Online Control of Cost-Minimizing Multi-source Power Supply for Datacenters with Uncertain Demand

Shengkai Shi (HKU)

October 18, 2013

Background

- Major power-related challenges.

Background

- Major power-related challenges.
  - Skyrocketing power consumption.

Background

- Major power-related challenges.
  - Skyrocketing power consumption.
  - Serious environment impact.

Background

- Major power-related challenges.
  - Skyrocketing power consumption.
  - Serious environment impact.
  - Unexpected power outage.

Datacenter Power Supply System (DPSS)

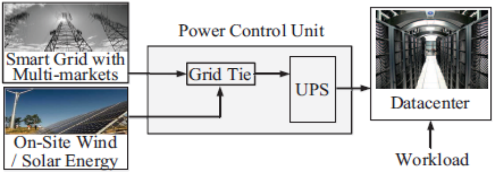


Figure: An illustration of the datacenter power supply system (DPSS).

Key Control Decisions

- Minimize long-term cost of running datacenters.

Key Control Decisions

- Minimize long-term cost of running datacenters.
  - How much demand should the DPSS serve at each time?

Key Control Decisions

- Minimize long-term cost of running datacenters.
  - How much demand should the DPSS serve at each time?
  - How much energy to be purchased from the long-term and real-time grid market respectively?

Key Control Decisions

- Minimize long-term cost of running datacenters.
  - How much demand should the DPSS serve at each time?
  - How much energy to be purchased from the long-term and real-time grid market respectively?
  - How to use the UPS to store and discharge energy?

Related Work

- Drawbacks of previous work.

Related Work

- Drawbacks of previous work.
  - Assume a priori knowledge of the power demand.

Related Work

- Drawbacks of previous work.
  - Assume a priori knowledge of the power demand.
  - Require substantial statistics of the system dynamics.

Related Work

- Drawbacks of previous work.
  - Assume a priori knowledge of the power demand.
  - Require substantial statistics of the system dynamics.
  - Limited to only a single-day optimization.

Salient Features

- Stochastic model.

Salient Features

- Stochastic model.
- Two-stage Lyapunov optimization.

Salient Features

- Stochastic model.
- Two-stage Lyapunov optimization.
- Online DPSS control algorithm.

Salient Features

- Stochastic model.
- Two-stage Lyapunov optimization.
- Online DPSS control algorithm.
- No priori knowlegde of system dynamics and demand pattern.

System Overview

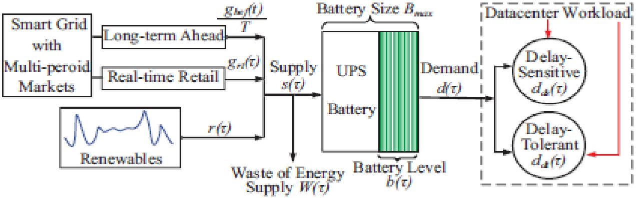


Figure: System model of multi-source energy delivery to serve demand.

Two Timescales

- K coarse-grained time slots.

Two Timescales

- K coarse-grained time slots.
  - Length T.

Two Timescales

- K coarse-grained time slots.
  - Length T.
- T fine-grained time slots.

Two Timescales

- K coarse-grained time slots.
  - Length T.
- T fine-grained time slots.
  - Length 1.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.
  - $d(t)$: demand.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.
  - $d(t)$: demand.
  - $r(t)$: renewable energy generation.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.
  - $d(t)$: demand.
  - $r(t)$: renewable energy generation.
  - $g_{bef}(t)$ and $p_{lt}(t)$: amount and price of energy purchased from long-term market.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.
  - $d(t)$: demand.
  - $r(t)$: renewable energy generation.
  - $g_{bef}(t)$ and $p_{lt}(t)$: amount and price of energy purchased from long-term market.
- At each fine-grained time slot $\tau \in [t, t + T - 1]$.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.
  - $d(t)$: demand.
  - $r(t)$: renewable energy generation.
  - $g_{bef}(t)$ and $p_{lt}(t)$: amount and price of energy purchased from long-term market.
- At each fine-grained time slot $\tau \in [t, t + T - 1]$.
  - $d(\tau)$: demand.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.
  - $d(t)$: demand.
  - $r(t)$: renewable energy generation.
  - $g_{bef}(t)$ and $p_{lt}(t)$: amount and price of energy purchased from long-term market.
- At each fine-grained time slot $\tau \in [t, t + T - 1]$.
  - $d(\tau)$: demand.
  - $r(\tau)$: renewable energy generation.

