



# 匿名化

## 次世代ネットワークにおける匿名化を用いたデータ利活用

### 匿名化

#### 個人情報と匿名化

ルータを通過するパケットを取得・解析した結果を公開する際、データに含まれる個人情報を保護する必要がある。この情報保護の手段としては匿名化による値の一般化が著名で、広く利用されている。しかしながら匿名化を用いた安全で実際的なデータ公開方法については研究・確立されておらず、また計算コストが膨大で汎用的な利用や低遅延な匿名化は困難であった。

本研究室では上記問題を解決するため、データ二次利用基盤の提案および、ハードウェア実装による高速な匿名化処理を目指している。

#### 匿名化処理 ( $k$ 匿名性, $l$ 多様性)

- 一般的に使用される匿名化手法

氏名	郵便番号	年齢	症状(注目属性)
Tom	25423	20	腹痛
Bob	25425	23	腹痛
Alice	25520	32	冷え性
Fred	25520	32	発熱



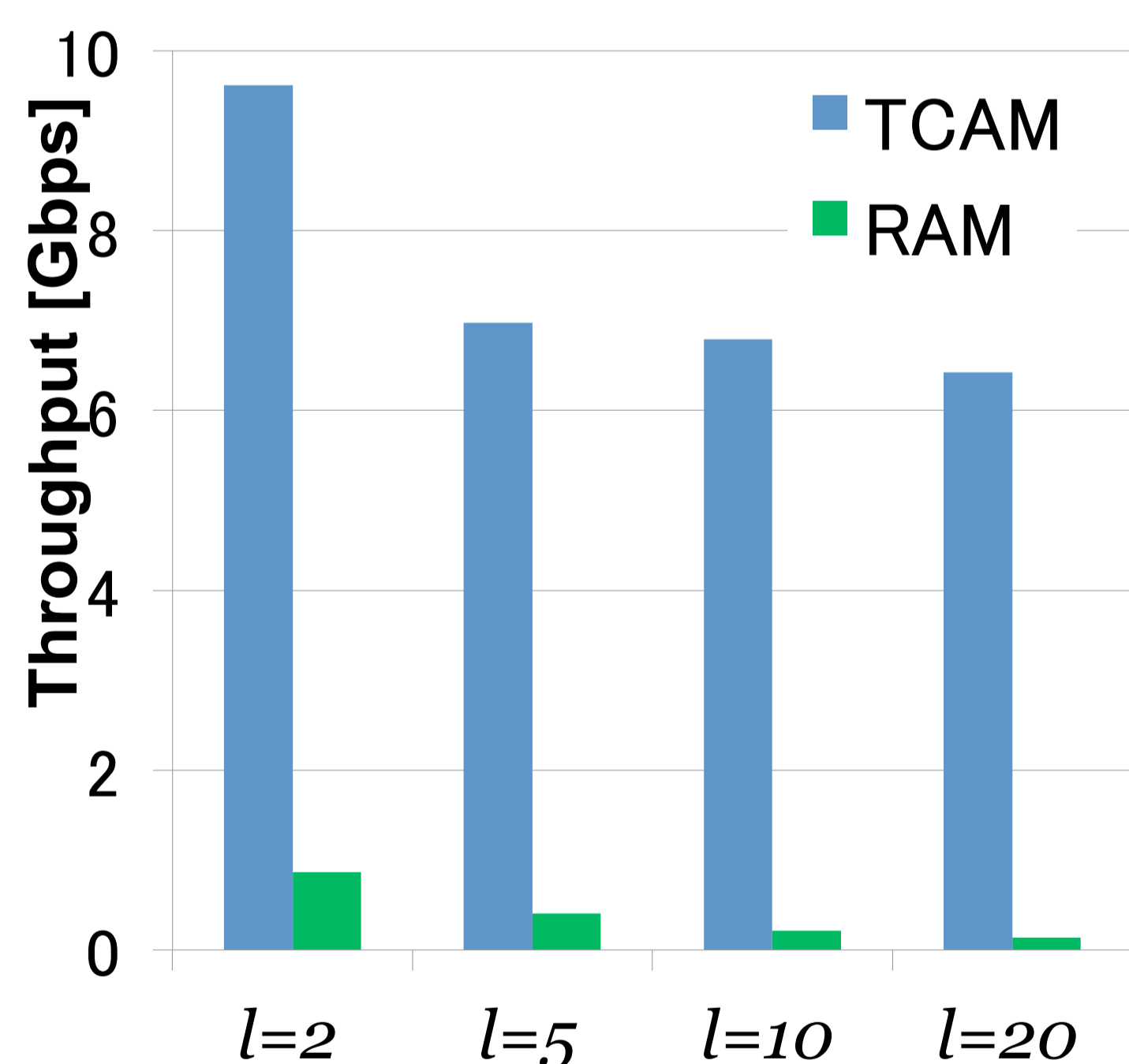
**$k$ 匿名性**  
同じ準識別子の項目が $k$ 個以上

**$l$ 多様性**  
準識別子グループ内の注目属性の種類が $l$ 個以上

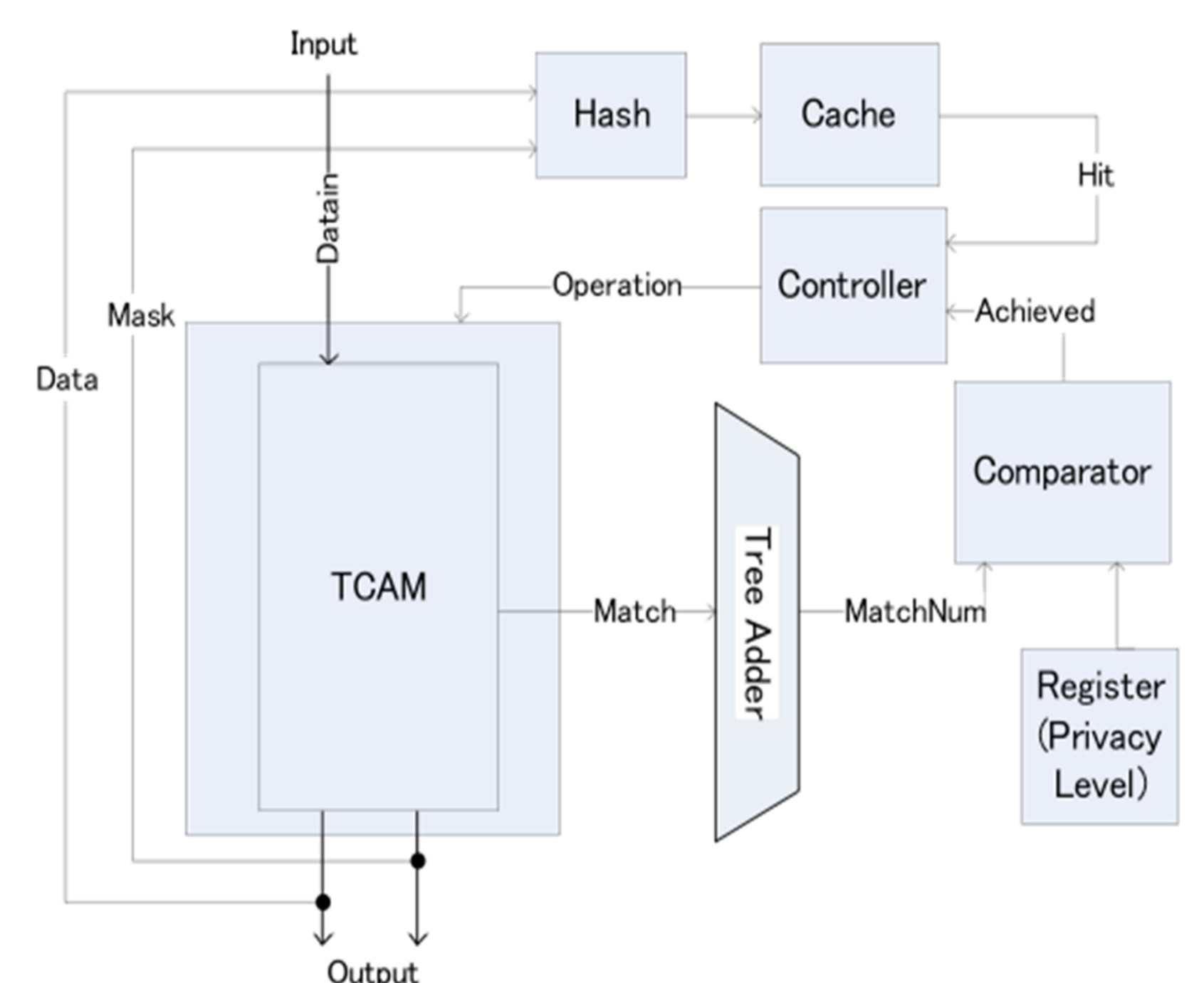
氏名	郵便番号	年齢	症状(注目属性)
Tom	2542*	2*	腹痛 $l = 1$
Bob	2542*	2*	腹痛 $l = 1$
Alice	25520	32	冷え性 $l = 2$
Fred	25520	32	発熱 $l = 2$

