SAPIENZA
UNIVERSITÀ DI ROMA

Optimization of satellite's onboard data processing workload and related system resources, by offloading specific CPUintensive tasks onto external computational nodes

Gianluca Palermo
Sapienza - Università di Roma

Paolo Gaudenzi
Sapienza - Università di Roma

Introduction

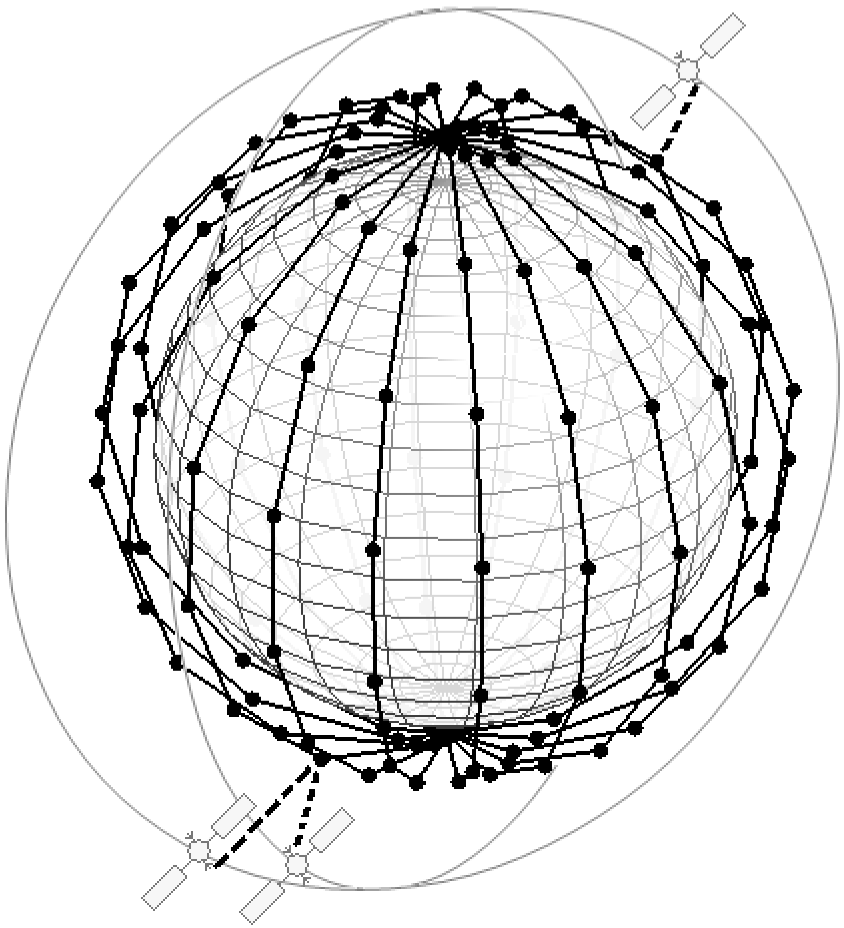
- Purpose of the study presented here is to introduce an innovative way of supporting Low Earth Orbit satellites in particular Earth Observation missions by deploying a LEO infrastructure made up of highperforming computational nodes which can assist clients in their most demanding computational needs.

- In particular, this work investigates the viability of such a solution, proposing, for specific categories of missions, an approach to reduce the use of some onboard resources and to save on satellite system’s costs by offloading computationally - intensive tasks from the onboard systems, while still optimizing the exchange of data through the radio link.

- Optimization of the computational workload is performed by taking into account the minimization of radio link utilization, of power consumption, of onboard memory usage, and of wait-time for the availability of the final data that result from the overall processing.

- The proposed method relies on the application of the Pareto downselection principle to a tradespace of architectural configurations, in order to identify optimal solutions worth being further investigated.

Introduction



Categories of mission that could benefit most from the use of such infrastructure:

- Missions concerning Earth Observation, which typically deal with huge volumes of data and challenging computational needs, would be the target application that benefits most from such a processing scheme.

- Also, the application of this principle could be especially useful in satellite missions that normally perform low-demanding computational tasks and only occasionally require intensive data processing.

Introduction

- Timeliness and efficiency are fundamental aspects when dealing with earth observation missions, in which critical tasks are often involved.

- In the vast majority of cases, satellite's onboard computational capabilities do not allow these operations to be performed locally and, above all, do not allow them to be completed within acceptable times.

- Nonetheless performing these tasks on board might be desirable in a number of specific situations, for example in all the cases where it would be advantageous that an autonomous decision is taken on board in order to avoid the delays intrinsic to LEO missions response times.

Introduction

The following two examples will help to shed light on the convenience in taking autonomous decisions and, consequently, on the need for onboard image processing (or space segment confined image processing):

EXAMPLE #1:

The onboard evaluation of the portions of an image acquired in the visible wavelength spectrum that result covered by cloud formations would allow an efficient transmission to the ground facilities only of the portions of the image not affected by clouds, with consequent savings on the sides of ground access times (in particular w.r.t. narrow visibility time-windows), radiofrequency spectrum occupation, and onboard energy consumption.

EXAMPLE #2:

In a crisis management situation it could be desirable that the space segment of the mission can take autonomous decisions on programming the spacecraft for its image acquisition during the next passes over the area of interest.

For example, when a fire on the ground is detected through an image acquisition, the programming of the same or another satellite for the acquisition of a subsequent high resolution image during its next pass over the fire area would face a series of delays:

1) delay in the transfer to the ground of the first detection image,
2) delay for the time needed in order to process image for detection and for operators to identify the most suitable satellite passing over the fire area,
3) delay for the reception by such satellite of the appropriate telecommand,
4) delay for the time needed for its next useful pass over the area
5) delay consisting in the time elapsing before it reaches an available ground station for the final transmission of the high resolution image to the crisis management center.

Moreover this process could need to be iterated multiple times, multiplying this series of delays.

