

# DATA PIPELINE ARCHITECTURE DOCUMENT

**Summit Digital Solutions, Inc.**

**Document Version: 2.0**

**Last Updated: January 9, 2024**

**Classification: Confidential**

## 1. INTRODUCTION

1. This Data Pipeline Architecture Document ("Architecture Document") sets forth the technical and operational specifications of Summit Digital Solutions, Inc.'s ("Company") proprietary data pipeline infrastructure that supports the Peak Performance Platform(TM) and related digital transformation solutions.

2. This document is considered Confidential Information as defined in the Company's Master Services Agreement and is subject to all applicable confidentiality obligations and restrictions.

## 2. DEFINITIONS

1. "Data Pipeline" means the Company's end-to-end system for collecting, processing, transforming, and delivering data across the Peak Performance Platform(TM).

2. "Processing Nodes" refers to the distributed computing infrastructure responsible for executing data transformation operations.

3. "Edge Collection Points" means IoT-enabled data gathering devices and interfaces deployed at client locations.

4. "Core Processing Layer" means the central data processing infrastructure hosted in the Company's secure cloud environment.

## 3. ARCHITECTURE OVERVIEW

### 1. System Components

The Data Pipeline consists of four primary layers:

- a) Data Collection Layer
- b) Data Integration Layer

- c) Processing & Analytics Layer
- d) Distribution & Access Layer

## 2. Infrastructure Configuration

- a) Primary Data Center: AWS US-East-1
- b) Backup Data Center: AWS US-West-2
- c) Edge Processing: Distributed kubernetes clusters
- d) Storage: S3-compatible object storage with encryption at rest

## **4. DATA COLLECTION SPECIFICATIONS**

### 1. Edge Collection Architecture

- a) Proprietary IoT sensor network utilizing IEEE 802.15.4 protocol
- b) Real-time data buffering with local storage capacity of 72 hours
- c) Automated failover and recovery mechanisms
- d) End-to-end encryption using AES-256

### 2. Data Ingestion Protocols

- a) REST API endpoints with OAuth 2.0 authentication
- b) MQTT broker network for IoT device communication
- c) Batch processing capabilities for historical data imports
- d) Real-time streaming via Apache Kafka

## **5. PROCESSING & ANALYTICS INFRASTRUCTURE**

### 1. Core Processing Components

- a) Distributed processing framework using Apache Spark
- b) Machine learning pipeline powered by TensorFlow
- c) Real-time analytics engine using Apache Flink
- d) Time-series optimization using InfluxDB

### 2. Scaling Mechanisms

- a) Auto-scaling based on processing load
- b) Resource allocation management

- c) Load balancing across processing nodes
- d) Fault tolerance and recovery procedures

## **6. SECURITY AND COMPLIANCE**

### **1. Data Protection Measures**

- a) Encryption in transit and at rest
- b) Role-based access control (RBAC)
- c) Multi-factor authentication
- d) Regular security audits and penetration testing

### **2. Compliance Framework**

- a) SOC 2 Type II certified
- b) ISO 27001 compliant
- c) GDPR and CCPA compliant
- d) HIPAA compliance capabilities where required

## **7. PERFORMANCE SPECIFICATIONS**

### **1. System Performance Metrics**

- a) Maximum throughput: 1M events per second
- b) Average latency: <50ms
- c) Data processing SLA: 99.99% uptime
- d) Recovery Time Objective (RTO): 4 hours

### **2. Monitoring and Alerting**

- a) Real-time performance monitoring
- b) Automated alerting system
- c) Historical performance analytics
- d) Capacity planning tools

## **8. PROPRIETARY RIGHTS**

- 1. All architectural designs, implementations, and related documentation described herein are the exclusive intellectual property of Summit Digital Solutions, Inc.

2. This document may not be reproduced, distributed, or modified without express written consent from the Company.

## **9. DOCUMENT CONTROL**

1. This document shall be reviewed and updated annually or upon significant architectural changes.

2. Change Management

a) All modifications require CTO approval

b) Version control maintained in internal documentation system

c) Change log maintained for audit purposes

## **10. EXECUTION**

IN WITNESS WHEREOF, this Data Pipeline Architecture Document has been executed by the duly authorized representative of Summit Digital Solutions, Inc.

SUMMIT DIGITAL SOLUTIONS, INC.

**By:**

Name: Michael Chang

Title: Chief Technology Officer

Date: January 9, 2024

APPROVED BY:

Dr. Alexandra Reeves

Chief Executive Officer

James Henderson

Chief Digital Officer