

Quest



CANACCORD|Genuity



ngg_shortcode_0_placeholder

GCP Case Study

Read the Google Cloud blog post: “Quest for a better investment platform leads to Google Containers, BigQuery and Sheets”

Presentation Download

- SenTai Canaccord Quest Migration – EN (Google Slides)
- SenTai Canaccord Quest Migration – EN (ppt)
- SenTai Canaccord Quest Migration – EN (pdf)
- SenTai Canaccord Quest Migration – FR (Google Slides)
- SenTai Canaccord Quest Migration – FR (ppt)
- SenTai Canaccord Quest Migration – FR (pdf)

Introduction

Quest® is a B2B Business Intelligence financial web application, owned and run by Canaccord Genuity Ltd. It is a proprietary analytical research system utilising historical

financial accounting data and consensus forecast data. Its end users are typically asset managers, using it as an equity analysis and valuation tool, for information and decision making purposes.

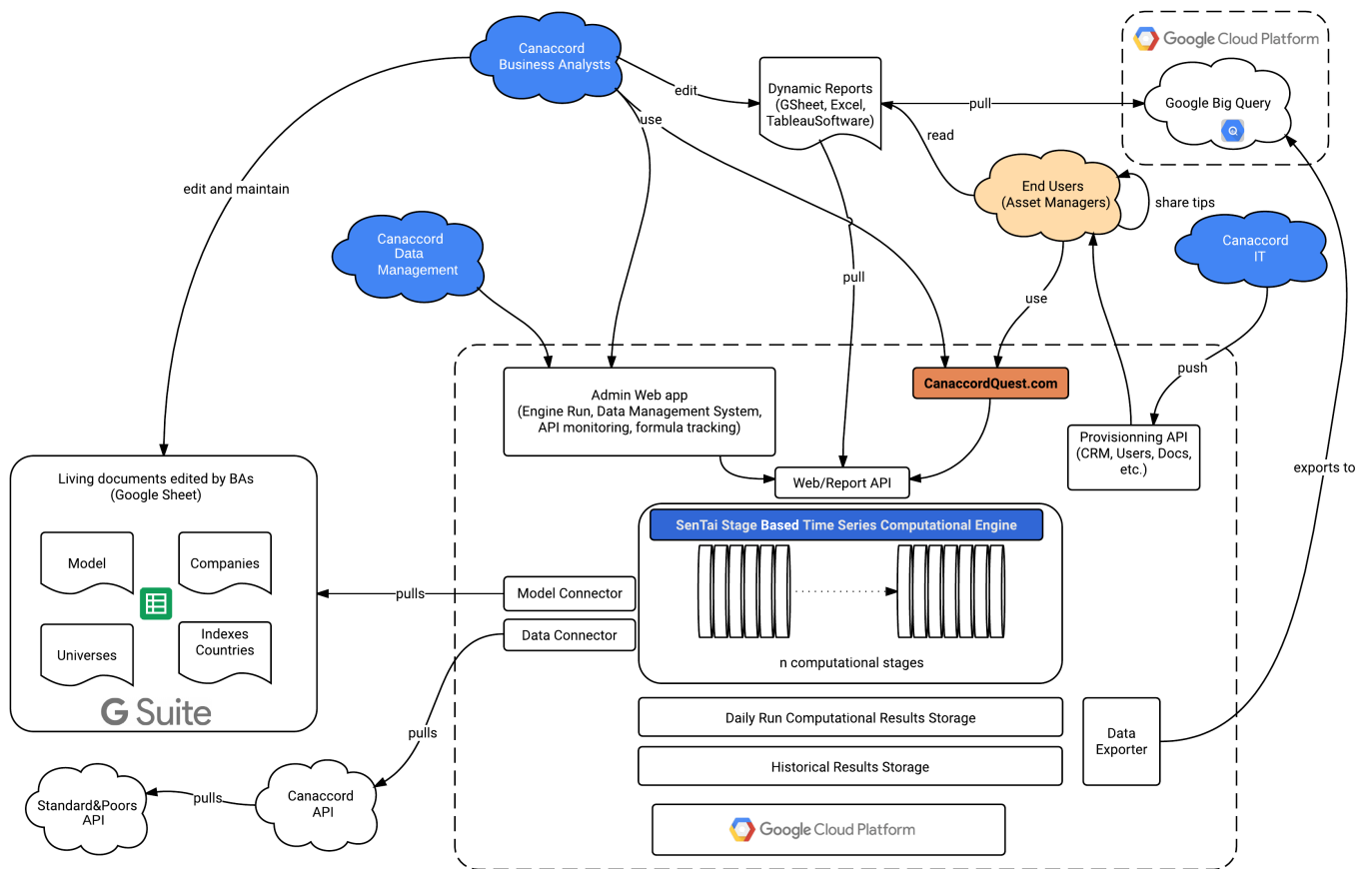
Web



Story

Quest®, originally built by Collins Stewart, has existed for nearly 20 years, and has evolved through many cycles of development and redevelopment. Since inception, it has been run and hosted in-house, based on Microsoft Server Technology (ASP.net). Like most financial applications, it is driven by the ways in which the processed data is made accessible to end users. This can be in the form of a website (primary use), via Google Sheets and centralized Google scripts (using REST APIs and a Google Managed Library), or via a Web Admin Interface for Canaccord Internal users. Like most financial web apps, it was hosted in-house, requiring costly and time consuming backup, maintenance, and support. The complex multiple stage-based model required a complex architecture and infrastructure which became very difficult to maintain, and reached a hard impasse in terms of scalability.

Architecture & Technique



Architecture & Scalability

The application was no longer scalable, in terms of overnight processing capacity or storage. Thus both the model and the application were too complicated to maintain. It had to be rewritten from scratch to allow for a much larger pool of companies to be processed during nightly computations. Moving from on-premise hosting to cloud based hosting was a necessity to allow for peak scaling of both the computational, memory, and storage capacity.

Computational Challenge

The difficulty lay in rethinking the architecture of the software to mimic the tree-like calculation of the Quest® valuation model within a few hours. This data intensive task requires the reading of millions of files, more than 30+ billions data values and the production of a vast number of

