Capacity Planning and Headroom Analysis for Taming Database Replication Latency

- Experiences with LinkedIn Internet Traffic

Zhenyun Zhuang, Haricharan Ramachandra, Cuong Tran, Subbu Subramaniam, Chavdar Botev, Chaoyue Xiong, Badri Sridharan

zhenyun@gmail.com

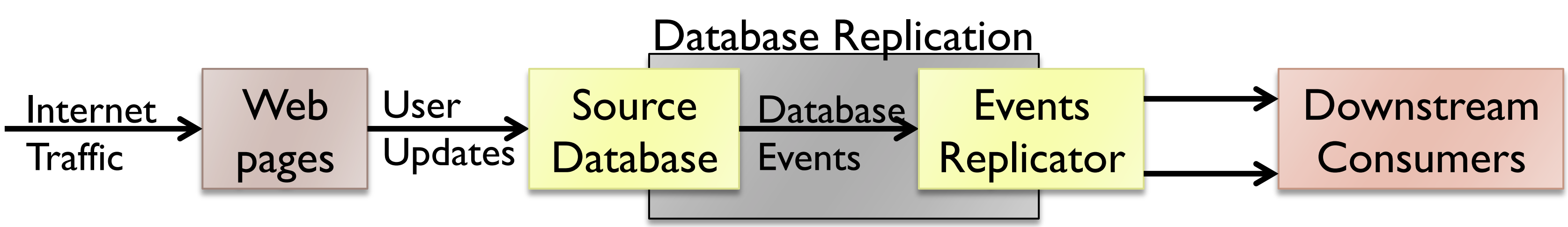
LinkedIn Corp.

Outlines

- Introduction
- Problem definition
- Observations of LinkedIn Internet traffic
- Solutions
- Evaluation

Introduction - Database replication

- Why replicating database events?
  - Source database protection
  - Inter-datacenter synchronization
- Dataflow
  - Source database (Espresso database)
  - Database replication component (Databus)
  - Clients (Downstream products)



Introduction – Capacity planning

- Importance
  - Determine SLA
  - Capacity planning (e.g., cluster size, replication capacity)
  - Reduce operation cost
- Questions in capacity planning
  - Future traffic rate forecasting
  - Replication latency prediction
  - Replication capacity determination
  - Replication headroom determination
  - SLA determination

