



Hewlett Packard
Enterprise

CADDY

Scalable Summarizations over Voluminous Telemetry Data for Efficient Monitoring



Presenter: **Saptashwa Mitra**

Authors: Saptashwa Mitra, Scott Ragland, Vanessa Zambrano, Dipanwita Mallick, Charlie Vollmer, Lance Kelley, and Nithin Singh Mohan

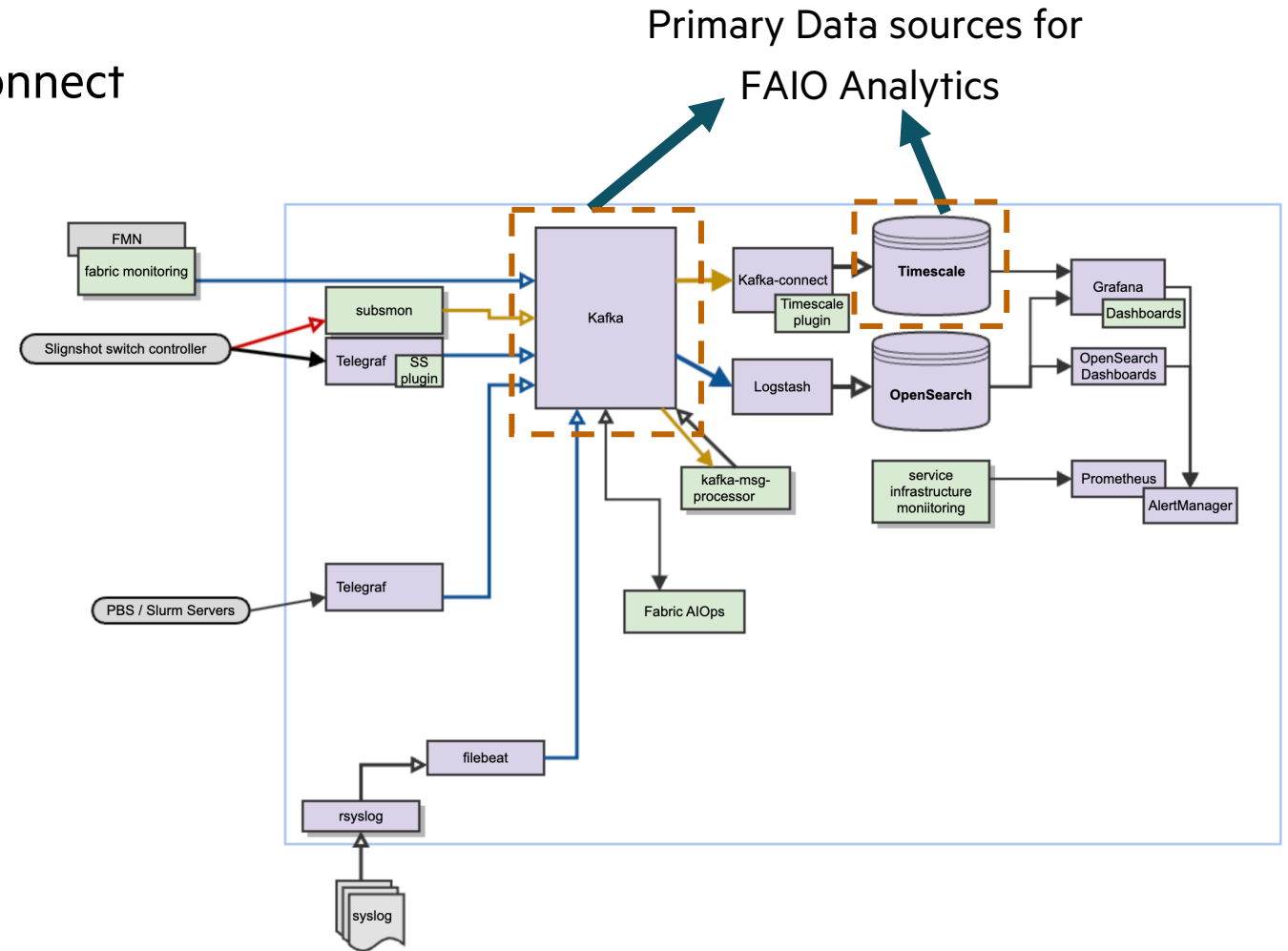
May 7th, 2024

Problem Definition



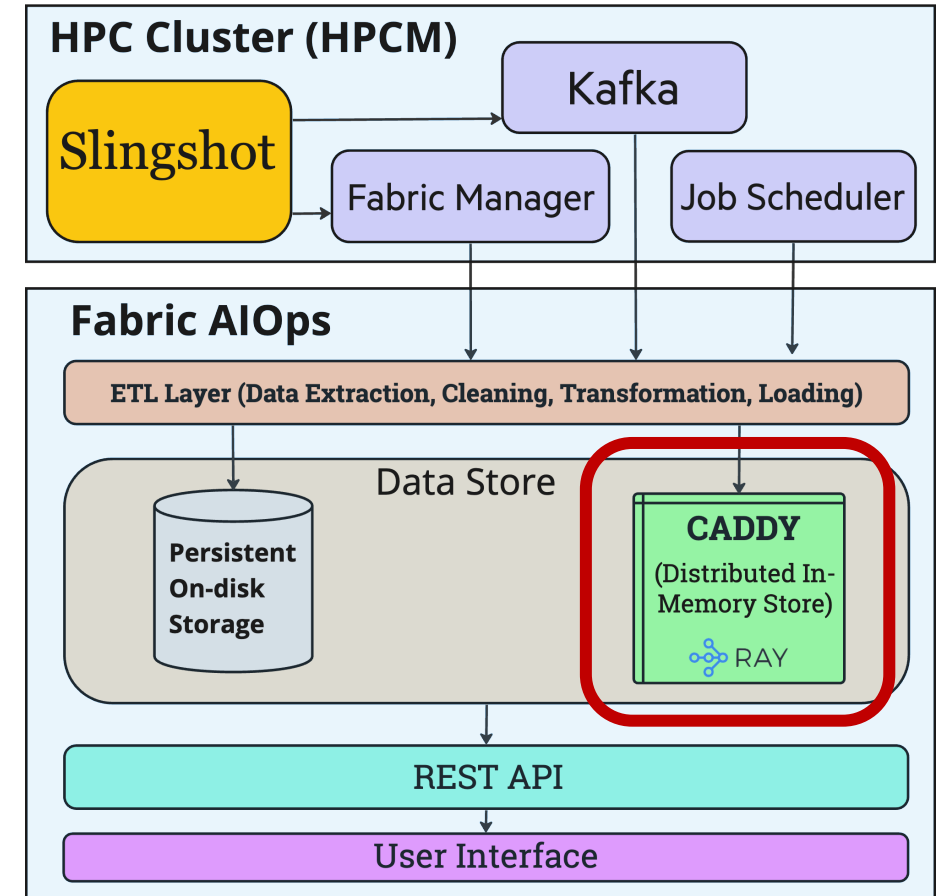
Overview: Fabric AIOps (FAIO)

- Monitoring framework over Slingshot interconnect
 - HPCM, CSM
 - telemetry data
 - High-Volume
 - High-Velocity
- System Performance analysis
 - Insight into fabric performance
 - Consolidated information
 - Visualize various cross sections of the fabric



Overview: Live Mode

- Live mode (*currently*)
 - Rapid analytics over most recent telemetry (10 mins)
 - Ray In-memory Store
 - Monitor traffic over the n/w for running workloads
 - Impact of job(s) over a network
- Enhance live mode
 - Allow sysadmins to **playback** fabric telemetry statistics
 - Diagnose fabric state
 - Interactive UI
 - In-memory storage
 - Improved storage capacity (*in the order of hours*)



