

SSH 接続した日数と実行中のプロセスとサービスをもちいた 仮想マシンの特徴の特定

鈴木 友也¹ 高橋 風太² 串田 高幸¹

概要: 仮想マシンの情報を共有することができる構成管理データベース（以下、CMDB）を作成することで、構成要素の保護管理につながる。仮想マシンの情報は、仮想マシンの名前、ホスト名、IP アドレス、ユーザ ID、最後に SSH 接続した日、使用目的とする。課題は、使用目的を取得するために仮想マシンの特徴が必要になることである。提案方式は、SSH 接続した日数による仮想マシンの分類、プロセスとサービスの抽出の 2 種類に分けられる。SSH 接続した日数による仮想マシンの分類では、仮想マシンを過去 30 日間の SSH 接続した日数からシルエット法によって決定したクラスタ数にクラスタリングする。プロセスとサービスの抽出では、プロセス一覧とサービス一覧に TF-IDF をもちいて、仮想マシンの特徴となるプロセスを抽出する。評価実験は、SSH 接続した日数による仮想マシンの分類、プロセスの抽出、サービスの抽出の 3 種類行った。対象の仮想マシンは、SSH 接続した日数による仮想マシンの分類では 79 台、プロセスの抽出では 79 台、サービスの抽出では 41 台である。SSH 接続した日数によるクラスタリングでは、18 のクラスタに分けられた。プロセスの抽出では使用されているソフトウェアは特定できる仮想マシンは 79 台中 8 台であった。サービスの抽出では、使用されているソフトウェアは特定できる仮想マシンは 41 台中 8 台であった。

1. はじめに

背景

IT サービス管理は、IT のビジネス上の利点を確保し、クライアントに対してコアビジネス価値を実行する機能を提供するために不可欠である [1]。IT サービス管理について、情報技術インフラストラクチャライブラリ（以下、ITIL）が英国商務省によって導入および配布されている [2,3]。ITIL には、プロセスが効果的に機能するために不可欠な情報源として CMDB というリポジトリが含まれている [4,5]。このリポジトリは、多くの組織で広く使用されている [6,7]。CMDB は、組織内の IT 部門に関連するすべてのものを記録するために作成される [8]。CMDB のメリットとして、インシデントの処理の支援、構成項目の関係性の把握、管理レポートの作成の支援、リソース管理の最適化があげられる [9]。

東京工科大学コンピュータサイエンス学部コンピュー

タサイエンス学科クラウド・分散システム研究室（以下、CDSL）では、7 台のハイパーバイザを使用している。CDSL に所属している学生は、自由にハイパーバイザを使用し、仮想マシンを作成、使用することができる。仮想マシンの作成が自由であることから、仮想マシンの情報が作成者本人以外に共有されていない。仮想マシンの情報を共有することができる CMDB を作成することで、構成要素の保護管理につながる [10]。仮想マシンの情報は、仮想マシンの名前、ホスト名、IP アドレス、ユーザ ID、最後に SSH 接続した日、使用目的とする。仮想マシンの名前、ホスト名、IP アドレスは、仮想マシンの識別に必要である。ユーザ ID は、仮想マシンの利用者を把握するために必要である。最後に SSH 接続した日、使用目的は、仮想マシンの削除するかしないかの判断に必要である。CMDB には、このような情報を記録しておくことができる。

図 1 に、CMDB の作成について示す。CDSL において、CMDB を作成するための情報源として、ハイパーバイザ、仮想マシン、ユーザが挙げられる。ハイパーバイザから仮想マシンの名前、ホスト名、IP アドレスを取得できる。仮想マシンからユーザ ID、最後に SSH 接続した日を取得できる。ユーザから使用目的を取得できる。

¹ 東京工科大学コンピュータサイエンス学部
クラウド・分散システム研究室
〒192-0982 東京都八王子市片倉町 1404-1

² 東京工科大学大学院バイオ・情報メディア研究科
コンピュータサイエンス専攻
クラウド・分散システム研究室
〒192-0982 東京都八王子市片倉町 1404-1

