A Demand-Response Calculus with Perfect Batteries

Jean-Yves Le Boudec and Dan-Cristian Tomozei[*]

EPFL - LCA2
Bâtiment BC
Station 14
CH-1015 Lausanne
{jean-yves.leboudec, dan-cristian.tomozei}@epfl.ch

Abstract. We consider an electricity consumer equipped with a perfect battery, who needs to satisfy a non-elastic load, subject to external control signals. The control imposes a time-varying upper-bound on the instantaneous energy consumption (this is called "Demand-Response via quantity"). The consumer defines a charging schedule for the battery. We say that a schedule is *feasible* if it successfully absorbs the effects of service reduction and achieves the satisfiability of the load (making use of the battery). Our contribution is twofold. (1) We provide explicit necessary and sufficient conditions for the load, the control, and the battery, which ensure the existence of a feasible battery charging schedule. Furthermore, we show that whenever a feasible schedule exists, we can explicitly define an online (causal) feasible schedule. (2) For a given arrival curve characterizing the load and a given service curve characterizing the control, we compute a sufficient battery size that ensures existence of an online feasible schedule. For an arrival curve determined from a real measured trace, we numerically characterize the sufficient battery size for various types of service curves.

1 Introduction

Growing requirements for integration of renewable energy in the power grid pose a difficult challenge to Distribution Systems Operators (DSO). The increase in the penetration level of renewables, coupled with the volatility and lack of predictability of such energy sources (e.g., wind turbines, photovoltaic cells) may produce instabilities in the transport and distribution grid, which in turn may result in blackouts. It is possible that such sources produce too little energy when demand is high, or too much energy when demand is low. Studies have shown [1] that in order to guarantee a smooth operation of the grid in its current state, operators need to react quickly to changes in renewable output via use of ancillary services (e.g., gas turbines). Such services are able to provide large quantities of energy on short notice; however they constitute an investment that actors in the energy market are reluctant to perform, since it indirectly increases the cost of "green" energy.

A method proposed in the next generation grid for dealing with variability of production is Active Demand Management, or Demand-Response. Specifically, at critical peak periods, Demand-Response mechanisms throttle flexible energy demands (called "loads"), such as Plug-in Electric Vehicles, heating, or air conditioning, thus adjusting demand to the available production. The control of electricity consumption is done either via *price*, or via *quantity*.

Demand-Response via price consists simply of informing consumers of an increase in electricity price, which provides an incentive for consumers to reduce their load. This method has the advantage of being decentralized and is widely embraced in the literature [6]. Its main drawback is the fact that it exposes consumers to the variability of electricity prices. It has been shown [2] that real-time wholesale prices can fluctuate wildly in a fair deregulated electricity market.

The second option, Demand-Response via quantity, has been implemented by companies such as Voltalis or Peaksaver [11, 8]. These companies have control over a large number of flexible loads (typically heating appliances) and act as virtual energy providers during peak consumption hours by throttling the loads of their customers, while guaranteeing hard limits on the service reductions. Such an approach can be modeled via service curve-based agreements between suppliers and consumers [5]. In this paper, we focus on Demand-Response via quantity.

In both cases, the effect of Demand-Response can be interpreted as a time-varying imposed upper-bound on electricity consumption from the grid. Indeed, when control is done via pricing, such an upper-bound can be computed locally by a Smart Home Controller, which aims to minimize electricity cost. When control is done via quantity, the upper-bound is explicit. We refer to this imposed upper-bound as *the control signal* received from the grid.

Demand-Response requires that loads be flexible, which is naturally the case for some appliances, as mentioned earlier, but not for most. One way to make a load flexible is to insert some form of energy storage (we call it a "battery") between the load and the electrical grid. The problem we study in this paper is how to size the battery, and how to schedule its charge, in order to make sure that the inflexible load can be served transparently, in presence of Demand-Response via quantity.

Contributions We consider a non-elastic load equipped with a battery (interposed between the load and the grid), subject to time-varying control signals (e.g., via throttling from Demand-Response mechanisms). Despite imposed reductions in energy consumption from the grid, the load should be satisfied (from the grid and the battery). In order to ensure this, the battery needs to act like a buffer and provide energy to the load when the grid is unable to do so. We make the hypothesis that the battery is perfect (i.e., there are no thermal losses, and its efficiency is 1).

In this paper we study battery charging schedules that ensure satisfiability of the load. We call such schedules *feasible schedules*. The first main contribution consists in providing explicit necessary and sufficient conditions on the load, the control, the battery capacity, and the initial charge level of the battery, that guarantee existence of such feasible charging schedules. Moreover, we show that there exist feasible schedules, if and only if there exist *causal* feasible schedules (i.e., schedules that can be computed online, without knowledge of the future).

In addition, we examine the problem of sizing the battery. We consider loads that satisfy an arrival curve and control signals that are constrained by a service curve. Such control signals have been considered for performing Demand-Response [5]. The second main contribution of this paper consists in determining a sufficient battery capacity that ensures existence of feasible online charging schedules for all loads (and control signals) characterized by a given arrival curve (and a given service curve, respectively). Using an arrival curve computed based on real measurement traces, we characterize numerically the dependence between the battery size and various types of service curves that can be supported.

Related Work The problem of determining feasible battery charging schedules is somewhat related to the problem of optimal multimedia smoothing [4, Chapter 5]. The analysis of the latter employs techniques from min-plus system theory. We provide a novel formulation of the former and analyze it using similar min-plus system techniques: we determine the set of all feasible charging schedules (Theorem 1) and compute sufficient battery capacity (Theorem 4), assuming a perfect battery.

