General:

1. What is your experience in Rails? 3
2. How many projects have you worked on?

*in general, more than 5 projects*

* 1. Please describe one of your best projects

*For me the best project is the one requires me to do something I never done. Make a worker to process csv files containing products & latest stocks sent via smtp to update an existing API. This projects required to be done within a week.*

1. What rails version do you use?

*The last RoR project I worked on use Rails 5*

1. What databases do you usually use?

*The last project I worked on use MongoDB. The last RoR project I worked on use SQLServer. But the most used database for overall projects I was in is PostgreSQL.*

1. What libraries do you usually use:
   1. Database => pg
   2. Background process => DelayedJob
   3. Caching => Redis
   4. Pagination => kaminari
   5. Upload processing => imagick, paperclip
   6. Cronjob => whenever
   7. Unit Test => RSpec
   8. Refactor and fix performance issues => Rubocop, rails best practice
   9. Automation tools for deployment process => capistrano
2. Are you familiar with these plugins, if so can you describe the function for each of these plugins:
   1. Sidekiq. A background processor + cronjob library
   2. Shrine
   3. Ransack
   4. Devise. A authentication library
   5. JWT. Tokenization library. Used for stateless session.
   6. Pagy
   7. Webpacker. Asset (js & css) bundler. Install, bundle, and to compress (minimize) the size.
   8. Sprocket
   9. Capistrano. For deployment. So you can deploy from your local to any server.
3. Are you familiar with Linux, which distributions do you often use?

*A bit. I used Ubuntu as a primary OS, but now I use it as a subsystem alongside windows. I use it as a development environment purpose only.*

1. Are you familiar with ubuntu, which command do you often use?

*cd. mkdir. Chown. Chmod. Sudo su. Apt install. Scp. Pwd. Which. Grep. Less. Nano. Ls -la.   
as many I need to get the project run. Since I rarely use docker.*

1. Which web server do you usually use? Are you familiar with Nginx?

*I know simple configuration in Nginx & apache2. But most of the time it set by Ops. So syntax to restrict access on such file or related to it, I need to google first.*

Technical:

1. Describe the difference between module and class?

*Module used to group related classes. You make a class from an object need to be manipulate.*

1. Describe the difference between the class variable and the class instance variable?

*To use class instance variable, you need to make an instance from the class first. For class variable, you can call it directly to the classname. Class instance variable value changes affect only to the instance of the class. If you change the class variable value, the value change for all the instances of that class.*

1. What is polymorphism and give an example for that in a model class?

*One example for polymorphism is toString() method in java model class. For each model you create which inherit Object Class (toString() defined here), will have access to toString() as default, and you can also override it to your custom.  
So to distinguish it from inheritance, since inheritance more like sharing attributes & methods, in polymorphism the child allowed to behave as the parent and also allowed to do different stuff for the same operation from the parent.*

1. How do you handle N+1 performance issues?

*Use Eager Loading. This method use explicit joins to load main data and all the related model to it. Combined it with caching & indexing will make the query significantly faster. Review the query for eager load, use subquery if needed, to process filtered data instead of process the data then filter.*

1. Describe what is Lazy Load, and when to use it?

*Lazy load when you get a data from a model, and all the related model not automatically included in the first query, but rather saved as another query that will be call whenever the data (the related data) needed.  
So lazy load is fast for retrieving the main data even if it have a lot of related data to it, when the related data not served the first time it loaded. For example, instagram posts, the comments only shows after you click the comments button.*

1. How to handle Unique Records in Concurrent Process?

*Define which column to have unique index. Use database transaction to keep persistency in data.*

Coding Test:

I want you to build a Ruby application without any frameworks, but you can use other stacks like databases if necessary. The application is to count how many files are inside a path with the SAME content. It could have the same name or not.

For example, I have folders with 5 files:

3 files have the same content, the content is “abcdef”, and the others have different content

In this case if I pass the folder's path, the script will return me the number of files that have the same content. On those folders content1 = content2 = content3, So the application will return content + number. So in this case is: abcdef 3

Also you need to return the bigger number of files if there are multiple files with the same content. For example, if there are 4 files with content “abcdef” and 5 files with content “abcd” then the return value should be: abcd 5

Please take note that the example files are in bytes level, but the code needs to be able to handle big files as well. Think of Megabytes, Gigabytes level. And the application needs to be able to handle tens/hundreds/thousands/millions of files.

The other requirement is I want this app to scan a path DYNAMICALLY. This means I should be able to scan any folder that I like, without changing anything on the code. It could be a parameter or a config file.

Please let me know the repo once it's done (Github/BitBucket), also please take note that the way you commit your code is being scored.

Here I put the code for this test: <https://github.com/ainun-jariya/ror_virtualspirit> on branch *master*, and later on *develop* branch.

Database Test:

1) I have the following Users Table:

|  |  |  |
| --- | --- | --- |
| ID | Name | Mark |
| 1 | John | A |
| 2 | Marissa | B |
| 3 | Bob | C |
| 4 | Britany | C |

Please create a query to get some users who have a specific mark.

|  |  |
| --- | --- |
| Mark | Count |
| A | 1 |
| B | 1 |
| C | 2 |
| A+B | 2 |

|  |  |
| --- | --- |
| A+C | 3 |
| B+C | 3 |

|  |  |
| --- | --- |
| A+B+C | 4 |

select mark as Mark, count(id) as Count from users where mark in ('A','B','C') group by mark;

2) I have users & books table as follows:

**Users Table**

|  |  |
| --- | --- |
| ID | Name |
| 1 | John |
| 2 | Bob |
| 3 | Britany |

**Books Table**

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Name | Date | user\_id |
| 1 | Bachelor of Information Systems | 01-01-2019 | 1 |
| 2 | Bachelor of Design | 02-02-2019 | 2 |
| 3 | Bachelor of Commerce | 03-03-2019 | 3 |
| 4 | Associate Degree in Health Science | 04-04-2019 | 3 |
| 5 | Master of Architectural Technology | 05-05-2019 | 2 |
| 6 | Bachelor of Psychology | 06-06-2019 | 2 |
| 7 | Associate Degree in Information Systems | 07-07-2019 | 1 |

Please create a query to get the latest book (based on date column) for each user. Below is the result example:

|  |  |  |
| --- | --- | --- |
| User | Book | Date |
| John | Associate Degree in Information Systems | 07-07-2019 |
| Bob | Bachelor of Psychology | 06-06-2019 |
| Britany | Associate Degree in Health Science | 04-04-2019 |

SELECT

users.name AS User, latest\_borrowers.name AS Book, date AS Date

FROM

(

SELECT

books.user\_id, books.name,

books.date AS date,

ROW\_NUMBER() OVER(

PARTITION BY user\_id

ORDER BY date DESC

) AS rn

FROM books

) AS latest\_borrowers

INNER JOIN users ON users.id = latest\_borrowers.user\_id

WHERE rn = 1;