Problem Definition - Terminology

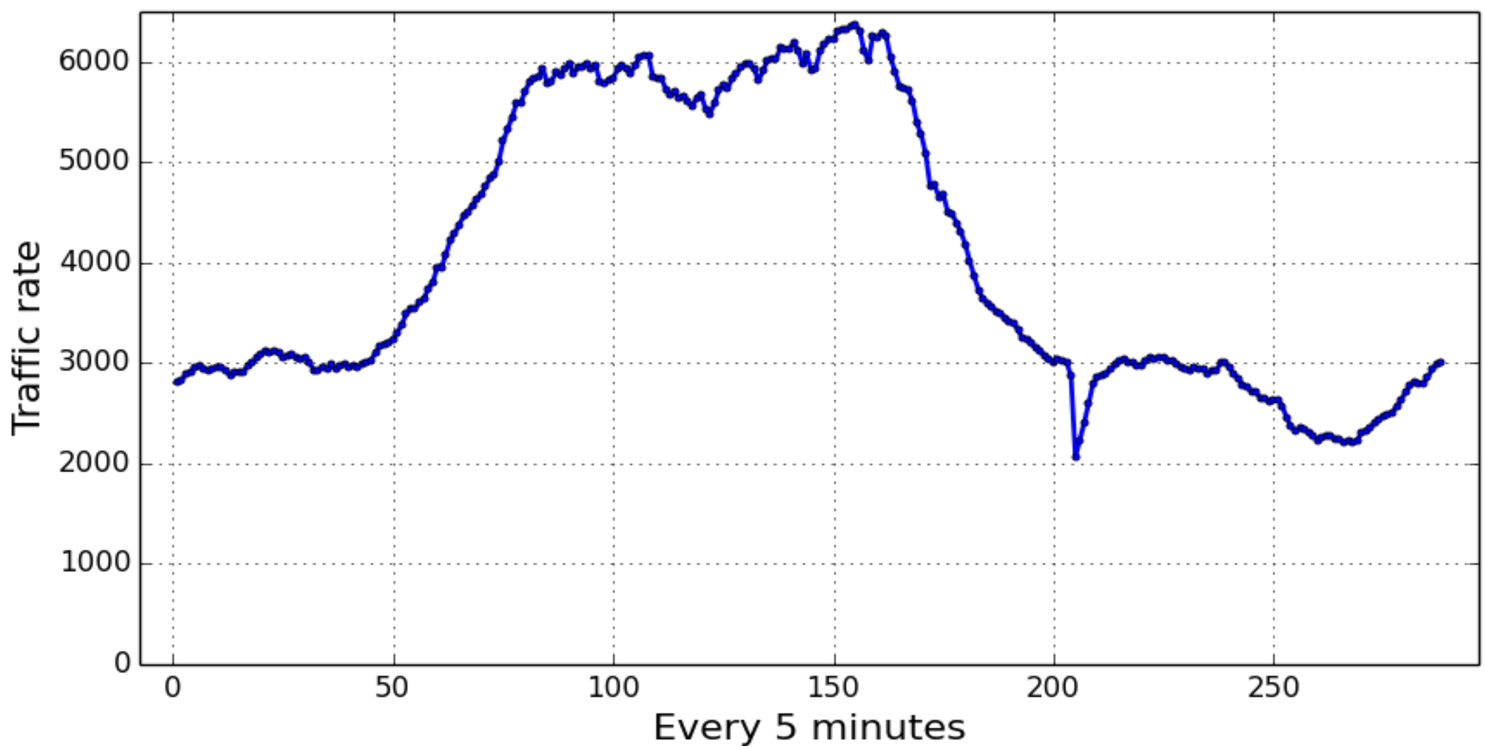
- Replication latency
  - Time difference between:
    - The event is inserted into source database
    - The event (after replication) is ready for downstream consumption
- Replication SLA
  - Service level agreements
  - E.g., Largest replication latency < 60 seconds
- Incoming traffic rate
  - Number of incoming web events per second
- Replication capacity
  - Number of events processed by replication component per second
  - Aka, Relay Capacity

Problem Definition

- Forecast future traffic rate
  - Given historical traffic rate of $T_{i,j}$, what is the future rate?
- Determine the replication latency
  - Given the traffic rate of $T_{i,j}$ and relay capacity of $R_{i,j}$, what is the replication latency $L_{i,j}$?
- Determine SLA
  - What is the largest replication latency? P99 value?
- Determine required replication capacity
  - Given SLA of L_{sla} and traffic rate of $T_{i,j}$, what is the required replay capacity of $R_{i,j}$?
- Determine replication headroom
  - Given L_{sla} and $R_{i,j}$, what is highest traffic rate $T_{i,j}$ it can sustain?
  - What is the expected data of d_k of that traffic rate?

