May, 2011

Architecture Evaluation without an Architecture: Experience with the Smart Grid

Rick Kazman, *Software Engineering Institute*
Len Bass, *Software Engineering Institute*
James Ivers, *Software Engineering Institute*
Gabriel A. Moreno, *Software Engineering Institute*

Architecture Evaluation without an Architecture: Experience with the Smart Grid

Rick Kazman
Software Engineering Institute/CMU
and University of Hawaii
4500 Fifth Ave.
Pittsburgh, PA, 15213
kazman@sei.cmu.edu

Len Bass, James Ivers, Gabriel A. Moreno
Software Engineering Institute
Carnegie Mellon University
4500 Fifth Ave.
Pittsburgh, PA 15213
{ljb, jivers, gmoreno}@sei.cmu.edu

ABSTRACT

This paper describes an analysis of some of the challenges facing one portion of the Electrical Smart Grid in the United States—residential Demand Response (DR) systems. The purposes of this paper are twofold: 1) to discover risks to residential DR systems and 2) to illustrate an architecture-based analysis approach to uncovering risks that span a collection of technical and social concerns. The results presented here are specific to residential DR but the approach is general and it could be applied to other systems within the Smart Grid and to other critical infrastructure domains. Our architecture-based analysis is different from most other approaches to analyzing complex systems in that it addresses multiple quality attributes simultaneously (e.g., performance, reliability, security, modifiability, usability, etc.) and it considers the architecture of a complex system from a socio-technical perspective where the actions of the people in the system are as important, from an analysis perspective, as the physical and computational elements of the system. This analysis can be done early in a system's lifetime, before substantial resources have been committed to its construction or procurement, and so it provides extremely cost-effective risk analysis.

Categories and Subject Descriptors

D.2.11 Software Architectures; D.2.2 Design Tools and Techniques;

General Terms

Design

Keywords

Architecture analysis, ultra-large-scale systems, socio-technical systems, Smart Grid.

ICSE 2011, May 21–28, 2011, Honolulu, HI, USA.
Copyright 2011 Carnegie Mellon University
978-1-4503-0445-0/11/05... $10.00.

1. INTRODUCTION

Architecture evaluation has been well established in industrial practice for over a decade [2] but one (quite reasonable) assumption that all evaluation methods have made is that there is an architecture to evaluate. In Ultra Large Scale (ULS) systems [10], the architecture may not yet exist for portions of the system or there may be competing architectures for the same functionality in different portions of the system.

In this paper, we describe an application of the principles underlying traditional architectural evaluation methods to the problem of analyzing an *architecture landscape*: a broad set of architectural decisions which represent a spectrum of potential architectures. Why would we want to analyze an architecture landscape, rather than a concrete architecture? We are motivated to do this in situations where there are many architectural decisions to be made, many stakeholders, many similar systems to be built, and where the architectural decisions are non-trivial and their consequences are far-reaching. ULS systems are examples of such landscapes: in a given domain there may be many different potential architectures and it may be difficult to assess the long-term consequences of the many seemingly minor architectural decisions that must be made.

We were faced with just such a situation in analyzing the emerging Smart Grid in the United States. Electric utilities are being asked to plan for the Smart Grid of the future, but there is little guidance and few examples of how to build the IT systems that will manage the Grid, and the examples that currently do exist are of modest scale. The existing electric grid is a very large system but the IT system that accompanies it has, in the past, been comparatively simple.

Consider the architectural challenges posed by just one portion of the emerging Smart Grid: residential Demand Response (DR). DR is a system that attempts to systematically manage consumers' demand for energy, typically by shedding load in situations of high energy demand. Such systems are highly beneficial to utilities who must attempt to plan for, and build for, peak demand. By lowering the peaks, utilities can save considerably on infrastructure that only gets used a few times a year. Utilities are being asked to plan for DR systems, but they must make architectural decisions in a vast decision-space with little guidance. On behalf of the United States Department of Energy, we have investigated the architecture landscape for DR, to guide utilities in making architectural choices when they implement these systems.

