



www.ociweb.com

High Volume Data Ingestion: A Best Approach to “Big Data”.

You have Data-- Lots and Lots of Data! High Volume Data Ingestion is the Next Best Step in Turning Big Data into Meaningful Business Intelligence.

The Answer is High Volume Data Ingestion! Largely the focus of just a few small technical circles until recently, High Volume Data Ingestion includes a series of solutions that prepare and optimize Big Data so that it may be analyzed, interpreted and applied in near real-time.

WHAT CAUSES BIG DATA BOTTLENECKS?

- Multiple clients trying to write/read to a single server, creating a “pinch point”
- Larger messages, which take up bandwidth
- Heavy-weight compression algorithms, which use up processor capacity
- Different systems using different data formats
- Hardware limitations

High Volume Data Ingestion Solutions

Some High Volume Data Ingestion Solutions include:

1. **Semantic Data Compression.** Today’s data compression solutions are being used sparingly (by using a message’s fixed structure or schema only) if at all. The Semantic Data Compression solution takes advantage of knowledge about the structure and context of data streams. This compression yields “leaner” messages and better prioritization of what data is sent.
2. **Cloud-Based Ingestion Gateways.** Secure storage/retrieval connections are managed by an ingestion gateway. These gateways have traditionally been a Big Data bottleneck. By virtualizing gateways and placing them in a Cloud, Big Data across multiple sources becomes more manageable, improves overall system stability, and eliminates the need for server load-balancing.
3. **Elastic Services.** Instead of relying on monolithic applications that do *everything*, elastic services (or “microservices”) are smaller software applications that may be used independently. Thus, individual parts of the system can be rolled-out, scaled-up, or upgraded as needed, in real time. This prevents delays and allows a smarter allocation of tech resources.

4. **Scalable Databases.** General-purpose databases (as opposed to databases designed to address a particular purpose) scale with need and enable organizations to accommodate future growth (and significantly reduce time/space/financial investment).
5. **On-the-Fly Data Analysis.** Traditionally, data analysis consists of a serial process, whereby data is (1) extracted from relevant sources, (2) normalized and cleaned, (3) transferred to a central database, and (4) loaded for use (also known as Extract/Transfer/Load, or ETL). Since these steps do not overlap, there is a considerable lag between extraction and use. More modern methods perform data analysis “on the fly,” whereby data is processed as soon as it is available, while other data is being gathered and prepared. This allows for the use of information in near real-time.
6. **Data Normalization.** Data normalization combines variants of the same bit of data, thereby reducing data redundancy (which in turn speeds up transfer and processing), and producing more accurate analytics.

WHERE ARE HIGH VOLUME DATA INGESTION SOLUTIONS VALUABLE?

- ✓ Financial services firms
- ✓ Energy Companies/Utilities
- ✓ Government and Defense
- ✓ eCommerce sites
- ✓ Smart Warehouses
- ✓ Social Media
- ✓ Mobile media applications
- ✓ Any organization with unstructured or semi-structured data

Getting Started with High Volume Ingestion

As described above, High Volume Ingestion Solutions optimize the flow and preparation of Big Data. These solutions are just a handful of the myriad options available to business to more quickly and accurately transform and monetize the vast business intelligence sprinkled throughout Big Data.

OCI – We are software engineers.

We build high performance, real-time, mission critical middleware systems and integration solutions. Our goal is to make solutions more open, scalable, reusable, interoperable, and affordable. Please visit www.ociweb.com to learn more about our engineering services, open source middleware technologies, and professional IT training.

If you would like to learn more about High Volume Ingestion Solutions and which Solutions are best for your organization, please contact us today.

12140 Woodcrest Executive Drive, Suite 250 Saint Louis, MO 63141, MO
tel: 01*314*579*0066
email: info@ociweb.com

© Copyright Object Computing, Inc. 1993, 2015. All rights reserved.

