Apprenticeship Portfolio

Oliver Matthew Bowker (220263618)

August 4, 2023

Contents

List of Figures

List of Tables

Todo list

Projects

I will now briefly describe the two projects that will be discussed during this document/KSBs. More detail will be added to these projects throughout the multiple KSBs.

Off Product Inspector (OPI)

The Off Product Inspector (OPI) is a tool used by internal members of the BBC to view what media we offer to both internal and external partners. The tool has two main functions:

1. To provide a search mechanism on all brand/series/episodes metadata that partners can access.
2. To provide a tool that generates per partner specific links to catered or promotional data.

The metadata we provide partners is written in JSON (JavaScript Object Notation) and can be hard to read, as a lot of the people are non-technical we provide a web based GUI for users to navigate and browse the data.

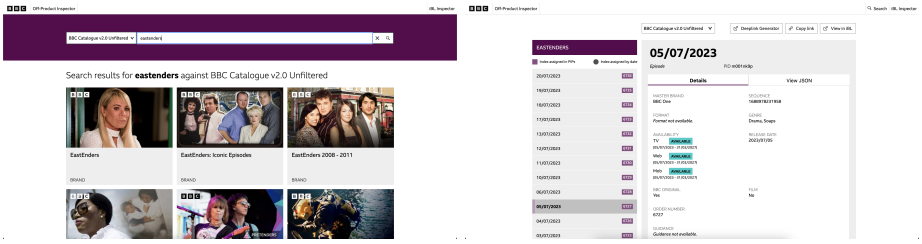

Figure 1: Screenshot of the search and episode/series/brand metadata viewer, (left-to-right)

The *'Deeplink Generator'* is another part of the OPI that allows users to generate links to promotional/catered content. In the past links have been malformed causing errors for partners, this tool aims to fix that by not having to manually write the links ourselves.

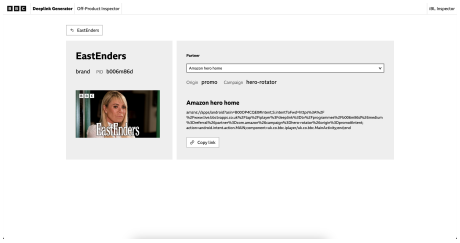

Figure 2: Screenshot of the Deeplink Generator tools

Schedules

The BBC plan to deprecate a service called *'Nitro'* in order to save money on running the service and licenses attached to the service. As part of this certain data the service provided now has to supplied elsewhere. My team is responsible for supplying external partners with the schedule/EPG information.

This data contains what is on linear television at what time, and also includes the aforementioned episode/series/brand metadata. In the real world this data is seen when you click 'guide' or corresponding button on the TV remote.

Figure 3: Example of how EPG/schedules data can be used

Schedules

We have two pipelines, the ***catalogue*** pipeline that is responsible for all the data that is currently accessible to partners on iPlayer and the new ***schedules*** pipeline. We first have to ingest the schedule data from another source, which consists of over 1000 files and then parse that data for what we want to share with partners. This parsed data is then stored in redis for future use. In the older catalogue pipeline this process was done by an EC2 (server), however with this pipeline we wanted less to manage and decided to go with a lambda (serverless) approach. To speed up ingestion of data we decided to experiment with multiple lambdas being invoked at the same time for concurrent processing of the files. This is the part of the project I was tasked to document/create.

B1

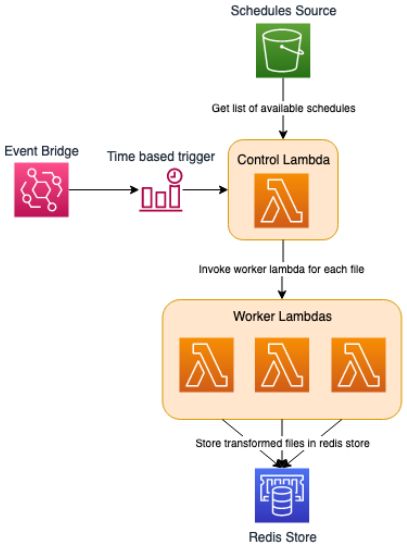

Figure 4: Ingester AWS architecture.

Figure 4 shows what the cloud infrastructure on AWS looks like. We get an event from event bridge, the available data is then retrieved from source and the parsing is then handed over to multiple concurrent lambdas. Finally this data is stored in redis for the next component in the pipeline to use. I have illustrated this flow in the sequence diagram below.

