ABSENCE: Usage-based Failure Detection in Mobile Networks

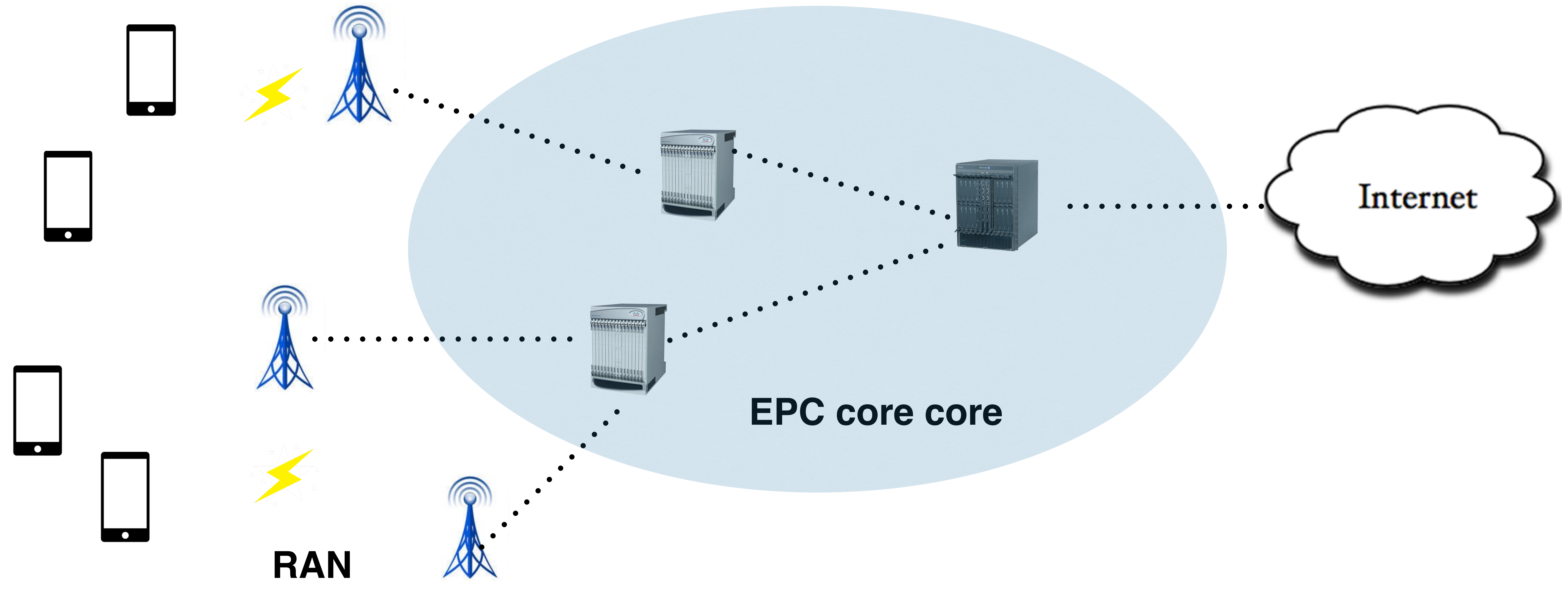
Binh Nguyen, Zihui Ge, Jacobus Van der Merwe, He Yan, Jennifer Yates

Mobicom 2015

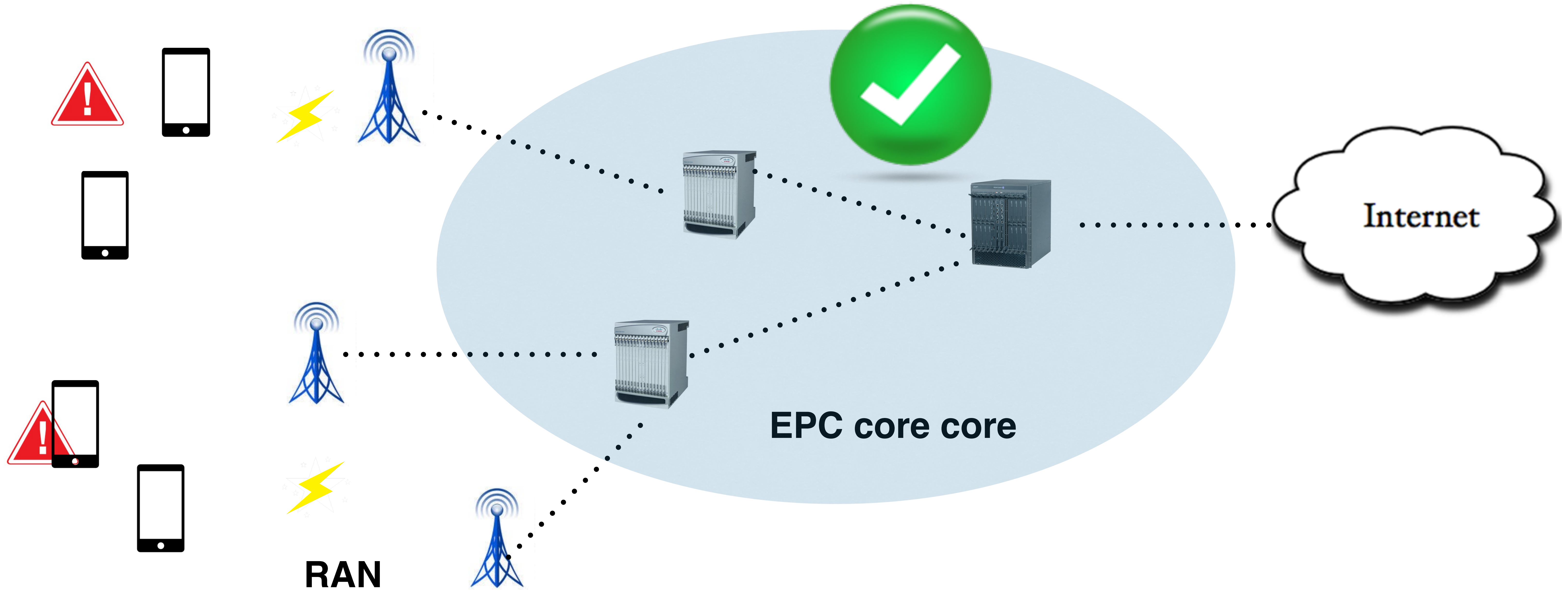
THE UNIVERSITY OF UTAH

AT&T Labs Research
The Science Behind The Service

Silent failures

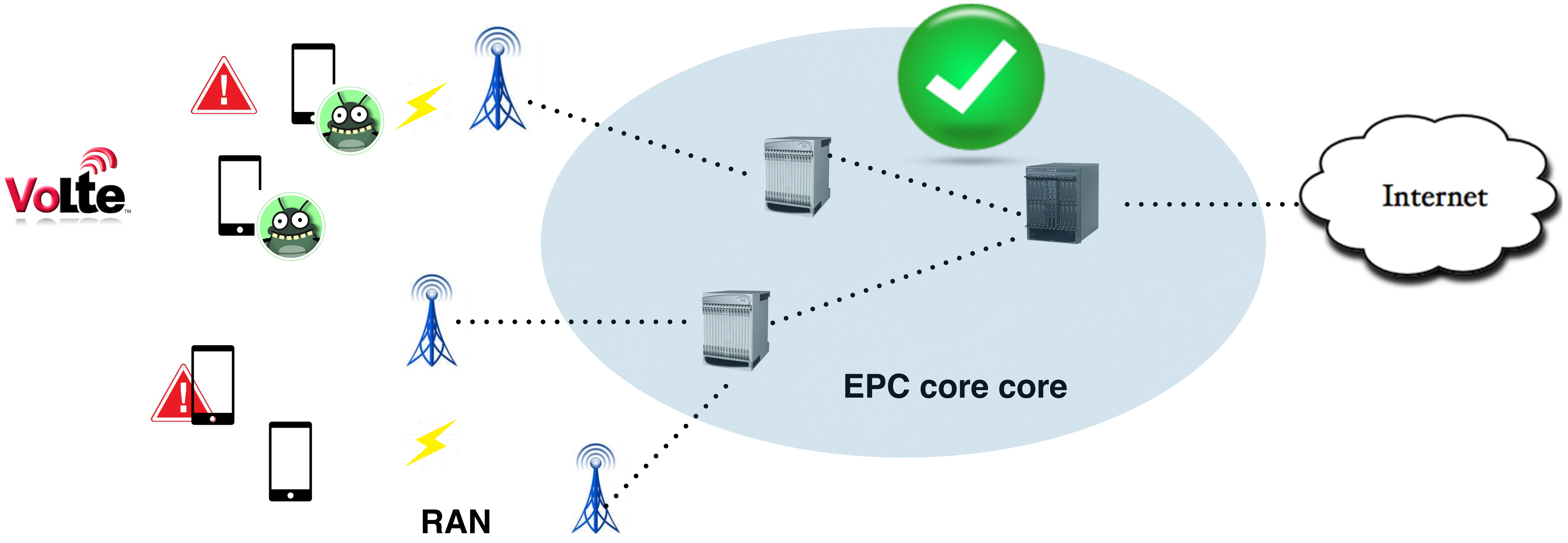


Silent failures



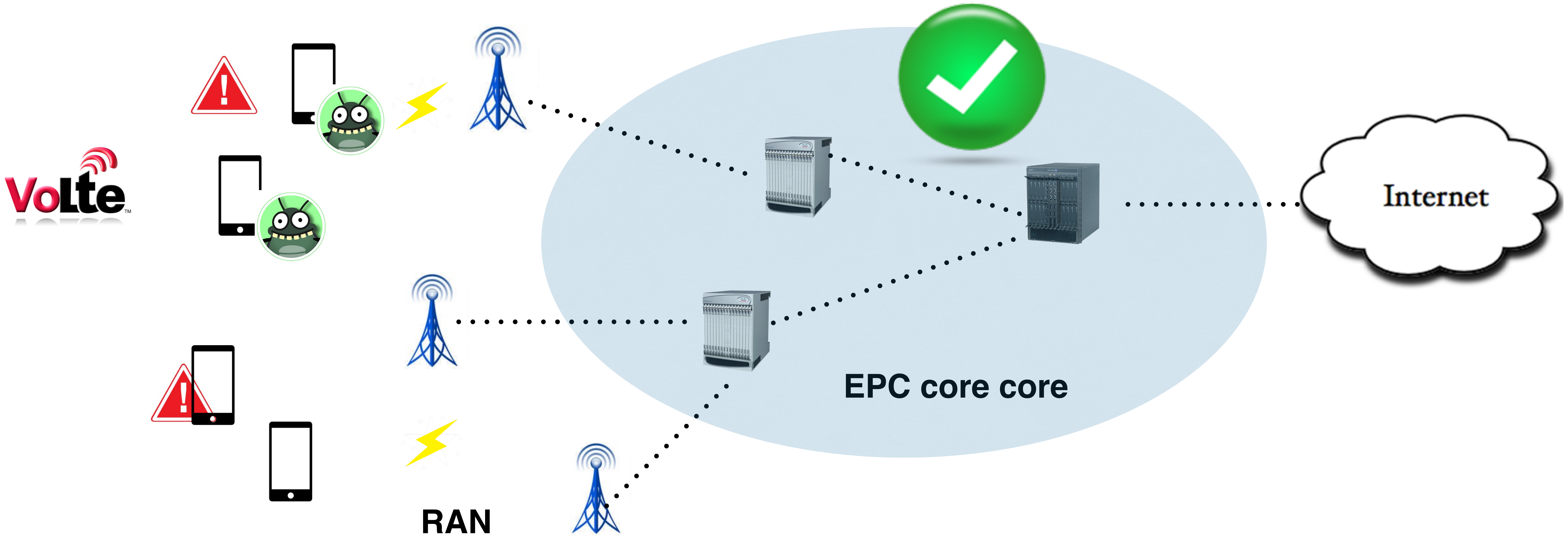
- Silent failures: service disruptions/outages that are not detected by current monitoring systems.
- New features rolled out, bugs on devices, or combination of both.

Silent failures



- Silent failures: service disruptions/outages that are not detected by current monitoring systems.
- New features rolled out, bugs on devices, or combination of both.

Silent failures



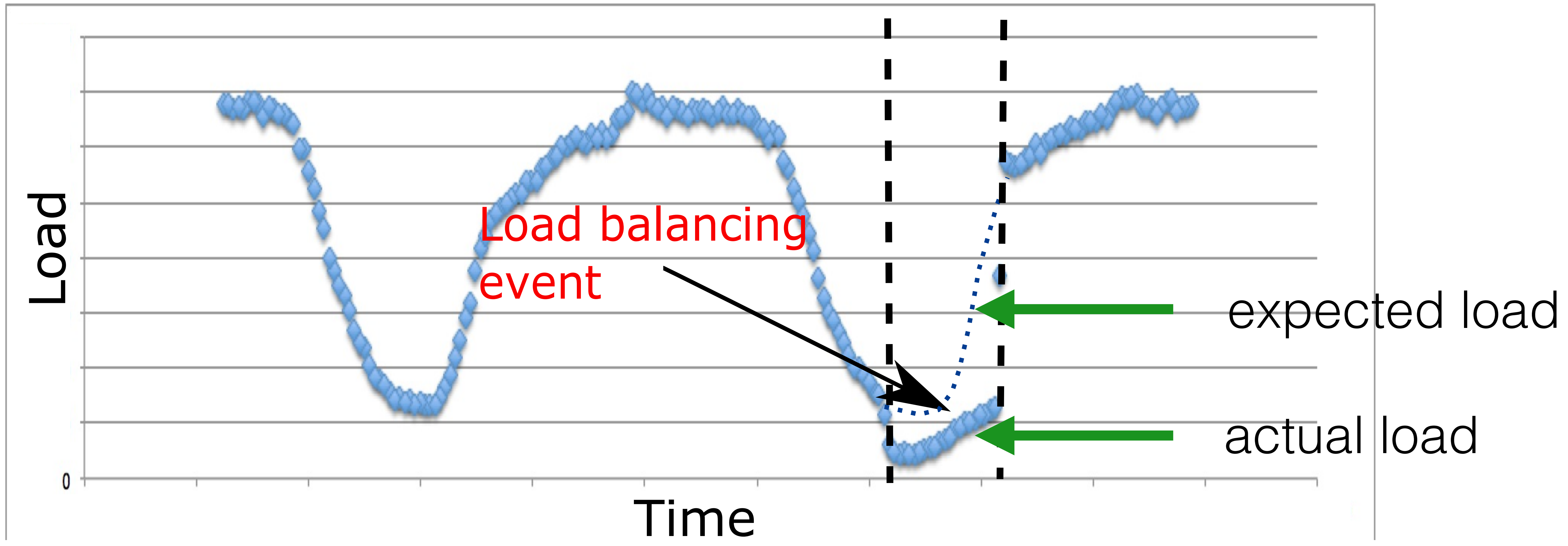
- Silent failures: service disruptions/outages that are not detected by current monitoring systems.
- New features rolled out, bugs on devices, or combination of both.

