International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Low Latency Data Lookups in Real - Time Data Engineering Pipelines

Girish Ganachari

Email: *girish.gie[at]gmail.com*

Abstract: ***Healthcare, finance, and the Internet of Things need real - time data engineering pipelines with low - latency data lookups for processing and decision - making. Incremental ETL pipeline scheduling, FPGA - based hardware acceleration, and on - GPU thread - data remapping minimise latency in this research. It covers low - latency data pipeline applications, benefits, drawbacks, and prospects. It improves flexibility, productivity, and decision - making. To improve real - time data pipelines, the research proposes resource management, hardware - software interaction, and data consistency.***

Keywords: Low - latency, Real - time data, Data engineering pipelines, FPGA acceleration, Incremental ETL, On - GPU processing, Data consistency

1. Introduction

Real - time data engineering pipelines need low - latency data lookups for quick decisions and insights. These pipelines handle continuous data for fast banking, healthcare, and IoT responses. Real - time data analysis can save lives in healthcare. Low - latency pipelines provide real - time risk management, high - frequency trading, and huge IoT network monitoring. This article explores low - latency data search algorithms, their benefits and limitations, and how they may enhance real - time data engineering. This extensive study shows how low - latency data lookups may improve productivity, accuracy, and responsiveness in real - time data pipelines across sectors.

2. Methodologies

a) FPGA - Based Hardware Acceleration

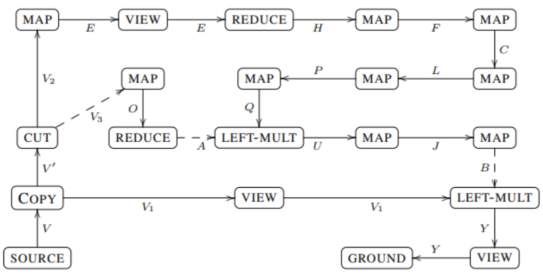

Figure 1.1: Road Boundary Detection

Data processing pipeline latency may be improved using FPGA - accelerated hardware. Using FPGAs may greatly cut down on the length of time that the data flow is delayed [17]. Hardware - accelerated FPGAs minimise lookup and processing time for parallel data processing [3]. Over CPU processing, FPGA processing may reduce latency. This low - latency technology processes massive data for high - frequency trading and real - time analytics.

b) Incremental ETL Pipeline Scheduling

Schedule incremental ETL to accelerate data processing. An incremental ETL pipeline handles data deltas instead of the whole dataset [19]. Selection speeds data retrieval and decision - making with near - real - time data warehouse updates and reduced processing. Dynamic data environments suit incremental ETL data changes. Monitoring financial systems with changing transaction data in real time [4]. These pipelines prioritise incremental updates, enhance speed and efficiency, use less resources, and keep accurate and current data without reprocessing. Incremental ETL improves real - time data engineering and work management.

c) On - GPU Thread - Data Remapping

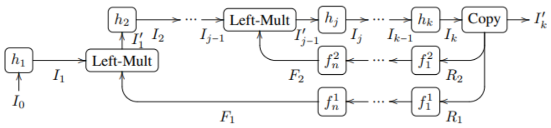

Figure 1.2: Multi - loop feedback – control system for controlling a set of operators in the streaming pipeline

Low - latency packet processing using On - GPU Thread - Data Remapping [10]. GPU parallel processing handles large data packets. GPU thread remapping accelerates real - time pipelines. Large - scale simulations and video streaming benefit from fast data processing.

d) RASP: Real - Time Network Analytics

RASP, a complicated system that swiftly stores and aggregates incoming and external data using NoSQL engines and distributed stream processing [23]. Storm, Kafka, HBase, and Phoenix accelerate network analytics and real - time execution in this novel method. RASP reduces data search latency and provides rapid insights from current network systems' high - velocity data [23]. System design assesses large streaming data sets fast.