Outline In Section 2 we formally define the problem. In Section 3 we characterize the feasible battery charging schedules. In Section 4 we study bounds on the required battery size which ensure existence of feasible online schedules. In Section 5 we present numerical evaluations on real traces. We conclude in Section 6. Technical lemmas are given in Appendix A.

2 Problem Definition

We consider a load $L(t) = \int_0^t \ell(s)ds$, i.e., energy required over the interval $[0, T]$ by a consumer who is subject to control signals $G(t) = \int_0^t g(s)ds$ (i.e., maximum available energy over the interval $[0, T]$). The consumer owns a battery of capacity B. We assume a perfect battery, with no thermal loss and perfect efficiency (i.e., all the energy stored in the battery can be retrieved at a later time).

The *actual load* $U(t) = \int_0^t u(s)ds$ is the load drawn from the grid. It is computed by the consumer, and may differ from $L(t)$, since the consumer may use the battery instead of the main supply; we say that $U(t)$ is a "schedule". Let B_0 be the battery level at time $t = 0$. An actual load $U(t)$, defined for $t \geq 0$, is feasible if and only if for all times s, t such that $0 \leq s \leq t$:

$$L(t) \leq B_0 + U(t), \tag{1}$$
$$U(t) - L(t) + B_0 \leq B, \tag{2}$$
$$0 \leq U(t) - U(s) \leq G(t) - G(s). \tag{3}$$

Eq.(1) expresses the no-underflow condition for the battery, i.e., all load comes either from the main supply $U(t)$ or from the initial buffer B_0. Eq.(2) expresses the no-overflow condition. Eq.(3) expresses that the supply is positive, i.e., there is no local production, and the supply is constrained by the control signals $G(t)$. Note that if U is feasible, the battery level at time t is

$$B(t) = U(t) - L(t) + B_0. \tag{4}$$

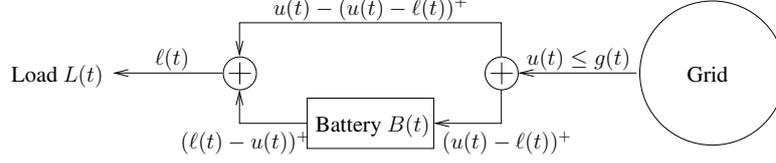The consumer setup is represented in Figure 1.

Fig. 1. The consumer setup. The load $L(t)$ can be satisfied from the battery $B(t)$ and the grid $U(t)$ under DR constraints given by control $G(t)$. We have denoted $(a)^+ = \max(a, 0)$.

The control signals $G(t)$ and the load $L(t)$ are given. The consumer-side problem is to compute an online schedule $U(t)$, i.e., a function $U(t)$ whose definition depends only on the past and present values of L, G and U. A typical example is the policy $\tilde{U}$ that greedily maximizes battery level subject to constraints given by Eq.(2) and Eq.(3).

Before turning to the online problem, consider an oracle that knows the past, present and future of the load $L(t)$ and the control signals $G(t)$ and asks the question whether there exists an actual load $U(t)$ that is feasible. We assume that $U(t) = 0$ for all $t \leq 0$, i.e., there was no load before time 0. We find necessary and sufficient conditions for the existence of such a schedule. We also show that any feasible schedule is bounded by a minimum $U^*_{\min}$ and a maximum $U^*_{\max}$. We show that the schedule $U^*_{\max}$ is computable online and corresponds to the policy $\tilde{U}$ that greedily maximizes the battery level.

Returning to the online problem, our result implies that if there exists a feasible schedule, then there is an online one. This comes as a surprise, since the set of online schedules is much smaller than the set of all possible schedules available to the oracle. In particular, our result implies that computing a feasible schedule is either impossible, or can be done online.

In Section 4 we consider the case where the control signals $G(t)$ are constrained by the service curve β. We derive from the previous results the values of minimum battery capacities that are required to guarantee existence of a schedule.

Our results use min-plus system theory [4, Section 4.1 and Theorem 4.3.1].

3 Feasible Schedules

Theorem 1. *Assume $L(t)$ and $G(t)$ are known and $0 \leq B_0 \leq B < \infty$. There exists a feasible schedule $U(t)$ if and only if both conditions hold:*

$$B_0 \geq \sup_{t \geq 0} \left(L(t) - G(t) \right), \tag{5}$$

$$B \geq \sup_{0 \leq s \leq t} \left(L(t) - L(s) - G(t) + G(s) \right). \tag{6}$$

If these conditions hold, there exist two feasible schedules $U^*_{\min}$ and $U^*_{\max}$:

$$U^*_{\min}(t) = \max \left[0, \sup_{\tau \geq t} \left(L(\tau) - G(\tau) \right) + G(t) - B_0 \right], \tag{7}$$

$$U_{\max}^*(t) = \min\left[G(t), \inf_{s:0\le s\le t}(L(s) + G(t) - G(s)) + B - B_0\right], \qquad (8)$$

such that any feasible schedule U satisfies

$$U_{\min}^*(t) \le U(t) \le U_{\max}^*(t) \text{ for all } t \ge 0. \qquad (9)$$

Proof. The idea of the proof is first to relax constraint (1); the relaxed problem has a maximal solution, $U_{\max}^*$, and the original problem is feasible if and only if $U_{\max}^*$ satisfies the constraint (1). Similarly, by relaxing constraint (2), we obtain a problem with a minimum solution $U_{\min}^*$.

