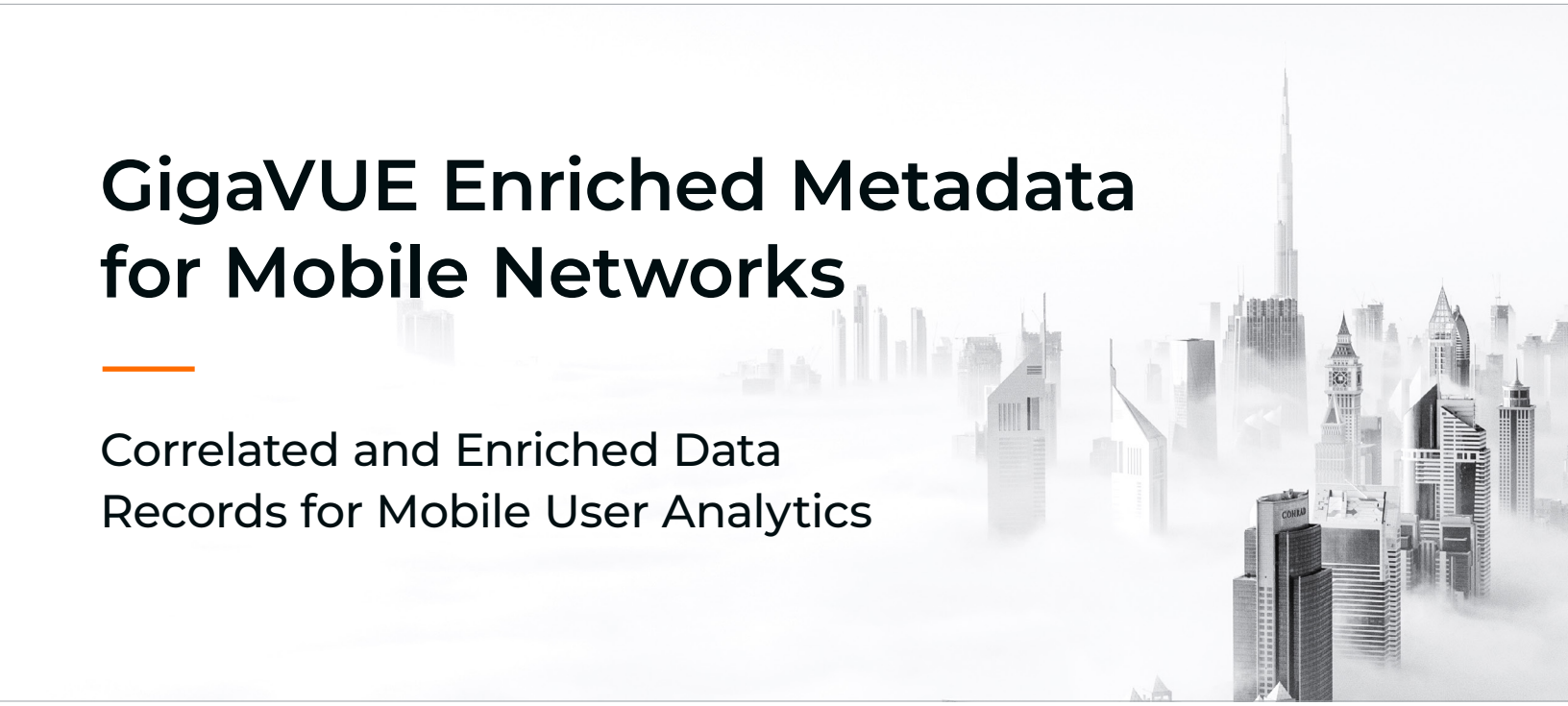
GigaVUE Enriched Metadata for Mobile Networks

Correlated and Enriched Data Records for Mobile User Analytics



KEY BENEFITS

- 99% volume reduction for user plane monitoring
- Transforms 4G and 5G packet data into one metadata record per user application flow
- Enriches activity metadata with subscriber, device, cell, QoS, network slice, and more
- Integrates natively with Apache Kafka bus for data lakes and multi-consumer deployments
- Produces metadata with sub-minute latency and granularity for analytics

Introduction

The need for efficient, high-quality data about mobile network users is self-evident. Marketing, advertising, infrastructure planning, customer service, and product development all benefit from better data about mobile users. But traditional approaches to collect this data have been costly. This is why Gigamon has developed GigaVUE® Enriched Metadata for Mobile Networks, to power analytics around mobile users, their activity, application usage, and patterns, for every subscriber, every device, and every location–but with efficiency designed in.

