Clock Skew Based Client Device Identification in Cloud Environments

Wei-Chung Teng
Dept. of Computer Science & Info. Eng.
National Taiwan University of Sci. & Tech.

Wei-Chung Teng

on behalf of Assoc. Prof. Yuh-Jye Lee

CLOUD SERVICE DEFENSE-IN-DEPTH SECURITY TECHNOLOGY RESEARCH AND DEVELOPMENT

Project Structure

Main project director:
Assoc. Prof. Yuh-Jye Lee

Sub-project 1: Anomaly Detection Based on Cloud Clients Behavior P
Director: A
Prof. Yuh-J

Sub-project 2: The Study of Software Testing in Cloud Service
Director
Hahn-M

Sub-project 3: Cloud application service security analysis mechanism based on intrusion e
analysis pl
Director: A
Prof. Hsin

Sub-project 4: Cloud application service communication security and infrastructur
protection
Director: Pro
Ren Jeng

科学技術振興機構
Japan Science and Technology Agency

行政院國家科學委員會
National Science Council

Project Organization

Cloud Service Defense-in-depth Security Technology

Prevention	Detection	Analysis
2. The Study of Software Testing in Cloud Service	1. Anomaly Detection Based on Could Clients Behavior Profiling	3. Cloud service security event analysis
Cloud service weakness analysis and detection	Anonymous user behavior profiling and prediction	Cloud malicious web application detection
Large scale cloud service penetration test	Data mining based analysis platform	Cloud malicious service scene and event analysis
Cloud service feedback oriented detection techniques	Online anonymous behavior detection mechanism	Sequence extraction and behavior similarity analysis

- Key technology of infrastructure security, data security, identification and access control
- Secured application example

4. Cloud application service communication security and infrastructure protection

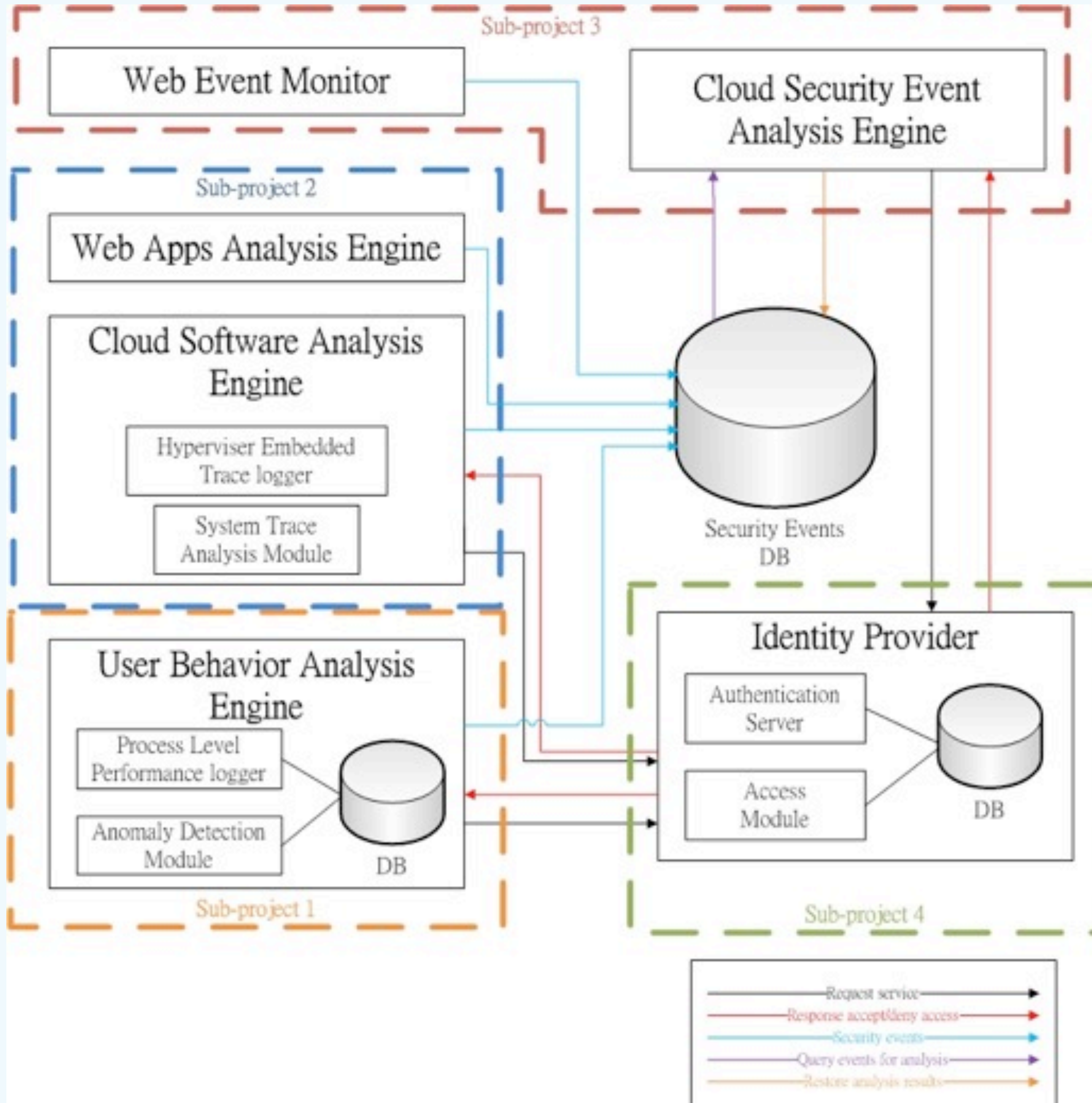
Key Features

- Image-based authentication & re-authentication
  - Protect users from automatic programming attack
  - Protect users from account hi-jacking with user behavior anomaly detection
- User behavior anomaly detection
  - System usage continuously monitoring for both hypervisor & VMs
  - Collect process-level information for build user profiles
  - Detect anomalous behaviors which differ from user profiles

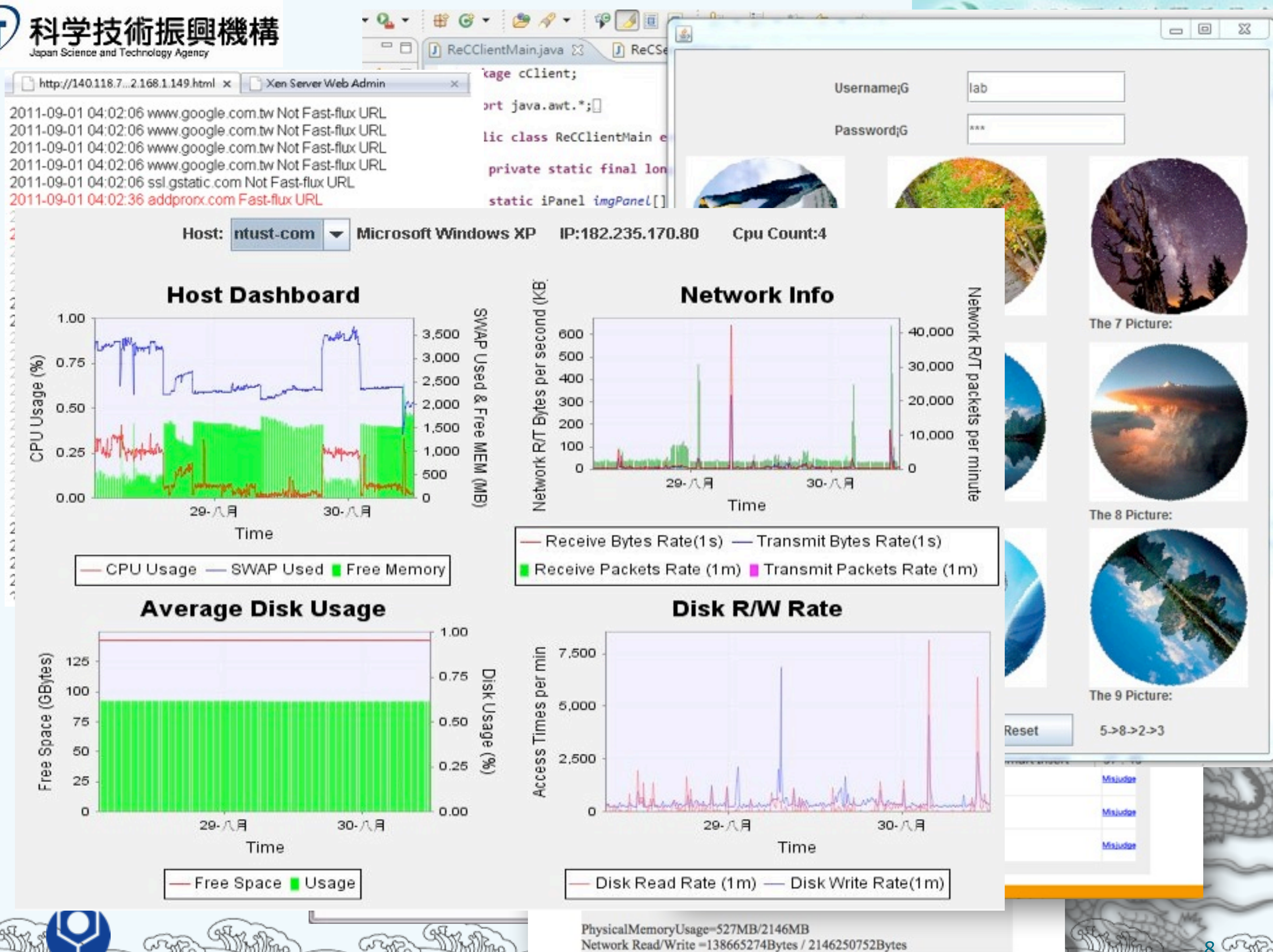
Key Features (cont.)