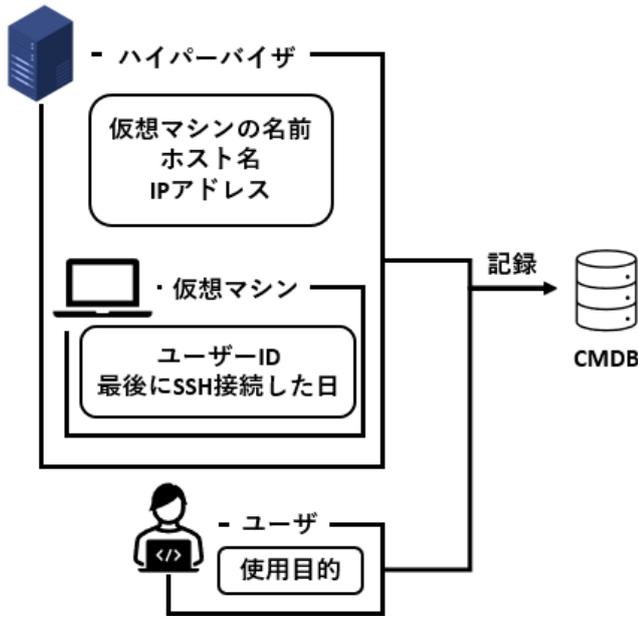


図 1 CMDB の作成

課題

課題は、使用目的を取得するために仮想マシンの特徴が必要になることである。図 2 に、課題の概要を示す。ユー

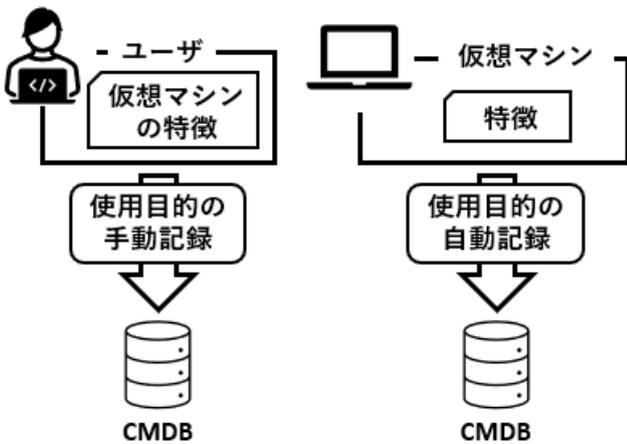


図 2 課題の概要

ザによる手動入力をする場合、ユーザが知っている仮想マシンの特徴をもとに使用目的を CMDB に入力すればよい。一方で、自動入力する場合は、自動で仮想マシンの特徴を取得し、使用目的を特定する必要がある。

各章の概要

第 2 章の関連研究では、関連する既存研究を説明する。第 3 章の提案では、課題を解決する提案方式、ユースケース・シナリオの説明をする。第 4 章の実装では、実装方法の説明をする。第 5 章の実験では、実験環境、実験結果の分析の説明をする。第 6 章の議論では、提案方式の議論をする。第 7 章のおわりにでは、全体をまとめる。

2. 関連研究

デバイス設定と CMDB 間の二重構成管理の量を減らすために、ネットワークの限られた部分のみが CMDB に保存されることで起こる、CMDB とデバイスステータスのギャップを埋めることを目指した研究がある [11]。デバイス設定の抽象化されたグラフ表現に適用されるグラフトラバース技術にもとづいて、デバイスごとの構成設定を自動的に分析し、テナントごとのネットワーク全体の構成を復元している。復元されたテナントごとのネットワーク構成の実行情報は、CMDB に自動的にアップロードされる。実装は、約 100 のテナントの管理に約 5000 の CMDB レコードが含まれるデータセンター環境に適用された。実験では、CMDB レコードの量を 2 倍にすることができた。この研究では、CMDB の情報量を増やすことはできるが、それぞれのデバイスの特徴を特定することはできない。

データセンタインフラストラクチャにおける CMDB の潜在的なデータモデル計画の作成に焦点を当て、構成項目を処理するデータベースモデルを提案している研究がある [12]。構成項目には、会社が所有及び使用するデータセンター、部屋、ラック、およびハードウェアが含まれている。調査にもとづいた 3 つの異なるレベルのソリューションが提示されている。最初の 2 つである最上位論理データモデルと論理データモデルは、Justin Group によって作成された CMDB の実用データモデルにもとづいている。3 番目のモデルである物理データモデルは、実装前のモデリングプロセスの最終ステップを表す。これらのモデルは、適応性、効率性、読みやすさを考慮して作成されており、Errillisverkot の既存システムである Netbox a Data Center Infrastructure Management と連携している。これらのモデルは概念的なものであり、実装を通じてテストする必要がある。また、構成項目の特徴を特定することはできない。

ネットワーク運用サービスの中断を引き起こす可能性のある問題である、人為的な設定ミスによる DNS レコードと IP アドレスの競合の両方を回避するためのプロビジョニングメカニズムを提案している研究がある [13]。仮想マシンはサーバとして運用され、静的に IP アドレスが与えられることを想定している。一意の名前レコードと一意のネットワークアドレスを用いて、一意の識別子を仮想マシンに与える。一意の識別子をもとに、仮想マシンの作成時に DNS サーバに接続し、DNS レコードを自動的に更新する。データルームまたはクラウドアーキテクチャに、数千の管理対象のサーバとなる仮想マシンがあることをユースケースとしている。このような量の仮想マシンを人間が介入して管理すると、エラーや大幅な遅延が発生する。この研究では、各仮想マシンの DNS レコードを手動で管理する必要がなくなり、DNS レコードの重複やエラーがなくな

る。しかし、CMDB を作成することはできず、仮想マシンの特徴を特定することはできない。

CMDB の構成項目の依存関係を記録する大きなグラフとして表すことができることを利用し、新しい変更セットはそれぞれ以前の変更セットのインスタンスと類似している可能性が高いという推測にもとづいて、変更セットを提案している研究がある [14]。アナリストが構成項目 x が新しい変更セットにあると判断した場合、CMDB に保存されている x を含む以前の変更セットを検索し、それらのセット内の構成項目を新しい変更セットに含めることを検討する必要がある。この研究では、サポートと信頼性の尺度を使用して、ノード x と y が過去の変更セットで一緒に出現した頻度にもとづいて、それらの関連性を推定している。特定の CMDB の 3 年間の履歴といくつかのフィルタリング手法にもとづいて、観測された再現率と精度の値はそれぞれ 69.8% と 88.5% であった。しかし、構成項目の特徴を特定することはできない。

