

電子情報通信学会
2012年8月AI/SC研究会

サービスレベルの維持自動化に向けた リソースキャパシティ予測方式

NEC 情報・ナレッジ研究所

○大野允裕 外川遼介 富沢伸行 八木真二郎 岡田勲

発表内容

背景

- クラウドサービスにおけるサービスレベル管理

アプローチと課題

- ばらつきが大きい不規則な監視情報ではリソースキャパシティの予測が困難

提案方式

- 頻度分布を用いたリソース使用率の特定

方式評価

考察

まとめ

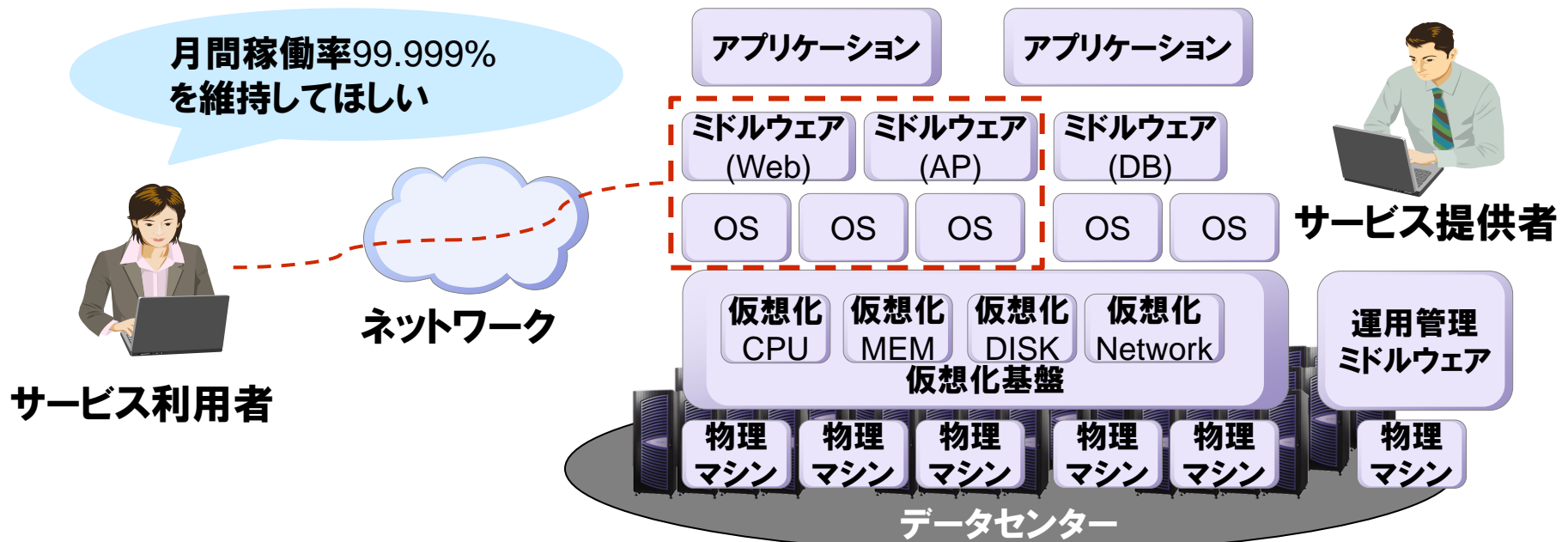
クラウドサービスにおけるサービスレベル管理

クラウドサービスの普及

- 仮想化されたシステムをネットワーク経由のサービスとして提供
- パブリッククラウド, プライベートクラウド, ハイブリッドクラウド

プライベートクラウドでは可用性や性能などのサービスレベルを重視

