MICHIGAN STATE UNIVERSITY

Project Plan

Force Platform Ingestion Tool

The Capstone Experience

Team Rook

Roy Barnes
Matt Hammerly
Will McGee
Chiyu Song
Mark Velez

Department of Computer Science and Engineering
Michigan State University

Spring 2017

From Students…
…to Professionals

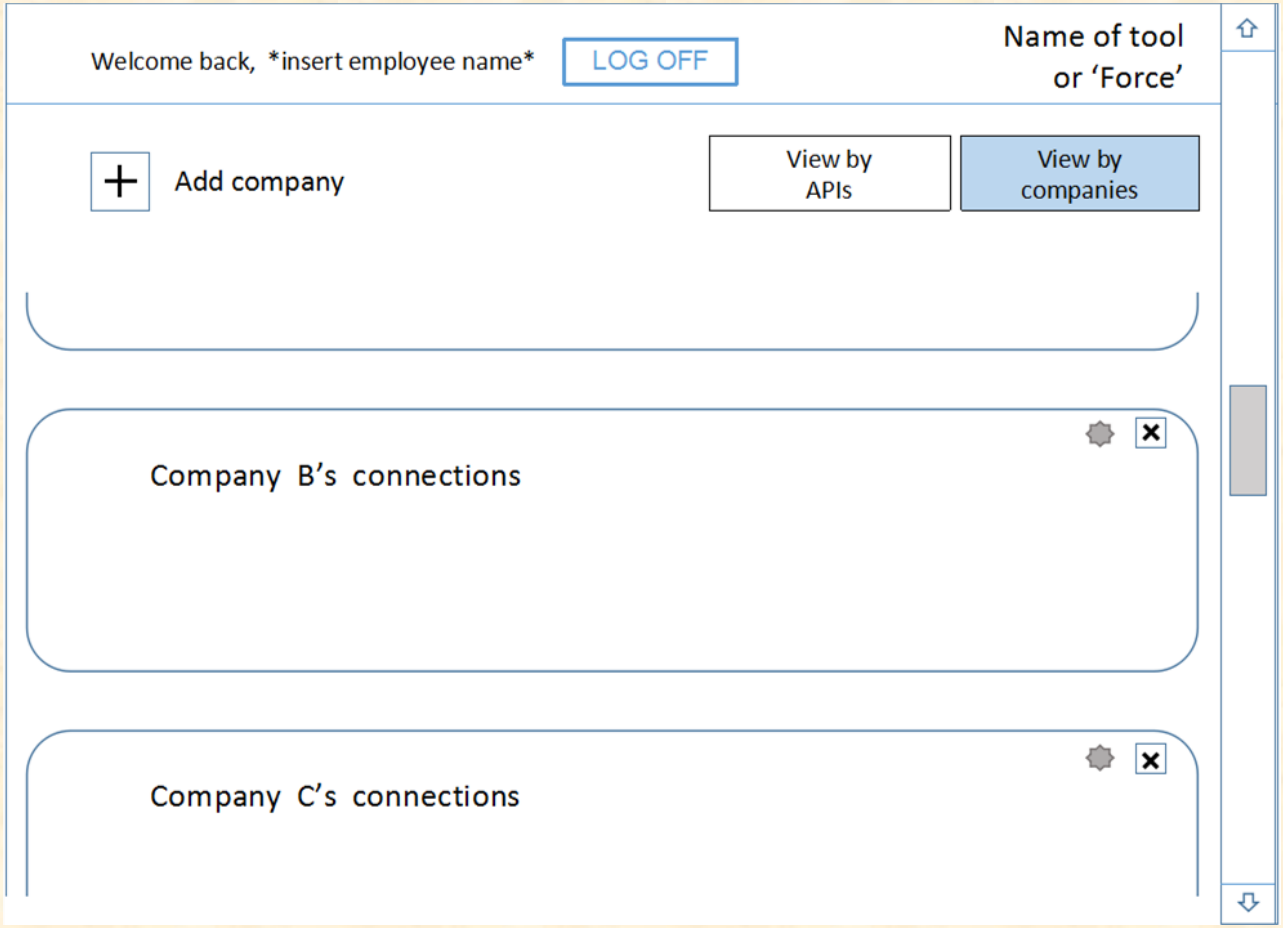
Functional Specifications

- Force platform for security alert management/analysis
- Force accepts data in one format, but clients send data in different formats
- Force PIT provides an easy way for clients to integrate existing monitoring tools with Force
- Promotes outcome-focused mission by allowing analysts to see related alerts