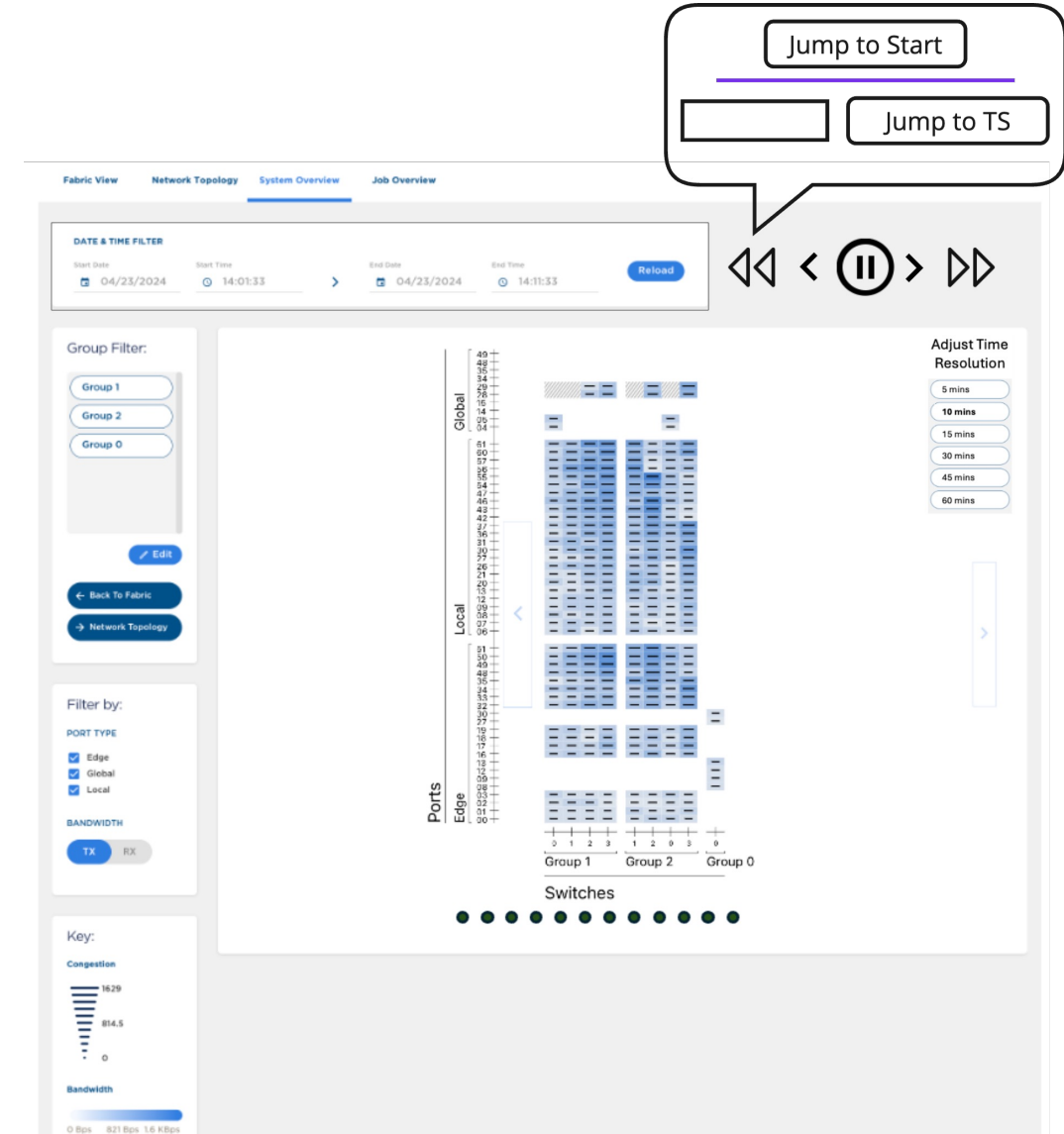
Scopes for Improvements over Live Mode

- Compressed representation of in-memory data
 - Better utilization of in-memory storage
 - *Eg. 500x compression for 10 mins snapshots*
- Optimized indexing
 - Faster identification of relevant data-segments
- Alleviate Redundant computation
 - Pre-aggregation of in-memory data
- Query analytics
 - Improved interactivity



Overall Goal

- Browse fabric state incrementally
 - Over *hours*
 - In small time windows (snapshot)
 - Easy panning (over a sequence of snapshots)
- Bird's eye view of the fabric state
- Analytics for better understanding of telemetry
 - min, avg, max, sum, std
 - Upto 4th order of moment
- **Configurability**
 - User-specified temporal resolutions for snapshots
 - Configurable Cache Management Strategies



Challenges

- Data compression
 - Loss in accuracy
 - Affects ingestion speed
 - Low overhead
- Sliding window
 - Identify and release *stale* data-elements
 - Low overhead
- Easy integration with current FAIO