- Fast-flux detection
  - Detect fast-flux URLs from all the http requests in the cloud
  - Protect cloud users from phishing & malware delivery attacks
- Malicious Software Analysis
  - Automatically build sandbox in hypervisor for analyzing software uploaded in the cloud
  - Protect cloud users from downloading malware
  - Prevent abusing cloud service as a malware spreading platform
- Graphic based security event correlation analysis
  - Collect security events from different sensors in the cloud
  - Automatically generate correlation graphs for analyzing

System Framework



科学技術振興機構
Japan Science and Technology Agency



Publications

- **CAPTCHA**
  - Albert b. Jeng, De-Fan Tseng, Chein-Chen Tseng ,"An Enhanced Image Recognition CAPTCHA Applicable to Cloud Computing Authentication," 2nd Annual International Conference on Business Intelligence and Data Warehousing (BIDW 2011), Singapore,2011.
- **Re-authentication**
  - Szu-Yu Lin, Te-En Wei, Hahn-Ming Lee, Albert B. Jeng, "A Novel Approach For Re-Authentication Protocol Using Personalized Information", ICMLC2012, China.
- **Anomaly Detection**
  - Yuh-Jye Lee, Yi-Ren Yeh and Yu-Chiang Frank Wang. "Anomaly Detection via Online Over-Sampling Principal Component Analysis", IEEE Transactions on Knowledge and Data Engineering (TKDE), (To appear).
  - **Ding-Jie Huang, Kai-Ting Yang, Chien-Chun Ni, Wei-Chung Teng*, Tien-Ruey Hsiang, and Yuh-Jye Lee "Clock Skew Based Client Device Identification in Cloud Environments," The 26th IEEE International Conference on Advanced Information Networking and Applications (IEEE AINA-2012), Fukuoka, Japan, March 26-29, 2012.**
- **Fast-flux detection**
  - Horng-Tzer Wang, Ching-Hao Mao, Kuo-Ping Wu and Hahn-Ming Lee, "Real-time Fast-flux Identification via Localized Spatial Geolocation Detection," IEEE Signature Conference on Computers, Software, and Applications (COMPSAC 2012), Izmir, Turkey, July 16-20, 2012.

Publications

- **Security events analysis**
  - Chien-Chung Chang, Hsing-Kuo Pao, and Yuh-Jye Lee. "An RSVM Based Two-teachers-one-student Semi-supervised Learning Algorithm", Neural Networks, Vol. 25: pp. 57-69, Jan., 2012. [SCI]
  - Hsing-Kuo Pao, Ching-Hao Mao, Hahn-Ming Lee, Chi-Dong Chen, and Christos Faloutsos. "An Intrinsic Graphical Signature Based on Alert Correlation Analysis for Intrusion Detection", Journal of Information Science and Engineering, Vol. 28, no. 2: pp. 243-262, March, 2012. [SCI]
  - Hsing-Kuo Pao, Junaidillah Fadlil, Hong-Yi Lin, and Kuan-Ta Chen. "Trajectory Analysis for User Verification and Recognition", Knowledge-Based Systems, (accepted). [SCI]
  - Hsing-Kuo Pao, Yan-Lin Chou, Yuh-Jye Lee. "Malicious URL Detection based on Kolmogorov Complexity Estimation", 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WI-IAT 2012), Macau, Macau, December 2012.
  - Danai Koutra, Tai-You Ke, U Kang, Duen Horng Polo Chau, Hsing-Kuo Pao, and Christos Faloutsos. "Unifying Guilt-by-Association Approaches: Theorems and Fast Algorithms", European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), Athens, Greece, Sep. 2011.

CLOCK SKEW BASED CLIENT DEVICE IDENTIFICATION IN CLOUD ENVIRONMENTS

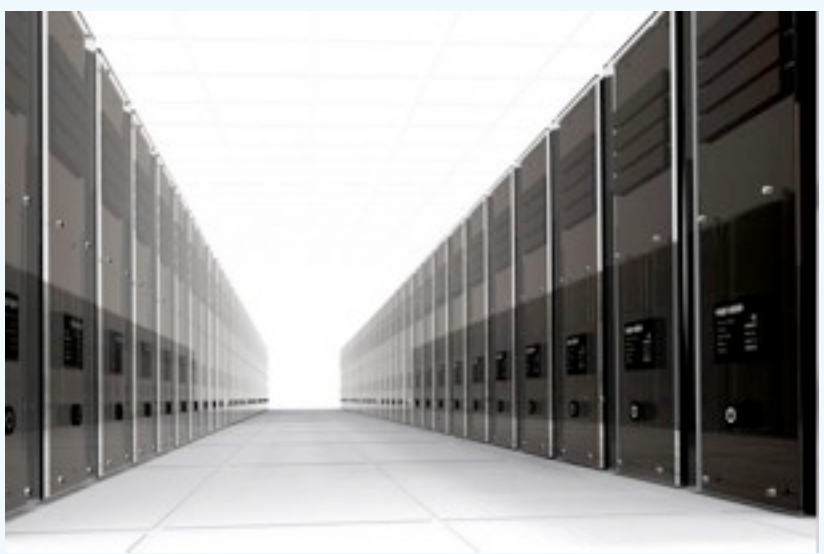
Why client device Identification?

