Characterizing Power Usage in Zero Reserved Power Data Centers to Enable Planned Maintenance

Hassan Ali Khan, Muhammad Shahzad
North Carolina State University, USA

***Abstract*—The prevalent convention in data centers (DCs) has been to provide reserved power to handle any failures of power device. To reduce the environmental impacts and to increase resource utilization, recently, DC providers have started shifting towards zero reserved power (ZRP), where a DC can safely consume power up to 100% of the capacity of the power sources. As DCs have only recently started shifting to ZRP, little is known about the trends in the power usage of ZRP DCs. In this paper, we fill this gap and present, for the first time, a thorough characterization of the power usage trends in operational ZRP DCs of a large data center provider (LDCP). We analyze various spatial, temporal, organizational, and statistical factors that impact the power usage in LDCP's ZRP DCs. Based on our observations, we further develop an approach that accurately predicts future time windows during which the power usage can be expected to stay below a given threshold. Such prediction is required to enable planned maintenance of power sources in ZRP DCs because as ZRP DCs do not have reserved power, when one power source is disconnected for maintenance, the power usage of the running workloads should not exceed the capacity of the remaining power sources. Otherwise, the ZRP DC could go into a blackout. Our evaluations on data from ZRP DCs show that the proposed prediction approach successfully identifies 91% of all the 1-hour time-windows that are safe for maintenance up to 14 days in the future.**

I. Introduction

Conventionally, data center (DC) providers limit the workloads in their DCs such that there is significant amount of reserve power to prevent any unavailability events due to power source failures. Unfortunately, by operating at less than 100% of the available power, the DC providers essentially underutilize their compute resources, resulting in lost revenue. To overcome this problem, recently, cloud providers have started converting their DCs to zero reserved power (ZRP), where a DC can safely consume power up to 100% of the capacity of the power sources. This was enabled by observations that most workloads could tolerate some degradation of computational resources. An automated system in each ZRP DC monitors the power usage to maintain a safe power consumption level, and if needed, reduces the clock speeds of selected computing resources, or shuts down some servers after moving their workloads to underutilized servers. Interested readers can read more about ZRP DCs in [1].

As ZRP DCs are quite new, little is known about the trends in the power consumption of such DCs. In this paper, we fill this gap and present, for the first time, a detailed measurement study on the power consumption trends in operational ZRP DCs of a large data center provider (LDCP). To conduct this measurement study, we collected per-second power utilization values (PUVs) for 10 weeks from a large number of power sources in several ZRP DCs of LDCP spread across multiple regions throughout the world. The resulting data set is comprised of close to a billion PUVs. We present an extensive analysis of this data set from four perspectives. First, we study the spatial aspect of power utilization in ZRP DCs, where we discuss how the power utilization varies in DCs across different geographical regions as well as within different rooms of any given DC. Second, we study the temporal aspect of power utilization in ZRP DCs, where we study the diurnal patterns in the PUV time-series of the power sources, identify various types of PUV time-series, and study how PUVs differ between weekdays and weekends. Third, we study the power consumption patterns of LDCP's various internal and public organizations. Last, we present statistical analysis of the PUV time-series of all power sources, where we study what distributions these time-series follow and what trend and seasonality components exist in them.

In our measurement study, we made several notable observations with significant implications for data center providers. We briefly mention three of those here and discuss them in more detail later in the paper. First, geographically distributing the workloads of different organizations across DC rooms can help achieve consistent power usage trends. Second, the PUV time-series of most power sources can be modeled with just three probability distributions. Third, the PUV time-series exhibit both trend and seasonality components, which span seven days and one day, respectively.

The observations that we present can enable several applications. We briefly describe four such applications next.

1) **Safe Maintenance Window Prediction:** This application would identify time windows in the future where the expected power usage of some of the power sources located in the same physical location (room) is below a certain threshold. This would enable the DC providers to avoid power blackouts during scheduled power source maintenance operations.
2) **Workload Relocation:** This application would perform power-aware workload relocation by using the predicted PUVs of various power sources along with patterns in PUV time-series to move workloads within DCs or across DCs to balance the load on various power sources.
3) **Renewable Energy Integration:** With increasing demand for sustainable energy, DC providers are exploring the use of renewable energy sources. However, the integration of renewable sources is challenging due to their intermit-

ISBN 978-3-903176-63-8 © 2024 IFIP

978-3-903176-63-8/24/$31.00 ©2024 IEEE

tent/unpredictable nature. DC providers can use the ability to predict PUVs to pre-plan how much power to import from which type of power source (e.g., solar, wind, grid etc.) at different times of day.

4) **Dynamic Pricing:** This application would perform power-aware dynamic pricing to adjust prices for clients based on, among other routine factors, expected power usage to maximize the revenue for public cloud providers and minimize the cost for cloud users.

To demonstrate that the observations from this study can indeed enable the development of such applications, as an example, we develop the Safe Maintenance Window Prediction application, described next.

The power sources in DCs need routine maintenance, such as software/hardware patching, replacing corrupt batteries, infrastructure upgrades, and so on. Such maintenance operations often need temporary shut down of power source(s) during which, the servers powered by those power sources need to be shifted to other collocated power sources. This was not a problem in conventional data centers due to the availability of reserve power. However, in ZRP DCs, there is a possibility that the servers that have temporarily been shifted to a collocated power source and the servers which that power source was already powering could collectively, due to new workloads, require more power than the capacity of the power source. This would result in a partial or complete blackout. To avoid this, ZRP DCs need to accurately predict time windows of appropriate duration where the power required by servers is expected to stay below a threshold.

Just like the lack of measurement studies on ZRP DCs, to the best of our knowledge, there is no prior work on identifying safe maintenance windows in ZRP DCs. To predict safe maintenance windows, we use insights from the observations we present from our measurement study to identify observable factors that correlate with PUVs. We use these factors to create a model that uses information from past several days to predict the future PUVs. From those predicted PUVs, we identify time windows of desired duration during which all predicted PUVs stay below the required threshold. Our evaluations on real data from operational ZRP DCs show that our approach successfully identifies 91% of all 1-hour safe maintenance time-windows up to 14 days in advance.

