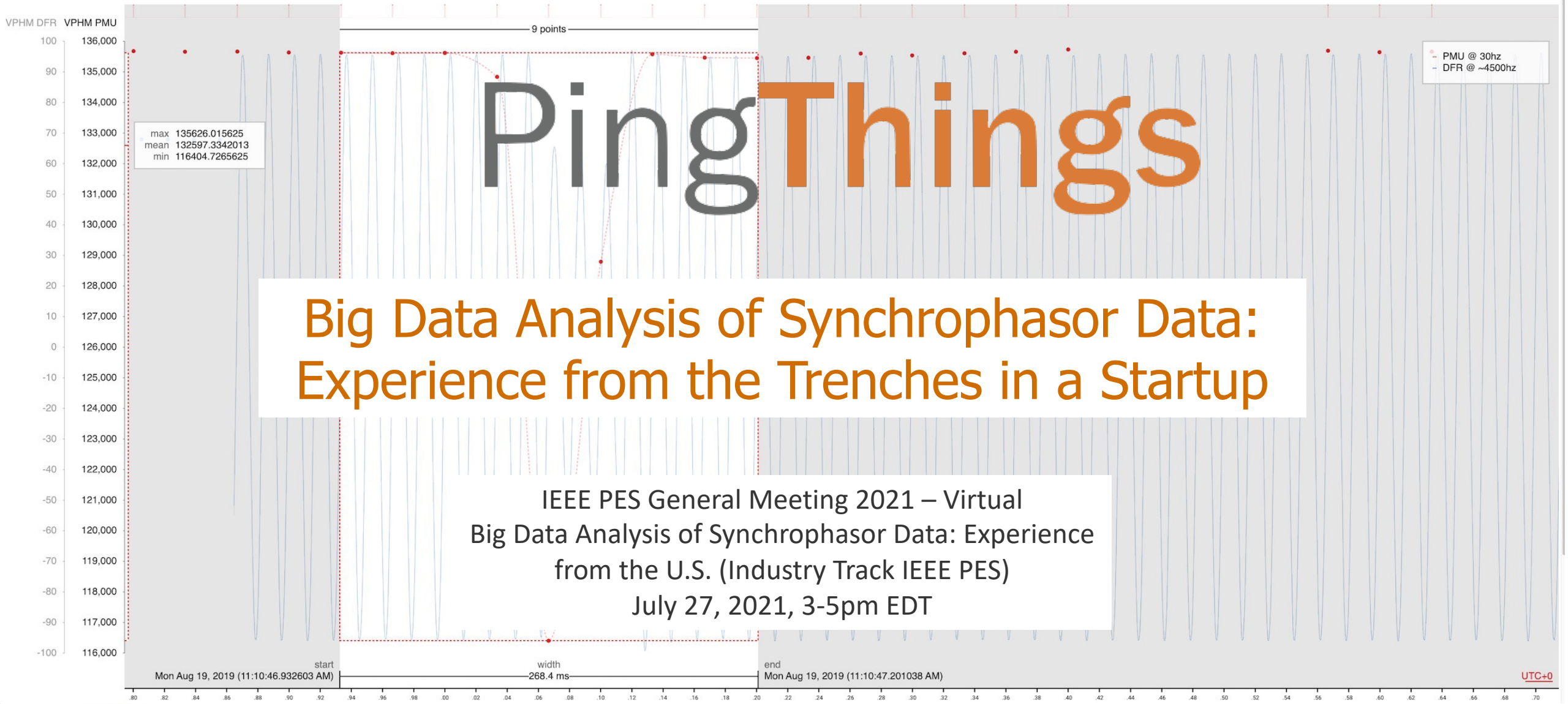


PingThings

Big Data Analysis of Synchrophasor Data: Experience from the Trenches in a Startup

IEEE PES General Meeting 2021 – Virtual
Big Data Analysis of Synchrophasor Data: Experience from the U.S. (Industry Track IEEE PES)
July 27, 2021, 3-5pm EDT

IEEE PES Power & Energy Society®

IEEE

Introduction

Sean Murphy, CEO

PingThings, Inc.

sean@pingthings.io

https://www.linkedin.com/in/seanpatrickmurphy1/

PingThings

Overview

- Go to market in 2017
- 3 full time employees in Jan 2020
- 17 full time employees in Jan 2021
- Entire team is technical, no sales or marketing (yet)
- Fully remote team – MD, VA, DC, CA, WA, and NY
- Expected to hit 30 full time by end of 2021