More precisely, consider the problem

$$(P1)\begin{cases} U \text{ is non decreasing and } U(0) = 0 \\ U(t) \le B - B_0 + L(t) \text{ for all } t \ge 0 \\ U(t) - U(s) \le G(t) - G(s) \text{ for all } t \ge 0 \text{ and } s \ge 0 \text{ with } s \le t, \end{cases}$$

where the unknown is the function $U : \mathbb{R}^+ \to \mathbb{R}^+$. This problem is equivalent to Problem $(P10)$ of Lemma 1 with

$$f(t) = B - B_0 + L(t) \text{ for all } t > 0,$$
$$f(0) = 0,$$
$$g(t) = G(t) \text{ for all } t \ge 0.$$

Thus $(P1)$ has a maximal solution, let us call it $U_{\max}^*$, given by

$$U_{\max}^*(t) = \inf_{0\le s\le t}(g(t) - g(s) + f(s))$$

which, after some re-arrangements, gives Eq.(8)

Now we show that there exists a feasible schedule if and only if $U_{\max}^*$ satisfies the relaxed condition Eq.(1). One implication is obvious: if $U_{\max}^*$ satisfies Eq.(1), then it is a feasible schedule. Conversely, assume that there exists a feasible schedule, say U. Then U is a solution of Problem $(P1)$ therefore $U(t) \le U_{\max}^*(t)$ for all t. Since U satisfies Eq.(1), so does $U_{\max}^*$.

Last, saying that $U_{\max}^*$ satisfies Eq.(1) is equivalent to saying that both

$$U(t) = G(t) \text{ and } U(t) = \inf_{s \text{ such that } 0\le s\le t}(L(s) + G(t) - G(s)) + B - B_0$$

satisfy Eq.(1), which gives Eq.(5) and Eq.(6).

It remains to show that there is a minimum solution. To this end, we consider the problem $(P2)$, obtained by relaxing Eq.(2):

$$(P2)\begin{cases} U \text{ is non decreasing and } U(0) = 0 \\ U(t) \ge -B_0 + L(t) \text{ for all } t \ge 0 \\ U(t) - U(s) \le G(t) - G(s) \text{ for all } t \ge 0, \end{cases}$$

where the unknown is the function $U : \mathbb{R}^+ \to \mathbb{R}^+$. $(P2)$ is equivalent to Problem $(P20)$ of Lemma 2 with

$$f(t) = \max\left(0, L(t) - B_0\right) \text{ for all } t \geq 0,$$
$$g(t) = G(t) \text{ for all } t \geq 0.$$

Thus $(P2)$ has a minimal solution, let us call it $U^*_{\min}$, given by

$$U^*_{\min}(t) = \sup_{s:s \geq t} \left(g(t) - g(s) + f(s)\right),$$

which, after some re-arrangements, gives Eq.(7). Note that there is a feasible solution to the original problem if and only if $U^*_{\min}$ satisfies the relaxed condition Eq.(2). After some algebra, this gives Eq.(5) and Eq.(6) which were already obtained. It follows that, if these conditions hold, any feasible schedule U is a solution of Problem $(P2)$ and therefore $U(t) \geq U^*_{\min}(t)$ for all $t \geq 0$. $\qquad\square$

Note that $U^*_{\max}$ is causal, i.e., can be computed online. In contrast, $U^*_{\min}$ is not causal; it is in fact *anti-causal*, i.e., it depends only on the present and the future of L and G. Also note that different schedules result in different energy flows, however, by Eq.(4), $U(t) - B(t)$, i.e., the bought energy minus the stored one, is the same for all feasible schedules, since we assume that there is no loss.

The maximum schedule introduced in Theorem 1 is causal but is defined in abstract terms; this was required to prove the theorem. We show next that it corresponds to a simple, implementable online policy. To see this, consider discrete time and define the policy $\tilde{U}(t) := \sum_{s=0}^{t} \tilde{u}(s)$, where

$$\tilde{u}(t+1) = \min\left[g(t+1),\ B - \tilde{B}(t) + \ell(t+1)\right], \tag{10}$$

with $\tilde{B}(t)$ denoting the battery level at time t when the policy $\tilde{U}$ is used. This policy greedily maximizes battery level subject to constraints given by Eq.(2) and Eq.(3), i.e., it consists in storing into the battery the maximum possible amount of energy, considering the constraints that (1) the load needs to be served, (2) the battery has a maximum capacity B and (3) the energy drawn from the grid cannot exceed the control signals.

Theorem 2. *The schedules $U^*_{\max}$ and $\tilde{U}$ are equal.*

Proof. Let us show by induction on t that

$$U^*_{\max}(t) - U^*_{\max}(t-1) = \tilde{u}(t). \tag{11}$$

Denote by $B^*(t) = U^*_{\max}(t) - L(t) + B_0$ the battery level at time t when the maximal schedule $U^*_{\max}$ is used.

At time $t = 1$, since $B - B_0 > 0$,

$$U^*_{\max}(1) - U^*_{\max}(0) = \min\left[G(1),\ L(1) + B - B_0\right] - 0 =$$
$$\min\left[g(1),\ B - B(0) + \ell(1)\right] = \tilde{u}(1).$$

Suppose that Eq.(11) holds at times $1, \dots, t$. We show that it also holds at time $t+1$:

$$
\begin{aligned}
U^*_{\max}(t+1) = \min \big[& G(t) + g(t+1), \\
& \inf_{s:0 \leq s \leq t} \left(L(s) + G(t) - G(s) \right) + B - B_0 + g(t+1), \\
& L(t) + \ell(t+1) + B - B_0 \big] \\
= & U^*_{\max}(t) + \min \left[g(t+1), \; \ell(t+1) + B - B^*(t) \right].
\end{aligned}
$$

Since $B^*(t) = \tilde{B}(t)$ by the induction hypothesis, we can conclude. $\qquad \square$

