecting and disclosure of user information – the project will adhere to this principle by clearly stating in the privacy policy page when they collect user information and how it will be used and managed throughout the project and also mentioned in the terms of conditions when registering.

Governance and accountability of the organization – the project will take reasonable steps to protect the personal information which from exploitation loss and unauthorized access and modification.

Information correction and integrity – allow authorized users to update their information to remove inaccurate. Outdated, incomplete and misleading information. For example, user need to be logged in before changing the phone number.

User rights – the project has to give a right that user can view their whole account information which are provided by the users. And after security check (login), user can also update the information.

Q9 Describe the structural aspects of the relational database model. Your description should include information about the structure in which data is stored and how relations are represented in that structure.

The relational database model uses a two-dimensional table structure to represent the relationship between entities. The basic data structure of the relational data model is the relation. A relationship is actually the

state or content of a relationship schema at a certain moment. That is, the relational schema is the type and the relation is its value. The relational schema is static and stable, while the relation is dynamic and changing over time, because relational operations are constantly updating the data in the database.

A Relation is usually said to be a Table

	Attribute (a column in table is attribute, name attribute is attribute name)	Key (An attribute group in a table that can uniquely identify a tuple)
Tuple (a row in table is a tuple)	Domain (the value range of the attribute)	

Relational schema: A row definition in a two-dimensional table, that is, a description of a relationship is called a relational schema. Generally expressed as (attribute 1, attribute 2, ..., attribute n), such as the relationship model of teachers can be expressed as teachers (teacher number, name, gender, age, title, department).

Q10 Describe the integrity aspects of the relational database model. Your description should include information about the types of data integrity and how they can be enforced in a relational database.

There are three main integrity constraints for relationships:

Entity integrity – defines the conditions that the primary key of each basic relationship in the database should meet, and can ensure the uniqueness of the tuple.

Referential integrity – defines the reference relationship between tables, that is, the reference and the referenced relationship. There is often some kind of connection between entities in the real world. In the relational model, this connection is described by the relationship. For example:

Teacher	Employ number	Name	Gender	Department No.
Department	Department No.	Department name	Location	

There is an attribute reference between the above two relationships, and the department Number is both the primary key of the department relationship and the foreign key of the teacher relationship. Then the value of the department number attribute in the teacher relationship needs to refer to the value of the department number in the department relationship or be NULL.

User-defined integrity – is a data rule formulated by a user for a specific application environment, reflecting the semantic requirements that the data involved in a specific application must meet.

Q11 Describe the manipulative aspects of the relational database model. Your description should include information about the ways in which data is manipulated (added, removed, changed, and retrieved) in a relational database.

The operations of the relational database model mainly include querying, inserting, deleting and modifying data; these operations must satisfy the integrity constraints of the relation.

INSERT – add data into table. Use as: INSERT INTO TableName (column1, column2, ...) VALUE (column1 column2, ...);

DELETE – delete all or select data from a table. Use as: DELETE TableName WHERE column=values;

UPDATE – update existing data from a table. Use as: UPDATE TableName Set column=values WHERE column=values;

SELECT – retrieve data from the database. Use as: SELECT column1, column2, ... FROM TableName WHERE column=values;

In practical, to manipulated data in relational database are basically completed with DBMS. When these basic manipulations are combined according to different conditions and sequences, various complex operations can be generated to meet various query requirements. The results of these basic manipulations are still presented in the form of a table (VIEW) which is not stored at database, only demonstrating during the querying.

Q12 Conduct research into a marketplace website (app) and answer the following parts:

Udemy is a global marketplace for online learning and teaching.

- a. List and describe the software used by the app.

Udemy uses the following software:

Python – general purpose programming language

jQuery and React – open-source JavaScript library for building user interface or UI components

Django – used as the framework for Python development

Zendesk – cloud-based help desk management solution to build customer service portal

Cart Functionality – ecommerce platform for shopping cart or checkout page

Stripe and Paypal – payment processing for internet business for fraud protection, invoicing and subscription management

Google Analytics – web analytics for tracks and reports website traffic

Memcached – High performance, distributed memory object caching system

Cloudflare – Web performance and security

OneTrust – cloud-based data privacy management compliance platform

Amazon SES – bulk and transactional email-sending service

Business tools – Slack, Jira, Trello and confluence

DevOps – GitHub, Git, Docker, Webpack, Pycharm, ESLint, Sentry and Jenkins.

Applications and Data – MySQL, Amazon S3, Amazon EC2, AWS Lambda, Redis, Google Drive and Kafka.

Utilities – Postman, Twilio SendGrid, Optimizely and Crowdin.

b. Describe the hardware used to host the app.

Udemy is service-based platform which does not require masses hardware. All the database are hosted at the cloud. Most of the courses and database are stored at Amazon S3, hosting by Amazon EC2 which handles high website traffic. Udemy use Redis to store in-memory data structure for hosting large size database.

c. Describe the interaction of technologies within the app

Backend – For the website like Udemy, there are always several programming languages to choose to build the site. It used PHP with customer MVC framework however, with increasing popularity of Python, the company switched to Python and is still processing.

Database – MySQL

Search and analytics –Udemy are using Elasticsearch as search engine and Jenkins as testing, building and delivering.

Frontend – HTML5, JavaScript, CSS3, Django, React and bootstrap are the main roles for the site structuring, presenting and formatting content. Cloudflare used as CDN produce content faster especially the larger amount users of Udemy. Fastly redefines content delivery. Django is the web framework for Python development.

Hosting – Amazon S3, Amazon EC2 to host the serve and store the its data

Other tools – Trello and Asana are using for the product design. Greenhouse tools is using for the HR works. Dropbox, Slack and Zendesk is for the productivity. These tools help the education marketplace achieve different business goals.

Security – Udemy's strategy is governed by a controls framework and uses industry-standard encryption methods (e.g., RSA asymmetric-Key algorithms).

d. Describe the way data is structured within the app

Udemy uses MySQL as database and uses relational database model which categories different entities into tables, that are then connected through the relationships they have with each other. The main structure of Udemy would be two group, Users and Courses.

User can be a Teacher and a Student; both have their own primary id link to the User id which is primary as well. Teacher and Student can have zero or many courses. An enroll course can be connected to Detail course when student pay or enroll the course. And course has its own source like video, discussing forum, comments and feedbacks.

Each one of these entities has a database table to specify and store them.

e. Identify entities which must be tracked by the app

- Users
 - Teacher
 - ◆ Join date
 - Student
 - ◆ Enroll-course
 - ◆ Comments
 - ◆ History Course
 - ◆ Join date
- Course
 - Category
 - Course name
 - Upload date
 - price

- Enroll Course
 - Enroll Course Date
 - source link
- Forum
 - Discuss
- Review and Comment
 - Reviews
 - Comments
 - Rating

f. Identify the relationships and associations between the entities you have identified in part (e)

Teacher belongs to User

Teacher has many Course

Teacher has many Review and Comment

Teacher has many forum

Student belongs to User

Student has many Enroll Course

Student has many Comments and Reviews

Student has many forum

Enroll Course belongs to Course

Course has one forum

Course has many Review and Comment

- g. Design a schema using an Entity Relationship Diagram (ERD) appropriate for the database of this website (assuming a relational database model)