Financial Background

- Raised ~$1.6M in seed funding since 2017
- Supported by multiple DOE, ARPA-E, and NSF grants totaling over $8M
- Large, multi-year contract with Dominion Energy
- Customers include transmission, distribution, telecom, and transportation

The Competitive Field

Performer	Team Members	
Iowa State	Electric Power Group (EPG), Google Brain, IBM	Robust Learning of Dynamic Interactions for Enhancing Power System Resilience
SEL	Oregon State University	Machine Learning Guided Operational Intelligence from Synchrophasors
University of California Riverside	EPG, Michigan Technological University	Discovery of Signatures, Anomalies, and Precursors in Synchrophasor Data with Matrix Profile and Deep Recurrent Neural Networks
University of Nevada, Reno	Arizona State University (ASU), IBM, Virginia Tech	Robust Event Diagnostics Platform: Integrating Tensor Analytics and Machine Learning Into Real-time Grid Monitoring
GE	GE Grid Solutions	PMU-Based Data Analytics using Digital Twin and PhasorAnalytics Software
Siemens	Southern Methodist University, Temple University	MindSynchro
Ping Things	NA	Combinatorial Evaluation of Physical Feature Engineering and Deep Temporal Modeling for Synchrophasor Data at Scale
Texas A&M	Temple University, Quanta Technology	Big Data Synchrophasor Monitoring and Analytics for Resiliency Tracking (BDSMART)

Talk Overview

"The goal of the panel session is to provide a venue and platform for the DE-FOA-0001861 awardees to share our experience, knowledge, and insights gained from mining the large-scale real-world PMU data from the U.S."

Order of the Problem

Group	Start	End	Days	Pts/Stream	PMUs
A	7/21/2018	8/24/2019	400	1,036,800,000	215
B	1/1/2016	12/31/2017	730	1,892,160,000	43
C	1/1/2016	12/31/2017	730	1,892,160,000	188

Estimated Upper Bound for Total Measurements

20,000,000,000,000

(20 Trillion)

The "Physics" of the Problem

Assume 10 Trillion Measurements

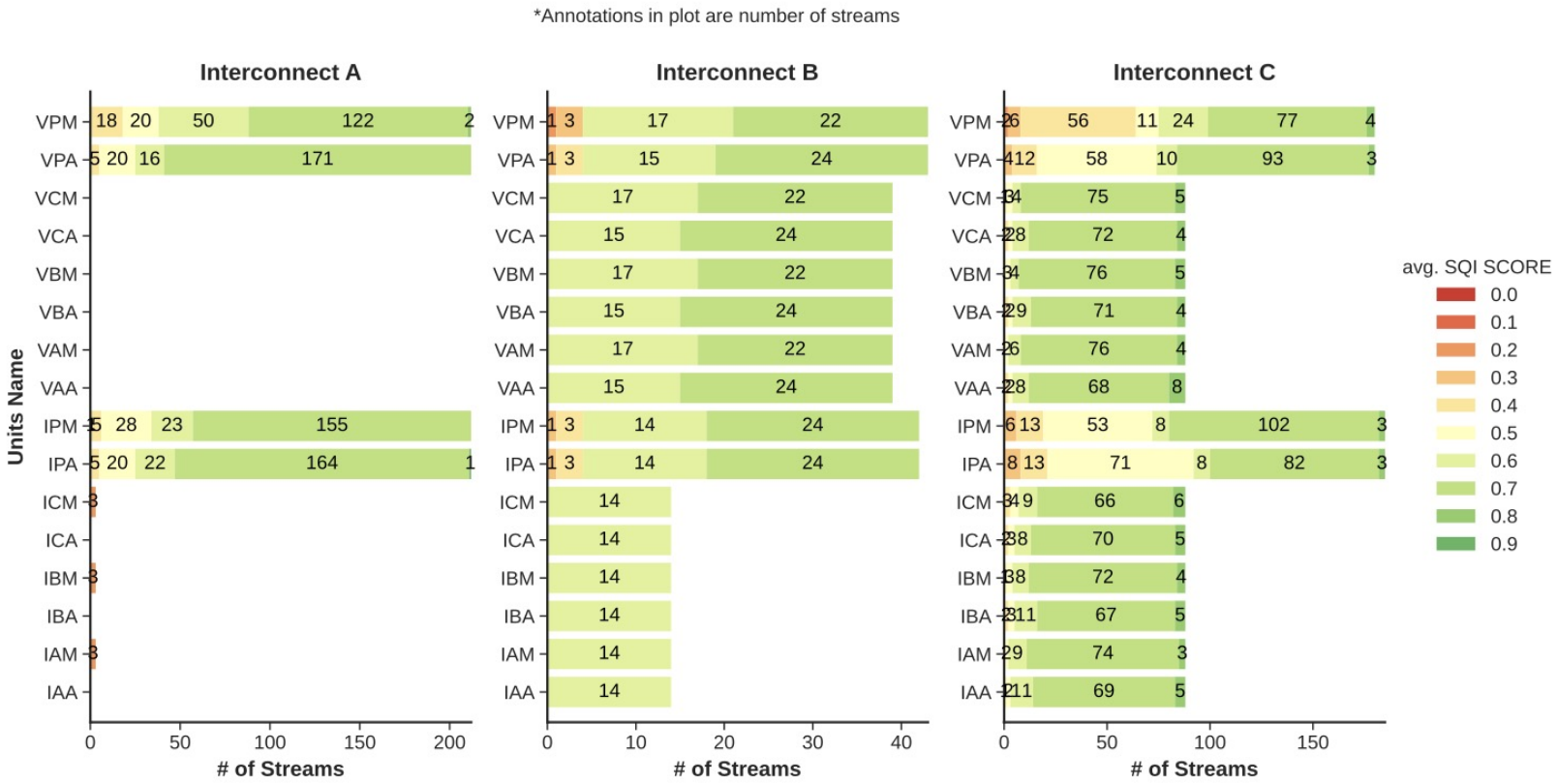
Read Speed	Time to Read All Measurements
10 kp/s	~ 3 Decades
100 kp/s	~3 Years
1 Mp/s	~3 Months
10 Mp/s	~ 2 Weeks
100 Mp/s	~ 1 Day