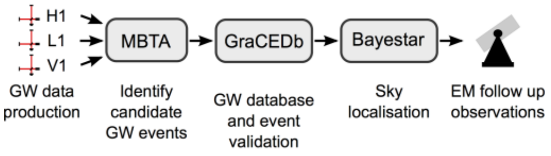

Figure 1.3: Overview of the GW - EM follow - up pipeline

Strong, versatile data analytics solutions from RASP use contemporary technology that Improves network data processing speed and accuracy [20]. It helps cybersecurity fight attacks and telecom manage network performance and traffic. Storm for real - time processing, Kafka for high - volume data intake, and HBase and Phoenix for scalable storage make RASP strong for contemporary network analytics in quickly changing data settings [23].

e) Edge - Based Data - Driven Pipelines

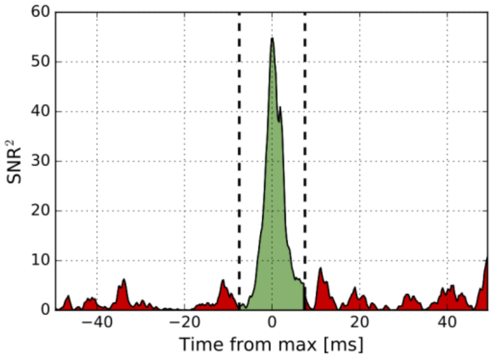

Figure 1.4: SNR2 time series from the matched filter output for a simulated GW in the presence of noise

The dependence of solutions on the periphery Improved velocity for transferring data in the cloud [20]. IoT technologies generate data peripherally and offer benefits. By processing data locally, edge - based industrial automation and smart city pipelines minimise latency and server burden.

f) Multi - Band Template Analysis for Gravitational Waves

Multi - Band Template Analysis gravitational wave detection [1]. MBTA proves low - latency scientific study by detecting gravitational waves in less than a minute at comparable processing costs [5]. This system detects and understands unusual events in real time via rapid data processing.

g) Low - Latency 3D Genomic Data System

A quick, big database for storing, retrieving, and presenting three - dimensional genetic data is presented by [2]. Fast technology is essential for genetic research. DB and 3DGB browser are system elements. For advanced genomics research, this approach accelerates genetic data availability.

h) Low - Latency Event - Based Filtering

Dynamics vision sensors use FPGAs for real - time event - based filtering and feature extraction [11]. Vision system responsiveness is improved by lowering visual data stream interpretation time. Event - based filtering detects and understands video events in real time for surveillance and autonomous automobiles.

i) Stream Algebra for Computer Vision Pipelines

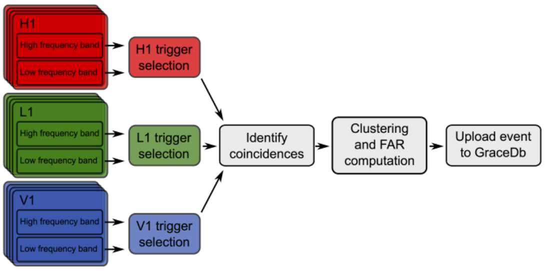

Figure 1.5: Overview of the MBTA pipeline

The stream algebra increased global computer vision pipeline efficiency [8]. This algebra speeds video and image analysis. Computer vision pipeline data flow optimization increases visual input processing.

3. Applications

a) Healthcare

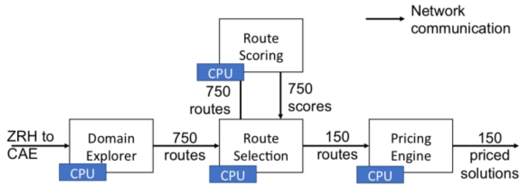

Figure 1.6: The Flight Availability Search and Pricing engine

Live data pipelines monitor alerts and react to healthcare issues. FPGA - based real - time patient data processing and assessment may speed up medical procedures [17]. Every minute counts in critical care, thus this instrument is crucial [6]. Clinicians may employ real - time data pipelines for telemedicine, tailored care, and remote monitoring.

b) Finance

High - frequency traders benefit from low - latency data searches. Instant data pipelines expedite financial processes and improve traders' market decisions. Real - time incremental ETL pipelines speed financial dataset processing [19]. Fast financial data pipelines provide consumer analytics, risk management, and fraud detection.

c) IoT Systems

In IoT systems, the amount of data generated by sensors and devices is enormous. Real - time data pipelines are needed to analyse this data. Edge - based architecture by Renart et al. (2019) analyses data at its source [21]. This technology reduces central computer data transfer latency for real - time environmental monitoring, industrial automation, and smart city decisions.

