Load Research Modernization at Duke Energy

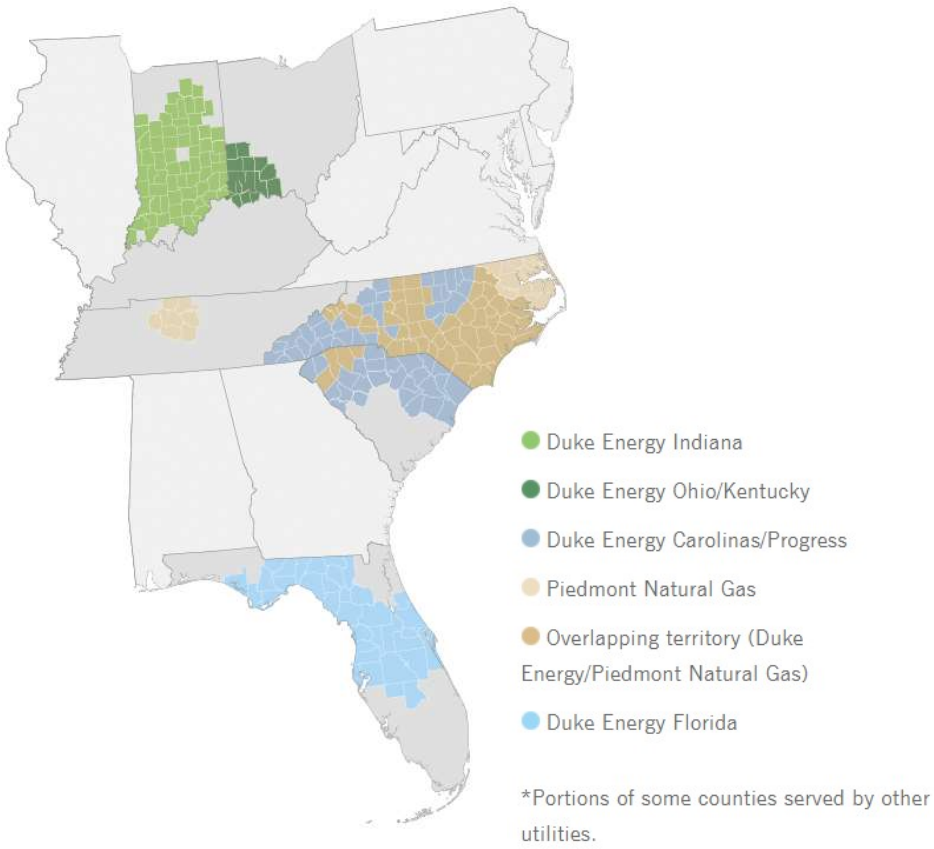
2021 AEIC and WLRA Annual Load Research and Analytics
Web Conference