multi-dimensional (tensor-like) metrics, per company, per industry, and per geographical region. Furthermore, intra-day recalculations have to be enabled, per company, or per geographical region. This can take place following a correction to the Model via formula alteration, or Data re-import. The computational challenge is therefore intertwined with the agility and ease of use of the entire platform.

Static vs Dynamic Valuation Model & Java transformation

A full reverse engineering analysis was performed on the static legacy model, and adapted to a tree-like valuation model consisting of several thousand formulas, 80+ levels deep of a large dependency tree, and a similar structure for aggregate (regional and industrial) type calculations. Altogether more than 200+ stages are sequentially executed, performing 7000+ rules, using non trivial mathematical / numerical procedures, and complex business rules.

Having a clean tree-like model was a first step, the second challenge was to allow Business Analysts to modify the model (add/modify/remove formulas) without any help from IT or developers. With the legacy application, a minor model modification involved a minimum of 3 weeks process.

This tour-de-force was accomplished thanks to Google Sheets, where Business Analysts can define all their formulas. This tool has 2 main advantages for Business Analysts:

- they can work collaboratively (simultaneously) with appropriate permissioning.
- they are pretty familiar with this kind of tool

The other obvious main advantage was to use Google Sheets as a ready-made GUI to manage these formulas.

Formulas are written in a specialized language (google sheet

syntax or excel-like syntax) that is formally defined by a grammar. The Google Sheet neither interprets any formula, nor does it calculate anything. It works as a living documentation.

The Valuation Engine (where the Quest valuation model resides) reads this spreadsheet, performs thorough validation checks, to detect and report cycles for example, and transforms each formula into Java code. This last operation is done thanks to the Antlr framework (language parser generator). Java code execution greatly speeds up the entire calculation (“Full VE Run”).

At a higher level, business analysts can configure which model (spreadsheet) the VE should use. This enables them to experiment with new models and manage themselves the model life cycle, independently of the VE life cycle.

Benefits

Agility - Flexibility of Valuation Model Changes

Before

Model Changes: minor modifications to the model required **1-2 months**. Changes could only be performed by chosen external vendor

Deployment by developers only:

- **in-DB model and computations:** entire compute engine carried out in DB calculations - SQL Stored Procedures (T-SQL) Very slow, difficult to fix model

After

Model Changes: **Live editable model online by analysts simultaneously** without programming knowledge

Deployment of new model done by business analysts themselves!

