Sumerian Analytics Platform

Our Analytics Platform is uniquely engineered to rapidly transform big data into intelligent insight.

Sumerian's Analytics Platform is built on 10 years of software engineering and data science expertise. It uses a range of proven mathematical algorithms and statistical techniques that have a long and influential history in science, engineering and medicine. We have applied these same founding principles to our Platform and analytics services – enabling today's IT teams to gain business-transforming value from their data and make a progressive leap to evidence-based decision making.

Dealing with Big Data

The IT data that Sumerian models and analyses falls into the category of data that is increasingly referred to today as "big data". Big data provides an unparalleled opportunity for IT organisations to gain a wealth of intelligence about their environments, however its sheer magnitude and complexity makes it difficult to exploit – most notably as it is extremely challenging to process using standard software tools or methodologies, due to the volume, velocity, variety and variability of its nature (The 4Vs of Big Data, Forrester Research, 2011).

Sumerian has been innovating and developing core IP in the big data analytics space since 2002 – long before the term "big data" was even coined. Our Analytics Platform has developed as a result, built on 10 years of solving our clients' big data challenges.

How the platform works

The Sumerian Analytics Platform is highly scalable and able to process vast quantities of heterogeneous data on a daily basis. It has a modular, 4 layer architecture that combines Sumerian's core IP with best-in-class third party technology and follows the Sumerian analytics methodology of Capture, Model, Analyse and Inform.

Capture layer

We have a large library of standard data integrations, and to date have integrated over 150 different data sources, including exports from common infrastructure monitoring tools and web server logs to bespoke application log files. The technology we have invested in means we can rapidly integrate any new source that's required.

To collect data, our Data Collector software sits in your environment and stages your data ready for transfer at whatever frequency is required. The data is then securely transferred via the internet to our datacentre where it is validated, cleansed and normalised by our Data Manager software. For clients whose security policy restricts the transfer of data in its raw form, we also offer data obfuscation and client hosted data capture solutions as alternatives.

Summary value

- **Big data analytics for enterprise IT** – captures, integrates, models and analyses big data from IT systems on an ongoing basis
- **Advanced modelling and analytics** – uses mathematical algorithms and statistical techniques to inform the current state of play across applications and predict future requirements
- **Rapid, repeatable analytics** – delivers ready to action, interactive reporting to drive tangible, quantified business outcomes

The Data Manager mines data from both structured and unstructured data types, extracting key information for our models. Structured data is typically generated by infrastructure monitoring tools, and will provide us with metrics around resource usage, utilisation of CPU, disk and memory on servers, storage, the network and so on. Unstructured data is typically generated by application log files, which provide multiple messages about what the application is doing at any one point in time. For example, these logs will provide us with metrics such as a time stamped, transaction identifiers, and raw text about what the application is processing. Once data has been processed by the Data Manager, it is then automatically transferred on for modelling.

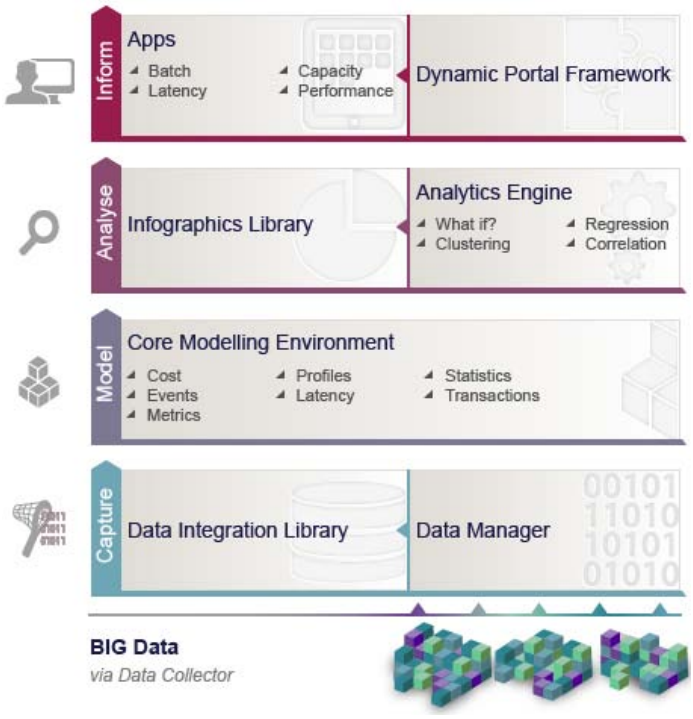

Fig. 1 – Sumerian Analytics Platform Architecture

Model layer

The Sumerian Core Modelling Environment (CME) consists of a focussed suite of interoperable models that can be rapidly applied to accurately model IT applications and environments. The CME is a dimensional data store that provides the ability to rapidly access and manage massive data sets. It is built on OLAP technology, which is both highly scalable and enables us able to deal with the huge volumes of metrics that are required for our analytics.

