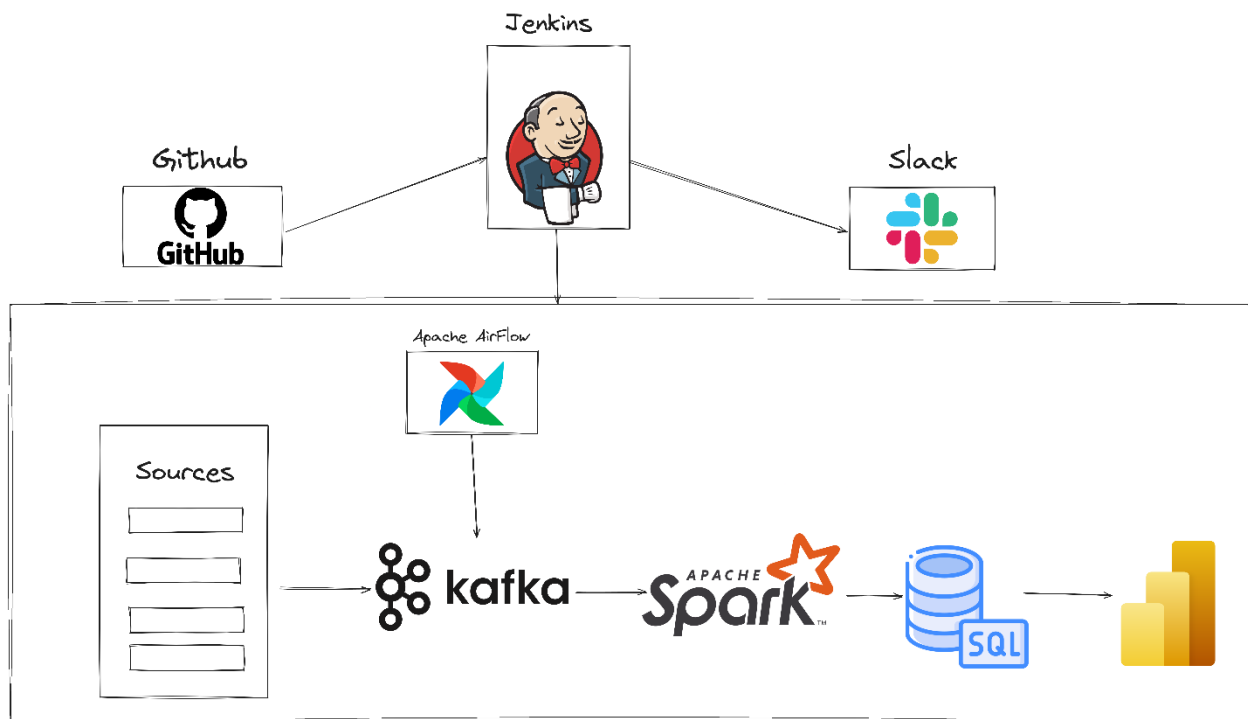


Movies Project

Project Overview

- **Introduction:** Our Project is aiming to integrate datasets from multiple streaming services (Disney, Amazon Prime, Hulu, Netflix) to gain insights.
- **Objective:** Our Goal is to process data from multiple sources, ingest it using Kafka, transform it with Apache Spark, and store it in SQL databases for streaming analysis and reporting.

Project Pipeline:



Architecture Explanation

- **Diagram Walkthrough:** Use the diagram to explain how the components interact:
 - **Sources:** Our datasets (Disney, Amazon Prime, Hulu, Netflix) are the starting point.
 - **Kafka:** Data from these sources is ingested and passed to Kafka for streaming.
 - **Apache Spark:** Spark processes and transforms the streamed data, applying the necessary transformations.
 - **SQL Database:** Processed data is stored in an SQL database for streaming analysis.
 - **Visualization:** The data visualized in Power BI to create reports and dashboards.

Data Sample:

the data is structured and stored in CSV.

