DC4400 Project Report

Processing of schedule updates/changes

Oliver Matthew Bowker (220263618)

February 2, 2024

Contents

List of Figures

List of Tables

1 Executive Summary

2 Introduction

The BBC is now over 100 years old (BBC, 2022) and is well known for it's TV channels and radio stations that are broadcast over the airwaves (Pilnick, Baer, 1973, p.3) to peoples homes across the UK. However this old way of broadcasting, sending out airwaves on a certain frequency to an antenna, is becoming less popular in the modern age of the internet. A study done by Ofcom showed that people *'watched on average about 16% less broadcast TV between 2019 ... and 2022'*, with viewing *'decreasing by 47%'* (Ofcom, 2023, p.7) between ages 16-24. In addition another study carried out by media analyst firm Ampere found that in 2021 37% of people claimed to watch no linear TV, this increased to 45% by 2023 (Ampere Analysis, 2023).

This fall correlates with the significant rise in internet enabled TVs in the home, with statista finding that *'In 2014 just 11 percent of households in the UK owned a Smart TV, whereas, in 2023, nearly 74 percent of households reported owning a Smart TV.'* (Statista, 2023).

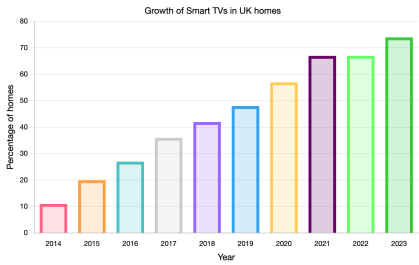


Figure 1: Bar chart showing growth of smart TVs in UK homes (Statista, 2023) *Created with Line Graph Maker* (Line Graph Maker, 2023)

Some of these devices still support OTA broadcasts, however devices like the Amazon Fire TV stick and Googles Chromecast, are purely internet based; However they do offer a *'guide/epg'* section with Amazon having a development guide (Amazon, 2021) on how to integrate with it. Director general of the BBC, Tim Davie, in 2022 stated:

> *'The vision is simple: from today we are going to move decisively to a digital-first BBC'* (Davie, 2022)

This statement highlights the goal to put more organisational focus on these new forms of media and internet enabled devices.

This report will discuss an upgrade carried out to the BBCs *'off-product'* schedules system, responsible for delivering up to date schedules to partners such as Freeview, Amazon and more. First I will give some background on the project, where I will discuss topics including storage Solutions and how they can work in parallel/multi-threaded systems, and strategies to protect live code

systems in a CI/CD environment. I will also give some background on the starting architecture of the system and how the changes align with the BBCs and teams OKRs (Sparks, 2024).

Following that, I will discuss the work that was done. This will be broken down into 5 sections that align with our teams ways of working flow.

1. Requirements and epic creation
2. Investigation and Spike
3. Slicing and task/ticket creation
4. Development of software
5. Releasing of software

I will then talk about the outputs of the project. Theses will include burn-up charts for the projects, dashboards created, documentation of the final architecture and a description of the final product.

Finally I will discuss potential improvements for future iterations. This will range between small code changes to a complete re-architecture of the system.

3 Background

In this section I will discuss the background work and research done for this project. I will start by disusing my teams place in the organisation and our OKRs, explaining how this project helps us hit these objectives. I will then outline the current architecture and the initial design for the project. Finally I will discuss some areas of interest around project, these include cloud computing, database parallelisation strategies and CI/CD challenges.

3.1 Organisation and Team

The BBC is broken into multiple layers with different responsibilities and goals. I am in a team called *SpaceChimps* which is partner of the partnerships group, which itself is in the product group.

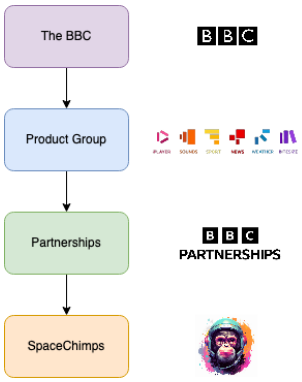


Figure 2: Image showing SpaceChimps place in the BBC (Bowker, 2023).

