



eTime Re-architecture

Recommendations for improving processing

07/28/2025



NOTICE TO THE RECIPIENT OF THIS PROPOSAL ("RECIPIENT")



The information contained herein, as well as any information shared by Capgemini in furtherance of this proposal, are the proprietary and confidential information ("Confidential Information") of Capgemini, and their release would offer substantial benefit to competitors offering similar services. The Confidential Information includes descriptions of methodologies and concepts derived through substantial research and development efforts undertaken by Capgemini. Recipient shall use this proposal and/or any Confidential Information solely for the purpose of evaluation for contract award. If Capgemini and Recipient are not otherwise parties to a non-disclosure agreement which protects Capgemini's Confidential Information herein, Recipient agrees to maintain in confidence and not to disclose the Confidential Information except to its employees or agents on a "need to know basis." Your acceptance of the submission of this proposal and evaluation thereof indicates your agreement to these terms. It is the position of Capgemini that this proposal and/or Confidential Information are not considered subject to release under the Freedom of Information Act, if applicable.

Copyright 2021 Capgemini America, Inc. All rights reserved. No part of this proposal may be reproduced by any means or transmitted without the prior written permission of Capgemini except internally by you for the purpose of evaluating this proposal. All copies of this proposal (or any portion hereof) should be returned to Capgemini or, at Capgemini's option, destroyed at the end of the proposal evaluation period if Capgemini is not selected.

Neither submission by Capgemini nor your acceptance of this proposal, in whole or in part, constitutes acceptance by Capgemini of any contractual terms contained in your Request for Proposal or similar document ("RFP"), if any, and shall not form a binding agreement between the parties, other than with respect to confidentiality as set forth herein. Such an agreement shall only exist upon the execution of a mutually acceptable contract by both parties ("definitive agreement"), which will contain the exclusive agreement between the parties regarding the scope and terms of Capgemini's engagement. Failure by Capgemini to comment upon any contractual terms in the RFP shall not constitute acceptance thereof. The RFP, and this proposal, will be superseded by the definitive agreement. Any proposed duties or obligations referenced herein should be understood as being subject to commercially reasonable efforts. Capgemini makes no representations, warranties, or assurances to you, except as may be mutually agreed in the definitive agreement. The terms of this Notice shall control over the terms in any RFP in the event of a conflict or inconsistency.

Notwithstanding anything contained in this proposal, any delay in Capgemini's performance or inability of Capgemini to perform as stated in this proposal due to conditions or circumstances related to COVID-19, regardless of foreseeability, will be excused and Capgemini reserves the right to modify its response accordingly. Similarly, a change in the approach or location for rendering the Services may be necessary to help mitigate COVID-19 impacts. If any of these situations arise, Capgemini will work collaboratively with you to reach mutual agreement regarding changes to the approach, schedule or other aspect of the Services that may be necessary.

The submission of this proposal is expressly subject to and conditioned upon the enforceability of the terms of this Notice. In the event that any of the terms of this Notice are held not to be enforceable, then, at Capgemini's option, Capgemini may withdraw this proposal without liability of any nature.

The term "Capgemini" appearing in this proposal may refer to Capgemini America, Inc., or to one or more of its global affiliates. However, this proposal is being submitted only by Capgemini America, Inc., which is solely responsible for its contents, and Capgemini America, Inc. shall be the contracting entity if its proposal is selected.

Capgemini America, Inc.

TABLE OF CONTENT

- 01 **Background & Scope Understanding**
- 02 **Solution Approach**
- 03 **Implementation Approach**
- 04 **Assumptions**
- 05 **Commercials**





Image Source: FirmX Truck Leasing

BACKGROUND & SCOPE UNDERSTANDING



Our Understanding Of FirmX's Objectives & Scope Of Work

eTime Data Estate was implemented as part of ADP modernization initiative.
After go-live, system had various issues with respect to scalability and data volume.

