

A tropical sunset scene featuring two large palm trees in the foreground, silhouetted against a bright orange and yellow sky. The sun is low on the horizon, creating a lens flare effect. The ocean is visible in the background, with a few small sailboats on the horizon. The overall mood is warm and serene.

Aloha!

Tapis-CHORDS Integration: Time-Series Data Support in Science Gateway Infrastructure

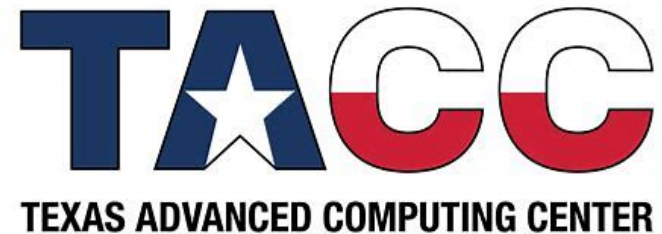
Gateways 19– San Diego, CA
September 25, 2019

Sean B. Cleveland, Anagha Jamthe, Smruti Padhy, Je'aime
Powell, Joe Stubbs, Michael D. Daniels,
Suzanne A. Pierce and Gwen A. Jacobs



Overview

- Project Background
 - Motivation
 - Hawai ' i EPSCoR - 'Ike Wai Gateway
 - Planet Texas 2050 – (Je'aime Powell up NEXT!)
 - The Challenge
 - Tapis
 - CHORDS
- Implementation
- Performance testing
- Future Work



Motivation - 'Ike (knowledge) Wai (water) Gateway

Climate Change, Incidents, Growth & Sustainability



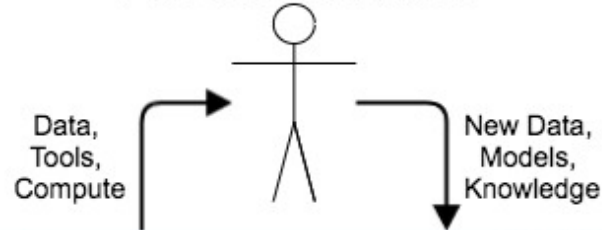
- **What is it?**
A data and modeling platform for Hawai'i volcanic hydro-geology, economic modeling, and decision support. Bringing data, models, computation, analysis and visualization into one place

- **What purpose does it serve?**
Enables decision and policy makers simpler access to better data and tools

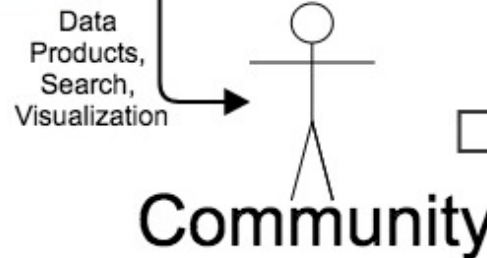
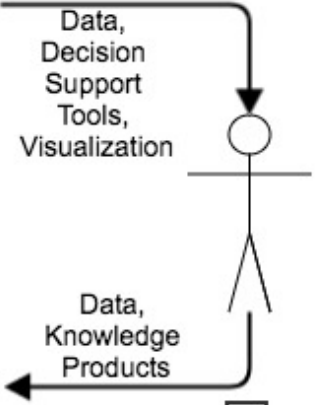
Data

GeoPhysics
Geo-Chemistry
Rainfall
Well Monitoring
Agency Data
Legacy Data

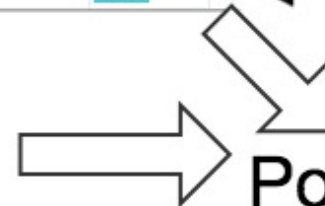
Researcher



Water Manager



Policy/Action



IKE

Q Search Ike Annotated Repository Files

Draw Bounding Box And Click "Search In Bounding Area" Button To Begin Searching...