Supply Side

- At each coarse-grained time slot $t = KT(k = 1, 2, ..., K)$.
  - $d(t)$: demand.
  - $r(t)$: renewable energy generation.
  - $g_{bef}(t)$ and $p_{lt}(t)$: amount and price of energy purchased from long-term market.
- At each fine-grained time slot $\tau \in [t, t + T - 1]$.
  - $d(\tau)$: demand.
  - $r(\tau)$: renewable energy generation.
  - $g_{rt}(\tau)$ and $p_{rt}(\tau)$: amount and price of energy purchased from real-time market.

Energy Supply

Energy supply from grid and renewable energy at τ:

$$s(\tau) = g_{bef}(t)/T + g_{rt}(\tau) + r(\tau), 0 \leq s(\tau) \leq S_{max}. \quad (1)$$

Demand Side

$d(\tau) = d_{ds}(\tau) + d_{dt}(\tau)$.

- $d_{ds}(\tau)$: delay-sensitive demand.

Demand Side

$d(\tau) = d_{ds}(\tau) + d_{dt}(\tau)$.

- $d_{ds}(\tau)$: delay-sensitive demand.
- $d_{dt}(\tau)$: delay-tolerant demand.

Demand Side

$d(\tau) = d_{ds}(\tau) + d_{dt}(\tau)$.

- $d_{ds}(\tau)$: delay-sensitive demand.
- $d_{dt}(\tau)$: delay-tolerant demand.
    - λ_{max}: maximal allowed delay.

Demand Side

$d(\tau) = d_{ds}(\tau) + d_{dt}(\tau)$.

- $d_{ds}(\tau)$: delay-sensitive demand.
- $d_{dt}(\tau)$: delay-tolerant demand.
  - λ_{max}: maximal allowed delay.
  - $Q(\tau)$: total delay-tolerant demand in the queue.

Demand Side

$d(\tau) = d_{ds}(\tau) + d_{dt}(\tau)$.

- $d_{ds}(\tau)$: delay-sensitive demand.
- $d_{dt}(\tau)$: delay-tolerant demand.
  - λ_{max}: maximal allowed delay.
  - $Q(\tau)$: total delay-tolerant demand in the queue.
  - $s_{dt}(\tau) = \gamma(\tau)Q(\tau)$: served delay-tolerant demand.

Update of the Delay-tolerant Demand Queue

Update of the delay-tolerant demand queue:

$$Q(\tau+1) = max(Q(\tau) - s_{dt}(\tau), 0) + d_{dt}(\tau). \tag{2}$$

Dynamics of UPS Battery Level

- $b_{rc}(\tau) = s(\tau) - d_{ds}(\tau) - s_{dt}(\tau)$: energy recharged to the battery.

Dynamics of UPS Battery Level

- $b_{rc}(\tau) = s(\tau) - d_{ds}(\tau) - s_{dt}(\tau)$: energy recharged to the battery.
- $b_{dc}(\tau) = d_{ds}(\tau) + s_{dt}(\tau) - s(\tau)$: energy discharged to the supplement the supply.

Update of UPS Battery Level

Update of UPS battery level:

$$b(\tau+1) = min(b(\tau) + b_{rc}(\tau)\eta_c - b_{dc}(\tau)\eta_d, B_{max}), \qquad (3)$$

where B_{max} is the maximum battery capacity, η_c is recharging efficiency and η_d is the discharging efficiency.

Matching Demand and Supply

At each fine-grained time slot τ,

$$s(\tau) + b_{dc}(\tau) - b_{rc}(\tau) = d_{ds}(\tau) + \gamma(\tau) Q(\tau). \qquad (4)$$

Maximal Energy Supply from Grid

The maximal amount energy from grid at τ is limited by P_{grid}.

$$0 \leq g_{bef}(t)/T + g_{rt}(\tau) \leq P_{grid}. \tag{5}$$

Guaranteeing Dealy-tolerant Demand Deadline

To guarantee the maximal deadline λ_{max}, any control policy should satisfy

$$Q(\tau) < Q_{max}, \tag{6}$$

where Q_{max} is the maximal backlog.

Ensuring Datacenter Availability

To avoid discretionary UPS discharging, UPS battery has a minimum energy level B_{min}.

$$B_{min} \leq b(\tau) \leq B_{max}. \tag{7}$$

UPS battery also has constraints on the maximal recharge and discharge energy.

$$0 \leq b_{rc}(\tau) \leq B^{c}_{max}, 0 \leq b_{dc}(\tau) \leq B^{d}_{max}. \tag{8}$$

UPS Battery Lifetime

The maximum discharging and charging number N_{max} during the time horizon t satisfies

$$0 \leq \sum_{\tau=0}^{t-1} n(\tau) \leq N_{max}, \tag{9}$$

where $n(\tau) = 1$ if $b_{rc}(\tau) > 0$ or $b_{dc}(\tau) > 0$, 0 otherwise.