3. 提案

SSH 接続した日数による仮想マシンの分類、プロセスの抽出、サービスの抽出の 3 種類の方式でそれぞれ特徴を特定する。

提案方式

提案方式は、SSH 接続した日数による仮想マシンの分類、プロセスとサービスの抽出の 2 種類に分けられる。SSH 接続した日数による仮想マシンの分類では、仮想マシンを過去 30 日間の SSH 接続した日数からシルエット法によって決定したクラスタ数にクラスタリングする。プロセスとサービスの抽出では、プロセス一覧とサービス一覧に TF-IDF をもちいて、仮想マシンの特徴となるプロセスを抽出する。

SSH 接続した日数による仮想マシンの分類

図 3 に SSH 接続した日数による仮想マシンの分類の概要を示す。各仮想マシンから過去 30 日間の SSH 接続した日数を集計する。刑事訴訟法 第百九十七条 ③に「通信履歴の電磁的記録のうち必要なものを特定し、三十日を超えない期間を定めて、これを消去しないよう、書面で求めることができる。」と定められていることから集計期間を過去 30 日とする。SSH 接続した日数から各仮想マシンのロバスト Z スコアを算出する。ロバスト Z スコアにもとづいて、仮想マシンをクラスタリングする。クラスタ数はシルエット法をもちいて決定する。クラスタリングは、k-means 法をもちいる。

ロバスト Z スコア

Z スコアとは、平均を 0、標準偏差を 1 とする標準化された値である [15]。しかし、正規分布していないデータに対

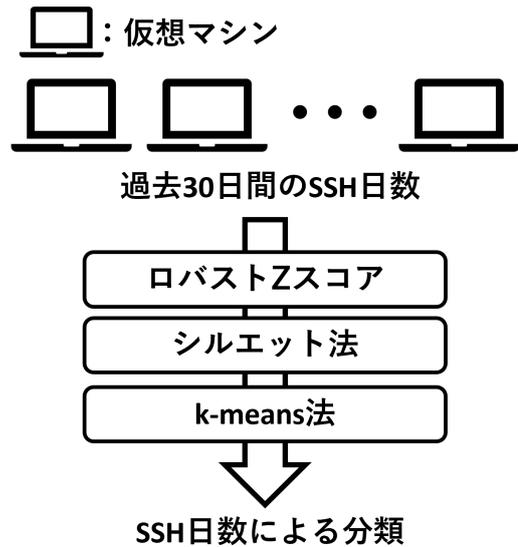


図 3 SSH 接続した日数による仮想マシンの分類の概要

しては、適切に機能しない。ロバスト Z スコアは、外れ値や非正規データに対して堅牢である [16]。図 4 に SSH 接続した日数と正規分布を示す。縦軸に確率密度、横軸に SSH

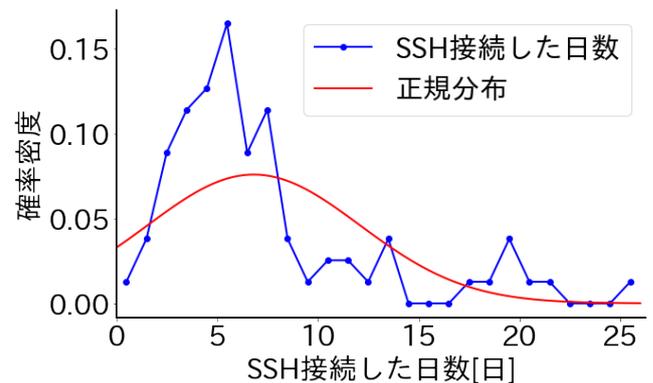


図 4 SSH 接続した日数と正規分布

接続した日数を示す。折れ線に SSH 接続した日数、曲線に正規分布を示す。SSH 接続した日数は、2024 年 11 月 6 日から 2024 年 12 月 5 日までの 30 日間に集計した。正規分布の曲線と SSH 接続した日数のヒストグラムが一致していないことから、SSH 接続した日数は正規分布していないことがわかる。SSH 接続した日数が正規分布していないため、Z スコアではなく、ロバスト Z スコアをもちいる。

式 (1) にロバスト Z スコアを求める式を示す。

$$z_{robust} = \frac{x - median(X)}{MAD \cdot 1.4826} \quad (1)$$

z_{robust} は、ロバスト Z スコアを表す。 x は、SSH 接続した日数を表す。 X は、SSH 接続した日数の集合を表す。 $median$ は、中央値を表す。 MAD は、中央絶対偏差を表す。中央絶対偏差は、標準偏差が 1 の正規分布において $1/1.4826$ になるため、1.4826 をかける。式 (2) に中央絶対

偏差を求める式を示す。

$$MAD = \text{median}(|X - \text{median}(X)|) \quad (2)$$

シルエット法

シルエット法とは、シルエットスコアをもちいてクラスタリング結果を評価する方法である [17]。式 (3) にシルエットスコアを求める式を示す。

$$s = \frac{b - a}{\max(a, b)} \quad (3)$$

s は、シルエットスコアを示す。 a は、クラスタ内の平均距離を示す。 b は、最近接クラスタの平均距離を示す。 SSH 接続した日数は最大で 30 日であるため、クラスタ数を 2 から 30 にしたクラスタリング結果を評価する。最もシルエットスコアが高かったクラスタ数でクラスタリングする。

k-means 法

k-means 法とは、データマイニングにおける典型的なクラスタリングアルゴリズムである [18]。この方法は、指定されたデータオブジェクトを反復処理によって k 個の異なるクラスタに分類する。各仮想マシンの SSH 接続した日数から算出されたロバスト Z スコアをもとにクラスタリングする。クラスタ数は、シルエット法をもちいて決定する。