Personal devices of private use

cloud services

two-factor authentication

clock skew as identity



account & password

Introduction of Clock Skew

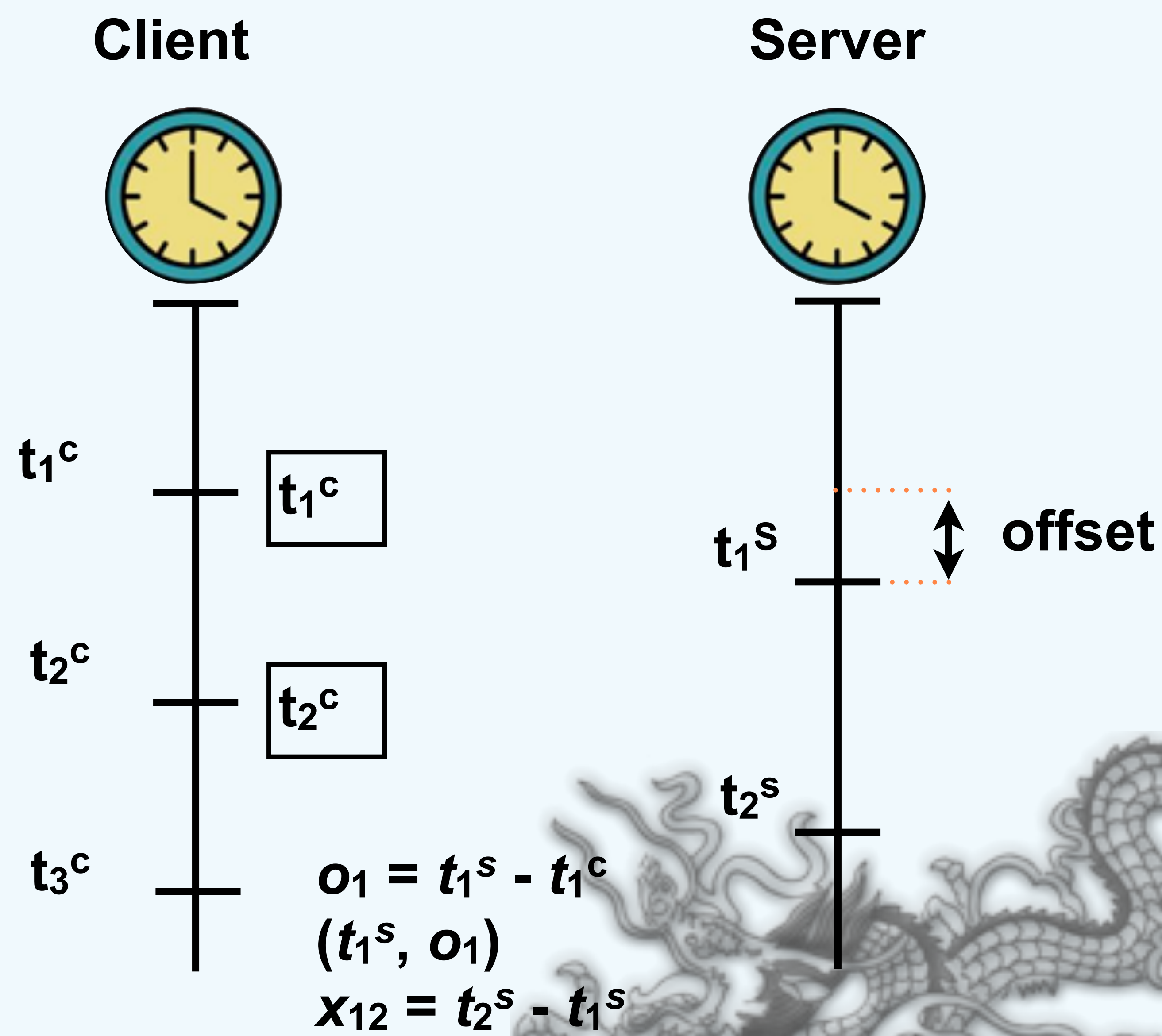
- Every client device has a clock (crystal oscillator), and Quartz crystal in every device works in slightly different frequency.
- Clock skew is stable under normal temperature.
- Basically, every clock skew measured remotely differs with others at 10^{-6} second precision. (Kohno, 2005)
- It is easy to alter clock skew, but hard to fake one if the target device change its time sync period from time to time.

Why Using Clock Skew as Identity?

- Clock skew is the relative speed of time passing, and both source and target device can be affected by temperature, but servers inside cloud are always maintained at stable temperature.
- Clock skews are measured in background, so users are unaware of the two-factor authentication going on.
  - legal users don't bother to pass the 2nd factor auth.

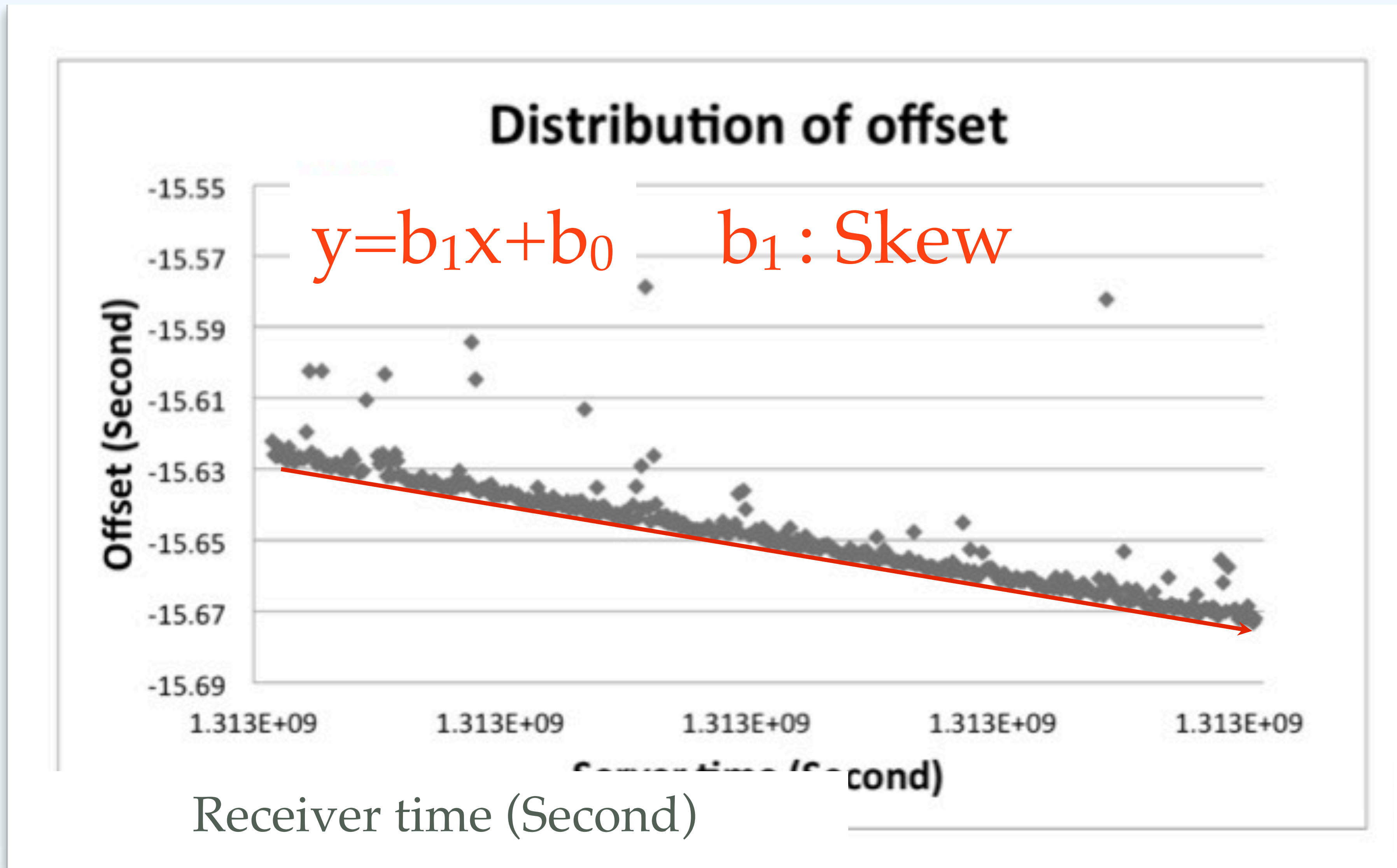
Clock skew measurement

- Let $C_x(t)$ be the time reported by the clock of device x. Let C_c and C_s be the clocks of client and server respectively.
- **Offset**: The difference between the time reported by C_c and C_s.
- **Frequency**: The rate at which the clock ticks. The frequency of C_c at time t is $C_c'(t)$.
- **Skew** (δ): The difference in the frequencies of two clocks, e.g., the skew of C_c relative to C_s at time t is $\delta(t) = C_c'(t) - C_s'(t)$.