d) Scientific Research

Science demands fast analysis and discovery utilising low - latency data pipelines. The Multi - Band Template Analysis (MBTA) pipeline detects gravitational waves in real time to respond to significant astronomical events [1]. This ability is needed to capture transitory experiences and make later observations that may lead to new discoveries. Low - latency data systems let genomic researchers swiftly access enormous genetic data sets.3D genomic data querying and visualisation for advanced study and discoveries [2]. Data allows tailored therapy and genetic abnormalities research. Real - time environmental data analysis helps climate specialists prepare for natural disasters, hence low - latency pipelines are crucial [20]. Innovation in data processing accelerates discovery and increases accuracy and reliability across disciplines.

e) Autonomous Systems

Drones and self - driving automobiles need low - latency pipelines for sensor data. These systems can swiftly detect and respond to environmental changes using event - based filtering and feature extraction [11]. Safe and successful autonomous operations need this.

f) Multimedia Processing

Real - time image processing, AR, and video streaming need low - latency multimedia processing data pipelines. Improved multimedia pipeline latency, data flow, and user experience using stream algebra [8]. VR, live streaming, and interactive gaming need real - time data processing.

4. Benefits

a) Enhanced Decision - Making

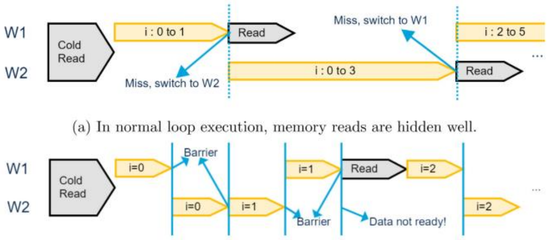

Figure 1.7: Effect of per - iteration synchronisation on the scheduling of two wavefronts (W1 and W2). A read by either wavefront brings in data for two iterations

Many firms use low - latency data searches to make decisions. Real - time patient data analysis saves lives and improves results [17]. Healthcare providers can quickly modify treatment plans or start life - saving medications in emergencies because to their fast access to current information [10]. A low - latency data pipeline can provide investors real - time market data and better decisions. By helping traders capture fresh opportunities, this knowledge reduces market risk and enhances profits [19]. Logistics monitors shipment data in real time to ensure delivery and supply chain management so this firm needs quick data retrieval [20]. Low - latency data pipelines let decision - makers monitor and optimise operations across domains using current and accurate data [8] [24]. Strategic planning, operational efficiency, and corporate growth increased by real - time choices.

b) Improved Efficiency

Data pipelines that save time and operate in real time are more efficient. Faster data processing and transactions enhance bank earnings. Real - time data analysis is a method that may be used to improve Internet of Things (IoT) systems [20]. Effective data processing saves money and resources.

c) Scalability

Real - time scalable data pipelines handle more data without slowing. The Internet of Things and banking require scalability due to data growth [19]. Scalable data pipelines let companies add sources and technology to meet demand.

5. Challenges

a) Data Consistency

High - velocity data streams may cause real - time pipeline data consistency issues. Data inconsistency may hamper analysis. Effective consistency solutions are required [2]. This can be handled via data versioning, eventual consistency, and dispute resolution.

b) Hardware and Software Integration

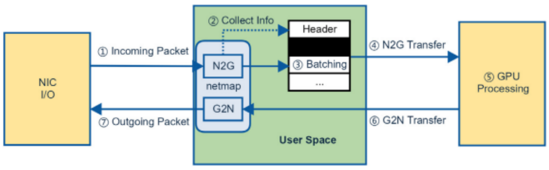

Figure 1.8: System Overview of Blink

Integrating FPGAs and GPUs with software frameworks takes time and talent. Hardware - software communication must be seamless to reduce latency [10]. Driver updates, hardware optimisations, and compatibility issues must be managed during integration.

c) Resource Management

Managing computer resources for shifting data loads without lag may be tricky. Dynamic resource allocation optimises workload performance [17]. Pooling, load balancing, and auto - scaling optimise resources and decrease latency.

d) Security and Privacy

Security and privacy of real - time pipeline data are tricky. Instant data processing and transmission risk security breaches and unauthorised access. Securing sensitive data and satisfying regulations requires robust encryption, access restrictions, and monitoring.