4   Battery Size with Service Curve Constraints

Assume now that the control signals $G(t)$ are constrained by the service curve β, i.e., $G(t) - G(s) \geq \beta(t-s)$ for all $0 \leq s \leq t$, where β is a positive super-additive function, known in advance. The super-additivity of β is implicit, since it defines a minimum guaranteed value over any sliding window, and thus $\beta(t + t') \geq \beta(t) + \beta(t')$, for all non-negative t, t' [5, Section IV]. A direct consequence of Theorem 1 is:

Theorem 3. *Assume $L(t)$ and the service curve β are known and $0 \leq B_0 \leq B < \infty$, with the service curve satisfying the condition*

$$
\bar{\beta}(v) := \sup_{u \geq 0} [\beta(u + v) - \beta(u)] < \infty. \tag{12}
$$

Then there exists a feasible actual load $U(t)$ for any realization of the control signal $G(t)$ compatible with the service curve β if and only if both conditions hold:

$$
B_0 \geq \sup_{t \geq 0} \left(L(t) - \beta(t) \right), \tag{13}
$$

$$
B \geq \sup_{0 \leq s \leq t} \left(L(t) - L(s) - \beta(t - s) \right). \tag{14}
$$

Proof. Consider a control signal G constrained by β. It satisfies $G(t) - G(s) \geq \beta(t-s)$ for all $0 \leq s \leq t$. If Eq.(13) and Eq.(14) are satisfied, then it follows that Eq.(5) and Eq.(6) are verified. Thus, by Theorem 1, there exists a feasible actual load $U(t)$.

Conversely, we need to show that if for all controls G constrained by β there exists a feasible actual load, then Eq.(13) and Eq.(14) hold. By contraposition, it suffices to show that if either Eq.(13) or Eq.(14) do not hold, then there exists a schedule $\bar{G}$ for which there is no feasible actual load.

If Eq.(13) does not hold, then there exists some $t_0 \geq 0$ such that $B_0 < L(t_0) - \beta(t_0)$. We define $\bar{G}(t) := \beta(t)$. This control is compatible with β via the super-additivity property of the service curve. Hence, $B_0 < L(t_0) - \bar{G}(t_0)$, and thus Eq.(5) does not hold for $\bar{G}$. By Theorem 1 it follows that there exists no feasible actual load.

If Eq.(14) does not hold, then there exist some $t_0 \geq s_0 \geq 0$ such that $B < L(t_0) - L(s_0) - \beta(t_0 - s_0)$. Consider some constant $A > 0$ and define the following control:

$$
\bar{G}(s_0 + u) = \begin{cases} A + \beta(u), & \text{if } u \geq 0 \\ A - \bar{\beta}(|u|), & \text{if } u < 0. \end{cases}
$$

Let us check that $\bar{G}$ satisfies the service curve β.

First, consider $0 \leq v \leq u$. Then $\bar{G}(s_0 + u) - \bar{G}(s_0 + v) = \beta(u) - \beta(v) \geq \beta(u - v)$, by super-additivity of β.

Moreover, $\bar{G}(s_0 - v) - \bar{G}(s_0 - u) = \bar{\beta}(u) - \bar{\beta}(v) \geq \beta(u - v)$. Indeed,

$$\bar{\beta}(v) + \beta(u - v) = \sup_{w \geq 0}[\beta(w + v) + \beta(u - v) - \beta(w)] \leq \sup_{w \geq 0}[\beta(w + u) - \beta(w)] = \bar{\beta}(u),$$

again by super-additivity of β.

Finally, consider $u, v \geq 0$ and let us check that $\bar{G}(s_0 + u) - \bar{G}(s_0 - v) \geq \beta(u + v)$. We can rewrite the above inequality as $\beta(u) + \bar{\beta}(v) \geq \beta(u + v)$. But this holds by definition of $\bar{\beta}(v) := \sup_{u \geq 0}[\beta(u + v) - \beta(u)]$.

To ensure the positivity of $\bar{G}$, it suffices to take a large enough constant, for example $A = \bar{\beta}(s_0)$. Thus, $\bar{G}$ is a valid control satisfying β. By definition, $\bar{G}(t_0) - \bar{G}(s_0) = \beta(t_0 - s_0)$. Since $B < L(t_0) - L(s_0) - \bar{G}(t_0) + \bar{G}(s_0)$, we have that Eq.(6) does not hold for $\bar{G}$. Again, by Theorem 1, there exists no feasible actual load for $\bar{G}$. $\square$

Note that the constraint on β expressed in Eq.(12) simply states that over any time window of finite length, the service curve β guarantees only a finite amount of energy, which depends on the length of the window. It is not a strong assumption in this context.

Let α be a sub-additive function of $t \geq 0$, nondecreasing and such that $\alpha(0) = 0$. We say that the load $L(t)$ is α-smooth iff

$$L(t) - L(s) \leq \alpha(t - s) \text{ for all } 0 \leq s \leq t. \tag{15}$$

The smallest such possible α for which a given load $L(t)$ is α-smooth is obtained via min-plus deconvolution:

$$\alpha^*(t) = (L \oslash L)(t) \overset{\text{def}}{=} \sup_{s \geq 0}[L(t + s) - L(s)].$$

It follows that we can compute the minimum required battery capacity:

Theorem 4. *Assume an α-smooth load and let control signals be constrained by β. Let*

$$B^* := \sup_{s \geq 0}\left(\alpha(s) - \beta(s)\right). \tag{16}$$

If $B_0 \geq B^*$ there exists a feasible load schedule, for any realization of the control signal $\bar{G}(t)$ compatible with the service curve β. One such schedule is computable online, using only past and present observations.

Note that the condition on the initial battery level $B_0 \geq B^*$ implies the condition on battery capacity $B \geq B^*$.

