**Logistics at the Speed of Data**

**Frank Blau – Data Architect**

**Current State of Data in Logistics**

* *Lots of inefficient manual processes*
* *Reliance on enterprise solution providers for development*
* *Islands of innovation*

Transportation networks and their management are becoming more complex and demanding for vendors, partners and consumers of the logistics software and services that support them. The volume, variety and velocity of the data required to operate within this ecosystem is also growing, as are the demands to develop rapid decision-making and optimization capabilities from this data. While the traditional worlds of traditional operational and analytical data are struggling to meet these demands, there are remarkable but disruptive technologies and techniques emerging that enable transformative solutions to be developed and leveraged to meet these rapidly evolving requirements.

As these technologies disrupt and enable logistics software and hardware vendors, there are still clear needs to increase both the reliability and the competitive quality while remaining cost-effective and trustworthy partners to the service consumers. While there have been significant gains and visible successes to this contemporary challenge, the logistics industry is still dependent on a lot of inefficient, labor-intensive traditions (paper-based and manual data collection, for example) that serve as the deep blue ocean around emerging islands of innovation. This paper will address some of these trends and opportunities within this fluid landscape and provide some mitigation strategy concepts for further exploration.

**Trends in Logistics Data**

* *Increased Networking Demands*

The networks and partnerships within a logistics ecosystem are becoming increasingly more complex, data-driven and competitive. The creation and operational management of these networks require the use of intelligent, and in many cases, automated approaches to their optimization within the supply chain.

*Increased regulatory demands*

*Growth in cargo*

*Adoption of clean technologies*

*Lower latency analytics*

**Disruptive Technologies**

*Cloud-based platforms*

The deployment of services, such as computing, storage and applications cannot continue to be locked within the hallowed crypt of the on premise data center. The demands for scalability, rich service tools, and vital elasticity of cloud-based resources makes the migration to a cloud platform a near imperative architectural choice for the deployment of logistics software. To deploy even something as “simple” as data gateways to 3rd party partners, is something that can take weeks or even months to provision within a traditional data center. Servers must be specified, ordered, racked, cabled, powered, cooled and configured in a process that consumes resources and budgets. On a cloud platform, these data gateways can be provisioned and integrated within hours, with their endpoints ready for integration in days.

The cloud also offers the use of highly elastic compute resources. This means that if a data scientist needs to process a significant calculation, say a network route optimization algorithm, it is not necessary to purchase the hardware to accomplish that and hope for cost recovery in the outcome. With elastic-compute, the data scientist can define, spin up and make use of a massive virtual cluster, that only needs to exist for the lifetime of the computation and then deliver the result to cost-effective analytic or storage service. In some cases, this can be provisioned down to the milliseconds of compute time needed to compute the result and no more.

Storage costs are also much cheaper on cloud platforms than on-premise discs will ever be. It isn’t just they are cheaper though, they are also more readily available, redundant, secure and scalable as needs increase. It is also possible to migrate data from “hot” high-speed data services as it is consumed, and then move to both “cooler” and cheaper platforms as the use of that data becomes less time-critical and the long term storage needs are realized.

*Schema-less data*

While the traditional relational database management system (RDBMS) will remain a key component of both the transaction processing (OLTP) and analytic (OLAP) solution playbook, there is an increased adoption of storage and processing architectures that are less dependent on a rigid database schema to provide rapid consumption and actionable insight from the operational data being created by logistics software. From open source tools like MongoDB, Cassandra and Redis, to the highly scalable enterprise tools from Crate, IBM, Amazon and Microsoft, there are many new approaches to data architecture that offer far less assumption of technical debt when developing and deploying data storage and analytic capabilities.

*Intelligent File Systems and Storage, the Data Lake*

File storage has evolved from expensive traditional storage on proprietary hardware and appliances, to faster, more intelligent and inexpensive devices within the data center. The rapid adoption of the Hadoop Distributed File System (HDFS) running on inexpensive commodity hardware, has provided a rich, highly available storage solution for up to petabyte scale demands. HDFS offers the combined benefits of both a low-cost storage solution and a mature data exploration and analysis platform that has both open source and enterprise vendor options for deployment. This has led to the adoption of the term Data Lake, a repository of data that was previously too expensive or unstructured to even consider storing for data mining and analysis. Given the growing volumes from sources like IoT and social media, the data lake is an essential landing zone for data that needs be retained for further exploration and consumption.

*Microservices and APIs*

*Blockchain*

Blockchain is emerging as perhaps the most-hyped but also the most promising technology for the logistics industry. The remarkably simple design pattern of an encrypted, immutable shared ledger and Smart Contract applications is proving to be a fertile inflection point for a dizzying array of innovative use cases for logistics. Developers and architects are investigating and socializing capabilities that were previously only imagined within the framework of expensive, monolithic applications. Use cases can now be explored and prototyped at a fraction of the cost of traditional logistics software solutions. This opportunity has not gone unnoticed by some of the largest players in the logistics world; From IBM’s $13 million deal with Maersk to the development partnership between Nestle and Walmart, there is considerable attention and resources being aligned with Blockchain-based opportunities.

The biggest challenge for Blockchain in the logistics industry is not a technological challenge. There are so many possible use cases that it is sometimes overwhelming to winnow out the ones that will provide the most business value and fit the design pattern for Blockchain applications.

Blockchain is also a relatively (4-5 years) new technology. There are still a lot of standards, protocols and software development tools that need to align into a coherent strategic platform to provide the secure and resilient transactional competence that is required of logistics software. But the time to investigate and understand this platform’s potential and patterns is today, as many of the leading logistics software players already are already heavily invested in this adoption.

*Internet of Things*

The Internet of Things (IoT) is the convergence of sensors, networks and data that is resulting in an almost untethered growth in the amount of data being generated about the location, usage and environmental conditions of logistics components. From GPS-equipped shipping containers to embedded environmental sensors, it is now possible to monitor the exact location, condition and status of parcels, package and transport vehicles as they move through the supply chain.

The deployment, monitoring and consumption of the unstructured and high velocity data from these devices is itself a significant challenge in today’s logistics world. It is abundantly clear that IoT represents one of the most important and impactful disruptive technologies for logistics in the coming decade. Compounding this challenge (and opportunity) is the sometimes-bewildering array of technologies and tools emerging to meet this challenge. From rapidly-evolving open source tools to the no-longer tentative steps of enterprise vendors into the IoT space, the decision to land on an IoT solution that is both flexible and scalable is a non-trivial process for logistics companies struggling to even understand this new kind of data. It is comforting that the large enterprise cloud vendors mentioned above have all been evolving their IoT platforms to meet this critical need for logistics.

*Artificial Intelligence and Machine Learning*

*Unmanned delivery platforms*

**Mitigation Strategies**

*Use Case Development for Pilot Projects*

*Collaborative development*

*Silicon-valley thinking*

**Conclusion**