

## SECOND ITERATION OF CLOUD-BASED ANALYSIS AND OPTIMIZATION FRAMEWORK

Volker Mueller<sup>1</sup>, Dru Crawley<sup>2</sup>, and Pratik Deb<sup>3</sup>

<sup>1</sup>Bentley Systems, Incorporated, Chapel Hill, NC, USA

<sup>2</sup>Bentley Systems, Incorporated, Washington, DC, USA

<sup>3</sup>Bentley Systems, Incorporated, Yorba Linda, CA, USA

<sup>1</sup>volker.mueller@bentley.com

### ABSTRACT

This research pursues an experimental implementation of an analysis framework in conjunction with an optimization framework for building design. The frameworks tie authoring and analysis tools together under one umbrella. In a prototype, the data flow uses a mix of proprietary and publicised file formats, exchanged through publicly accessible interfaces. An analysis framework brokers between the parametric authoring tool and the analysis tools. An optimization framework controls the processes between the authoring tool and parametric engine on one side and the optimization algorithm on the other. The prototype allowed testing assumptions about workflow, implementation, usability and general feasibility of the pursued approach during the SmartGeometry 2012 event.

### INTRODUCTION

In building design there has been increasing interest in integrating analyses into the design process at earlier stages based on the premise that changes early in design generate greater value with less effort than if similar scale changes were attempted in later design phases or as late as during construction. There are several major challenges that need to be addressed in order to establish a framework-based approach to this problem. One challenge is for example the often questionable interoperability between design authoring tools that are used to create the digital models of building designs and analysis tools that are used to attempt a prediction of how these modelled buildings might perform their various functions, like providing a comfortable indoor climate or a structurally sound environment. Another challenge is the distance between the cognitive skills and tasks involved in conceptual design on one side and on the other side the level of detail required of the information provided to analysis software.

This research aims at finding solutions to such challenges. Following an agile development approach, solutions are implemented iteratively, and assumptions are tested in formal and informal usability settings with immediate or close to immediate impact on subsequent development. In this paper we report about the progress of this process, intermediate findings and conclusions for

further development. The overarching goal is that these capabilities may enable design teams to arrive at better performing designs when compared to traditional design approaches. This seems to be an appropriate response to today's demands for high performance buildings.

The next section expands prior work (Mueller et al., 2013) and is included here to provide the proper background for the presented research. It describes challenges of design and analysis integration as they have been identified based on practice experience, research, or are otherwise self-evident. It characterizes how this research intends to respond to these challenges. Subsequent sections describe the first iteration in form of a prototype implementation, what was learned from this implementation and conclusions for the second iteration of this work.

### CHALLENGES OF DESIGN AND ANALYSIS INTEGRATION

#### **Interoperability:**

The building design software industry is similarly fragmented as the building industry at large. There are many incompatible software programs and data formats. Various approaches have been proposed to overcome or bypass this obstacle to seamless collaboration between design team members (1) A tight, closed proprietary system of software applications, unified through a proprietary data format. (2) a loose system of many software tools using as many data formats with individual translation modules resolving data transformations as needed resulting in many one-to-one mappings (Open Systems and Methods for Built Environment Modelling initiative [1]; Janssen et al. 2012). (3) A loose system of tools using an open or otherwise publicised, standardised data format like the Industry Foundation Classes (IFCs). Latter two approaches frequently employ workflow control systems to glue the pieces together (Flager et al. 2008; Toth et al. 2012). This research project uses a mix of the first and third approaches by using a shared data representation while initially focusing on the implementation of frameworks that will allow loose association of authoring and analysis tools through an application programming interface (API).

### **Data discrepancy:**

During design the design team develops the project from sketchy ideas to a detailed description of the building. Especially early in the design process a lot of the information required for construction is not known, yet, to the design team. While this is expected, this problem persists in various forms throughout the design process. Performance prediction for a project under design frequently require more detailed information than the design team is able to procure, because conceptual design moves through large ideas and concepts, and is less concerned with the detailed specification of constructions. However, analysis software often requires very detailed specifications in order to compute performance of a building (Bleil de Souza 2012). For example, the quality of the steel and exact structural cross sections of beams and columns, as well as how they connect to each other, needs to be specified in order to execute structural performance analysis.

As a response to this challenge of data discrepancy, we propose an analytic framework as the intermediary between authoring tools and analysis tools. The analytic framework API allows analysis engines or services to register through a schema the type of analysis they offer as well as the data required by the analysis services. The framework checks the design and analysis models against the registered schemas and reports any differential between available and required data to the user. The user then can provide the data or dispense with whichever analysis service lacks required inputs. With increasing choices in analysis tools registered to the framework, users gain increasing flexibility of matching the data they have generated to corresponding analysis services. However, if the required information is not available to the designer, yet, or if the design team has not made a decision about the necessary response to that information request, yet, the data discrepancy cannot be resolved and the corresponding type of analysis service cannot execute. A complementary remedy is the search for types of analyses applicable to early design requiring only a minimal amount of information in order to reduce the occurrence of data discrepancy. This is a separate research project.