October 28, 2021

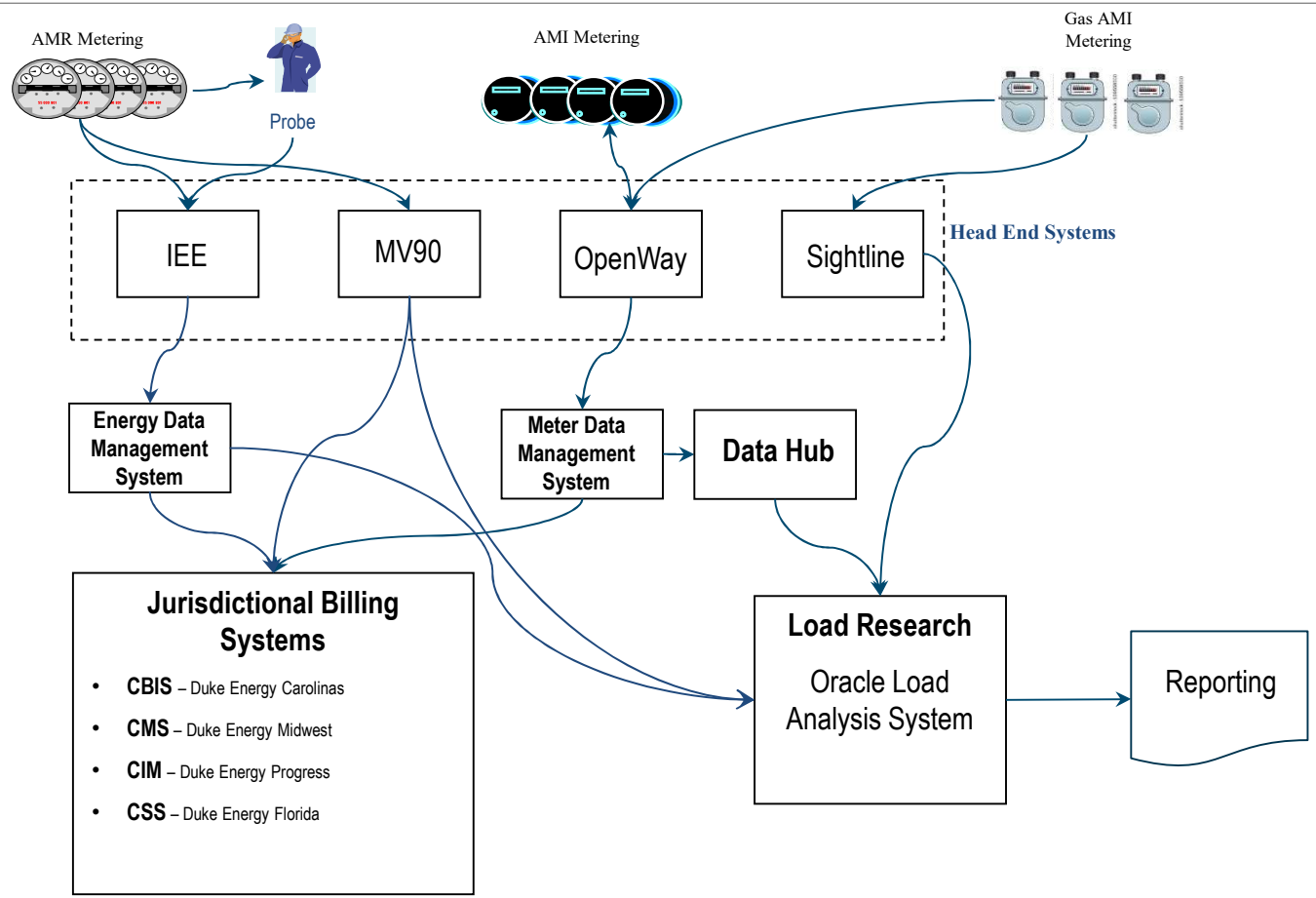
William Baker – Manager Load Research
Arnie Richardson – Dir Strategic Solutions Analytics & Innovation

Duke Energy Load Research Overview

- Performs Load Research for six electric jurisdictions and Midwest Gas
- ~ 7.8 million electric customers & 1.6 million natural gas customers across all jurisdictions
  - Electric customers produce over half a billion rows of new interval data every day
- Load Research samples use ~17,000 meters (0.2%)
  - 90% confidence that the true value is within +- 10% of our estimate (90/10)
- Primary Customers:
  - Rate department for Cost of Service purposes
  - Load forecasting for modeling
  - Rate design for various projects/pilots



Data Flow


AMR Metering
Probe
AMI Metering
Gas AMI Metering
IEE
MV90
OpenWay
Sightline
Head End Systems
Energy Data Management System
Meter Data Management System
Data Hub
Jurisdictional Billing Systems
CBIS – Duke Energy Carolinas
CMS – Duke Energy Midwest
CIM – Duke Energy Progress
CSS – Duke Energy Florida
Load Research
Oracle Load Analysis System
Reporting