Search In Bounding Area

Sensor Data and time-series legacy data (SGD, Chemistry, Precipitation, Well)

Filter With Selections

Matching Files: 58

Filter Current Local Search Results

Path	Filename
new_data/	3-1642-001--pdf
new_data/	3-1642-001.pdf
new_data/	3-1642-001.pdf-cp7
new_data/	3-1642-001.pdf5

Matching Water Quality Data: 115

Water Quality Site Name

- NF WAHEEE STREAM NR HEEIA, OAHU, HI 637
- SF WAHEEE STR NR HEEIA, OAHU, HI 641
- Kahaluu Str nr Ahumahu, Oahu, HI 602
- Kahaluu Stream near Heea, Oahu, HI 611
- SF Kapunahala Str at Kaneohe, Oahu, HI 659B
- Kamooali Str bw Luluku Str nr Kaneohe, Oahu, HI 6587

Matching Timeseries Observations: 99

Download Observations

Matching Timeseries Observations: 99

Download Observations

2016-05-19T16:00:00Z	its-fountain1	<ul style="list-style-type: none"> 1,1,1-Trichloroethane:milligrams per liter(mg/L) = 2 Chloride, dissolved:percent(%) = 69 Temperature:degree(deg) = 69 Wave height:meter(m) = 2,5
2016-05-19T17:00:00Z	punchbowl-wtr	<ul style="list-style-type: none"> 1,1,1-Trichloroethane:milligrams per liter(mg/L) = 4 Chloride, dissolved:percent(%) = 72 Temperature:degree(deg) = 58 Wave height:meter(m) = 2.4
2017-05-18T16:00:00Z	its-fountain1	<ul style="list-style-type: none"> 1,1,1-Trichloroethane:milligrams per liter(mg/L) = 2 Chloride, dissolved:percent(%) = 68 Temperature:dearee(dea) = 64

Download Observations

2016-05-19T16:00:00Z	its-fountain1	<ul style="list-style-type: none"> 1,1,1-Trichloroethane:milligrams per liter(mg/L) = 2 Chloride, dissolved:percent(%) = 69 Temperature:degree(deg) = 69 Wave height:meter(m) = 2,5
2016-05-19T17:00:00Z	punchbowl-wtr	<ul style="list-style-type: none"> 1,1,1-Trichloroethane:milligrams per liter(mg/L) = 4 Chloride, dissolved:percent(%) = 72 Temperature:degree(deg) = 58 Wave height:meter(m) = 2.4
2017-05-18T16:00:00Z	its-fountain1	<ul style="list-style-type: none"> 1,1,1-Trichloroethane:milligrams per liter(mg/L) = 2 Chloride, dissolved:percent(%) = 68 Temperature:dearee(dea) = 64



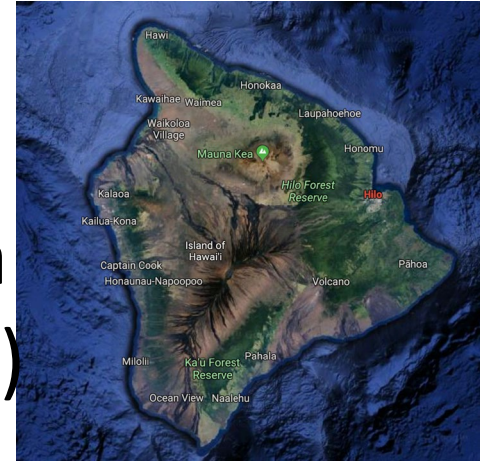
The Challenge

- Support:
 - the storage of streaming data from sensors and legacy time-series data
 - advanced annotation of streaming data/products
 - spatial queries to discover this streaming data.



Intelligent Systems Research to Support Geosciences (IS-GEO)

- Hosted Workshop for sensors on the Big Island
- Used the Cloud-Hosted Realtime Data Services for the Geosciences (CHORDS) for hosting the streaming data



This seemed like something we should leverage...