### **Data equivalency:**

Even if the design team knows all the required information for a specific analysis, authoring tools may not allow representation of all those data, leading to gaps in the data that prevent successful analysis. A conceptual geometric modelling tool might for example provide an elegant surface model of the building design and permit material assignments for visualization purposes while it does not permit to properly define enclosed spaces and wall or window constructions that would be required to simulate the thermal performance of the design.

There is no general remedy to this problem. As described in the proposed solution for meeting the data discrepancy challenge, missing data may be indicated to the user based on the differential between analysis services schemas and model contents; however, unless the authoring tool is supporting creation of these additional data or any necessary amendment of the existing data, users will not be able to easily resolve this situation.

### **Speed of feedback:**

Design is an iterative process with frequent and fast iterations. Analysis results feeding back into these design iterations have to be provided fast enough that they remain relevant for the current iteration (Hetherington et al. 2011). In order to accelerate feedback about the design, various strategies are being pursued. In the case of this research project, opportunities in cloud computing are being investigated in order to accelerate turnaround of analysis results.

### **Performance proxies:**

As mitigating strategy to accurate but slow forms of analysis, science often uses proxy indicators, which are sufficiently reliable, well enough aligned with the behaviour under investigation, and quickly computed. As of yet there is not sufficient research available to permit use of such performance proxies. This area of investigation proper is outside of the scope of this research.

### **Results display:**

Visualization of analysis results is not always intuitive and often not visually related to the geometry model of the design (Dondeti and Reinhart 2011). There are no well-developed conventions how to compare design variants by their predicted performance when working with parametric design tools. Opportunities have been provided within this research project to expand investigation of results display but have not been deeply explored, yet. In general, GC offers the capabilities for users to invent their own visualizations because analysis results are directly accessible in the model and easily matched with the entities to which they apply. Figure 1 shows a mock-up displaying top results of an optimization process side by side, although they are generated sequentially from the parameters associated with the top solutions.

Figure 2 shows a quick prototype developed by Dr. Woodbury visualizing a solution space with three fitness criteria, i.e. a three-dimensional fitness space, expanding the idea of a two-dimensional Pareto plot. A Pareto plot maps solutions into a two-dimensional graph with each one of two fitness criteria spanning x and y, respectively. In the plot, the non-dominated solutions establish a Pareto frontier, a curve along which designers are able to explore the trade-offs between the two fitness criteria. Instead of abstract symbols marking the locations of solutions, Dr.

Woodbury's prototype uses thumbnail renderings to mark the locations of the non-dominated solutions. Because this prototype uses three fitness criteria, the Pareto frontier spans a surface rather than a curve. The thumbnail renderings can also convey additional information. In the context of design feedback our initial focus has been in displaying analysis results in the context of the model, for example through colorization of analyzed elements (Figures 1 and 3).

#### **In-context results:**

Because analysis results are commonly only visualized in a representational fashion, they are not available in the digital model in a form permitting them to be accessed by other computational processes. This prevents automation of refinement iterations or multi-objective optimization routines, for which it is important to process analysis results in relation to the analyzed entity, allowing reactions to be computed in response to the analysis results. Relevant to the idea of integration of analysis feedback into the parametric design workflow is the actual parametric accessibility of the analysis results to the computational system. In this research, analysis results are associated with the analyzed elements so that results can be used in parametric dependencies in order to change the design because of how it performs.

#### **Human-machine balance:**

Especially in early design, not all design goals are measurable, or commensurate, or even quantifiable at all, latter also because qualities are a strong factor in successful design. This poses the challenge how to balance computed performance metrics with qualitative aspects whose performance evaluation is based on the designers' judgments.

There are several approaches possible to resolve this dilemma: (1) avoid any type of automation and fully focus on supporting design team decisions. The design team may utilize analyses and in-context result display to inform its decisions. (2) Implement some complex, pre-defined workflow that mixes design refinement automation with interaction between the system and the design team (Geyer and Beucke 2010). (3) Rely on the multi-level influence the design team can exert on the entire system, from the reasoning behind the parametric model with its behaviours to the selection of parameters that are accessible by the optimization engine or the metrics for the fitness of a specific design. Additional decision support for the team may be available through an interface that allows analysis of all computed solutions, or exploration of the Pareto-optimal set of optimisation runs.

All scenarios are suited to support the design team in its tasks; however, the latter approach makes the full power of computational design available to the design team. This approach is described in more detail in the next section.

## SYSTEM IMPLEMENTATION FOR SMARTGEOMETRY 2012

### **Prototype for SmartGeometry 2012**

For the SmartGeometry event in Troy, NY, in March of 2012, a first implementation step for this system was taken (Mueller et al. 2013). The workshop cluster "Material Conflicts" used a prototype implementation of the design system (Figure 4).

