Clusterpoint Monetizes Digital Assets at LETA

Overview

Clusterpoint has been in production since 2010 at LETA (Latvian News Agency); delivering high-speed, online access to their media archive as part of a highly profitable subscription service.

This paid service is targeted at media monitoring agencies, PR firms, researchers and analysts and enables LETA to generate incremental revenues from historical media assets that would otherwise have been seen as pure archive material.

Clusterpoint continues to run at the heart of this service and now serves thousands of subscribers issuing tens of thousands of search and retrieval requests against tens of millions of media assets every day. The service delivers consistent sub-second response times and runs on ten entry-level, off-the-shelf data servers.

The Business Challenge

As a business LETA wanted to monetize access to existing media assets in their archive in the form of a new subscription service that could be cost-effectively deployed at a competitive price point.

To achieve that LETA needed to build out a platform that would provide high performance search and retrieval of the most relevant media assets (including editorial, photo and video content).

The key requirements for the data management component were:

- High performance search and retrieval with superior ranking technology to ensure maximum relevance of results; maximizing value to the subscriber.
- Natural scalability to let them test the market by starting small while growing to hundreds and thousands of users without re-engineering or 'rip and replace'.
- Inbuilt high availability to ensure the service would run 24x7 without disruption; maximizing subscriber satisfaction and protect against churn.
- Deployment on commodity, off-the-shelf hardware to keep infrastructure costs to a minimum and leverage existing datacenter investment.
- Scalable and predictable pricing model to ensure expansion could be funded incrementally from subscriber uptake; avoiding punitive up-front costs
- Wrapped product incorporating all key enterprise features speeding up time-to-market and keeping integration and development costs to a minimum.

The Clusterpoint Answer

LETA initially evaluated an enterprise SQL-based platform but suffered from slow response times, complex setup /administration and a very real shortfall in search functionality (particularly the ability to control ranking at a granular level).

These problems prevented the team from building a compelling service prototype to sell back to the business.

Clusterpoint subsequently worked with two engineers at LETA and delivered a compelling prototype application within one month. The business signed off on the prototype immediately and the initial commercial service was rolled into production less than two months later; initially running on two commodity servers.

Since then the service has expanded organically. It now runs with tens of millions of asset objects under management, thousands of users and tens of thousands of search requests and asset retrievals per day. Al of this is delivered on ten commodity servers in high availability mode and the service highly profitable for LETA.

Key Clusterpoint Features

From a LETA perspective, the following capabilities were core to their decision to choose Clusterpoint.

- **Speed and scalability** - ultra-fast search with the capability to scale out incrementally with zero disruption.
- **Optimal ranking** - ranking of results based on subscriber relevance to ensure best user experience and make subscribers 'sticky'.
- **Faceted search** - fast and simple navigation of related results enabling subscribers to easily explore additional content of interest.
- **Language stemming** - inbuilt capability to search for common inflexions of subscriber search terms e.g. plurals, declensions etc.
- **Multi-structured data support** – integrated management of the variety of asset types commonplace in news and media e.g. editorial, photos, videos.
- **Open standards** – an XML based data store with support for the most commonly used languages through a rich and clearly documented API.
- **Speed of deployment** – a single solution with integrated feature set that enabled fast prototyping and equally fast deployment into production.
- **Predictable economics** – a pricing model and drag-along infrastructure cost that scaled with subscriber uptake and was fundable from revenue