In this paper we show how principles that underlie architecture analysis can be broadly applied to the less well-defined and fluid situation that is found in DR, and is found in ULS systems in general. The characteristics of ULS systems that are exhibited by DR systems are

- *Continuous evolution and deployment*. A DR program, for the foreseeable future, will continue to evolve as new technology and smarter appliances become available and as consumer behavior and expectations change.
- *Heterogeneous, inconsistent, and changing elements*. Each utility will have its own DR program with its own variations. Different programs will be inconsistent, different sets of incentives will be tried. Residents who sign up for one program may at some later date be governed by the rules of a quite different program. DR managers, whether for the utility or the resident will come into being, merge, and disappear all of which will lead to heterogeneity, inconsistency, and changing elements.
- *Erosion of the people/system boundary*. The effectiveness of DR programs depends on incentivizing residents to enroll in a program and then on ensuring that they continue to be enrolled in the program.

Our approach to evaluating an ill defined architecture is to identify architectural options and find risks by considering the consequences of combinations on a system's ability to meet its goals in different situations. The architectural option identification is useful to any organization defining a new DR system (as many utilities are) or for organizations that have an existing DR program and wish to understand the potential risks.

We begin this paper by describing how we perform an evaluation when there is no defined architecture. We then describe the domain of residential DR which we have evaluated. This is followed by a specific description of the evaluation and some examples of the results. We conclude by describing the validation procedure we used for the results.

2. EVALUATING UNDERSPCIFIED ARCHITECTURES

1. Architectural evaluation methods [4] assume a specific problem to be solved and a specific proposed solution. In this section we discuss the techniques we used in adapting architecture evaluation to the case where a common architecture is not precisely defined, but enough commonality exists among likely variants (e.g., based on common goals and general solution directions) to allow some shared analysis. Though a specific architecture that would be used by all utilities does not exist in our case, a specific problem does exist. The problem needs to be specified in a fashion that is amenable to the evaluation process. We did this by generating a small collection of use cases that characterized the important aspects of the problem.
2. To perform an architectural evaluation, we hypothesized a collection of architectures. Generating a complete collection of possible architectures is, of course, infeasible; it leads to combinatorial possibilities. What we did, instead, was to enumerate the most important architectural decisions, as guided by our use cases. These are the architectural decisions that *must* be made by utilities. The choice of a particular architectural decision, or collection of decisions, is what is we focused on, and judged to be problematic (risky) during the evaluation.
3. The yardstick for measuring a potential architecture is the set of business goals for the system being analyzed. Potential problems are couched as risks to the achievement of the business goals for DR systems.
4. As with traditional architectural analysis, concrete scenarios are used as our testing mechanism. A scenario leads to a set of steps through a potential architecture, and it is that sequence that is examined for risks.
5. Risks are determined by looking for failures in one of the quality attributes while examining a scenario. At each step of testing, based on the scenario, the question is asked "What can go wrong to jeopardize satisfaction of the quality attribute requirements for some choice of options?"

These five techniques are related. Before performing an actual evaluation, we must define the problem and define a yardstick, manifesting it as testable quality attribute requirements. In addition, we provide a concrete set of test cases and generate the decisions to be tested. The order of the first four steps are basically independent. Finally, we use the output of the first four steps to discover possible failures or risks. We then consolidate risks into risk themes (as in ATAM [2]) to provide consolidated information for managers.

3. RESIDENTIAL DEMAND RESPONSE

The Smart Grid is the term used to refer to the enhancement of the electric power grid with digital technology to make it more reliable, secure and efficient [12]. One of the problems the Smart Grid intends to solve is the problem of peak demand, which typically happens on hot summer days in which an increased air conditioning load is added to the normal electric consumption profile [9]. To satisfy peak demand, stand-by power plants have to be brought on-line. These peaking plants, which are used for only about 1 or 2 percent of the hours in a year [3], have the highest marginal cost, and are the most inefficient and more polluting of all power plants [7].

Demand response is a key component of the Smart Grid whose main objective is to reduce peak load during periods of high demand for electricity. Currently, most consumers pay a fixed price for each kWh of electricity consumed, regardless of the cost of delivering that electricity. Therefore, they do not have a strong incentive to reduce load when there is high demand and (consequently) high cost. Demand response motivates "changes in electric use by end-use customers in response to changes in the price of electricity over time, or to give incentive payments designed to induce lower electricity use at times of high market prices or when grid reliability is jeopardized." [13] It should be noted that the effectiveness of any DR program depends on the positive response of the consumers. This exemplifies the erosion of the people/system boundary, one of the characteristic of ULS systems.