- 企業向けのクラウドサービスでは, **サービスレベルの維持**が求められる



リソースキャパシティ管理

サービス提供者は、サービスレベルに設定した目標値を満たすため、必要となるリソースキャパシティ(割当CPU数、割当メモリ量)の過不足がないように、管理する

サービス提供前

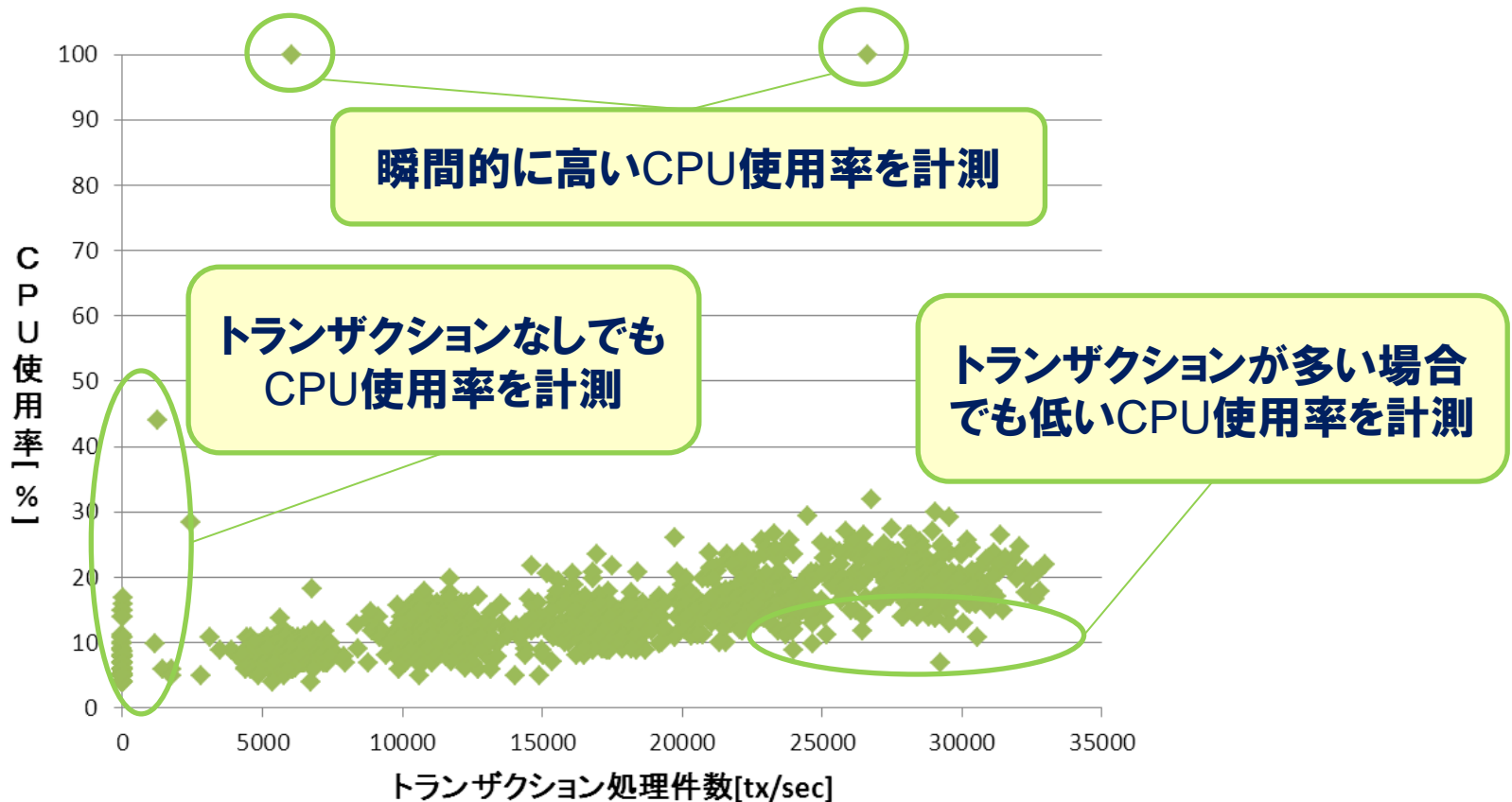
- **小規模なテスト環境**で計測されるリソース使用率の値から、リソース使用率の値の変動を推定し、必要なリソースキャパシティを設計する

サービス提供後

- **本番環境**で計測されるリソース使用率の値から、リソース使用率の値の変動を再推定し、必要なリソースキャパシティを見直す
- ⇒ アクセス状況や運用状況などによって、計測されるリソース使用率の値のばらつきが大きくなるため、リソースキャパシティの見直しが困難