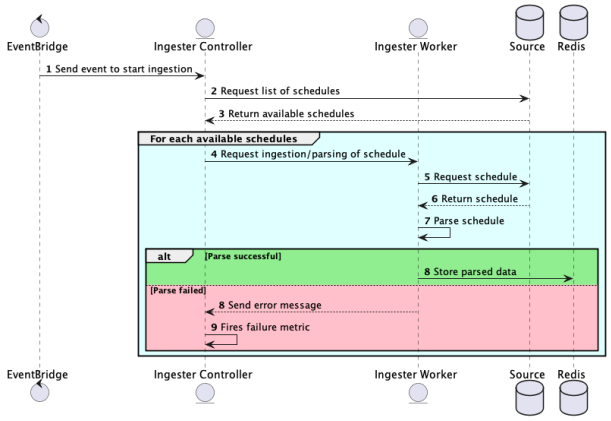

Figure 5: Ingester basic execution sequence diagram.

For completeness I have also drawn a sequence diagram for the removal of old schedules from the redis store. This process is done just before the parallelised lambdas are called to do their work. Removal is important here as if they were not removed it would have a knock on effect down the pipeline as there would be unnecessary processing done for schedules that are no longer being used or accessible to partners.

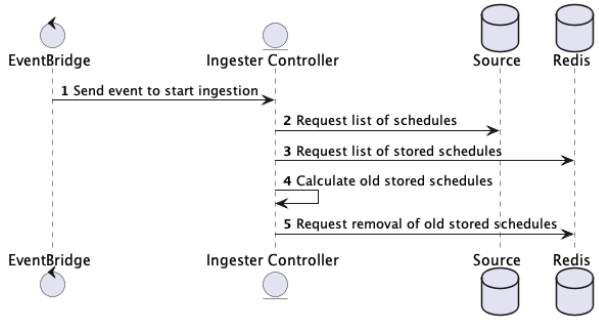

Figure 6: Ingester old data removal sequence diagram.

In the pipeline there is a second component called the *Schedule Generator* that is responsible for outputting the data in a format partners will receive. For this component we did not have the worker lambda, and instead had a single lambda do all the processing. The two graphs belows show the difference in lambda runtime between the two components.

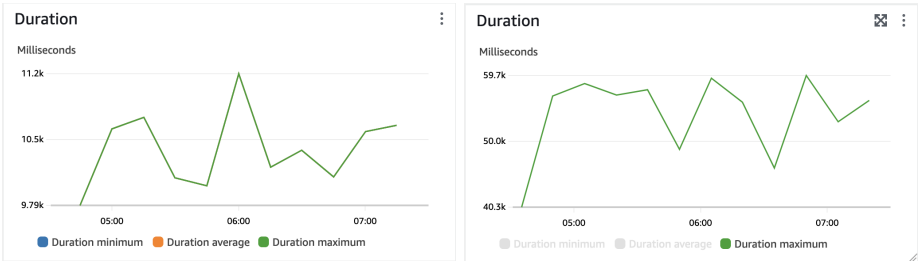

Figure 7: Screenshot of AWS lambda duration for schedule Ingester (left) and Generator (right).

As can be seen, using parallelised lambdas is 4x quicker on average than using a single lambda. This is important as the quicker this pipeline can run the 'fresher' the data is for partners. After seeing these results there is now some talks about also changing the Schedule Generator to use the same pattern.

This was also the first project where we upgraded to using the new version of the AWS SDK. This was something we hadn't looked into however was something that would be beneficial when we had to upgrade to node 18. AWS would no longer package version 2 of their SDK by default in lambdas. This would result in an extra 70MB of unpacked data being added to the final build which would slow down our build time and cost us a little bit more money, albeit fractions of a penny. I lead the investigation into using the new SDK and if it was feasible. This upgrade then also lead to the upgrade of our credentials provider which had security vulnerabilities that needed fixing.

L3

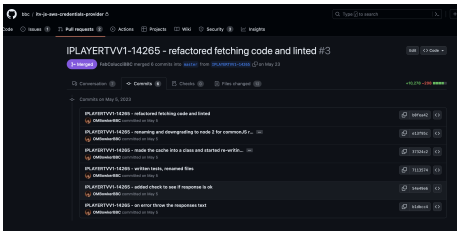

Figure 8: Screenshot commits for credentials refactor.

Off Product Inspector Tooling

The Off Product Inspector (OPI) was made so that people in editorial and partner marketing and non technical team members, could view the data that should be available to partners at that moment. It allows users to search for pids (unique identifiers) and titles of brands, series and episodes and then view the data of these items. I worked on two main things to extend this project, upgrade to support our new v2 catalogues and add the Deeplink Generator.