Proof. Since $\alpha(t) \geq L(t) - L(0)$ and $L(0) = 0$, it follows that

$$B^* \geq \sup_{t \geq 0}(L(t) - \beta(t)).$$

Additionally, for any $0 \leq s \leq t$, we have that $\alpha(t - s) \geq L(t) - L(s)$, and hence,

$$B^* \geq \sup_{t \geq 0}(L(t) - L(s) - \beta(t - s)).$$

Thus, $B \geq B_0 \geq B^*$ implies Eq.(13) and Eq.(14). By Theorem 3 we conclude. $\square$

5 Numerical Evaluations

Consider now a data center that subscribes to Demand-Response via quantity to reduce its electricity costs. This allows an external controller to reduce power consumption of the data center via load control signals at peak hours. The DR contract guarantees an upper bound on the amount of service interruption per day. We consider that the load of the data center is not controllable, and we wish to render the DR mechanism transparent. To this end, we need to install batteries in the data center. We make use of Theorem 4 to determine the required capacity of the batteries. The service curve β is agreed upon in the DR contract. To determine a reasonable arrival curve α we use measured loads. We again assume that the batteries are perfect.

5.1 Voltalis-like Service Curve

We consider the specific example of Voltalis's BluePod [11], which disconnects appliances from the grid for at most 30 minutes per day. In addition, we impose a maximum allowed power consumption $z_{\max}$ which can never be overcome. Previously [5] we have proposed the following service curve: for some $0 \leq t_0 < t_1$ and some $z_{\max} > 0$, define

$$\beta_1(t) := z_{\max}(t_1 - t_0)\lfloor t/t_1 \rfloor + z_{\max}(t - \lfloor t/t_1 \rfloor t_1 - t_0)^+, \tag{17}$$

where $(a)^+ := \max(a, 0)$. The Voltalis on/off control satisfies the service curve β_1 with $t_0 = 30$ mn and $t_1 = 24$ hours (see Figure 2). However, the service curve β_1 allows more complex controls, such as limiting the power consumption to $z_{\max}/2$ over a time interval of at most $2t_0$ per day.
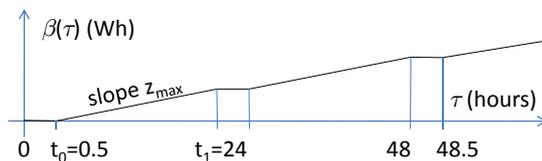

Fig. 2. The service curve β_1, shown here, allows the distributor to switch off the load for at most 30 mn every day, or (for example) to reduce the load to $z_{\max}/2$ for 60 mn every day.

5.2 Akamai Arrival Curve

We measured the power consumption of a desktop PC under different loads. The PC is equipped with a dual core Intel Pentium processor (2.8GHz) and is running Fedora Linux. We used the SPECpower_ssj2008 benchmark software developed by SPEC [10] and a power analyzer. The benchmark models a server application with a large number of users. It targets 11 possible request rates, from the maximum supported rate (100%

load) to 0-rate (idle, 0% load). Requests arrive at random intervals with lengths following an exponential distribution. Bursty arrivals result in requests queuing up. The process starts with a calibration period, during which the maximum supported request rate is determined, i.e., the maximum number of requests per second that can be treated. Subsequently, for each load, the consumption is measured as follows: a ramp-up period during which the specific load is targeted (as a fraction of the measured maximum load) is followed by a measurement period, and then by a ramp-down period. During the measurement period, power consumption values are stored, and average power consumption is computed. The resulting values are shown in Figure 3(a).
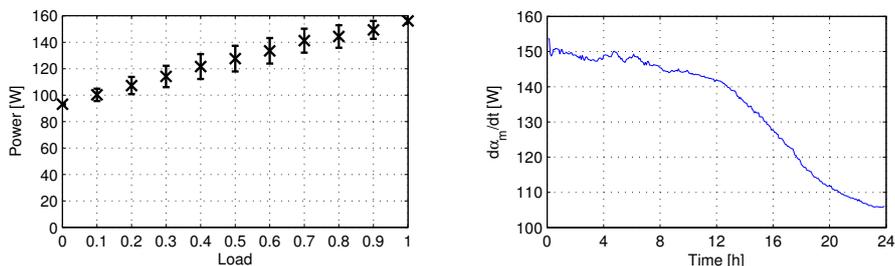
(a) Measured power for load varying from 0 to 100%. The error bars represent empirical standard deviation.

(b) Plot of $\frac{d\alpha_m}{dt}$ determined via min-plus deconvolution for the Akamai data set.

Fig. 3. Measured arrival curve.

We used the real traffic data from the Akamai data set presented in the work of Qureshi et al. [9]. There, the authors plotted incoming traffic at roughly half the servers of Akamai (in millions of hits per second) measured over a period of 24 days. We regard this data as a representative real-world workload. We assume that all the servers in a cluster are given roughly the same workload. This assumption allows us to map the Akamai traffic data to the load (from 0 to 1) of a single server.

Several models for estimating dynamic power usage [9, 3] assume that the power consumed by a cluster is roughly proportional to its utilization. We use a similar model: we map server load to power consumption using the roughly-linear power measurements in Figure 3(a) obtained as described above. We thus obtain a realistic estimation of the power consumption of a single server over 24 days (Figure 5).

We use this estimated power consumption to derive a minimum measured arrival curve $\alpha_m(\cdot)$ by applying min-plus deconvolution. Since the length of the trace is of $T_{\max} = 24$ days, we determine $\alpha_m(t) := \sup_{0 \leq s \leq T_{\max} - t}[L(t+s) - L(s)]$ for t less than some value $\tau \leq T_{\max}$ (i.e., we determine the maximum energy consumption over time windows of size less than τ). For visualization ease, we plot the derivative of the measured arrival curve (rather than the curve itself) in Figure 3(b) for $\tau = 24$ hours.