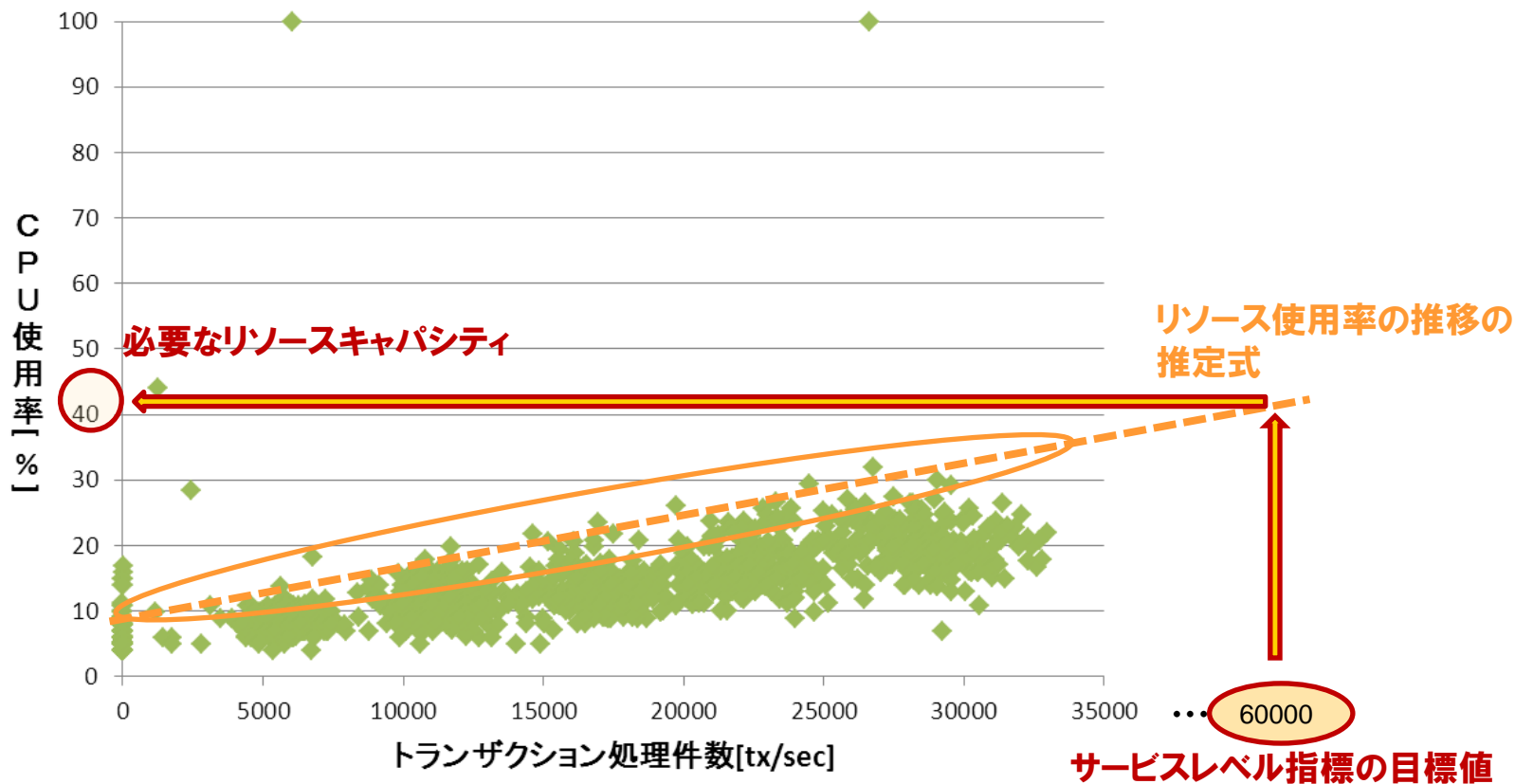
本番環境の不規則な監視情報

アクセス状況や稼働状況, 計測状況, 運用状況によって,
ばらつきが大きい不規則な監視情報になる



アプローチ

平常時に高負荷なリソース使用率を特定し、その推移の推定式に基づいて、サービスレベル指標の目標値に必要なリソースキャパシティを予測する



アプローチと課題

平常時に高負荷なリソース使用率を特定し，その推移の推定式に基づいて，サービスレベル指標の目標値に必要なリソースキャパシティを予測する

運用カレンダーや運用状況を考慮した運用経験則によって，平常時に高負荷な監視情報を特定する