Our main aim as a team is to provide data to partners that they can use on their devices to promote BBC content. The project described in this report does just that by providing schedules for live content to partners. This aim fits directly into Partnerships objectives:

- It helps drive growth as we are able to get content out to more people on more devices, increasing exposure to the BBC.
- It helps us improve our partner experience by working with them on integrating the data into their feeds.
- This project reduces the total time processing data, which therefore reduces our costs and makes us more sustainable.

All objectives can can be seen in **Appendix A** (BBC Partnerships, 2023).
As well as schedules we also provide a *'catalogue'* of episodes, series and brands that are currently available on iPlayer.

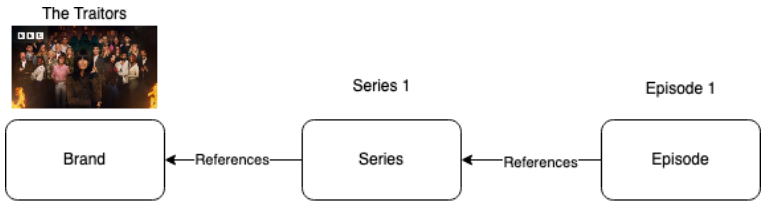


Figure 3: Image showing catalogue hierarchy and how they reference each other.

We provide both a 0 day and 8 day catalogue, these containing programme data that are available now or up to 8 days in the future. We also have an *'unfiltered'* catalogue that is not available to partners which contains all programme data with no availability limits. This unfiltered catalogue is what is used by the schedule pipeline to get it's data about episode/series/brands within the schedule.

3.2 Original Architecture

The original solution was composed of AWS services that created a pipe and filter architecture which transformed inputs into multiple outputs (Somerville, 2016, pp.182-183) that can be used by partners. The figure below shows this architecture.

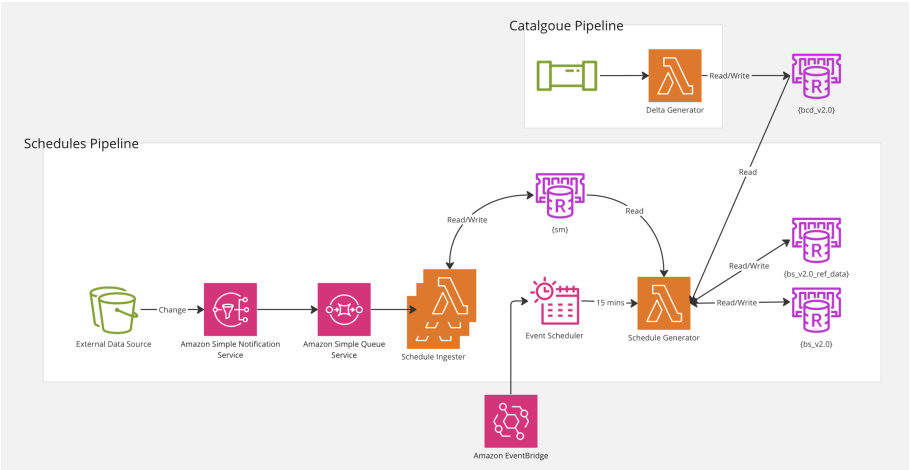


Figure 4: Image showing the initial architecture.

A pipe and filter architecture is achieved with the combination of AWS lambda (Amazon Web Services, 2024a) with updates being triggered by a publish/subscribe model (Somerville, 2021, pp.179) achieved through AWS' Simple Notification Service (SNS) (Amazon Web Services, 2024b) and Simple Queue

Service (SQS) (Amazon Web Services, 2024c) with the former publishing and the latter subscribing. The lambda only runs when a message is published to the SQS, this triggers the lambda and the pipeline begins precessing the new message.

The data is updated by an external system in AWS S3 (Amazon Web Services, 2024d), this publishes a message that the data has changed, our first lambda (the ingester) receives this message and processes/stores it into a *'common'* model that's used internally only. Our internal model is then up to date, however our partner facing model, created by the schedule generator, is not. In the original system this lambda was not driven by events, but instead ran every 15 minutes and processed all the schedules in one go, whether they had updated or not.