GigaVUE Enriched Metadata begins with high-quality packet data derived from the network, without impacting the network. Through passive taps and out-of-band processing, this solution starts with an independent, immutable, and trustworthy source of data. This data is then cleansed, summarized, and transformed into high-quality metadata records, enriched with key attributes well-suited for data mining analytics, and AIOps.

Mobile data per subscriber is **expected to triple** to 47GB over the next 5 years.

Dealing With Large Uncorrelated Datasets

The first-order challenge with understanding subscriber activity is the ability to see it and capture it at scale without overloading systems. Mobile data per subscriber is expected to triple to 47GB over the next 5 years.[1] That translates to roughly 250Gbps at busy hour for every 1 million subscribers. Traditional databases, and even modern data analytics systems, generally aren't equipped to handle this type of raw volume. Something needs to be done to simplify the problem.

The second, but related, challenge, is matching subscriber usage with meaningful information about that subscriber. Mobile networks are optimized for throughput and security, not reporting. When a subscriber connects to the internet from a mobile phone, their subscriber information is not transmitted on every packet. Reporting subscriber usage by IP address alone is incomplete for modern analytics objectives. Therefore, a second task is matching, or correlating, the large dataset of subscriber usage (aka. user plane) with the large dataset about the subscriber context (aka control plane), which includes who they are, what device they are using, and their location.

A Flood of Data, Easily Summarized

Rather than deal with raw packets, databases and data lakes prefer to store information in summary records. A database does not need to store 100,000 packets of the YouTube video, it just needs one record which essentially says, "subscriber A watched YouTube video B from location C on device D, and they used 100,000 packets over E seconds."

This is precisely what Gigamon does. With GigaVUE Enriched Metadata, all user plane flows are first summarized, correlated in near-real time, and fed to a data lake or Kafka bus for analytics. These records also stay consistent across the infrastructure, which can be extra beneficial for operators who deploy best-of-breed networks utilizing network elements from various suppliers.

This metadata can then address critical business questions, such as:

- Who are the top talkers? What apps are they using?
- Which subscribers spend the majority of their time on social media? Which subscribers only ever do email and web? Which subscribers are into gaming?
- What is the breakdown of apps used on a particular cell tower? Does this change throughout the day?
- Who is attending live sporting events, concerts, and other arena/stadium events? Who is shopping at particular retails stores or malls? (Perhaps such venues might be interested in this data.)
- Are any users violating terms of service by operating cryptocurrency mining or cloud computing?

How It Works: Leveraging the Deep Observability Pipeline

GigaVUE Enriched Metadata leverages existing solutions within the Gigamon Deep Observability Pipeline, providing one common infrastructure. This ensures you can both:

- Optimize packets to performance and troubleshooting tools, reducing monitoring costs
- Transform packets into GigaVUE Enriched Metadata, for data lakes, analytics, and AIOps

This solution is designed to serve all mobile network architectures, both classic packet cores, as well as modern 4G and 5G networks with high levels of scalability, versatility, and distributed functions.