Key Contributions. In this paper, we make three key contributions. 1) We present the first in-depth visibility into the power usage behaviors in ZRP DCs and study their spatial, temporal, organizational, and statistical aspects. 2) We develop a method to identify future time windows when power source maintenance can be performed safely. 3) We extensively evaluate our approach using data from operational ZRP DCs.

II. Data Set

In this section, we first describe the components of ZRP DC architecture that are relevant to this work. After that, we describe the data that we collect.

A. Zero-Reserved Power DCs

A ZRP DC is comprised of multiple rooms. Fig. 1 shows the layout of a typical DC room. Each DC room contains multiple power sources and multiple tiles, where a tile is a collection of multiple server racks. Each tile is powered by one of the four power sources. Each power source is connected with five telemetry devices, which capture PUVs from the power source every second and send them to a remote monitor. The remote monitor first verifies the correctness of these PUVs and then stores them in a persistent storage.

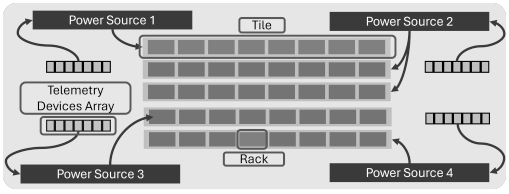

Fig. 1: A typical ZRP DC room.

DC providers contains multiple internal organizations that provide different services, such as e-mail, public cloud, DevOps etc. The DC providers assign all tiles powered by a given power source to exactly one internal organization to run its workloads. This enables a DC provider to measure the power consumed by that organization. Internal organizations with larger computational requirements often have multiple power sources powering their workloads, either within a DC room, or across multiple DC rooms in the same ZRP DC, or even across rooms in different ZRP DCs.

B. Data Collection and Aggregation

LDCP granted us access to a subset of its power sources located across multiple rooms in 13 of its several ZRP DCs. For confidentiality, we do not list the exact number of power sources, rather represent it with M, where $M > 100$, and thus large enough to provide statistically significant observations. Our data set comprised of around 1 billion PUVs collected at one-second frequency. Given the size of the data set, we first examined whether we could aggregate the PUVs to a more coarse time resolution to reduce the size of the data set (to make it easy to process) while still preserving the power consumption patterns. To determine the appropriate time resolution to aggregate to, we performed the following analysis on the PUV time-series of each power source. Let the i^{th} PUV time-series, S^i, contains n values, where $1 \leq i \leq M$. We divide S^i into temporally ordered subsets such that each subset contains t seconds worth of PUVs. This results in $\lceil n/t \rceil$ subsets. Next, for each subset k, where $1 \leq k \leq \lceil n/t \rceil$, we calculate the standard deviation ($\sigma^i_{k,t}$), mean ($\mu^i_{k,t}$), and coefficient of variation ($cv^i_{k,t}{=}\sigma^i_{k,t}/\mu^i_{k,t}$) of the PUVs contained in that subset. We repeat the steps above on S^i for all values of t from 2 seconds to 1 hour and pick that largest value of t as the aggregation duration for which $cv^i_{k,t} < 3\%$ for all k subsets. We set the threshold to 3%. Thus, a $\pm$3% deviation in our readings poses no risk of blackout. We represent this aggregation duration for S^i with T^i. We repeat the steps above to obtain a T^i for each of the M S^i and pick the smallest of these M T^i values as the final aggregation interval to

be used. This aggregation interval ensures that the standard deviation observed in all values measured within this interval will be less than 3% of the mean, *i.e.*, all PUVs are fairly similar and consistent within this interval. From applying the method above, the aggregation interval turned out to be slightly larger than 1 minute. We eventually chose 1 minute as the aggregation interval as that provided a nice time boundary. To aggregate, we took average of every 60 consecutive values in each PUV time series and stored these aggregated PUVs for further processing.

III. Characterizing the Power Usage in ZRP DCs

We start our discussion with the observations about power usage in LDCP's 13 selected ZRP DCs. After that, we study the temporal and diurnal patterns within individual ZRP DCs. Following that, we study the power usage by workloads of different organizations within LDCP. Last, we present statistical analyses of the power usage patterns in ZRP DCs.

A. Spatial Observations

We do not present the absolute PUVs of various DCs and the rooms in DCs due to the sensitive nature of that information. Instead, we present each PUV as a fractional multiple of the largest PUV, $P_{\max}$, observed across all ZRP DCs. We emphasize that the presentation of PUVs as a fractional multiple of $P_{\max}$ does not have any impact on any trends and observations that we present in this paper.

Fig. 2 shows the average PUVs of the 13 selected ZRP DCs as well as of the individual rooms in these DCs. We observe from this figure that the ZRP DCs in Regions 4 through 7 consume the most power. As providers carefully select the locations of their DCs to be geographically close to their users, we can infer from Fig. 2 that the majority of the demand for DC resources originates from and around Regions 4 through 7. The ZRP DCs in Region 2 showed the lowest average PUVs. This is because Region 2 DCs have only 1 room in per DC.

The PUVs of individual rooms provide the opposite observation. The PUVs of DCs, normalized with their respective number of rooms, are the highest in Region 2 and lowest in Region 4 through 7. Across the selected ZRP DCs, the average PUVs of individual rooms lied between $0.003P_{\max}$ and $0.31P_{\max}$. The rooms with very low PUVs were new rooms, and had very few servers setup in those at the time of data collection. Across rooms of data centers, the average PUVs of individual rooms is dissimilar because within LDCP, various organizations use the resources of the ZRP DCs. LDCP allocates various tiles in different rooms to different organizations and usually does not switch them around. As different organizations offer different workloads, they draw different amounts of power, resulting in diverse average PUVs across rooms.