5.3 Battery Sizing

Let us now determine feasible parameters of the service curve β_1 (with $t_1 := 24\text{h}$). Namely, we seek pairs $(z_{\max}, t_0)$ of maximum allowed power and interruption time per day which allow satisfiability of α_m-smooth loads using a battery of finite capacity.

Theorem 4 states that taking the supremum over the difference between the arrival curve α and the service curve β provides a sufficient battery size. Theorem 5 below gives a sufficient condition for the finiteness of this supremum. Namely, the guaranteed delivered energy needs to be greater or equal than the maximum required energy on time windows of some length T. Moreover, the same theorem allows us to compute B^* by taking the supremum over the interval $[0, T]$, instead of the semi-axis $[0, \infty)$.

Theorem 5. *Consider $\alpha(\cdot)$ and $\beta(\cdot)$, arrival and service curves respectively. If there exists some $T > 0$ such that*

$$\alpha(T) \leq \beta(T), \tag{18}$$

then

$$\sup_{0 \leq s \leq T} [\alpha(s) - \beta(s)] = \sup_{s \geq 0} [\alpha(s) - \beta(s)]. \tag{19}$$

Proof. Take $s > T$. We can write $s = kT + s'$, where $0 \leq s' < T$ and $k \in \mathbb{N}$. Then, by subadditivity of α and superadditivity of β, we have that $\alpha(s) - \beta(s) \leq k[\alpha(T) - \beta(T)] + \alpha(s') - \beta(s') \leq \alpha(s') - \beta(s')$. $\qquad\square$

After determining the empirical arrival curve $\alpha_m(t)$ for $t \leq t_1$ in Section 5.2, we wish to determine service curves β_1 with parameters $(z_{\max}, t_0)$, such that Eq.(18) is satisfied for some $T \leq t_1$ (i.e., we wish to guarantee a fully charged battery at some time during each day). We obtain a feasible region given by

$$D := \left\{ (z_{\max}, t_0) : 0 \leq t_0 \leq t_0^*(z_{\max}) := \sup_{0 \leq T \leq t_1} \left[T - \frac{\alpha_m(T)}{z_{\max}} \right] \right\}.$$

We plot the resulting feasible region in the first plot of Figure 4(a). We compute B^* for the points on the boundary of the feasible region using Eq.(16) and Theorem 5, and we plot it in the second plot of Figure 4(a) (we assume a 12V battery). The latter plot tells us that in the case in which there are no service interruptions (i.e., $t_0 = 0$), a maximum power of 134W and a 12V battery of 13.96Ah suffice to satisfy any α_m-smooth load.

Let us now determine for a fixed "reasonable" maximum power $z_{\max}$ what are the interruption times and required battery sizes which ensure satisfiability of the load. In Figure 4(b) we pick 4 values for $z_{\max}$, we vary the interruption time t_0, and we plot the required corresponding battery charge B^* for a 12V battery.

For example, pick $t_0 = 1\text{h}$. Then with $z_{\max} = 155\text{W}$, a 12V battery of roughly 12.55Ah suffices. However, since a high Depth of Discharge (DoD) shortens the lifetime of a battery [7], we choose a higher charge, say 40Ah. This corresponds to a standard car battery, which (for α_m-smooth loads) is guaranteed to have a charge level of roughly above two thirds of its maximum charge at all times.

In Figure 5 we show the evolution of the battery charge over time, when the greedy policy $\tilde{U}$ is used. The control G is binary (on/off) and is compatible with the service
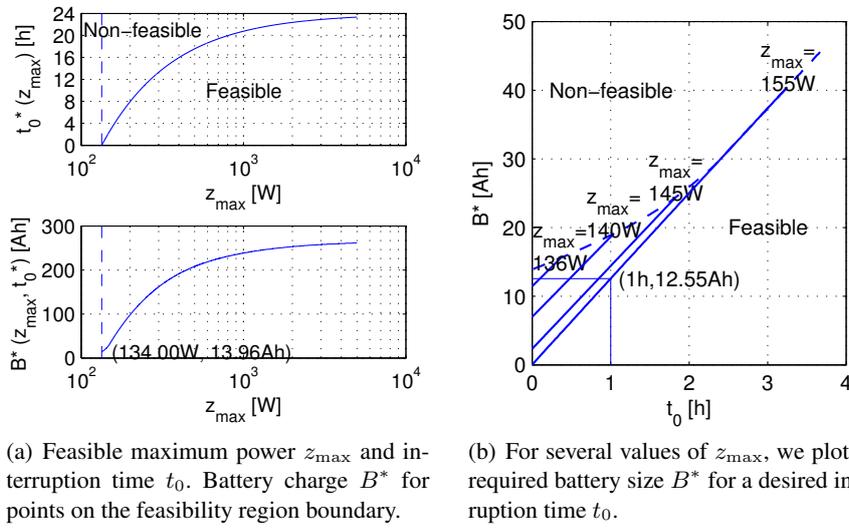
(a) Feasible maximum power z_{max} and interruption time t_0. Battery charge B^* for points on the feasibility region boundary.

(b) For several values of z_{max}, we plot the required battery size B^* for a desired interruption time t_0.