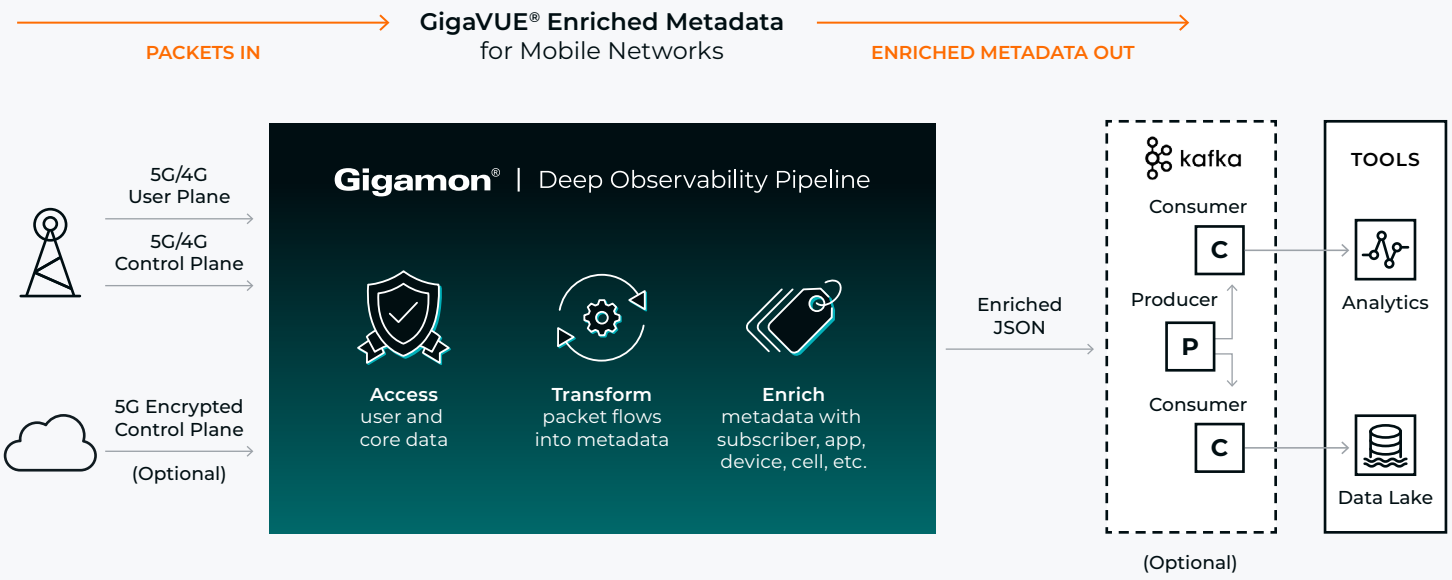


Figure 1. Diagram of Deep Observability Pipeline Components.

Components In Harmony

GigaVUE Enriched Metadata leverages many capabilities already present within the Deep Observability Pipeline, including tap landing, aggregation, and filtering, as well as plaintext access to encrypted 5G control plane, necessary for many modern 5G standalone or converged cores.

Inside the pipeline, the Gigamon metadata transformation process converts user plane packets (S1-U interface for LTE and N3 interface for 5G) into user plane metadata for every application flow, including default attributes for over-the-top application identification, byte volume, duration, and more. These metadata records are further enriched with the control plane data (S11 interface for LTE and N11 interface for 5G), to include subscriber id, device id, cell id, QoS id, and more.

This solution further leverages the Gigamon metadata export capabilities, which include JSON conversion, writing to local files for pickup, and/or producing to a Kafka bus.

Lastly, the Kafka Bus often represents a key component in the overall solution. While not part of the Gigamon Deep Observability Pipeline, Kafka is a popular industry-grade solution for brokering metadata. It follows a producer-consumer model, where a data source can produce, or publish, data to the Kafka bus, and others on the same network can consume, or subscribe to, the data that was published.

This is commonly used across a wide range of data sources, and for a variety of consumers. In the case of GigaVUE Enriched Metadata, Gigamon acts as the producer, and the end tools, such as the analytics engine or data lake, would be the consumers. Kafka is optional. When Kafka is not present, tools can pick up data directly from Gigamon. Note that Gigamon does not offer, install, nor operate the Kafka bus.

End-to-End Latency

An important consideration for many analytics use cases is the latency of the data. With GigaVUE Enriched Metadata, a record is delivered to Kafka or tools within 1 minute of the first packet on the network. In some cases, especially for reliable flows (for example, TCP), it can be considerably less time.