The coefficient of variation (cv) in PUVs was quite low for most DC rooms, except for room 5 of DC 9, room 5 of DC 7, and room 4 of DC 7, where the cv was 161%, 81%, and 40%, respectively. For the remaining rooms across all ZRP DCs, cv lied in the range of 0.6% to 39% with an average

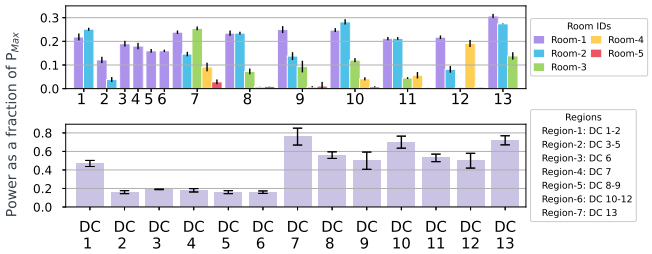

Fig. 2: Average power usage of ZRP DCs and their rooms.

of 11%. This small value of the average cv indicates that the power usage stays fairly stable and does not experience drastic spikes. When we present the temporal analysis, we will show that the PUVs indeed demonstrate stable diurnal patterns.

B. Temporal Observations

We observed that the PUV time-series of all power sources exhibited diurnal patterns, which were more pronounced on weekdays than weekends. Fig. 3 shows three representative PUV time-series. Most other PUV time-series were fairly similar to one of these three. We name them normal, power-law, and gamma time-series based on the probability distributions they follow. We will discuss the distributions in more detail later in Sec. III-D. The common aspect of all three types of time-series was that the power draw started peaking around 7 AM and dropped significantly after 3 PM during the weekdays. On weekends, the average PUVs were 4 to 11% lower than the average PUVs on weekdays for every distribution type across all power sources. We further observed that for more than 66% PUV time-series, the day of the week on which we observed the maximum average PUV was Monday, Tuesday, and Wednesday for Gamma, Power-law, and Normal time-series, respectively.

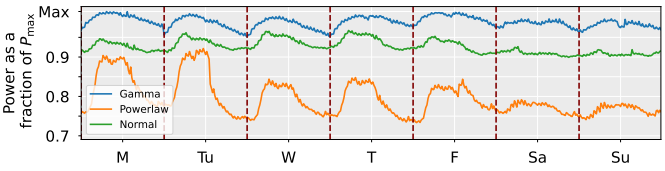

Fig. 3: PUV time-series of three power sources.

The three types of time-series differed in a few important ways. For example, the power-law time-series showed two peaks, whereas the others showed a single peak. The power-law time-series showed the highest standard deviation in PUVs while the gamma time-series showed the smallest. The PUV patterns and intensities are rather similar across the five days of weekdays and across the two days of weekends for normal and gamma time-series. However, for power-law time-series, the PUVs are 5% higher on first two weekdays as compared to the remaining weekdays and 13% higher compared to the weekends. Finally, except for the normal time-series, on weekends, all other time-series had similar temporal patterns as weekdays (but with smaller peaks), as seen in Fig. 3.

We further noted that aggregate PUVs of entire data centers or even entire rooms do not necessarily represent the PUV behavior of individual power sources. Different power sources

in the same DC room often did not produce the same type of time-series. To formally quantify this observation, we compute the Pearson temporal correlation coefficient [2], which we represent with r (where $|r| \geq 0.7$ indicates strong correlation), between all pairwise combinations of the PUV time series of the M power sources. Our analysis revealed that the average value of r from the pairwise combinations of all M power sources was just 0.28, from the pairwise combinations of the M power sources within each data center was 0.31, and from the pairwise combinations of all power sources within each room was 0.33. This shows that the correlation is weak not just across power sources around the world but even across the power sources within the DC rooms. This is because the tiles that different power sources are providing power to run different workloads (*e.g.*, game streaming, e-commerce, storage) from different organizations, and thus exhibit different power usage patterns. This shows that whenever a DC provider needs insights into power consumption behaviors, they must always study the PUV time-series of individual power sources. This insight will become crucial in our design of the approach to predict safe maintenance windows.

Now, we study whether the diurnal patterns in PUV time-series repeat across different days of the week. For this, we take the 70 daily PUV time-series of each power source. Each time-series is comprised of 1440 values, one per minute. For any given power source, we compute r for each pairwise combination of these 70 time-series of that power source. Fig. 4 shows the average value of r obtained between the daily PUV time-series for all pairs of days. We see that the correlation between all pairs of weekdays is moderate and is stronger between the same days across weeks. This shows that while the workloads that the DC organizations offer throughout the day are somewhat similar across weekdays, the amounts of workloads are highly similar on the same day across weeks. We further observe that the correlation between a weekday and a weekend day is relatively low. This is expected because, as we saw in Fig. 3, on weekends the data center utilization is relatively lower.

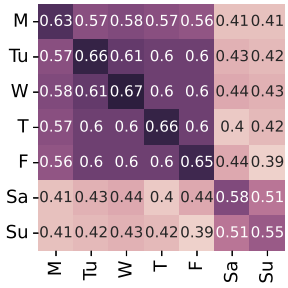

Fig. 4: PUV correl. across days

C. Organizational Observations

LDCP has multiple internal organizations, represented with G, that operate their workloads in its ZRP DC rooms. We sort the organizations based on the amount of power they utilize and number them from 1 to G. Table I lists the five highest and lowest power-consuming organizations. Table I also shows the number of rooms across the selected ZRP DCs in which the workloads of any given organization are running. Fig. 5 shows the box plots for the organizations listed in Table I. A box plot for any organization is made using 70 data points, one per day, normalized using the largest daily value observed across all of the G organizations.

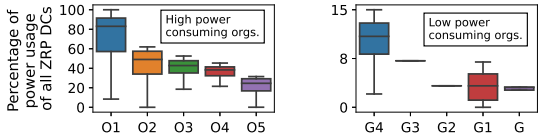

Fig. 5: Box plots of the %age of power used by select organizations.