Because there is significant overlap in the architecture of the first prototype and the second iteration, the following description uses present tense although it describes a system that is now defunct. The first prototype implementation uses the parametric dependency and relationship modelling and propagation software GenerativeComponents (GC) from Bentley Systems as design authoring tool and as workflow design tool, as well [3]. GC has been extended with dynamically linked libraries (DLLs) containing feature or node classes for analytical contents, e.g. structural nodes, columns and beams. These DLLs also contain analysis node classes as GC's interface to energy and structural analysis. The energy analysis engine used in this implementation is EnergyPlus [4]. The analysis engine is Bentley Systems' STAAD structural analysis engine [5]. An analytic framework sits as intermediate broker between the analysis nodes in GC and the analysis engines. This applies to both modes of analysis execution, analysis executed locally on the client machine and analysis executed in the cloud. Latter requires a few additional services, like authentication service, the cloud-side counterpart of the analytic framework as a service, file upload and download services, etc.

The prototype workflow and still current workflow is that designers create a parametric model of their designs, including GC nodes that represent structural members (nodes, beams, columns) or energetically active building elements (surfaces as walls, floors or windows) if inclusion of those analyses is intended. In combination with complementary project level information provided through a project node (project name, location, weather data, building type and foundation type), the user then links these data to their respective analysis nodes in GC. The analysis nodes collect the analytical contents from the parametric model and transmit these data packaged as analysis requests to the client-side analysis framework service. Based on the user's selection of either "Local" or "Cloud" execution mode the client-side analysis framework routes the analysis requests to a locally installed analysis engine or to the cloud-side analysis framework service using applicable authentication and file upload services. As cloud resource, the prototype and the current second iteration of the implementation use Microsoft Azure. The cloud-side analytic service framework initiates an analysis job, with the compute node that picks up the job accessing the corresponding data package,

processing the analysis and depositing the result files in cloud storage. Once the analysis job is completed, a post-processing script may extract only salient result data and prepare them for download to the client. On the client-side, the analysis framework receives the analysis results and passes them back to the analysis node in GC. The change propagation then can trigger any downstream behaviours of the parametric model to consume these results, including for visualization in the model, or to close a design feedback loop in case of an optimisation workflow.

For the optimisation workflow, an optimisation node (Design Evolution node or DE node) is added downstream of the analysis nodes and any computations performed on the analysis results in order to generate appropriate fitness values (Figure 4). The DE node is the interface to the optimisation framework, which in turn communicates with the optimisation engines. In the prototype two flavours of evolutionary algorithms or genetic algorithms (GAs) have been implemented which mimic evolutionary processes in order to converge the performance of successive generations of individuals towards some posited “optimum.” Designers identify the contents for two major types of input: (1) the set of design variables or parameters that the GA can manipulate in order to generate design variations, and (2) the set of one to several decision values or fitness values that let the GA evaluate what the performance of the current design variation is. The DE node connects from GC to the Darwin Framework [6] and its evolutionary optimisation algorithms and transmits range and resolution of the design variables to it. These design variables that are considered changeable for design iterations constitute the design’s genome, which the GA needs to consider. After evaluating the genome, the optimisation framework generates genotypes for all individuals in a generation’s population and then waits for all fitness values to arrive. It evaluates the fitness of the individuals to determine the next parent genome and to develop the next generation’s phenotypes.

#### **User Explorations at SmartGeometry 2012**

Material Conflicts cluster participants explored various approaches with the prototype system. Optimisation schemes that were examined comprised an urban study for Penang by Greig Paterson of Aedas (Figure 1) dealing with potentially conflicting goals of maximizing usable floor area for residential use, maximizing overshadowing to minimize cooling loads in a tropical climate, and maximizing views to the ocean with the potential pressure of decreasing building volume and increasing inter-building distances; a natural ventilation scheme utilizing a modulated stack effect by Lem3a and Kristoffer Negendahl (Figure 3) which varies openings in the building to modulate cross-ventilation while controlling distances to a shadow-shroud and assessing shading on the building; and Dr.

Woodbury’s exploration of an alternative representation of a Pareto front (Figure 2). This has been described in the section “Results display.”

#### **CONCLUSIONS FOR THE SECOND IMPLEMENTATION ITERATION**

Prototype implementations by definition are limited. Some of these limitations will overshadow other observations. In the case of the described prototype, analysis models were kept at the minimal implementation necessary to allow analyses to execute while possibly achieving sufficient completeness of the models for conceptual design. Such balance is difficult to strike and can only be reached by implementing several defaults in case of the structural analysis, and many default assumptions in case of the energy analysis. The resulting loss of control on the user’s side may be justified with the ease of use in conceptual design; with the resultant increased speed of constructing analysis models or integrating them into the overall parametric model; and with the lower demand for accuracy in the conceptual design phase. The prototype also suffered from a general lack of robustness.