6. Future Directions

a) Advanced Hardware Acceleration

For complex data, research should focus on fast hardware accelerators. FPGA and GPU improvements will aid this breakthrough [10]. Neuromorphic and quantum computing may reduce latency by processing data quicker.

b) Enhanced Data Consistency Mechanisms

More reliable and economical data consistency solutions are required for real - time pipelines [12]. Study consistency zones and thread coordination, proposed by Qu and Deßloch [19]. New consistency models and distributed consensus may improve real - time data consistency.

c) Integration of Edge and Cloud Computing

Cloud - based edge computing may reduce latency by processing data locally. A hybrid system for the distribution of workloads between the edge and the cloud [20]. This link may boost real - time data pipeline reliability, efficiency, and scalability.

d) Machine Learning Integration

Real - time data pipelines with machine learning models may enhance predictive analytics and automated decision - making [16]. These models should be improved for applications that need minimal latency [8]. Online learning, federated learning, and edge AI may improve data pipelines and real - time analytics.

e) Real - Time Data Visualization

Powerful real - time data visualisation technologies may help users gain insights and make quicker decisions [7]. Live data is available via interactive dashboards, AR interfaces, and real

- time notifications. Innovative visualisation and UI technologies may improve data analysis.

f) Integration with Blockchain Technology

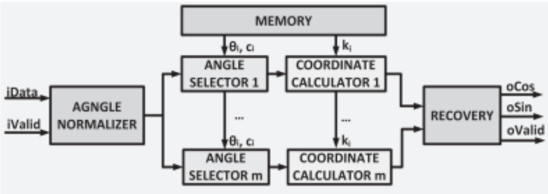

Figure 1.9: The hardware architecture of parallel pipeline CORDIC

Blockchain - powered real - time data pipelines may boost security, integrity, and transparency [15]. Permanent blockchain technology may secure scattered network data and transaction records [9]. Blockchain - integrated real - time data pipelines may need consensus and scalable platforms.

7. Conclusion

Some commercial applications need real - time data engineering pipelines with low - latency data retrieval. Latency may be reduced via hardware acceleration, thread - data remapping, and progressive ETL pipeline scheduling. This problem needs ongoing study to enhance decision - making, productivity, and scalability. Real - time data pipelines benefit from resource management, hardware - software interaction, and data consistency. AI, blockchain, and fast hardware accelerators allow real - time data engineering. Application and use vary per connection.

References

[1] Adams, T., Buskulic, D., Germain, V., Guidi, G. M., Marion, F., Montani, M., Mours, B., Piergiovanni, F. and Wang, G., 2016. Low - latency analysis pipeline for compact binary coalescences in the advanced gravitational wave detector era. *Classical and Quantum Gravity*, *33* (17), p.175012.

[2] Butyaev, A., Mavlyutov, R., Blanchette, M., Cudré - Mauroux, P. and Waldispühl, J., 2015. A low - latency, big database system and browser for storage, querying and visualization of 3D genomic data. *Nucleic acids research*, *43* (16), pp. e103 - e103.

[3] Cappellari, P., Roantree, M. and Chun, S. A., 2017. A scalable platform for low - latency real - time analytics of streaming data. In *Data Management Technologies and Applications: 5th International Conference, DATA 2016, Colmar, France, July 24 - 26, 2016, Revised Selected Papers 5* (pp.1 - 24). Springer International Publishing.

[4] Doherty, C. and Orenstein, G., 2015. Building Real - Time Data Pipelines.

[5] Goel, A. K., Pound, J., Auch, N., Bumbulis, P., MacLean, S., Färber, F., Gropengiesser, F., Mathis, C., Bodner, T. and Lehner, W., 2015. Towards scalable real - time analytics: An architecture for scale - out of OLxP workloads. *Proceedings of the VLDB Endowment*, *8* (12), pp.1716 - 1727.

[6] Guo, C., Yuan, L., Xiang, D., Dang, Y., Huang, R., Maltz, D., Liu, Z., Wang, V., Pang, B., Chen, H. and Lin, Z. W., 2015, August. Pingmesh: A large - scale system for data center network latency measurement and analysis. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication* (pp.139 - 152).

[7] Gupta, A. and Shute, J., 2019. High - availability at massive scale: Building google's data infrastructure for ads. In *Real - Time Business Intelligence and Analytics: International Workshops, BIRTE 2015, Kohala Coast, HI, USA, August 31, 2015, BIRTE 2016, New Delhi, India, September 5, 2016, BIRTE 2017, Munich, Germany, August 28, 2017, Revised Selected Papers 9* (pp.63 - 81). Springer International Publishing.

[8] Helala, M. A., Qureshi, F. Z. and Pu, K. Q., 2020. A stream algebra for performance optimization of large scale computer vision pipelines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44* (2), pp.905 - 923.