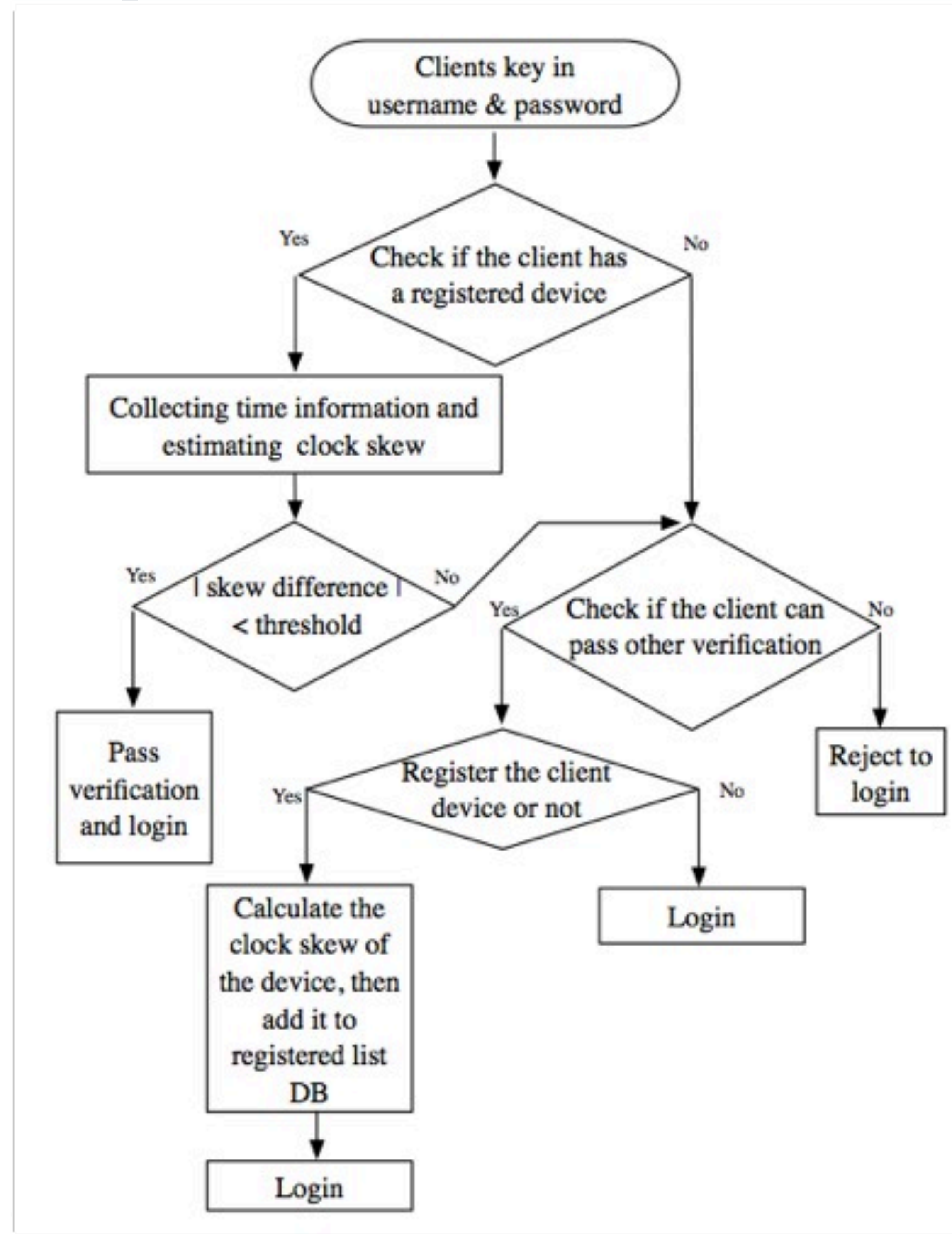
Measured Offsets vs. Clock Skews

The value of offset fluctuates is considered due to transmission jitter. The bottom line should be the closest estimation to the real skew.



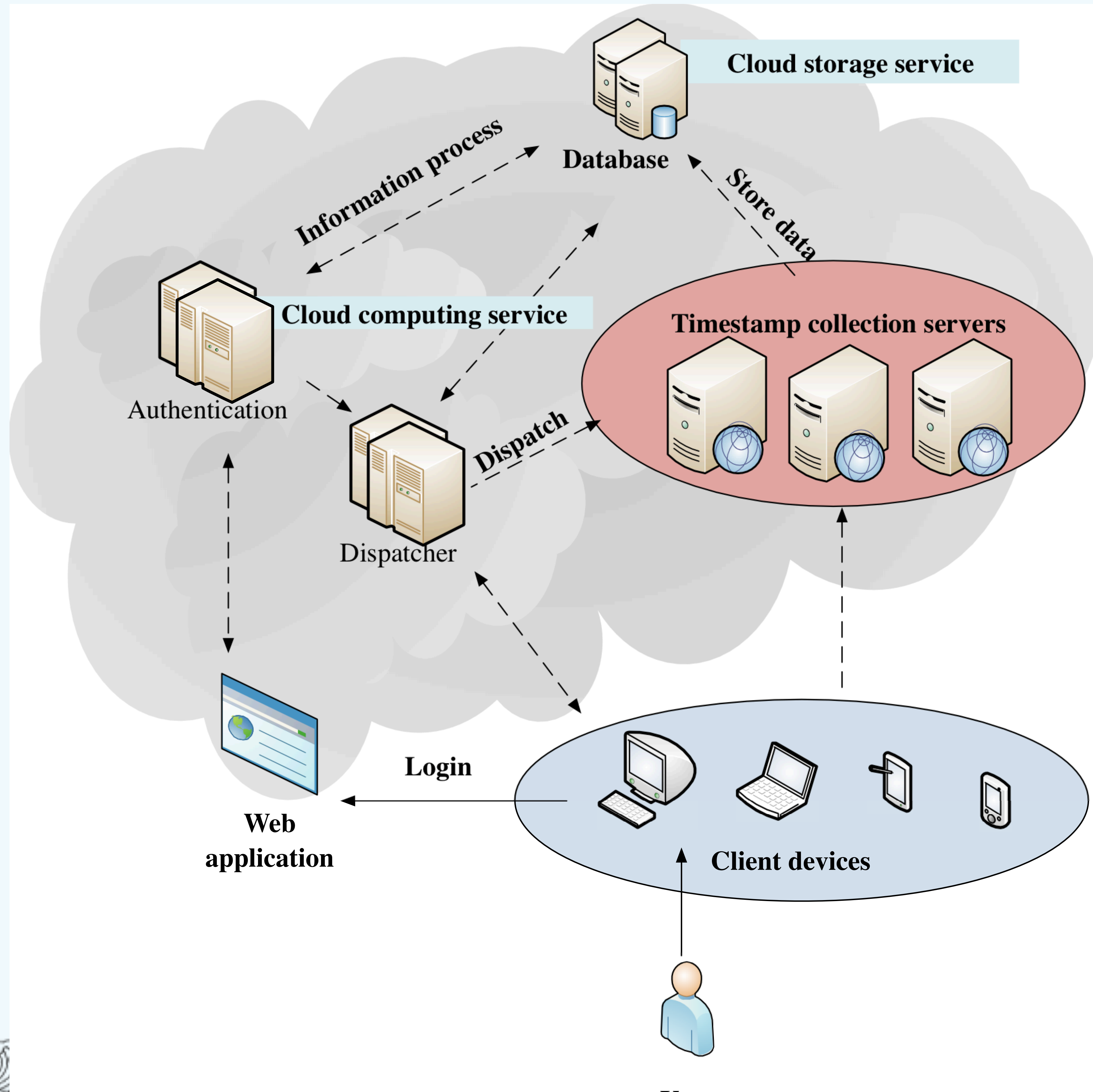
Flowchart of clock skew based host identification system

- Login procedure
  1.Register device
  2.Clock skew measurement
  3.pass verification or call other method