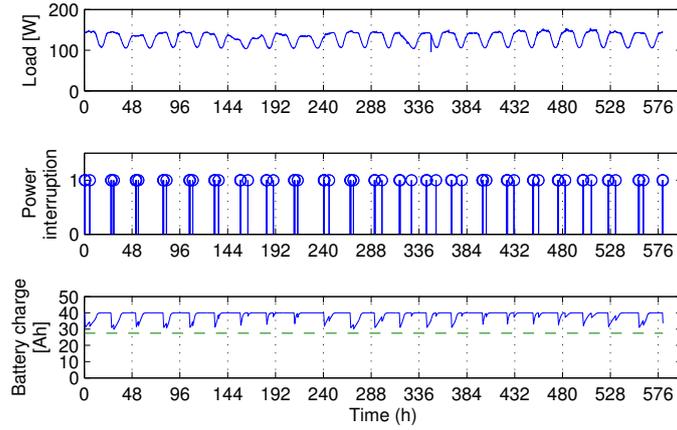**Fig. 4.** Finding battery charge B^* for measured arrival curve α_m, for various parameters of β_1.

Fig. 5. Evolution of a 12V - 40Ah battery charge, using greedy schedule $\tilde{U}$, subject to control satisfying β_1 with $z_{\mathrm{max}} = 155\mathrm{W}$ and $t_0 = 1\mathrm{h}$. The charge level remains above roughly $^2/_3$.

curve β_1 ($z_{\mathrm{max}} = 155\mathrm{W}$ and $t_0 = 1\mathrm{h}$). Time is discrete and we plot only the time steps during which an interruption is triggered. For verification, we plot the battery level $(40\mathrm{Ah} - B^*(155\mathrm{W}, 1\mathrm{h}))$. As expected, the battery charge never drops below this level. Additionally, an oversized battery provides protection against unexpected blackouts.

Note that the deployment of such a system depends only on monetary incentives. Namely, the savings generated by using a Demand-Response service compared to the normal cost of electricity need to justify equipping each server with a battery. In this section we have provided the order of magnitude of the required battery size.

6   Conclusion

We have studied the dimensioning of a perfect battery in a scenario where an inflexible load is subject to Demand-Response via quantity. We have shown that the techniques used in Network Calculus can easily be applied and provide interesting results. Namely, we have shown that when the load is characterized using a known arrival curve, and when Demand-Response is constrained by a known service curve, a large enough battery charge can be determined such that there exists an online battery charging schedule.

The perfect battery assumption allows the derivation of closed form expressions for the minimal and maximal charging schedules. In upcoming work we will focus on models for non-perfect energy storage, which do not seem to posess the same property.

References

[1] CAISO and GE Consulting. Integration of renewable resources. Operational requirements and generation fleet capability at 20% RPS. Technical report, CAISO, August 2010.

[2] I. Cho and S. Meyn. Efficiency and marginal cost pricing in dynamic competitive markets with friction. *Theoretical Economics*, 5:215–239, 2010.

[3] X. Fan, W.-D. Weber, and L. A. Barroso. Power provisioning for a warehouse-sized computer. In *Proceedings of ISCA*, 2007.

[4] J.-Y. Le Boudec and P. Thiran. *Network Calculus*. Springer Verlag Lecture Notes in Computer Science, Vol.2050, http://infoscience.epfl.ch/record/282, July 2001.

[5] J.-Y. Le Boudec and D.-C. Tomozei. Demand response using service curves. In *Proceedings of PES ISGT - Europe (To appear)*. IEEE, 2011.

[6] N. Li, L. Chen, and S. H. Low. Optimal demand response based on utility maximization in power networks. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1 –8, july 2011.

[7] D. Linden and T. Reddy. *Handbook of batteries (3rd ed.)*. McGraw-Hill, 2002.

[8] Peaksaver. Peaksaver. http://www.peaksaver.com/.

[9] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs. Cutting the electric bill for internet-scale systems. In *Proceedings of the ACM SIGCOMM 2009*, pages 123–134, New York, NY, USA, 2009. ACM.

[10] SPEC. Design document SSJ workload. http://www.spec.org/power/docs/SPECpower_ssj2008-Design_ssj.pdf, 2011.

[11] Voltalis. Bluepod. http://www.voltalis.com/.

A   Technical Lemmas

The proof of Theorem 1 is based on min-plus and max-plus system theoretic lemmas [4, Chapter 4], which are given in this section. We denote with $\mathcal{G}$ the set of functions $f : [0, \infty) \to \mathbb{R}$ that are wide-sense increasing, and with $\mathcal{F}$ the subset of all $f \in \mathcal{G}$ such that $f(0) = 0$. Note that the functions L, G and U are all in $\mathcal{F}$.

A.1 Maximal Solution

Lemma 1. *Let $f \in \mathcal{F}$ and $g \in \mathcal{F}$. Consider the problem*

$$(P10) \begin{cases} U \in \mathcal{F} \\ U(t) \leq f(t) \text{ for all } t \geq 0 \\ U(t) - U(s) \leq g(t) - g(s) \text{ for all } t \geq 0 \text{ and } s \geq 0 \text{ with } s \leq t, \end{cases}$$

where the unknown is the function U. This problem has a maximal solution $\bar{U}$, defined by

$$\bar{U}(t) = \inf_{s:\, 0 \leq s \leq t} \left(g(t) - g(s) + f(s) \right), \tag{20}$$

i.e., $\bar{U}$ is a solution and any other solution U satisfies $U(t) \leq \bar{U}(t)$ for all $t \geq 0$.

Proof. Note that the problem always has one trivial solution, $U(t) = 0$ for all $t \geq 0$.