3.1 プロセスとサービスの抽出

図 5 に特徴となるプロセス、サービスの抽出を示す。各

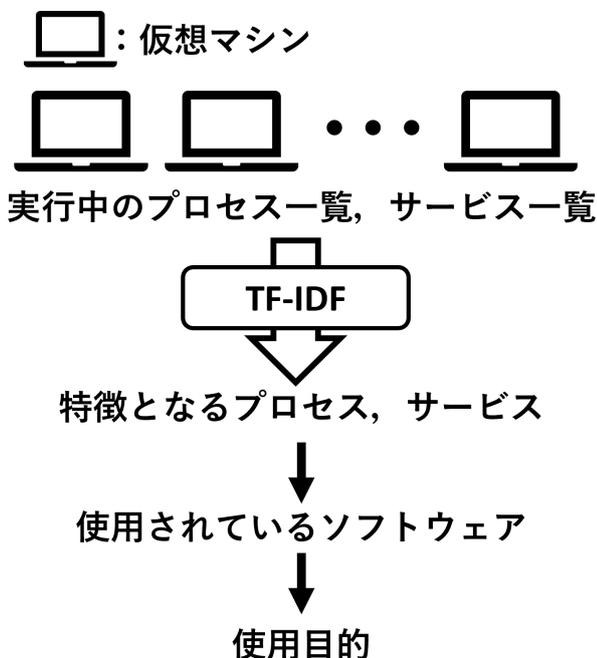


図 5 特徴となるプロセス、サービスの抽出

仮想マシンから実行中のプロセス一覧とサービス一覧を取得する。プロセス一覧とサービス一覧に対して、TF-IDF をもちいて特徴となるプロセス、サービスを抽出する。プ

ロセス、サービスからそれぞれ使用されているソフトウェアを特定し、仮想マシンの特徴の特定を行う。

TF-IDF

TF-IDF とは、情報検索システムで使用されている用語重み付けスキームである [19]。式 (4) に TF 値を求める式を示す。

$$tf(t, d) = \frac{f(t, d)}{N(d)} \quad (4)$$

tf は、TF 値を表す。 t は単語、 d は文書を表す。 $f(t, d)$ は、文書 d における単語 t の出現回数を表す。 $N(d)$ は、文書 d における総単語数を表す。式 (5) に IDF 値を求める式を示す。

$$idf(t) = \log\left(\frac{N}{df(t)}\right) \quad (5)$$

idf は、IDF 値を表す。 N は、総文書数を表す。 $df(t)$ は、単語 t を含む文書の数を表す。

ユースケース・シナリオ

ユースケース・シナリオとして、CDSL で行われる、仮想マシンのアーカイブ作業をあげる。図 6 にユースケース・シナリオの概要を示す。CDSL では、不要になった仮想

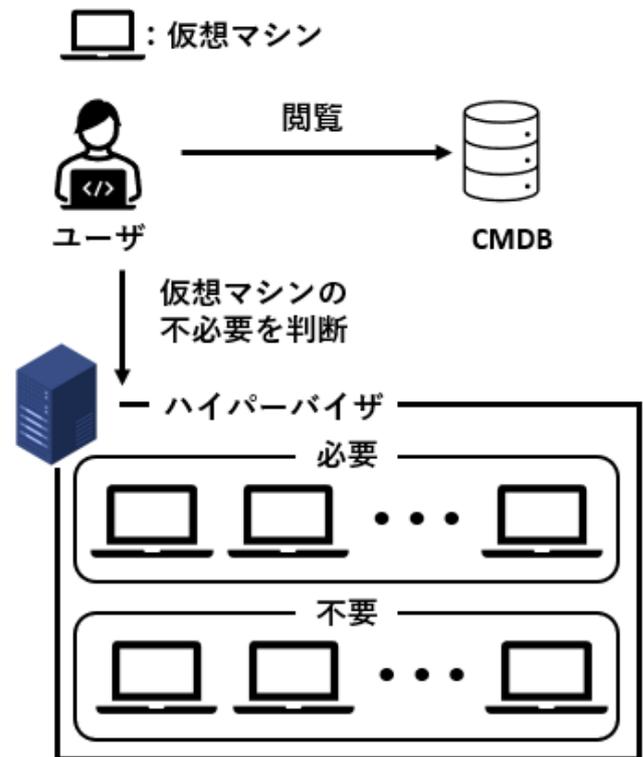


図 6 ユースケース・シナリオの概要

想マシンをアーカイブする。ユーザは、CMDB を閲覧し、仮想マシンの不必要性を判断する。CMDB を閲覧することで、各仮想マシンが誰に、どのように、どの程度使われているか把握できる。これらの情報は仮想マシンの不必要性の判断に役立てられる。