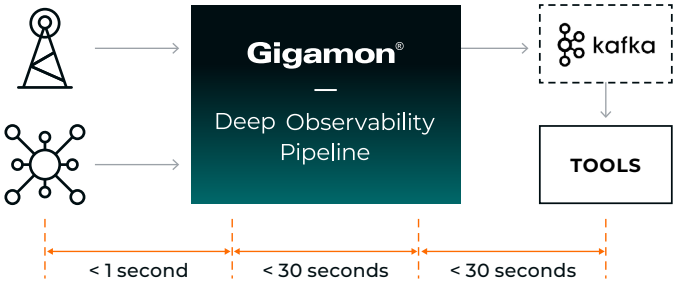


Figure 2. End-to-End Solution Latency.

The first packet in any flow will reach the Deep Observability Pipeline almost instantaneously. Interim records are generated for active flows every 30 seconds; however, TCP flows can be forwarded quicker, as soon as the TCP connection closes.

The data export interval is set to the minimum value of 30 seconds by default. (Alternatively, data can also be exported based on maximum record count in a batch, instead of number of seconds.)

In this way, the total time between when the mobile user starts an app or flow, and the time when a record is received at Kafka or the tool, is 60 seconds or less.

Mapping Metadata to Use Cases

The following is a list of several example use cases for GigaVUE Enriched Metadata, including background information about each use case, and a highlighted set of basic information on how to query it.

A. Network Improvement

- Network Planning and Optimization
- Automated Resource Management for Self-Optimizing Networks
- Network Upgrade: Before & After
- Quality of Service (QoS) Improvement
- Network Security

B. Customer Experience

- Service Personalization and Optimization
- Customer Experience Scoring
- Roaming Subscriber Usage
- Enterprise Customer (MPN) Usage Reporting
- Terms of Service Violation Detection

C. Revenue Generation

- Revenue Sharing Agreements
- User Profile Data for Large Venues
- User Profile Data for Application Providers
- User Activity Data for Advertisers

In all cases, operators should take care to ensure compliance with local laws, regulations, user agreements, and any other data governance considerations.

A. Network Improvement Use Cases

Network Planning and Optimization

GigaVUE Enriched Metadata aids in network planning and optimization by providing insights into user traffic trends, coverage areas, and service demand. This information helps Mobile Network Operators to make informed decisions regarding optimal configuration, network expansion, capacity upgrades, and the deployment of new infrastructure to meet growing service demands. A few examples are:

- Usage volume by cell site by application for optimal cell site configuration
- Usage volume by region/core data center for core network planning

Cell Site Usage Analysis
This data can be used to analyze the usage per cell site to aid in optimizing network configurations. Examples of queries are as follows:

- Query to group by cell id and aggregate by data volume percentage
- Query to list all the records in a selected time interval by selecting the cell id, subscriber id (IMSI)
- Query to group by cell id and aggregate by subscriber count

Cell Site Application Analysis
Additional analysis can be performed to understand data volume and application mix for a given cell, by grouping on application name.

Time-Based Cell Site Analysis
Additional analysis can be performed to understand data volume and application mix for a given cell, by using the start_time and end_time attributes and arranging into buckets such as hour-of-day and day-of-week. This information could be used to map out cell configuration optimization that changes throughout the day and week.

Core Network Usage Analysis
Like cell usage analysis, core network usage analysis can be performed using the same data volume and subscriber counts but grouping on the control tunnel core IP address.

Automated Resource Management for Self-Optimizing Networks

GigaVUE Enriched Metadata can be used to power algorithms and automation around resource allocation and management. Through automated analysis of historical data patterns and current data usage, Mobile Network Operators can improve overall network experience while minimizing resource waste. Specific areas that benefit from automation include self-optimizing cell configuration, and virtual network element allocation or de-allocation.

Automated Cell Optimization

This use case builds upon the "Time-based Cell Site Analysis" capability outlined in use case 2, to prevent network bottlenecks by understanding given cell usage by time-of-day and day-of-week, but further leverages the sub-minute latency and granularity of GigaVUE Enriched Metadata to deliver near-realtime information to automation and self-optimization systems.