Detecting silent failures is challenging!

Detecting silent failures is difficult - passive network monitoring

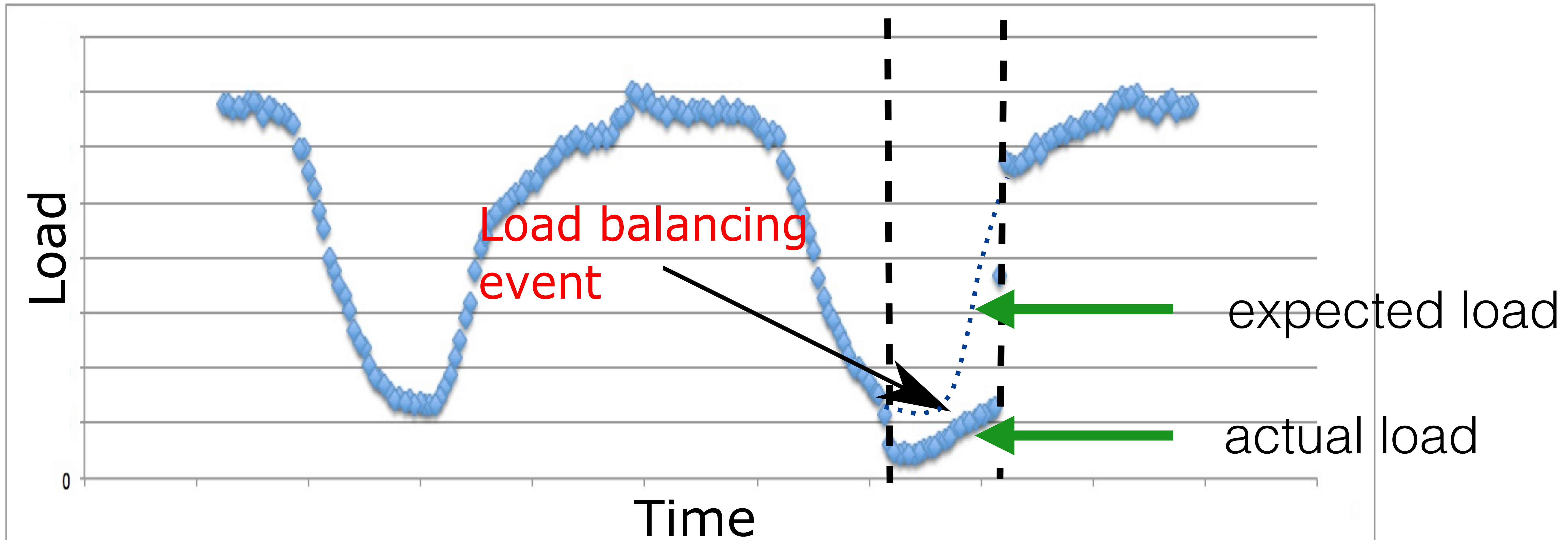
Detecting silent failures is difficult - passive network monitoring

- Drops in traffic/usage on network elements ***do not*** imply service disruptions:
  - Load balancing/maintenance activities.
  - Dynamic routing/Self-Organizing Network (SON).



Detecting silent failures is difficult - passive network monitoring

- Drops in traffic/usage on network elements ***do not*** imply service disruptions:
  - Load balancing/maintenance activities.
  - Dynamic routing/Self-Organizing Network (SON).
- Key Performance metric Indicators (KPI) **may not** reflect service issues:
  - E.g., accessibility KPI looks good even when only a subset of users can access the network.

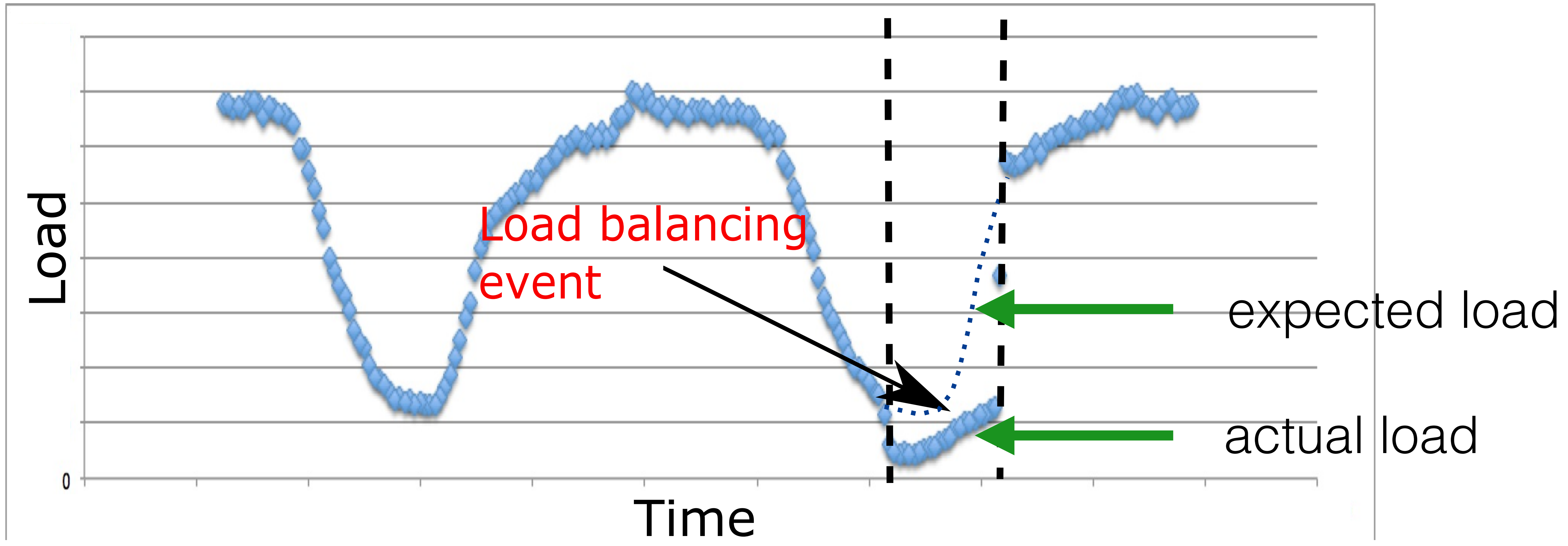


Detecting silent failures is difficult - passive network monitoring

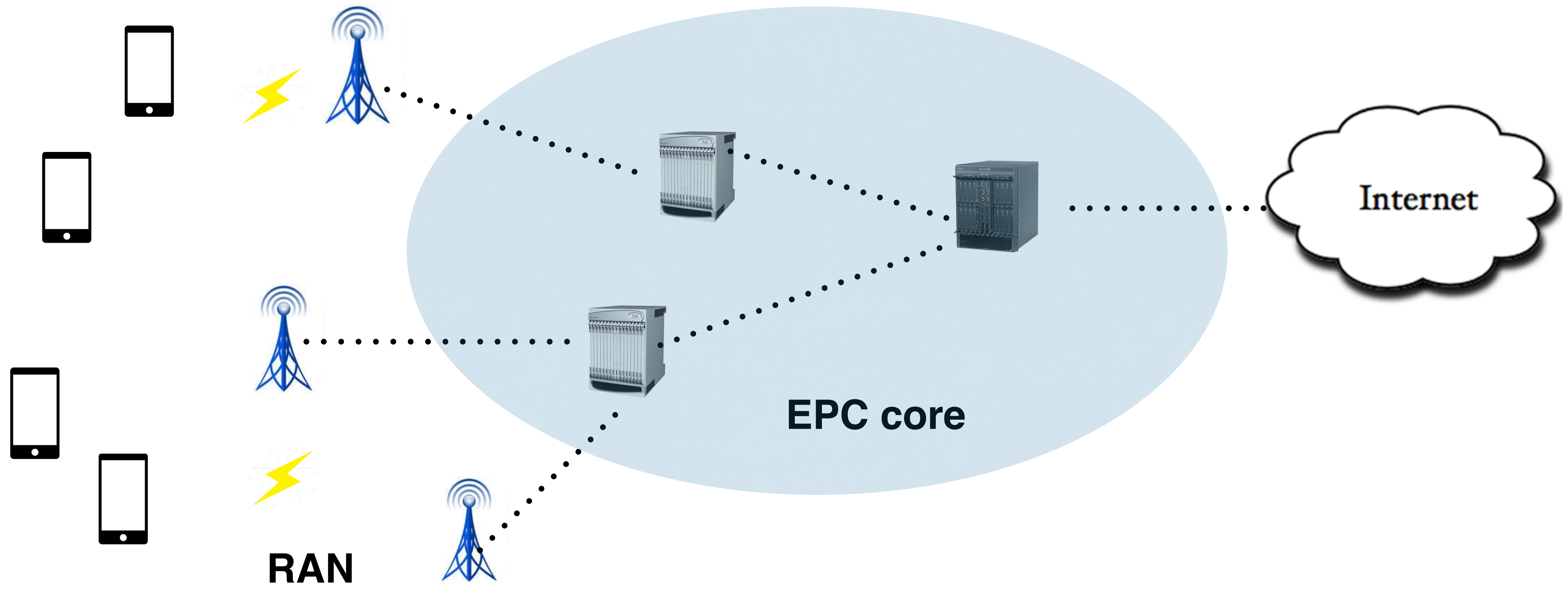
- Drops in traffic/usage on network elements ***do not*** imply service disruptions:
  - Load balancing/maintenance activities.
  - Dynamic routing/Self-Organizing Network (SON).

A "healthy network" (from a monitoring perspective) does not guarantee service experience of users!

  - E.g., accessibility KPI looks good even when only a subset of users can access the network.

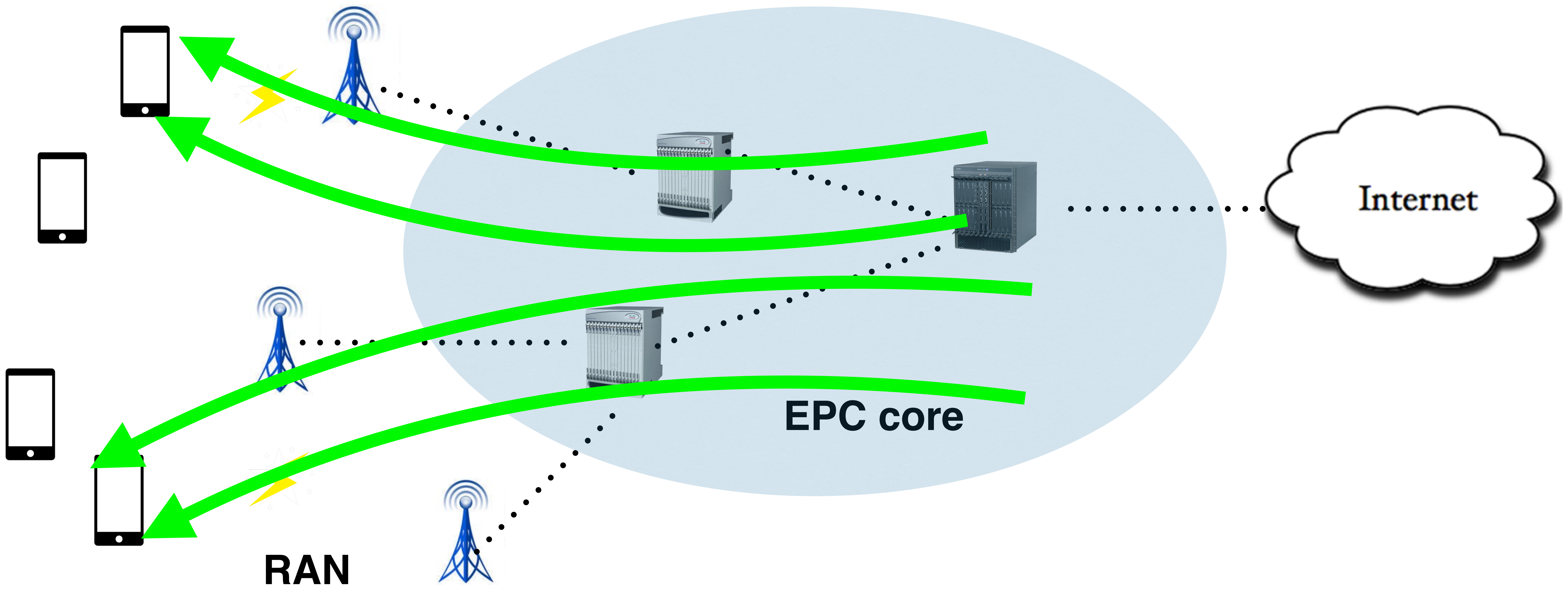


Detecting silent failures is difficult - active service monitoring



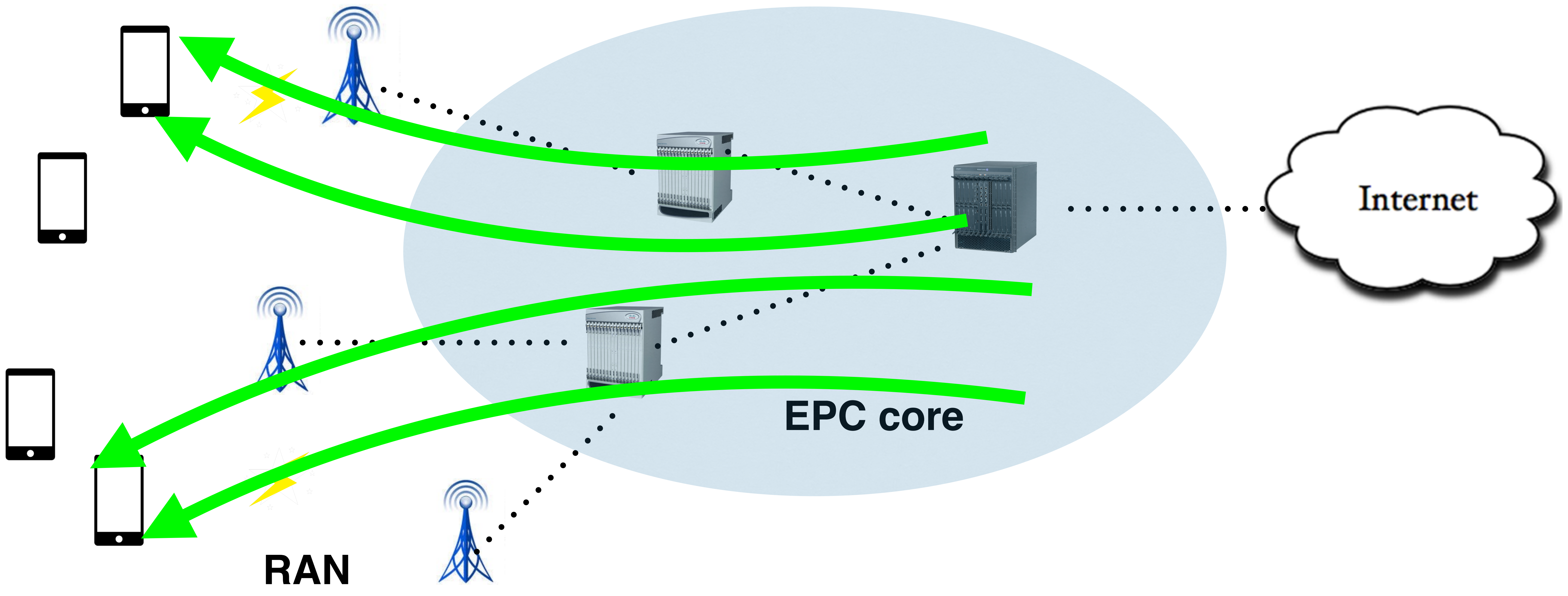
Detecting silent failures is difficult - active service monitoring

- Sending test traffic across the network on **all** service paths.