4. 実装

実装は、仮想マシンの情報の収集と提案の実装の2つの段階に分けられる。

仮想マシンの情報の収集

図7に情報収集プログラムの実装を示す。情報収集プロ

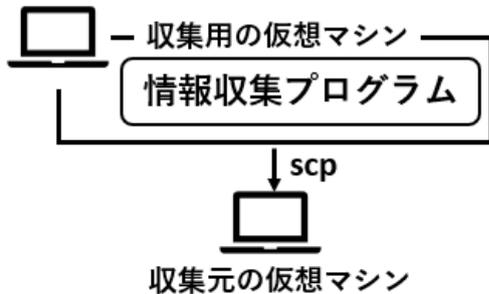


図7 情報収集プログラムの実装

グラムは、収集用の仮想マシンから収集元の仮想マシンに scp コマンドで送信する。収集元の仮想マシンは、SSH 接続した日数による仮想マシンの分類、プロセスの抽出、サービスの抽出を行う仮想マシンである。

図8に仮想マシンの情報の収集の概要を示す。収集する

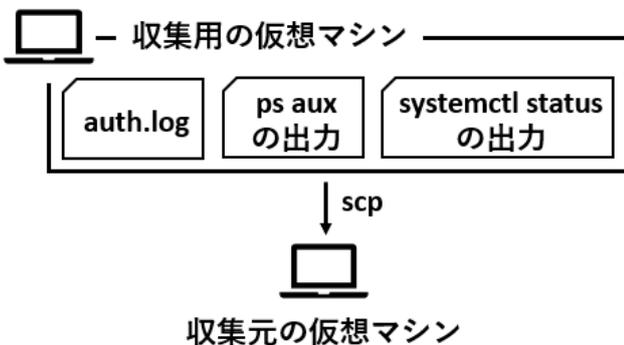


図8 仮想マシンの情報の収集の概要

情報は、ユーザが SSH 接続したログ、実行中のプロセス一覧、サービス一覧である。ユーザが SSH 接続したログは、auth.log に記録されている。実行中のプロセスは、ps コマンドで取得することができる。サービス一覧は、systemctl status コマンドで取得することができる。auth.log は、コピーを作成し、scp コマンドで収集元の仮想マシンから収集用の仮想マシンに送信する。ps aux の出力と systemctl status の出力は、テキストファイルに出力してから scp コマンドで収集元の仮想マシンから収集用の仮想マシンに送信する。

図9に、auth.log を示す。auth.log にはユーザのログイン履歴が記録されている。ユーザが SSH 接続したすると、auth.log に session opened for user と記録される。

```
c0a21069@c0a21069-vm-info:/var/log$ cat auth.log
2024-12-15T00:05:01.424250+00:00 c0a21069-vm-inf
o CRON[246222]: pam_unix(cron:session): session
opened for user root(uid=0) by root(uid=0)
2024-12-15T00:05:01.425844+00:00 c0a21069-vm-inf
o CRON[246222]: pam_unix(cron:session): session
closed for user root
●
●
●
2024-12-16T06:45:01.327044+00:00 c0a21069-vm-inf
o CRON[297436]: pam_unix(cron:session): session
opened for user root(uid=0) by root(uid=0)
2024-12-16T06:45:01.329346+00:00 c0a21069-vm-inf
o CRON[297436]: pam_unix(cron:session): session
closed for user root
```

図9 auth.log

```
c0a21069@c0a21069-vm-info:~$ ps aux
USER      PID %CPU %MEM    VSZ   RSS TTY
STAT  START   TIME COMMAND
root         1  0.0  0.4  22196  8896 ?
Ss   Dec09   0:58 /sbin/init
root         2  0.0  0.0      0     0 ?
S    Dec09   0:00 [kthreadd]
●
●
●
root      292068  0.0  0.0      0     0 ?
I    05:29   0:00 [kworker/1:2]
c0a21069 292069 100  0.2 10884  4480 pts/4
R+   05:29   0:00 ps aux
```

図10 ps aux の出力

図10に、ps aux の出力を示す。COMMAND に、プロセスの内容が示されている。

図11に、systemctl status の出力を示す。8行目からサー

```
● c0a21069-vm-info
  State: running
  Units: 427 loaded (incl. loaded aliases)
  Jobs: 0 queued
  Failed: 0 units
  Since: Mon 2024-12-09 08:04:13 UTC; 6 days
systemd: 255.4-1ubuntu8.4
CGroup: /
├─init.scope
│   └─1 /sbin/init
├─system.slice
│   └─ModemManager.service
│       └─41405 /usr/sbin/ModemManager
●
●
●
└─user@1000.service
    └─init.scope
        └─285683 /usr/lib/systemd/sy
            └─285684 "(sd-pam)"
```

図11 systemctl status の出力

ビス一覧が記述されている。

提案の実装

提案方式をもとに、Python 3.12.3 でソフトウェアを作成した。提案の実装は、SSH 接続した日数による仮想マシンの分類、プロセスの抽出、サービスの抽出の3種類に分