Automated Virtual Resource Optimization

Similar to cell usage analysis, core network usage traffic patterns can be analyzed by the time of the day for resource usage, by grouping on the gateway ID (Core Addr). For virtualized network elements serving these connections, including the Serving Gateway (SGW), User Plane Function (UPF), or the OpenRAN Cloud Unit (CU), virtual resources can be deallocated or reallocated on an automated basis.

Detection of Roaming Issues

For many operators around the globe, a change in seasons can bring about a change in roaming subscriber traffic. In some cases, a sudden increase in subscriber count can cause adverse effects on the network. GigaVUE Enriched Metadata can be used to report on roaming subscriber increases in specific locations, or to detect issues related to increased control plane usage on the network due to roaming.

Network Upgrade Before & After

GigaVUE Enriched Metadata can be used to analyze the impact on the network after a network upgrade or change control. After the upgrade or change is completed, a correlated view of metadata will be able to offer deeper insights into changes. Significant differences can be spotted at the application and cell level, to indicate potential problem areas resulting from the upgrade or change control.

Network Performance by Application

- Query for mix of top N applications by volume
- Query for mix of uplink and downlink

Network Performance by Network Element

- Query for usage volumes at the gateway level
- Query for usage volumes at the cell level or detect cells with no traffic
- Query the control plane transaction counts by the session event types to ensure the counts are not dropping significantly

Quality of Service (QoS) Improvement

GigaVUE Enriched Metadata provides insight into application traffic patterns, allowing Mobile Network Operators to understand network performance by QoS class, to ensure the intended levels of performance are being met at the various QoS levels, or whether improvements are needed to enhance experience quality for different types of applications and services. By analyzing application throughput, combined with other network performance datasets, network operators can allocate resources to ensure the right QoS for the applications.

Application Performance by QoS

Here the application is combined with QoS identifier:

- Query to group by application name and QoS Class Id (QCI), aggregating uplink (src) and downlink (dst) data volumes for a selected time interval
- Optionally join to other data sets which group applications into an application type

This provides the application usage throughput, separately for uplink and downlink, which is important for validating performance for certain types of QoS. This data may additionally be combined with other attributes to derive additional information like packet loss, and packet delay, to ensure the quality of service for different applications. This enables the network to prioritize critical applications over less time-sensitive applications.

Application Performance by APN/DNN

Additional analysis can be performed by grouping or filtering on apn, in order to look at QoS performance for a given network name.

Network Security

GigaVUE Enriched Metadata can be used for security monitoring and threat detection. By analyzing patterns and anomalies in data flows, spotting expired certificates, and detecting port spoofing, Mobile Network Operators can identify potential security threats, including malicious attacks or other abnormal behavior, and take appropriate measures to safeguard the network and user data.

Unusual Behavior

Examples of queries are as follows:

- Query to group by application type and aggregate by the uplink data volume for a selected time interval
- Query to group by application type and aggregate by downlink data volume for a selected time interval

This detects unusual behavior in the applications used in the network. By monitoring and analyzing the traffic pattern, operators can spot patterns, such as unusually high amounts of upload on unknown applications. This enables the implementation of proactive security measures, thereby enhancing the overall security posture of the 5G network.

Port Spoofing

Port spoofing can be a red flag for security, and can be detected by looking at extensible attributes. This is particularly important because traditional log data often only relies on port number to identify protocol, giving threat actors cover for malicious activity.

Expired Certificates

Extensible attributes also allow for the detection of expired certificates.

B. Customer Experience Use Cases

Service Usage and Personalization

GigaVUE Enriched Metadata provides information about application usage for each individual subscriber, or user, enabling Mobile Network Operators to gain insight and tailor their product and service offerings to specific customer segments. By understanding usage patterns, providers can further offer personalized services, promotions, and recommendations. Usage reporting includes, but is not limited to:

- Top N Over-The-Top (OTT) Applications by device type for device promotions
- Top N OTT Applications by region to inform for regional offerings
- Top N OTT Applications by subscriber groups for prepaid/postpaid or other subscriber group plans
- Application usage by roaming subscribers for roaming plans
- Mix of time spent on social media apps, vs. productivity, vs. gaming, vs. video
- Time of day or week different apps are used

Subscribers by Device Type
This data, combined with the Type Allocation Code (TAC) listing service providers already possess, will further identify the device make and model. Examples of queries are as follows:

- Query to enrich the TAC id and device id (IMEI) with device make and model
- Query to group by device type and aggregate by subscriber count

These queries together provide a breakdown of unique subscribers per device type, to identify the most popular devices used by the subscribers. This can be further broken down by region for regional promotions, by the Gigamon labeling capability.