Goal of this project is to address the critical issues observed as given below:

1 System Assumptions Mismatch

- The source system can provide data in semi-structured form, and there are no controls and overall knowledge on the internals for what can cause a change
- Observing higher incremental volume of data than what was expected originally. (Increase of 3X-4X)

2 Small File Issue

- Too many small files are generated by current process
- Many single record files for query by Athena (AWS S3 GetObject Rate limit)

3 Execution Cycle for API Process

- API Pagination is set to 500 records, which causes slowness when there are large volumes of change in source system.
- This causes a looping within the process, and in cases where there are 50,000 records can take over 3.5 hours.

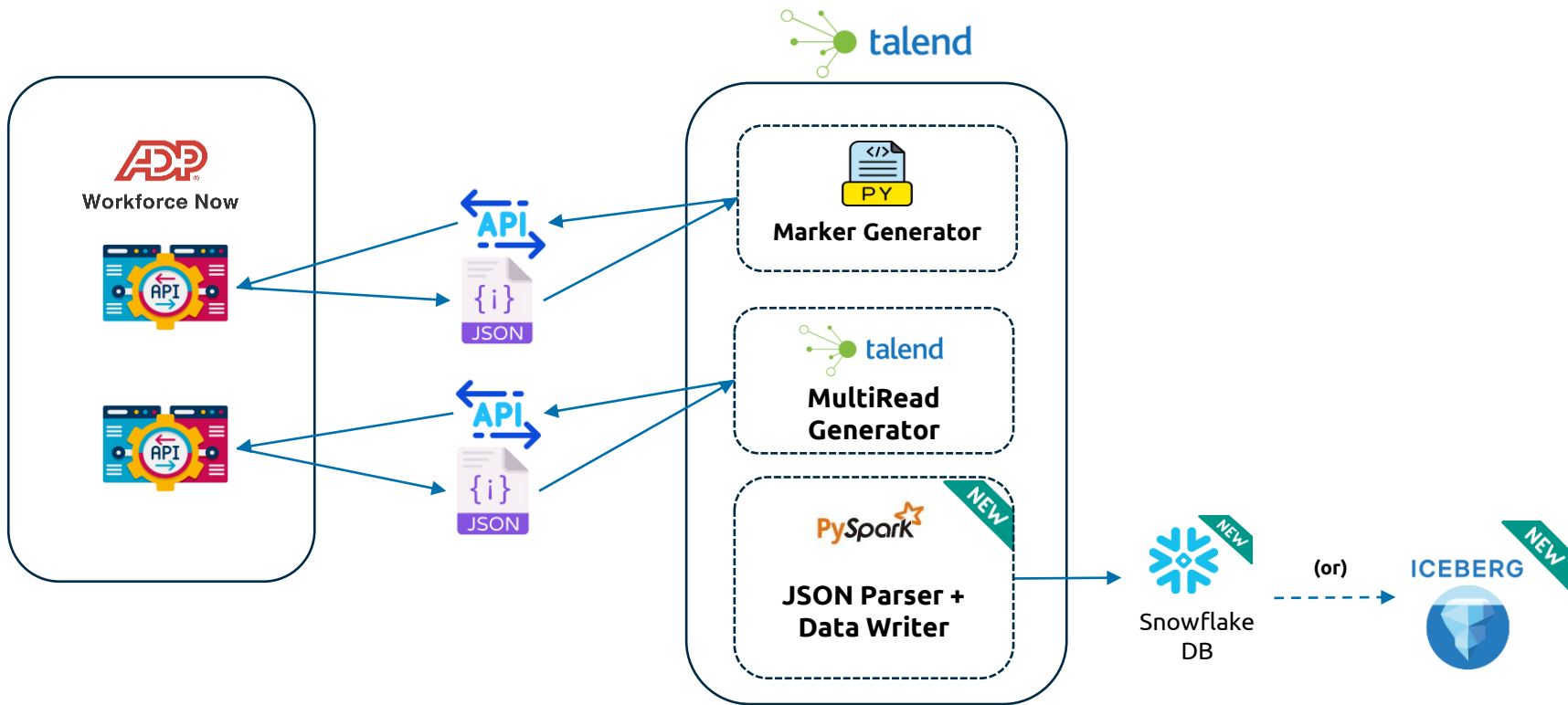


GET THE FUTURE YOU WANT

To improve scalability we are proposing two viable options...(1/2)

Redesign Approach #1 - Snowflake

Reuse existing Python and Talend code, **Leveraging** Spark on Serverless for scalable processing, **Adopt** Snowflake as the target storage for its performance and flexibility.