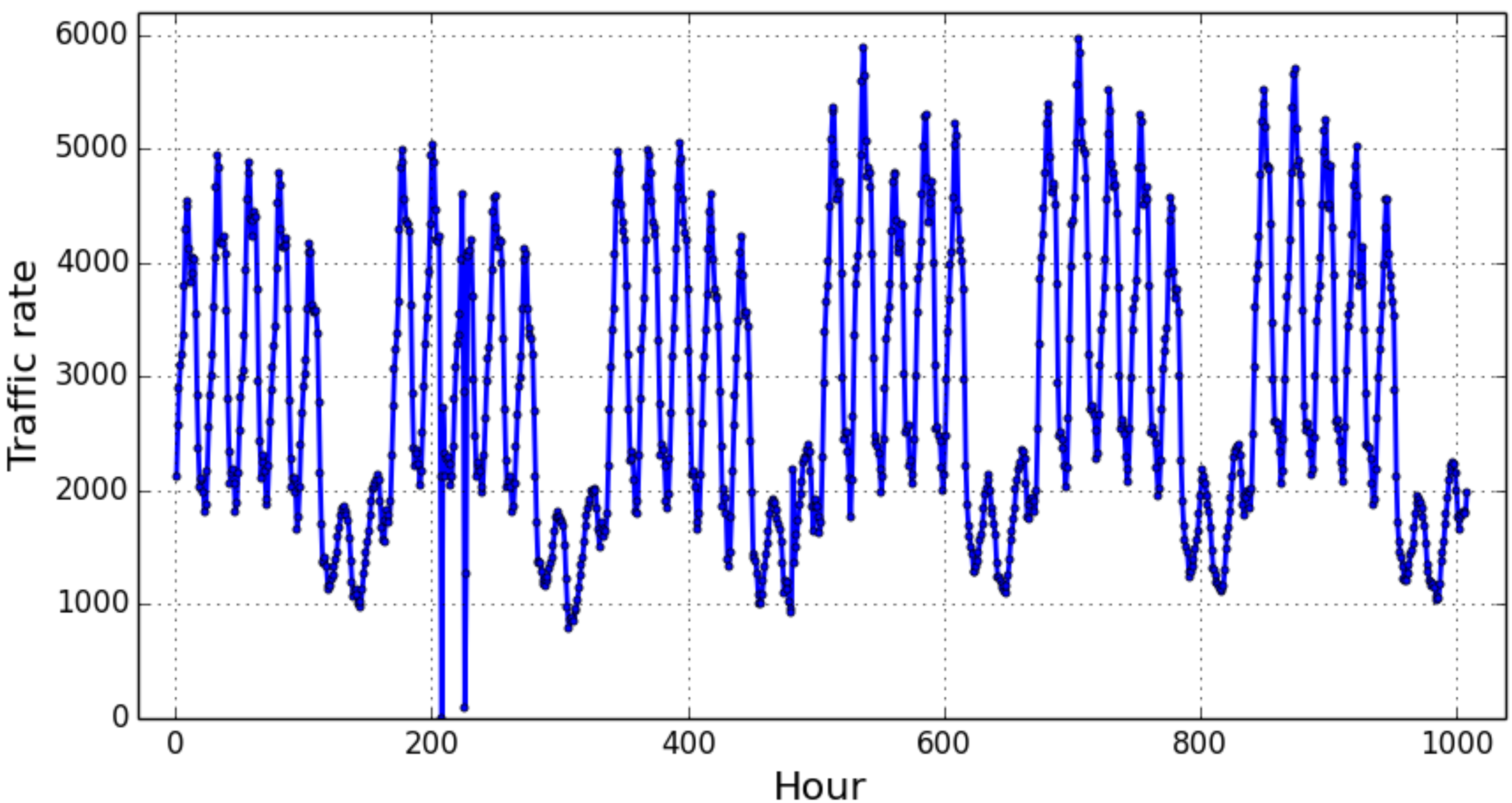
Observations of LinkedIn Internet traffic