#	B	C	D	E	F	G	H	I	J	K	L	M	N
	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	source	
1	Movie	The Grand Seduction	Don McKellar	Brendan Gleeson, Taylor Kitsch, Gordon Pinsent	Canada	30-Mar-21	2014		113 min	Comedy, Drama	A small fishing village must procure a loci	Amazon Prime	
2	Movie	Take Care Good Night	Girish Joshi	Mahesh Manjrekar, Abhay Mahajan, Sachin Khed India		30-Mar-21	2018 13+		110 min	Drama, International	A Metro Family decides to fight a Cyber C	Amazon Prime	
3	Movie	Secrets of Deception	Josh Webber	Tom Sizemore, Lorenzo Lamas, Robert LaSardo, R	United States	30-Mar-21	2017		74 min	Action, Drama, Suspense	After a man discovers his wife is cheating	Amazon Prime	
4	Movie	Pink: Staying True	Sonia Anderson	Interviews with: Pink, Adele, Beyonc	United States	30-Mar-21	2014		69 min	Documentary	Pink breaks the mold once again, bringing	Amazon Prime	
5	Movie	Monster Maker	Giles Foster	Harry Dean Stanton, Kieran O'Brien, George Cost	United Kingdom	30-Mar-21	1989		45 min	Drama, Fantasy	Teenage Matt Banting wants to work with	Amazon Prime	
6	Movie	Living With Dinosaurs	Paul Welland	Gregory Chisholm, Juliet Stevenson, Brian Hensor	United Kingdom	30-Mar-21	1989		52 min	Fantasy, Kids	The story unfolds in a an English seaside	Amazon Prime	
7	Movie	Hired Gun	Fran Strine	Alice Cooper, Liberty DeVitto, Ray Parker Jr., Dav	United States	30-Mar-21	2017		98 min	Documentary, Special Interest	They are the "First Call, A-list" musicians,	Amazon Prime	
8	Movie	Grease Live!	Thomas Kail, Alex Rudzinski	Julianne Hough, Aaron Tveit, Vanessa Hudgens, I	United States	30-Mar-21	2016		131 min	Comedy	This honest, uncompromising comedy chri	Amazon Prime	
9	Movie	Global MeltDown	Daniel Gilboy	Michael Par	Canada	30-Mar-21	2017		87 min	Action, Science Fiction, Suspi	A helicopter pilot and an environmental sc	Amazon Prime	
10	Movie	David's Mother	Robert Allan Ackerman	Kirstie Alley, Sam Waterston, Stockard Channing	United States	1-Apr-21	1994		92 min	Drama	Sally Goodson is a devoted mother to her	Amazon Prime	
11	Movie	Forest Fairies	Justin G. Dyck	Emily Wilder, Adrian Cowan, Gary McKendzie, Jere	Canada	4-Apr-21	2016		88 min	Adventure, Kids	Amanda stumbles upon a hidden village o	Amazon Prime	
12	Movie	Take Care	Liz Tuccillo	Leslie Bibb, Kevin Curtis, Nadia Dajani	United States	10-Apr-21	2014		93 min	Comedy	When a car crash leaves Frannie immobili	Amazon Prime	
13	Movie	The Night Eats The W	Dominique Rocher	Anders Danielsen Lie, Golshifteh Farahani, Denis	France	17-Apr-21	2018		94 min	Horror, Suspense	After waking up in an apartment the night	Amazon Prime	
14	Movie	Resiliencia	Jep Barcelona	Rafinha Alcantara, Marc-Andr	Spain	24-Apr-21	2020		46 min	Documentary, Sports	The documentary follows the midfielder	Amazon Prime	
15	Movie	Elon Musk: The Real	L Sonia Anderson	Elon Musk, Per Wimmer, Julie Anderson-Ankenbr	United Kingdom	2-May-21	2018		74 min	Documentary	Discover the meteoric rise of Elon Musk, t	Amazon Prime	
16	Movie	Summer '03	Becca Gleason	Joey King, Jack Kilmer, Andrea Savage, Paul Sche	United States	3-Jun-21	2019		96 min	Comedy, Drama	Summer '03 is the hilarious story of 16-ye	Amazon Prime	
17	Movie	Zoombies	Glenn Miller	Marcus Anderson, Kahlil Lyman, Andrew Asper			2016 13+		87 min	Horror, Science Fiction	When a strange virus quickly spreads thro	Amazon Prime	
18	TV Show	Zoo Babies		Narrator - Gillian Barlett			2008 ALL	1 Season		Kids, Special Interest	A heart warming and inspiring series that	Amazon Prime	
19	TV Show	Zoë Coombs Marr: Bossy Bottom		Zoë Coombs Marr			2020 18+	1 Season		Comedy, Talk Show and Vari	Zoë Coombs Marr has been on hiatus. Sc	Amazon Prime	
20	Movie	Zoe	Drake Doremus	Ewan McGregor, L			2018 R		104 min	Science Fiction	ZOE tells a tale of forbidden love between	Amazon Prime	
21	TV Show	Zoboomafoo		Alex Cazares, Ben Diskin, Cristina Vee, Erika Ishi			2001 TV-Y	1 Season		Kids	Chris and Martin Kratt bring their enthus	Amazon Prime	
22	TV Show	Zo Zo Zombies: Mini-Series					2020 TV-Y7	1 Season		Action, Anime, Comedy	Zo Zo Zombies is our undead hero who, wit	Amazon Prime	
23	Movie	Zis Boom Bah	William Nigh	Grace Hayes, Mary Healy			1941 NR		62 min	Arts, Entertainment, and Cul	"Hey, kids, let's get together and put on a	Amazon Prime	
24	TV Show	ZeroZeroZero		ANDREA RISEBOROUGH, HAROLD TORRES, CLAI	Italy		2020 18+	1 Season		TV Shows	ZeroZeroZero series follows the journey o	Amazon Prime	

Columns:

- ➔ Show_id.
- ➔ Type.
- ➔ Title.
- ➔ Director.
- ➔ Cast.
- ➔ Country.
- ➔ Date_added.
- ➔ Release_year.
- ➔ Rating.
- ➔ Duration.
- ➔ Listed_in.
- ➔ Description.
- ➔ Source (Netflix , Hulu, Amazon_prime, Disney)

Challenges:

- The datasets didn't have a well-defined Fact table structure.
- Since an OLAP setup wasn't possible, the system was implemented using an OLTP.
- Using SQL Database Instead of NoSQL (MongoDB or Cassandra)
- The SQL database was optimized to handle the processed data efficiently