NETWORK TRAFFIC REPORTING

TEAM MEMBERS: Nikita Sirandasu, Pranav Chinthakuntla, Yatharth Jain, Maggie Gao, Carlos Reyes, Derek Peng, Lakshay Mahajan, Ashwin Tarinikanti, Manas Jha, Riley Noggle, Adrian Maliackel

OVERVIEW

About Caterpillar:
- Global industry leader in construction and mining equipment, off-highway engines, turbines, & locomotives

Project Objective:
- Our goal is to create a user-friendly and accessible enterprise traffic reporting system to replace Cat's legacy system. This system has many benefits, such as:
  - Providing **insights** and **global data analysis**, which will help efficiently **identify trends** and **anomalies**
  - Helping support the IT department by **improving efficiency** and **effectiveness** of network monitoring and engagement

PLATFORM TEAM

- **Objective: Develop a pipeline** system for collecting, processing, and storing network traffic data with:
  - GoFlow2 – network traffic data **collection**
  - Kafka / Confluent - **processing and batching**
    - Kafka – buffer for data streams
    - Confluent – uses JDBC Connector to integrate Kafka to database
  - PostgreSQL DB with Timescale extension for **storage**
- Set up logging system and developed PostgreSQL database schema for maintenance

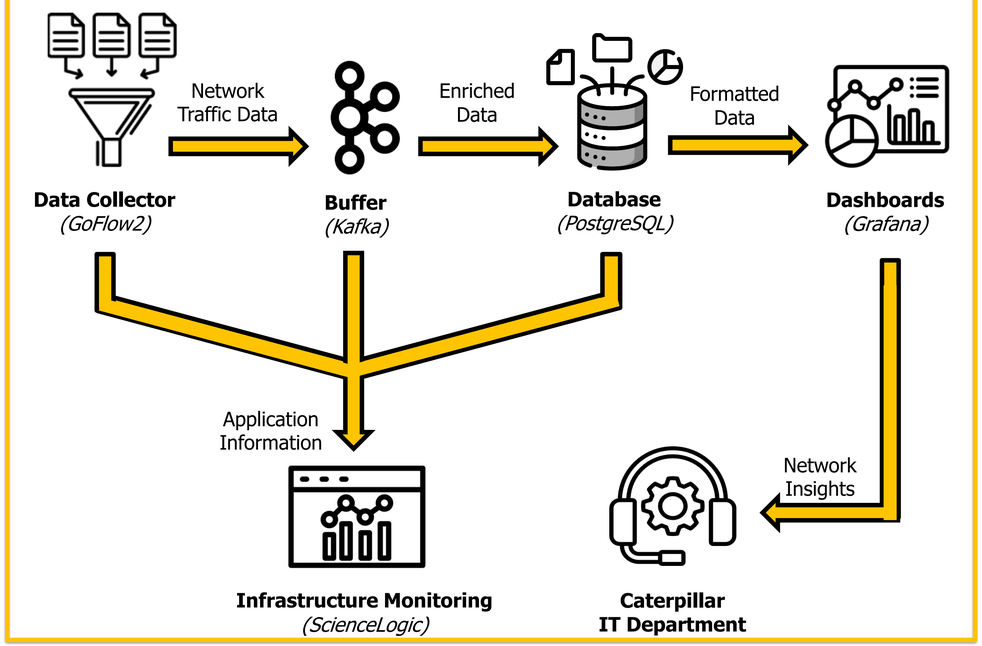
Network Pipeline

REFERENCES

- Apache Kafka Documentation
- GoFlow2 Documentation
- How the OSI Model Works | Network Fundamentals
- Timescale Documentation

METHODOLOGY

Network Traffic and Application Insights



ANALYTICS TEAM

Objective: Monitoring, cleaning, and analyzing
- Investigated multiple methodologies across our various databases to extract **variable insights**
- Used **ScienceLogic**—a monitoring platform—to integrate our database and monitor database processes in real-time
- Created Python scripts to **analyze and clean ScienceLogic data** for improved database analysis

USER INTERFACE TEAM

Objective: User-friendly, intuitive, accessible interfaces
- Dashboards made with **Grafana**–a data visualization platform—with functions such as:
  - Displaying **network activity trends** during business and non-business hours globally
  - Monitoring **device** and **process memory** usage
  - Displaying **peak utilization patterns** during the day

Peak Utilization Points Dashboard

CONCLUSION & FUTURE PLANS

- **Platform Team:** Built a **network data pipeline** and added **database monitoring** for efficiency.
- **Analytics Team:** Developed **visualizations** and Python **scripts** for ScienceLogic data analysis.
- **UI Team:** Created **dashboards** and improved network traffic queries.

- **Future Plans**: These efforts have strengthened CAT's data infrastructure, **improving performance** and **analytical capabilities.** CAT plans to use our system as a supplemental tool to **diagnose** and **troubleshoot issues** and **enhance logging analytics** next year

ACKNOWLEDGEMENTS

- We would like to thank our CAT Mentors Mike Osborne, Mo Alwaheidi, Nathan Wilson, and Rohan Mhatre for their invaluable support and guidance. We would also like to thank Tom Linn, our Scrum Master
- An additional thank you to the Data Mine staff for all the support and feedback

Caterpillar: Confidential Green