Volume by Device Type or Application
Additional analysis of the devices can be performed to understand data volume by device, and/or by popular application types. This can be achieved by aggregating the data volume in the previous queries as the data volume in bytes and the application name.

Customer Experience Scoring

GigaVUE Enriched Metadata provides insights into usage patterns for the Top N over-the-top applications by subscriber, enabling operators to monitor the Key Quality Indicator (KQI) for application throughput for validation of performance from other sources. This data may additionally be combined with other attributes, to derive additional information like packet loss and packet delay, to ensure the quality of experience score for different applications. A few examples are:

- Quality of experience score per subscriber data usage used by the customer satisfaction and the network experience teams
- Quality of experience scores for popular and trending applications for the top devices in the network
- Quality of experience scores for popular and trending applications by region

This KQI combined with other network KQIs of packet loss, latency, and retransmission can be used to compute or validate the quality of experience scores for the subscribers.

Throughput KQI
This data can be used in the following queries:

- Query to group by IMSI and aggregate by application type and by the uplink data volume for a selected time interval
- Query to group by IMSI and aggregate by application type and by downlink data volume for a selected time interval

This provides the KQI for per subscriber, top N application throughput. This KQI combined with network performance KQIs gives a quality of experience score per subscriber in the network. Quality of experience scores are key metrics for analyzing the subscriber experience and are used by different organizations like Customer Satisfaction, and Network Experience teams.

Roaming Subscriber Usage

Roaming subscribers are a key source of revenue for mobile network operators. Roamer-specific data plans can be offered and targeted with deeper understanding of roamer usage patterns. Metadata is a great way to get a summary of subscribers roaming in a network, including how they use the network by application, location, and time.

Volume By Roaming Subscribers

Roaming subscribers can be distinguished from non-roaming subscribers by examining the carrier identifier present in the metadata, as well as the Mobile Country Code identifier.

- Query top applications used by the roamers to identify usage patterns and offer personalized roaming plans
- Query total uplink and downlink volume used by hour of the day to identify peak hours and plan for the network resources.
- Query application mix to build resource profiles for roaming subscribers

Enterprise Customer (Mobile Private Network) Usage Reporting

GigaVUE Enriched Metadata provides detailed usage data about subscribers belonging to Mobile Private Network (MPN). MPNs are often used by large enterprises and government agencies, leveraging a mobile network to provide work-related mobile service for their employees, technicians, or other members. This usage information is an attractive feed of data for which such organizations would be willing to pay, to ensure that usage of such enterprise-provided devices and resources are not being abused.

Application Usage by Subscriber

This record will provide usage information by subscriber, using a few different methods:

- Query to filter by list of subscribers belonging to a MPN, grouped on subscribers, application name, volume, and time
- Query to filter by APN designating a specific MPN, grouped on subscribers, application name, volume, and time
- [5G only] Query to filter by network slice id designating a specific MPN, grouped on subscribers, application name, volume, and time

Detect Users Violating Terms of Service

GigaVUE Enriched Metadata provides detailed usage data about subscribers using the mobile network to run banned applications, or activities which otherwise violate terms of service. Examples of such activities may include:

- Operating Crypto Mining
- Operating Cloud Computing
- Visiting Banned Internet Destinations

Detecting Banned Applications
This record will provide usage information about specific banned applications:

- Query by banned application name, returning a list of subscribers and time of use

Detecting Banned Internet Destinations
This record will provide usage information about specific banned destinations, subject to local laws and regulations:

- Query destination IP against known list of banned IPs, returning a list of subscribers and time of use

C. Revenue Generation Use Cases

Revenue Sharing Agreements

GigaVUE Enriched Metadata provides detailed usage data for specific Over-The-Top (OTT) applications, including volume, subscriber counts, location, and more, and can be leveraged to manage or enforce revenue-sharing agreements with application providers, as well as provide insight into opportunities for new revenue agreements.

