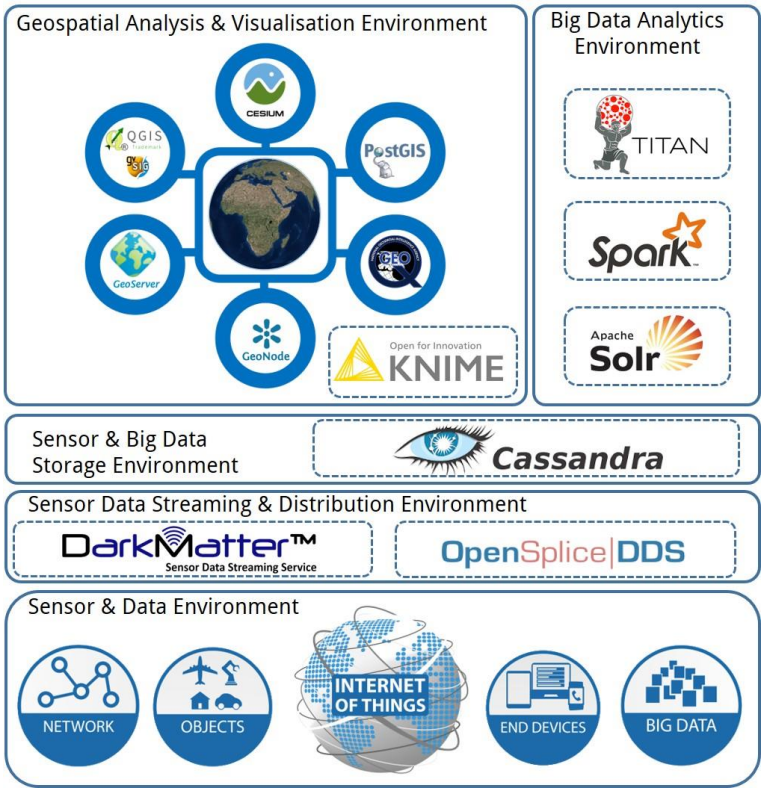
Geo-spatial & Data Framework

IoT Stack Framework

Location Intelligence is a capability to visualise geo-spatial data in order to analyse and identify patterns and trends - it brings together geospatial analysis and business intelligence. It enables organisations to view and analyse data in space and time. Making sense of geospatial data patterns has become a crucial element in the Internet of Things and the resulting Big Data it is generating. Location is a nexus factor of intelligence platforms driving decision-making in business-critical processes and operations.

Value opportunities are created for businesses by linking sensors, end-devices, objects, data networks and analysing the resulting Big Data it is generating. In order to provide solutions for our clients in harnessing these opportunities and creating business value we have created a product framework to support the "Thing Stack" as coined by PwC.



Cost is always a factor, and for this reason we have selected best-of-breed open source software as building blocks in this framework. We have created very successful and unique systems for some of our clients through the integration of these technologies into solutions to harness the IoT. The only proprietary element in our framework is DarkMatter™ an in-house developed sensor agnostic data streaming software service.

Big Data Storage & Analytics

Apache Cassandra™ is a proven high performance open source NoSQL database. It is fault-tolerant, data is automatically replicated to multiple nodes, with no single point of failure. Cassandra™ is highly durable and elastic with read and write throughput increasing linearly as new hardware nodes are added. Due to its high write speeds it is specifically suited for sensor data storage requirements. DarkMatter™ is fully integrated with Cassandra as a standard storage environment for the sensor agnostic data streaming service.

Apache Solr™ is a highly scalable and fault tolerant open source enterprise search server built on Apache Lucene™. It is optimised for high volume traffic and has standards based open interfaces for JSON, XML, PHP, Python and REST API's amongst other. One of its key features is its geo-spatial search capability that includes multiple points per document and polygons. It pairs very well with Cassandra™ and has the same linear scalability. It is a search engine of choice for major Hadoop distributions as well as commercial and open source content management systems. We have found that SOLR has excellent capabilities in use as a standalone NoSQL data store in analytical projects as well.

Apache Spark™ is a Big Data processing framework focussing on sophisticated analytics, speed, and ease of use. Spark differs from Hadoop in that it is a data-processing tool that functions on top of distributed data collections. It has four main libraries:

- Spark SQL - exposes Spark datasets over JDBC API to allow SQL like queries
- Spark Streaming - for performing real-time streaming analytics
- Spark Machine Learning Library (MLlib) - distributed machine learning framework which includes summary statistics, hypothesis testing, regression, cluster analysis, principle component analysis (PCA) and other.
- Spark GraphX - an API for graphs and graph-parallel processing.

Apache Spark™ supports a host of languages such as Java, Python, Scala and R. It requires a cluster management tool (Mesos or Yarn) and interfaces seamlessly with Cassandra for distributed data storage.

Titan is a distributed graph database that interfaces very well with Cassandra™, Solr™ and Spark™, making it the ideal platform for graph based analytics in our "Thing Stack" framework. It builds on top of and is linearly scalable like Cassandra (distribution & replication) and allows for thousands of concurrent users performing complex graph queries. It differs from Spark GraphX in the way it stores nodes an edges (vertices & edges) of network visualisations (graphs). It is distributed under the Apache 2 licence. Titan forms a key element in our Activity Based Intelligence solutions.

Harfield Village Centre, 48 2nd Avenue, Claremont, South Africa

Tel: +27(0)21 683-3624
Fax: +27(0)21 683-3626

www.suritecgeospatial.co.za
Info@suritecgeospatial.co.za