Other lessons learned as described by Mueller et al. (2013) were the scalability challenge of adding new node types to the parametric model vocabulary for each new analysis type; danger of oversimplification of analysis models; lack of validity checks for the supported analysis data and models; improving messaging to the user; the need to move all parts of the optimisation loop into the cloud to avoid or reduce various performance penalties associated with the cloud.

#### **ONGOING AND FUTURE WORK**

Based on the results of the experiments at SmartGeometry 2012 and subsequent experiments work has focused in several areas. Most crucial is a redesign of the software architecture so that all components can run as services in the cloud in order to minimize cloud upload and download penalties while at the same time maximizing parallelization in the cloud (Figure 5). While the local architecture on the client computer is similar to the architecture of the prototype, and actually could still function in the same way with only the analyses executed in the cloud, the number of service components in the cloud and their architecture has changed. The services in the cloud now mirror more completely the software applications on the client. However, instead of a parametric design authoring tool as a full client application, a modified version now runs on the cloud as parametric modelling engine. This is required so that the optimisation algorithm can generate the necessary individual models in each generation by instantiating them in the parametric engine via specific value sets for the design variables, i.e. via specific genomes. In the engine each genome is inserted into the respective parametric base model

and thus generates a specific individual instance of that model, with the appropriate dependency propagation then triggering any analyses included in the model. These in turn send processing requests via a cloud-based analysis framework to the corresponding analysis engines. The analysis results are post-processed on the cloud and analysis results collected into the parametric model instance to which they belong. This triggers any computations in the model subsequent to the analyses and before returning the decision or fitness values back to the optimisation framework, which after receiving all fitness values for all individuals in a generation determines the parents for the next generation. Once the optimisation reaches its termination criteria the overall optimisation results are passed back to the client application from which the optimisation was initiated.

Crucial is also an increase in the robustness of the analytic framework. That goes hand in hand with an increase in the robustness of communications between software components and services on the cloud.

Compared to the prototype all code has been refactored or rewritten to meet production level standards and use standardized communication protocols in form of RESTful APIs and services are implemented as RESTful services [7]. This is fundamental requirement to warrant adequate quality assurance on the software modules.

Most of these pieces are in place and work on the optimization framework is progressing. The refactored analytic framework is already usable and has been extended to be accessible from an increasing number of authoring tools which either benefit from off-loading multiple concurrent analysis scenarios to the cloud, or which can use the resources in the cloud due to their analysis processes being “embarrassingly parallel,” i.e. allowing several dozen or several hundred analysis threads to process independently in parallel. The biggest challenge for this type of framework is flexibility. In the current real life design processes, users use various analysis engines from different vendors as well as analysis engines developed in-house. Therefore, the success of this approach is heavily dependent upon how easily users can add analysis engines to the system. If the setup time is large and beyond the comfort zone of users the benefit of expected time savings would be neutralized.

This work will be described separately.

The optimization framework also requires refactoring for robustness. The user interface to the optimization node in the parametric model needs thorough redesign. Both of these are still in progress.

An increase in the “completeness” of the analysis and simulation models will allow pursuit of more design questions without increasing the required model complexity. Shading elements have been added to the

energy model. Explicit marking of structural nodes as foundation nodes and structural surfaces for control of wind loading have been added to the structural model.

Further future work includes an increase in capabilities of the analytical models, for example by adding automated zone creation for energy analysis. Separate research is required for finding proxy analyses or simulations that may guide conceptual design with sufficient reliability while executing at speeds superior to traditional analyses.

## ACKNOWLEDGEMENTS

This work was supported by contributions from various groups within Bentley Systems, Incorporated, especially Applied Research, Special Technology Projects, Design and Simulation’s Structural Analysis team, and the BentleyCONNECT team.

## ENDNOTES

1. Open Systems and Methods for Built Environment Modelling initiative. Available from [http://www.linkedin.com/groups?home=&gid=4000091&trk=anet\\_ug\\_hm](http://www.linkedin.com/groups?home=&gid=4000091&trk=anet_ug_hm). Accessed Dec. 7, 2012.
2. “Parti” in architectural design means reduction of the design idea into a diagrammatic representation, either as drawing/sketch or as narration.  
par·ti [pahr-tee, pahr-tee] noun *Architecture*.  
the basic scheme or concept of an architectural design.  
Definition from <http://dictionary.reference.com/browse/parti>. Accessed May 23, 2013.
3. GenerativeComponents from Bentley Systems: <http://www.bentley.com/en-US/Products/GenerativeComponents/>. Accessed May 23, 2013.
4. EnergyPlus from the U.S. Department of Energy, available at <http://www.energyplus.gov>. Accessed June 12, 2013.
5. STAAD.Pro user interface to the STAAD analysis engine: <http://www.bentley.com/en-US/Products/STAAD.Pro/>. Accessed May 23, 2013.
6. Darwin Optimization (version 0.91) by Dr. Zheng Yi Wu. Available from: Be | Communities [http://communities.bentley.com/communities/other\\_communities/bentley\\_applied\\_research/w/bentley\\_applied\\_research\\_wiki/6584.aspx](http://communities.bentley.com/communities/other_communities/bentley_applied_research/w/bentley_applied_research_wiki/6584.aspx). Accessed Dec. 7, 2012.
7. RESTful APIs and RESTful services: REST:

[http://en.wikipedia.org/wiki/Representational\\_State\\_Transfer](http://en.wikipedia.org/wiki/Representational_State_Transfer).

RESTful services at Microsoft developer network:

<http://msdn.microsoft.com/en-us/magazine/dd943053.aspx>.

Both accessed May 23, 2013.

## REFERENCES

Bleil de Souza, C.: 2012, Contrasting paradigms of design thinking: The building thermal simulation tool user vs. the building designer, *Automation in Construction*, Volume 22, 112–122.

Chachere, J. and Haymaker, J.: 2008, Framework for Measuring Rationale Clarity of AEC Design Decisions, *CIFE Technical Report TR177*, Stanford University.

Dondeti, K. and Reinhart, C.F.: 2011, A “Picasa” for BPS – An interactive data organization and visualization system for building performance simulations, in Soebarto, V. and Williamson, T. (eds.), *Proceedings of Building Simulation 2011: 12th Conference of International Building Simulation Association*, Sydney, Australia.

Flager, F.; Soremekun, G.; Welle, B.; Haymaker, J. and Bansal, P.: 2008, Multidisciplinary process integration and design optimization of a classroom building, *CIFE Technical Report TR175*, Stanford University.

Geyer, P. and Beucke, K.: 2010, An integrative approach for using multidisciplinary design optimization in AEC, in Tizani, W. (ed.), *Proceedings of the International Conference on Computing in Civil and Building Engineering*, Nottingham University Press.

Hetherington, R.; Laney, R.; Peake, S. and Oldham, D.: 2011, Integrated building design, information and simulation modelling: the need for a new hierarchy, in *Proceedings of Building Simulation 2011*, Sydney.

Janssen, P.; Stouffs, R.; Chaszar, A.; Boeykens, S., and Toth, B.: 2012, Data transformations in custom digital workflows: Property graphs as a data model for user-defined mappings, in *International Workshop: Intelligent Computing in Engineering 2012 - EG-ICE 2012*, Munich, Germany.

Mueller, V.; Crawley, D. B., and Zhou, X.: 2013, Prototype implementation of a loosely coupled design performance optimisation framework, in Stouffs, R.; Janssen, P. H. T.; Roudayski, S., and Tuncer, B. (eds.), *Open Systems: Proceedings of the 18<sup>th</sup> International Conference of the Association of Computer-Aided Architectural Design Research in*, Association for Computer-

Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

Toth, B.; Boeykens, S.; Chaszar, A.; Janssen, P. and Stouffs, R.: 2012, Custom digital workflows: A new framework for design analysis integration, in Fischer, T.; De Biswas, K.; Ham, J.J.; Naka, R. and Huang, W.X. (eds.), *Beyond Codes and Pixels: Proceedings of the 17th International Conference on Computer-Aided Architectural Design Research in Asia*, Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong, 163–172.