- A weekday traffic across time
- Weekday vs weekend
- Traffic volume is growing



Observations of LinkedIn Internet traffic

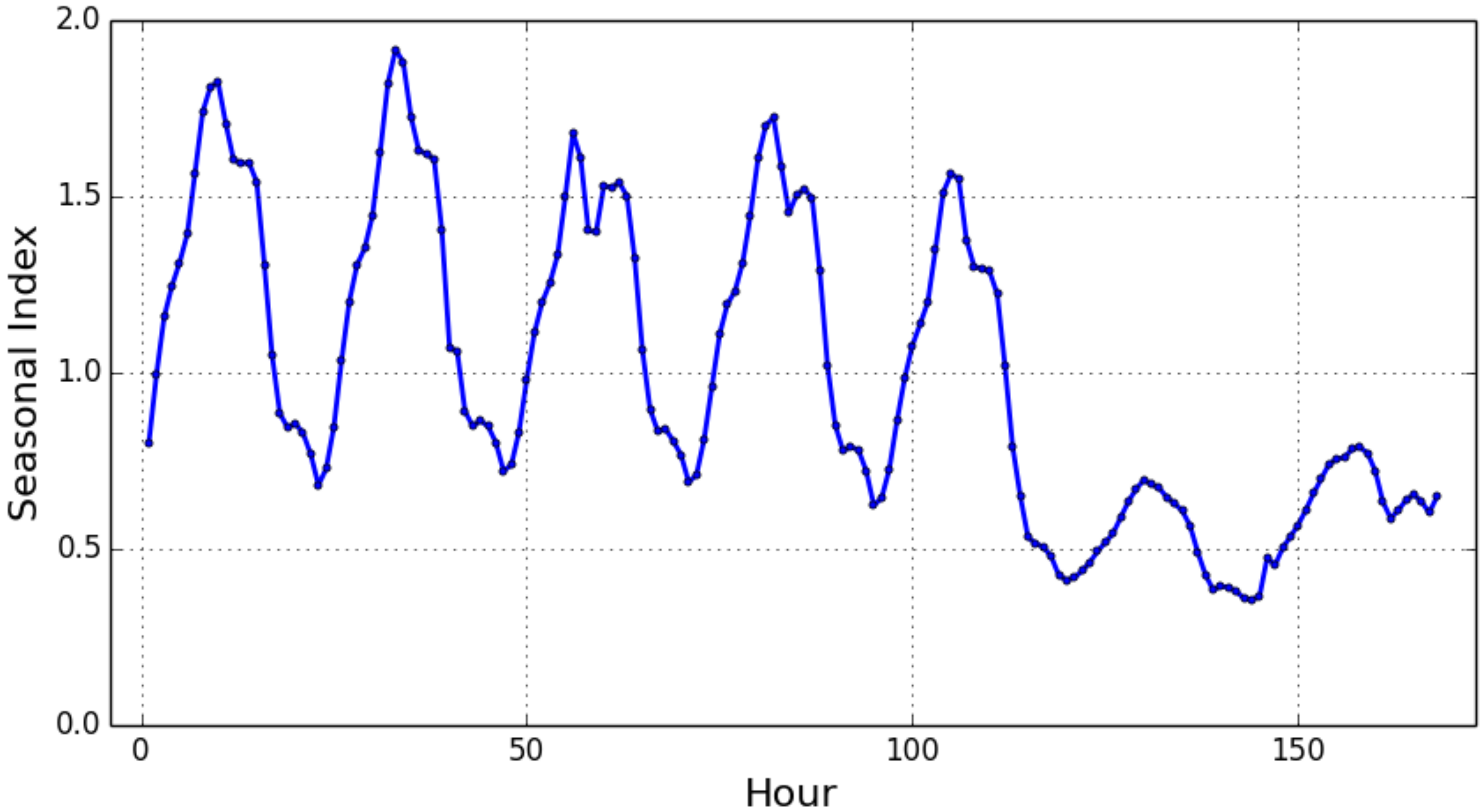
- Strong periodical patterns at day, week, month level



Design – Forecasting future traffic

- Two models
  - Time series model (ARIMA)
  - Regression analysis model
- Challenges
  - Goal: forecast per-hour (or per-minute, per-second) rate
  - ARIMA: not suitable for long period seasonality (e.g., 168)
  - Regression analysis: works well on weekly (or monthly) traffic
- Two step approach
  - Forecasting future **Daily/weekly** traffic
    - Both ARIMA and Regression analysis
  - Converting daily/weekly traffic to **hourly** traffic
    - Seasonal index (hourly)

Design – Seasonal Index



Design – Forecasting with ARIMA

- ARIMA(p,d,q)
  - P=7, d=1, q=0
- Historical traffic is aggregated on a daily/weekly basis
  - E.g., 42 days or 6 weeks
- Forecasting into daily/weekly traffic
  - E.g., 21 days or 3 weeks
- Computing hourly seasonal index
  - Totally 168 values (for a week)
- Converting daily traffic to hourly traffic

Design – Forecasting with Regression Analysis

- Linear fitting
  - $Y = aW + b$
- Traffic is aggregated on a weekly basis
  - E.g., 6 weeks
- Forecasting into weekly traffic
  - E.g., 3 weeks
- Using hourly seasonal index
  - Totally 168 values (for a week)
- Converting weekly traffic to hourly traffic

Design – Predicting replication latency

- Iterating each hour of a day
  - Starting from the lowest traffic rate
  - If traffic rate > relay capacity: Accumulated latency
  - If traffic rate < relay capacity: Decreased latency

$L_{i,j} = predict(T_{i,j}, R_{i,j})$: // Latency in seconds.