Support v2 Catalogues

The OPI originally was designed to support the v1.1 catalogue, however the catalogue had been extended to v2 which included different data and also had 3 *horizon catalogues*.

- **v2.0 catalogue** - Contains all data, both what is currently available and unavailable on iPlayer.
- **v2.0 0-day catalogue** - Contains data that is on iPlayer right now.
- **v2.0 8-day catalogue** - Contains data that is on iPlayer right now and also contains things that will be available within 8 days.

Partners are being encouraged to move over to the new v2 version of the catalogue so its good to be able to see what theses catalogue should offer. Me and a fellow junior software engineer began by working on the *spike* for the project. A spike is a task thats aim is to gain knowledge and information on ahead of doing the work. This can often times mean creating a small MVP to show that it is possible. I talk about this and other software lifecycle issues in the **Software Engineering-1.pdf** document. The spike we created was an MVP that allowed the switching between catalogue, however the code was not refined, there was no tests written and zero documentation, this would all come later. We then showed what we had come up with to the team and then began the process of breaking the work down into slices/tickets.

B5, P3

I had never done or heard of a *spike* before doing this for the OPI. It's easy to get carried away and go beyond the scope of the spike. Towards the end of this spike I began writing some tests for certain things but learned that this was not part of what a spike was. When doing another spike for the schedules ingester, I took this lesson and stayed within the scope. This was also the first time I had been given a task that was UI based. This is an area I have much more knowledge in from my own teachings and allowed me to help/teach my colleague I was spiking with learn about how things like React hooks work and the interplay between React and Next.js. Further along in the project I also suggested that we move to a newer way of writing css and components called *styled components*. Some benefits of these are:

L2, L3, L7

- Styles are not separate from the component itself, making it easier to debug.

- Styled components can take props to easily customise aspects of a components style.
- Default html tags are given a more meaningful name, making it easier to understand the role of a tag.

Below is an example of the difference between using styled components and styling using SaSS (Syntactically Awesome Style Sheets).

```jsx
import styled from '@emotion/styled'
import PropTypes from 'prop-types'

const AvailabilityStatus = styled.div`
  font-size: 13px;
  font-weight: bold;
  padding: 4px 8px;
  background-color: ${props => props.available ? '#10ccca' : '#d7d7d7'};
`

const AvailabilityStatusExample = ({ available }) => (
  <AvailabilityStatus available={available}>
    {available ? 'AVAILABLE' : 'UNAVAILABLE'}
  </AvailabilityStatus>
)

AvailabilityStatusExample.propTypes = {
  available: PropTypes.bool
}

export default AvailabilityStatusExample
```

```scss
.availability-status {
  font-size: 13px;
  font-weight: bold;
  padding: 4px 8px;
  background-color: #d7d7d7;

  .available {
    background-color: #10ccca;
  }
}
```

Figure 9: Screenshots of how not using styled components would work, JSX on left, SCSS on right.

```jsx
import '../styles/example.scss'
import PropTypes from 'prop-types'

const AvailabilityStatusExample = ({ available }) => (
  <div className={`availability-status${available ? ' available' : ''}`}>
    {available ? 'AVAILABLE' : 'UNAVAILABLE'}
  </div>
)

AvailabilityStatusExample.propTypes = {
  available: PropTypes.bool
}

export default AvailabilityStatusExample
```

Figure 10: Screenshot of how it looks using styled components.

Deeplink Generator

Deeplink generator stuff here.

P2, P3

Deeplink Generator Presentation

On the 18th of July 2023 I presented at the Partnerships show n' tell meeting on behalf of my team. I presented the work we did for the deeplink generator and have included the slides in a separate document.

This was the first time I've presented anything to a larger group of people and I was nervous going in. The presentation was short, due to the work being

discussed not being very large, however it still took around 5 -7 minutes for the whole presentation. Overall I'm happy with how I managed the presentation and with the presentation itself. It's something I should do more of in the future to become more comfortable with these kinds of things. For future I would take my time more when going through the slides. I felt like I rushed a little bit due to my nerves, however I feel like this is something that comes with practice and being the the situation.

Personal feedback on OPI

I show the questions and feedback I received in the document **Feedback Form Output.docx**.

B1 (Skill)

'Identify, document, review, and design complex IT-enabled business processes that define a set of activities that will accomplish specific organisational goals and that provide a systematic approach to improving those processes.'