APPROACH

Approach

- The study makes use of Systems Architecting techniques, which are based on the Pareto-optimality principle.

- This method consists in indentifying some key architectural features (variables) to vary in order to assess how some key aspects of utility and cost are affected (the metrics).

- The aim is to identify, in the variation of the architectural variables (i.e. the **enumeration** of all the possible architectural combinations), the best trade-offs achievable between two or more different metrics of choice.

- This is done by using the Pareto non-dominance principle: given the set of all the solutions, (**tradespace**), an architectural configuration is optimal when it lies on the **Pareto-front** which is the set of points (solutions) for which no solution exists which can outperform, w.r.t. each of the metrics considered, those in such set.

- Among the optimal solutions on the Pareto-front, a downselection is made to identify, those trade-offs which seem to better satisfy the design requirements.

- It is up to decision makers to select one architecture among the Pareto-optimal ones, to be used as a start-point to proceed in a more detailed design.

Approach

An integrated approach has been developed to characterize the trade space for the proposed infrastructure. An overview of such approach is presented in the figure below:

1. Formulation

1. Development of customer database
2. Identification of relevant utility metrics for customers
3. Identification of relevant cost metrics
4. Inference of the most influencing architectural aspects
5. Architectural design matrix definition

2. Enumeration

All architectural configurations are explored, discarding unfeasible architectures

3. Simulation

1. Propagate customers & infrastructure-nodes' orbits
2. Evaluate visibility & range for each customer/node pair
3. Evaluate number of available CPUs for the considered nodes
4. Select the best performing customer/node pair
5. Calculate datarate and transferred data volume for each customer
6. Calculate client's average RF power emission

4. Evaluation

1. Calculate Average Response time from simulation results
2. Calculate client's average power consumption
3. Calculate infrastructure's costs
4. Evaluate architectures with the metrics of interest

5. Downselection

Perform a downselection of solutions applying the Pareto-optimality principle

Approach

By the proposed analytic process, architectures of interest are identified and characterized as a set of major decisions, providing alternatives to decision-makers for subsequent choice of which design points carry through more detailed design.

> FORMULATION

The table below shows the architectural design matrix considered:

Architectural decisions	Decision alternatives			
Altitude	700 km	800 km	900 km	
Number of satellites per plane	8	10	12	14
Number of orbital planes	3	4	5	6
Number of CPUs per node	5	10	20	

Approach

The following assumptions have been made:

- nodes' communication payloads have the ability to track customer satellite during operations.
- contacts between spacecraft can be established instantaneously as a function of the line of sight;
- intersatellite connectivity is assumed among nodes, allowing an optimal computational workload balance and memory allocation distribution among all nodes.

The following rules are implemented into the simulation algorithm:

- the best node for each client is selected after verifying the reciprocal visibility on the basis of the smallest distance and the CPU availability.
- Client/node coupling solutions that provide too low data rates (due to excessive client/node distance) are discarded in order to avoid occupying and wasting with negligible workloads CPU resources which can instead be profitably used by other clients.

Approach

- Lifecycle simulation is performed for each architecture. Simulation consists in modelling the lifecycle of the architecture to evaluate the average response time for the processing of a standard image (5000*5000 pixels * 32 bit) and the average RF power required to the client spacecraft to exchange outbound and inbound data.

- The response time as defined in the simulation includes:
    - time for the client's outbound transmission of data to be processed towards a node of the infrastructure
    - wait time when no node visibility or availability occurs
    - node's data processing time
    - time for the client's inbound transfer of the processed data from the node of the infrastructure employed for the computational task.

- Average RF power takes into account only the instants in which a data transfer occurs (the time during which the trasmitter is off does not contribute to the calculation of the average RF power).

Approach

The following metrics have been used to evaluate architectures:

Constellation cost, including: • RDT&E • Launch • Ground segment • Operations
Average Response Time = outbound transmission time + processing time+ inbound transmission time
Average Client's Power Consumption (RF power)

Economies of scale have been taken into account and the incidence of architectural choices on cost has been roughly estimated using a typical learning curve factor of 75%.

Launch costs (based on Falcon 9) have been calculated starting from the satellite mass and by determining the number of satellites that can be grouped in one single launch (depending on mass and designated orbits).

Also, ground station and operation costs have been estimated approximately using typical values from the literature.