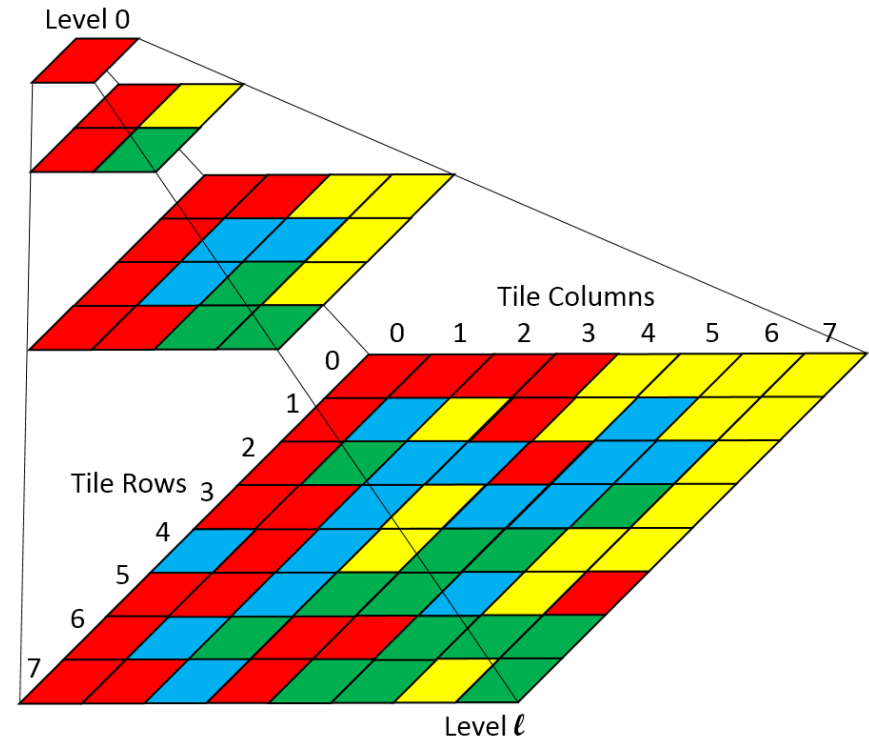


Methodology

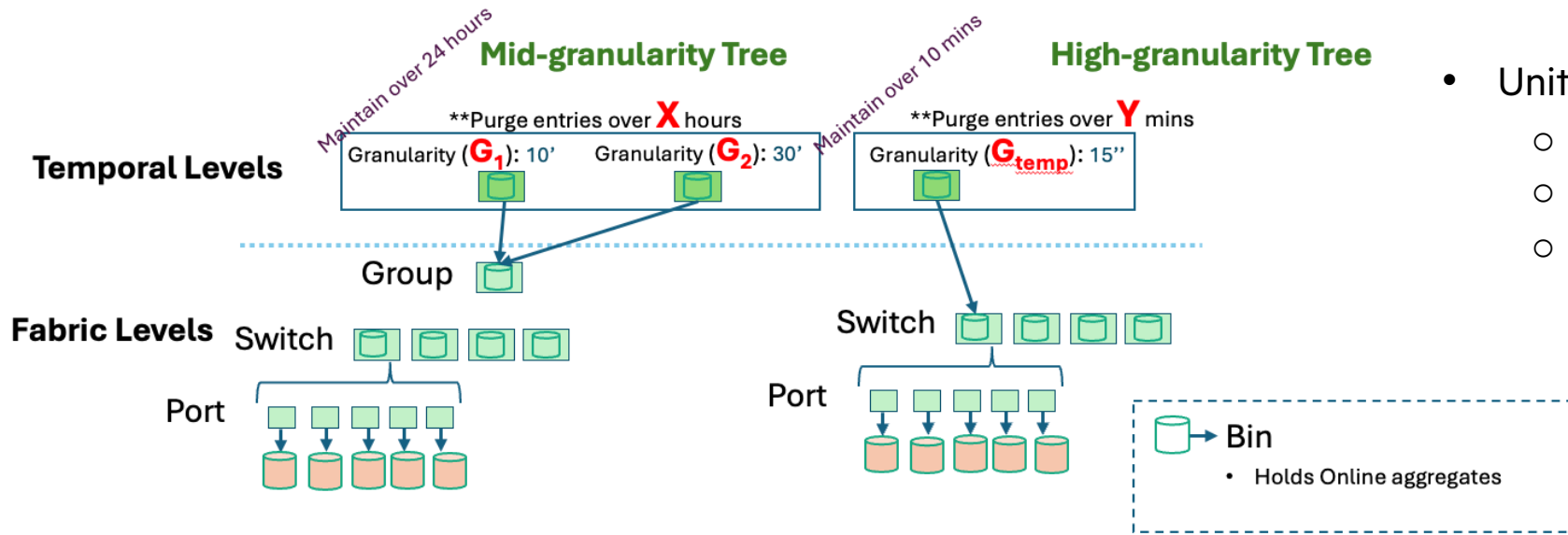


Background

- Pre-aggregation of data
 - Faster Analytics
 - Avoids redundant computations
 - Pre-processing overhead
 - Accuracy for streaming data
- Tile Layer
 - Aggregations at multiple resolutions of data
 - Faster query times
 - Links grow exponentially in number



Caddy: In-memory Telemetry



- Unit of Data: **Bin**
 - Aggregations at multiple levels of telemetry
 - Facilitates fast merging
 - Dynamic Hierarchical Aggregations

- Supported aggregations
 - min, avg, max, sum, std
 - Up to 4th order of moment

Bins: Compression and Accuracy

- Bounds
 - Timerange
 - Telemetry bounds {group, switch, port_range}
 - Key: (timestamp, group, switch, port)
- Store only fixed set of attributes (compression)
 - Irrespective of underlying telemetry events
- Statistically accurate online aggregation
 - **Welford's online statistics**
 - Fast updates
 - Statistically accurate
 - Constant memory footprint
 - Irrespective of data size

Initialize:

$m1 \leftarrow 0$ (mean)

$m2 \leftarrow 0$ (variance)

$m3 \leftarrow 0$ (skewness)

$m4 \leftarrow 0$ (kurtosis)

$n \leftarrow 0$ (number of samples seen)

for each new data point x **do**

$n \leftarrow n + 1$

$\delta \leftarrow x - m1$

$m1 \leftarrow m1 + \delta/n$

$m2 \leftarrow m2 + \delta * (x - m1)$

$\delta2 \leftarrow \delta * \delta$

$m3 \leftarrow m3 + \delta2 * (x - m1) - 3 * m2 * \delta/n$

$\delta3 \leftarrow \delta2 * \delta$

$m4 \leftarrow m4 + \delta3 * (x - m1) - 6 * m2 * \delta2/n + 4 * m3 * \delta/n^2$