B2 (Skill)

'Design and develop technology roadmaps, implementation strategies, and transformation plans focused on digital technologies in order to achieve improved productivity, functionality, and end-user experience in an area of technology specialization.'

B3 (Skill)

'Deliver workplace transformations through planning and implementing technology-based business- change programmes, including setting objectives, priorities, and responsibilities with others in an area of technology specialization.'

B4 (Knowledge)

'The strategic importance of technology enabled business processes, and how they are designed and managed in order to determine a firm's ability to compete effectively.'

B5 (Knowledge)

'The principles of business transformation and how organisations integrate different management functions in the context of technological change.'

B6 (Knowledge)

'Own employer's business objectives and strategy, its position in the market, and how it adds value to its clients through the services and/or products it provides.'

B7 (Knowledge)

'How to justify the value of technology investments and apply benefits management and realisation.'

P1 (Skill)

'Negotiate and agree digital-and-technology- specialism delivery budgets with those with decision- making responsibility.'

P2 (Skill)

'Develop and deliver management level presentations which resonate with senior stakeholders, both business and technical.'

P3 (Skill)

'Professionally present digital-and-technology- solution-specialism plans and solutions in a well- structured business report.'

P4 (Skill)

'Demonstrate self-direction and originality in solving problems, and act autonomously in planning and implementing digital-and-technology-solution-specialist tasks at a professional level.'

P5 (Skill)

'Be competent at negotiating and closing techniques in a range of interactions and engagements, both with senior internal stakeholders and external stakeholders.'

P6 (Knowledge)

'The role of learning and talent management in successful business operations.'

L1 (Skill)

'Evaluate the significance of human factors to leadership in the effective implementation and management of technology-enabled business processes.'

L2 (Skill)

'Develop own leadership style and professional values that contribute to building high-performing teams.'

L3 (Behaviour)

'Inspire and motivate others to deliver excellent technical solutions and outcomes.'

L4 (Behaviour)

'Establish high levels of performance in digital and technology solutions activities.'

L5 (Behaviour)

'Be results and outcomes driven in order to achieve high key performance outcomes for digital and technology solutions objectives.'

L6 (Behaviour)

'Promote a high level of cooperation between own work group and other groups in order to establish a technology-change led culture.'

L7 (Behaviour)

'Develop and support others in developing an appropriate balance of leadership and technical skills.'

L8 (Behaviour)

'Create strong positive relationships with team members to produce high-performing technical teams.'

L9 (Knowledge)

'The role of leadership in contemporary technology-based organisations.'

L10 (Knowledge)

'The personal leadership qualities that are required to establish and maintain an organisation's technical reputation.'

L11 (Knowledge)

'The role of leaders as change agents, and the identity of contributors to successful implementation.'

T1 (Skill)

'Apply broader technical knowledge, combined with an understanding of the business context and how it is changing, in order to deliver to the company's business strategy.'

T2 (Skill)

'Demonstrate effective technology-leadership and change-management skills for managing technology- driven change and continuous improvement.'

T3 (Skill)

'Create and implement innovative technological strategies in order to support the development of new products, processes, and services that align with the company's business strategy, and develop and communicate compelling business proposals to support the strategies.'

T4 (Knowledge)

'How to monitor technology related market trends and research and collect competitive intelligence.'

T5 (Knowledge)

'Technology road-mapping concepts and methods, and how to apply them.'

SE-S01 (Skill)

'Architect, build, and support leading-edge concurrent-software platforms that are performant to industry standards and that deliver responsive solutions with good test coverage.'

SE-S02 (Skill)

'Drive the technology-decision-making and development process for projects of varying scales, considering current technologies including DevOps and Cloud Computing, and evaluate different technology-design and implementation options, making reasoned proposals and recommendations.'

SE-S03 (Skill)

'Develop and deliver distributed or semi-complex software solutions that are scalable, and that deliver innovative user experiences and journeys that encompass cross-functional teams, platforms, and technologies.'

SE-S04 (Skill)

'Update current software products, improving their efficiency and functionality, and build new features to product specifications.'

SE-S05 (Skill)

'Accomplish planned software-development tasks that deliver the expected features within specified time constraints, security, and quality requirements.'

SE-S06 (Skill)

'Be accountable for the quality of deliverables from one or more software-development teams (source code quality, automated testing, design quality, documentation, etc.), and following company-standard processes (code reviews, unit testing, source code management, etc.).'

SE-K01 (Knowledge)

'The rationale for software-platform and solution development, including the organisational context.'

SE-K02 (Knowledge)

'The various inputs, statements of requirements, security considerations, and constraints that guide solution architecture and the development of logical and physical systems' designs.'