[9] Kipf, A., Pandey, V., Böttcher, J., Braun, L., Neumann, T. and Kemper, A., 2019. Scalable analytics on fast data. *ACM Transactions on Database Systems (TODS)*, *44* (1), pp.1 - 35.

[10] Lin, H. and Wang, C. L., 2019. Efficient low - latency packet processing using On - GPU Thread - Data Remapping. *Journal of parallel and distributed computing*, *133*, pp.51 - 62.

[11] Linares - Barranco, A., Perez - Peña, F., Moeys, D. P., Gomez - Rodriguez, F., Jimenez - Moreno, G., Liu, S. C. and Delbruck, T., 2019. Low latency event - based filtering and feature extraction for dynamic vision sensors in real - time FPGA applications. *IEEE Access*, *7*, pp.134926 - 134942.

[12] Liß, C., Ulbricht, M., Zia, U. F. and Müller, H., 2017, June. Architecture of a synchronized low - latency network node targeted to research and education. In *2017 IEEE 18th International Conference on High Performance Switching and Routing (HPSR)* (pp.1 - 7). IEEE.

[13] Meehan, J., Aslantas, C., Zdonik, S., Tatbul, N. and Du, J., 2017, January. Data Ingestion for the Connected World. In *Cidr* (Vol.17, pp.8 - 11).

[14] Najafi, M., Zhang, K., Sadoghi, M. and Jacobsen, H. A., 2017, June. Hardware acceleration landscape for distributed real - time analytics: Virtues and limitations. In *2017 IEEE 37th international conference on distributed computing systems (ICDCS)* (pp.1938 - 1948). IEEE.

[15] Nguyen, H. T., Nguyen, X. T. and Pham, C. K., 2017. A low - latency parallel pipeline cordic. *IEICE Transactions on Electronics*, *100* (4), pp.391 - 398.

[16] Nidever, D. L., Holtzman, J. A., Prieto, C. A., Beland, S., Bender, C., Bizyaev, D., Burton, A., Desphande, R., Fleming, S. W., Pérez, A. E. G. and Hearty, F. R., 2015. The data reduction pipeline for the apache point observatory galactic evolution experiment. *The Astronomical Journal*, *150* (6), p.173.

[17] Owaida, M., Alonso, G., Fogliarini, L., Hock - Koon, A. and Melet, P. E., 2019. Lowering the latency of data processing pipelines through FPGA based hardware acceleration. *Proceedings of the VLDB Endowment*, *13* (1), pp.71 - 85.

[18] Patra, P. G. K., Cesen, F. E. R., Mejia, J. S., Feferman, D. L., Csikor, L., Rothenberg, C. E. and Pongracz, G.,

2018. Toward a sweet spot of data plane programmability, portability, and performance: On the scalability of multi - architecture P4 pipelines. *IEEE Journal on Selected Areas in Communications*, *36* (12), pp.2603 - 2611.

[19] Qu, W. and Deßloch, S., 2017. Incremental ETL pipeline scheduling for near real - time data warehouses.

[20] Renart, E. G., Balouek - Thomert, D. and Parashar, M., 2018. Edge based data - driven pipelines (technical report). *arXiv preprint arXiv: 1808.01353*.

[21] Renart, E. G., Balouek - Thomert, D. and Parashar, M., 2019, May. An edge - based framework for enabling data - driven pipelines for iot systems. In *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp.885 - 894). IEEE.

[22] Schales, D. L., Hu, X., Jang, J., Sailer, R., Stoecklin, M. P. and Wang, T., 2015, April. FCCE: highly scalable distributed feature collection and correlation engine for low latency big data analytics. In *2015 IEEE 31st International Conference on Data Engineering* (pp.1316 - 1327). IEEE.

[23] Touloupas, G., Konstantinou, I. and Koziris, N., 2017, December. RASP: Real - time network analytics with distributed NoSQL stream processing. In *2017 IEEE International Conference on Big Data (Big Data)* (pp.2414 - 2419). IEEE.

[24] Zhang, K., Hu, J. and Hua, B., 2015. A holistic approach to build real - time stream processing system with GPU. *Journal of Parallel and Distributed Computing*, *83*, pp.44 - 57.

[25] Zhang, P., 2019. Performance Analysis of Cloud - Based Stream Processing Pipelines for Real - Time Vehicle Data.