Org #	Organization Name	Rooms	PUV%
1	Cloud services	33	82%
2	Virtualization services	31	50%
3	DC Networking services	21	46%
4	Business e-mail services	4	44%
5	Storage services	17	28%
$G-4$	Virtual networking	2	11%
$G-3$	DevOps	2	8%
$G-2$	Platform services	1	3.5%
$G-1$	Database services	2	3.5%
G	High performance computing (HPC)	1	3.1%

TABLE I: Five highest and lowest power using organizations. PUV% is the median value normalized using the largest daily value observed across all of the G organizations.

We observe from Table I and Fig. 5 that approximately 80% of the total power is used by merely seven organizations, 1 through 5, online file hosting (6), and GPU computing (7). Organization 1 stands out as the most power-consuming organization, accounting for an average of 26% of total power usage. The organizations with fewer customers consume less power. For example, organization numbered G serves customers primarily engaged in research simulations and thus consumes just 0.1% of total power.

We observe from Table I that organizations with higher power usage have their workloads running in more DC rooms. A peculiar case is that of organization 4, which consume 13% of the total power but operates out of just four rooms (albeit uses all tiles within those rooms). This is because business emails often contain sensitive information, necessitating strict security measures. At the same time, they are not latency sensitive and thus need not be serviced very close to the customer. These two factors resulted in email services organization getting placed in four dedicated rooms across the selected ZRP DCs. Except organization 4, we observed that any organization that consumes $\geq 6\%$ of total power had workloads running in at least 10 rooms across the selected DCs, while organizations consuming $\leq 1\%$ had workloads running in no more than 4 rooms.

Another interesting observation from our data is that the cv averaged across the all organizations is 21.76%, where only six organizations exhibited cv below 10%. In contrast, recall from Sec. III-A that the cv averaged across all rooms of our selected ZRP DCs was just 11%. This difference suggests that LDCP understands the power usage patterns of various organizations and orchestrates the allocation of tiles to different organizations in such a way that the collective power usage of workloads across these tiles experiences lower fluctuations compared to workloads of individual organizations.

Next, we discuss temporal trends in the power usage of these organizations. We observed that the power usage by any organization was higher on weekdays than on weekends. The organization $G-4$ exhibited the maximum 11% increase in daily power usage on weekdays compared to weekends,

while organization G showed the lowest increase of 3.2%. Except organization 1, all organizations used most power on Tuesdays and the least on Sundays. This observation regarding organization 1 could be influenced by online e-commerce markets, which bring peak traffic on weekends when people have more time for shopping. Except for organizations $G-2$ and $G-3$ the remaining organizations exhibited diurnal power usage patterns, with peaks occurring on weekdays between 6 AM and 9 AM, and again between 12 PM and 3 PM. These peaks were relatively modest though, with peak usage only 2 to 5% higher than the average for that day. Weekends, however, showed no discernible peak hours for most organizations.

D. Statistical Observations

Now we study two statistical aspects of PUV time-series: distributions that PUV time-series follow and temporal autocorrelations within PUVs. Our observations will be useful in designing the approach to predict safe maintenance windows.

1) PUV Distributions

The knowledge of the distribution of a PUV time-series enables one to learn about the frequency of various PUVs in any given time-series, which is instrumental in predicting safe maintenance windows. To determine the distribution that the PUV time-series of any given power source follows, we calculate the sum of squared errors (SSE) of that PUV time-series against seven common probabilistic distributions using the Fitter library [3]. The seven distributions we selected include normal, power-law, gamma, exponential-power, uniform, rayleigh, and exponential. Our motivation behind choosing these distributions is that they are frequently observed in human and computing environments.

The fitting results revealed that the most common distributions that the PUV time-series followed include normal, power-law, and gamma distributions, which described 84% of all time-series. Each of the three distributions represented almost the same percentage of time-series, *i.e.*, 27 to 29%. While some other distributions better described the remaining time-series, the SSE values for gamma, power-law, or normal distributions were not too large compared to the best-fit distribution. Similar to what we saw in Sec. III-B, we noted that the PUV time-series of different power sources within the same DC room often did not fit the same distribution.

2) PUV Autocorrelation

Next we analyze the autocorrealtion within PUV time-series. This will enable us to identify patterns, such as seasonality and trends, in PUV time-series, which, in turn, will enable us to predict future PUVs with better accuracy. The seasonality component refers to the regular and periodic fluctuations that occur at fixed intervals in a time-series, while the trend component represents the overall direction in which the time-series is moving, such as upward, downward, flat, etc.

Before conducting the autocorrelation analysis, we performed the augmented Dickey-Fuller test [4] on the PUV time-series of each of the K power sources to ensure that the PUV time-series are not stationary, which turned out to be the case. This non-stationarity is a pre-condition to perform any autocorrelation based analysis. Next, we apply the STL algorithm [5] on each PUV time-series to decompose the time-series into its seasonal and trend components. Fig. 6 plots the PUV time-series of a power source in a room of a selected ZRP DC for the entire 70 days. This figure also shows the decomposed trend and seasonal components. We observe from this figure that the power consumption trend exhibits a repeating cycle of 7 days while the duration of the seasonal component is 1 day. This demonstrates that the time of day and the day of the week are essential features for predicting safe maintenance windows. We made similar observations from the PUV-times series of all other power sources.

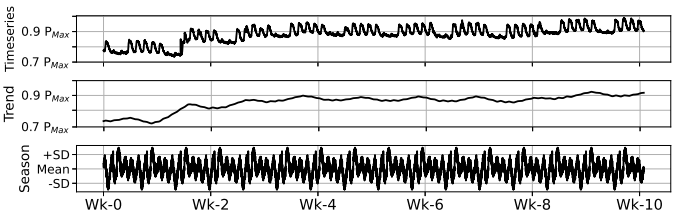

Fig. 6: The trend and seasonal components of PUV time-series of a selected power source.