Detecting silent failures is difficult - active service monitoring

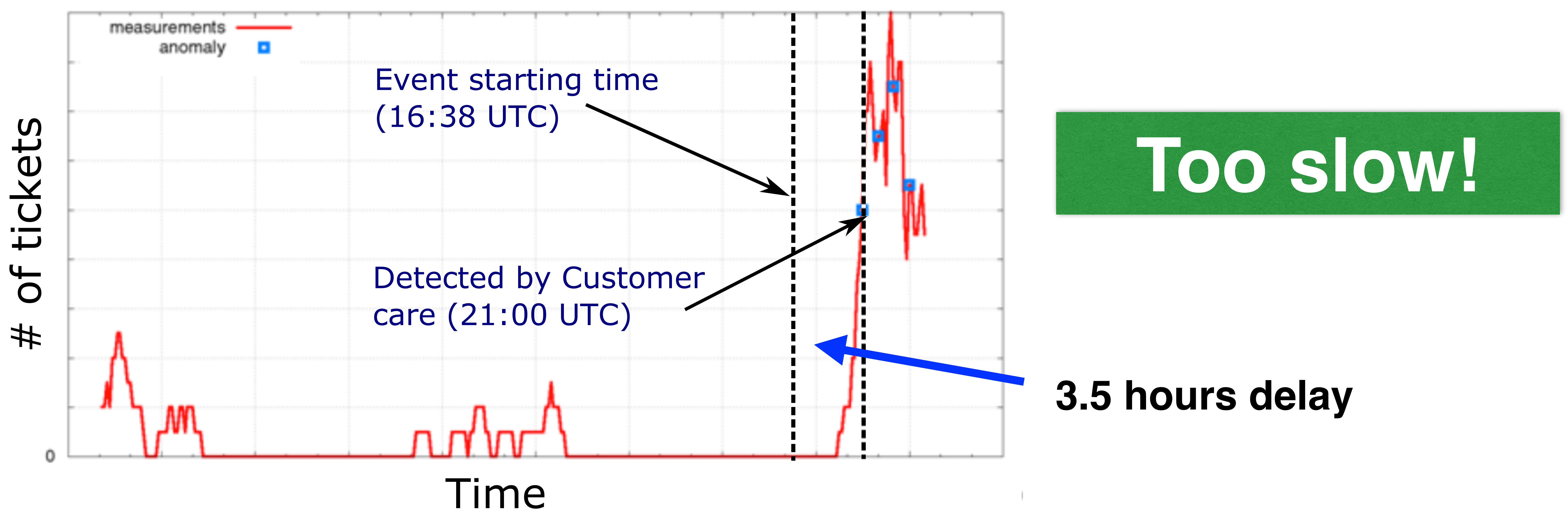
- Sending test traffic across the network on **all** service paths.
- Many types of customer devices, applications, huge geographic environment to probe.



Active monitoring does not scale!

Relying on customer feedback

- It takes time for customers to give feedback.
- Relying on customer feedback is **too slow**: hours of delay.
  - E.g., failure happens at **16:38 UTC** but manifests in customer feedback at **21:00 UTC**, **3.5 hours of delay**.



ABSENCE: usage-based failure detection

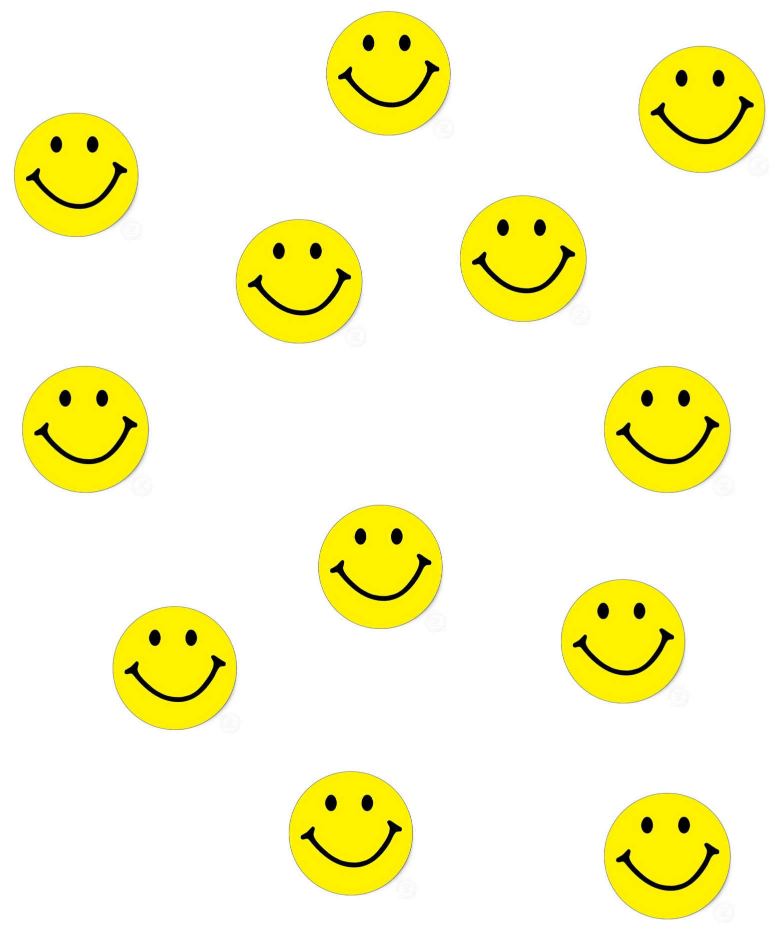
ABSENCE: usage-based failure detection

- ABSENCE: **Passive service monitoring** approach - monitor usage of users in a passive manner.

ABSENCE: usage-based failure detection

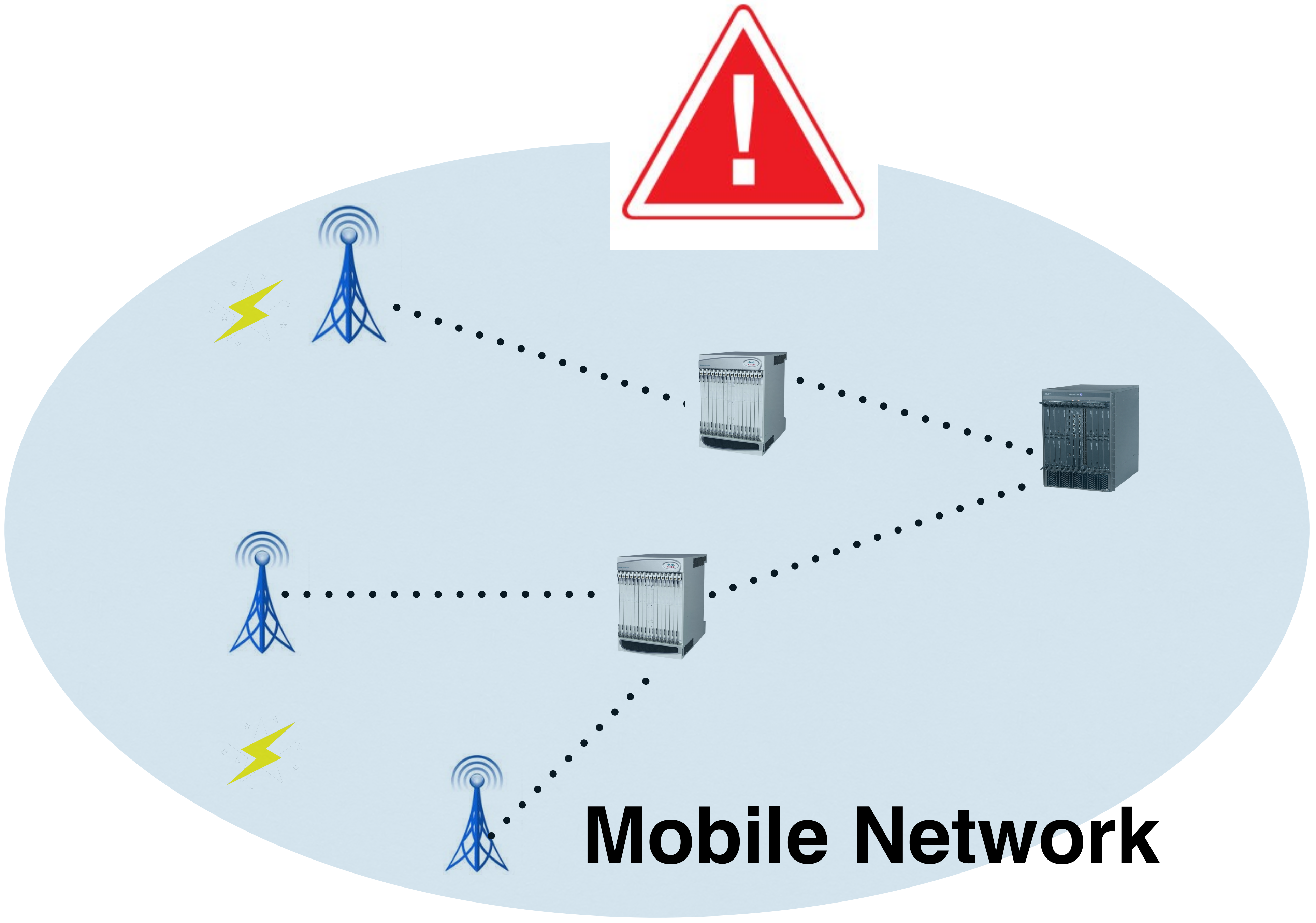
- ABSENCE: **Passive service monitoring** approach - monitor usage of users in a passive manner.
- ***Absence of customer usage*** is a reliable indicator of service disruptions in a mobile network.