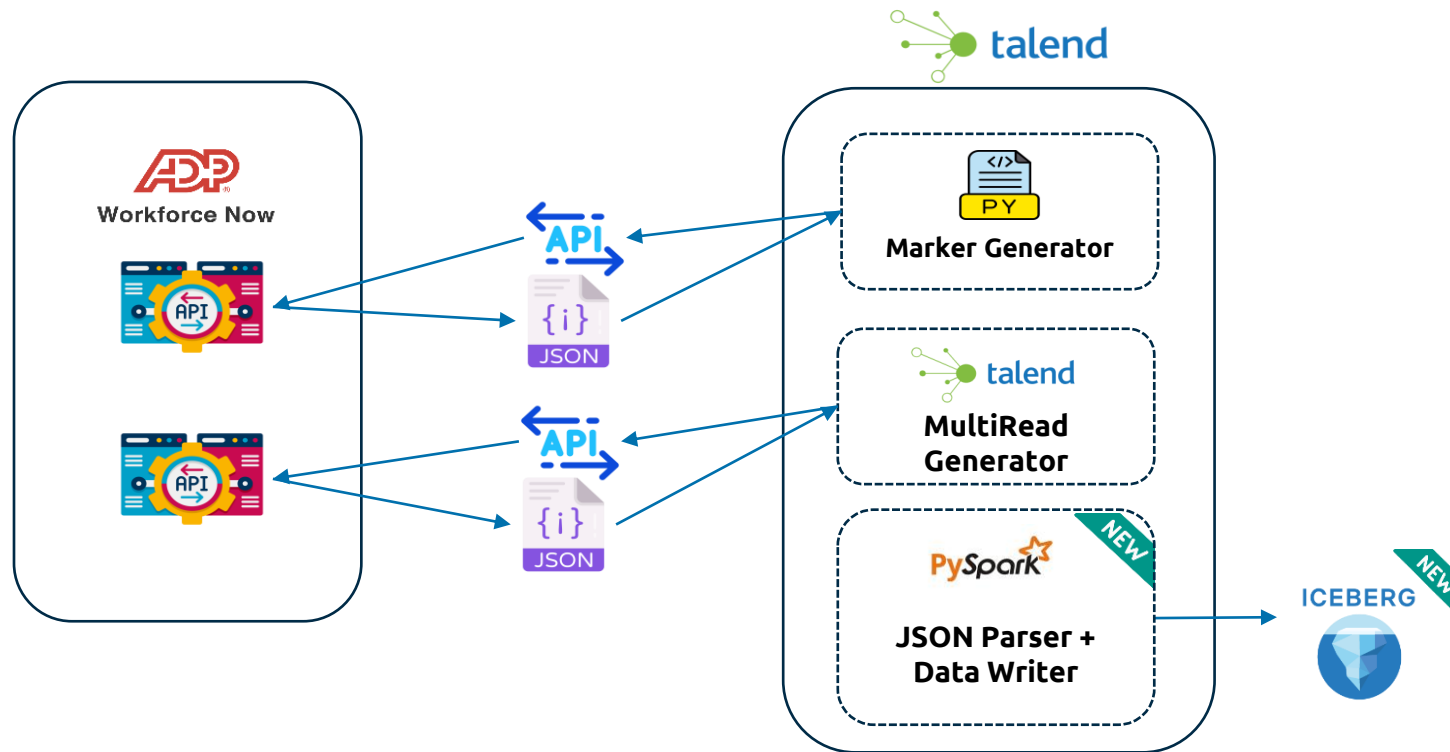


Solution
<ul style="list-style-type: none">Leverage Existing Assets: Reuse Python and Talend codebase to accelerate development and maintain consistency.Modern Execution Framework: Utilize Spark on Serverless architecture for scalable and cost-efficient data processing.Target Storage Strategy: Adopt Snowflake as the preferred data storage solution for performance and scalability