#### TCAMを用いた匿名化機構

- TCAMによる探索機能による高速匿名化処理
- キャッシュを用いることでデータ損失率を削減
- 従来の匿名化機構に比べ10-50倍高速



TCAM匿名化機構のスループット



TCAM匿名化機構のアーキテクチャ

#### ハードウェア・アクセラレータによる高速化

- FPGAによる専用回路を用いた高速な匿名化処理
- ソフトウェアによる処理と比較し350~400倍のスループットを実現
- BloomFilterを用いることにより、 $k$ 匿名性と $l$ 多様性の適合を同時に確認可能

研究者名

慶應義塾大学理工学部 西研究室





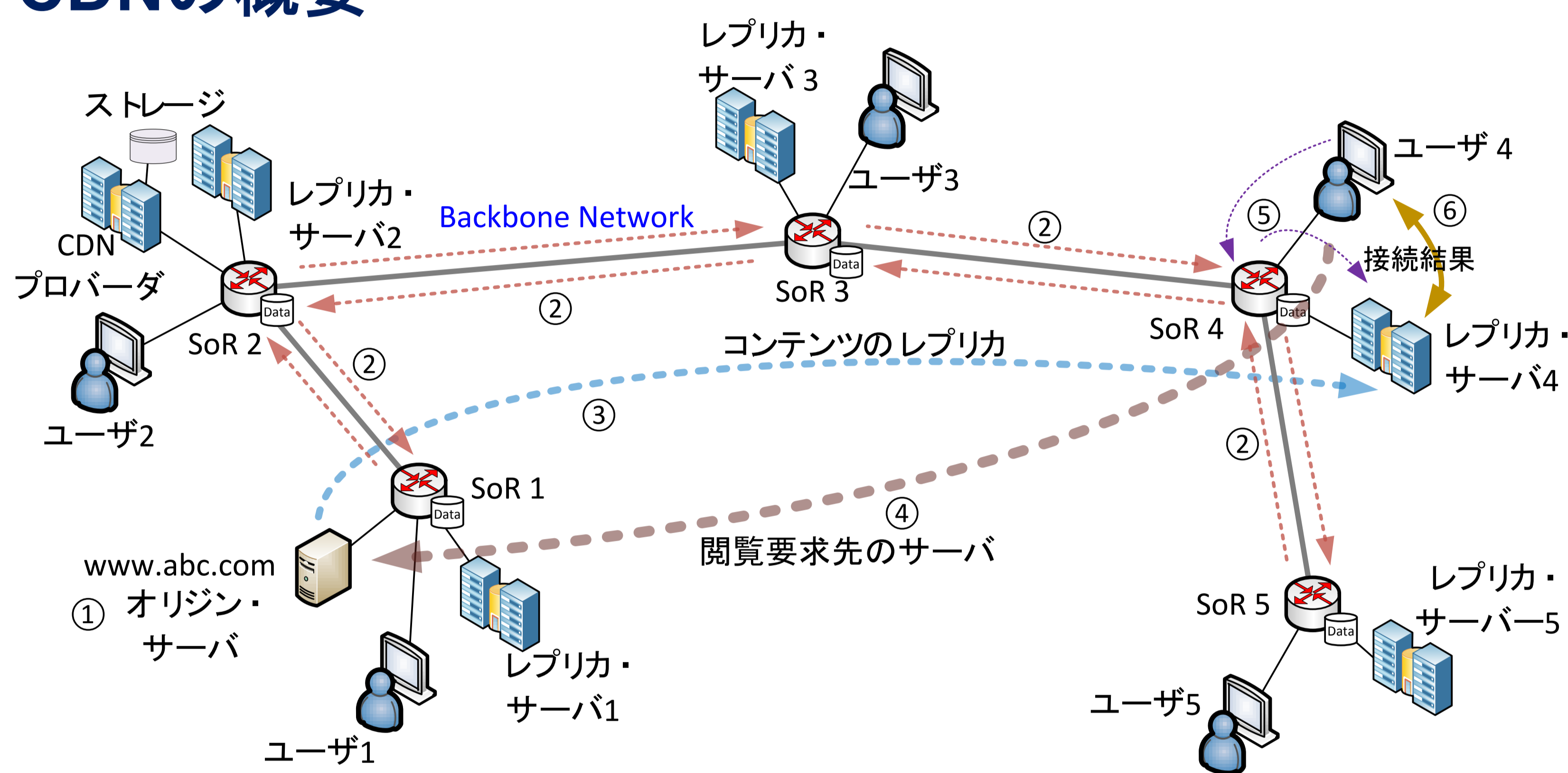
# SoRによるコンテンツ・デリバリ・ネットワーク

## 次世代CDNのSoRに基づいたモデリングおよび実装

### SoRによる次世代CDN

インターネットは、複数のネットワーク間の通信をルータによって中継することにより、世界規模での通信を可能にしている。本研究は、特に近年利用の増加が著しい動画などのコンテンツ配信にサービス指向ルータを利用することで、ルーティング最適化、低遅延化、耐障害化し、質の高いサービスを提供する研究を行っている。

### CDNの概要



- ① オリジン・サーバはSoR1を經由して、www.abc.comを配信
- ② SoRが互いにレプリカ・サーバの情報を有し、最適なレプリカ・サーバを判断
- ③ 最適なレプリカ・サーバにwww.abc.comのコンテンツをコピー
- ④ ユーザ4がwww.abc.comの閲覧を要求し、その要求をSoR4が取得
- ⑤ SoR4が要求内容をDPIにより判断し、要求先をより近いレプリカ・サーバに変更
- ⑥ レプリカ・サーバ4からユーザ4へ要求内容を送信

### 応答時間のモデリング

- MATLABによるシミュレーション手法の開発
- 任意のユーザにおいて、想定される応答時間を待ち行列としてモデリングし予想する

Let  $R$  be the response time given by:

$$R = E[T] + E[D]$$

Where Processing time  $E[T]$ , delay  $E[D]$

$$E[T] = \frac{1}{\Gamma} \left( \frac{L}{X^{SoR} - L} + \frac{AL}{X^{SS} - AL} \right)$$