Building the Foundation to Enable Success

Foundational Investments

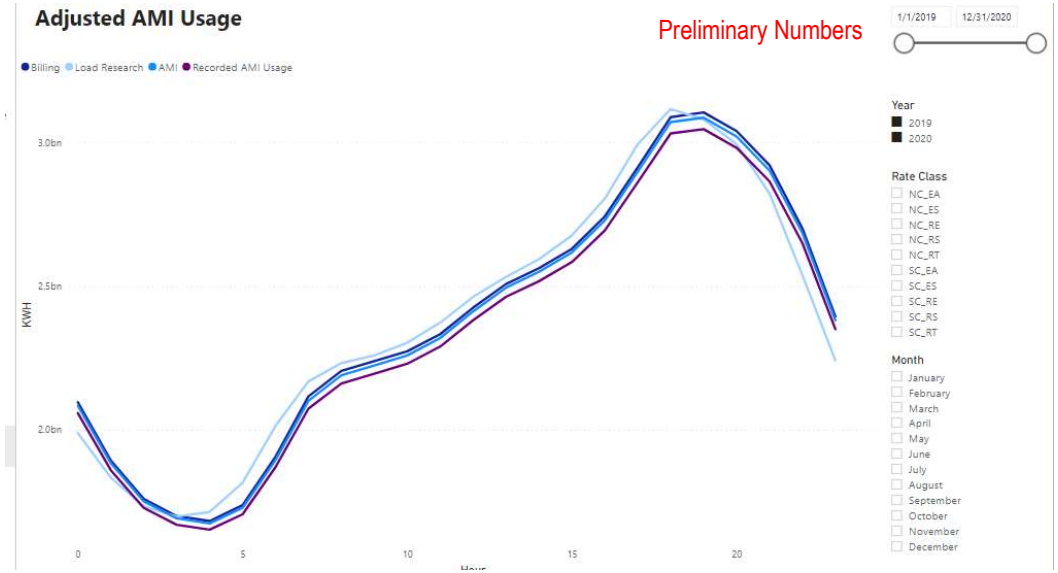
- AMI Rollout completed in all jurisdictions
- Multiple analytics efforts needing clean and efficient interval data for analytics
- Invested in Big Data Platforms (Hadoop & AWS)
- Organizational strategy to combine Load Research with Advance Analytics

Value of Near Population Analysis

- Increase accuracy, automation, and time to deliver
- Additional flexibility in customer segmentations for future use cases
- Dynamic reporting in PowerBI
- Commission requirement for next rate case

Testing Near Population Load Research

- Started by just looking at residential rates in our Carolina jurisdictions
- Calculated load shapes 4 ways:
  - Historic Load Research Samples
  - Raw AMI data
  - Aggregated AMI data scaled to billing data
  - Aggregated AMI data scaled using a “mean per customer” approach for missing data



Key Findings

Data Quality

- We received 99.44% of all AMI data, with only 3 days with data below 95% complete (identified as an issue with the data pipeline)
- Issues with DST and Hour Ending vs Hour Beginning
- We could have issues with C&I customers, especially totalized accounts

Accuracy

- Monthly kWh showed some differences due to billing cycles
- Overall, near population load shapes were 'smoother' than traditional LR
- Summer and winter peaks had differences ranging from 1.95% to 6.62%

Speed

- Analysis was able to run and populate a PowerBI with minimal manual intervention
- No sample design/maintenance needed

Conclusion:

Using our near-population interval data will enhance the precision of the analysis as long as we build robust processes for addressing data quality

Future Process

- Load Forecasting
- Rate Design
- Cost of Service
- Solutions Development
- Marketing/Research