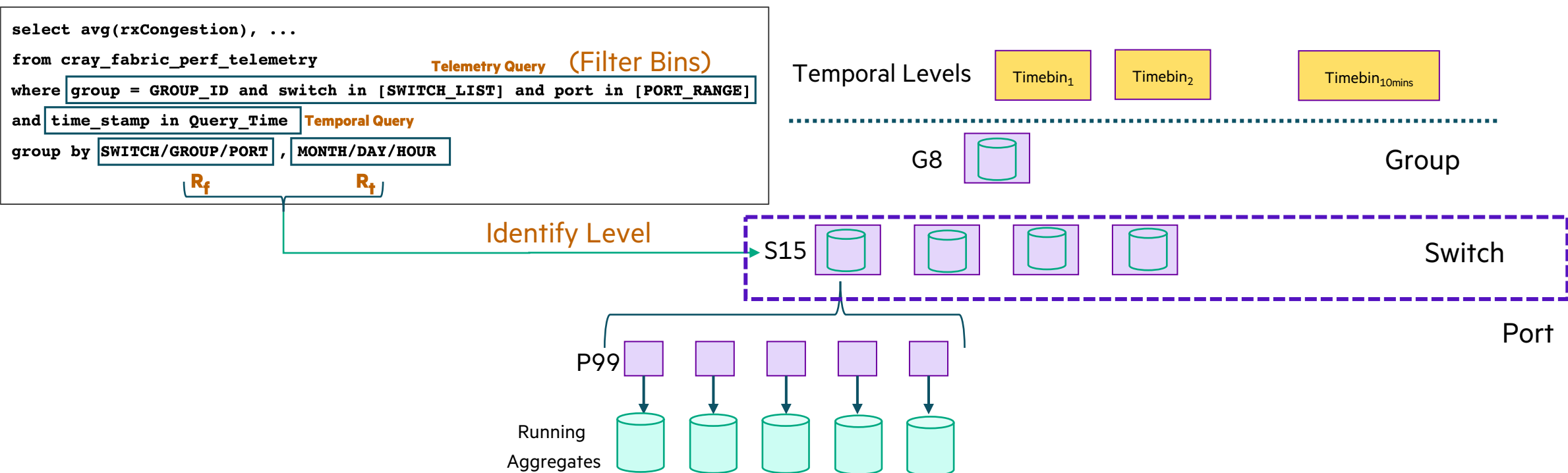
end for

Caddy Implementation

- Organization
 - Logically as a hierarchical tree of tiles
 - Physically
 - A hierarchical list of hashmaps
 - Each key has sufficient information to infer
 - Neighborhood
 - Position in hierarchy
 - Avoid need for links

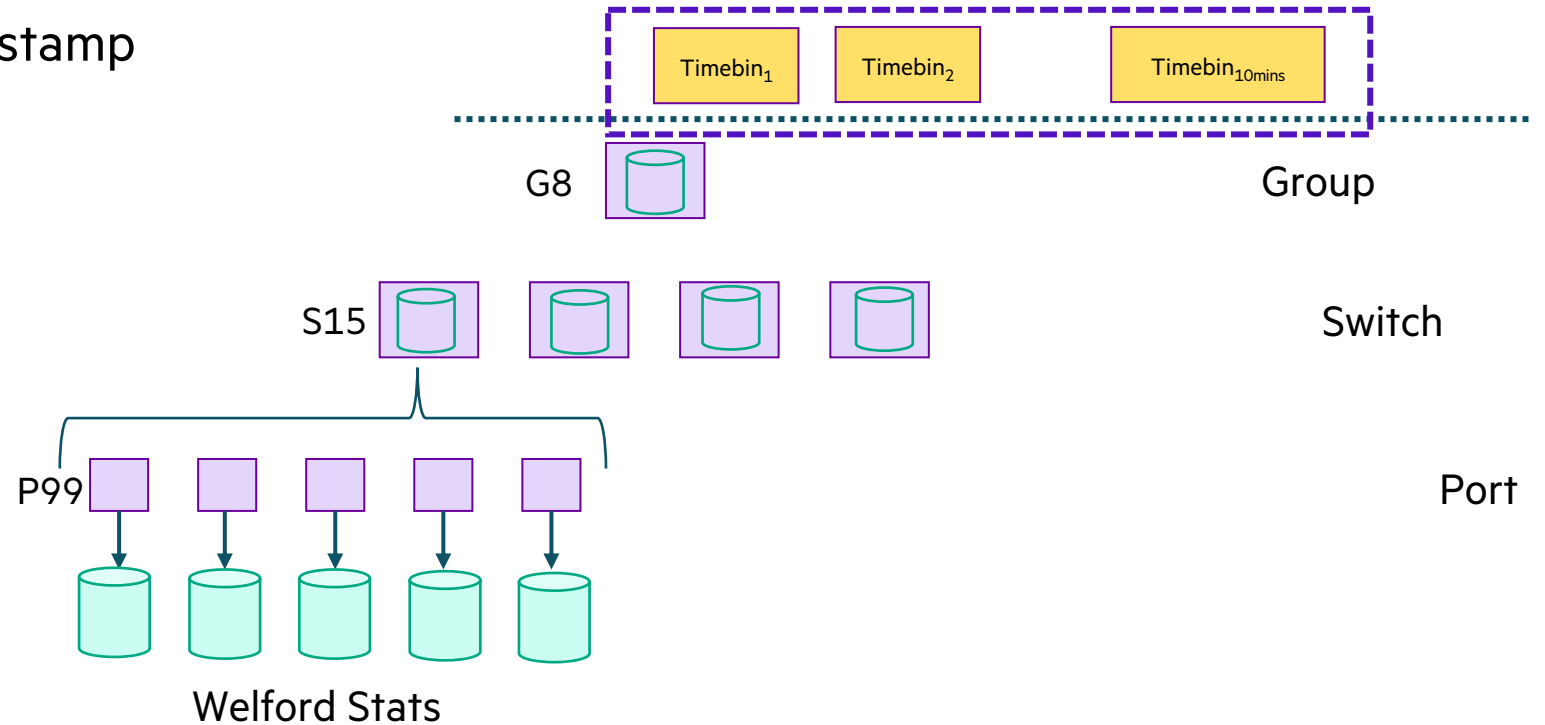


Fetching Snapshots (Query Evaluation)