- **Google Sheet:** Rewrite model within GS, adding a proprietary syntax for handling of Time-Series
- **Model Translation:** Into Java code from Google Sheet use ANTLR (language builder)
- **Java based computations on GCE**

App Production Cycle

The application is based on a production cycle involving:

- Core Valuation Model: a core tree-like computational model represented in a vast **Google Sheet** with 80+ stages both at the company level as well as the Aggregate level (Industry & Geographical region).
- Data Import: Raw Financial Accounting Data from S&P
- Data Integrity, validation, correction, filling, restatements and approximations
- **Google Sheets**: expression of the 7000+ business rules in custom languages and rules used by a team of 10+ business analysts working on the same model (possibly simultaneously).
- Engine: A Java engine parsing the model rules from **Google Sheet** and dynamically compiling to raw bytecode for maximum performance
- Data Storage Optimization: for multi-dimensional data storage and high performance access.
- Web: 300+ dashboards, charts and dynamic searching within the billions of results done with AngularJS and CoffeeScript.
- Dynamic **Google Sheet**: dynamic representation of pre and post calculation data using our API.

Each one of these components had to be optimized both on the software side, as well as on the infrastructure side, taking into account the various stages and their requirements in terms of RAM, CPU and Storage. The ease of deployment and dynamic rescaling of these cloud ready options made the management of the web application extremely agile.

Collaborative Challenge

The core model was (re)written by a team of 10+ business and financial analysts, based on a proprietary syntax we developed, and reviewed and augmented by our team of developers, on a very large collaborative **Google Sheet**. The ability to work collaboratively on this **Google Sheet** was key to the success of this project, allowing both the business and

developer sides to communicate, review, share, and improve the model in a highly agile way.

For more details, please see G Suite and Quest.

Remote and Online Collaboration

Remotely working developers in various parts of France, graphic designers in London, together with Canaccord's Business analysts and IT teams collaborating via **Google Hangout**, was the key element making the flow of information and exchange of ideas possible. Screen sharing via **Google Hangout** helped resolve misunderstandings much faster than any chain of mails. Fast chats for shorter conversations.

The complexity of the model was extremely challenging: 7000+ rules using non trivial mathematical / numerical procedures, and complex business rules. Business analysts were held back by the difficulty in expressing correctly their needs in terms of the business model (logic). Interaction within this model using comments in the **Google Sheet** proved to be perfectly adapted to the needs of a continuous Q&A resolution cycle. Just like a developer bug tracking software.

The fact that we could share remarks via the comment feature of **Google Sheet** was also key: building a conventional backend application to support the input of such a model would have been way too costly. Worse, it would have diverted our focus to technicalities, rather than the creation of business value.

Our usage of **Google Sheet** helped us focus solely on a flowing and flawless implementation of business rules.

A critical component of building such a complex model is the validation process: being able to connect **Google Sheets** to our backend API, and easily show clear graphs and tables to the Business Analyst (BA) team, helped us efficiently discover errors and discrepancies.

Technical benefits

Persistent Storage and Application Performance

We have extensive needs for SSD persistent storage during the computational phase to optimize the performance of the application. Thus the ability to gradually ramp up the SSD persistent storage to 4Tb per machine has proven key to reduce our overnight app run time by a factor of 2.5. What once required 10 hours+ for significantly less data, we can now compute in the new platform in less than 30 minutes (see Performance & Scalability Gains for more details). The time increase not being linearly proportional to the number of companies, we hope to be able to compute up to 50K companies overnight, on a single 32 core CPU VM instance with 200+GB RAM.

Performance & Scalability Gains

In the legacy system, a full VE Run took 10 hours to manage 2,200+ companies, way fewer industrial sectors (only the first two MSCI Industrial sector levels), as well as a limited number of geographical regions (10). In the new system, we are running 8,800+ companies, the full 275 MSCI industrial sector classification, as well as 25 different geographical areas. Finally the model accommodates for a much greater number of company and aggregate metrics. All done in under 30 minutes. The performance and scalability gains are staggering, at more than **50x**.