Even though DR is not a new concept and is already used with some industrial and large commercial consumers, it has been recognized that the residential sector presents the largest untapped potential for significant cost-savings in electricity production [6][11]. However, tapping the residential sector implies dealing with a much larger scale of users: 125 million out of the 144 million electricity customers in the U.S. are residential customers [5]. Furthermore, in the near future, smart appliances will be able to interact with DR systems on behalf of the user. The smart appliance market in the U.S. is expected to grow at a 40% compound annual growth rate between 2011 and 2015 [14]. This will add to the scale of DR not only in number of elements but also in the number of interactions.

4. PERFORMING THE EVALUATION

In this section, we describe how we applied the general principles for evaluating under-specified architectures that were described in Section 2.

4.1 Determine the Goals for the System

For ULS systems, determining the goals of the stakeholders can be difficult. For residential DR, the goal of the utility companies is primarily to be responsive to their regulators. The goal of the regulators is to balance the desire of consumers with the needs of the utility companies. The consumers wish to have uninterrupted power at low cost. There is an additional stakeholder, the US Department of Energy, which has no direct involvement in residential DR but which is providing technical guidance and moral suasion to encourage its adoption.

Through the National Energy Technology Laboratory (NETL), the Department of Energy has defined key success factors for the Smart Grid [8]. These factors are the "business goals" for the grid: reliable, secure, economic, efficient, environmentally friendly, and safe. Furthermore, the NETL defined the characteristics the Smart Grid should have to achieve those success factors:

- Self-heals
- Motivates and includes the consumer
- Resists attack
- Provides power quality for 21st century needs
- Accommodates all generation and storage options
- Enables markets
- Optimizes assets and operates efficiently.

But these goals are for the Smart Grid as a whole and are too high-level and abstract to support detailed analysis for residential DR. The following list shows how some of these goals were refined to express more specific goals for residential DR.

- Optimizes assets and operates efficiently
  - Load reduction is achieved in spite of intermittent network failures
  - Can scale to accommodate a large number of devices and protocols, including new ones
  - Is resilient to common mode failures of participating devices (e.g., programmable thermostats)
  - Can gracefully react to information network overload situations
  - System failures do not negatively impact consumers
- Motivates and includes the consumer
  - Provides relevant and timely feedback to consumers about their behavior
  - Gives a positive payoff to participating consumers
  - Does no harm to consumers that rely on electricity for critical needs such as medical conditions
  - Does not damage consumers' devices
  - Allows modification and evolution to induce sustained or greater levels of consumer enrollment
  - Allows consumers to participate with minimal effort (e.g., through automation)
  - Does not burden the consumer with impact of external change
- Resists attack
  - Prevents consumers and the system from being affected by attacks
  - Prevents uncooperative consumers from benefiting from wrongdoing
- Self-heals
  - Load reductions can be quickly enacted (or induced) to deal with loss of generation
  - Can recover from a power outage
- Enables markets
  - Supports growing numbers of consumer devices
  - Allows participation of a wide variety of DR service providers

These refinements of the NETL goals serve as the detailed business goals against which potential risks are evaluated in our subsequent analysis.

4.2 Document the Most Common Use Cases for the System

There is no standard architecture for residential DR. To construct the architecture landscape, we needed to understand the major architectural variants. We determined this through an analysis of existing pilot DR programs, through the knowledge of industry domain experts, and through vision documents produced under the auspices of the Department of Energy. In particular, we needed to understand the types of incentives proposed for residential DR systems as these drive the way that these systems work, and hence constrain many of the most important architectural choices. An understanding of the types of incentives, in turn, led us to develop four use cases that provide the basis for developing a list of common architectural elements.

The first use case represents the current situation for almost every utility, where no residential DR program is in place, as shown in Figure 1.

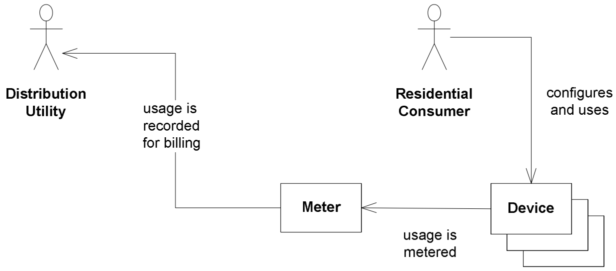

Figure 1: Electricity delivery without DR

Note that this, and the following figures are not architectural diagrams, but merely elicitation devices where stick figures represent actors, lines represent interactions, and rectangles represent key architectural elements.

The diagram in Figure 2, in contrast to that of Figure 1, shows an architecture in which a DR program based on direct load control (DLC) has been implemented. In this case, the utility directly controls some of a consumer's home appliances when load reduction is needed.