UPS Battery Operation Cost

If a UPS costs C_{buy} to purchase and it can sustain C_{cycle} charge/discharge cycles, then the cost of battery charging and discharging per time is $C_b = C_{buy} / C_{cycle}$. At time τ, the UPS operation cost is $n(\tau)C_b$

Stochastic Constrained Cost Minimization Problem

At each fine-grained time slot τ, the operation cost is

$$Cost(\tau) = g_{bef}(t)/Tp_{lt}(t) + g_{rt}(\tau)p_{rt}(\tau) + n(\tau)C_b. \tag{10}$$

The objective is to solve the following stochastic cost minimization problem $P1$

$$\begin{array}{ll} \min\limits_{g_{bef}, g_{rt}, \gamma} & Cost_{av} \triangleq \liminf\limits_{t \longrightarrow \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}[Cost(\tau)] \\ \text{s.t.} & \forall \tau : \text{constraints } (4)(5)(6)(7)(8)(9). \end{array}$$

An Optimal Offline Algorithm

Lemma 1: In every optimal solution of the optimization problem **P1**, it holds that $\forall\tau$, $g_{rt}(\tau) \equiv 0$ or $p_{rt}(\tau) \equiv 0$ *.

Workload Delay-Aware Virtual Queue

In order to guarantee the maximum delay λ_{max}, a delay-aware virtual queue $Y(t)$ is defined.

$$Y(t+1) = max[Y(t) - s_{dt}(t) + \epsilon 1_{Q(t)>0}, 0], \tag{11}$$

where $\epsilon 1_{Q(t)>0}$ is an indicator variable that is 1 if $Q(t) > 0$ and 0 otherwise.

Datacenter Availability-Aware Virtual Queue

In order to constraint (7), $X(t)$ is defined to track the battery level.

$$X(t) = b(t) - U_{max} - B_{min} - B^{d}_{max}\eta_d. \tag{12}$$

The dynamics of $X(t)$ is given as

$$X(t+1) = min[B_{max}, X(t) + b_{rc}(t)\eta c - b_{dc}(t)\eta_d]. \tag{13}$$

Two-Timescale Lyapunov Optimization

$\Theta(t) = [Q(t), X(t), Y(t)]$ is a concatenated vector of the actual and virtual queues. The quadratic Lyapunov function is defined as

$$L(\Theta(t)) \triangleq \frac{1}{2}[Q^2(t) + X^2(t) + Y^2(t)].$$

The T-slot conditional Lyapunov drift is defined as

$$\Delta_T\,(\Theta(t)) \triangleq \mathbb{E}[L(\Theta(t+T)) - L(\Theta(t))|\Theta(t)].$$

The drift-plus-penalty term every T slots is defined as

$$\Delta_T\,(\Theta(t)) + V\mathbb{E}\{\sum_{\tau=t}^{t+T-1} Cost(\tau)|\Theta(t)\},$$

where V is a control parameter.

Drift-plus-Penalty-Bound

Theorem 1: (Drift-plus-Penalty Bound) Let $V > 0$, $\epsilon > 0$, $T \geq 1$ and $t = kT, \tau \in [t, t+T-1]$. To ensure two-timescale power purchasing $0 \leq g_{bef}(t)/T + g_{rt}(\tau) \leq P_{grid}$, demand management decision $\gamma(\tau) \in [0,1]$, battery level $b(\tau) \in [B_{min}, B_{max}]$ and demand backlog $Q(t) < Q_{max}$, under any operation actions, the drift-plus-penalty satisfies:

$$
\begin{aligned}
& \Delta_T\left(\Theta(t)\right) + V\mathbb{E}\{\sum_{\tau=t}^{t+T-1} Cost(\tau)|\Theta(t)\} \quad (19) \\
\leq \; & H_1 T + V\mathbb{E}\{\sum_{\tau=t}^{t+T-1} Cost(\tau)|\Theta(t)\} \\
- \; & \mathbb{E}\{\sum_{\tau=t}^{t+T-1} Q(\tau)[s_{dt}(\tau) + d_{dt}(\tau)]|\Theta(t)\} \\
+ \; & \mathbb{E}\{\sum_{\tau=t}^{t+T-1} X(\tau)[b_{rc}(\tau) - b_{dc}(\tau)]|\Theta(t)\} \\
+ \; & \mathbb{E}\{\sum_{\tau=t}^{t+T-1} Y(\tau)[\epsilon - s_{dt}(\tau)]|\Theta(t)]\}
\end{aligned}
$$

where $H_1 = {S_{max}^{dt}}^2 + \frac{1}{2}[{D_{max}^{dt}}^2 + {B_{max}^{c}}^2\eta_c^2 + {B_{max}^{d}}^2\eta_d^2 + \epsilon^2]$.

Relaxed Optimization Problem

Approximate near future statistics as the current statistics: $Q(\tau) = Q(t)$, $X(\tau) = X(t)$ and $Y(\tau) = Y(t)$ for $t < \tau \leq t + T - 1$.