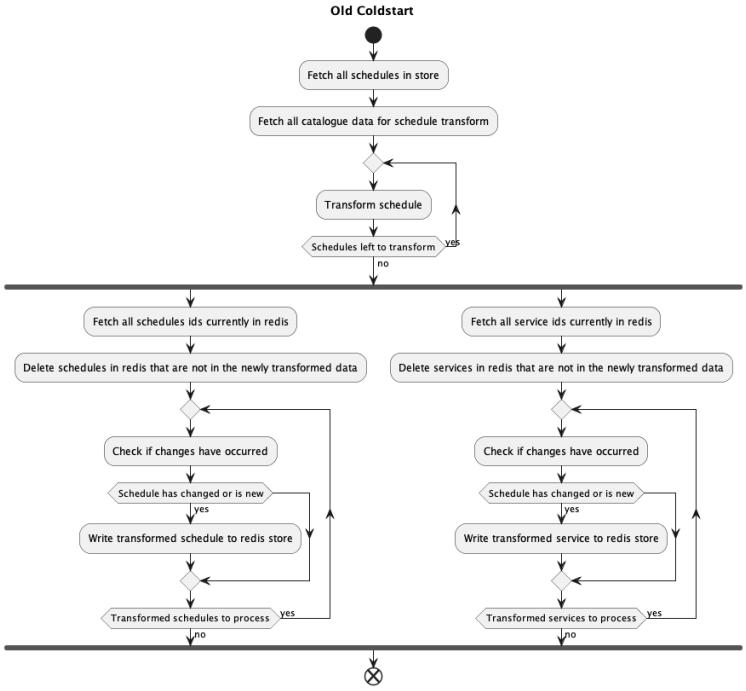


Figure 5: Activity diagram showing schedule generators logic.

This could leave partners with data up to 15 minutes out of date, dependant on when they retrieve the data, resulting in end users being shown incorrect schedules.

3.2.1 Initial design solution

The schedules pipeline also requires data from our catalogue pipeline, for titling, descriptions, viewer discretion warnings and subtitles. For this reason it needs

to be alerted when catalogue data changes as well as when schedules change. An initial design had already been complete before the work started by another member of the team.

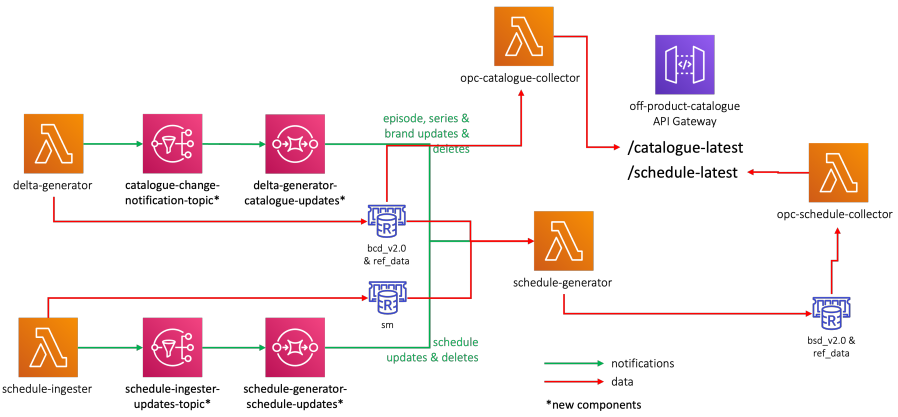


Figure 6: Image showing the initial updated schedules design (Lloyd, 2023).

The above diagram has been edited down to include only work related to the project, a full diagram including notifications being directly sent to partners can be found in **Appendix B**. This diagram also includes our endpoints and collectors that partners use to retrieve the data. This proposed system uses the pub/sub model described earlier and will subscribe to both catalogue and schedule events.

Alongside the architectural design an algorithm was proposed for these events.