Scenario of time information collection

- collected info.
  - client time
  - server time
  - IP address

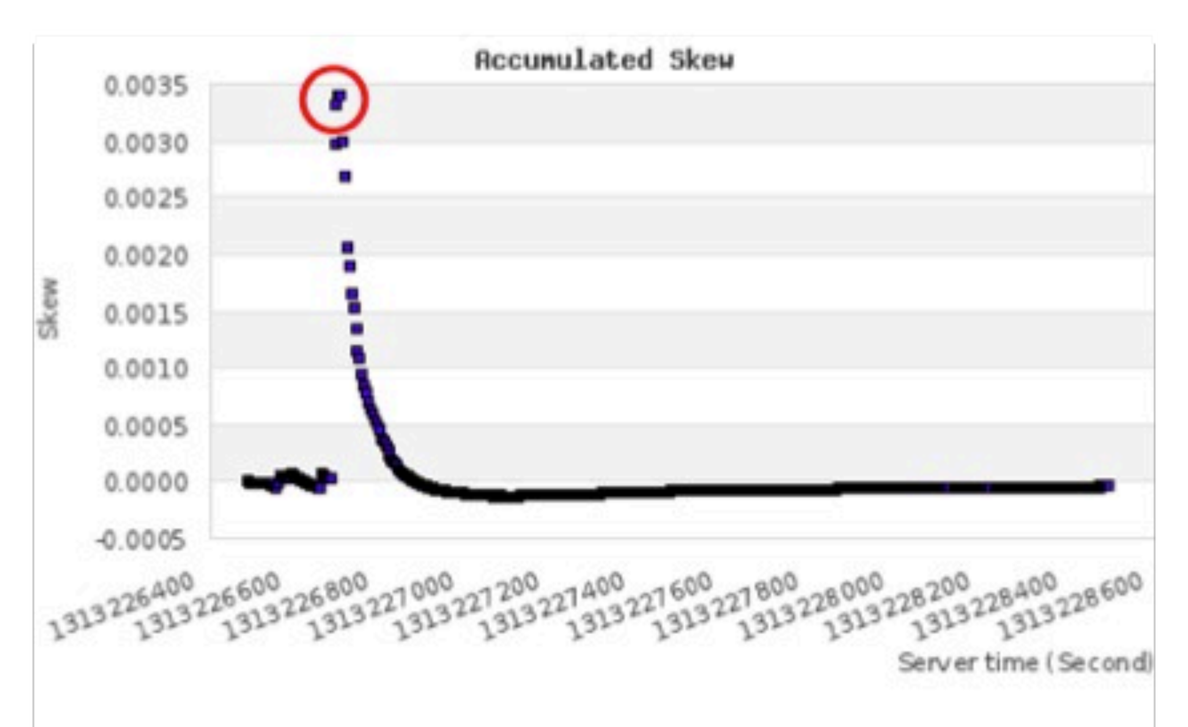


Challenges and Tools

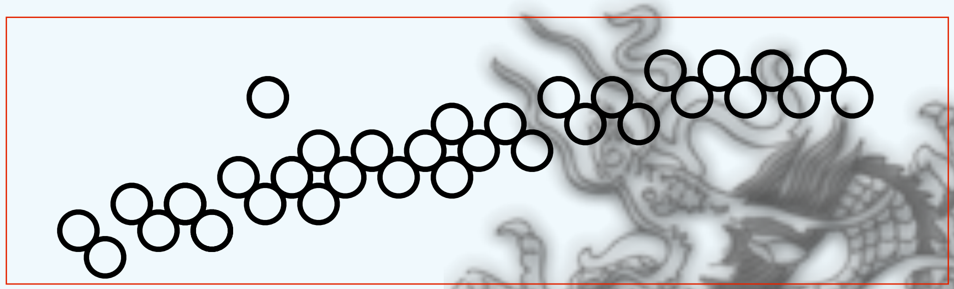
- Problems when I want a quick-n-dirty skew
  - spikes: temporary high offsets due to e.g. network congestion
  - outliers: happens occasional (network congestion, time sync etc)
  - jump points: change base station during mobile communication sessions
- Methods
  - Linear regression
    - Sliding-Windows Skew with Lower-Bound Filter
    - Accumulated Sliding-Windows Skew with Lower-Bound Filter
  - Quick Piecewise Minimum Algorithm
  - Jump point detection

Accumulated Skew

- For accumulated skew, while packets sent from the client are received by the server, the server computes the estimated skew immediately. The estimated skew can be represented as LR(N_{1i}), while receiving i^{th} request from the client.



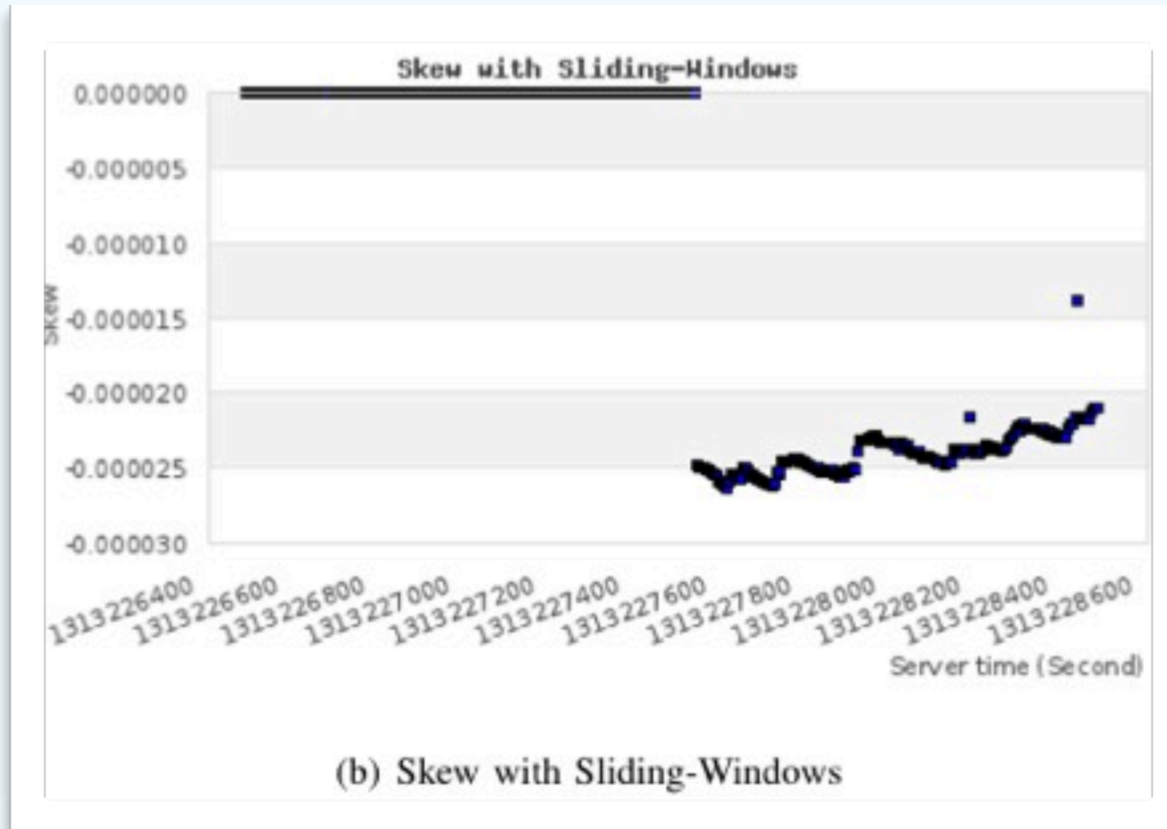
(a) Accumulated Skew



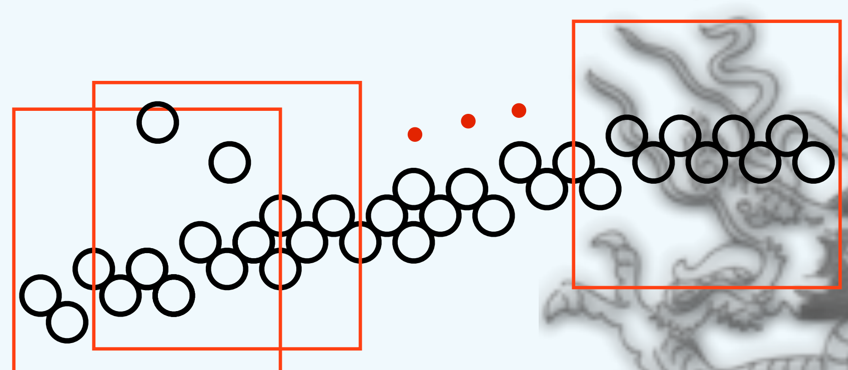
-

Skew with Sliding-Windows

- A sliding-windows computation that only sampling part of the data set can prevent the effect caused by previous fluctuated data.
- For sampling windows with size w, the sliding-windows skew LR(N_{ij}) must satisfy $j - i = w$.

