FUSE SUITE

OVERVIEW

DATA PLATFORM

WWW.PROMES.CO.TH

WE ARE
TECHNOLOGY SOLUTION
FOR ENTERPRISE

FUSE SUITE

We're a cutting-edge research organization focused on advancing artificial intelligence, machine learning and Data analytic in a way that benefits society as a whole. Our mission is to develop and promote friendly AI, NLP and Data Insight that is safe and beneficial for Thailand and the world.

WHY CHOOSE FUSE SUIT DATA PLATFORM ?

Streamlined Integration

Our platform seamlessly integrates a range of open-source tools and technologies, making it easier than ever to manage and process data within the Hadoop ecosystem.

Scalability

Whether you're a startup or an enterprise, our platform scales to meet your evolving data needs, ensuring optimal performance and efficiency at any scale.

Flexibility

Tailor the platform to suit your specific requirements and workflows, with customizable configurations and deployment options to accommodate diverse use cases.

BENEFITS OF FUSE SUIT : DATA PLATFORM.

Enhanced Performance

Reliability and Security

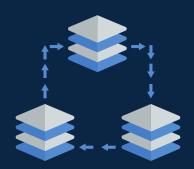### Cost - Effectiveness

Ease of Use

KEY FEATURES

Scalable Storage and Processing

HDFS: Reliable, scalable, and distributed storage system. Parallel Processing: Powerful framework for processing vast amounts of data efficiently.

Comprehensive Data Integration

Data Transfer Tools: Efficiently move bulk data between various systems and Hadoop. Stream Data Ingestion: Seamlessly ingest data from multiple sources into HDFS.

Real-time Data Analytics

Cluster Computing: Fast and general-purpose system for real-time data analytics. Data Warehousing Solutions: Facilitate easy data summarization, querying, and analysis.

Enhanced Data Security

Secure Authentication: Ensures data privacy and integrity through robust authentication methods. Centralized Security Management: Framework for data authorization and auditing.

Advanced Data Management

Workflow Scheduling: Manage and schedule complex data processing jobs. Coordination Services: Efficiently manage distributed applications.

DATA PLATFORM FEATURES

1. Advanced Distributed Architecture

- **Scalable Design :** Developed with a distributed architecture to ensure high availability and scalability.
- **Fault Tolerance :** Built-in mechanisms for data replication and fault tolerance to prevent data loss.

2. Centralized Management Interface

- **Unified Control :** Manage all aspects of the platform through a centralized, web-based interface. Cluster
- **Management :** Efficiently manage and monitor cluster performance and health.

3. Hadoop Distributed File System (HDFS)

- **Reliable Storage :** Use HDFS for reliable, scalable, and distributed storage of large datasets. Data
- **Redundancy :** Ensure data redundancy and quick recovery from failures.

4. Comprehensive Data Handling

- **Structured and Unstructured Data:** Store and manage both structured and unstructured data seamlessly.
- **Query Support :** Perform complex queries using SQL-like languages (Hive) and NoSQL databases (HBase).

5. Powerful Querying and Analysis

- **Hive Query Language (HQL) :** Utilize HQL for querying data stored in Hive without additional costs.
- **Advanced Analytics :** Support for distributed query engines for efficient data analysis.

6. Standard Connectivity Options

- **ODBC and JDBC :** Support for ODBC and JDBC standards for seamless database connectivity.
- **API Integration :** Easily integrate with other systems and applications using standard APIs.

7. Extensive Language Support

- **Programming Languages :** Support for R, Python, and Spark ML for advanced data analysis and machine learning.
- **Custom Scripting :** Enable custom scripting and automation to meet specific business needs.

8. Notification and Alerting System

- **Real-time Alerts :** Receive real-time alerts via email, SNMP, or HTTP based on predefined conditions.
- **Custom Notifications :** Configure custom notification channels as per organizational requirements.

9. Configuration and Cluster Management

- **System Configuration :** Manage system settings and configurations through a user-friendly interface.
- **Resource Allocation :** Efficiently allocate and monitor resources such as CPU and memory.