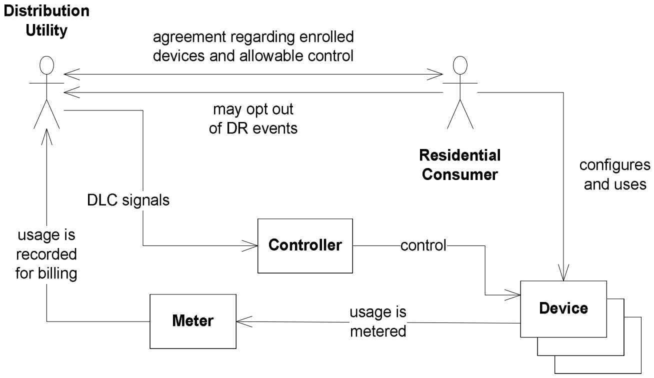

Figure 2: Direct load control of devices

Figure 3 shows a use case in which the utility provides pricing signals and it is up to the consumers to determine how their devices will react, if at all, to these signals. Note that in this case, the DR system uses an advanced meter, capable of recording not only how much electricity was consumed, but also when it was consumed.

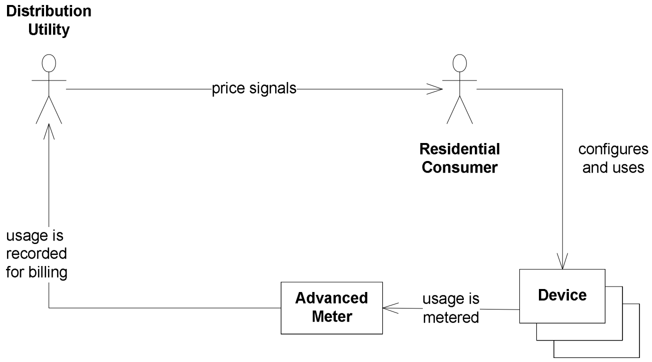

Figure 3: DR architecture with price signal without automated response

And finally, Figure 4 shows a use case in which the utility provides pricing signals and devices react automatically to these signals, with no human intervention other than initial configuration.

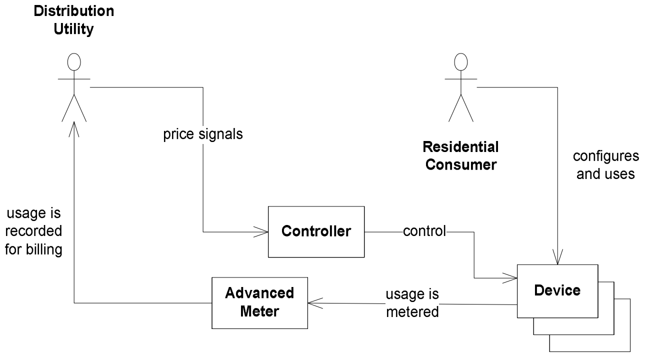

Figure 4: Control of devices through pricing signals

4.3 Enumerate Architectural Decisions

By considering the use cases, an architectural landscape may be drawn, showing the possible major alternatives for the creation and operation of the system. Each alternative represents an important architectural decision that must be made. The alternatives may be based upon logical options (e.g. push versus pull communications, acknowledgement of messages or not), commercially available components (e.g. types of networks available), or design decisions within an architectural element (e.g. frequency of communication).

We identified 23 different architectural decisions, each of which must be considered, instantiated, and refined by an architect when creating a specific system:

1. The protocol of interaction between the utility and any DR utility program manager must be established.
2. The pricing model and the rules for the DR program must be established.
3. If the DR rules are to be automatically executed, their location(s) must be decided.
4. The mechanisms for recording opt-out events from the resident and enforcing non-opt out circumstances must be decided.

5. The utility must decide which types of devices are to be supported and how they are to be commissioned and registered.
6. The utility must decide on the enrollment mechanism.
7. Residents must decide if they wish to have a DR resident manager operate the program for them.
8. The resident or their DR resident manager must decide whether an energy management system (EMS) is to be used to manage the devices registered in the DR program.
9. The utility, the utility DR program manager, the resident program manager, and the resident must decide on the type of feedback to be available to the resident.
10. Installation options must be chosen.
11. Controller options must be chosen
12. Commissioning options must be chosen.
13. Registration options must be chosen.
14. Inputs to the forecasting model must be chosen.
15. The type of carriers to be used for different portions of DR event propagation must be chosen.
16. How a DR signal is transformed as it progresses from a utility to a device must be decided.
17. Which portions of the DR signal transmission are push and which are pull must be decided.
18. Frequency of DR signals must be decided.
19. Granularity of data returned to the utility must be decided.
20. What data is available to the user instantaneously or retroactively must be decided.
21. Where personally identifiable information is removed from data must be decided.
22. Whether signals must be confirmed and/or recorded must be decided
23. How errors in devices are detected and reported must be decided.