Data Quality: Stream Quality Index (SQI)

SQI encapsulates & quantifies several data quality issues per stream.

- Point count density
- Obviously bad values
- Repeated values
- Timestamp jitter

This enables quick filtering of poor-quality streams before analysis.



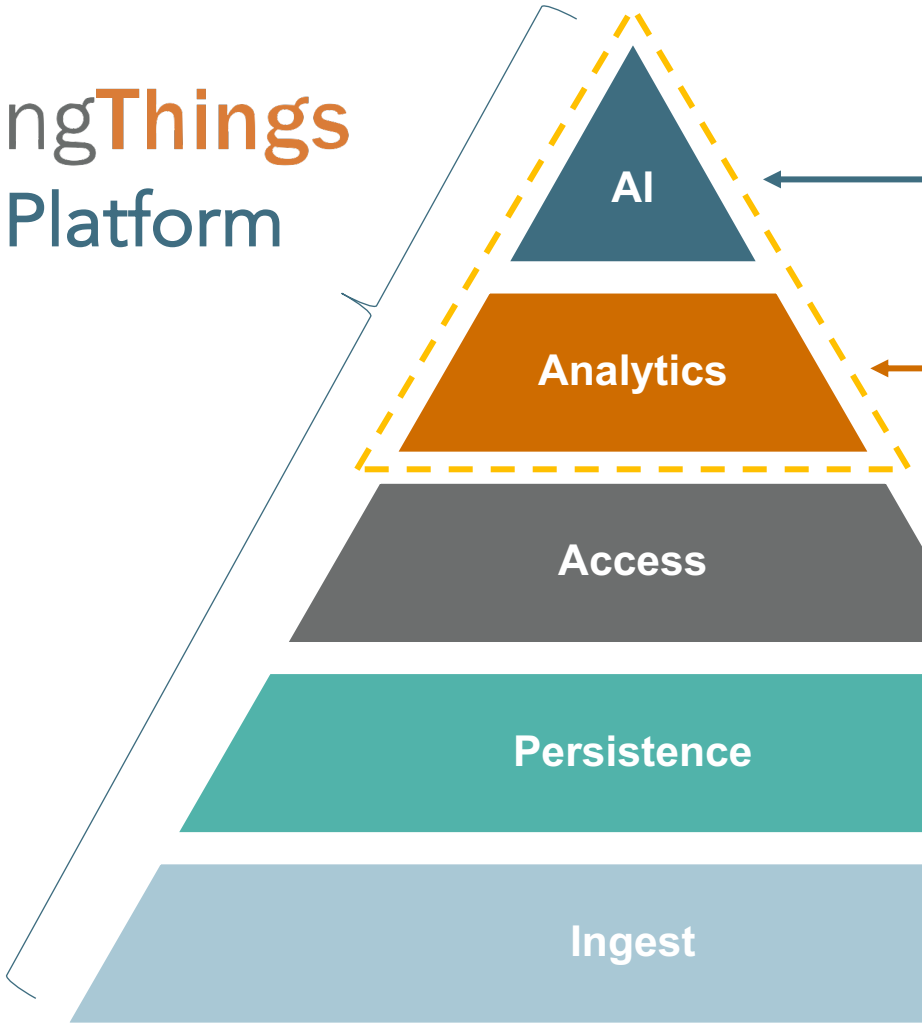
SQI across stream types in each interconnect.