Approach

> SIMULATION

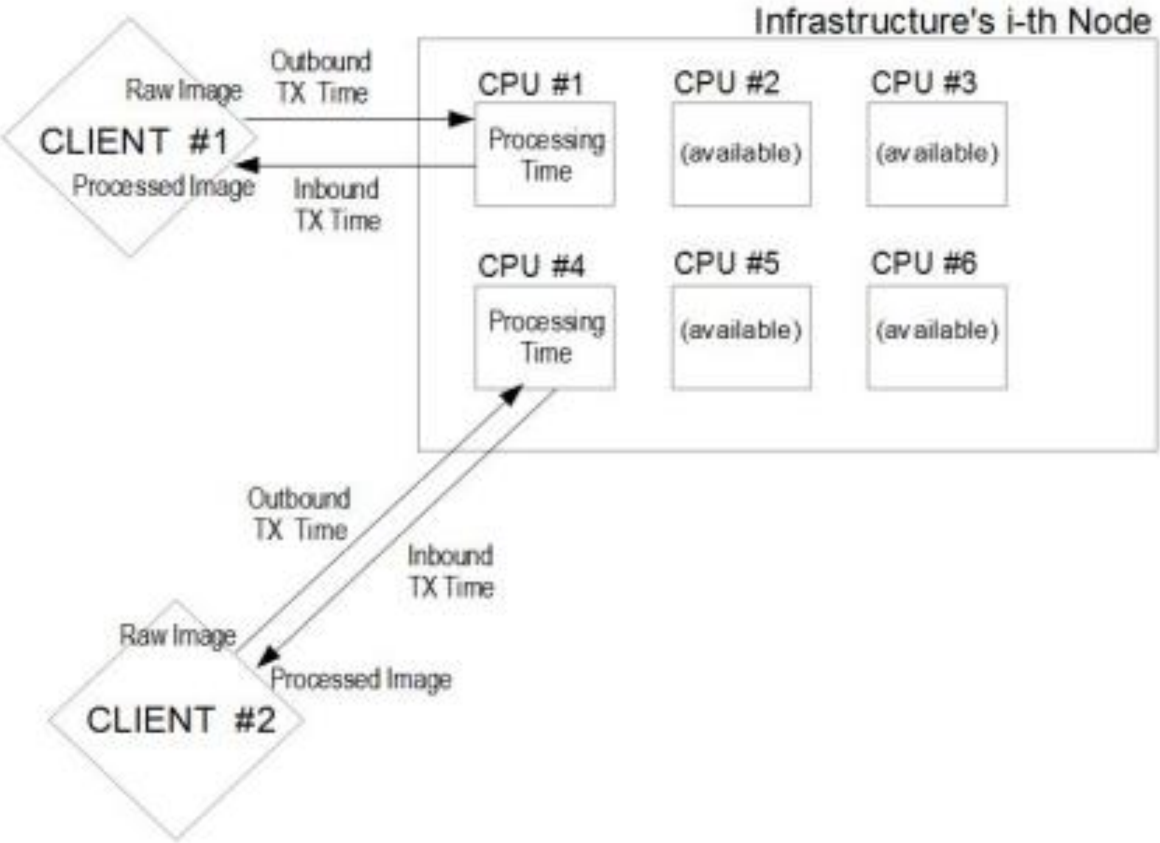
-The orbital motion of all the node and client satellites is simulated with an orbital propagator

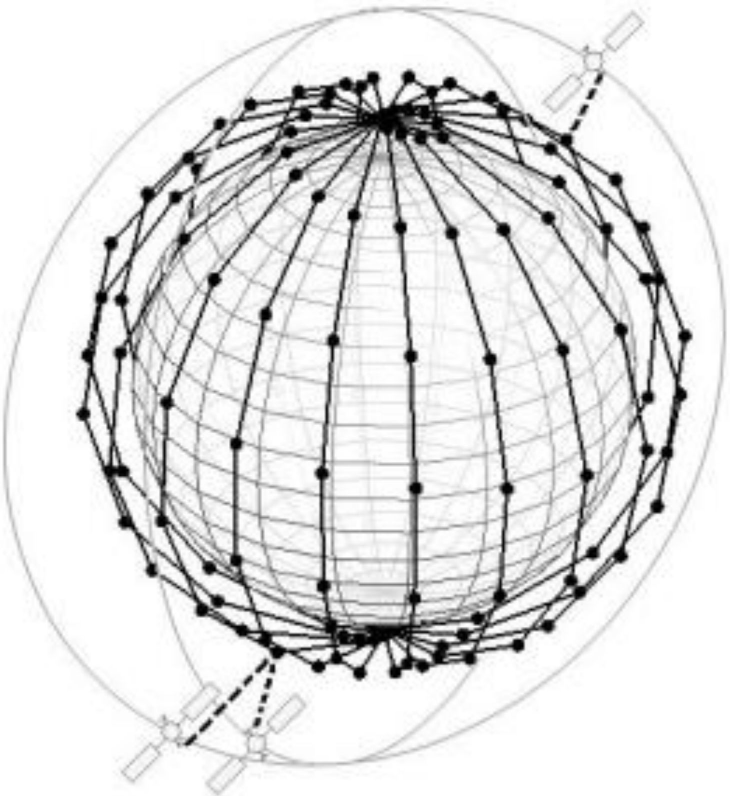
-All the link aspects are evaluated according to the relative positions of the spacecraft (visibility and distance);

-The concurrent access dynamics are simulated during each time-step of the orbital simulation to provide a realistic assessment.

- Two different scenarios are simulated: 30 and 100 customers.

Approach


Infrastructure's i-th Node
CLIENT #1
Raw Image
Processed Image
Outbound
TX Time
Inbound
TX Time
CPU #1
Processing
Time
CPU #2
(available)
CPU #3
(available)
CPU #4
Processing
Time
CPU #5
(available)
CPU #6
(available)
Outbound
TX Time
Inbound
TX Time
Raw Image
Processed Image
CLIENT #2



CASE STUDY

Case Study

The case study for both 30 and 100 clients scenarios will assume that the task to be performed by the infrastructure is the same for all clients and consists in an image processing for the change detection.