⇒ 運用管理者の負担が大きい

⇒ 経験則によっては，予測の正確さが低くなる

サービスレベル指標の目標値を達成したときのリソース使用率の値を計測し，それらの監視情報から必要とされるリソースキャパシティを確率的に予測する

⇒ 本番環境は過剰なリソースキャパシティが設計されることが多いため，確率的に予測可能な十分な情報を収集できない

提案方式

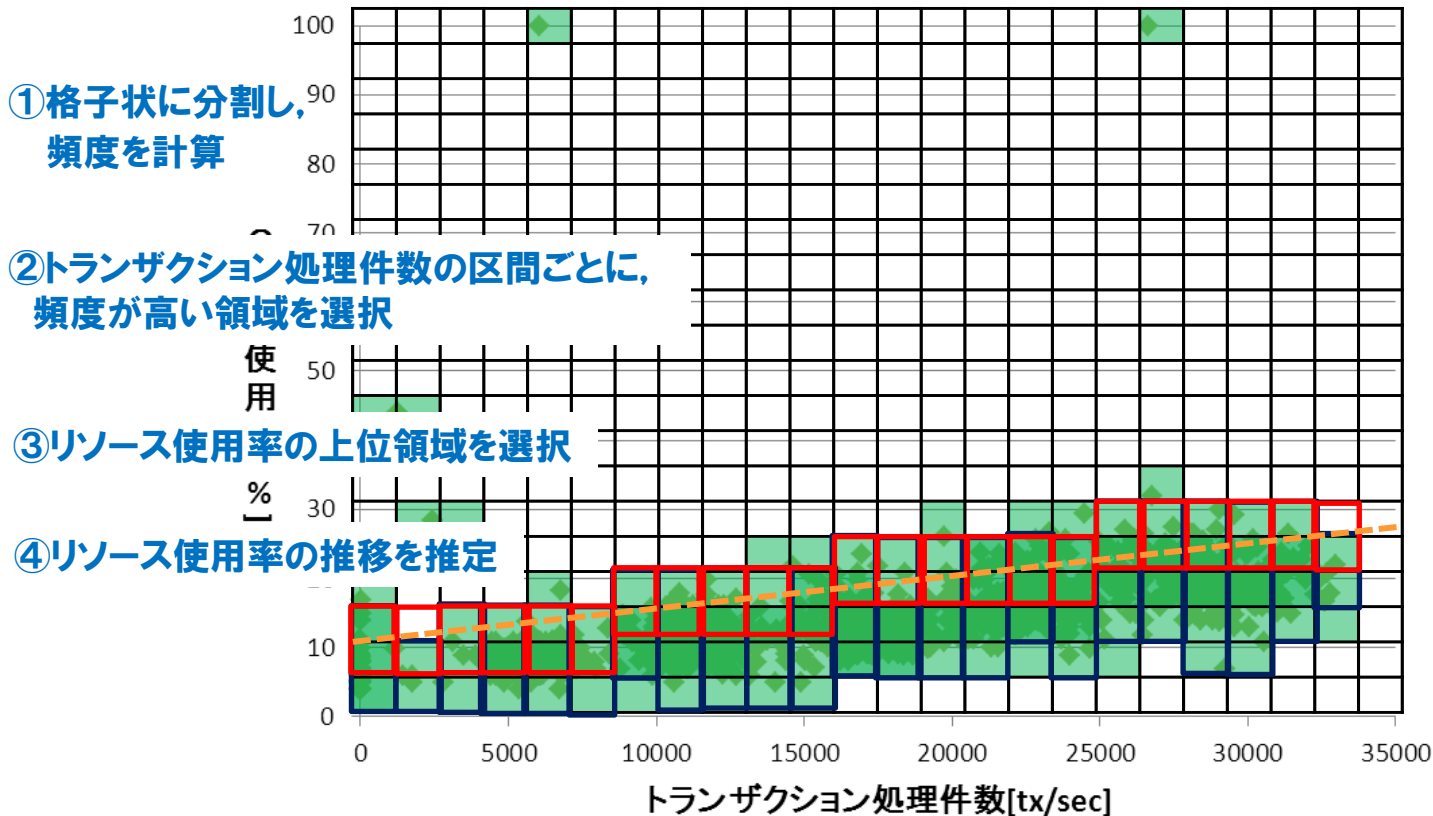
サービスレベル指標とリソース使用率との頻度分布を用いて、頻度が高いリソース使用率の上位領域を機械的に特定し、リソースキャパシティを予測する