To improve scalability we are proposing two viable options...(2/2)

Redesign Approach # 2 – AWS Iceberg

Reuse existing Python and Talend code, **Leveraging** Spark on Serverless for scalable processing, **Replace** AWS S3 /AWS Athena with AWS Iceberg



Solution

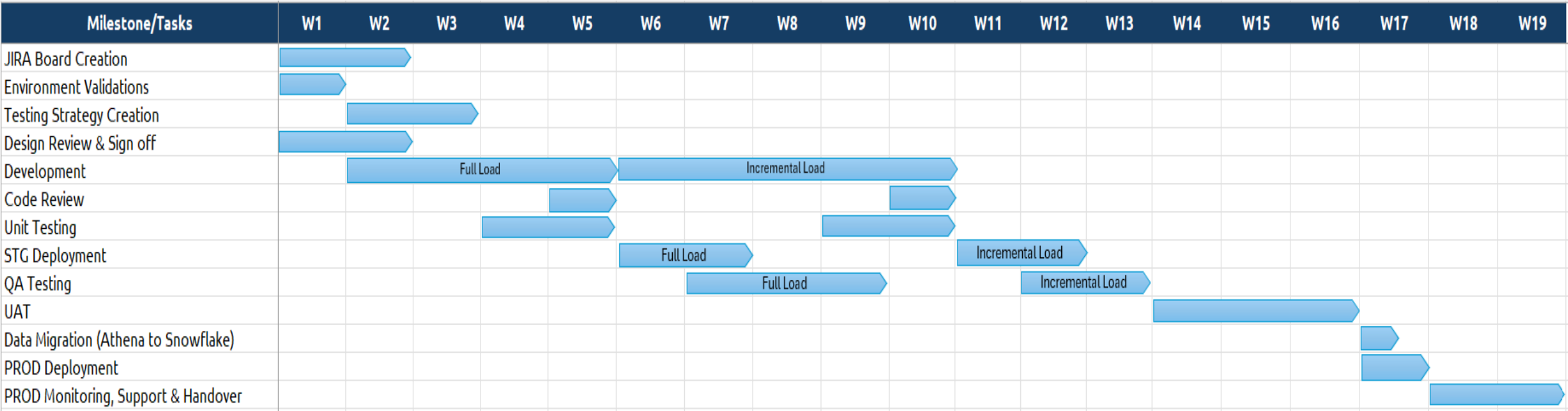
- Reuse the Python and Talend code we have
- Spark on Serverless to write data to the target
- Lakehouse approach can handle CRUD data processing.
- Needs daily maintenance cycles.

COMPARISON OF OPTIONS

	Snowflake	Lakehouse (Iceberg)
Pros	<ul style="list-style-type: none">• Simplified management strategy for data processing.• Provides scaling capabilities to meet sporadic volume increases• Aligns better with overall team skillsets.	<ul style="list-style-type: none">• Improvement over file processing for parquet files.• Leverages more of the existing data processing design.• Scaling can be managed within processing engine (Spark Cluster).
Cons	<ul style="list-style-type: none">• Requires more development over moving to Lakehouse approach to retool supporting reporting needs.• Slight deviation from data lake strategy which maybe mitigated with more discussions and design considerations not yet explored.	<ul style="list-style-type: none">• High-volume events generating small file issue remains unresolved.• Does not resolve continued daily support to processing necessary to deliver reporting.• More management and monitoring is needed to support.• Full read/write supported only by using Spark.
Our Recommendation & reasons	Snowflake: <ul style="list-style-type: none">• Offers excellent performance for analytical queries due to its unique architecture that separates storage and compute.• Elasticity allows for cost optimization by scaling resources up and down as needed.• Provides robust data sharing capabilities and strong security features, making it ideal for enterprise-level analytics.	