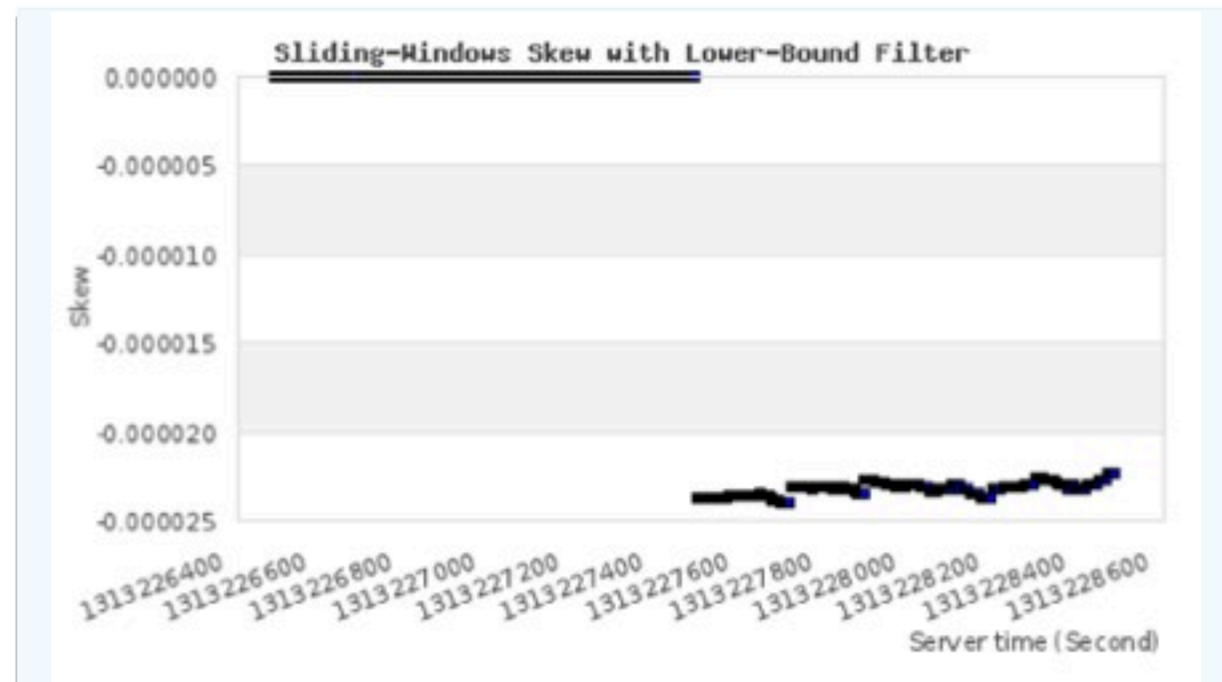


(b) Skew with Sliding-Windows

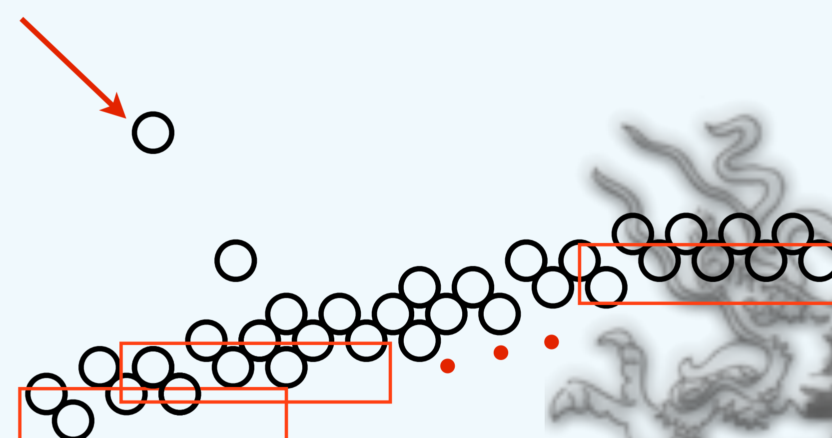


Sliding-Window Skew with Lower-Bound Filter

- To disassemble the effect caused by outliers, the most effective method is to filter them out.
- The local minimum offset is picked for every m packets in each sliding window w.
  - the amount of sampling data for skew estimation is reduced to $\lfloor w/m \rfloor$.

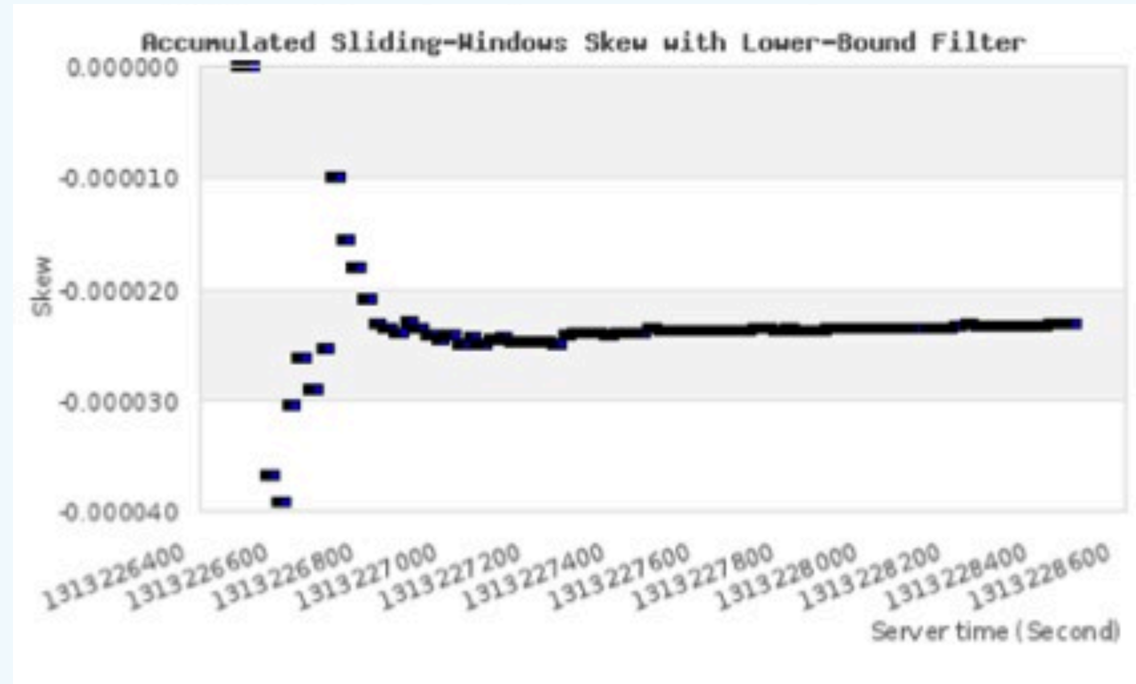


(c) Sliding-Windows Skew with Lower-Bound Filter



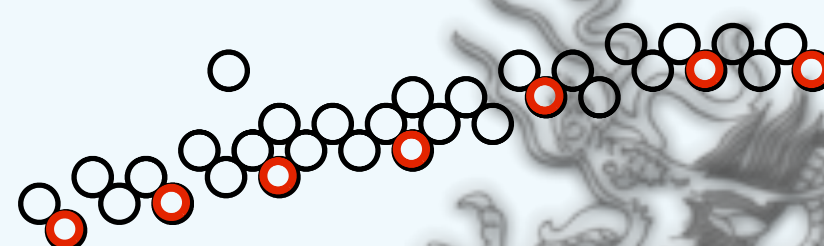
Accumulated Sliding-Windows Skew with Lower-Bound Filter

- Since the local minimum offset is useful to find the lower-bound skew, we further calculate the accumulated skews with these local minimum dataset.