ABSENCE’s key ideas

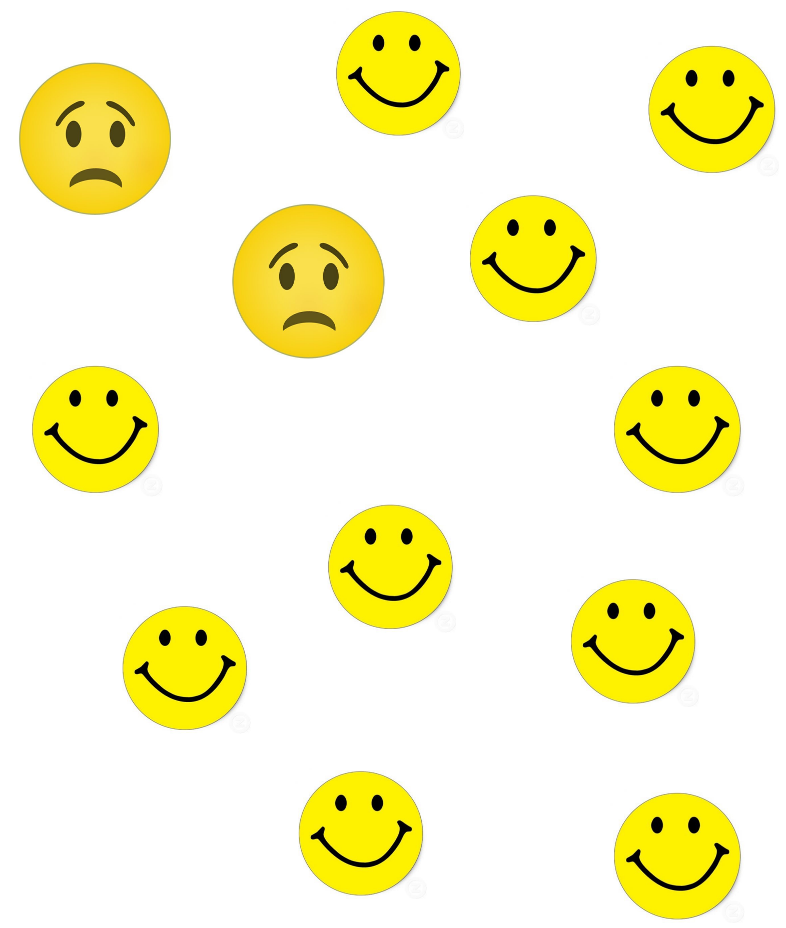


A group of users

Usage

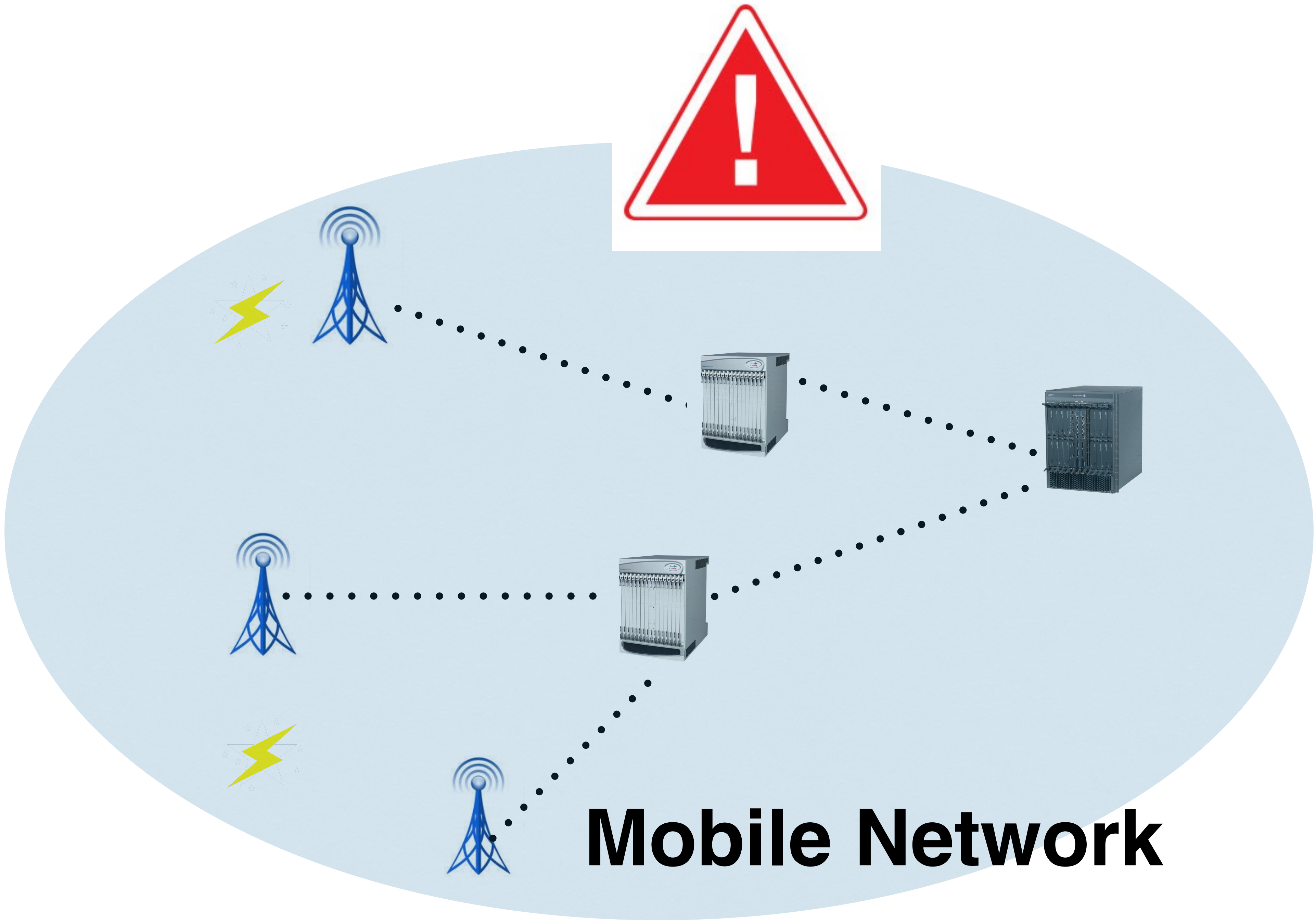


ABSENCE’s key ideas



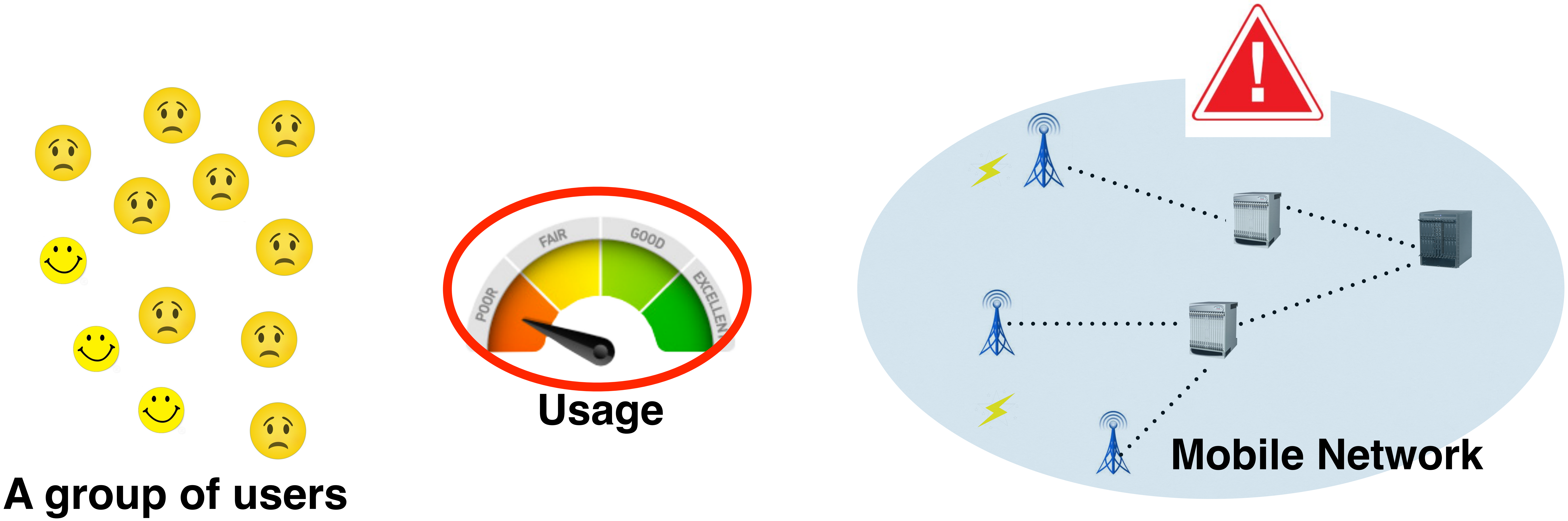
A group of users

Usage



- If failure happens, users are not able to use the network as normal.

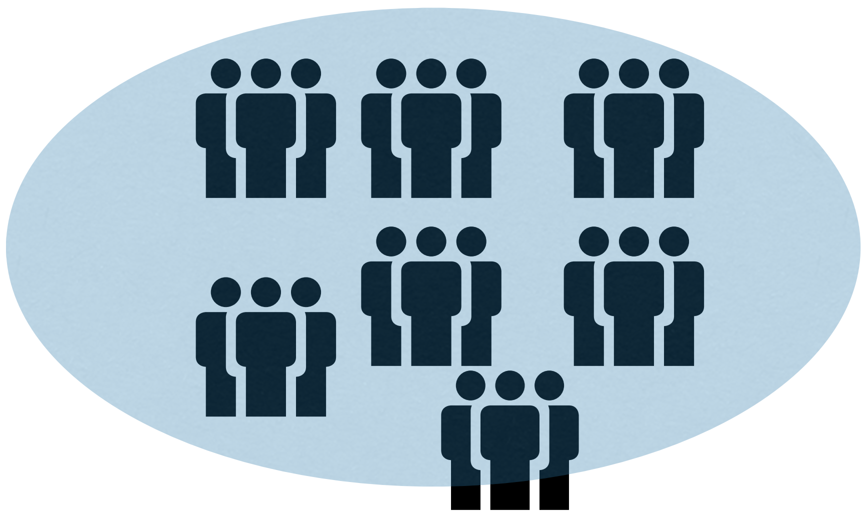
ABSENCE’s key ideas



- If failure happens, users are not able to use the network as normal.
- Large number of users cannot use the network leads to a drop in usage.
- Could detect both hard failures (outages) and performance degradations.

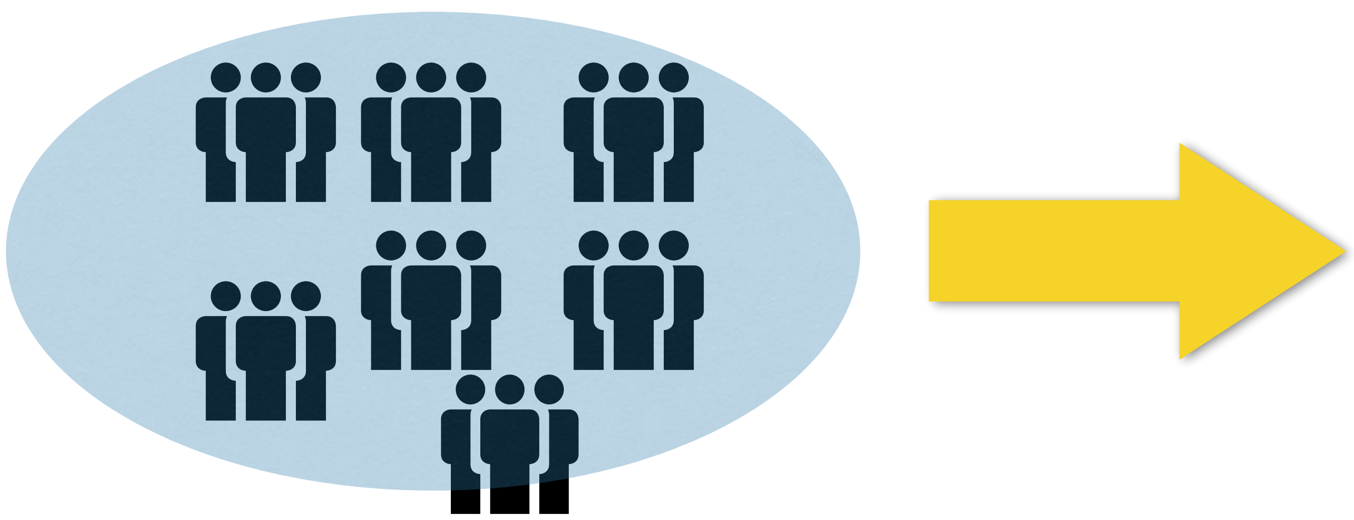
ABSENCE overview

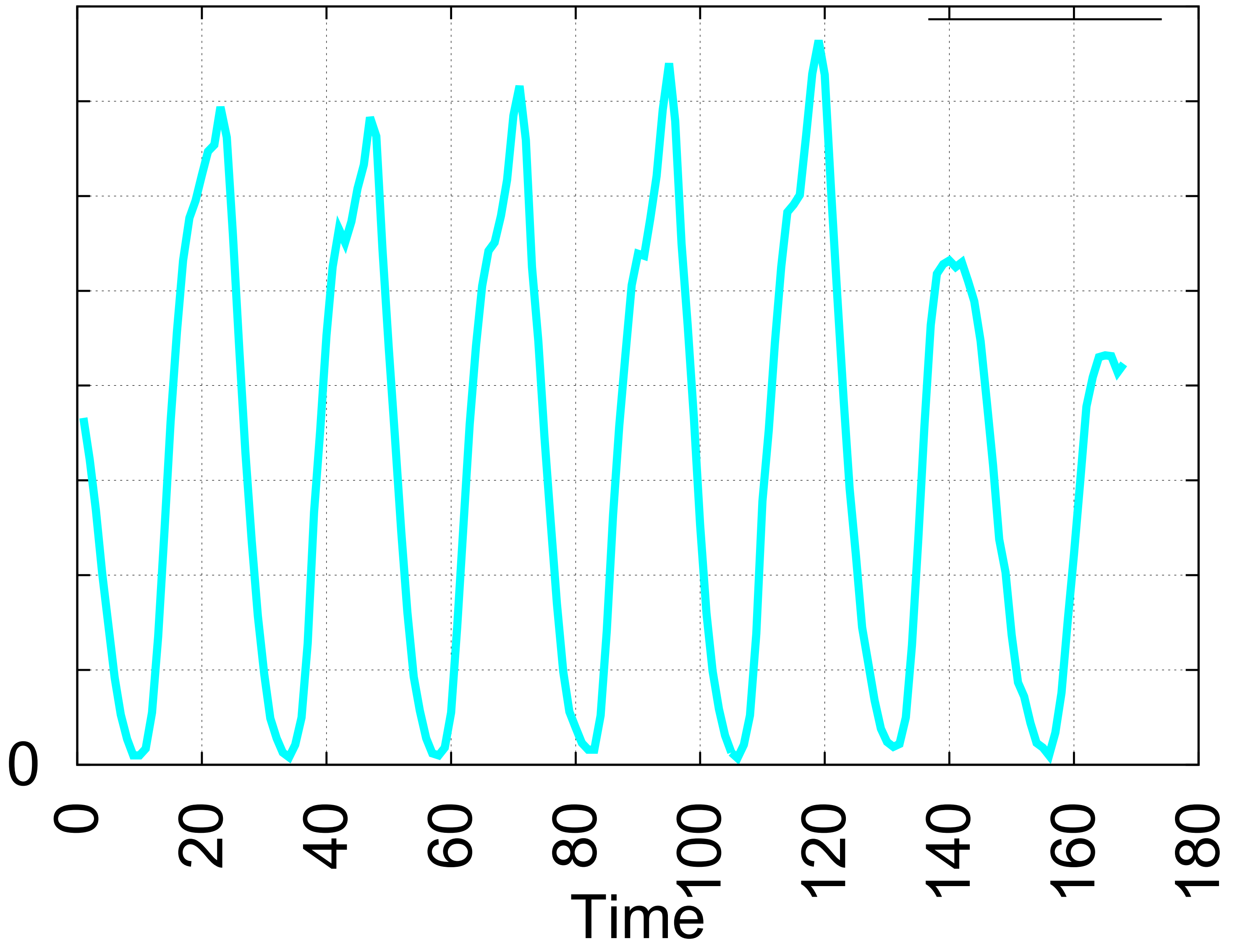
Use anonymized and aggregated Call Detail Record (CDR) collected in real time from an U.S. operator.



ABSENCE overview

Use anonymized and aggregated Call Detail Record (CDR) collected in real time from an U.S. operator.