PROPOSED EXECUTION PLAN

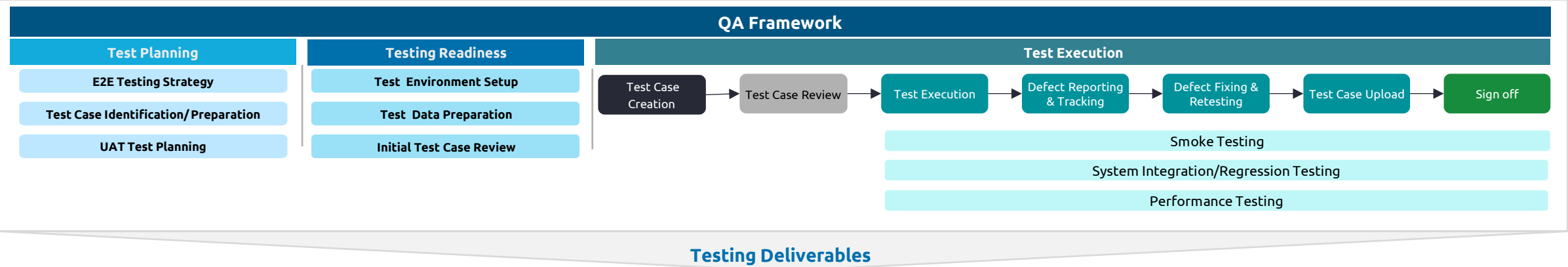
Capgemini recommends approach #1 and the implementation timeline below



Milestones	Responsibility	By Week
Environment Availability	FirmX Infra	1
ADP Clarifications	FirmX Business Team	1
Design Review & Sign-off	FirmX IT Team	2
Test Cases Signoff	FirmX Business Team	3
Test data for QA and Performance Testing	FirmX Business Team	6
Development & Code Review	Capgemini Project Team	10
Unit Testing Signoff	Capgemini Project Team	10

Milestones	Responsibility	By Week
Stage Deployment	Capgemini Project Team	12
QA Testing & Signoff	Capgemini Project Team	13
UAT Testing Signoff	FirmX Business Team	16
Data migration (Athena To Snowflake)	Capgemini Project Team	17
Data migration Signoff (Athena To Snowflake)	FirmX Business Team	17
Production Deployment	FirmX IT Team	17





E2E Test Strategy Document, Test Case, Performance Scripts, Test Execution Reports, Test Case Jira Upload

Testing Phases	Functional/System Testing	Integration Testing	Performance Testing	Regression Testing
Test Scenarios	<ul style="list-style-type: none">Match record counts (source vs. target)Exception HandlingValidate data types, formats	<ul style="list-style-type: none">End-to-End data flow validation across layersError Handling: Test error propagation, execution termination and graceful recoveryMatch record counts (source vs. target).	<ul style="list-style-type: none">System Load: Stress test with max load +25% and concurrent jobs.Execution Timing within SLA limitsDashboard refresh rate for adequately populated data within SLA limits	<ul style="list-style-type: none">Failure RecoveryRestart ValidationScheduled RUN Tests

ASSUMPTIONS

Category	Description
ADP related Responses	ADP specific questions should be clarified before Design completion
Test Data for QA Testing Cycles	Accurate and adequate test data should be made available to project team for all valid business scenarios
Test Data for Performance / Regression Testing Cycles	Performance testing will be carried on peak volume plus 25% of test data.
UAT Testing	UAT of 3 weeks is planned before PROD deployment
Production Support	Post-Production Support of 2 weeks is planned after PROD deployment