Usage by Application Provider
This record will provide usage information by application.

- Query to group by application type and aggregate by the traffic volume.

This can be used to generate reports as per the revenue sharing agreements, or search for new opportunities.

Usage For Regulatory
In addition to revenue sharing opportunities, this same type of query can be used to satisfy government regulations around reporting user traffic to specific websites.

User Profile Data for Large Venues

GigaVUE Enriched Metadata provides detailed usage data about subscribers connected on specific cell sites, which can be mapped to large venues, and combined with other datasets like subscriber profile information, to provide an attractive feed of data for which such venues would be willing to pay. Data may need to be anonymized depending on local laws, end user agreements, and other considerations.

Example venues may include:

- Sporting event arenas and stadiums
- Concerts and other event venues
- Retail shopping malls
- Airports and seaports
- Military bases
- Large warehouses
- University campuses
- Office campuses

Usage at Large Venues
This record will provide usage information, mapped to subscriber, and filtered on cell id, to match specific cell sites tied to a large venue.

- Query to filter on a list of cell ids, grouped by subscriber, and optionally grouped by application type, with aggregation on the traffic volume.

User Activity Data for Application Providers

GigaVUE Enriched Metadata provides detailed usage data about subscribers using specific applications, when combined with other datasets like subscriber profile information, creates an attractive feed of data for which application providers, either the provider of application or their competitors, would be willing to pay. Data may need to be anonymized depending on local laws, end user agreements, and other considerations.

Usage by Application
This record will provide usage information by application.

- Query to filter by application type, group on subscribers, and optionally group on location (ECGI) and aggregate on traffic volume.

This can be used to generate reports or feeds as per the data agreement.

User Activity Data for Advertisers

GigaVUE Enriched Metadata provides detailed usage data about subscribers using specific applications or visiting specific websites or other internet destinations. When combined with other datasets like subscriber profile information, this creates an attractive feed of data for which advertisers would be willing to pay.

Destination Visitation Usage By Subscriber
This record will provide usage information by application.

- Query to group on subscribers, returning application usage, and/or internet destination IP. Optionally group on location (ECGI) and aggregate on traffic volume.

This can be used to generate reports or feeds as per the data agreement.

Conclusion

GigaVUE Enriched Metadata is designed to be a *best value* for data network and user activity, providing a level of detail greater than log-based data, but without the high cost of probe-sourced metadata. This "goldilocks" solution provides tremendous opportunity for data mining, revenue generation, and more. Lastly, the powerful and extensible framework can be customized to meet specialized business needs and serves as a foundation to what is next—powering the emerging business needs of tomorrow.

Support and Services

Gigamon offers a range of support and maintenance services. For details regarding Gigamon Limited Warranty and its Product Support and Software Maintenance Programs, visit gigamon.com/support and-services/overview-and-benefits.

About Gigamon

Gigamon offers a deep observability pipeline that harnesses actionable network-derived intelligence to amplify the power of observability tools. This powerful combination helps IT organizations to assure security and compliance governance, speed root-cause analysis of performance bottlenecks, and lower operational overhead associated with managing hybrid and multi-cloud IT infrastructures. The result: Modern enterprises realize the full transformational promise of the cloud. Gigamon serves more than 4,000 customers worldwide, including over 80 percent of Fortune 100 enterprises, nine of the ten largest mobile network providers, and hundreds of governments and educational organizations worldwide. To learn more, please visit gigamon.com.

[1] Ericsson Mobility Report, June 2023, pg. 20

Gigamon®

Worldwide Headquarters
3300 Olcott Street, Santa Clara, CA 95054 USA
+1 (408) 831-4000 | gigamon.com

05.24_02