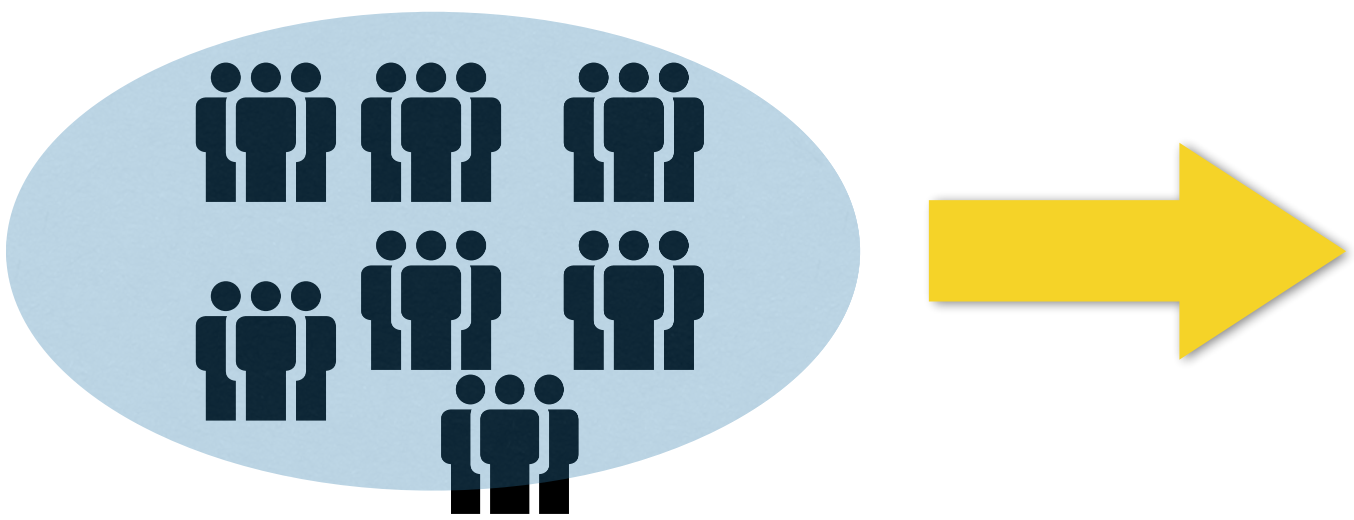


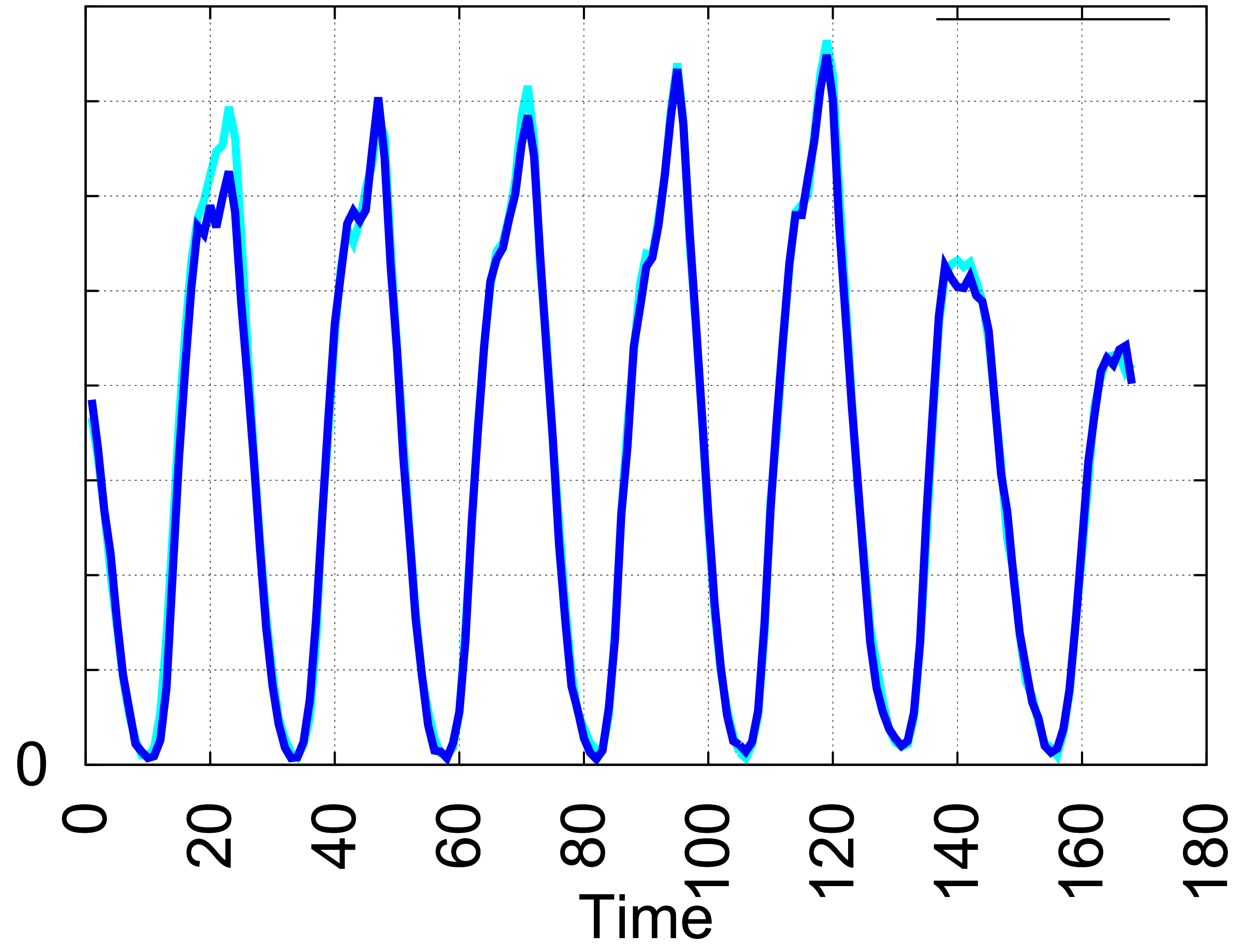


Week 1

ABSENCE overview

Use anonymized and aggregated Call Detail Record (CDR) collected in real time from an U.S. operator.

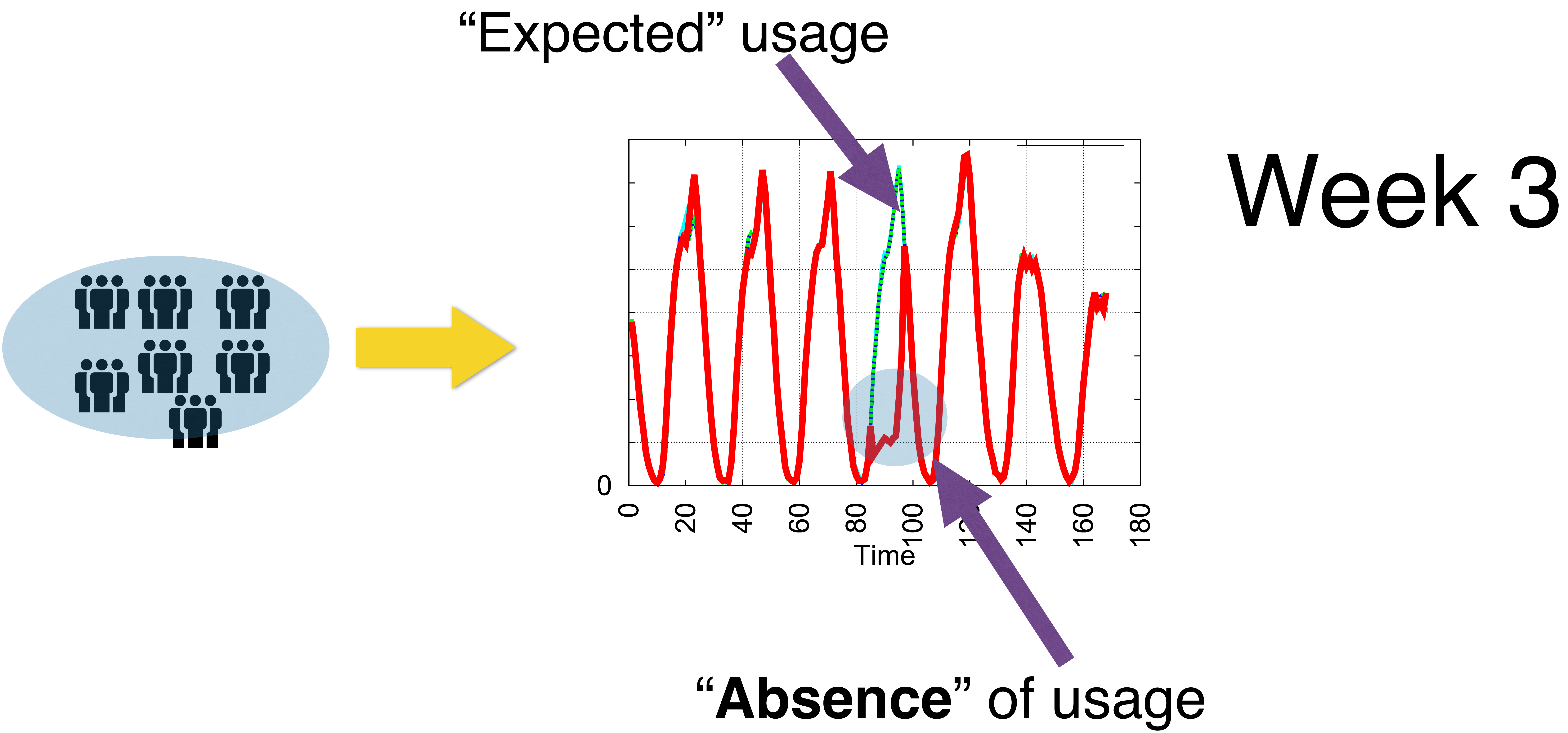




Week 2

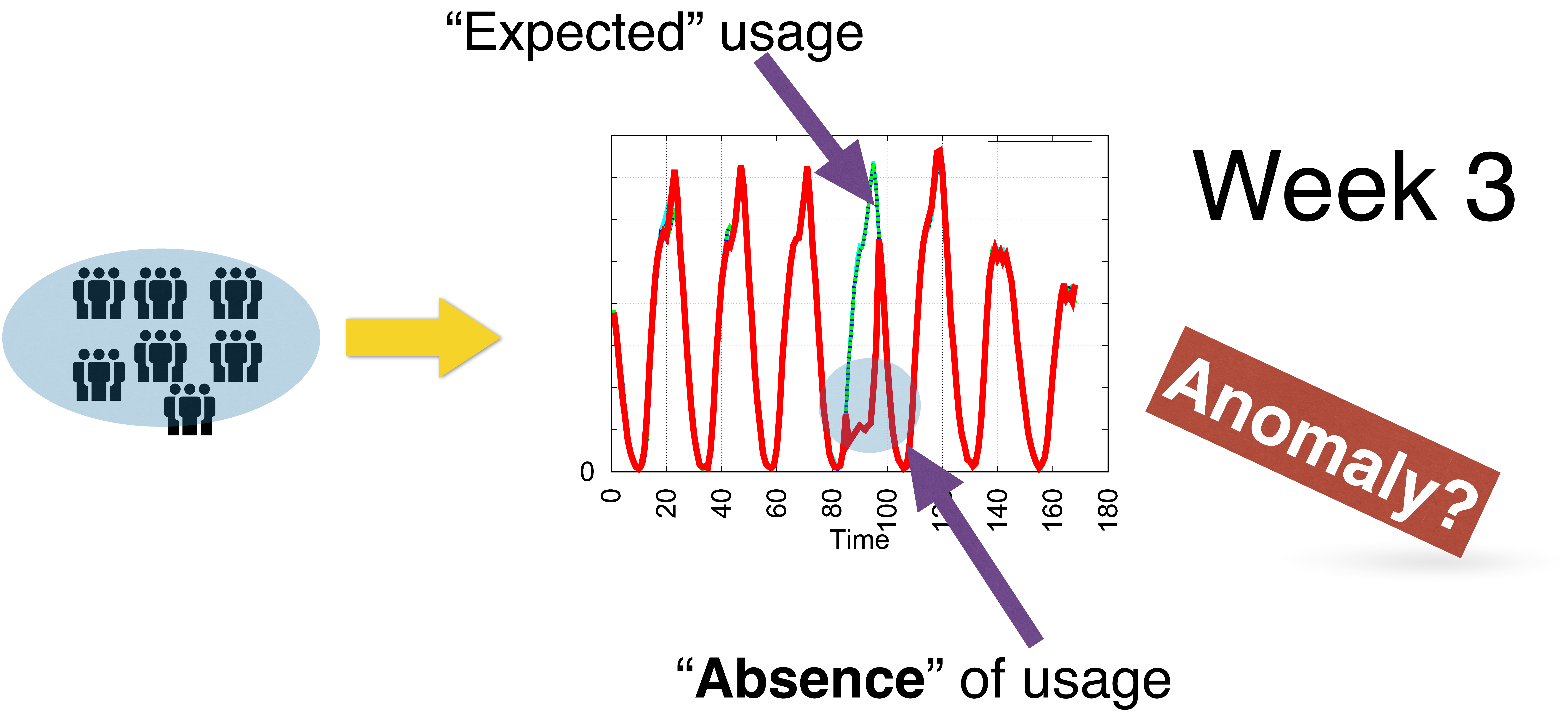
ABSENCE overview

Use anonymized and aggregated Call Detail Record (CDR) collected in real time from an U.S. operator.



ABSENCE overview

Use anonymized and aggregated Call Detail Record (CDR) collected in real time from an U.S. operator.



Outline

- Motivation.
- ABSENCE overview.
- **Is ABSENCE feasible?**
- ABSENCE's challenges.
- ABSENCE's event detection.
- Synthetic workload evaluation.
- Operational validation.

Is ABSENCE feasible?

Is usage predictable enough for anomaly detection?