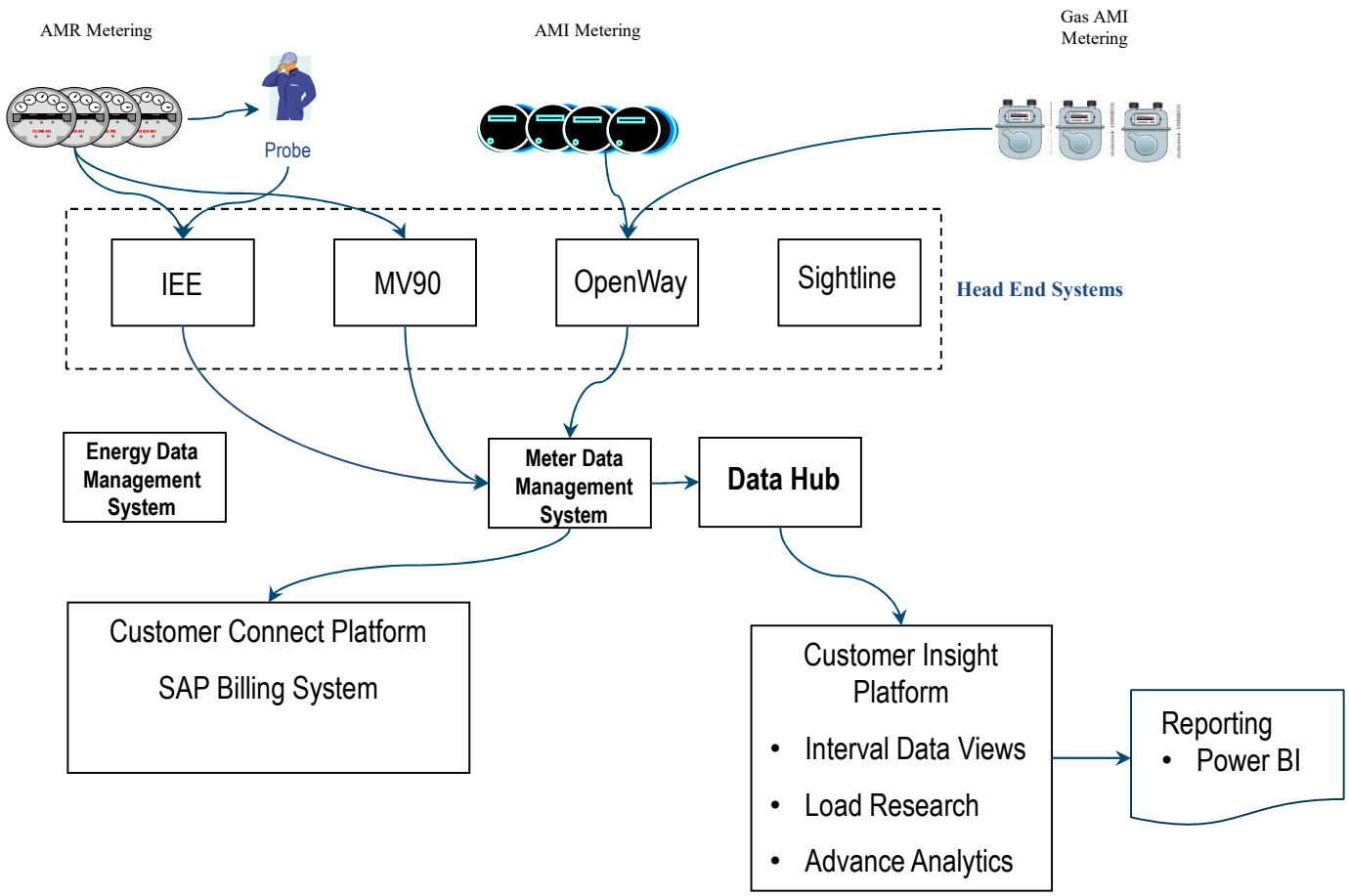
- Subset customers (rates, demographics, housing, usage, etc.)
- Define time period

- Database for Ad Hoc queries
- Automated data quality checks, estimation & reporting

- Core COS analysis runs monthly on an automated schedule
- Includes load shape & statistics

- Internal PowerBI for LR to review, validate & publish
- External self-serving PowerBI for core customers

Modernization Data Flow



Enhancements:

AMI, MV90 & IEE all flows to the MDM

New billing system consolidating 4 historic systems

Robust platform for load research and advance analytics

Building the Digital Customer

Demographics
Income, Race, Age, Education, Number of People in Home

Housing
Dwelling type, Sq Ft, Heating Type, Age of Home

Psychographics
Personas, Propensities, Motivation, Lifestyle

Customer Journeys
Start Service, Bill Payments, Etc.