EXPECTATIONS FROM FIRMX / ADP

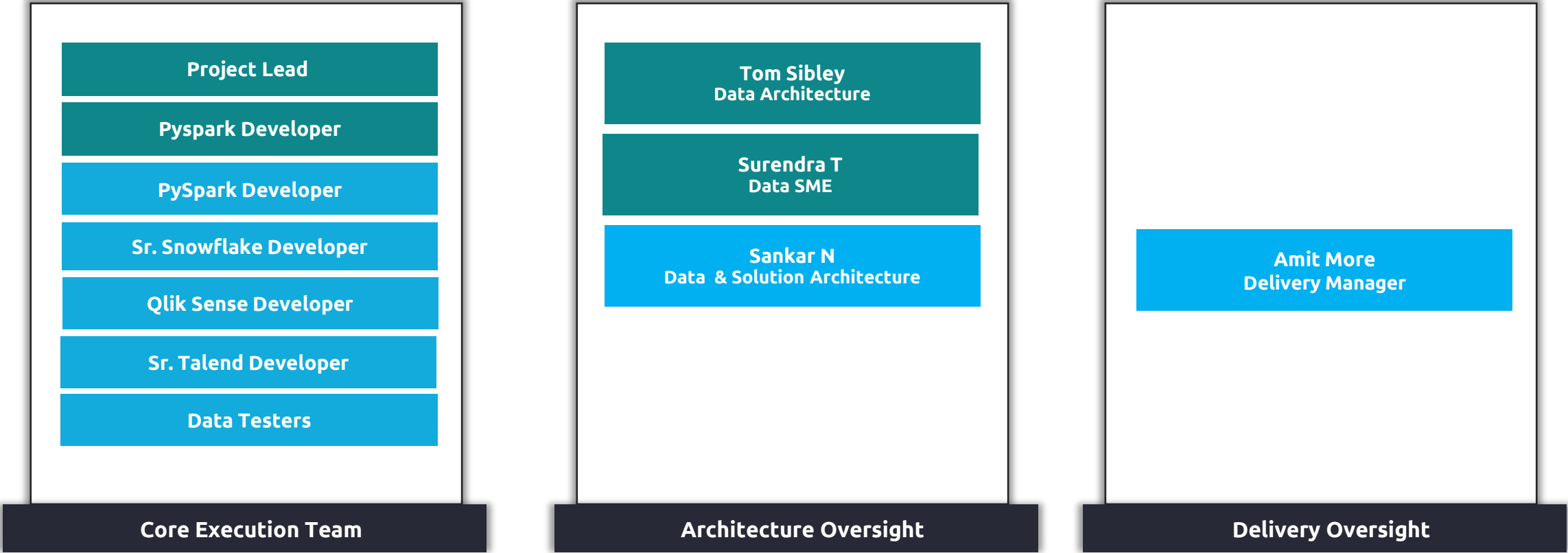
Task / Description	Who	By When?
Does ADP support file-based data transfer (SFTP)?	ADP	??
Can ADP provide Change Markers for all Tables to reduce the amount of data that flows out? <i>"AccuralTransactions" doesn't have change markers and we must download 21GB of data daily</i>	ADP	??
Can we increase the rate limit on the calls from FirmX?	FirmX	??



PROJECT TEAM

**| GET THE FUTURE
YOU WANT**

OUR PROPOSED TEAM STRUCTURE



Offshore Onshore



COMMERCIALS

| GET THE FUTURE
YOU WANT

INFRASTRUCTURE COST PROJECTIONS

Per Month Cost Estimates	Option 1 – (Recommended) Talend + Snowflake	Option 2 – Talend + Iceberg
Snowflake	\$2000	
EMR Serverless	Same as today	Same as today
Iceberg		\$20
S3 Table		\$15
TOTAL	~\$2000	~\$20

** Snowflake cost includes ad-hoc queries & dashboard queries as well.*

Options	Resourcing	Timelines
Option 1: Snowflake	Developers: 1 Snowflake, 2 PySpark , 1 Talend, 2 Testers	16 weeks to PROD
Option 2: AWS Iceberg	Developers: 1 PySpark , 1 Talend, 1 Big Data, 2 Testers	14 weeks to PROD



This presentation contains information that may be privileged or confidential and is the property of the Capgemini Group.

Copyright © 2021 Capgemini. All rights reserved.

About Capgemini

Capgemini is a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided everyday by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of 270,000 team members in nearly 50 countries. With its strong 50 year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fuelled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering and platforms. The Group reported in 2020 global revenues of \$19.3 billion.

Get the Future You Want | www.capgemini.com

This message is intended only for the person to whom it is addressed. If you are not the intended recipient, you are not authorized to read, print, retain, copy, disseminate, distribute, or use this message or any part thereof. If you receive this message in error, please notify the sender immediately and delete all copies of this message.