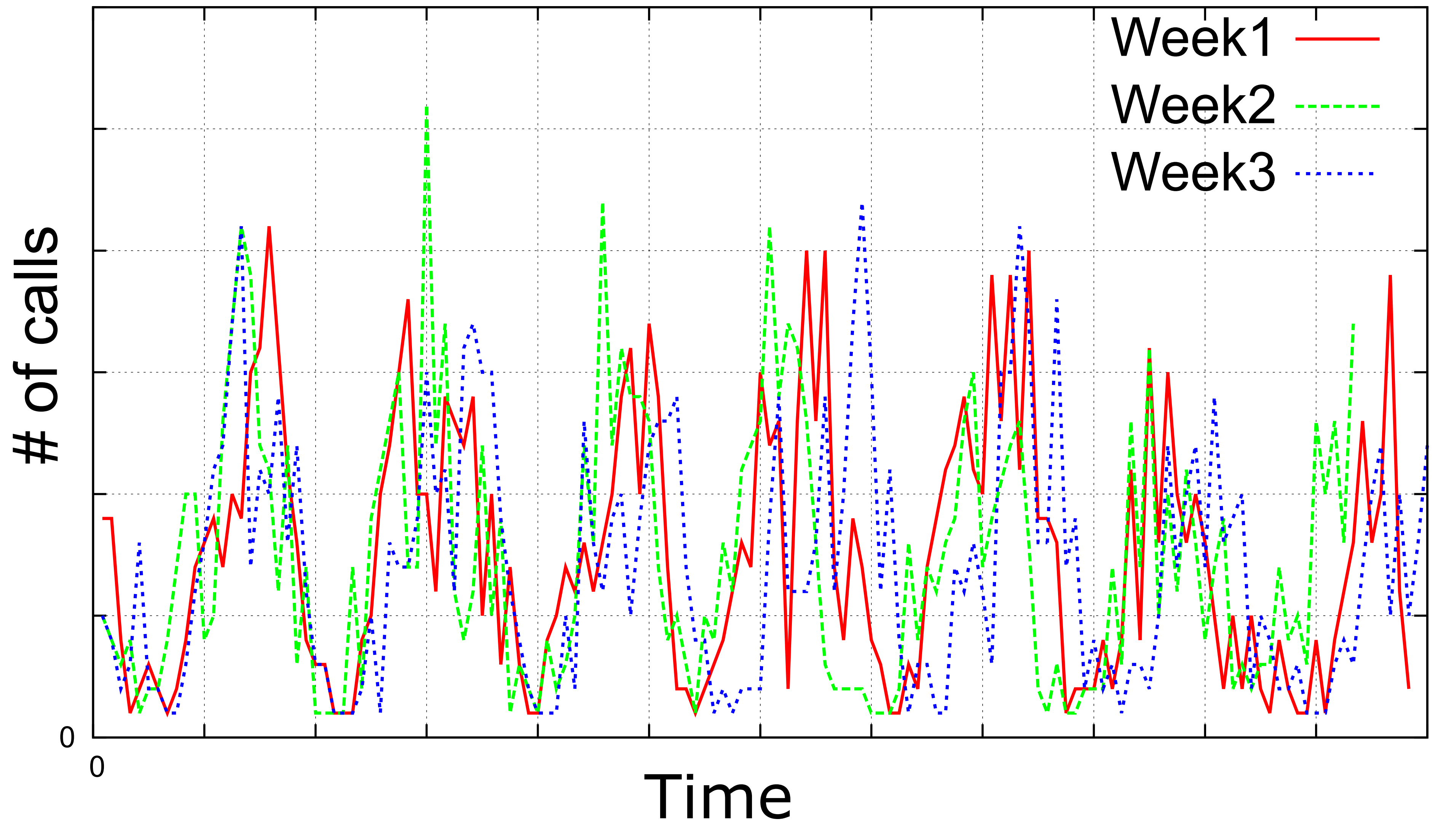
Is usage predictable enough?

Is usage predictable enough?

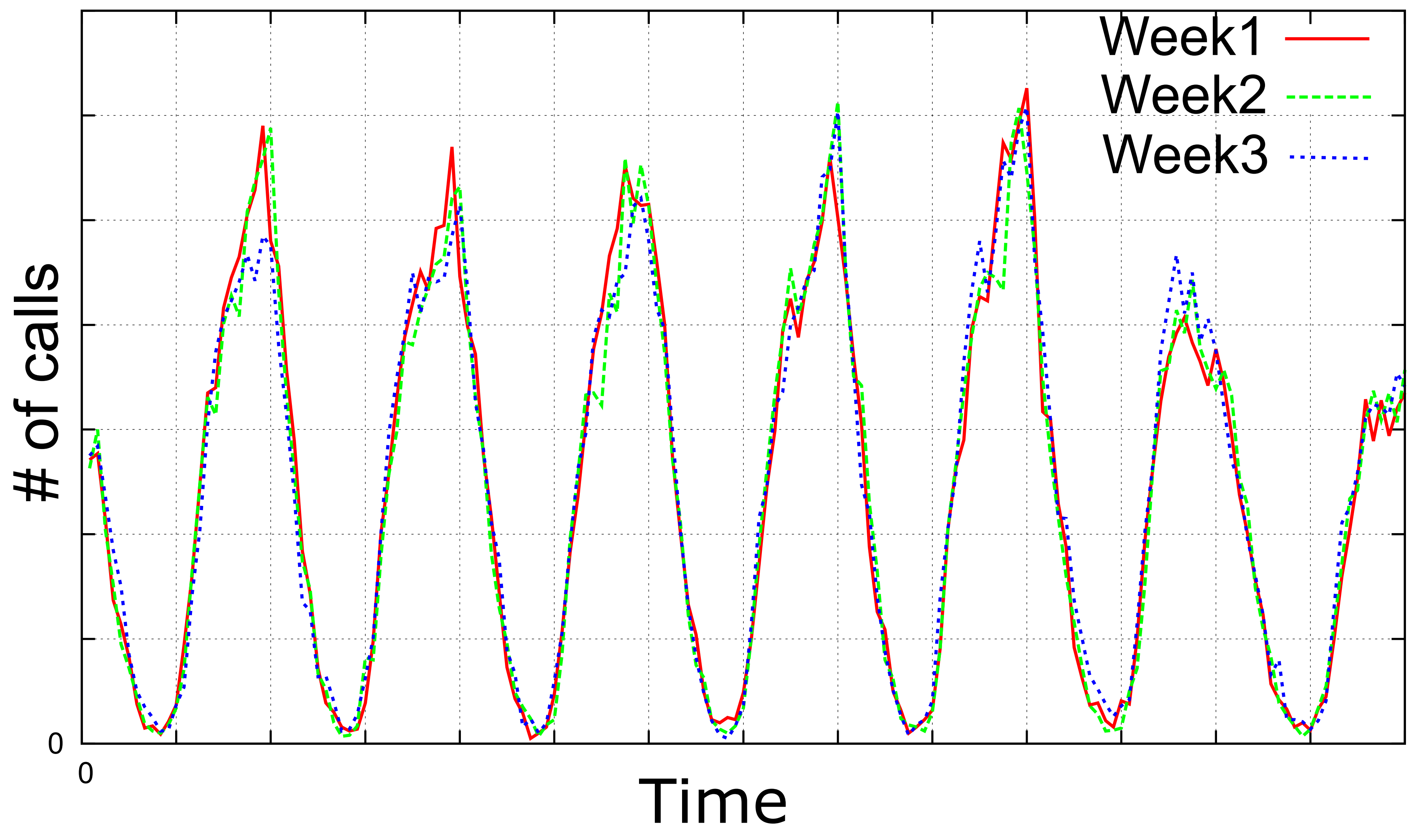
- While individual user usage is not predictable, usage of a large group of users is predictable.

Is usage predictable enough?

- While individual user usage is not predictable, usage of a large group of users is predictable.
- For example: 3 weeks of usage overlapped, usage of a small group is less predictable than usage of a large group.



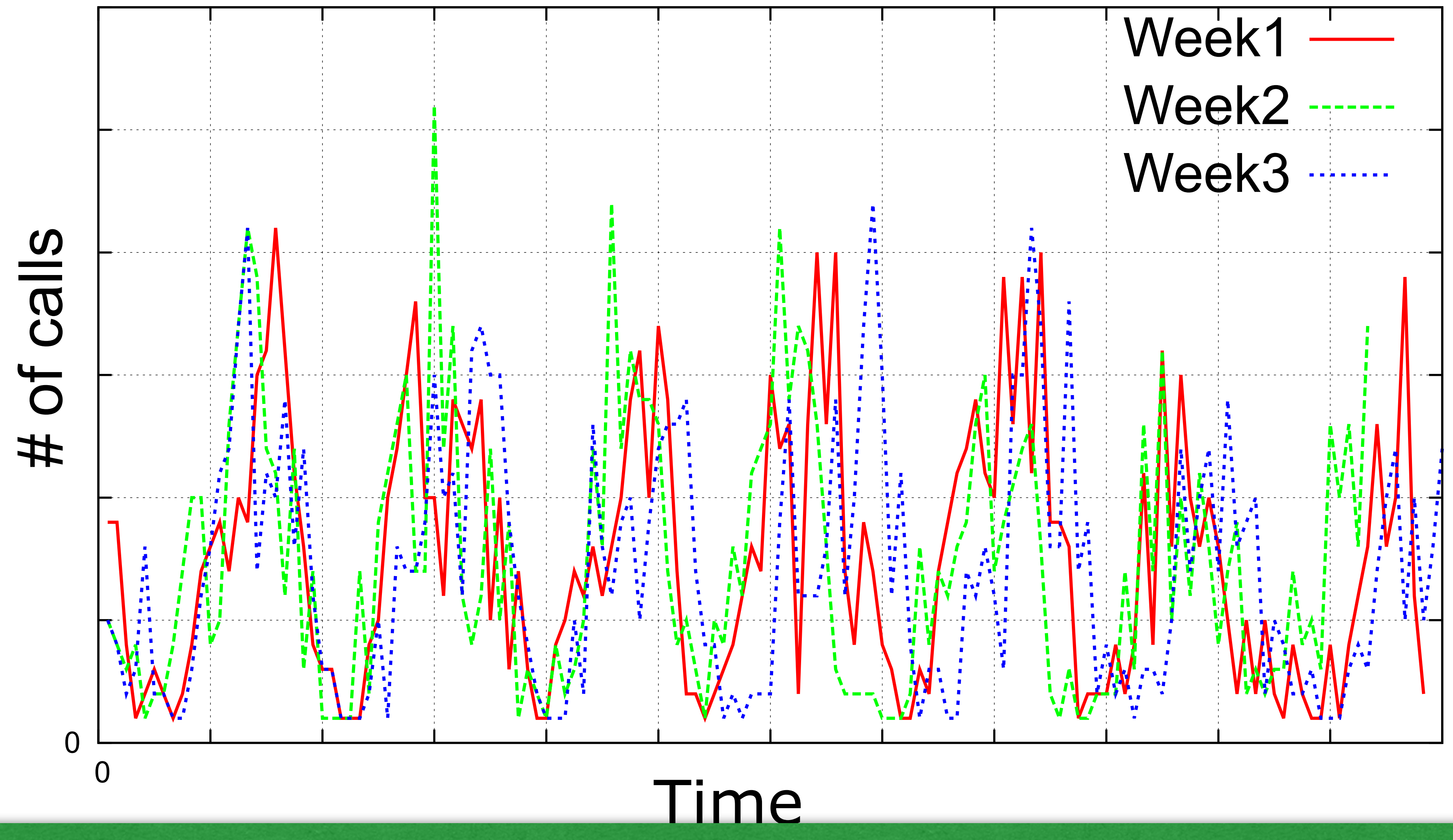
70 users

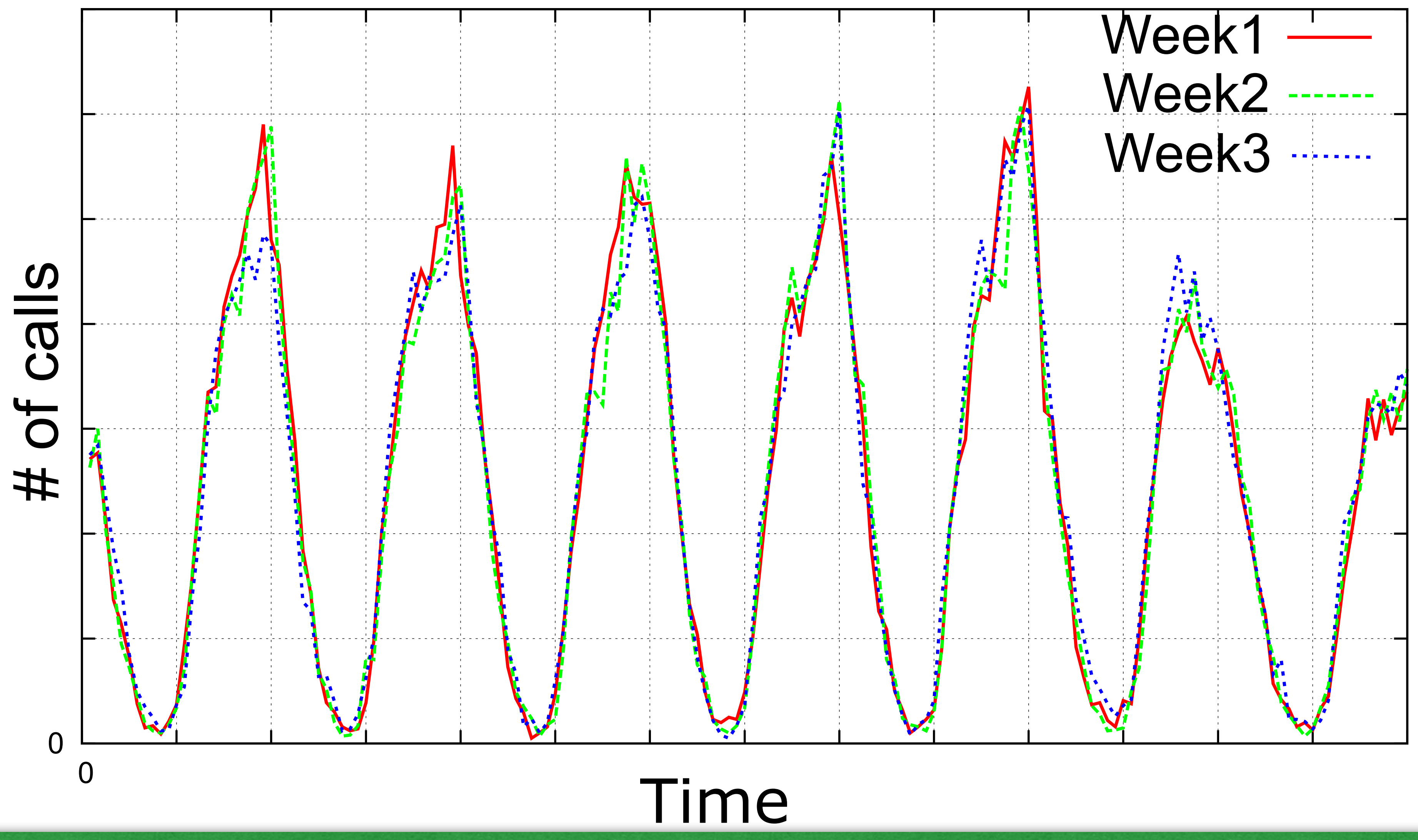


3000 users

Is usage predictable enough?