手順

- ① サービスレベル指標とリソース使用率との二次元空間を正方形の格子状に分割し、各格子の頻度を計算することで、頻度分布を生成
- ② 頻度が高いリソース使用率の領域を選択
⇒ 瞬間的に発生するリソース使用率の極端な値を除く
- ③ リソース使用率の上位領域を選択
⇒ 過小なリソースキャパシティとなるリソース使用率の低い値を除く
- ④ 選択された領域の値から、リソース使用率の推移を推定

提案方式

サービスレベル指標とリソース使用率との頻度分布を用いて、頻度が高いリソース使用率の上位領域を機械的に特定し、リソースキャパシティを予測する



リソースキャパシティ予測ツール

不足が想定される割り当リソース, 余剰な割り当リソースを特定

分析対象とする
サービス, 期間,
SLOを表示

サービスレベルで,
キャパシティ超過
があるか否かを
表示

ノードレベルで,
キャパシティ超過
見込みの大きい
リソースTop5を,
一覧表示



方式評価

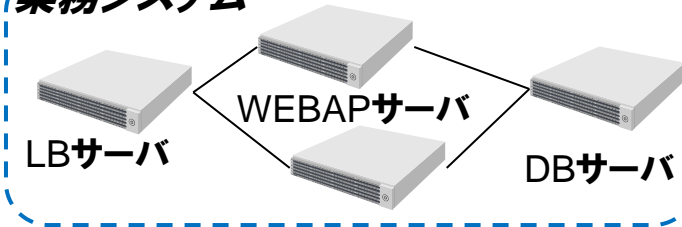
評価環境

- オークションサイトの振る舞いを模倣した環境を構築

高負荷と低負荷の2種類の
トランザクションを発生



業務システム



LBサーバのTx理件数と、
各サーバのCPU使用率と
メモリ使用率を計測

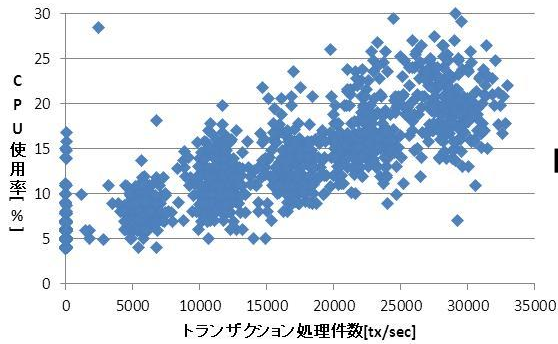
運用管理サーバ

評価方法

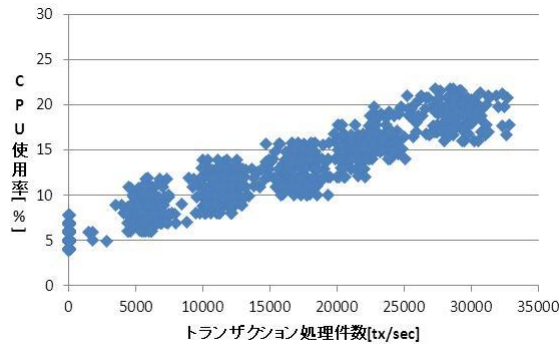
- 平常時に高負荷な監視情報の特定精度
 - ・ノイズを除いた高負荷かつばらつきが少ない監視情報を特定できるかを評価する
⇒フィルタリング前後の統計情報(変域、平均値、標準偏差)を比較する
- リソースキャパシティの正解率
 - ・どの程度の正確なリソースキャパシティを予測できたか否かを評価する
⇒正解とするリソースキャパシティの平均値の95%信頼区間の上限値と比較する