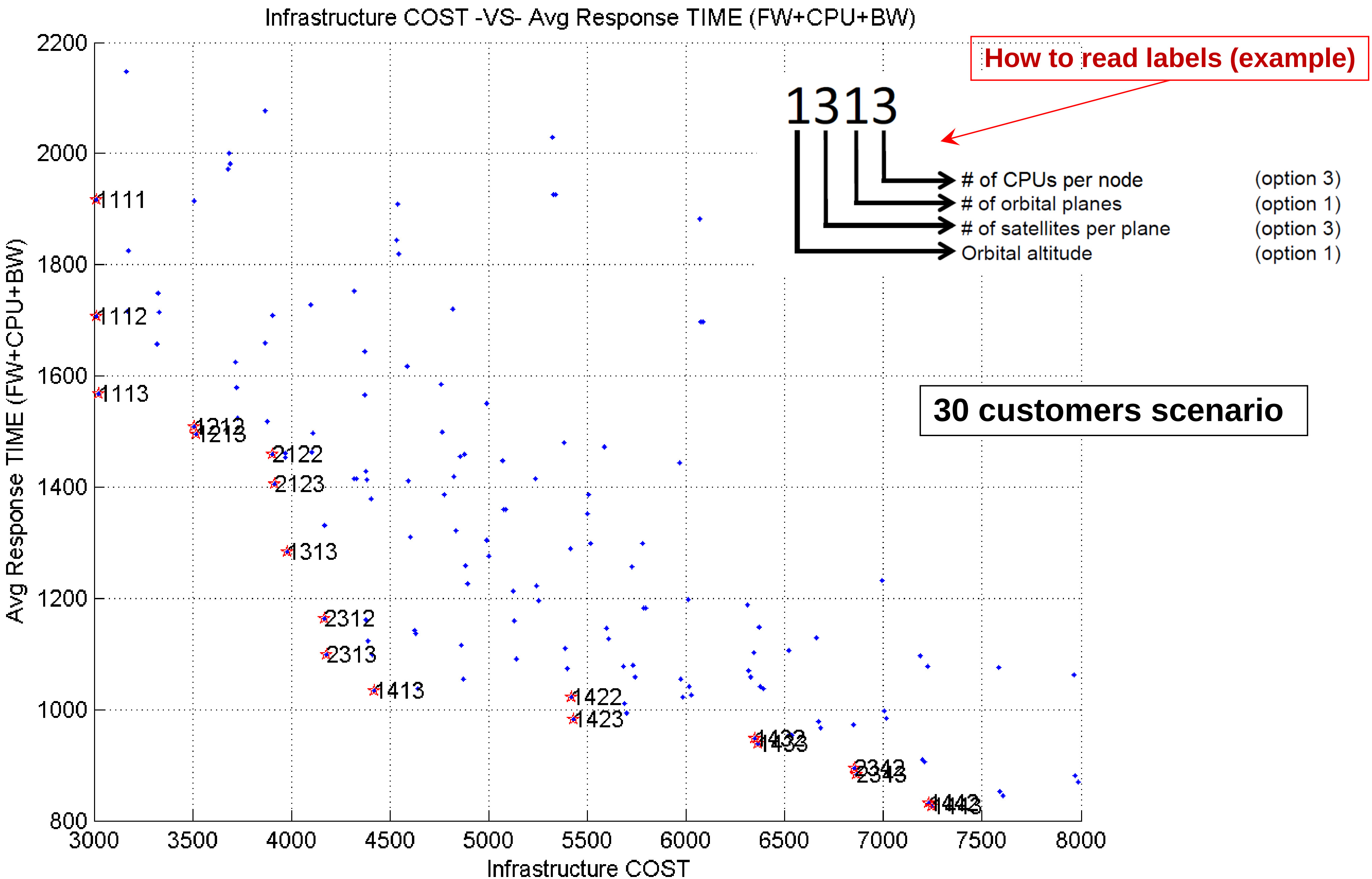
Details on the assumptions for both scenarios, along with some technical details are reported in the table below:

Node-Spacecraft's Mass	450 Kg
Node-Spacecraft's Total Power	800 W
MAX client RF Power (Peak)	140 W
Carrier Frequency **(client/node)**	35GHz
Bandwidth **(client/node)**	20MHz

Nominal Data Rate	600 Mbps
Standard Image Size	5000*5000 pixels, 32 bits/pixel
Type of Processing	Change Detection (PRINCIPAL COMPONENT ANALYSIS)
CPU Processing time (@2000 MIPS)	150 seconds

Case Study

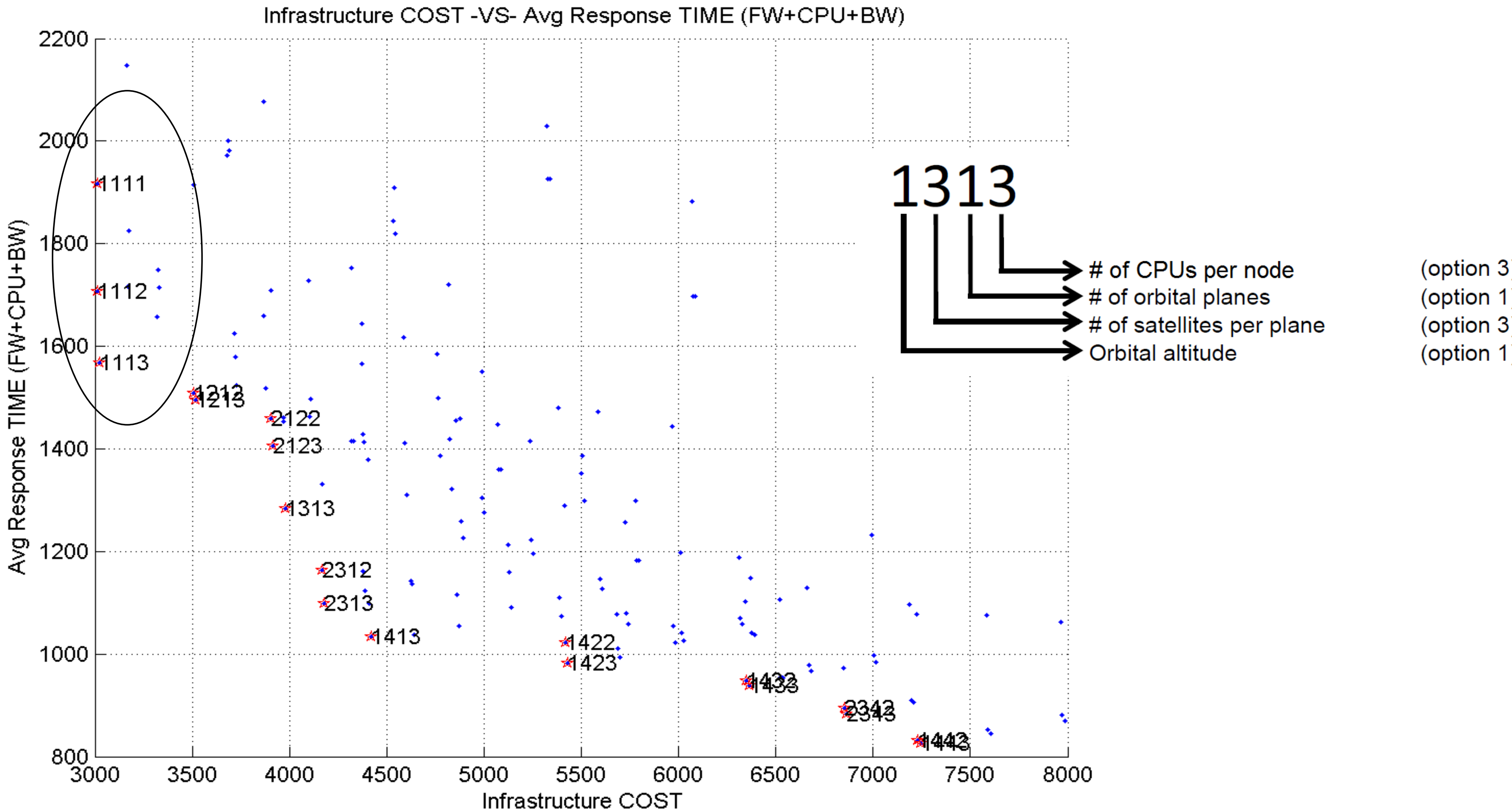
Pareto efficient architectures can be found on the highlighted front:



Case Study

Three regions can be identified on the Pareto front:

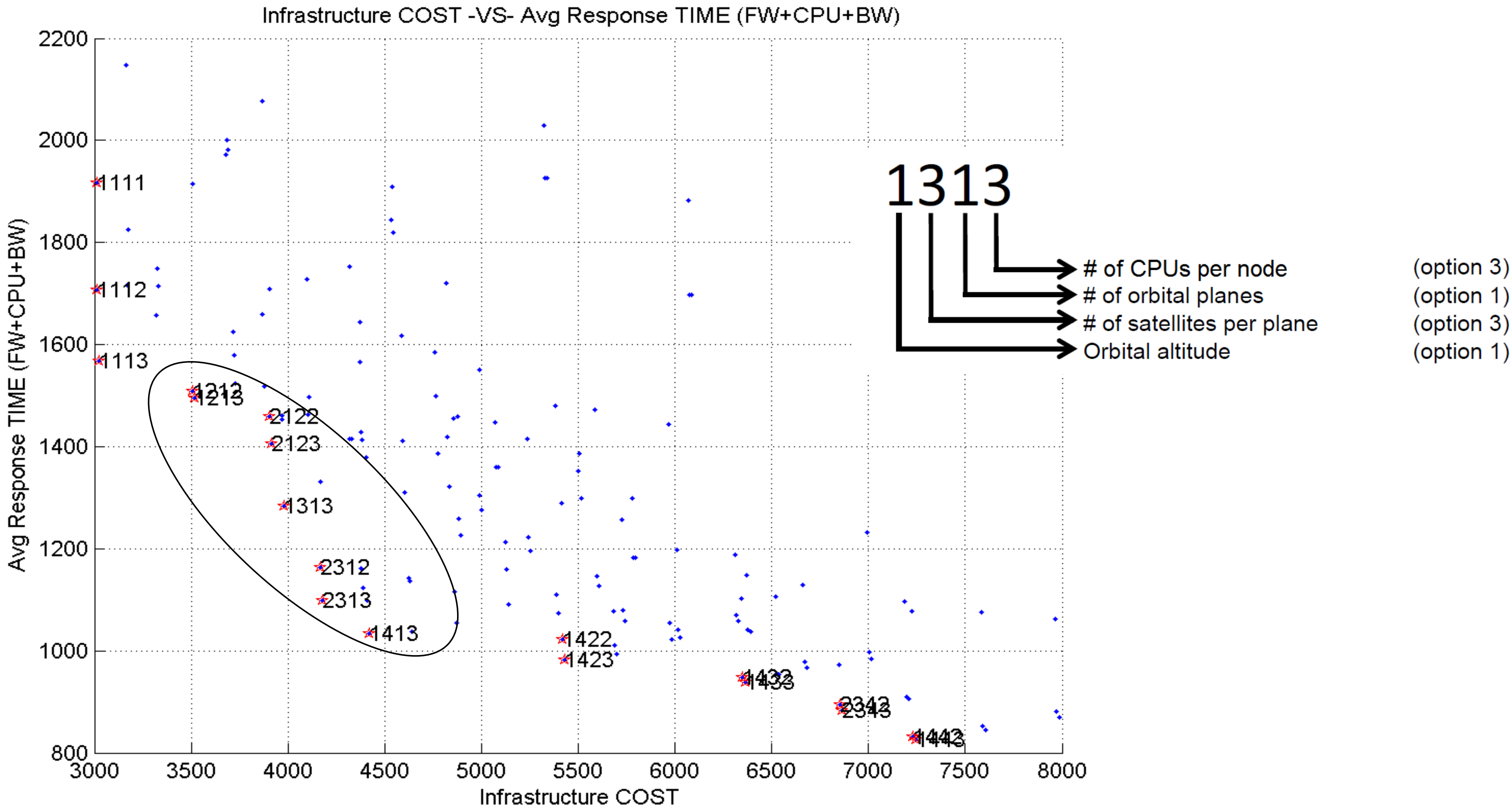
1) **<u>Northwest region of the pareto front</u>**: small increases in cost lead to significant reduction in response time. Offers low cost architectures, which however could be significantly improved by modest increases in budget available.