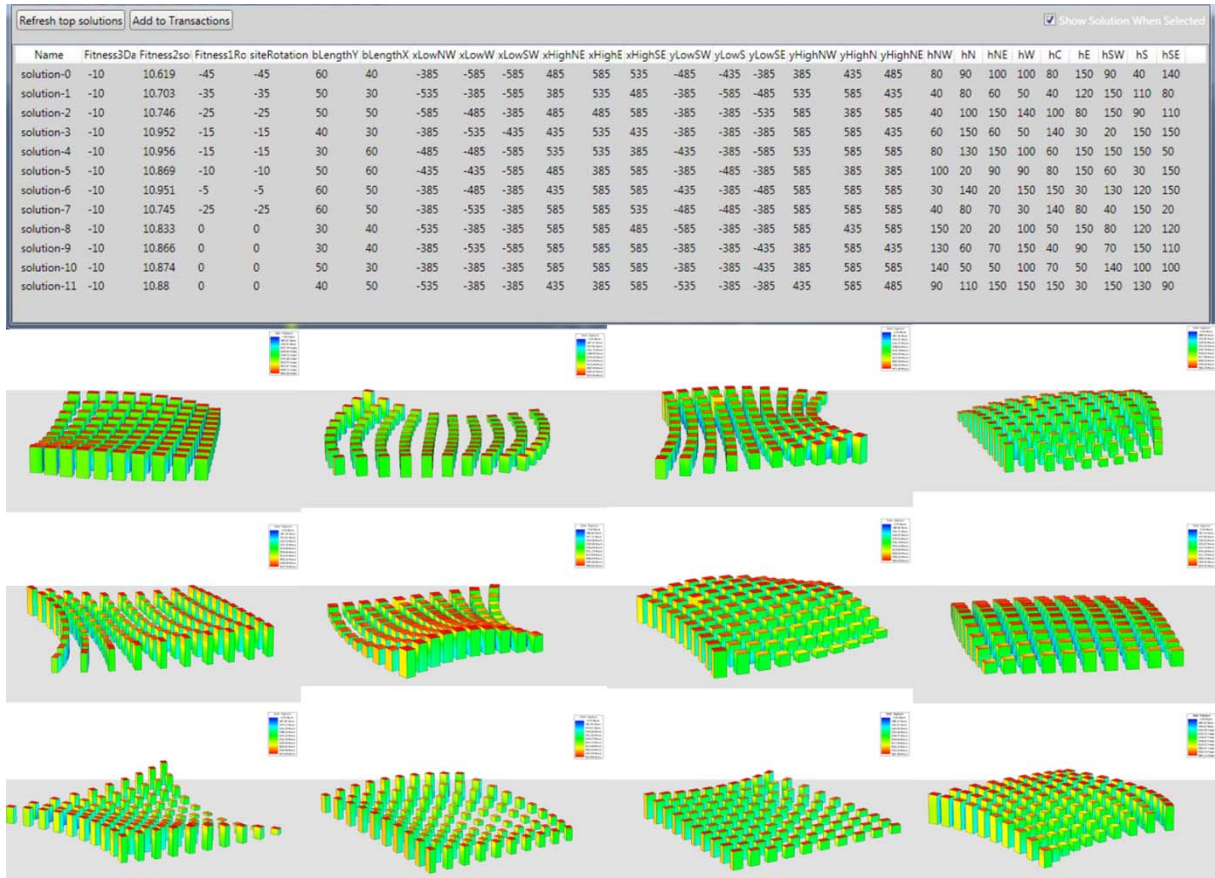


Figure 1: Greig Paterson's urban scheme for Penang - twelve top solutions.

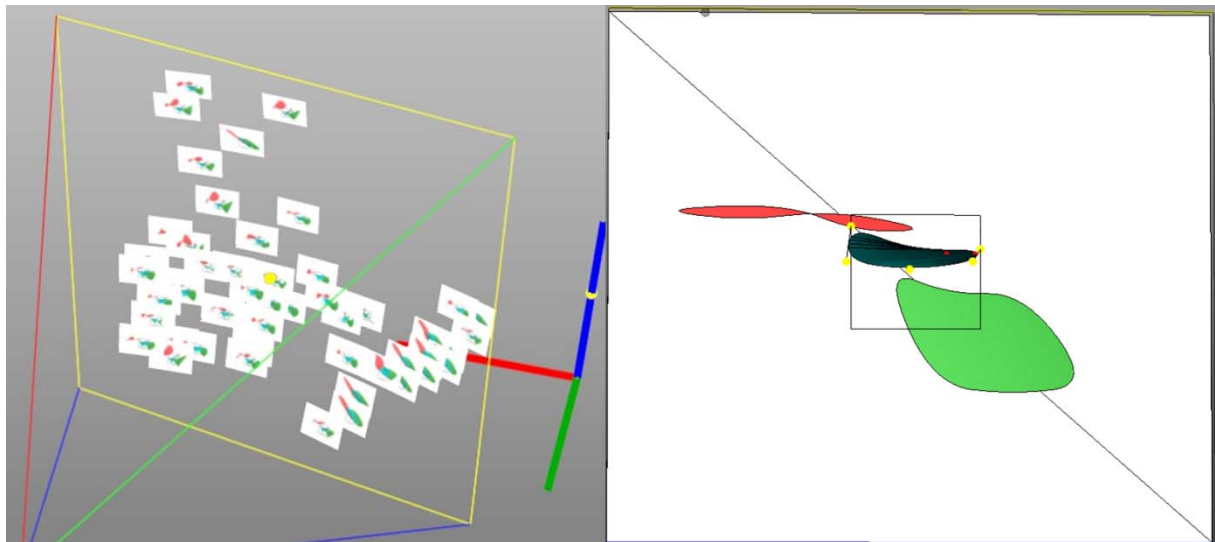


Figure 2: Dr. Woodbury's experiment in Pareto-front visualisation.