The presence of a strong seasonality component suggests that the previous PUVs of a power source significantly influence its current and future PUVs. To determine the lag, *i.e.*, the number of previous PUVs that influence the current PUV, we calculate the autocorrelation function (ACF) [6] and partial ACF (PACF) [7] of the PUV time-series of each power source. While ACF considers all past observations, PACF removes the effects of intermediate observations. Fig. 7 shows the ACF and PACF of the PUV time-series in Fig. 6. The x-axis represents the number of observations, where each observation is the average PUV over a 5-minute window. The y-axis represents the confidence interval of the correlation. Notice that, unlike ACF, we do not observe a cyclic pattern in PACF because PACF removes the intermediate relationships between observations. Our analysis reveals that each PUV is influenced by the preceding 4 to 12 observations (*i.e.*, 20 to 60 minutes as each observation is aggregated from 5 minutes of PUVs). This dependency is visible in the confidence values shown in the PACF plot. We observe a significant decrease in autocorrelation with the confidence score dropping below 5% after 11 previous observations (which correspond to 55 minutes of PUVs). Thus, we can confidently assert that the PUV recorded in any given minute of the time series in Fig. 6 is heavily influenced by the PUVs observed in the preceding 55 minutes. This value lied in the range of 20 to 60 minutes across all PUV time-series in our data set. The take away is that the PUVs of the last 1 hour will serve as important features to predict safe maintenance windows.

IV. Safe Maintenance Window Prediction

A safe maintenance window is a time window during the entirety of which, the power usage of a DC room stays below $(n-1)/n$ times the total power that all n power sources can provide. Next, we first describe the features that we use to

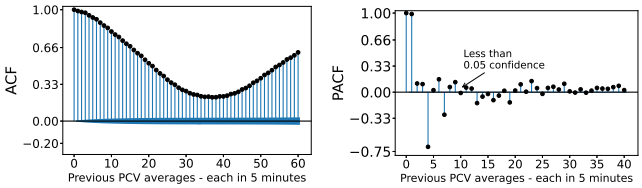

Fig. 7: ACF and PACF of PUV time-series in Fig. 6

predict the safe maintenance windows. We have derived these features based on the observations in Sec. III. After that, we analyze the effectiveness of these features and describe the machine learning (ML) approach that we used. Finally, we present results from evaluation of our prediction approach.

A. Feature Selection

As a safe maintenance window needs to be identified for any given DC room, the features we present here focus on n power sources of any given room. An ML prediction model is generated using feature vectors. For any power source in a given DC room, the number of feature vectors we have equals the number of PUVs we collected from that source. A feature vector is comprised of multiple elements, described next.

PUV: The last element of the feature vector is a PUV of the power source identified in the first element (described next). This is the element that the ML model will need to predict when looking for safe maintenance windows. Note that we know the value of this element for all feature vectors, which we use during training and as ground truth during testing.

1. Power Source ID: This is an identifier of the power source in the given DC room. When a DC room has n power sources, we identify each power source with a value from 1 to n.

2–4. Timestamp: These three elements represent the time when the PUV listed in the last element was measured (or needs to be predicted). These include the time (a value between 1 to 1440 minutes), day (1∼7), and week (1∼52).

5–12. Biweekly Statistics: These eight elements represent various statistics of the power source identified in element#1, calculated using all PUVs collected during the last 14 days. All times and durations that we mention in this section are relative to the timestamp in elements # 2–4. We chose 14 days for two reasons. First, as noted in Fig. 6, the trend component had a duration of one week. Thus, we should use at least one week of data. Second, the ZRP DCs usually need to know the safe maintenance window within the next 14 days, and our analysis of the data set revealed that on average, the past 14 days of PUVs showed the highest correlation with the next 14 days of PUVs. The 8 statistics include minimum, maximum, average, mode, standard deviation, and 25$^{\text{th}}$, 50$^{\text{th}}$, and 75$^{\text{th}}$ percentiles. The choice of these 8 statistics was motivated by our observations in Sec. III-A where we saw that these values were unique for different power sources in different rooms of ZRP DCs. The inclusion of these 8 statistics enables the ML model to learn how any observed value of PUV relates to the running statistics of the PUV time-series.

13–108. Hourly Statistics: We noted a strong seasonality component in Fig. 6 that repeated every 24 hours. Thus, we include minimum, maximum, average, and standard deviation of PUVs for each of the past 24 hours. These statistics enable the ML model to learn the how their values in the previous 24 hours relate to the next PUV.

109–114. Organizational Elements: As we saw in Sec. III-C, different organizations have different power consumption trends. Thus, including information about statistics of the PUVs of the organization allocated to the power source is imperative. We add 6 elements that provide information about the organization. These include organization code (a value between 1 and the the number of organizations in DC), mean and standard deviation of all PUVs (over the last 14 days) from all power sources assigned to the organization, and the top three peak hours of the organization in the past 24 hours.

115–122. Distribution Parameters: As different PUV time-series follow different distributions (Sec. III-D1), we include 8 elements about the distribution that the time-series (of which the PUV in the last element is part of) follows. These include a distribution identifier (a value of 1 to 7 corresponding to the 7 distributions we observed) and up to 7 parameters that define the distribution. For the distributions that are fully defined by fewer than 7 parameters, we set the remaining elements to 0.

123–182. Lagged PUVs: Recall from Sec. III-D2 that for any given power source, any of its PUVs was significantly impacted by the PUVs of the last 60 minutes. Thus, we include the last 60 PUVs (averaged per minute) of the power source in the feature vector.

B. Feature Analysis

To study the effectiveness of various elements in the feature vector, we calculate mutual correlation (MC), a statistical technique known for its effectiveness in assessing the relationship between features and predictands, independent of the ML algorithm [8]. MC of 0 suggests no association between the given feature and predictand, while 1 indicates a complete overlap between the two. Fig. 8 shows the averaged MC values for all 7 types of features discussed in Sec. IV-A. We see from this figure that while the highest association exists between lagged PUVs and the predicand (MC=0.79), all 7 types of features have non-zero MC values. This indicates that all 7 types of features contain useful information for an ML algorithm to learn and make accurate predictions.