評価結果(1) 監視情報の特定精度

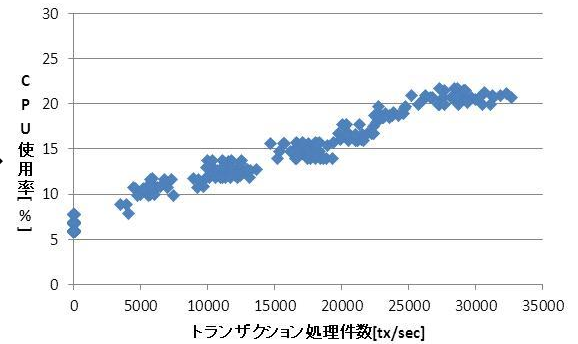
頻度が高いリソース使用率の上位領域を機械的に特定
変域や標準偏差値から、ばらつきが少なくなっていることを確認



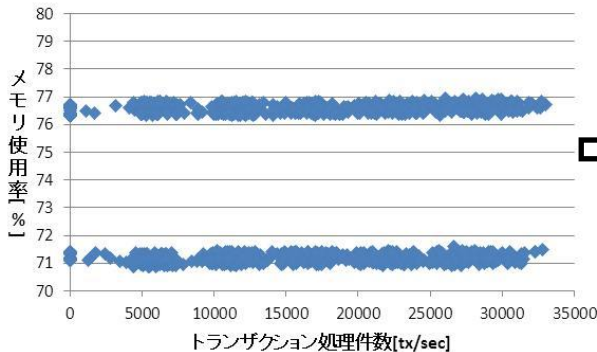
計測された監視情報



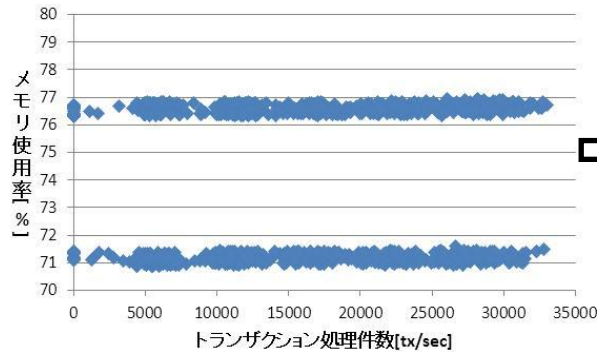
頻度が高い領域の選択後



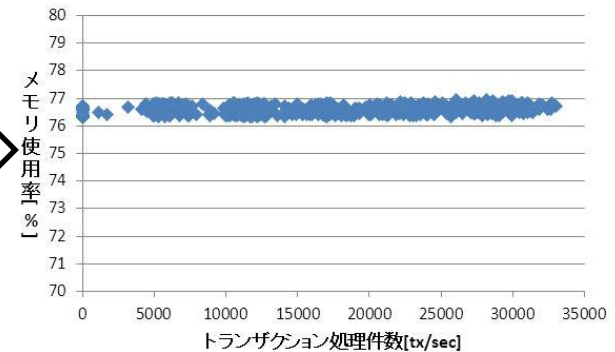
上位領域の選択後



計測された監視情報



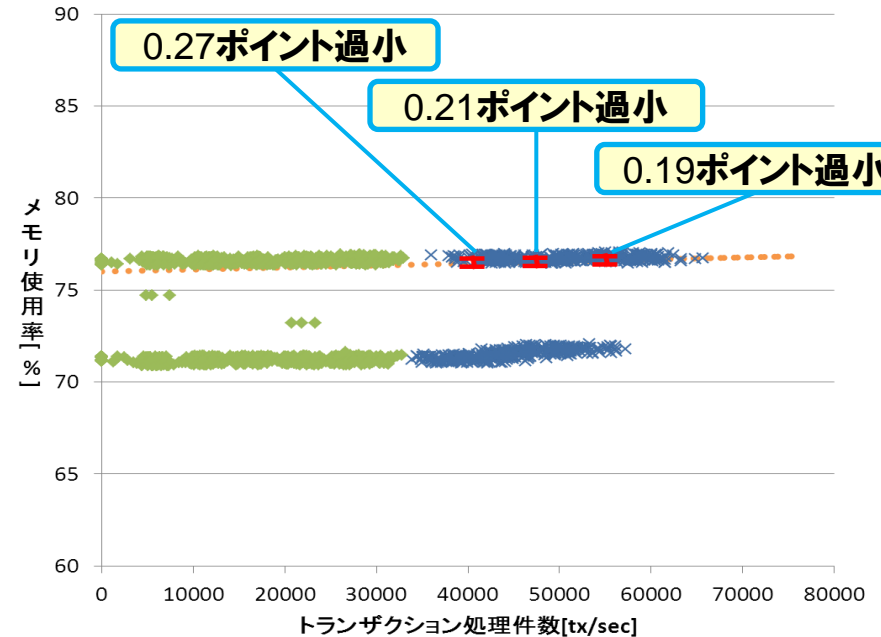
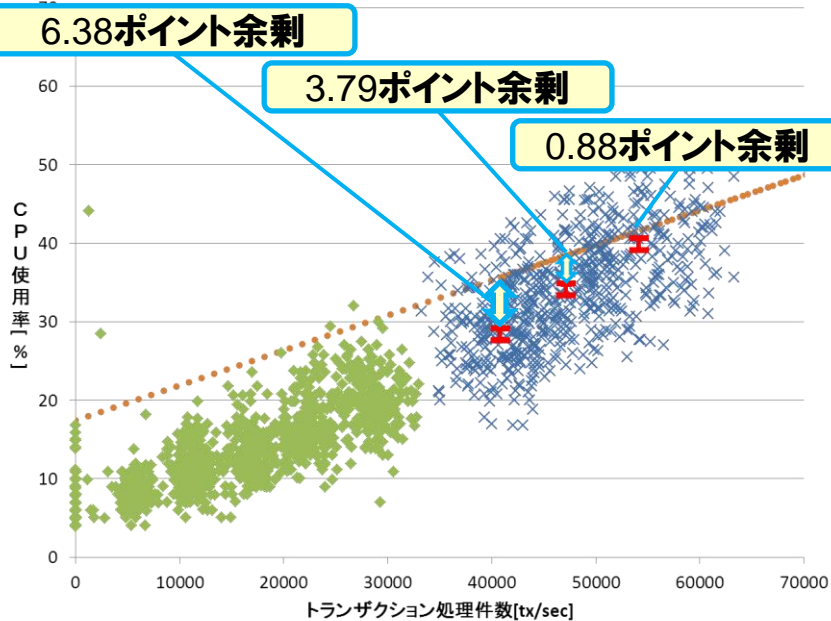
頻度が高い領域の選択後



上位領域の選択後

評価結果(2) リソースキャパシティの正解率

CPU使用率では、正解から1~6ポイント程度のキャパシティを予測
メモリ使用率では、正解から1ポイント未満のキャパシティを予測



◆...計測可能とするリソース使用率の計測値

■...正解とするリソースキャパシティ(平均値の95%信頼区間)

×...予測対象とするリソース使用率の計測値

---...提案方式によるリソース使用率の推移の推定式

考察

CPU使用率とメモリ使用率との種別に関わらずに、適用可能

- トランザクション処理件数の増加に対し、該当リソース使用率も増加する傾向があるため、適用可能であった。
- 他のリソース使用率も、今回の評価結果で示したCPU使用率やメモリ使用率と同様の推移であれば適用可能で、同様の結果が得られるものと考ええる。

おわりに

まとめ

- 本番環境では、アクセス状況や稼働状況、計測状況、運用状況によって、ばらつきが大きい不規則な監視情報になるため、リソースキャパシティの予測が困難
- 頻度分布を用いて頻度が高いリソース使用率の上位領域を特定し、リソースキャパシティを予測する方式を提案**
- 提案方式は、正解とする95%信頼区間の上限値と比較し、6ポイント程度のわずかな余剰があるリソースキャパシティを予測できることを確認**

今後の取り組み

- より頑健な推移式の推定方法への改良
- 実際の運用管理環境での検証

NECグループビジョン2017

人と地球にやさしい情報社会を
イノベーションで実現する
グローバルリーディングカンパニー



Empowered by Innovation

NEC