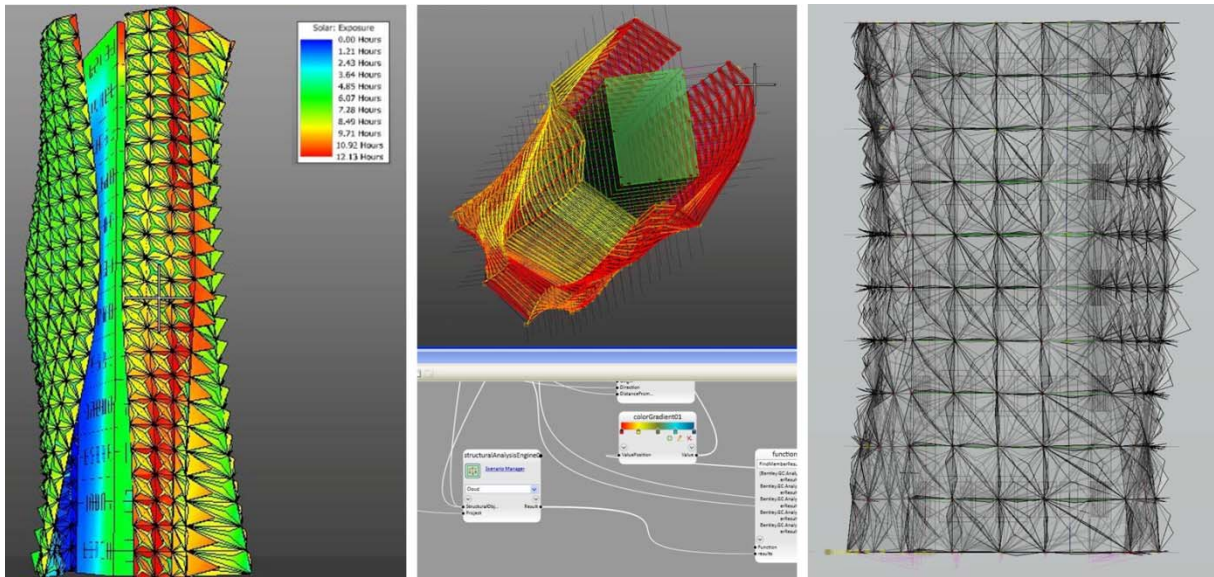


Figure 3: Lem3a's and Kristoffer Negen Dahl's stack effect utilization for natural ventilation, optimization of envelope to modulate stack effect in relation to window openings.