Case Study

Three regions can be identified on the Pareto front:

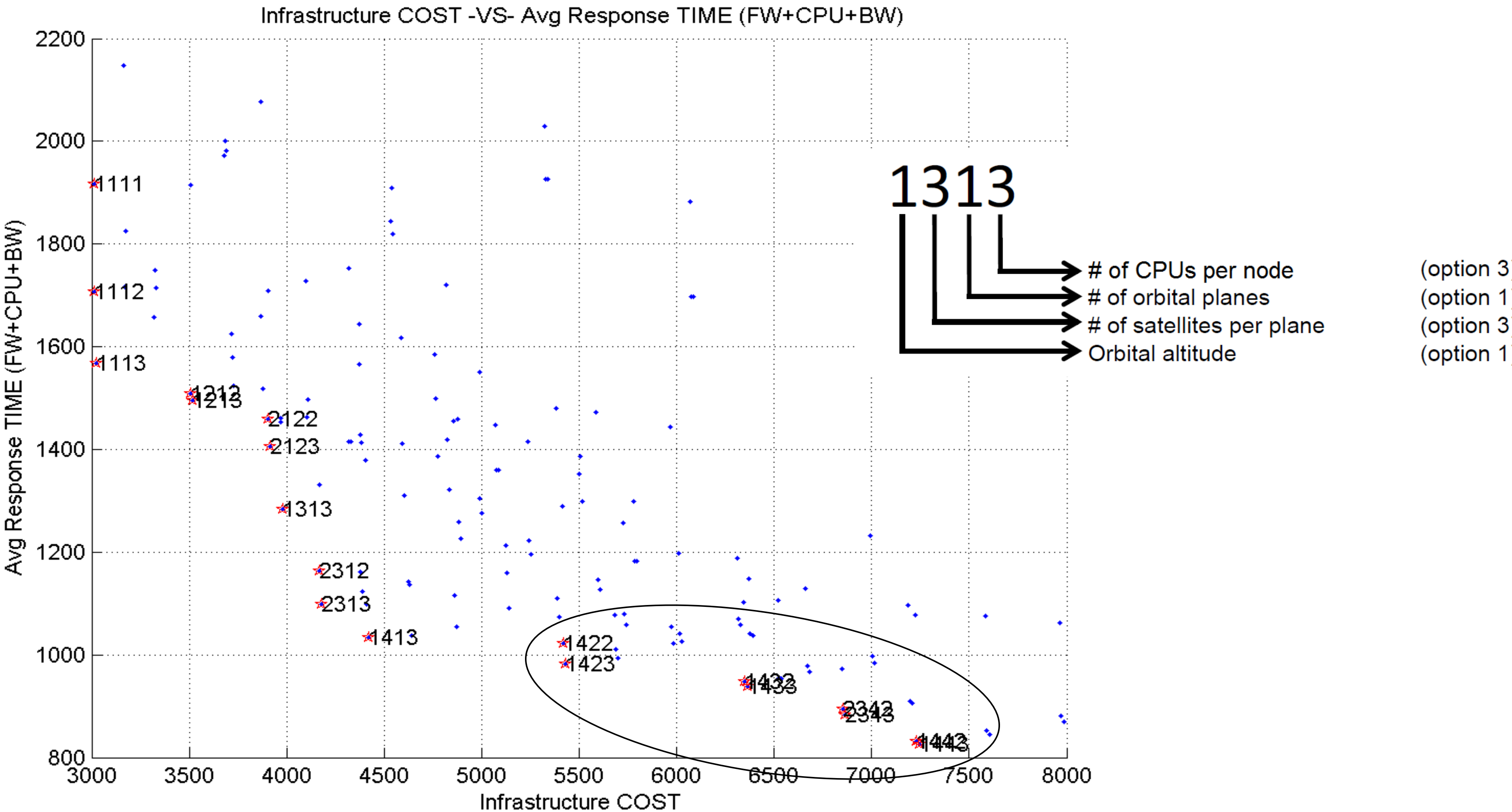
2) **Southwest region**: increases in cost are commensurate to reduction in response time. This region includes architectures that represent best compromises in the tradespace, offering an adequate performance w.r.t. the associated total cost.



Case Study

Three regions can be identified on the Pareto front:

3) **<u>Southeast region</u>**: small reduction in response time correspond to large increases in cost. This region offers architectures that are high performing, although a significant cost burden is required to achieve a marginal increase in performance.