DATA PLATFORM FEATURES

10. Activity Monitoring and Reporting

- **Job Monitoring :** Track the status of jobs and system activities, with real-time updates.
- **Reporting Tools :** Generate comprehensive reports on system usage and performance.

11. Resource Management

- **Dynamic Allocation :** Automatically allocate resources based on workload demands.
- **Performance Monitoring :** Monitor resource utilization and optimize system performance.

12. Workflow Management

- **Job Scheduling :** Schedule and manage complex workflows with a built-in job scheduler.
- **Workflow Automation :** Automate repetitive tasks to improve efficiency and accuracy.

13. Security and Compliance

- **User Authentication :** Ensure secure access with robust user authentication methods.
- **Access Control :** Define user roles and permissions to control access to sensitive data.
- **Data Encryption :** Encrypt data in transit and at rest to safeguard against unauthorized access.

14. Performance and Priority Management

- **Workload Management :** Balance and prioritize workloads to ensure efficient processing.
- **Resource Optimization :** Optimize resource usage to maximize performance and minimize costs.

15. Audit and Logging

- **Activity Logs :** Maintain detailed logs of user activities and system operations.
- **Audit Trails :** Create audit trails to ensure compliance with regulatory requirements.

16. Session Management

- **Session Timeout :** Configure session timeouts to enhance security and resource management.
- **User Session Tracking :** Track active user sessions and manage system access.

17. Ecosystem Compatibility

- **Integration :** Seamlessly integrate with other distributed architecture ecosystems.
- **Scalability :** Support for expanding the ecosystem with additional software without disruption.
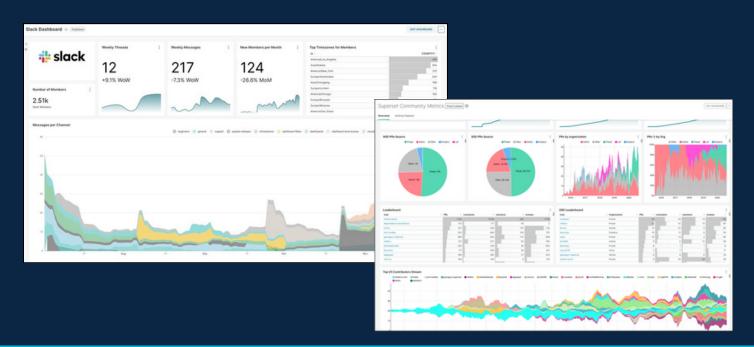
18. Extensibility

- **Modular Design :** Add new features and functionalities as needed without affecting existing operations.
- **Third-party Tools :** Support for integrating third-party tools and applications to enhance platform capabilities.

FUSE SUITE : # TECHNICAL SPECIFICATION

Data Visualization

Services

Hadoop Distributed File System (HDFS)

- Store and manage large volumes of data efficiently with this distributed file system.

Apache Hive

- Query and analyze data using SQL-like syntax, simplifying data exploration and analysis tasks.

Apache YARN

- Efficiently manage resources across distributed applications, ensuring optimal performance and resource utilization.

Apache Spark

- Process large-scale data sets rapidly and efficiently with this fast and versatile data processing engine.

Apache Kafka

- Facilitate real-time data streaming and processing for mission-critical applications.

Apache Airflow

- Orchestrate complex data workflows with ease, automating data pipelines and scheduling tasks efficiently.

Apache Ambari

- Simplify cluster management and monitoring with this intuitive web-based tool.

Trino

- Execute interactive SQL queries across multiple data sources in real-time, enabling fast and efficient data analysis.

BUSINESS PROCESSES

Data Sources

- Customer Data
- Sales Data
- Marketing Data

STEP 01

Data Collection and Integration

- Batch Collection
- Real-time Data
- Collection

STEP 02

Centralized Data Repository

- Structured Data
- Unstructured Data

STEP 03

Centralized Data Analysis

- Data
- Transformation
- Data Cleansing
- Data Enrichment
- Business Intelligence
- Analytics Machine Learning

STEP 04

Business Insights and Decision Support

- Dashboards Strategic
- Planning Operational
- Reports Performance
- Management Advanced
- Analytics Customer
- Engagement

STEP 05