The dimensional nature of our models allows us to rapidly aggregate and transform data across multiple hierarchies – such as geography, transaction type, time and so on. Our core models have been specifically designed to manage different types of data, including:

- **Metrics** - optimised to efficiently store large volumes of scalar numerical data with configurable aggregations (sum, average, max/min etc.) across multiple dimensions.
- **Statistics** - designed to calculate and store statistical measures (mean, median, variance, standard deviation and so on) as well as empirical distributions across a range of metrics.
- **Events** - models point-in-time information about events that either are generated by applications (e.g. user logins) or are contextual (e.g. market events), enabling these to be used in analysis.
- **Transactions** - sequences of actions and events with a common underlying feature (e.g. a trade or customer identifier) can be joined together to determine information such as duration and path features.
- **Cost** - our cost models are designed to capture and aggregate costs for physical and logical service components, including servers, storage, licenses and personnel.
- **Latency** - optimised to store, aggregate and interrogate response time data across large volumes of transactions, providing a breakdown of time spent in various service components.
- **Profiles** - enables deeper insight into resource consumption and performance by exposing information on user and application features, to answer questions of who uses what functionality, from where and when.

Analyse layer

Once data is modelled, the Sumerian Analytics Engine applies the appropriate analytical techniques and algorithms to the data depending on the analytics required. The engine employs a pluggable architecture that accesses the data sets persisted in one or more of the dimensional models, to execute the required algorithm and compute its results at high-speed, persisting those results again as needed. Sumerian's algorithms are extensible and continually being refreshed to integrate the latest innovations from our development team. The following details a selection of the most commonly used:

- **Correlation** - quantifies the strength of relationships between any pair of metrics. For example, correlation enables Sumerian to determine which demand variables are having an impact on resource utilisation, and ranks these in order of impact if there are multiple demands.
- **Regression** - quantifies the nature of the relationship between sets of metrics. For example, a regression between a demand variable such as volume of concurrent users and a resource variable such as CPU utilisation will determine the amount of CPU required to support each additional concurrent user, thereby enabling headroom to be predicted.
- **Queuing** - allows us to quantify the relationship between throughput and delay, either at the level of a single node or across full end to end transaction processing. This capability allows situations where a seemingly underused resource, such as CPU, is nevertheless adversely impacting performance (e.g. by introducing latency into a trading flow).
- **Chargeback** - allows us to apply configurable cost allocation rules to data covering service usage, asset costs or business entities and profiles. This advanced technique enables accurate chargeback based on consumption or other measures.
- **Predictive "What if?" scenario modelling** - all of the above algorithms are automated and configured to run across any portion of data and provide a basic level of predictive modelling, such as increases in business volumes or a change in the mix of transactions.

For each of our dimensional models and algorithms, we have a suite of associated visualisations that are applied automatically from our Infographics Library. The Library is constantly being updated with new and enhanced visualisations, designed to convey simple "at a glance" understanding of complex data findings. For example, Sumerian's Capacity Flowpipe visualisation (see below) distils millions of correlated data points from terabytes of heterogeneous data down to one simple graphic.

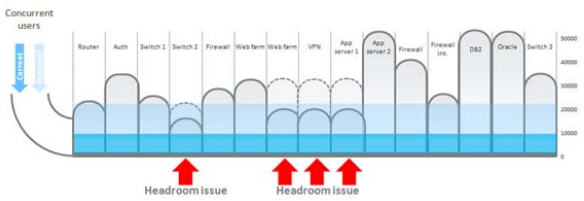

Fig. 2 – Sumerian Capacity Flowpipe visualisation

All Sumerian's visualisations are automatically fed to a dedicated, secure Web portal for immediate access as soon as data is analysed.

Inform layer

As part of our service we host a dedicated, secure web portal that allows quick access to interact, explore and export your modelled and analysed data. The portal uses our Dynamic Portal Framework and suite of analytics Apps, to present a range of targeted views. Apps currently available include:

Capital Markets

- Latency Analytics
- Performance Analytics
- Capacity Analytics
- Batch Analytics
- Fill Rate Analytics
- Grid Analytics

Enterprise Cloud

- Capacity Analytics
- Performance Analytics
- Usage Analytics
- Invoice Verification
- Chargeback Analytics

The portal contains a range of tools that make it easy for you to export data in Microsoft Excel, Word and PDF formats. Our powerful Clipbook facility allows you to create your own PowerPoint presentations, automatically exporting the charts and associated commentary into PowerPoint.

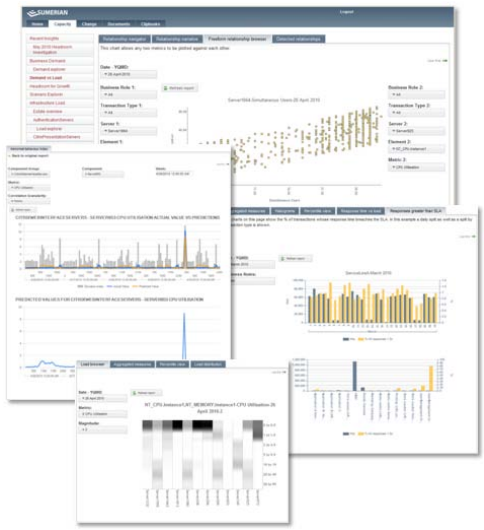

Fig. 2 – Sumerian portal views

The portal also facilitates a collaborative environment where you can work in conjunction with Sumerian's Data Scientists – allowing you to access support as and when you need it. Our team is also on hand to help configure the portal to your exact requirements.

More information

For further information on Sumerian or to arrange a demonstration of our services, contact us on 0141 229 7580, e-mail us at **info@sumerian.com** or visit our Web site at **www.sumerian.com**