- While individual user usage is not predictable, usage of a large group of users is predictable.
- For example: 3 weeks of usage overlapped, usage of a small group is less predictable than usage of a large group.





Yes, usage of a large enough group of users is predictable!

Outline

- Motivation.
- ABSENCE overview.
- Is ABSENCE feasible?
- **ABSENCE’s challenges.**
- ABSENCE’s event detection.
- Synthetic workload evaluation.
- Operational validation.

Challenges

- Failures happens to different scopes: geo-area, device makes/models, service types.

- How to deal with users mobility?

- How to improve predictability of aggregate usage?

- How to make ABSENCE scalable, given a large amount of data in the network?

Challenges

- Failures happens to different scopes: geo-area, device makes/models, service types.
- How to deal with users mobility?
- How to improve predictability of aggregate usage?
- How to make ABSENCE scalable, given a large amount of data in the network?

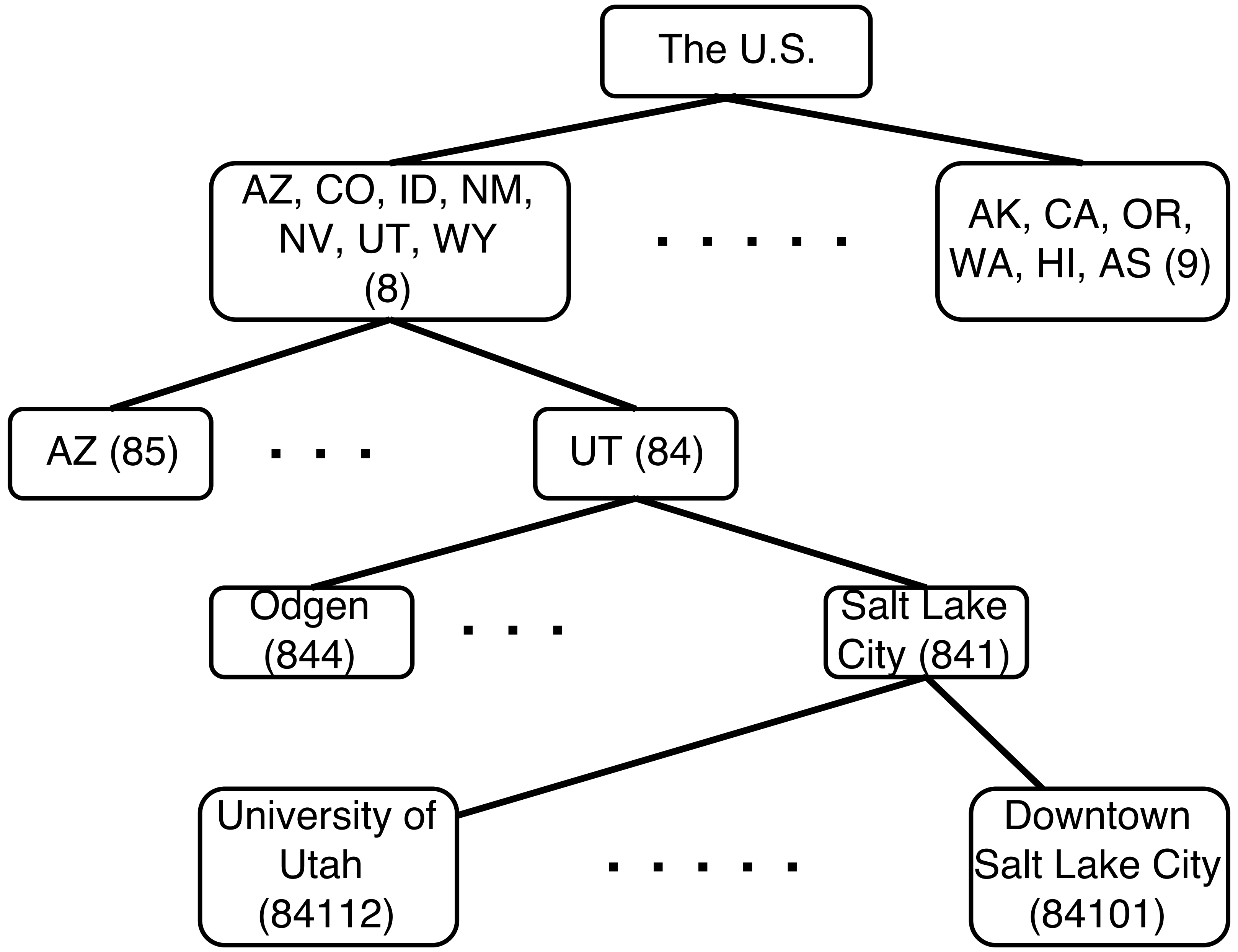
How to detect failures with different scopes?

How to detect failures with different scopes?

- Group users based on their geographical information: ZIP code area, city, state.
- A user could belong to multiple geographical groups in the same time.
- Under each geographical group: further divided to device OS, make.

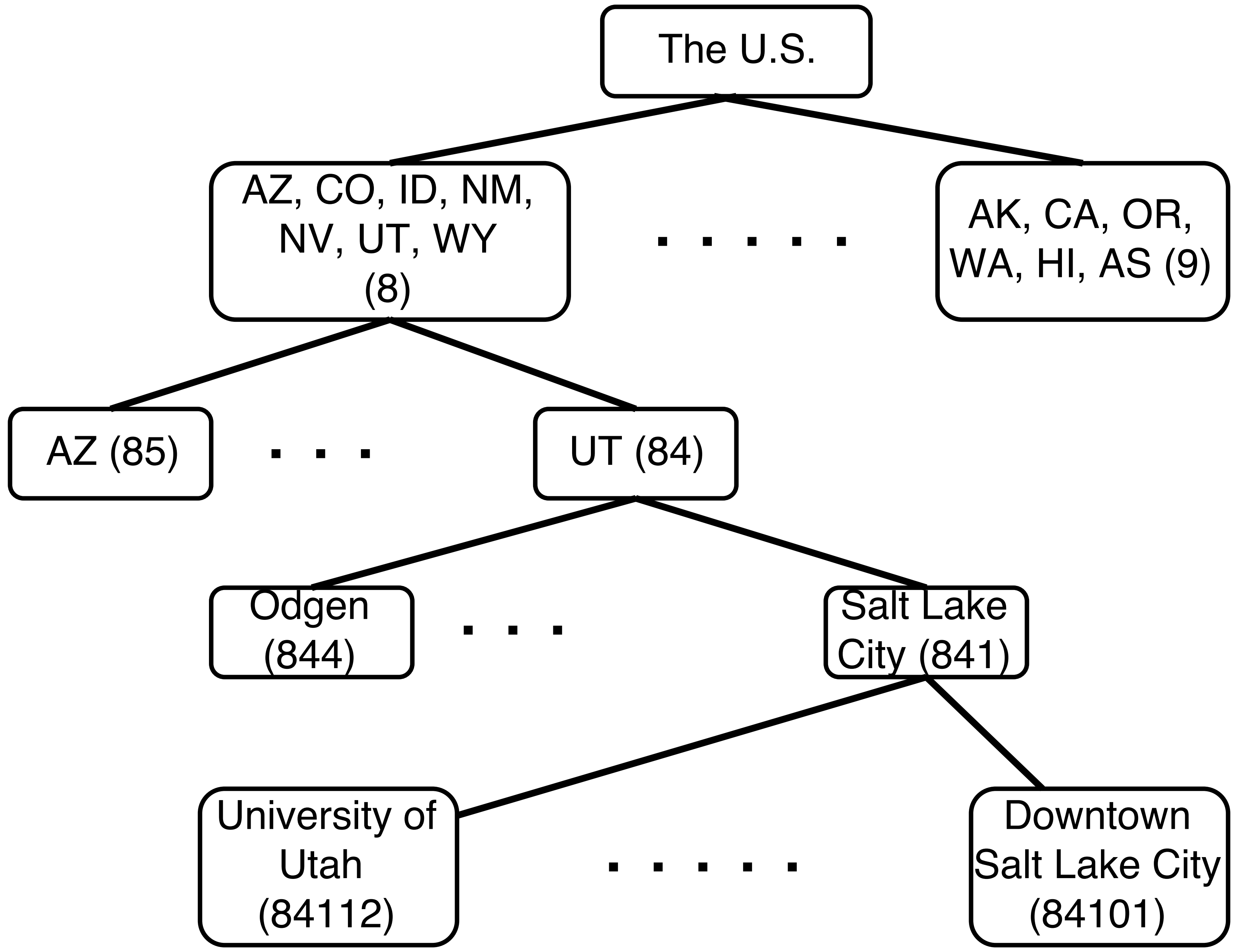
How to detect failures with different scopes?

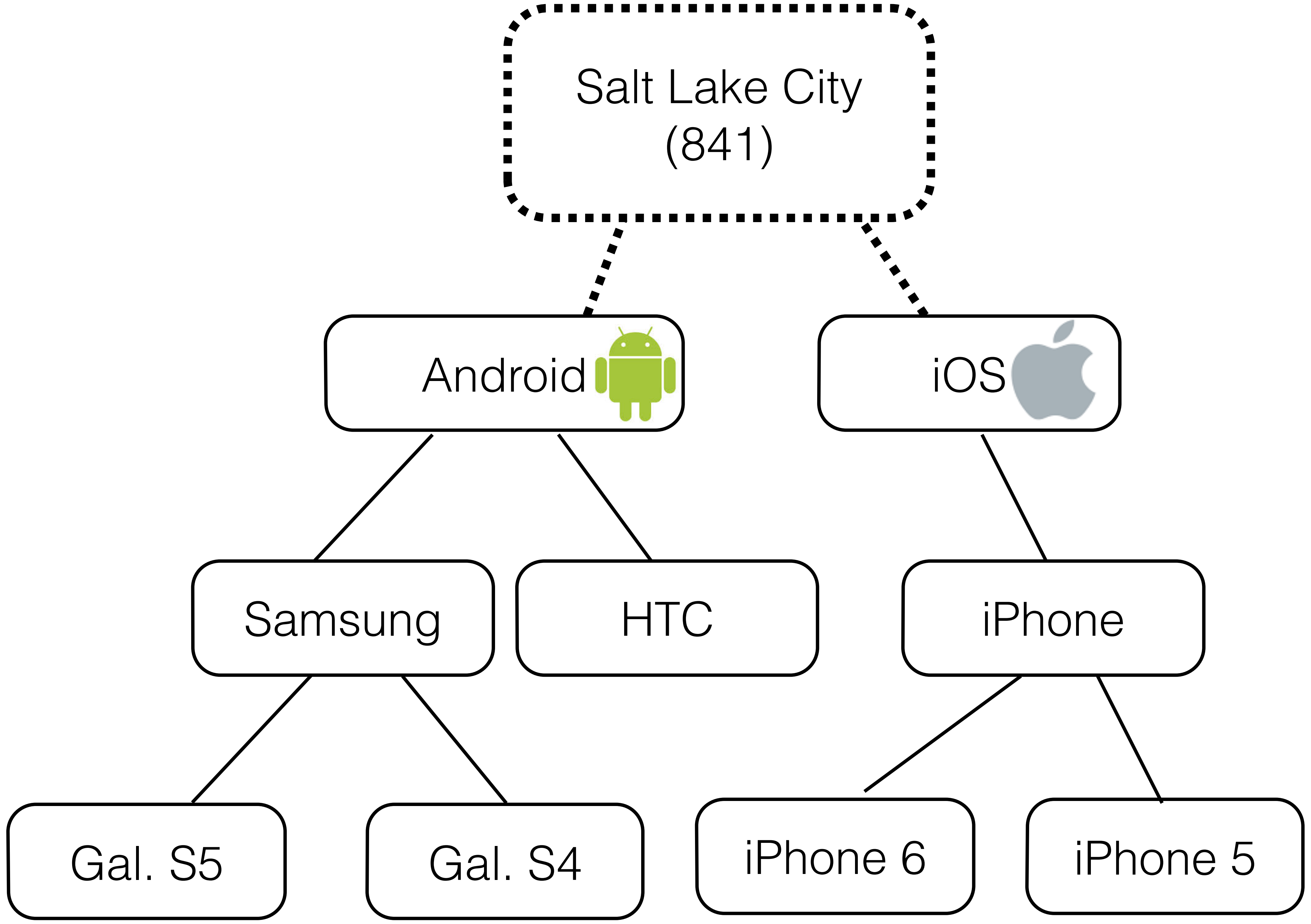
- Group users based on their geographical information: ZIP code area, city, state.
- A user could belong to multiple geographical groups in the same time.
- Under each geographical group: further divided to device OS, make.



How to detect failures with different scopes?

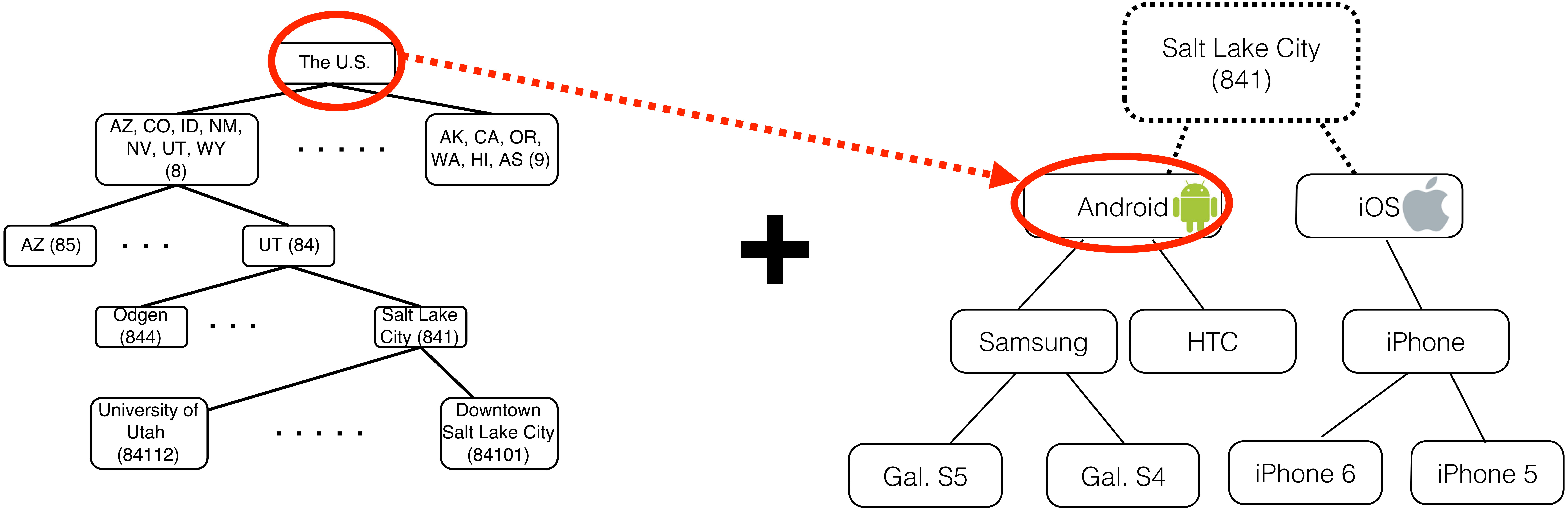
- Group users based on their geographical information: ZIP code area, city, state.
- A user could belong to multiple geographical groups in the same time.
- Under each geographical group: further divided to device OS, make.