// assuming no latency buildup in d_{i-1}, so $L_{i,0} = 0$.

for each hour of h_j, where $1 \leq j \leq 24$:

  if $T_{i,j} > R_{i,j}$: // increase latency buildup

    $L_{i,j} = L_{i,j-1} + \frac{3600(T_{i,j} - R_{i,j})}{R_{i,j}}$

  else: //decrease latency buildup if any

    $L_{i,j} = L_{i,j-1} - \frac{3600(R_{i,j} - T_{i,j})}{R_{i,j}}$

    if $L_{i,j} < 0$:

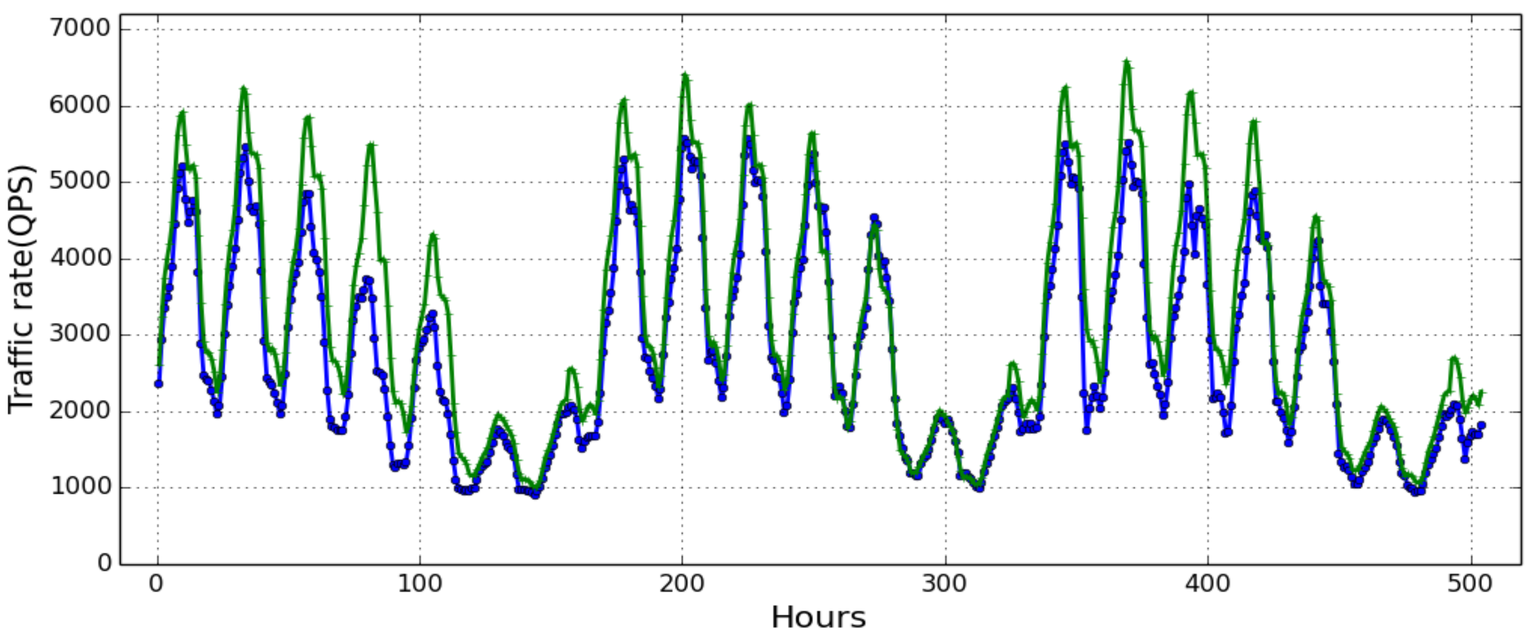
      $L_{i,j} = 0$

Design – Determining replication capacity

- Input:
  - SLA and Traffic rate
- Output:
  - Required replication capacity
- Binary searching
  - Starting with a (very) small capacity and a (very) large capacity
  - Get the middle capacity, determine the corresponding replication latency
  - Reset small or large capacity