Purging

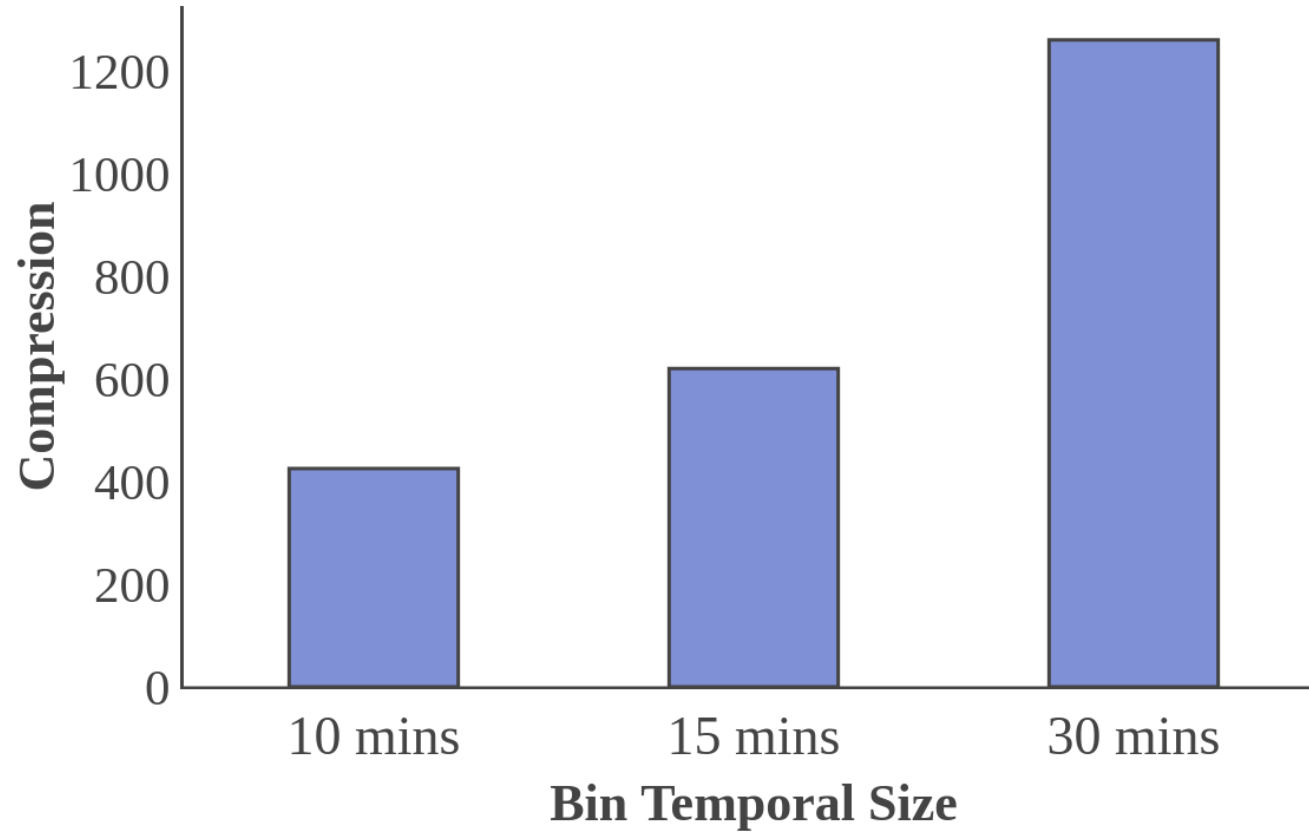
- Sort temporal level based on timestamp
 - Ascending order
- Remove entire subtree
- Faster purging