けられる。

SSH 接続した日数による仮想マシンの分類

図 12 に分類の実装の概要を示す。収集した auth.log か

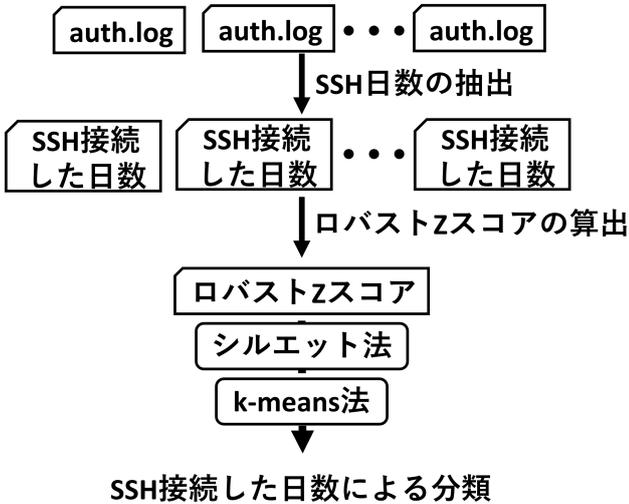


図 12 分類の実装の概要

ら SSH 接続した日数を抽出する。抽出した SSH 接続した日数をもとに、ロバスト Z スコアを算出する。算出したロバスト Z スコアをもとに k-means 法をもちいて分類する。

プロセスの抽出

図 13 に特徴となるプロセスの抽出を示す。ps aux の出

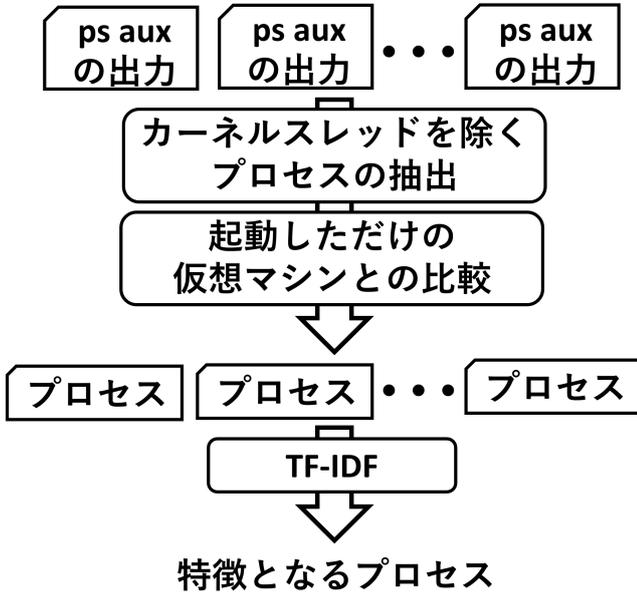


図 13 特徴となるプロセスの抽出

力から [] で囲われたカーネルスレッドを除外してプロセス抽出する。抽出されたプロセスを起動しただけの仮想マシンと比較し、デフォルトで実行されるプロセスを除外する。残ったプロセスに TF-IDF をもちい、特徴となるプロセスを抽出する。

サービスの抽出

図 14 に、特徴となるサービスの抽出を示す。systemctl

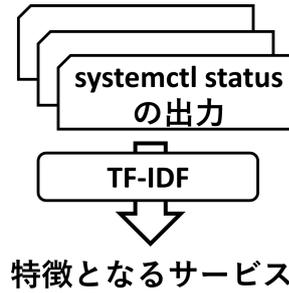


図 14 特徴となるサービスの抽出

status の出力に対し、TF-IDF をもちいて特徴となるサービスを抽出する。

5. 評価実験

評価実験では、SSH 接続した日数による仮想マシンの分類の実験結果とプロセスの抽出、サービスの抽出の評価実験結果を示す。

実験環境

VMware ESXi 8.0U2 を使用して作成された仮想マシンを対象に実験を行った。対象の仮想マシンは、SSH 接続した日数による仮想マシンの分類では 79 台、プロセスの抽出では 79 台、サービスの抽出では 41 台である。SSH 接続した日数は、2024 年 11 月 6 日から 2024 年 12 月 5 日までの 30 日間に集計した。プロセスとサービスは、2024 年 12 月 16 日に収集した。プロセスとサービスからの使用されているソフトウェアの特定は、著者自身が行った。