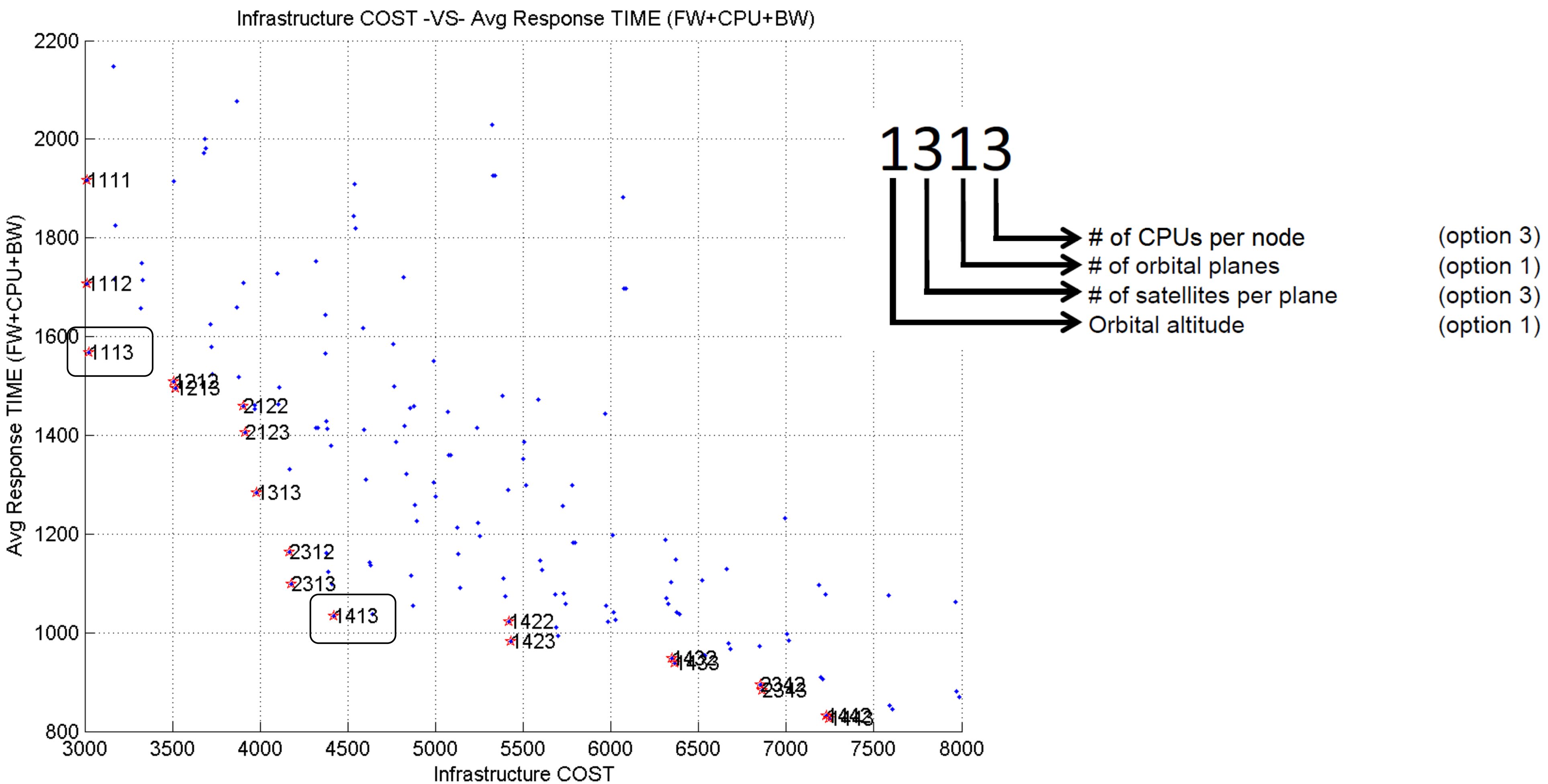
Case Study

Comparison between two example architectures lying on the Pareto front :

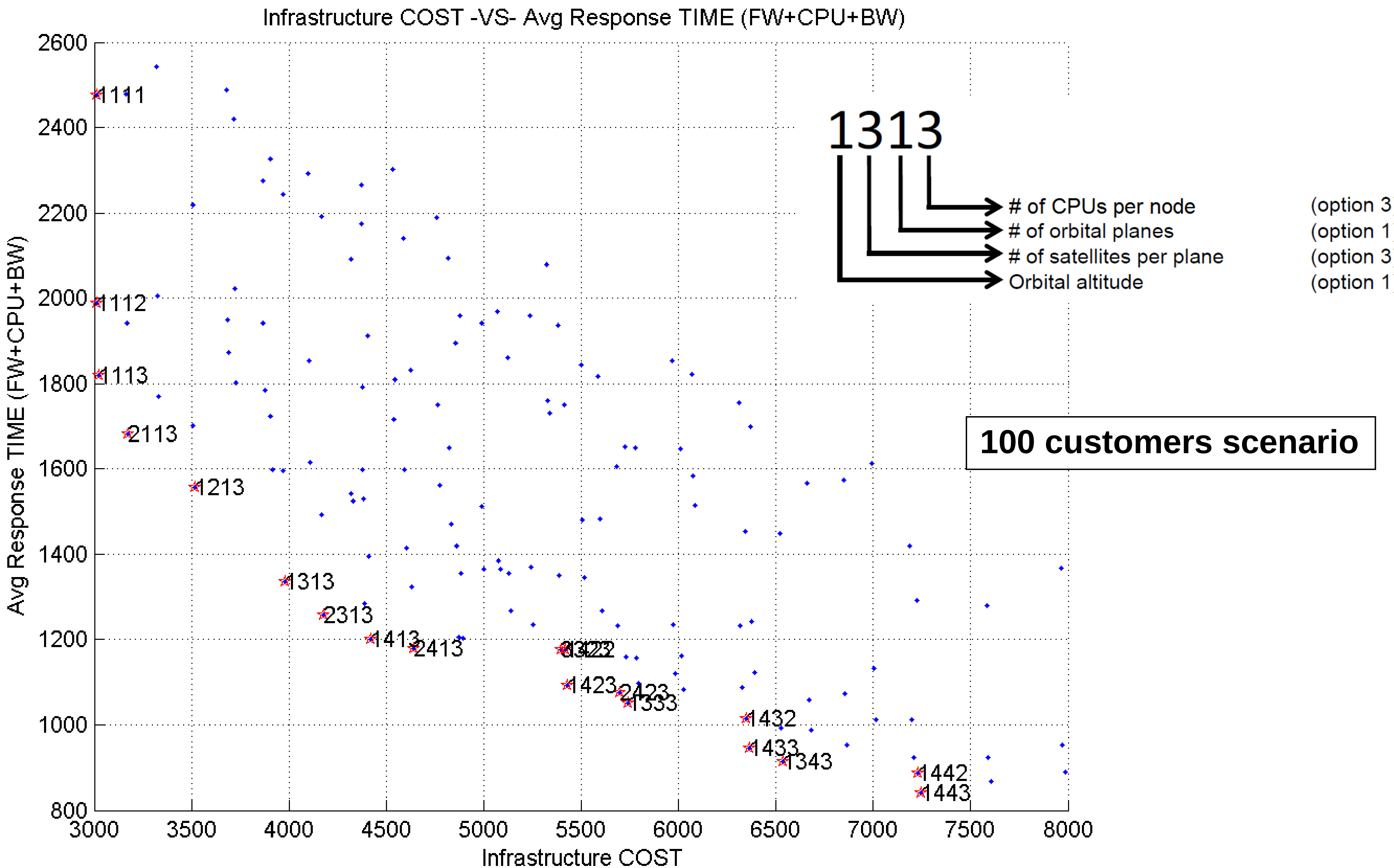
1) Architecture **ID 1113**: 1580 seconds average response time, lifecycle cost of 3,000 MUSD.