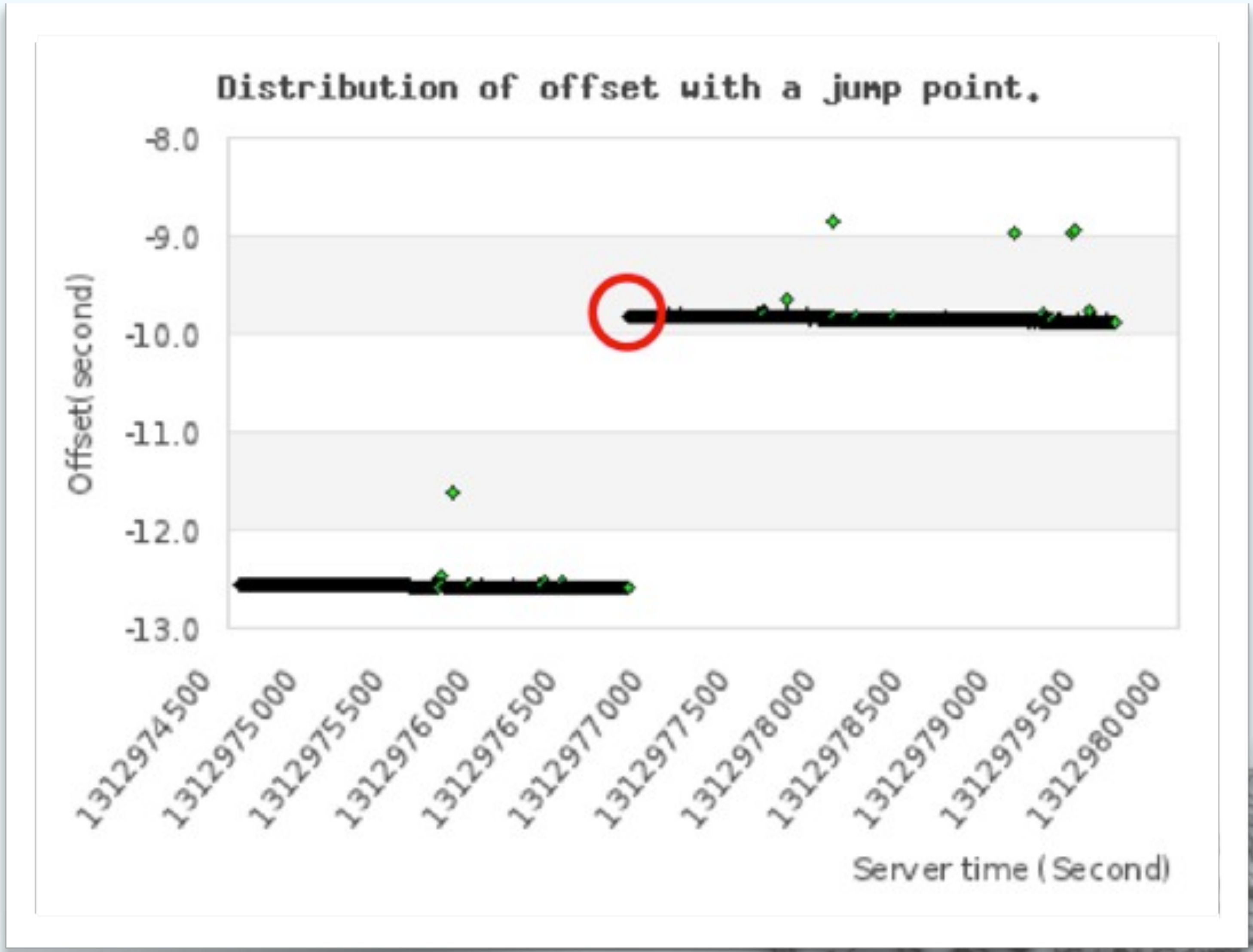
(d) Accumulated Sliding-Windows Skew with Lower-Bound Filter

- We find that this method can both reduce the effect of huge network delay and calculate an approximate skew rapidly within 20 packets.

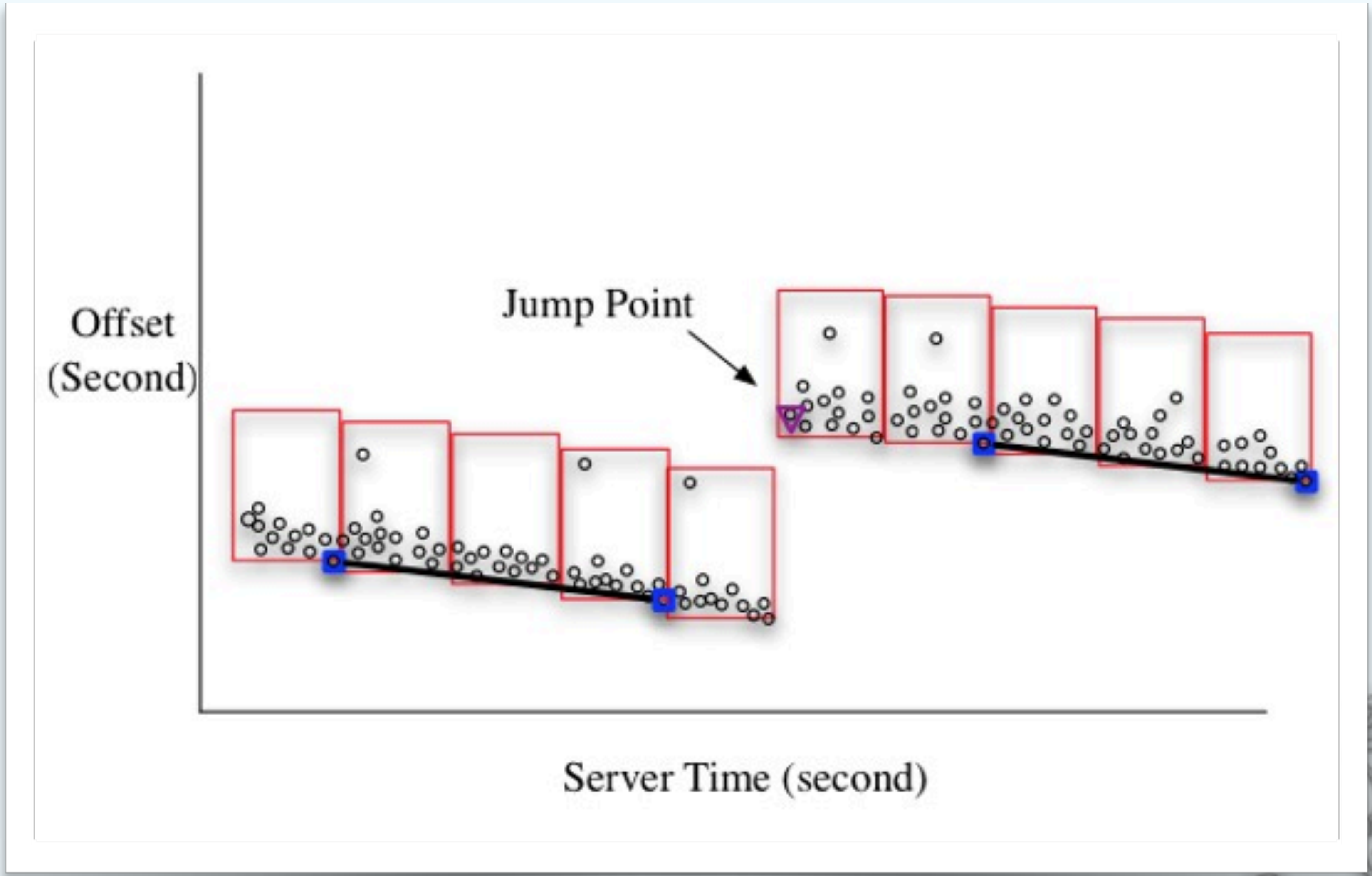


Jump point detection & handling

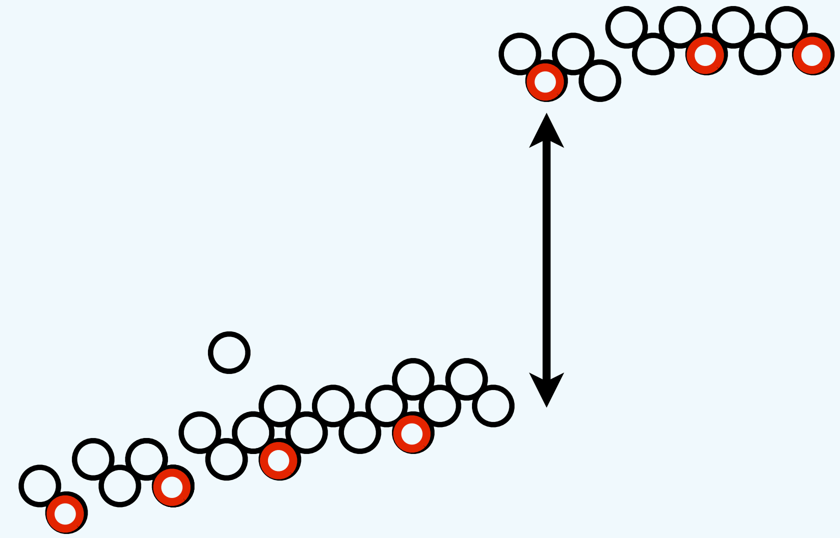
- A jump point of offset occurs if the client is performing time synchronization with a time server or roaming between different network providers.