SSH 接続した日数による仮想マシンの分類

図 15 に、SSH 接続した日数によるクラスタリングの結果を示す。

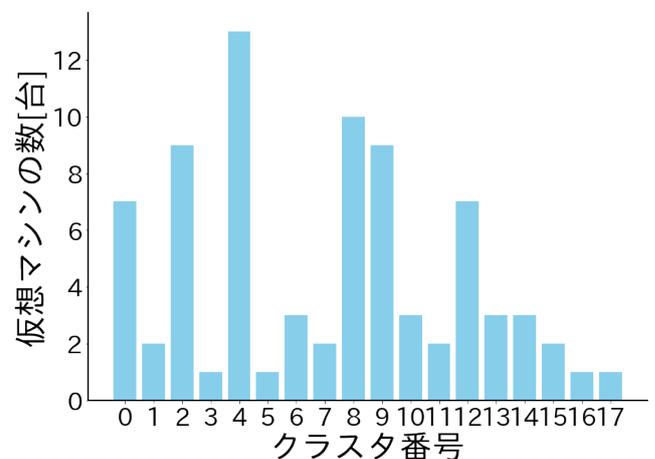


図 15 SSH 接続した日数によるクラスタリングの結果

シルエット法をもちいて決定したクラスタ数は 18 であった。クラスタ番号が大きくなるほど SSH 接続した日数は多くなる。

プロセスの抽出

図 16 に、抽出されたプロセスを示す。縦軸にプロセス、

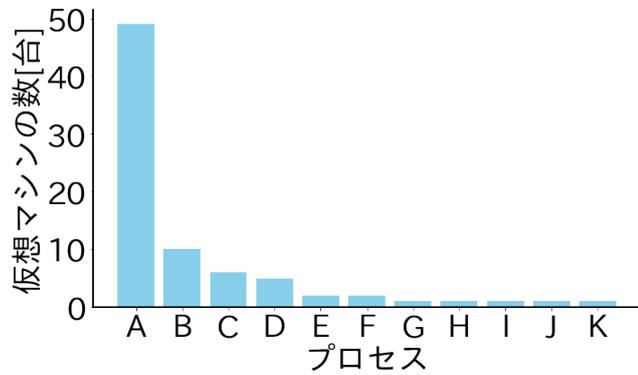


図 16 抽出されたプロセス

横軸に仮想マシンの数を示す。表 1 にプロセスと対応する記号を示す。使用されているソフトウェアが特定できるプ

表 1 プロセスと対応する記号

対応する記号	プロセス
A	usr
B	systemd
C	bin
D	apache2
E	3fcd4fc3ae2ba4d5 77d4ee08ad7092538cd7a7f0 da701efa2a8807d44a25f66
F	e50868881d9744d0d00 27dda983507e867b3787482eb 00005d97239d9aa501a5
G	bitnami
H	k0s
I	microk8s
J	e82313669fe2739df5 3b3870076163d1fe7785336a68b 4771685219e51c9785d
K	ac0baecab6b7fa399482b 08daa7117e7f2a0b1a739da5c31131b ea4ebfaedfec

プロセスは、apache2、bitnami、k0s、microk8s であり、仮想マシン 79 台中 8 台であった。apache2 と bitnami は web サーバのため、k0s と microk8s は Kubernetes のために仮想マシンを使用していると推測できる。

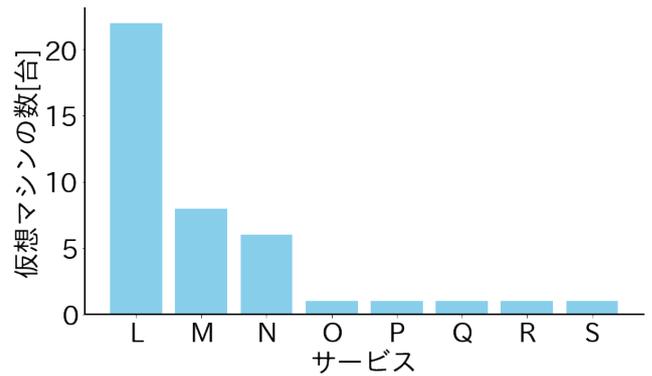


図 17 抽出されたサービス

サービスの抽出

図 17 に、抽出されたサービスを示す。縦軸にサービス、横軸に仮想マシンの台数を示す。表 2 にサービスと対応する記号を示す。使用されているソフトウェアが特定できる

表 2 サービスと対応する記号

対応する記号	サービス
L	service
M	systemd
N	containerd
O	snap
P	bitnami
Q	pki
R	nginx
S	apt

サービスは、containerd、bitnami、nginx であり、仮想マシン 41 台中 8 台であった。containerd は Docker のため、bitnami と nginx は web サーバのために仮想マシンを使用していると推測できる。

6. 議論

提案方式では、ps aux の出力の COMMAND に対して TF-IDF をもちいて特徴となるプロセスを抽出している。カーネルスレッドやデフォルトで実行されるプロセスは除外して処理を行っているが、表 1 の E のような COMMAND に出力される id は除外できていないため、使用されているソフトウェアを推測できないプロセスが含まれてしまった。TF-IDF をもちいる前に COMMAND に出力されるプロセスを分析し、不要な出力を除外することで使用されているソフトウェアを推測できないプロセスを除外できる。

提案方式では、systemctl status の出力に対して TF-IDF をもちいて特徴となるサービスを抽出している。systemctl status の出力に対してそのまま TF-IDF をもちいているため、表 2 の L のようななどの仮想マシンにも含まれている特徴として捉えられないサービスが結果に含まれてしまった。systemctl status の出力に対して分析を行い、どの仮

想マシンにも含まれるサービスを除外してから TF-IDF をもちいることで特徴として捉えられないサービスを除外できる。

提案方式では、特徴となるプロセスと特徴となるサービスの抽出している。特徴となるプロセスと特徴となるサービスから使用されているソフトウェアの特定は、自動的に行うことができず、人間による確認が必要となった。抽出されたプロセスとサービスに対し、事前に使用されているソフトウェアを定義しておくことや人工知能を使用することで自動で特定ができる。