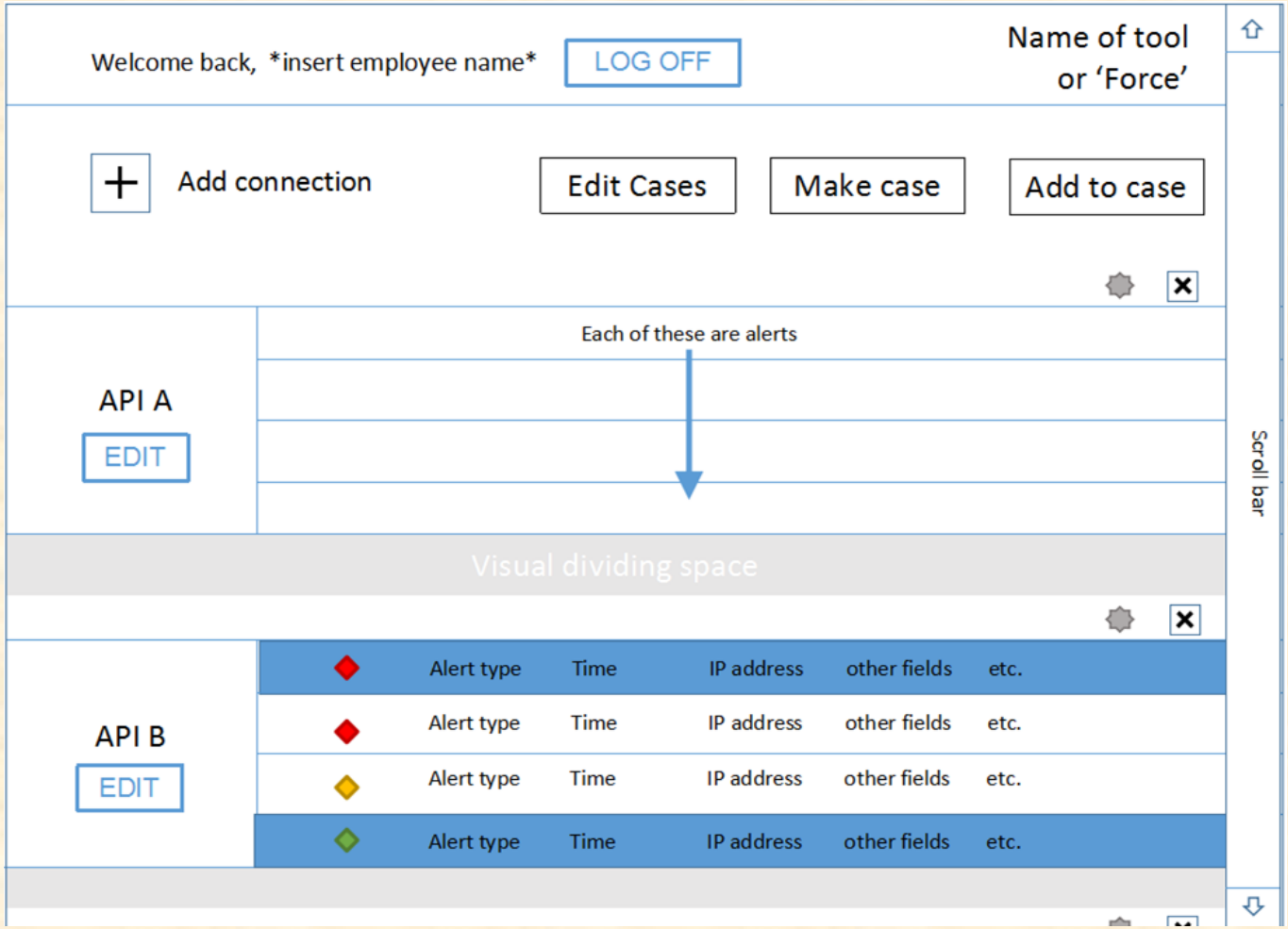
Design Specifications

- Mirrors Force's outcome-focused design
- Filters context options for viewing alerts
- Lists alerts by alert severity
- Allows grouping of alerts into suggested cases