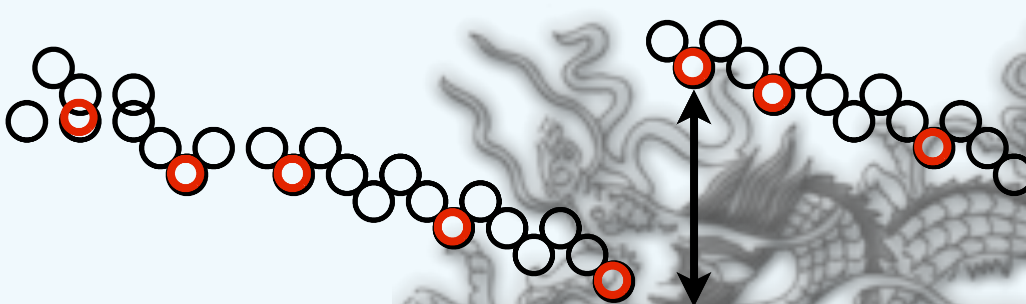
Detection of jump point between two segments



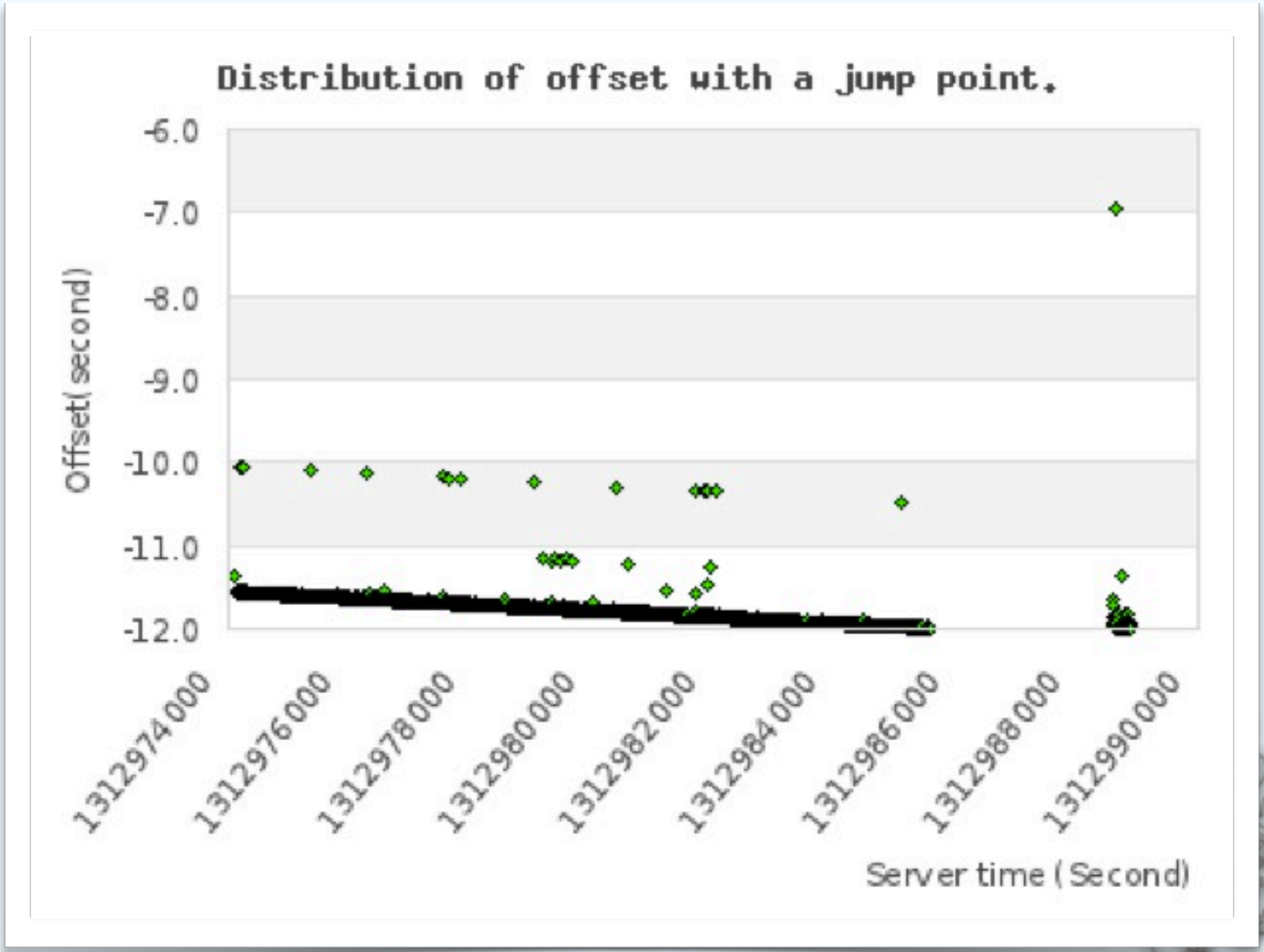
Jump point detection algorithm



1) Pick the local minimum offsets for every p packets.
2) Compute the *diff* between every pairs of contiguous local minimum offsets.
3) Following detection process is divided into two classes:
   a) If derived *diffs* are all positive or all negative:
      - Denote the median of derived *diffs* as $Med(diff)$.
      - If there exists a *diff* that $diff > k \cdot Med(diff)$, a jump point exists inside these p packets.
   b) If only part of derived *diffs* are positive:
      - If positive *diffs* are followed by one negative *diff* at x, this x is the jump point.
      - Similarly, negative *diffs* followed by one positive *diff* is processed vice versa.



Another form of jump point: time gap



EXPERIMENT RESULTS

The estimated skews for the same device under different environments

- The estimated skews vary from -21.08 ppm to -23.71 ppm. However, skews of the same network type differ no more than 1.31 ppm.
- Notice that skews of virtual machine change every time the virtual machine reboots.

Network type	Skew estimation	Packets	IP amount
LAN	-21.91 ppm	1001	1
	-23.24 ppm	207	1
	-22.74 ppm	13322	1
ADSL	-21.48 ppm	5837	1
	-21.08 ppm	1400	1
3G	-23.24 ppm	951	1
	-23.71 ppm	1027	1
Wi-Fi	-21.79 ppm	9810	1
	-23.06 ppm	1470	1
Tor	-22.53 ppm	15007	55
	-23.22 ppm	12922	57
	-22.88 ppm	24120	108
VM	-113.19 ppm	868	1
	-114.22 ppm	1001	1
	-6.40 ppm	1001	1
	-6.83 ppm	890	1

Conclusions

- A web based skew measuring system and related technologies are introduced. Even the precision of timestamp is millisecond, limited by Javascript, the estimated clock skew is able to reach microsecond precision after at least 1000 seconds.
- According to experiment results, clock skew is a potential candidate that can be used alongside with other properties to serve as fingerprints of physical devices.
- skew estimation should be able to improved further by linear programming method and/or with more precise timestamps.

THANK YOU FOR YOUR ATTENTION

Q&A