Let us examine some of these decisions in more detail. For example, consider the architecture decisions that must be made to implement the communication from the utility to the devices in the second and fourth use cases. Some of these decisions are:

- The type of carriers (networks) to be used for different portions of DR event propagation must be chosen.
- How a DR signal is transformed as it progresses from a utility to a device must be decided.
- Frequency of DR signals must be decided.
- Inputs to the forecasting model must be chosen.
- Which portions of the DR signal transmission are push and which are pull must be decided.

Digging deeper, let us consider the implications of the last point, about push versus pull. A few of the considerations in the push/pull decision are:

- *Acknowledgment.* Reliability (and non-repudiation) can be enhanced by having explicit acknowledgements of DR messages. Acknowledgements introduce a tradeoff with performance: introducing more traffic on the network.
- *Redundancy*: The communication infrastructure and the repository may be unavailable for extended periods, especially if stressed because of high traffic. Reliability can be enhanced by having redundant networks or repositories, creating a tradeoff with complexity and cost.
- *Knowledge of recipient.* Non-broadcast push requires that the sender know the identity of the recipient. Adding or removing recipients from the list of recipients may be error prone. Push through broadcasting, on the other hand, does not require maintaining a list of recipients.
- *Data location.* Pull can be effected by having pricing values or DR signals on a central repository. These values can be the same for all users or can be customized for individuals or classes of users.
- *Effects of scale.* Pushing many messages simultaneously may stress the communication structure. Pulling may stress the repository if the pull is from a repository. There is little effect of scale on broadcast messages.
- *Device knowledge.* If the utility or the utility DR manager controls devices directly, it must have facilities for registration/deregistration of devices and the rules for controlling each device. This can be established during a commissioning/decommissioning activity.

If the architect chooses to have explicit acknowledgment messages, this will improve non-repudiation (for example, a consumer cannot claim that they didn’t receive a price change message). However, it will do so at the expense of having substantially greater network traffic. And the architect may choose to have redundant networking between parts of the system, or redundant servers or repositories, to avoid the risks of a single point of failure. But these will increase the cost of the system, and increase the complexity of building and evolving the system. In each case, these seemingly low-level decisions within the architecture landscape involve substantial quality attribute consequences and non-trivial tradeoffs.

At this point it should be obvious that the architectural landscape for this system is quite large, the architectural decisions are many and non-trivial, and these decisions may have far-reaching consequences for the many competing quality attributes that the architect will want to optimize.

4.4 Develop Scenarios Describing Challenges to the System from Multiple Quality Attribute Perspectives

It is not enough that an architecture works well under normal conditions (the use cases), but it must work well when stressed, when faced with unexpected demands or unexpected failures, or when faced with evolutionary pressures. Scenarios are chosen to understand the implications of such challenges on architectural decisions. The scenarios were generated by considering reliability, scalability, usability, performance, modifiability, the characteristics of ULS systems, and the goals enumerated in step 1 (see Section 4.1). We did not do an analysis for security because security is being treated quite extensively elsewhere in the Smart Grid community, although we did consider tradeoffs that are made to implement security in the grid.

We identified 30 distinct scenarios, covering stability, performance, modifiability, interoperabililty, reliabililty, evolvability, usability, safety, and several other quality attributes. We illustrate a portion of our set of scenarios with the seven shown below.

1. User is incentivized to change behavior, does so, and gets feedback on cost and environmental impact within minutes.
2. Utility deploys DR and time-of-use pricing simultaneously. A consumer signs up for a DR program, and the bill goes up. Consumer sues utility (class action lawsuit).
3. Consumer enrolls air conditioner in DR when well and later develops asthma, making the air conditioner safety critical.
4. DR is deployed, but only 10% of consumers enroll. Consequently a new DR program must be put in place.
5. A new DR program comes out that does not work with existing DR device: the options are 1) leave the device unsupported, 2) utility or consumer must replace the device; 3) vendor patches the device; 4) utility supports multiple variants of the device.
6. Shared network becomes stressed (due to one or more emergency situations). Network traffic due to DR does not negatively affect more critical functions (e.g., situational awareness).
7. Common software failure causes 50% of some devices (e.g., thermostats) to fail simultaneously in some way: 1) to stop responding; 2) to initiate a massive synchronous event; 3) to fail live and send large numbers of messages.