C. Model Generation

While there are several ML approaches that one can use to generate the models, such as SARIMAX [9], RNN [10], and LSTM [11], we chose XGBoost [12] because it can fit a family of equations to a given data set during training. This perfectly aligns with our setting as we expect the mathematical model of each power source to be unique. XGBoost automatically creates activation functions during the training that, when presented with test data, execute specific equations based on the values of certain elements of the feature vector to predict a PUV. For more details on how XGBoost accomplishes this, we refer the reader to [12].

Due to its intrinsic activation functions, XGBoost enables us to generate a single model for predicting PUVs for all

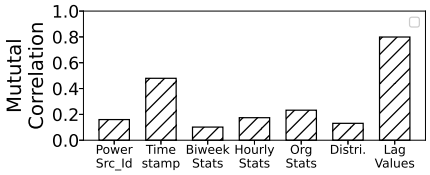

Fig. 8: Usefulness of features for prediction.

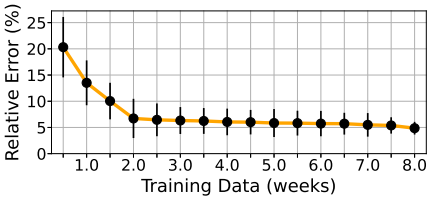

Fig. 9: Impact of training data size on RE.

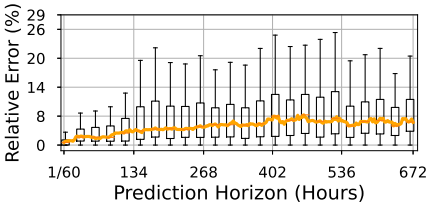

Fig. 10: Impact of time into future on RE.

power sources. This is advantageous because independent models for different power sources cannot learn the subtle interdependencies that exist between their time-series as all those power sources, after all, belong to the same company. Our experiments showed that the single XGBoost model outperformed individualized models by 11% on average in prediction accuracy 14 days in the future. Due to their lower accuracy and limited space, we do not discuss technical details and results for the individualized models further in this paper.

Recall from Sec. IV-A that we already have power source ID as the first element in our feature vector. We add two more elements to the feature vector:

183 Data center ID: an integer between 1 and the number of ZRP DCs that identifies the DC in which the power source identified in the first element is located.

184. Data Center Room ID: an integer between 1 to d that identifies the room in the DC containing the power source. d is the total number of rooms in the DC listed in element 184.

These two elements, along with the power source ID in the first element of the feature vector, enable XGBoost to create separate activation functions for different DCs, different rooms in a given DC, and different power sources in a given room that exhibit different behaviors in their PUV time-series.

The process of training the XGBoost model is straightforward. We feed the feature vectors in the training set for the given scenario to the standard XGBoost library [13] and perform a grid search with 10-fold cross-validation to select the hyperparameters. We will describe how we select the feature vectors for training and testing in the next section.

D. Evaluation

1) Individual PUV Prediction

We quantify the prediction performance for individual PUVs using relative error (RE), which is defined as $|(predicted\ value - actual\ value)|/actual\ value$.

How much past data do we need? Fig. 9 plots the RE in predicting the next value using PUV data from past w weeks, where w ranges from 0.5 to 8 weeks. For any given value of w, we randomly selected 1000 PUVs to predict such that for each PUV, at least w weeks of PUV data existed right before that PUV was collected. Next, for each of these 1000 PUVs, we first generated a prediction model using the w weeks of data collected right before this PUV, obtained an estimate of the PUV using the model, and calculated the RE, which gave us Fig. 9. Any point corresponding to any value of w in Fig. 9 is the average of these 1000 RE values and the bars above and below that point represent the standard deviation of these 1000 values. One modification we made to the feature vectors in generating the models that gave us this figure is that instead of using 14 days of data to calculate various statistics for feature vector for any value of w, we used w days of data to calculate those statistics.

We observe from Fig. 9 that as w increases, RE decreases. The decrease is more rapid initially; however, as w reaches 2 weeks, the rate of decrease in the average RE reduces. This empirically reinforces the arguments we gave in Sec. IV-A to use biweekly statistics. With $w = 2$ weeks, the average RE is just 6%. The RE as well as the standard deviation both continued to decrease as w increased. Thus, if one is not constrained by the amount of available past data and by the computational power, one can use more data from the past to get lower RE. However, as we will show, the RE of 6% suffices to predict safe maintenance windows with high accuracy.

How far in advance can we predict? Fig. 10 plots the RE in predicting a power value h hours into future using data from 2 immediate past weeks, where h ranges from 1/60 (one minute) to 672 (four weeks). For any given h, we randomly selected 1000 sets of two-week PUV data. Next, we generated a prediction model from each of these 1000 sets and iteratively predicted the next $h \times 60$ PUVs (recall that we have one PUV per minute per source) for each of our M sources. Finally, we calculated the RE of the $h \times 60^{\text{th}}$ predicted PUV of each power source with the actual PUV of that power source. This gave us 1000 values of RE per power source and $M \times 1000$ values across all power sources. A box plot for any value of h in Fig. 10 is made from these $M \times 1000$ values. We observe from this figure that the RE is close to 2% on average for data predicted up to five days into future. Beyond five days, the RE starts increasing but is still $< 6\%$ for 2 weeks into future.

The reason for observing low error in very near future is that the model has been trained on recent observations. The reason for seeing less increase in RE between 2 to 4 weeks into the future is that various internal factors remain relatively stable. For example, long-term power usage averages, peak power usage hours, and seasonal effects persist, allowing the model to maintain consistency. Observing higher standard deviation with increasing value of h is because some external factors become more unpredictable over longer periods. For example, policy changes, weather changes, and national or regional holidays introduce unexpected variability that cause RE for some predictions to get rather large.