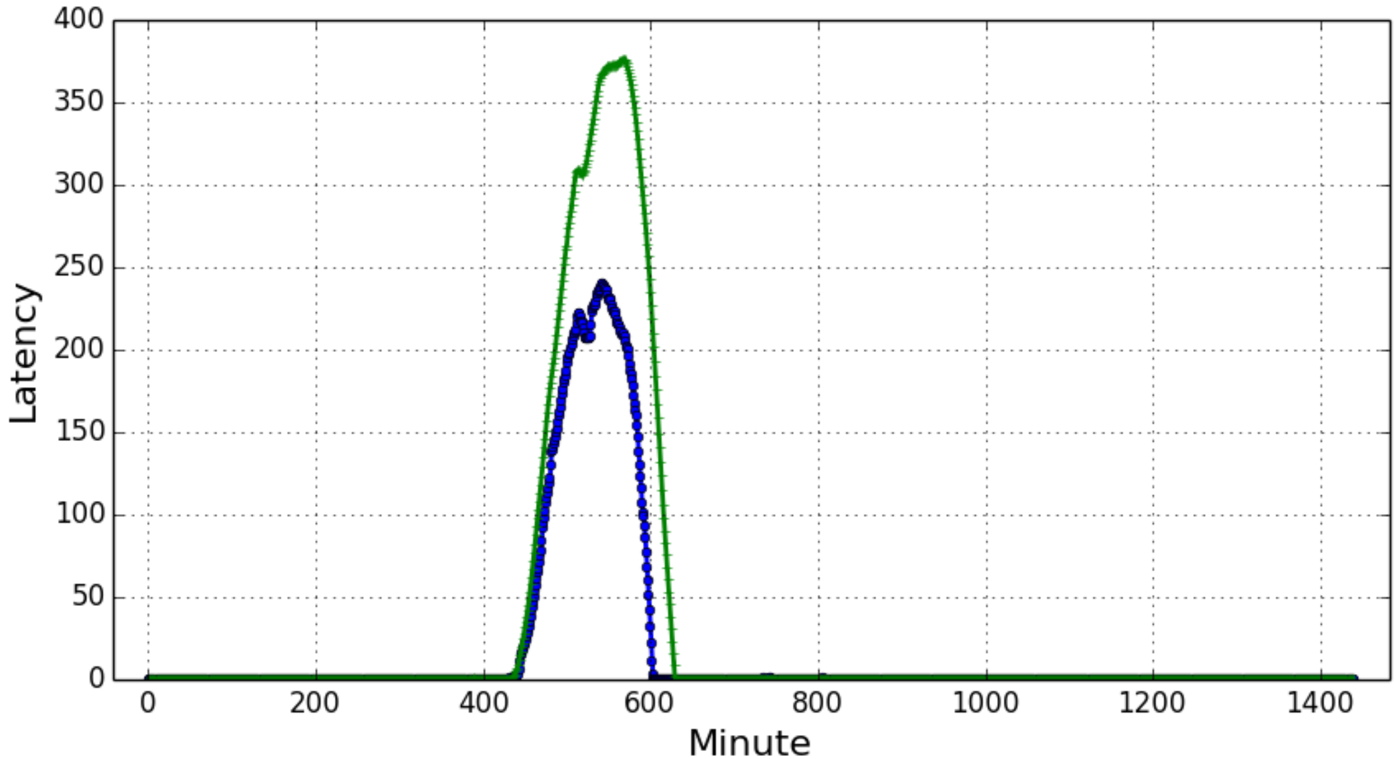
Evaluation - Forecasting

- Regression Analysis and ARIMA
  - Forecasted traffic rates have similar accuracies
- Reasons
  - Little dependency between neighboring data points (hourly)
  - Regression analysis works on weekly data, even less dependency



Evaluation – Determining replication latency

- Methodology
  - Choosing the busiest server; Reset offset
- Comparing the calculated relay lag
  - Shape is almost identical; peak value is 1.6X (376 vs 240 sec)



Evaluation - Others

- Replication capacity determination
  - Traffic rate of 2386 event/s; SLA 60 seconds
  - Takes 12 steps to get capacity of 3374 event/s
- Replication headroom determination
  - Capacity of 5000 event/s; SLA 60 seconds
  - Takes 9 steps to find it can sustain 8000 event/s traffic rate
  - Or taking 13 months to reach
- SLA determination
  - Capacity of 6000 event/s
  - Finds the maximum replication latency of 1135 seconds
  - P99 of replication latency is 850 seconds

Thanks!

- Questions ?
- zhenyun@gmail.com