4.5 Identify Potential Risks

When scenarios are mapped onto an architectural landscape, the assumptions behind individual architectural decisions become evident. Some of these assumptions, alone or in combination, pose potential risks for the achievement of a system's quality attribute goals. This mapping, along with a model of each quality attribute, is the basis for the analysis in an ATAM [2].

In some cases those architectural decisions pose risks for the achievement of one or more scenarios. For each scenario, we walked through it, mapping it onto various architectural alternatives that might be used to achieve that scenario. During the walkthrough, we asked questions derived from models of the quality attributes that we were evaluating [1].

For example, when analyzing the scenario "DR is deployed, but only 10% of consumers enroll. Consequently a new DR program must be put in place" we arrived at the following potential risks (subject to the specific decisions a utility would make).

- *modifiability*: DR rules are not encapsulated, making new rules expensive and time-consuming to add
- *modifiability*: changes in DR rules may ripple to other parts of the DR architecture and to other parts of the enterprise (e.g., billing)
- *interoperability*: changes to the nature of communication may reduce the number of supported devices or require coordination with vendors and consumers to update devices
- *usability*: consumers may be confused about rapidly changing rules

4.6 Consolidate the Risks into Risk Themes for Strategic Planning

The final step of architectural analysis is to consolidate the risks that we find as a result of the scenario mapping process in Step 5 into *risk themes*. In mapping substantial numbers of scenarios, we often see the same kinds of risks emerging over and over. Such risk themes need to be explicitly identified as these pose the greatest risks to the success of the system. An architectural analysis exercise always locates many potential risks [2] but not all risks are equally likely and not all of them have the same set of consequences. The commonalities in the risks found have led us to "roll up" many of the risks into themes so that these may be made the focus of future investigations.

We developed four over-arching risk themes when analyzing residential DR architectures:

1. *Does too little*: A DR program is unable to control devices or influence consumers to reduce load in a timely fashion or to a sufficient degree to achieve goals.

   This could be due to resource contention, common mode failures, poor scaling with consumer and device enrollment, inadequate incentives, or adversarial interference.

2. *Does too much*: Grid operations are complicated or placed at risk by unpredicted large-scale load reductions related to DR program operations.

   This could be due to synchronized responses caused by automation that is triggered by a common global event (like time or simultaneous notification), common mode failure, or adversarial interference.

3. *Effects degrade over time*: Controlled load decreases over the life of a DR program.

   This could be due to consumer confusion or dissatisfaction resulting in unenrollment, stranded devices or protocols as the program evolves, or changes in the incentive mechanisms that reduce participation.

4. *Operational costs increase excessively*: DR program operating costs increase excessively (beyond predictions or expectations) over the life of the program.

   This could be due to costs to make changes as the program changes (e.g., to modify incentives or support new regulations), costs relating to providing support for an increasing numbers of devices, protocols, and versions, costs in coordinating changes with other stakeholders (e.g., vendors or third party service providers), or costs related to addressing consumer complaints or legal challenges.

4.6.1 Unpacking Risk Theme 1

Let us consider the first risk theme—"*Does too little*"—in some detail. One of the risks that contributes to this theme is a risk from one of the scenarios: "DR signal does not reach sufficient consumers: due to bandwidth limitations, bottlenecks, or a failure of a critical network, server, or class of devices". How did we arrive at this risk in the first place? It was motivated by a large number of scenarios.

Let us focus on just one of these, "Common software failure causes 50% of some devices (e.g., thermostats) to fail simultaneously in some way: 1) to stop responding; 2) to initiate a massive synchronous event; 3) to fail live and send large numbers of messages". Again, focusing in on just one part of this scenario, let us consider option 3: "to fail live and send large numbers of messages". What architectural elements are involved in this and what architectural decisions might lead to this risk being realized?