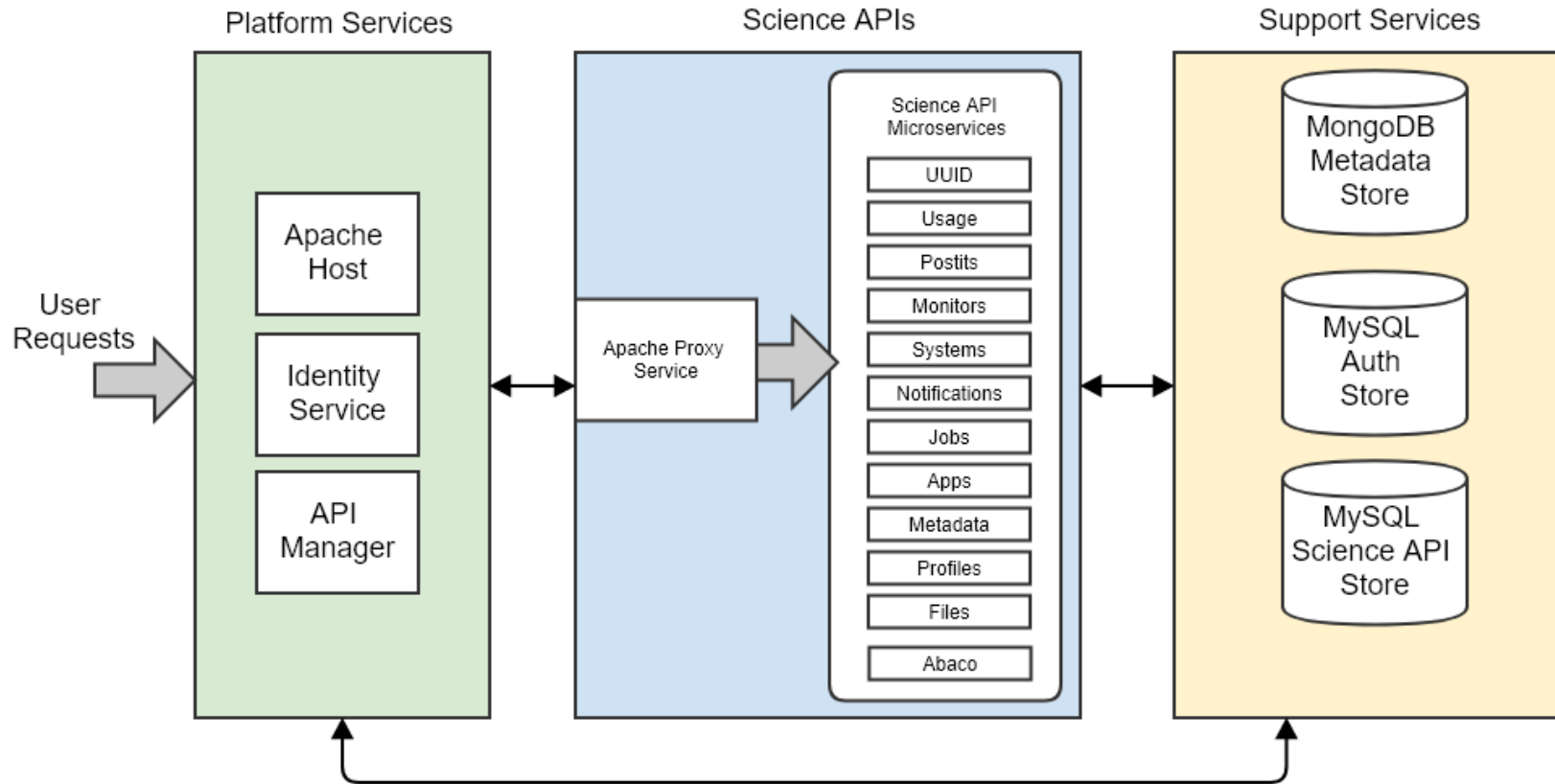
Our Challenge With Technologies

- Tapis offers almost everything necessary for all the use cases. BUT
 - There was a gap for:
 - Granular time-series data support.
- CHORDS support time-series data for instruments/sensors. BUT
 - There was a gap for:
 - Spatial data query support
 - Advanced/custom metadata annotation.
 - Multi-tenancy support.