SE-K03 (Knowledge)

'The methodologies designed to help create approaches for organizing the software-engineering process, the activities that need to be undertaken at different stages in the life-cycle, and techniques for managing risks in delivering software solutions.'

SE-K04 (Knowledge)

'The approaches used to modularise the internal structure of an application, and to describe the structure and behaviour of applications used in a business, with a focus on how they interact with each other and with business users.'

SE-K05 (Knowledge)

'How to design, develop, and deploy software solutions that are secure and effective in delivering the requirements of stakeholders, and the factors that affect the design of a successful code.'

SE-K06 (Knowledge)

'The range of metrics which might be used to evaluate a delivered software product.'

Appendix

Appendix A - Full diagram of schedule pipeline infrastructure.

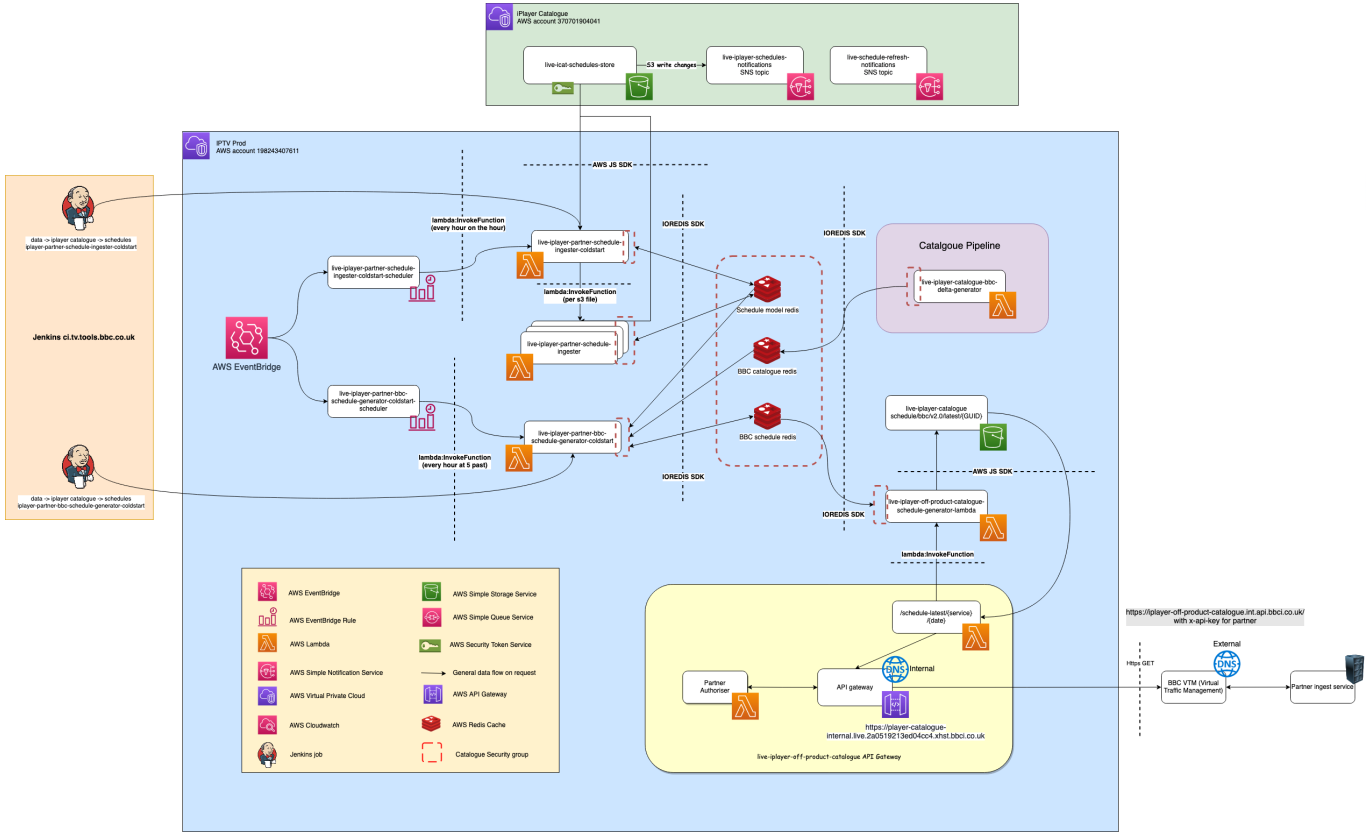

Figure 11: Schedules threat model showing the entire pipeline.