Benefits

Scalability

Before (10 hr)  Windows Server 2003

Weekly computations (frequency)
20 year history
2,200 companies
10 regions
100 industrial sectors in 2 level structure
Inability to scale up (F,Y,C,R,I.S.)
Static valuation model
Antiquated website
No new features

After (30mn)  Google Cloud Platform

Daily computations (even intra-day)
30 year history (**1.5x**)
8,800 companies (**4x**)
25 regions (**2.5x**)
275 industrial sectors in 4 level structure (**2.75x**)
Scalable (F,Y,C,R,I.S.)
Dynamic valuation model
Web 2.0 user friendly
New features

Benefits

Performance



Data (Quantity/Speed)

2.5x the amount of computed data (input)
1/20x of the time (30mn vs 10 hrs)
10x more aggregate computations (output)
120 billion data points per full run
7,000+ complex business rules using non trivial mathematical and numerical procedures
200+ computational stages

Price Competitiveness & Benefits

Significant cost savings have been made thanks to:

- An extremely agile development process
- A very low cost of commoditized elastic computing cloud infrastructure
- A very stable platform with low maintenance cost

Cloud Savings

Frequent price drops and most competitive offering are driving the whole industry. Commoditization is driven by Google's know-how, and benefits the end users such as Canaccord. We've been able to start small, and scale up only when ready. The benefits compared with previous hosting providers is substantial. Canaccord was used to in-house infrastructure, with a very slow pace of change: they were amazed at the speed at which we could create/delete/update instances to meet our requirements. The affordability of both GCP and GA4B made them the right combination for the dev team, as well as the business team.

Rationale for Canaccord Genuity

Cost Reduction (4X) - Benefits

Total running costs vs. legacy costs

Infrastructure costs @ GCP: Low monthly costs (less than \$6500/month)

Maintenance: No maintenance of bare metal servers hosted in Canaccord datacenter

Autonomy: Canceling of costly technical outsourcing contract to maintain legacy in-house infrastructure

Resources: Reduction of data cleansing team by automation rules (ETL). Allows to reallocate resources

Benefits

Performance: Stability, scalability, maintainability, evolution, agility, security

Integration: API technology between GSuite & GCP allowed to develop a fully integrated and efficient environment for both business and technology needs

Autonomy: Recover technical knowledge: appropriation of platform by in-house developers trained by us

Resources: Focus its resources on core business (i.e. model improvement)

Risk Mitigation



Security

Several layers of security provided by Google infrastructure



Backups

Use of high speed Disk Cloning, Snapshots, and Imaging. Several layers of backup processes available via managed GCP



Disaster Recovery

Automatically redeploy from scratch entire production environment **in 5mn**, via single command line. **24h previously**



Confidentiality

Customer information is a part of CG's customer base, and minimal data is pushed to cloud platform for credentials, site preferences, site permissions



Intellectual Property

Permissioning (sharing) & restrictions to access to Valuation Model handled by Google Apps



Cloud Restrictions

Cloud restrictions imposed by S&P (financial data provider) required separate in-house SQL server

Training and Handover

Through the entire development process, we have accompanied the business analysts very closely to help them master the new valuation methodology, enabling them to master the novel Time-Series base language we developed. Additionally, we also started training the internal IT team @ Canaccord to prepare them with the handover of the platform. Training in-house Open Source developers is a key to the long-term success large cloud computing and web platforms.

Success

Customer Feedback

Successful Relaunch Institutional Investor Feedback and Usage



- Investment in Quest has created a state of the art platform in the cloud
- Relaunch has been a huge success. Already we are witnessing increased broker votes in, and incremental revenues.

Portfolio Manager:

"It's an excellent tool and we love it. It fits all of our requirements and we appreciate all of the hard work done by the Quest® team."

"New Quest® is now better than Holt."

Senior Portfolio Manager:

"In my personal opinion, the system really is outstandingly powerful. The system is very quick and looks very good. This is a website capable of winning an award for speed, user-friendliness and - if I may say - intelligence of icon design / subtle function prompt."

Director of European Equities:

"CITN (Quest Feature) is the best product on the street by a country mile and new Quest® is a fantastic tool."

Research Analyst:

"Quest® product truly excellent. Product upgrade looks great."

Fund Manager:

"Quest® refresh excellent."

IT Feedback



“

With Google Cloud Platform, our product has deeper data, runs better analysis, is more customisable, handles more than 2.5x the load, and runs 20x faster than the old version.

”

Rob Garwood
Head of IT Development, Canaccord Genuity

CANACCORD GENUITY