- **For schedule updates** - We want to update the partner facing model of the schedule linked in the notification. This schedule has a list of broadcasts that map to episodes in the catalogue. A list should be maintained within each episode to create a link between both, allowing episode updates to trigger updates to schedules that reference them. All catalogue data referenced in a schedule should be copied over to the schedules keyspace, leaving the catalogue keyspace and items untouched.
- **For schedule deletes** - Remove schedule from store, and remove schedule reference from list contained within each episode that schedules references in its list of broadcasts.
- **For catalogue/programme updates:**

  For episodes - Update episode in store, update each schedule that is linked in episodes broadcast list.

  For series/brands - Get all episodes linked to either the series or brand, update each schedule that is linked in each episodes broadcast list.

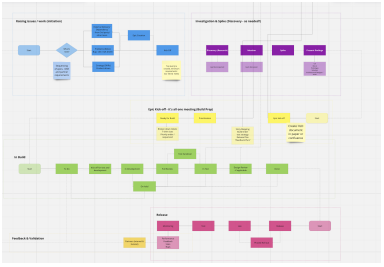
Flow diagrams can be found in **Appendix C** (Lloyd, 2023).

4 Research

4.1 Storage Solutions for Parallelism

4.2 CI/CD and it's challenges

5 Work Done



5.1 Investigation and Spike

5.2 Slicing and ticket creation

5.3 Build software

5.4 Release

6 Outputs

7 Future Work

8 Conclusion

9 References

BBC. (2022) *Timeline - History of the BBC* [online]. Available at `https://www.bbc.co.uk/historyofthebbc/bbc-100/timeline/` (accessed on 18th January 2024).

Pilnick, C. Baer, S, W. (1973) *Cable television: A guide to the technology* [online]. Available at `https://www.rand.org/content/dam/rand/pubs/reports/2006/R1141.pdf` (accessed on 18th January 2024).

Ofcom. (2023) *Media Nations report 2023* [online]. Available at `https://www.ofcom.org.uk/__data/assets/pdf_file/0029/265376/media-nations-report-2023.pdf` (accessed on 18th January 2024).

Ampere Analysis. (2023) *Almost half of Internet users say they have switched off broadcast TV* [online]. Available at `https://www.ampereanalysis.com/press/release/dl/almost-half-of-internet-users-say-they-have-switched-off-broadcast-tv` (accessed on 18th January 2024).

Statista. (2023) *Household penetration of smart TV sets in the United Kingdom (UK) from 2014 to 2023* [online]. Available at `https://www.statista.com/statistics/654074/smart-tv-penetration-in-uk-households/` (accessed on 18th January 2024).

Line Graph Maker. (2023) *Bar Graph Maker* [online]. Available at `https://linegraphmaker.co/bar-graph` (accessed on 19th January 2024).

Amazon. (2021) *Linear TV Integration Guide Overview* [online]. Available at `https://developer.amazon.com/docs/fire-tv/linear-tv-integration-guide-overview.html` (accessed on 18th January 2024).

Davie, T. (2022) *A digital-first BBC* [online]. Available at `https://www.bbc.co.uk/mediacentre/speeches/2022/digital-first-bbc-director-general-tim-davie` (accessed on 18th January 2024).

Sparks, R. (2024) *OKRs: the ultimate guide to objectives and key results* [online]. Available at `https://www.atlassian.com/agile/agile-at-scale/okr` (accessed on 18th January 2024).

Bowker O. (2023) *Analysis of BBC process and proposal for a new API key management system.* Assignment for DC4000, MSc Digital and Technology Solutions Specialist, Aston University, Unpublished.

BBC Partnerships. (2023) *Partnerships - one year on.* BBC. Unpublished.

Somerville, I. (2016) *Software engineering.* Tenth edition, Global edition. London: Pearson Education.

Somerville, I. (2021) *Engineering Software Products: An Introduction to Modern Software Engineering.* Global edition. London: Pearson Education.

Amazon Web Services. (2024a) *AWS Lambda* [online]. Available at `https://aws.amazon.com/lambda/` (accessed on 2nd February 2024).