Let Π be the operator $\mathcal{G} \to \mathcal{G}$ defined by $\Pi U(t) = \inf_{s:\, 0 \leq s \leq t} \left(g(t) - g(s) + U(s) \right)$. The fact that $\Pi U \in \mathcal{G}$ whenever $U \in \mathcal{G}$ follows from [4, Chapter 3], where this operator is called $\Pi = h_g$. Also note that $(\Pi U)(0) = U(0)$ thus Π maps $\mathcal{F}$ to $\mathcal{F}$.

Consider the problem $(P10')$ obtained by relaxing the condition $U(0) = 0$:

$$(P10') \begin{cases} U \in \mathcal{G} \\ U(t) \leq f(t) \text{ for all } t \geq 0 \\ U(t) - U(s) \leq g(t) - g(s) \text{ for all } t \geq 0 \text{ and } s \geq 0 \text{ with } s \leq t. \end{cases}$$

The problem $(P10')$ is equivalent to

$$U(t) \leq \min(f(t), (\Pi U)(t)), \text{ for all } t \geq 0.$$

The operator Π is min-plus linear [4, Chapter 4], therefore it is isotone and upper-semi continuous. Thus, by [4, Theorem 4.3.1], the problem $(P10')$ has a maximum solution $\bar{U} = \bar{\Pi} f$, where $\bar{\Pi}$ is the min-plus closure of Π. The min-plus closure of Π is the infimum of the identity and the iterates of Π.

Now note that Π is idempotent, i.e., for all $U \in \mathcal{F}$, $\Pi(\Pi U) = \Pi U$; thus the iterates of Π are equal to Π. Further, $(\Pi U)(t) \leq U(t)$. Therefore the min-plus closure of Π is $\bar{\Pi} = \Pi$. Thus the maximal solution of $(P10')$ is $\bar{U} = \Pi f$, which gives Eq.(20).

Last, note that $\bar{U} \in \mathcal{F}$ because $f \in \mathcal{F}$. Thus $\bar{U}$ is a solution to $(P10)$; further, any solution U of $(P10)$ is solution of $(P10')$, thus is upperbounded by $\bar{U}$. $\qquad \square$

A.2 Minimal Solution

Lemma 2. *Let $f, g \in \mathcal{F}$ such that $f(t) \leq g(t)$ for all $t \geq 0$.*
Consider the problem

$$(P20) \begin{cases} U \in \mathcal{F} \\ U(t) \geq f(t) \text{ for all } t \geq 0 \\ U(t) - U(s) \leq g(t) - g(s) \text{ for all } t \geq 0 \text{ and } s \geq 0 \text{ with } s \leq t, \end{cases}$$

where the unknown is the function U. This problem has a minimal solution $\underline{U}$, defined by

$$\underline{U}(t) = \sup_{s:\, s \geq t} \left(g(t) - g(s) + f(s) \right), \tag{21}$$

i.e., $\underline{U}$ is a solution and any other solution U satisfies $U(t) \geq \underline{U}(t)$ for all $t \geq 0$.

Proof. Note that the problem always has at least one solution, $U(t) = g(t)$.

Let Q be the operator $\mathcal{G} \to \mathcal{G}$ defined for $U \in \mathcal{G}$ by

$$(QU)(t) = \sup_{s \geq t} \left(g(s) - g(t) + U(t) \right).$$

Q is the max-plus equivalent of Π. More precisely, define $H(t, s)$ for $s \geq 0$ and $t \geq 0$ by

$$H(t, s) = \begin{cases} g(t) - g(s) \text{ if } t \leq s \\ 0 \text{ if } t \geq s, \end{cases} \tag{22}$$

so that

$$(QU)(t) = \sup_{s \geq 0} \left(H(t, s) + U(s) \right),$$

and further, $t \mapsto H(t, s)$ is nondecreasing and $s \mapsto H(t, s)$ is nonincreasing therefore Q is a max-plus linear operator from $\mathcal{G}$ to $\mathcal{G}$ ([4, Theorem 4.1.3]).

Now consider the problem $(P20')$ obtained by relaxing the condition $U(0) = 0$:

$$(P20') \begin{cases} U \in \mathcal{G} \\ U(t) \geq f(t) \text{ for all } t \geq 0 \\ U(t) - U(s) \leq g(t) - g(s) \text{ for all } t \geq 0 \text{ and } s \geq 0 \text{ with } s \leq t. \end{cases}$$

The problem $(P20')$ is equivalent to

$$U(t) \geq \max(f(t), (QU)(t)), \text{ for all } t \geq 0.$$

We have shown that the operator Q is max-plus linear, therefore it is isotone and lower-semi continuous. Thus, by [4, Theorem 4.3.2], the problem $(P20')$ has a minimal solution $\underline{U} = Q^* f$, where Q^* is the max-plus closure of Q. The max-plus closure of Q is the supremum of the identity and the iterates of Q.

Next, it is straightforward to check that Q is idempotent, i.e., $Q \circ Q = Q$ and that $(QU)(t) \geq U(t)$ for all $t \geq 0$ and $U \in \mathcal{G}$. It follows that $Q^* = Q$. Therefore, Problem $(P20')$ has a maximal solution, equal to $\underline{U} = Qf$, which is the same as Eq.(21).

It remains to check that $\underline{U} \in \mathcal{F}$, i.e., that $\underline{U}(0) = 0$. Note that Q does not map $\mathcal{F}$ to $\mathcal{F}$. However, we have assumed that $f(0) = g(0) = 0$ and $f(s) \geq g(s)$ for all $s \geq 0$; therefore

$$\underline{U}(0) \stackrel{\text{def}}{=} \sup_{s \geq 0} \left(-g(s) + f(s) \right) = 0.$$

Thus $\underline{U}$ is a solution of $(P20)$. As in the end of the proof of Lemma 1, this shows that $\underline{U}$ is the minimal solution of $(P20)$. $\qquad\square$