We identified the architectural elements that might be involved in realizing this risk: Devices, Controllers, Carriers, and EMSs (Energy Management Systems). Is it possible for devices to "fail live", thereby emitting large numbers of messages? Are devices tested and qualified for inclusion in a DR program? Does the DR program require, or even accept, any feedback from the devices? Similarly, are the Devices monitored? If so then Devices might provide some feedback. And how are the monitors themselves tested and qualified for inclusion in a DR program? Many of these same considerations apply to the EMS: how is it tested and qualified? An EMS typically does provide feedback as to the operation of the devices that it controls, which implies the possibility that it could fail live. Finally the Carrier will carry the signals from one architectural element to another. Does the Carrier monitor traffic? Can it detect anomalous behavior (in much the same way that an intrusion detection system monitors internet traffic, looking for denial of service attacks and other potential network disruptions)? Can network traffic be throttled and/or prioritized?

4.6.2 Unpacking Risk Theme 4

Let us consider another example, this one from risk theme 4: "*Operational costs increase excessively"*. One of the risks that contributes to this theme is the following: "DR rules change requiring consequent changes that are either slow or expensive: due to the inability of providers to easily modify their system software in response to program changes". How did we arrive at this risk in the first place?

It was motivated by three scenarios. We will focus on one of them: "A new DR program comes out that does not work with an existing DR device: 1. leave the device unsupported; 2. utility or consumer must replace the device; 3. vendor patches the device; 4. utility supports multiple variants of the device." What architectural elements are involved in this and what architectural decisions might lead to this risk being realized? Obviously Devices are involved, but our focus is really on the parts of the system that must interface with the Devices, namely: Controllers, EMSs, and potentially advanced Meters. Each of these elements could have an interface with the Device and when the Device changes there is the possibility that this interface changes, causing "ripple effects" to the other identified architectural elements. These ripple effects, if not contained, will lead to one of the three identified consequences of the scenario: obsolescence, new equipment, patching, or support for multiple variants. The degree to which ripple effects are contained—by standardized interfaces or by intermediaries that hide the details of the Devices, for example—will determine the degree to which this risk is realized.

4.6.3 Consequences for Strategic Planning

The answers to each of these questions—the architectural choices made or not made—will determine the extent to which the scenario might turn out to be an actual risk to the operation of a residential DR program as implemented by a specific utility for their specific market. Risks are always *potential* problems, and we strive for early identification of risks as a means of preventing them from being realized.

Each risk might then be further analyzed in more detail, e.g., by building a queuing model of performance, by building a Markov model of availability, or by creating a simulation, experiment, or prototype. And each of the architectural decisions that goes in to this risk should get particularly keen scrutiny by any architect designing a residential DR system.

5. VALIDATION

What we have now presented is a description of an architectural landscape for residential DR and an analysis of the potential risks lurking within that landscape, their consequences, and their interactions and tradeoffs. But how do we validate the methods and conclusions of this analysis?

There are two aspects to validating the risks discovered in *any* architecture analysis: coverage and correctness.

1. Did we find a majority of the most important risks?
2. Were the risks that we found truly significant challenges to the achievement of some important system goal?

When performing a normal architectural evaluation such as an ATAM, correctness and coverage are reasonably easy to achieve, as long as the method is faithfully prosecuted. Coverage is achieved in the process by ensuring that the appropriate stakeholders are involved, by capturing their concerns as scenarios, and by tracing the most highly ranked scenarios through the architecture. Correctness is achieved because the determination of risk is done in real time, in front of all the stakeholders. If risks are misunderstood or misidentified, these mistakes are immediately apparent and corrected.

When evaluating an underspecified, ULS architecture with multiple unknown and unknowable stakeholders, it is not possible to get agreement on the risks. So we must turn to notions of statistical coverage to validate the efficacy of our method. This has not yet been done.

Instead we have relied on input and feedback from experts within the Department of Energy and the utility industry to validate our results. The consensus has been that our process has revealed substantial architectural risks and issues "that are either glossed over or completely overlooked".

A proper statistical validation of the results found via our architecture analysis remains, however, as future work, requiring the engagement of a broad community of utility companies and their key stakeholders.

6. CONCLUSIONS

In this paper we had two goals, a narrow one and a broad one. The narrow goal was to analyze potential system architectures for residential DR. In doing so we revealed the landscape of architectural decisions that may be considered. We further showed how this early analysis can discover latent risks in the architectural decisions made (or not yet made) that may affect the key quality attributes of such systems: performance, usability, reliability, and evolvability.