Amazon Web Services. (2024b) *Amazon Simple Notification Service* [online]. Available at `https://aws.amazon.com/sns/` (accessed on 2nd February 2024).

Amazon Web Services. (2024c) *Amazon Simple Queue Service* [online]. Available at `https://aws.amazon.com/sqs/` (accessed on 2nd February 2024).

Amazon Web Services. (2024d) *Amazon S3* [online]. Available at `https://aws.amazon.com/s3/` (accessed on 2nd February 2024).

Amazon Web Services. (2024e) *Amazon DynamoDB* [online]. Available at `https://aws.amazon.com/dynamodb/` (accessed on 2nd February 2024).

Lloyd, K. (2023) *iPlayer Schedules - BBC Partner Schedule Generator - Delta Update Technical Design*. BBC. Unpublished.

10 Appendix

10.1 Appendix A

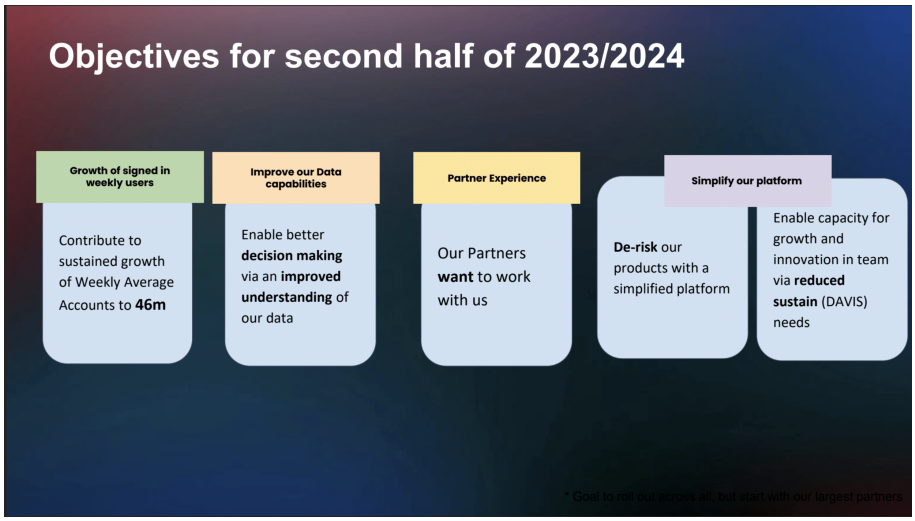


Figure 7: Image taken from a presentation given at a partnerships context setting event (BBC Partnerships, 2023).

10.2 Appendix B

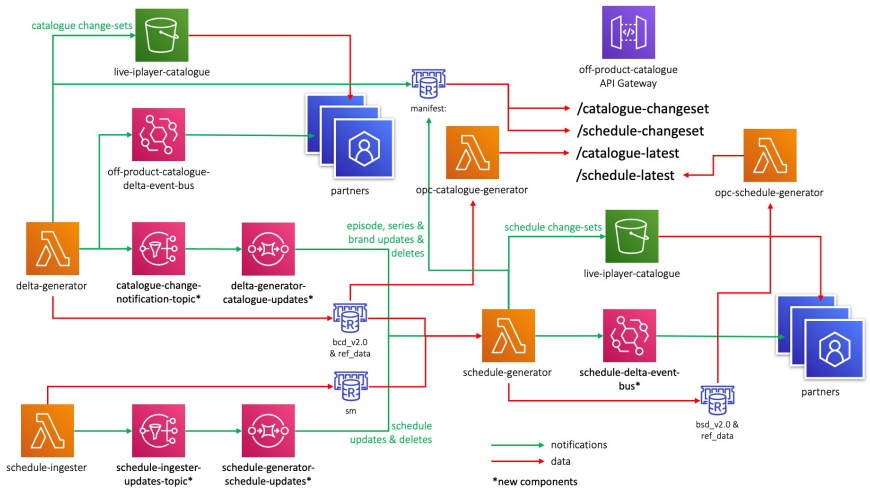


Figure 8: Full diagram of design for schedules pipeline, including future notifications to partners work (Lloyd, 2023).