Customer Offerings
Rates, Program Enrollment, Program Impacts, Event Response

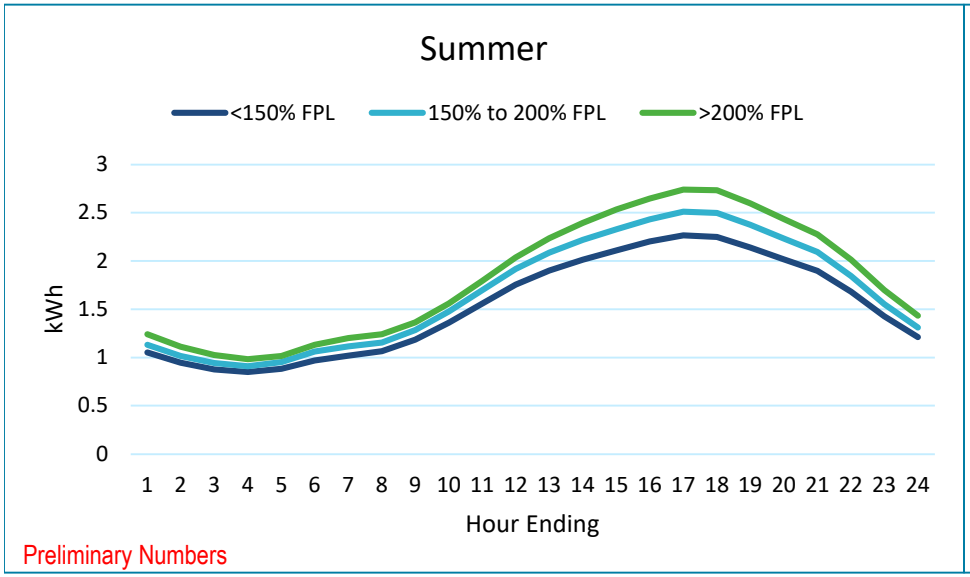
Grid & Reliability
Grid Location, Outages, GT&D Peak Contributions, Reliability

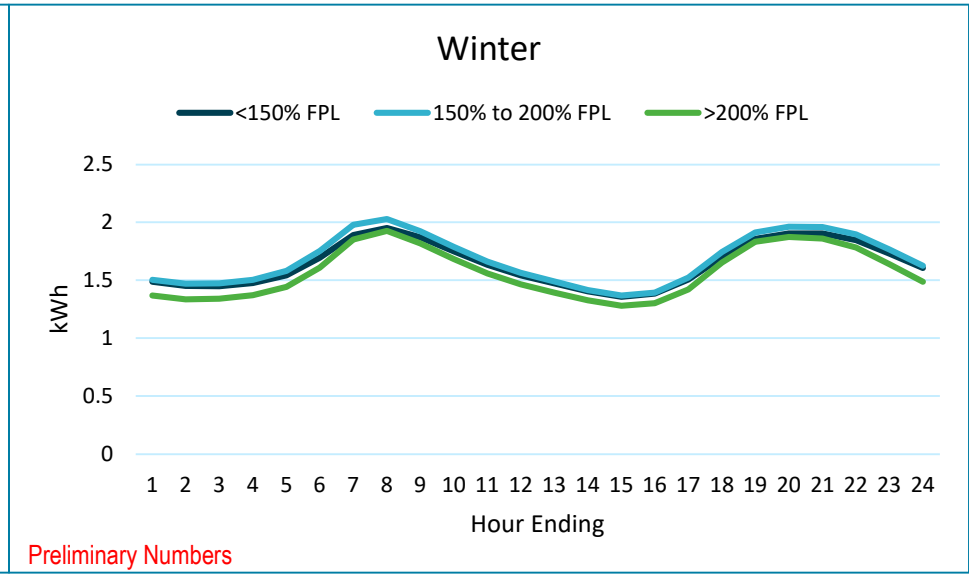
Communication & Feedback
Call Center, NPS, Surveys, Clickstream Data, Social Media

Usage
Load Shapes, Disaggregation, Weather Sensitivities, Charges

- Data Driven Rates & Program Design
- Locational Impacts
- Targeted Marketing

Example: Understanding Low Income Customers





Using third party data to get customer level income and # of people in the household we could segment customers by their Federal Poverty Line and calculate their load shapes to ultimately be data driven around program and rate design for struggling customers