Results

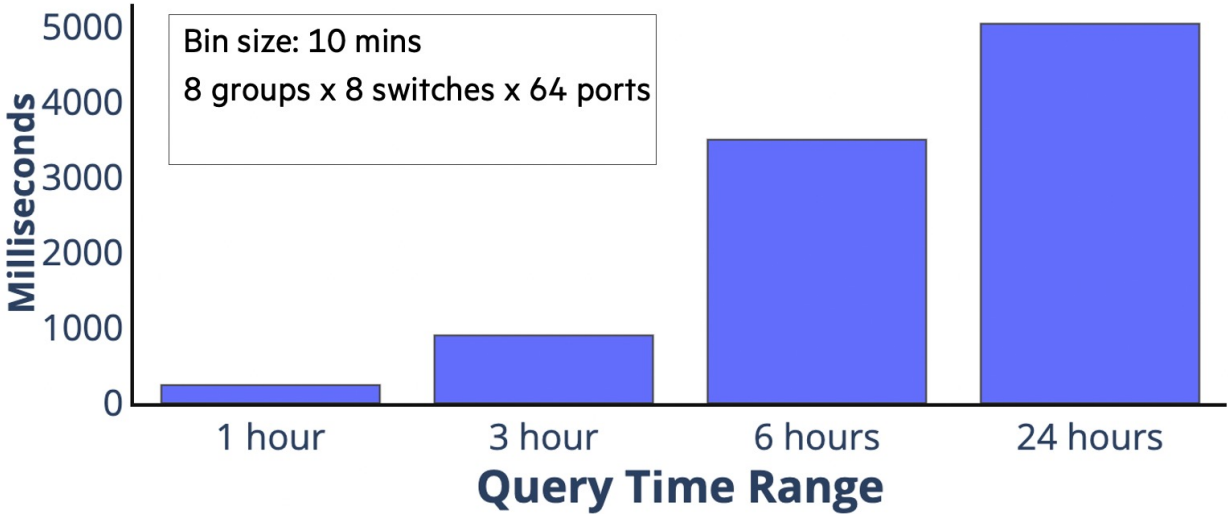


Compression Factor



Query Latency

Single Frame Fetch Time		
Mode	Time (ms)	
Live Mode	355	
Live Mode + Caddy	Group Level	151
	Port Level	267.51



Maintenance Overhead

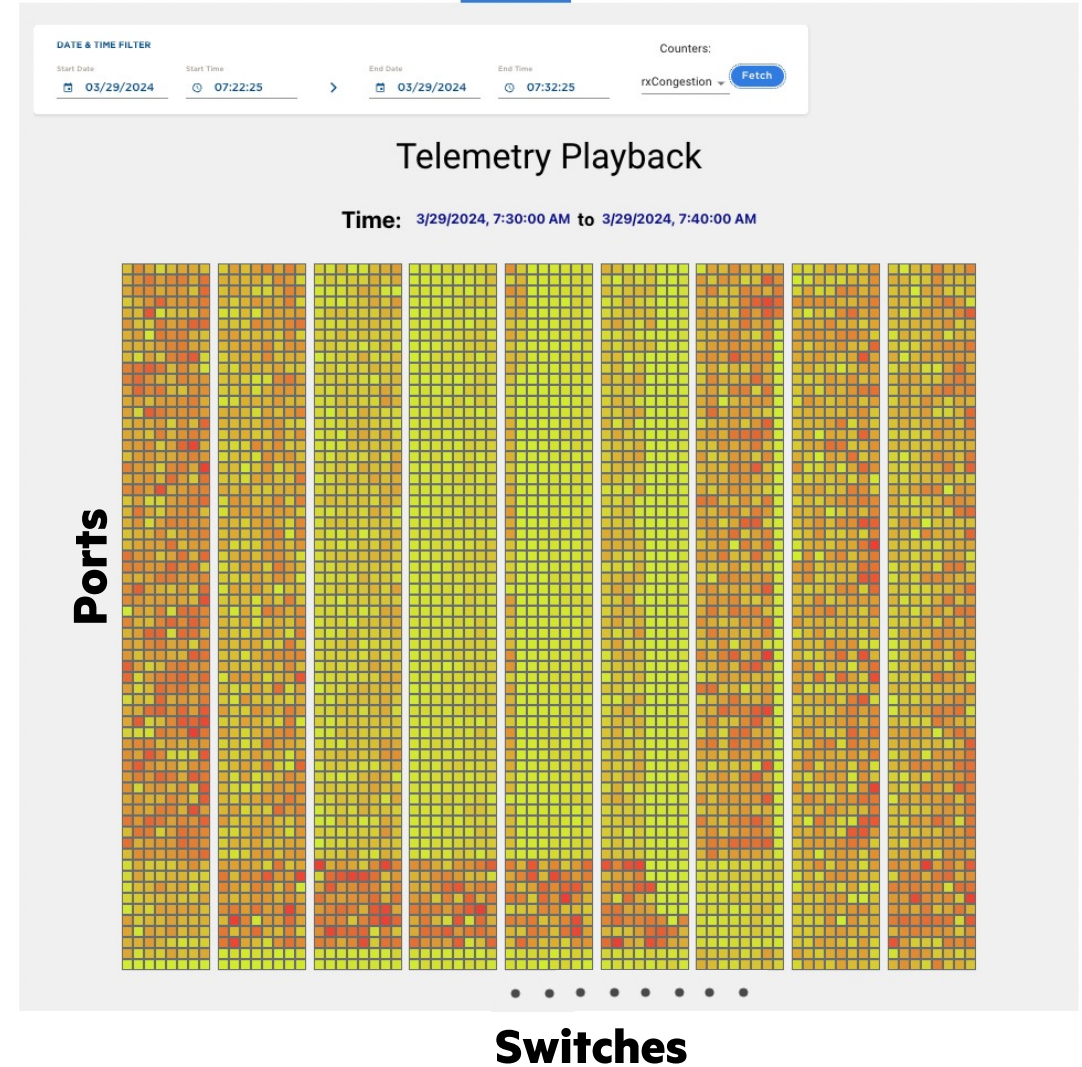
- Data Ingestion
 - Ingestion of 1M telemetry events
 - Over 1 hour period
- 18 - 32.5% overhead