10.3 Appendix C

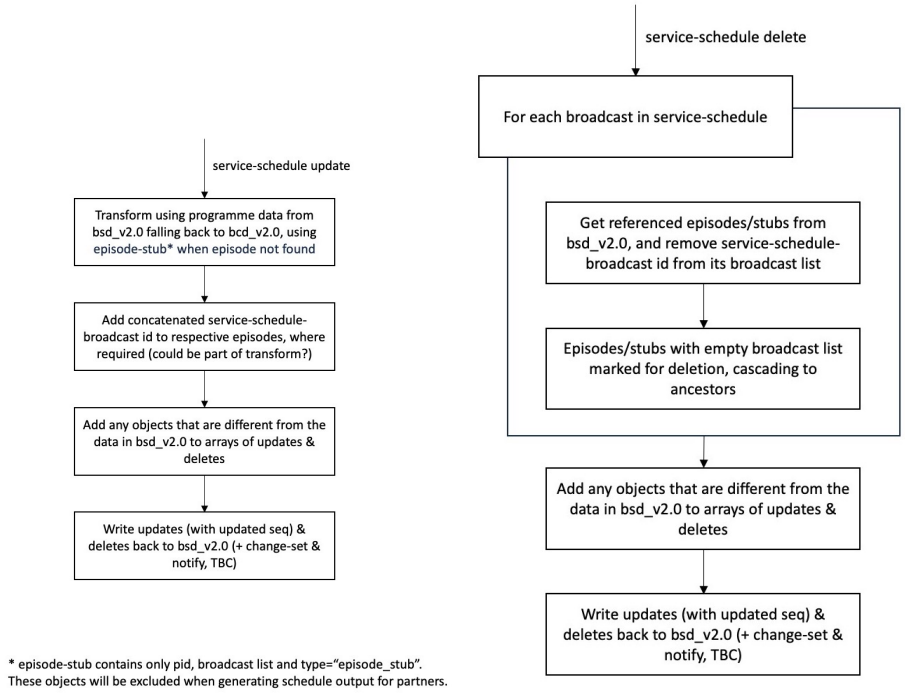


Figure 9: Flow diagrams for schedule events (Lloyd, 2023).

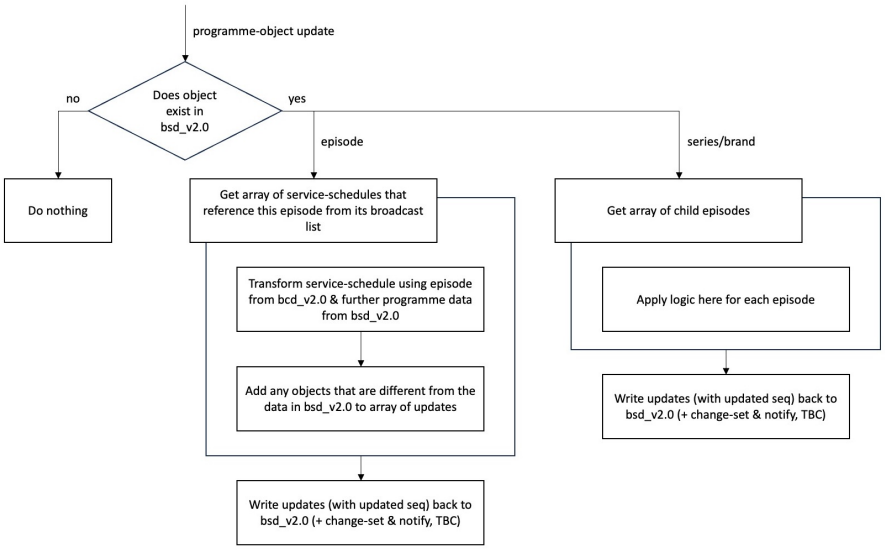


Figure 10: Flow diagram for catalogue/programme events (Lloyd, 2023).