Tapis Framework

Platform as a Service, REST APIs



Auth, Data Management, Compute

Project Tapis: Next Generation Software for
Distributed Research NSF CSSI #1931439, #1931575



CHORDS



EARTH CUBE

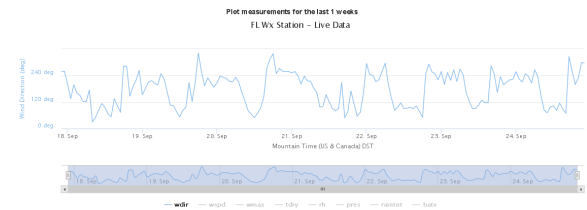
- Ruby on Rails APIs and web UI
- InfluxDB for time series data
- MySQL for Auth & Metadata
- Optional Grafana visualizations



FL Wx Station (id: 1) (sensor_id: n/a) located at NCAR Foothills Lab
 Description: Vaisala WXT510 weather station. See http://www.ed.ucar.edu/weather/weather_station.html.
 This instrument is designated as: ACTIVE
 (If 'INACTIVE', the instrument will not appear in the dashboard.)

Measurements
 2194039 measurements were reported.
 This instrument is expected to report a measurement every 300 seconds.
 The first measurement was measured at 2017-09-25 00:02:40 UTC.
 The last measurement is 2 minutes old. It was measured at 2019-09-25 04:15:50 UTC.

Data Access and Downloads
 Data ingest tool
[Data ingest tool](#)
 Download current day as
[GeoJSON](#), [GeoCSV](#)
 Instrument Description
[SensorML](#)



Data Start Time: 2019-09-19T04:15:50:00 Data End Time: 2019-09-26T04:15:50:00 [Adjust Data Range]
[Reset Zoom](#)

Short Name	Name	Units	Min/Max (Plot)	Measured Property
wdir	Wind Direction	degree deg	/	Wind Direction Angle
wspd	Wind Speed	meters per second m/s	/	Wind Speed
wmax	Wind Max	meters per second m/s	/	Wind Speed
tdry	Temperature	degree celsius degC	/	Temperature
rh	Humidity	percent %	/	Humidity Value
pres	Pressure	millibar mbar	/	Air Pressure Value
raintot	Rain Total	millimeter mm	/	Precipitation
battv	Battery	volts v	/	Temperature

[Back](#)

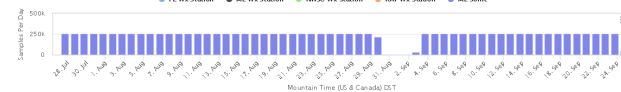
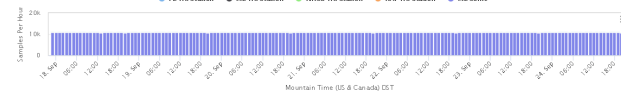
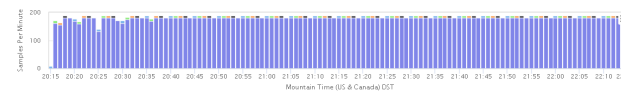


Dashboard

Measurements: 159,088,647
 Instruments: 5
 Sites: 4
 Storage: 591.0 MB
 DB Entries: 2017-09-27
 Uptime: 6 (days)

FL Wx Station (1) ML Wx Station (2) NWSC Wx Station (3) RAF Wx Station (5) ML Sonic (6)

http://portal.chordst.com/measurements/instrument_id=6&wid=228&wspd=0&4.7&station=608&day=median



Release: development (6996768) (2019-05-10 21:24:55 UTC)

Release: development (6996768) (2019-05-10 21:24:55 UTC)



UNIVERSITY OF HAWAII

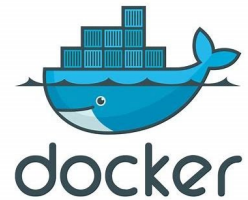
Our Solutions



Create a Proof of Concept **Streams API** that ties Tapis Metadata services together with CHORDS time-series data support.