ZRP DCs typically schedule most maintenance tasks 1 week into future. Thus, our prediction approach is accurate at forecasting safe maintenance windows, resulting in a minimal RE of only 3% when predicting PUVs a week in advance. In

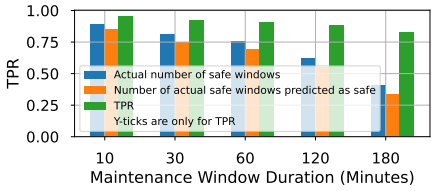

Fig. 11: True positive observations

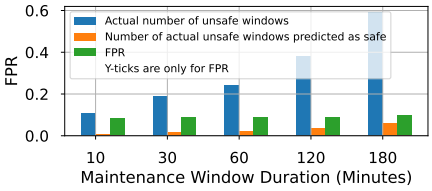

Fig. 12: False positive observations

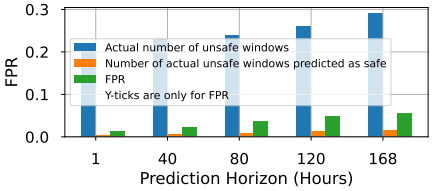

Fig. 13: 1-hour false positive observations

rare instances, ZRP DCs may need to schedule maintenance up to two weeks ahead. Even in such extreme scenarios, our method maintains a low RE of just 6%.

2) Safe Maintenance Window Prediction

To quantify the performance, we use false positive rate (FPR) and true positive rate (TPR). FPR is the ratio of the number of maintenance windows predicted as safe but were actually unsafe to the number of maintenance windows predicted as safe. TPR is the ratio of the number of maintenance windows predicted as safe and were indeed safe to the number of actual safe maintenance windows.

To calculate FPR, we first randomly selected 10,000 sets of two-week PUV data from the first 8 weeks of data (left out the last 2 weeks as we evaluate safe maintenance window prediction performance up to two weeks in future). For each set of two-week data, we generated a model and predicted next two weeks of PUVs for each power source. Next, we traversed the predicted PUVs of the power sources in each room and identified all non-overlapping time windows of duration D such that throughout each time window, the power utilization of the DC room was predicted to stay under 75%.We repeated this for five values of D: 10 min, 30 min, 1 hour, 2 hours, and 3 hours. We represent the number of maintenance windows of duration D predicted as safe with P^{D}. Next, for each predicted safe maintenance window for any given DC room, we looked at the corresponding actual PUVs of the power sources of that DC room and determined whether the actual power utilization of the room indeed stayed below 75%. We represent the number of predicted safe maintenance windows of duration D that indeed were safe with TP^{D}. This enabled us to calculate FPR as the ratio of TP^{D} with P^{D}.

To calculate the TPR, for each of the 10,000 sets of two-week data, we traversed the actual PUVs of each power source for the next two weeks and identified all non overlapping time windows of duration D in which the power utilization of the DC room stayed under 75%. Next, we counted all such time windows of duration D that had no overlap with any of the TP^{D} safe maintenance windows predicted by our approach and represented this count with TN^{D}. This enabled us to calculate TPR as the ratio of TP^{D} with TP^{D}+TN^{D}.

Figs. 11 and 12 plot the TPRs and FPRs, respectively, for all values of D. We observe from this figure that as D increases, the FPR increases and the TPR decreases. This was expected because the larger the time window to predict, the more chances of predicting an erroneous PUV and declaring an unsafe maintenance window as safe and vice versa. Nonetheless, even for the longest maintenance window of D = 3 hours, the FPR is just 0.1, and TPR is 0.82. The misclassification of windows in terms of False Positives and True Positives can be attributed to observed national and regional holidays during the data collection period. Although our paper does not evaluate the impact of holidays on the PUVs (though, we plan to do it in future), prior studies in different domains have highlighted the significant influence of holidays on altering supply and demand trends [14]. A network administrator can further supplement our approach with their subjective knowledge/experience to pick one of the many safe maintenance windows that our approach predicts.

Figs. 11 and 12 showed TPRs and FPRs from all actually safe and predicted safe maintenance windows over the entire two-week period after any selected two-week period in the 10,000 selected sets of two-week data. From our observations in Sec. IV-D1, we expect that as PUV predictions have lower RE when prediction is made in near-future than in far-future, the FPR would be lower for time windows identified in the near future. Fig. 13 plots the FPR for one-hour maintenance windows (most maintenance operations take less than an hour) predicted h hours in advance. We indeed observe that the FPR approaches 0 when maintenance window prediction is done close in the future. This is very useful: if a network administrator plans maintenance, for example, 7 days in advance, they can predict multiple safe maintenance windows 7 days in the future. One or two days before the maintenance day, the network administrator can leverage the near 0 FPR of our prediction approach and predict PUVs again using the latest data to determine which of the windows predicted as safe 7 days earlier are still safe and perform the maintenance during one of those.

V. Related Work

The inception of ZRP DCs is quite recent, with the first ZRP DC deployed by Microsoft in 2020 [1]. Thus, to the best of our knowledge, there is no prior work that either characterizes the PUVs of ZRP DCs or provides methods to predict safe maintenance windows. Nonetheless, there is prior work on predicting power usage in conventional data centers, which we summarize next.

Lin et al. [15] utilized workload related information to predict power usage trends within data centers using ML models, leveraging detailed information on each workload type. Deepika et al. [16] focused on forecasting the power usage of virtual machines in data centers to mitigate sudden fluctuations that could lead to power outages. While these studies offer valuable insights into different aspects of data center power prediction, the approaches mentioned are not

suitable for the specific problem addressed in this paper. This is because the prediction horizons of prior approaches typically range from minutes to a few hours, while our problem needs prediction up to 14 days into future. Furthermore, prior studies often rely on workload or hardware state information to train machine learning models in their experiments. Unfortunately, obtaining such detailed information for each workload, virtual machine, and server in a DC room is impractical due to the sheer scale, with hundreds of thousands of servers hosting multiple VMs in each DC room. Accessing such detailed information further raises legal concerns, especially when the servers are running public workloads and data center provider lacks authorization to access their source code or binaries.