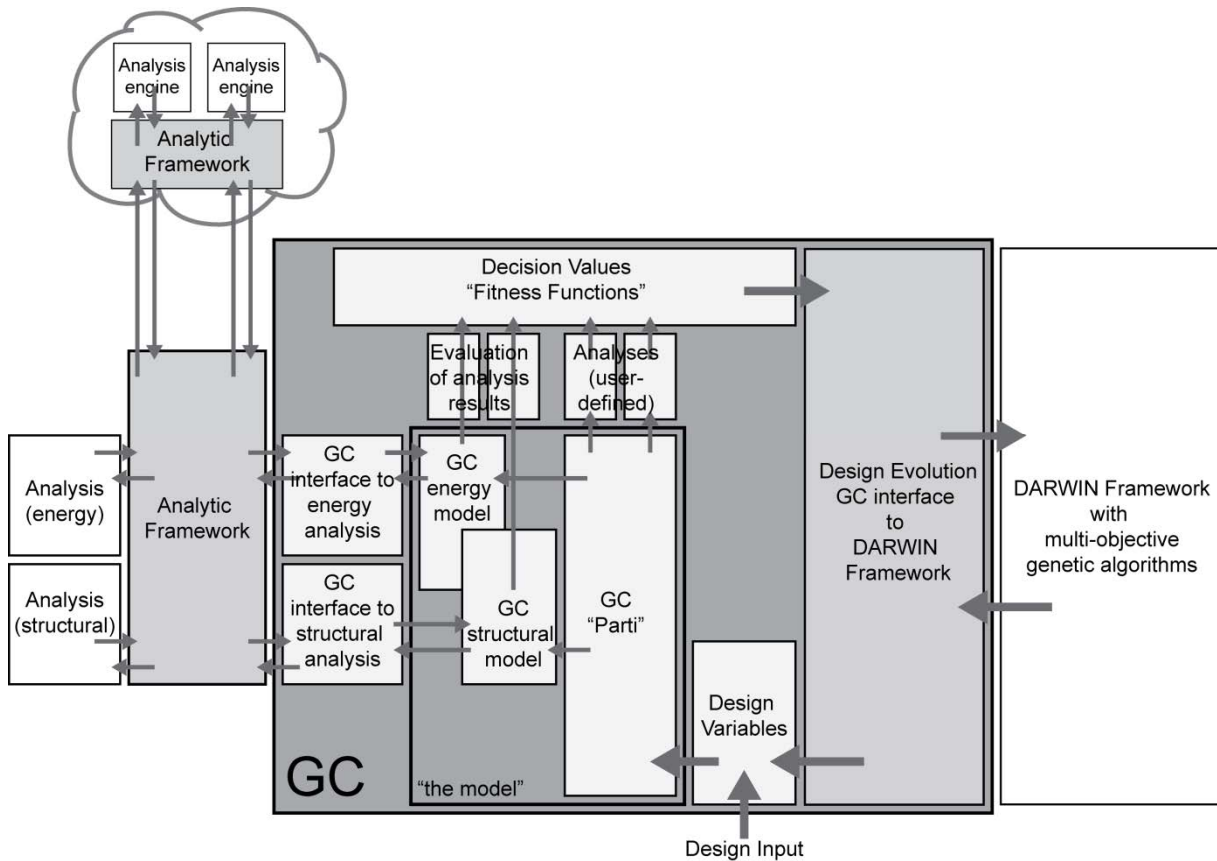


Figure 4: the architecture for the prototype at SmartGeometry 2012.



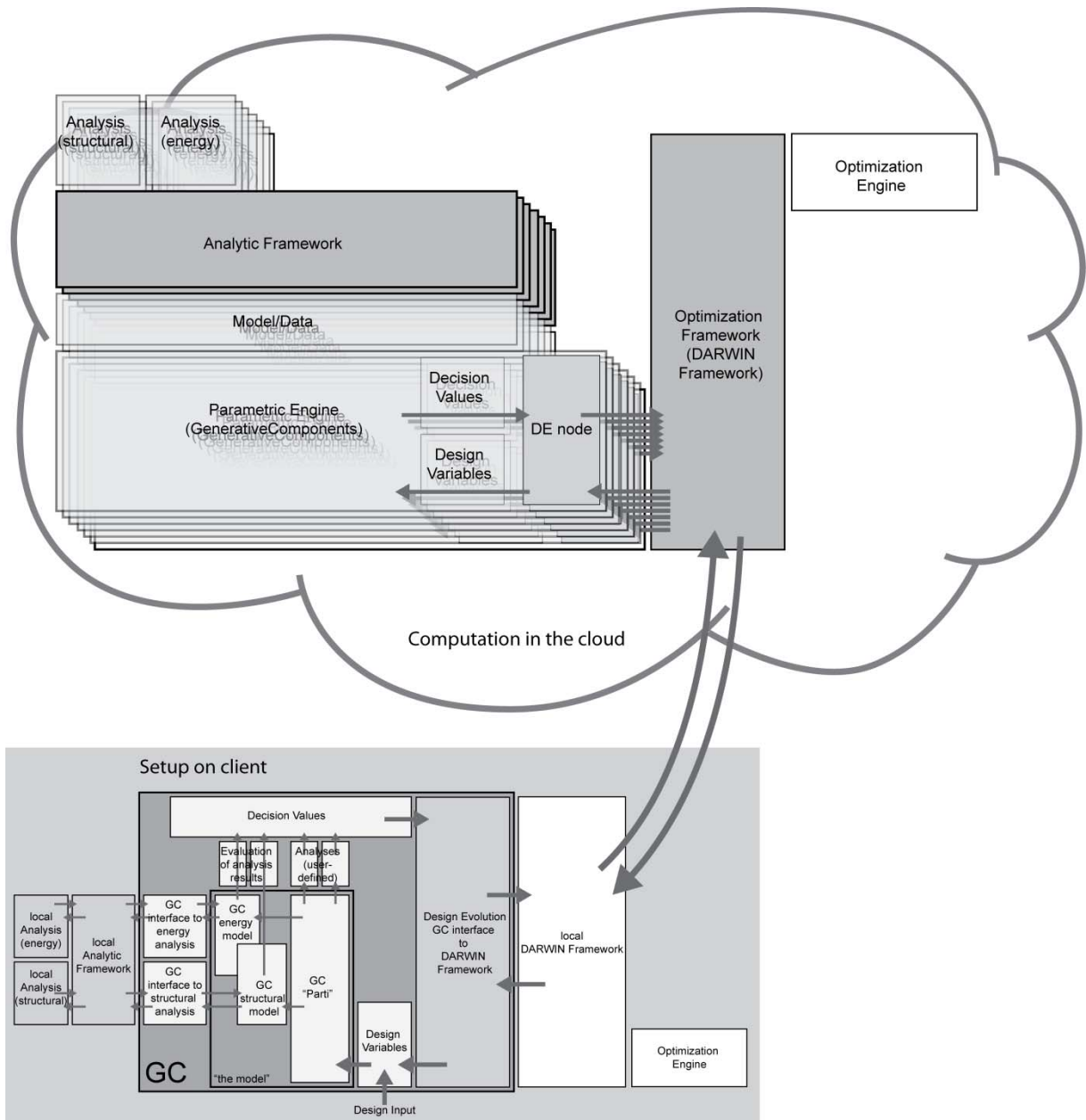


Figure 5: fully cloud-based optimisation architecture.