Our initial implementation uses a NodeJS API

- Tapis Authentication and Authorization features to secure data and provide multi-tenancy support
- Tapis Metadata for Annotations
- CHORDS for measurement storage



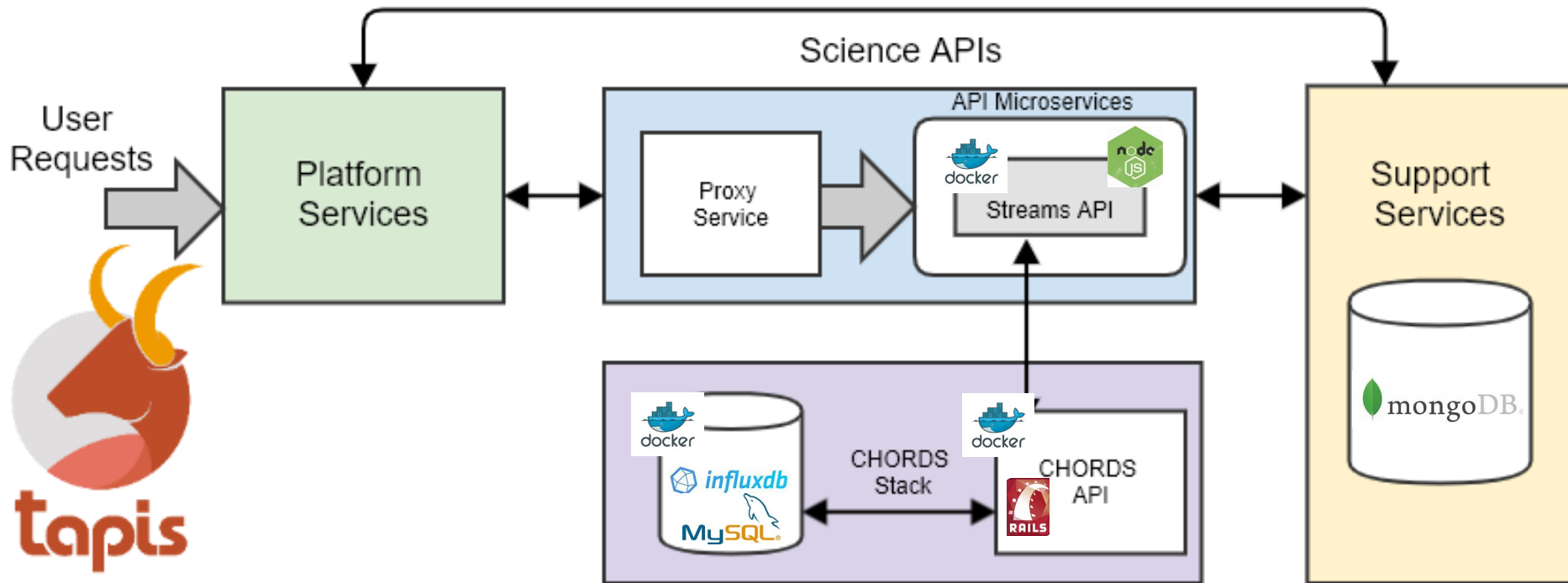
REST API Endpoints

- **Sites** – geographical location hosting 1-n instruments (WGS 84 GeoJSON)
 - **Instruments** – A source of measurements
 - **Variables** – A description of a measurement
 - **Measurements** – Recorded observation(s)
- **Spatial** - Query uses a GeoJSON polygon & returns matching Sites.

All requests require a Tapis API token with sufficient permissions in the inherited hierarchy(site/instr/var)



How does it work?



- All metadata(Site/Instrument/Variable) recorded in both Tapis metadata & CHORDS
- Utilizes a CHORDS service account for read/write
- CHORDS UI interface continues to work.