Unlike this paper, most prior literature did not have access to the actual PUVs from real DCs. Thus, a significant amount of work has focused on estimating the power consumption of various computer programs. For example, Cao *et al*. [17] presented a source code instrumentation framework to estimate power consumption using fine-grained resource utilization data. Yeung *et al*.. [18] developed a method to predict GPU utilization for different categories of workloads and then estimate the power consumption. Liang *et al*. [19] developed a method to estimate the power consumption of servers by measuring how much of which hardware (such as disk, CPU, memory, GPC etc.) is being used. Since, neither of these works had access to telemetry devices inside DCs, they demonstrated the accuracy of these approaches using data from mobile devices and lab machines. Furthermore, these approaches did not focus on predicting the power consumption, rather focused on *estimating* the power consumption, information about which is already available to DC providers. Thus, the prediction problems addressed in these works, while interesting and often novel, are orthogonal to the problem addressed in this paper.

Finally, a large amount of work exists on predicting various other aspects of computing and non-computing systems, such as query execution times [20], computational resource utilization [21], power consumption in hospitals [22], construction [23], and electric vehicles [24], and several more. While such works exist, no two prediction problems are the same. Solution to one prediction problem does not trivially translate to another, which necessitates the prediction work presented in this paper.

VI. Conclusion

In this paper, we presented an extensive characterization of the power consumption trends in ZRP DCs of LDCP using 70 days of real data. We presented the impacts of spatial, temporal, organizational, and statistical factors on power usage in these ZRP DCs and presented several important observations. For example, the PUV time-series most frequently follow normal, power-law, and gamma distributions. The past one-hour of PUVs have significant impact on the current and future PUVs. The PUV time series contained a 7-day trend component and a 1-day seasonal component. These and several other observations we presented have important implications and applications. We demonstrated this by leveraging these observations to solve the critical problem of accurately predicting safe maintenance windows in future. Our proposed approach accurately identifies 91% of all one-hour maintenance windows up to 14 days in advance while achieving near 0 FPR. In future, we plan to leverage the insights from this work to develop applications such as power-aware workload relocation, optimal renewable energy integration, and power-aware dynamic pricing.

References

[1] Chaojie Zhang, et. al. Flex: High-availability datacenters with zero reserved power. In *ISCA*, 2021.

[2] Alessandro Rovetta. A guide on using Pearson and Spearman coefficients to detect hidden correlations. In *Cureus*, 12(1), 2020.

[3] Thomas Cokelaer. Fitter 1.7.0 documentation. https://fitter.readthedocs.io/en/latest/, 2019.

[4] Rizwan Mushtaq. Augmented dickey fuller test. 2011.

[5] Taejoon Kim, David J Love, Mikael Skoglund, and Zhong-Yi Jin. An approach to sensor network throughput enhancement by PHY-aided MAC. *IEEE TWC*, 14(2), 2014.

[6] Laura L Methot. Autocorrelation in single-subject data: a meta-analytic view. Western Michigan University, 1995.

[7] Joao Henrique F Flores, Paulo Martins Engel, and Rafael C Pinto. Autocorrelation and partial autocorrelation functions to improve neural networks models on univariate time series forecasting. In *IJCNN*, 2012.

[8] Robin Ince, et. al. A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Human Brain Mapping*, 38(3), 2017.

[9] Fahad Alharbi and Denes Csala. A Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) Forecasting Model-Based Time Series Approach *Inventions*, 7(4), 2022.

[10] Alex Sherstinsky. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Nonlinear Phenomena*, 2020.

[11] Xijun Zhang and Qirui Zhang Short-term traffic flow prediction based on LSTM-XGBoost combination model. *Computer Modeling in Engineering & Sciences*, 15(1), 2020.

[12] Tianqi Chen and Carlos Guestrin Xgboost: A scalable tree boosting system. In *SIGKDD*, 2016.

[13] XGBoost – PyPI. https://pypi.org/project/xgboost/.

[14] Hassan A. Khan, Hassan Iqbal, Muhammad Shahzad, and Guoliang Jin, RMS: Removing Barriers to Analyze the Availability and Surge Pricing of Ridesharing Services In *CHI*, 2022.

[15] Weiwei Lin, et. al. An adaptive workload-aware power consumption measuring method for servers in cloud data centers. *Computing*, 2020.

[16] T Deepika and P Prakash. Power consumption prediction in cloud data center using machine learning. *IJECE*, 10(2), 2020.

[17] Yi Cao, et. al. Analyzing the power consumption of the mobile page load. In *SIGMETRICS*, 2016.

[18] Gingfung Yeung, et. al. Towards {GPU} utilization prediction for cloud deep learning. In *USENIX HotCloud*, 2020.

[19] Yang Liang, Zhigang Hu, and Keqin Li. Power consumption model based on feature selection and deep learning in cloud computing scenarios. *IET Communications*, 14(10), 2020.

[20] Silvery Fu,et. al. On the use of {ML} for blackbox system performance prediction. In *USENIX NSDI*, 2021.

[21] Michael Borkowski, Stefan Schulte, and Christoph Hochreiner. Predicting cloud resource utilization. In *UCC*, 2016.

[22] Elena Ruiz, et. al. Energy consumption modeling by machine learning from daily activity metering in a hospital. In *ETFA*, 2017.

[23] Hai-Xiang Zhao and Frédéric Magoulès. Feature selection for predicting building energy consumption based on statistical learning method. *Journal of Algorithms & Computational Technology*, 6(1), 2012.

[24] Millend Roy, et. al.. Reliable energy consumption modeling for an electric vehicle fleet. In *SIGCAS/SIGCHI COMPASS*, 2022.