~~Data Sharing~~ Data Collaboration

1. Intra-utility
2. Inter-organizational

A Hierarchy of Needs

PingThings Platform



Data becomes smarter
Train machine learning models for next-gen use cases

Put stored data to work
Realize previously-impossible engineering and industrial analytics at scale

Keep up with the speed of industry
Read 15M pts/s/node; write 30M pts/s/node

Keep and control *all* your data
Scales to Petabytes; best in class compression

Normalization across sensor types

The PingThings Platform

Data Sources

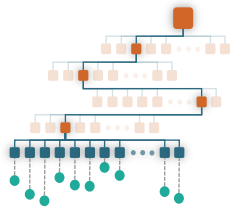
- *T&D Electric Utility*
- *RC / ISO*
- *Solar Fleet*
- *Wind Farm*
- *EV Charging Fleet*
- *Research Partners*
- *Sensor Fleets*
- *Data Concentrators*
- *Historians*
- *Individual Sensors*
- *Edge Devices*

PredictiveGrid™

Streaming Ingest

ETL

BTrDB
High performance, extremely scalable time series database for real world sensor data at scale.



Geospatial Data Store

Topology Store

Metadata Store

DISTIL
High performance stream and signal processing for time series measurements.

User Defined Functions

Jupyter Hub
jupyterhub

API and Access Control Layer
GraphQL
{ REST:API }
gRPC

Optional Managed Services

Apache Spark — Analytics

TensorFlow — Deep Learning

RAY — Compute

Applications

General Purpose

- Data Explorer
- Dashboards (Grafana)
- Notebooks (jupyter)
- Alerting
- Admin
- Platform Mgmt
- BI Tool Integrations

Industry-Specific

- Solar Integration
- Model Validation
- Small Signal Analysis
- Wildfire Detection
- GMD/GIC
- Capital Optimization

Customer and 3rd Party

App3 · App 6 · App 9

Secure, Scalable, and Reliable Infrastructure

IEEE PES Power & Energy Society®

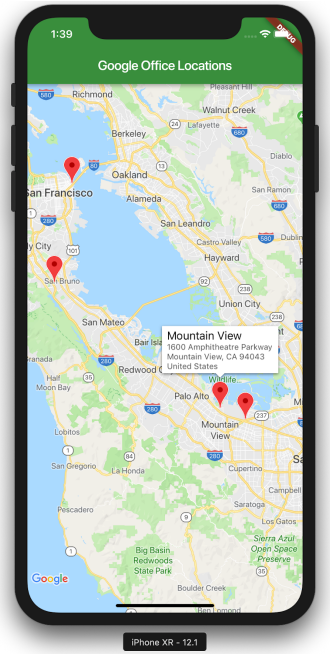
IEEE

The Future of PMU Data

Having all data at your fingertips fundamentally changed travel…

Small data with slow tools imprisons progress in invisible assumptions and processes

Data at scale allows new possibilities



Hundreds of billions in value created from unforeseen use cases

Uber
lyft
DOORDASH
DiDi
BIRD
VIA
yelp
instacart
Lime
Gett.
Flipkart
FOURSQUARE
BlaBlaCar
cabify
Rappi

… some utilities will undergo this transformation

IEEE PES Power & Energy Society®

IEEE