2) Architecture **ID 1413**: 1020 seconds average response time, lifecycle cost of 4,450 MUSD.

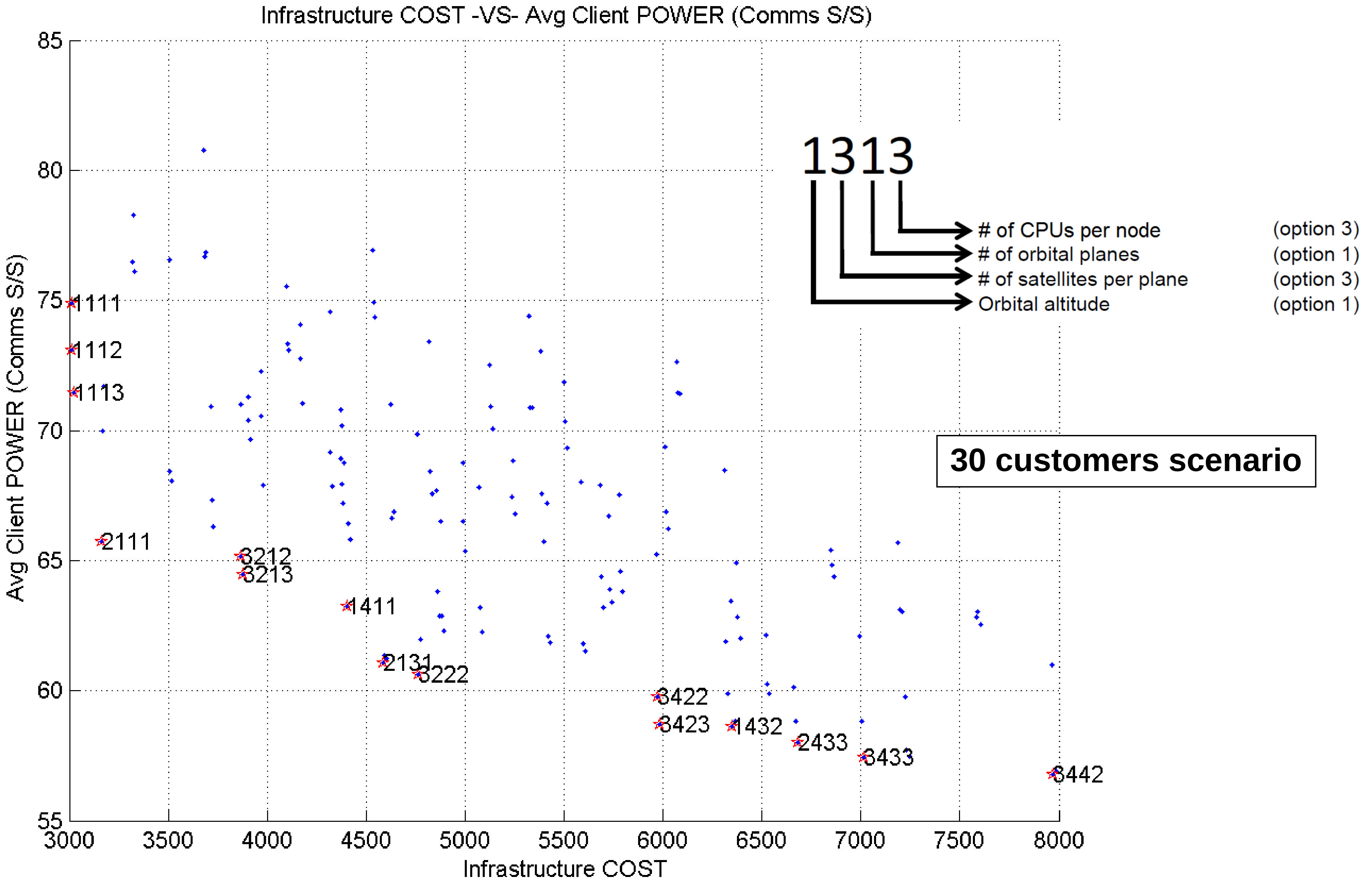
<u>A ~35% reduction of response time corresponds to a ~48% increase in cost.</u>



A Different Scenario: 100 Client Satellites



A Different Metrics: Client's Average RF Power



CONCLUSIONS

Conclusions

- This paper illustrated a systems architecture methodology to explore the tradespace of an infrastructure meant to assist client satellites in their most demanding computational needs.

- The analysis shown here identified the opportunities associated with the deployment of a LEO satellite constellation as an infrastructure intended to support LEO satellites by offloading computationallyintensive tasks from their onboard systems, while still optimizing their use of the radio link.

- Two scenarios 30 and 100 client spacecraft have been considered to evaluate how efficient architectures for the proposed infrastructure look like under different load conditions and to assess how the investment strategy should change accordingly.

Conclusions

- The cost model has been developed by contemplating costs for RDT&E (including economies of scale analysis), launch, Ground segment and Operations.

- Utility metrics has been identified in computational response time obtainable from the infrastructure

- Requirements on the client's side have been evaluated w.r.t. its average RF power emissions

- Designers can use the proposed model to better understand tradeoffs between cost, system performance and client constraints, and to architect the constellation accordingly, assessing a variety of Pareto optimal alternatives.

- Subsequently, decision makers can choose to elect the infrastructure's system performance as an architectural design driving criteria, or alternatively on the basis of the intended customer target, they can focus -for the choice of the most convenient architectural configurations- on the constraints to be imposed on the client side with the adoption of each optimal design point.

SAPIENZA
UNIVERSITÀ DI ROMA

Thank you. Questions?

Optimization of satellite's onboard data processing workload and related system resources, by offloading specific CPU-intensive tasks onto external computational nodes

Gianluca Palermo
Paolo Gaudenzi