$$E[D] = D^{SoR} + AD^{SS}$$

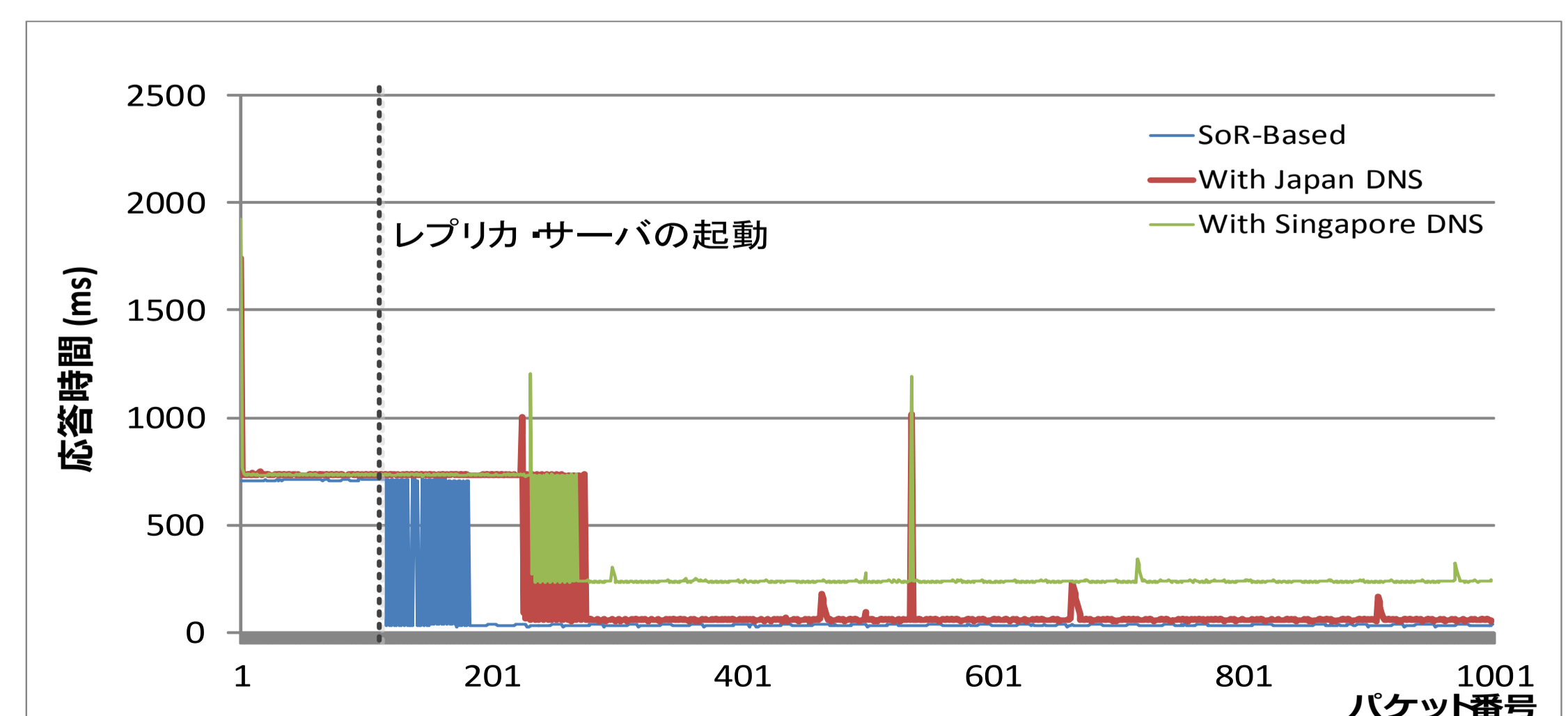
### PlanetLabへの実装

- 応答時間モデルをPlanetLabへ実装し、世界規模ネットワークにおける振る舞いを評価



### 応答時間の評価

- DNSによる経路選択の場合、メイン・サーバとコネクションを確立するまで1600~1800ms要する
- SoRによる経路選択の場合、約800msでメイン・サーバとのコネクションを確立可能
- コネクションの確率に要する時間を50%削減



研究者名

慶應義塾大学理工学部 西研究室





# サービス指向ルータ

## 情報抽出とサービスへの応用

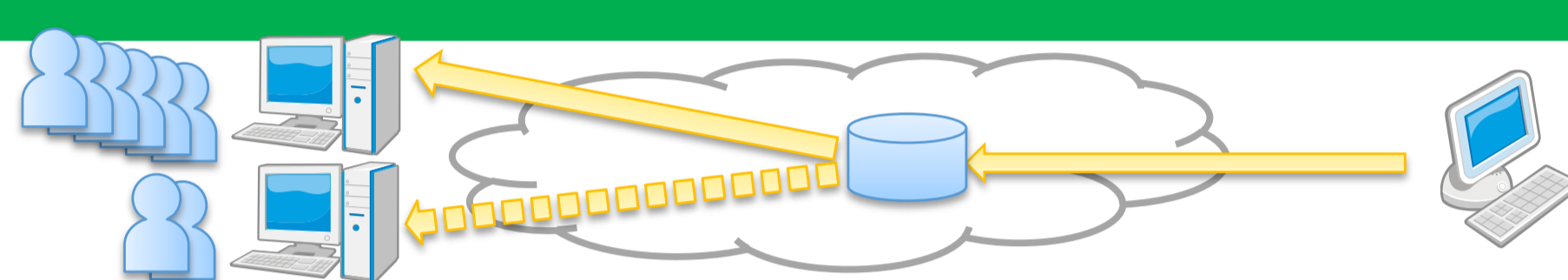
### サービス指向ルータ

目的: ルータを通過する情報の「中身」から必要な部分をデータベースに蓄え、共通APIを通して蓄えた情報をサービスとして利用者に提供する次世代インフラストラクチャ(情報オープンプラットフォーム)の中心デバイス。プライバシーに配慮しつつ、従来よりも高度かつ利用者に有益なサービスを展開し、現在の、また未来の様々な問題に対応する。