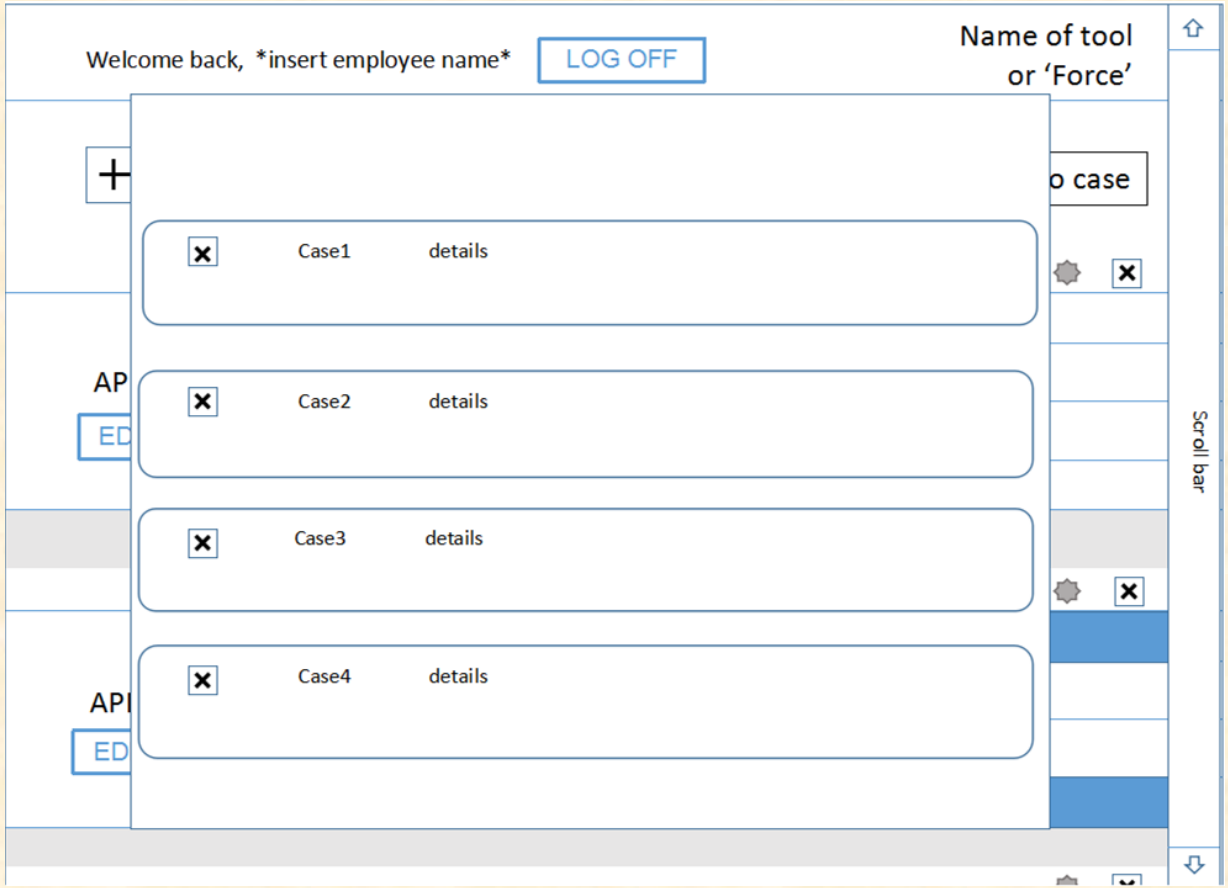
Screen Mockup: Connection Page



Screen Mockup: Alert Page



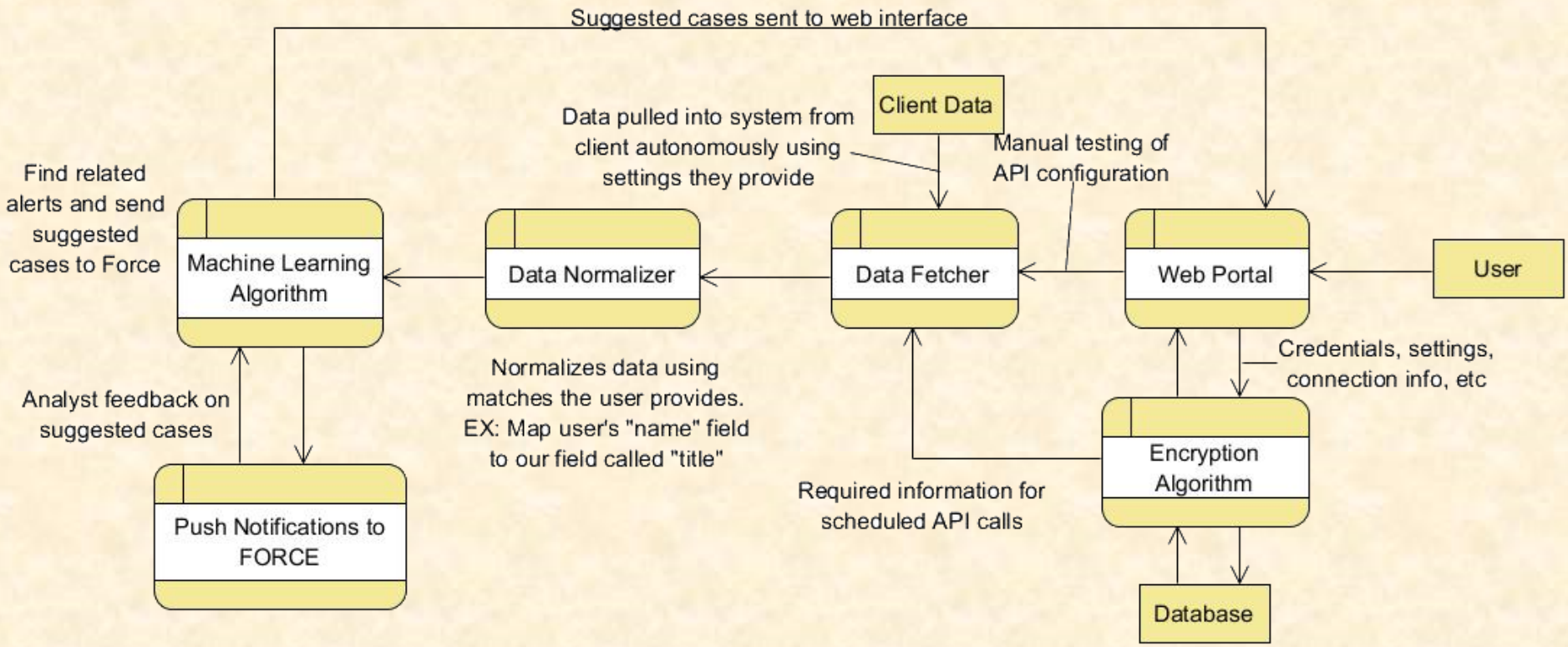
Screen Mockup: Case Page



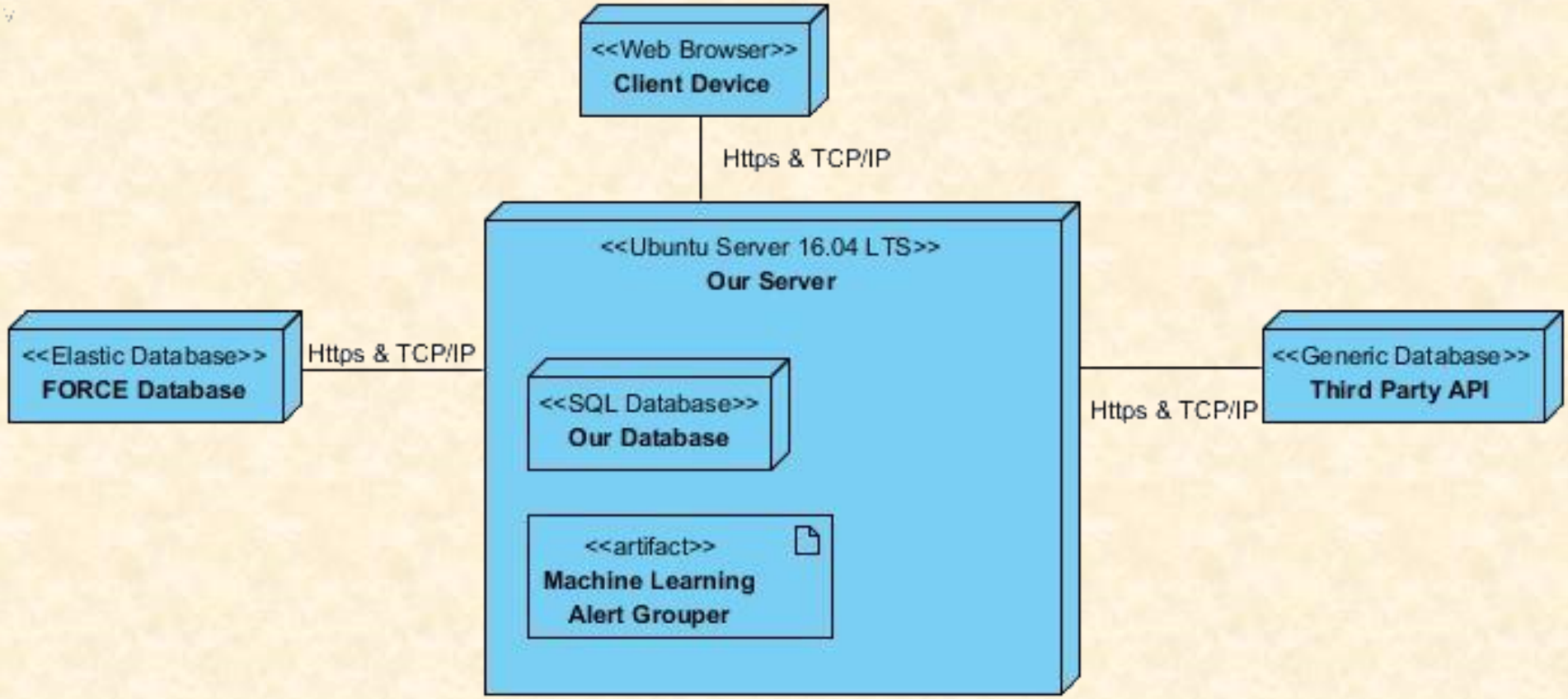
Technical Specifications

- Web Interface
  - Users configure new API connections
  - Analysts view machine learning suggested cases
- Data fetcher
  - Periodically polls each configured API connection
  - Normalizes API output and sends Force database
- Machine Learning component
  - Suggests groups of possibly related alerts
  - Analysts confirm relation to further train the model

System Architecture



System Architecture



System Components

- Hardware Platforms
  - Capstone Lab Server
  - Existing Rook Infrastructure
- Software Platforms / Technologies
  - Web application server
    - Ubuntu 16.04 LTS, nginx, uwsgi, Python (Django)
  - Data storage/retrieval
    - MySQL, Elasticsearch, DynamoDB
  - Development tools
    - Git, MyCLI, Visual Studio, PyCharm, Vim

Testing

- Compare machine learning algorithm against current statistical analysis
  - Track number of suggested cases validated by analysts
- Utilize Development/Master Branches
  - Pull requests must pass unit tests and review
- Code review with area partner prior to merge
  - Each area has two experts

Risks

- R1: Data Normalization
  - Various input data types -> Unified JSON format
  - Only certain APIs and template formats will be supported
- R2: Unsupervised Machine Learning Algorithm
  - Algorithm must improve based off of analyst feedback
  - Research unsupervised learning and utilize Rook contact
- R3: Web Portal UI
  - Front-end skills are required for a satisfactory result
  - Best practice research and client feedback

Questions?