7. おわりに

課題は、使用目的を取得するために仮想マシンの特徴が必要になることである。SSH 接続した日数による仮想マシンの分類、プロセスの抽出、サービスの抽出の3種類の方式でそれぞれ特徴を特定する。提案方式は、SSH 接続した日数による仮想マシンの分類、プロセスとサービスの抽出の2種類に分けられる。SSH 接続した日数による仮想マシンの分類では、仮想マシンを過去30日間のSSH接続した日数からシルエット法によって決定したクラスタ数にクラスタリングする。プロセスとサービスの抽出では、プロセス一覧とサービス一覧にTF-IDFをもちいて、仮想マシンの特徴となるプロセスを抽出する。評価実験は、SSH接続した日数による仮想マシンの分類、プロセスの抽出、サービスの抽出の3種類を行った。対象の仮想マシンは、SSH接続した日数による仮想マシンの分類では79台、プロセスの抽出では79台、サービスの抽出では41台である。SSH接続した日数によるクラスタリングでは、18のクラスタに分けられた。プロセスの抽出では使用されているソフトウェアは特定できる仮想マシンは79台中8台であった。サービスの抽出では、使用されているソフトウェアは特定できる仮想マシンは41台中8台であった。使用されているソフトウェアの特定は、自動的に行うことができず、人間による確認が必要となった。

参考文献

[1] Na-Lampang, N. and Vatanawood, W.: Development of an ontology-based configuration management system, *2016 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1–6 (online), DOI: 10.1109/ECAI.2016.7861164 (2016).

[2] Sharifi, M., Ayat, M. and Sahibudin, S.: Implementing ITIL-Based CMDB in the Organizations to Minimize or Remove Service Quality Gaps, *2008 Second Asia International Conference on Modelling Simulation (AMS)*, pp. 734–737 (online), DOI: 10.1109/AMS.2008.144 (2008).

[3] AlShamy, M. M., Elfakharany, E. and ElAzem, M. A.: Information technology service management (ITSM) implementation methodology based on information technology infrastructure library ver. 3 (ITIL V3), *International Journal of Business Research and Management*

(*IJBRM*), Vol. 3, No. 3, pp. 113–132 (2012).

[4] Brenner, M. and Gillmeister, M.: Designing CMDB data models with good utility and limited complexity, *2014 IEEE Network Operations and Management Symposium (NOMS)*, pp. 1–15 (online), DOI: 10.1109/NOMS.2014.6838375 (2014).

[5] Office, C.: *ITIL Continual Service Improvement 2011 Edition*, The Stationery Office, GBR (2011).

[6] Steinberg, R. A.: *Implementing ITIL: Adapting Your IT Organization to the Coming Revolution in IT Service Management*, Trafford (2005).

[7] Atherton, M.: Deploying CMDB Technology Pragmatism and realism will deliver the benefits, *Freeform Dynamics Ltd*, pp. 185–191 (2009).

[8] Ayat, M., Sharifi, M., Ibrahim, S. and Sahibudin, S.: CMDB Implementation Approaches and Considerations in SME/SITU's Companies, *2009 Third Asia International Conference on Modelling Simulation*, pp. 381–385 (online), DOI: 10.1109/AMS.2009.113 (2009).

[9] Wu, M.-S.: The Benefit and Cost Factors of CMDB Implementations: An Investigation of three Organizations in Taiwan, *Procedia - Social and Behavioral Sciences*, Vol. 147, pp. 64–69 (online), DOI: <https://doi.org/10.1016/j.sbspro.2014.07.104> (2014). 3rd International Conference on Integrated Information (IC-ININFO).

[10] Berggren, R. and Londögård, D.: Benchmarking and comparison of a relational and a graph database in a CMDB context, *LU-CS-EX* (2020).

[11] Himura, Y. and Yasuda, Y.: Bridging the Gap between Tenant CMDB and Device Status in Multi-Tenant Datacenter Networking, *IEICE Transactions on Communications*, Vol. 98, No. 11, pp. 2132–2140 (2015).

[12] Fallström, S.: CMDB: Data Model Plan for Suomen Erillisverkot Oy (2024).

[13] Marian, C. V.: DNS Records Secure Provisioning Mechanism for Virtual Machines automatic management in high density data centers, *2021 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pp. 1–5 (online), DOI: 10.1109/BlackSeaCom52164.2021.9527811 (2021).

[14] Nadi, S., Holt, R. and Mankovskii, S.: Does the Past Say It All? Using History to Predict Change Sets in a CMDB, *2010 14th European Conference on Software Maintenance and Reengineering*, pp. 97–106 (online), DOI: 10.1109/CSMR.2010.14 (2010).

[15] Abdi, H.: Z-scores, *Encyclopedia of measurement and statistics*, Vol. 3, pp. 1055–1058 (2007).

[16] Bilal Khan, W.: Differences Between Standard Deviation and Robust NIQR Method During Z-Score Evaluation, *2018 15th International Conference on ElectroMagnetic Interference Compatibility (INCEMIC)*, pp. 1–4 (online), DOI: 10.1109/INCEMIC.2018.8704599 (2018).

[17] Shutaywi, M. and Kachouie, N. N.: Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering, *Entropy*, Vol. 23, No. 6 (online), DOI: 10.3390/e23060759 (2021).

[18] Na, S., Xumin, L. and Yong, G.: Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm, *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, pp. 63–67 (online), DOI: 10.1109/IITSI.2010.74 (2010).

[19] Aizawa, A.: An information-theoretic perspective of tf-idf measures, *Information Processing Manage-*

ment, Vol. 39, No. 1, pp. 45–65 (online), DOI:
[https://doi.org/10.1016/S0306-4573\(02\)00021-3](https://doi.org/10.1016/S0306-4573(02)00021-3) (2003).