Performance Benchmarking

PERFORMANCE BENCHMARK: DATA INGESTION

Service	Number of POST requests per test run	Data Accuracy (%)	Avg. request response time (seconds)
CHORDS	100	100	0.11
CHORDS	200	100	0.11
CHORDS	500	100	0.12
Streams	100	100	0.65
Streams	200	100	0.6
Streams	500	100	0.6

The discrepancies in the times are related to the additional auth/authorization checks we have added on top of CHORDS (we plan to improve on these)



Performance Benchmarking

PERFORMANCE BENCHMARK: DATA FETCH

Number of GET requests per test run	Avg. response time CHORDS (seconds)	Avg. response time STREAMS (seconds)
100	0.61	0.98
200	0.67	1.0
500	0.64	0.91

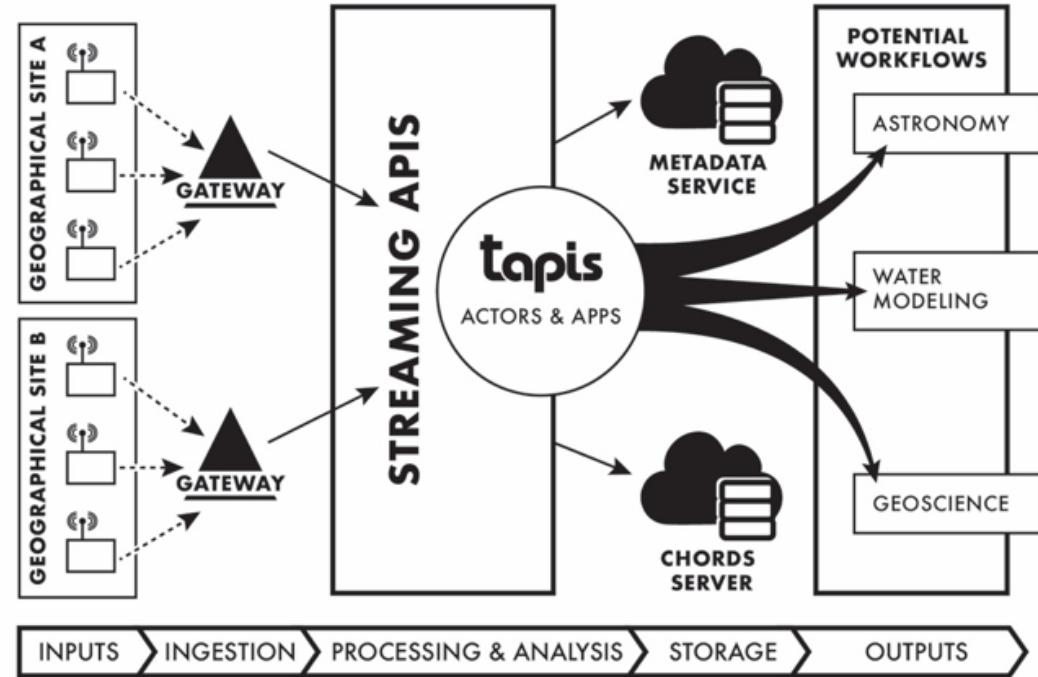
Some recent benchmarks for fetching **50K+** measurements clock in around **2.5** second across 500 repeated requests.(We plan to improve on that – we may go directly to InfluxDB in some cases)



Future Work



- Enhance performance of the Streams API
- Production release of the Streams API in Tapis coming in 2020
- Integration of Streams API with data triggers for executing applications and actors for workflow support.



Funding Acknowledgement

NSF CSSI #1931439, #1931575: Project Tapis: Next Generation Software for Distributed Research

NSF #1557349: RII Track 1: 'Ike Wai Securing Hawai'i's Water Future NSF OIA

NSF #1450413: Collaborative Research: SS2-SSI: The Agave Platform: An Open Science-As-A-Service Cloud Platform for Reproducible Science

NSF ICER #163221: EarthCube RCN IS-GEO: Intelligent Systems Research to Support Geosciences



EARTH CUBE



A vibrant tropical landscape featuring a double rainbow arching over a valley. The foreground is a lush green lawn with various tropical plants, including a prominent palm tree and several pandanus trees. In the background, dense green hills are partially shrouded in mist. A small white house is visible on the right side of the image.

Mahalo!

Questions?