Corollary 1: (Loosening Drift-plus-Penalty Bound) Let $V > 0$, $\epsilon > 0$ and $T \geq 1$. Considering Theorem 1 under approximation, the drift-plus-penalty term satisfies:

$$
\begin{aligned}
& \Delta_T\,(\Theta(t)) + V\mathbb{E}\{\sum_{\tau=t}^{t+T-1} Cost(\tau)|\Theta(t)\} \quad (20) \\
\leq\ & H_2 T + V\mathbb{E}\{\sum_{\tau=t}^{t+T-1} Cost(\tau)|\Theta(t)\} \\
-\ & \mathbb{E}\{\sum_{\tau=t}^{t+T-1} Q(t)[s_{dt}(\tau) + d_{dt}(\tau)]|\Theta(t)\} \\
+\ & \mathbb{E}\{\sum_{\tau=t}^{t+T-1} X(t)[b_{rc}(\tau) - b_{dc}(\tau)]|\Theta(t)\} \\
+\ & \mathbb{E}\{\sum_{\tau=t}^{t+T-1} Y(t)[\epsilon - s_{dt}(\tau)]|\Theta(t)\},
\end{aligned}
$$

where $H_2 = H_1 + T(T-1){B^c_{max}}^2\eta_c^2 + T(T-1)\epsilon^2$.

Online Control Algorithm

Algorithm 1: SmartDPSS Online Control Algorithm.

1) *Long-term Ahead Planning*: At each time $t = kT (k \in \mathbb{Z}_+)$, observing system states $Q(t), Y(t)$, renewable energy $r(t)$, power demand $d_{ds}(t)$, maximum available battery level $b(t)$ and energy prices $p_{lt}(t)$, DPSS decides the optimal power procurement in the long-term market $g_{bef}(t)$ to minimize the following problem **P4**:

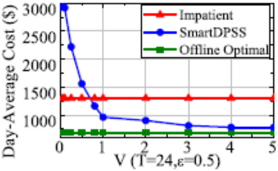
$$\begin{aligned} \min_{g_{bef}} \quad & g_{bef}(t)\Big[V p_{lt}(t) - Q(t) - Y(t)\Big] \\ \text{s.t.} \quad & g_{bef}(t)/T + r(t) + b(t) \geq d_{ds}(t), \\ & 0 \leq g_{bef}(t)/T \leq P_{grid}, B_{min} \leq b(t) \leq B_{max}. \end{aligned}$$

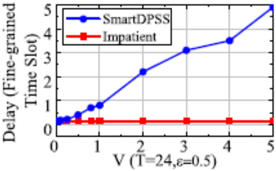
2) *Real-time Balancing*: At each fine-grained time slot $\tau \in [t, t+T-1]$, with system statistics $Q(t), X(t), Y(t)$, renewable production $r(\tau)$, power demand $d(\tau)$ and energy prices $p_{rt}(\tau)$, DPSS performs real-time procurement $g_{rt}(\tau)$ and delay-tolerant demand management decision $\gamma(\tau)$ to minimize the following optimization problem **P5**:

$$\begin{aligned} \min_{g_{rt},\gamma} \quad & \sum_{\tau=t}^{t+T-1} \Big\{ g_{rt}(\tau)\Big[V p_{rt}(\tau) - Q(t) - Y(t)\Big] \\ & + \gamma(\tau)\Big[Q(t)^2 - Q(t)Y(t)\Big] + V n(\tau) C_b + V W(\tau) \\ & + [Q(t) + X(t) + Y(t)][b_{rc}(\tau) - b_{dc}(\tau)] \Big\} \\ \text{s.t.} \quad & (4)(5)(6)(7)(8)(9). \end{aligned}$$

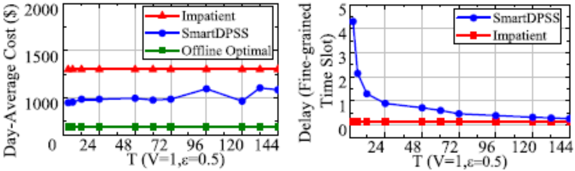
3) *Queue Update*: Update the actual and virtual queues using Eq. (2) (3) (12) (15).

Impact of Control Parameter V

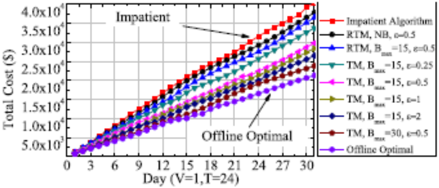




Impact of Long-term Time Slot T



Impact of Battery Size and Two-timescale Markets



Possible Application

- Cloud Service

Possible Application

- Cloud Service
  - How to minimize long-term cost by choosing among different pricing options.

Thanks!