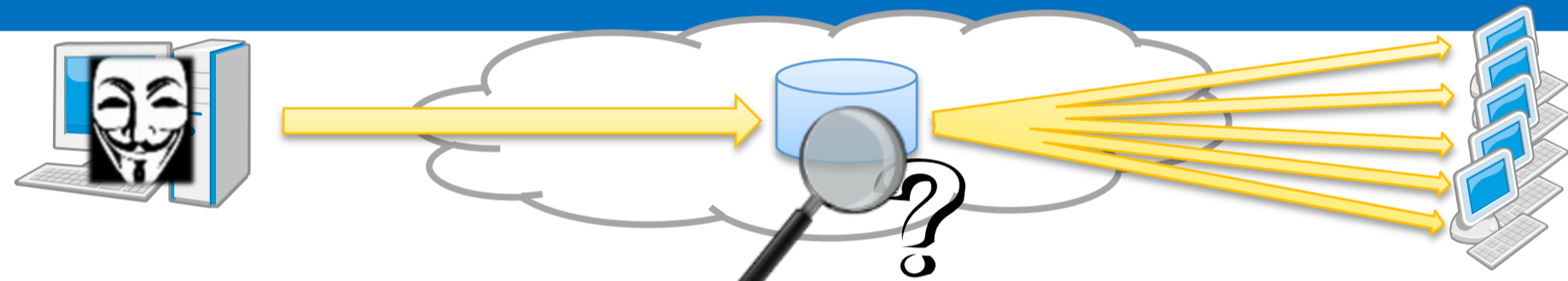
インターネットが攻撃耐性を備えて小型端末をサポート



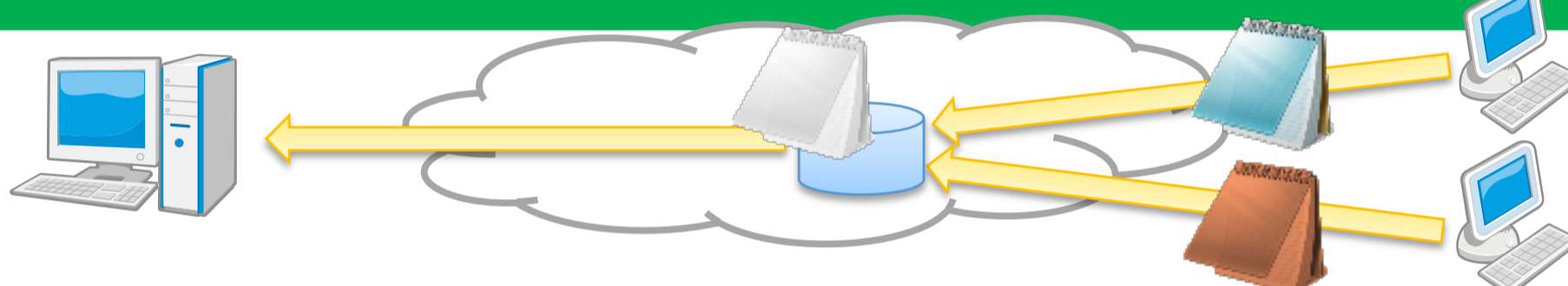
ユーザ行動に基づくレコメンデーションで新しいサービスを提供



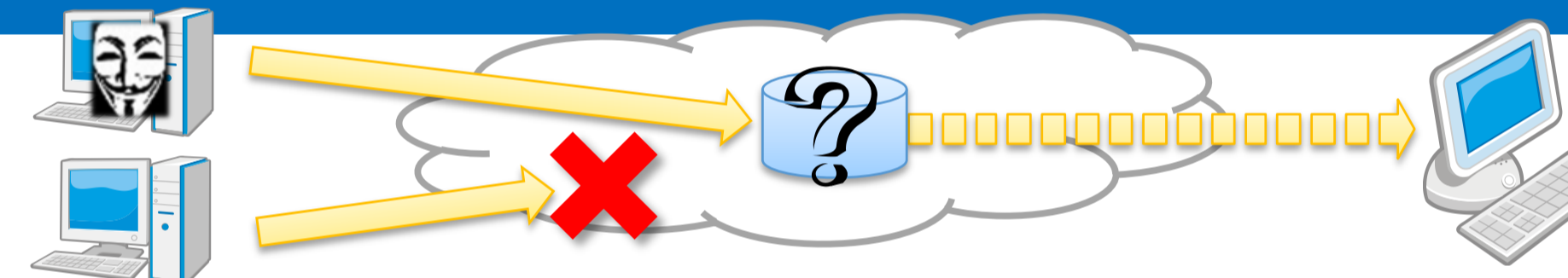
インターネットがフィッシング詐欺を解決、排除



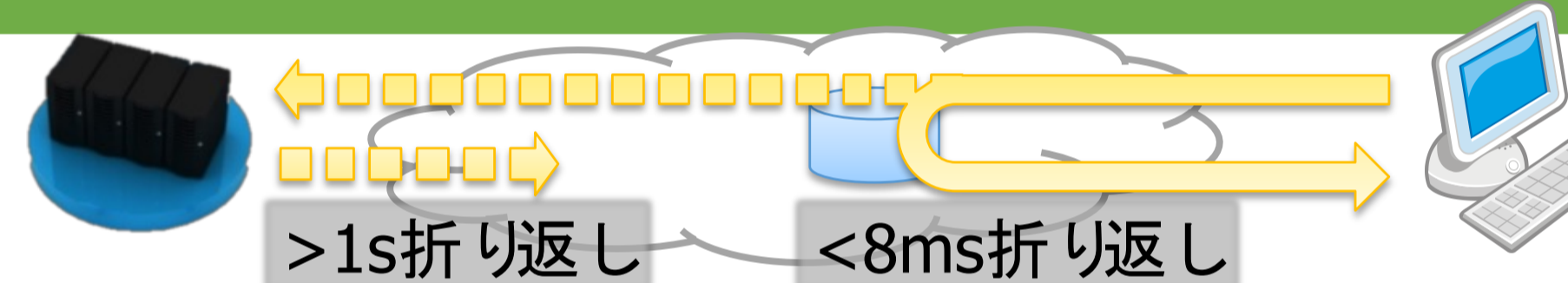
サービス授受において個人情報の詳細を消去し参入を容易化



挙動不審ユーザを監視することで、事前にアタックを排除



SG・EV-ADなど新インフラ通信・計算のクラウド集中問題を解決



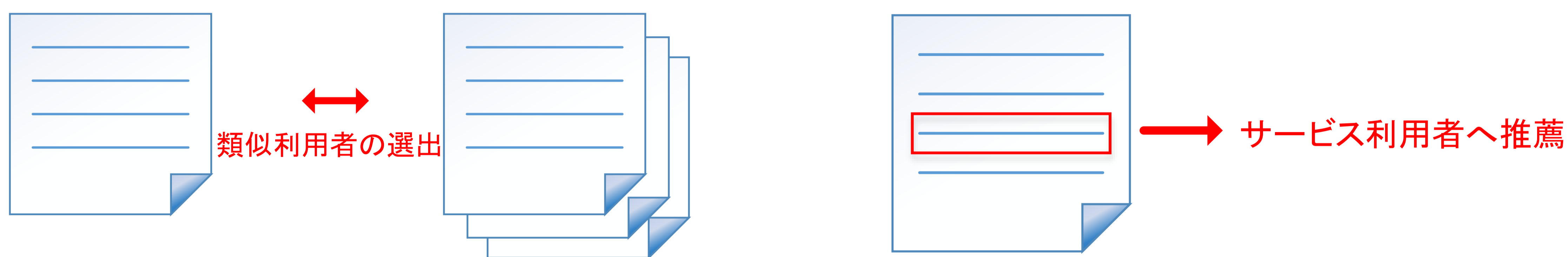
### ウェブページリコメンデーションサービス

サービス指向ルータのトラフィック解析により取得した利用者別閲覧履歴を利用し、サービス利用者に対しウェブページのリコメンデーションをおこなう。

サービス利用者の閲覧履歴

他の利用者の閲覧履歴

類似利用者の閲覧履歴



### Webページの滞在時間に基づくランキングを利用したWeb検索

平均滞在時間をルータで算出し、「現在、平均滞在時間の長いページは重要な情報がある」と考え、通常のランキング情報に加味して、Web検索結果に反映させる。

### 利用者のサーバ利用履歴に基づくロードバランサ

利用者の過去のサーバ利用負荷を学習、その結果から将来のサーバ利用負荷を予測しつつ負荷分散を行うコンテンツベースロードバランサを実現する。

研究者名

慶應義塾大学理工学部 西研究室