	Time (s)
Live Mode + Caddy	1.007
Live Mode	0.7589



Telemetry Playback

- Prototype interface



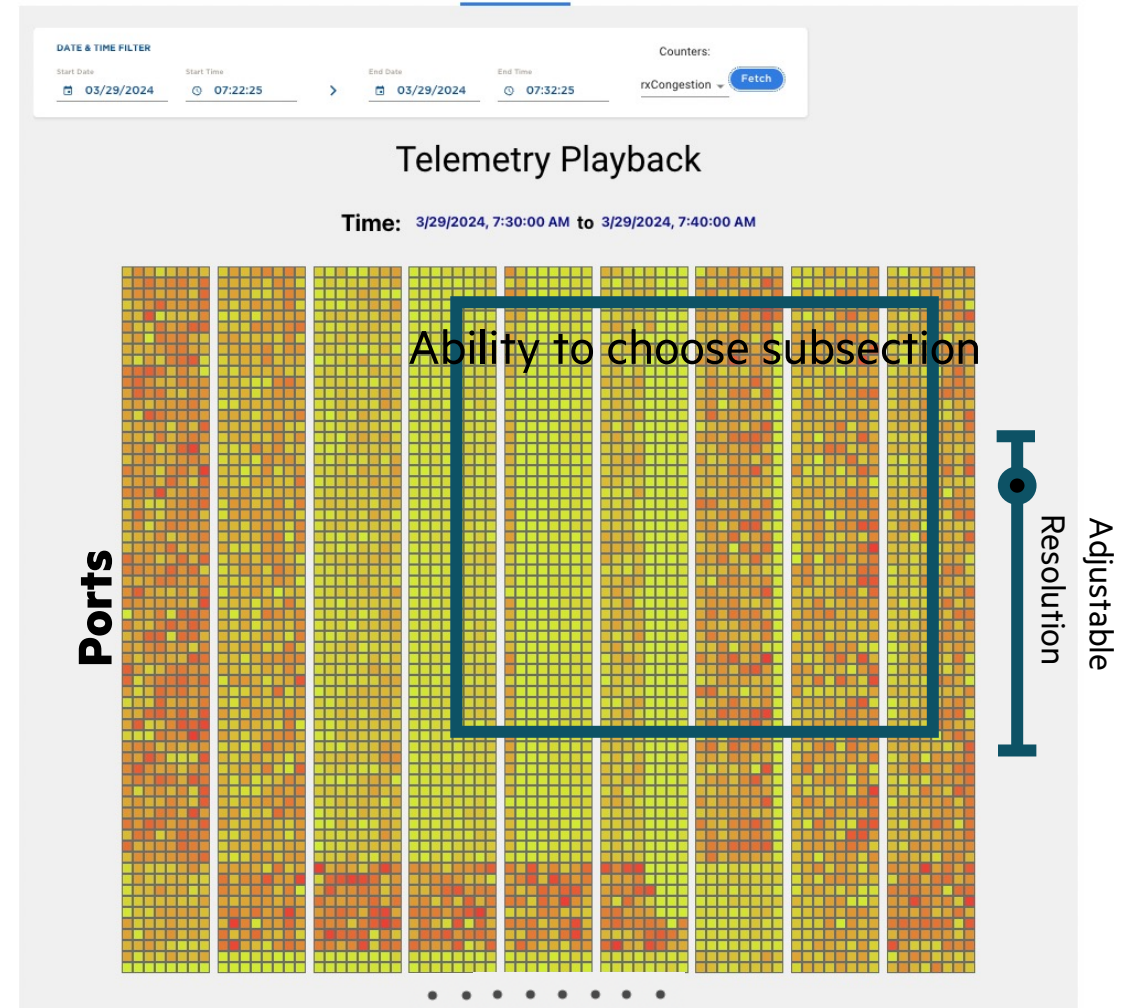
Overview of Features

- New changes are transparent to the UI
 - No loss in accuracy
- Improved interactivity
 - Adjustable queries



Utilities & Upcoming Features

- Allows for “exploratory browsing”
- View the overall fabric at a coarse resolution
 - Limited range of temporal resolutions
- Zone-in on problematic temporal/fabric sections
 - View detailed statistics over problematic regions
 - Through historical mode data from persistent storage
 - Or from lower-level in-memory nodes
- View subsections of the fabric
- Support more complex multivariate analysis
- Use granular aggregates for predictive models



CADDY: Scalable Summarizations over Voluminous Telemetry Data for Efficient Monitoring

Contact: saptashwa.mitra@hpe.com

Thank you