The broader goal was to describe an approach for systematically finding such risks in any architecture that has ULS characteristics: many stakeholders (users, operators, owners) with sometimes conflicting goals, important socio-technical considerations (where humans are a key part of the system and have an substantial role in its success), ubiquitous evolution, and—most importantly—no single architecture.

Of course, this work is not without its limitations and could be extended in several interesting ways. Two of the most important limitations of this work are:

1. All of the analysis presented here was made based on publically available documents. We are dependent on their accuracy for the accuracy and breadth of coverage of the results reported here.
2. We did not investigate how to mitigate the majority of the risks discovered. The focus of this investigation was on analysis, and not (re-)design.

These limitations are all easily addressed—we certainly could interview a broader set of stakeholders and we could investigate mitigations for each of the risks found. These extensions are consistent with an architecture-centric life-cycle approach to the Smart Grid. As such, these two limitations point to two obvious extensions of this paper:

1. Examine a set of non-trivial pilot residential DR programs from the point of view of their scalability, modifiability, and reliability, as well as their user satisfaction and retention.
2. Explore strategies for mitigating the scalability, modifiability, reliability, and usability risks that we identified. Prototypes, modeling, and simulation are all techniques that could be used to explore mitigation strategies.

In addition, the architectural analysis technique we used is, as we have said, applicable to other portions of the Smart Grid (e.g. Microgrids) or to other complex portions of critical national infrastructure such as water and sewer systems, networks of gas pipelines, or emergency responder systems and, indeed, to any system that exhibits ULS characteristics.

7. ACKNOWLEDGMENTS

We would like to thank the U.S. Department of Energy, and the RTSS program at the Software Engineering Institute, for their support for this research.

We would also like to thank Charles McParland (Lawrence Berkeley National Laboratory), Ed Koch (Akuacom), Michael Ingram (Tennessee Valley Authority), Wayne Longcore (Consumers Energy) for their input to this research.

8. REFERENCES

[1] L. Bass, P. Clements, R. Kazman, *Software Architecture in Practice*, 2nd ed., Addison-Wesley, 2003.

[2] P. Clements, R. Kazman, M. Klein, *Evaluating Software Architectures: Methods and Case Studies*, Addison-Wesley, 2001.

[3] Robert Earle, Ahmad Faruqui, Toward a New Paradigm for Valuing Demand Response, The Electricity Journal, Volume 19, Issue 4, May 2006, Pages 21-31, ISSN 1040-6190, DOI: 10.1016/j.tej.2006.03.006.

[4] D. Falessi, G. Cantone, R. Kazman, P. Kruchten, "Decision-making Techniques for Software Architecture Design: A Comparative Survey", *ACM Computing Surveys*, to appear, 2011.

[5] Faruqui, Ahmad and Sergici, Sanem, *Household Response to Dynamic Pricing of Electricity - A Survey of the Empirical Evidence* (February 2010).

[6] Federal Energy Regulatory Commission, *A National Assessment of Demand Response Potential*, June 2009.

[7] Koomey, Jonathan, & Brown, Richard E.(2002). *The role of building technologies in reducing and controlling peak electricity demand*. Lawrence Berkeley National Laboratory: Lawrence Berkeley National Laboratory. LBNL Paper LBNL-49947.

[8] National Energy Technology Laboratory, "A Systems View of the Modern Grid", http://www.netl.doe.gov/smartgrid/referenceshelf/whitepapers/ASystemsViewoftheModernGrid_Final_v2_0.pdf, Retrieved August 20, 2010.

[9] Newsham, Guy R. and Bowker, Brent G., (2010), The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review, *Energy Policy*, 38, issue 7, p. 3289-3296, http://dx.doi.org/10.1016/j.enpol.2010.01.027.

[10] Northrop, L., Feiler, P., Gabriel, R., Goodenough, J., Linger, R., Longstaff, T., Kazman, R., Klein, M., Schmidt, D., Sullivan, K., and Wallnau, K. *Ultra-Large-Scale Systems: The Software Challenge of the Future*. Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 2006.

[11] Salmi-Klotz, Jason. *FERC Policy on Demand Response and Order 719*. Proceedings of Grid Interop 2009.

[12] U.S. Department of Energy. *Smart Grid System Report*. July 2009.

[13] U.S. Department of Energy, *Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them - A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005.*

[14] Zpryme Research & Consulting, *Smart Grid Insights: Smart Appliances.* March 2010. http://www.zpryme.com/SmartGridInsights/2010_Smart_Appliance_Report_Zpryme_Smart_Grid_Insights.pdf