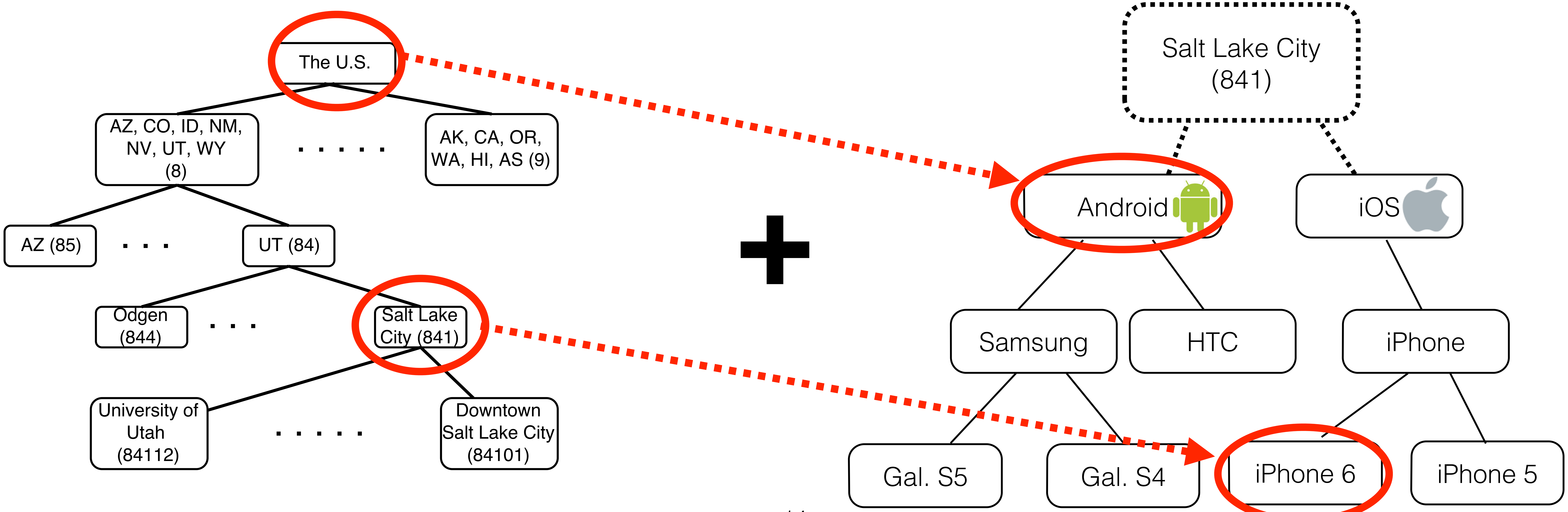
How to detect failures with different scopes?

- Group users based on their geographical information: ZIP code area, city, state.
- A user could belong to multiple geographical groups in the same time.
- Under each geographical group: further divided to device OS, make.



How to detect failures with different scopes?

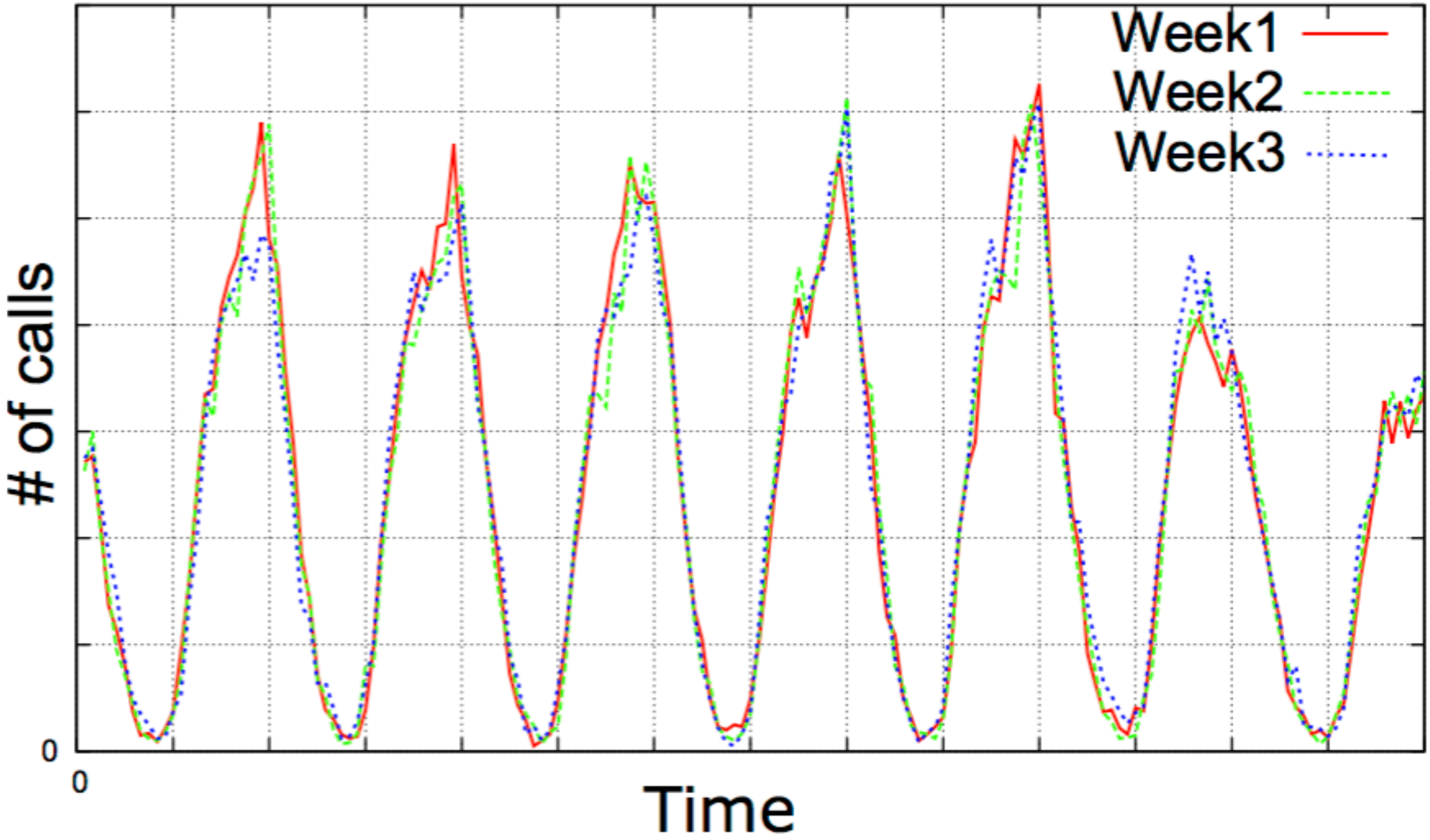
- Group users based on their geographical information: ZIP code area, city, state.
- A user could belong to multiple geographical groups in the same time.
- Under each geographical group: further divided to device OS, make.



Timestamp	Usage	State	City	Area	OS	Make	Model
2016/05/10 10:00	5000	Utah	Salt Lake City	U. Of Utah	Android	Samsung	Galaxy S6

Hierarchical attributes

Hierarchical attributes

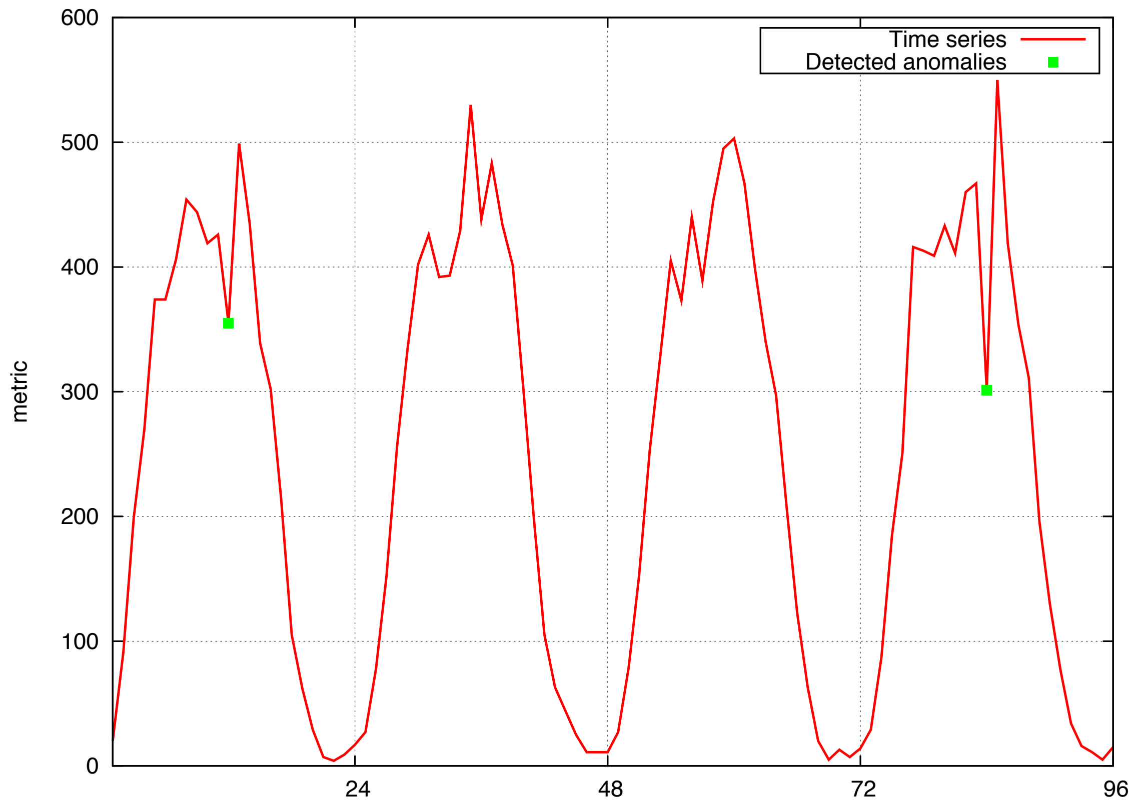


Need temporal aggregation to deal with sparse data during

Outline

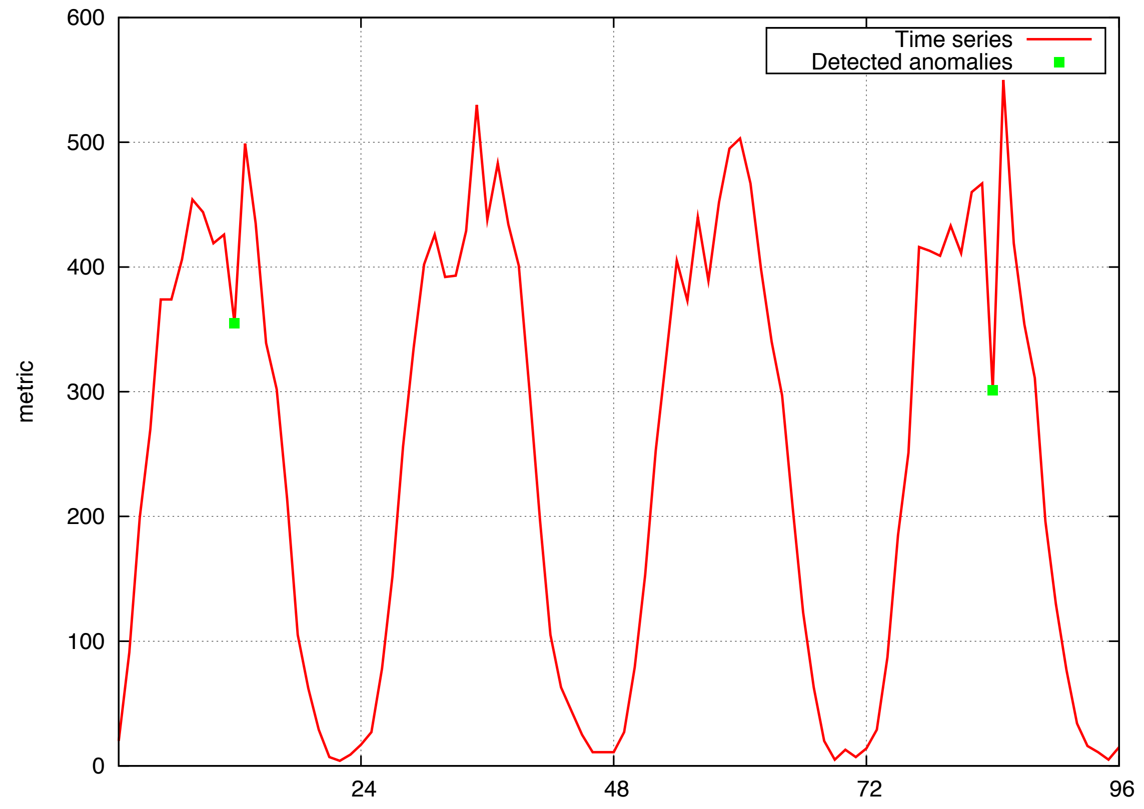
- Motivation.
- ABSENCE overview.
- Is ABSENCE feasible?.
- ABSENCE's challenges.
- **ABSENCE's event detection.**
- Synthetic workload evaluation.
- Operational validation.

Event detection algorithm

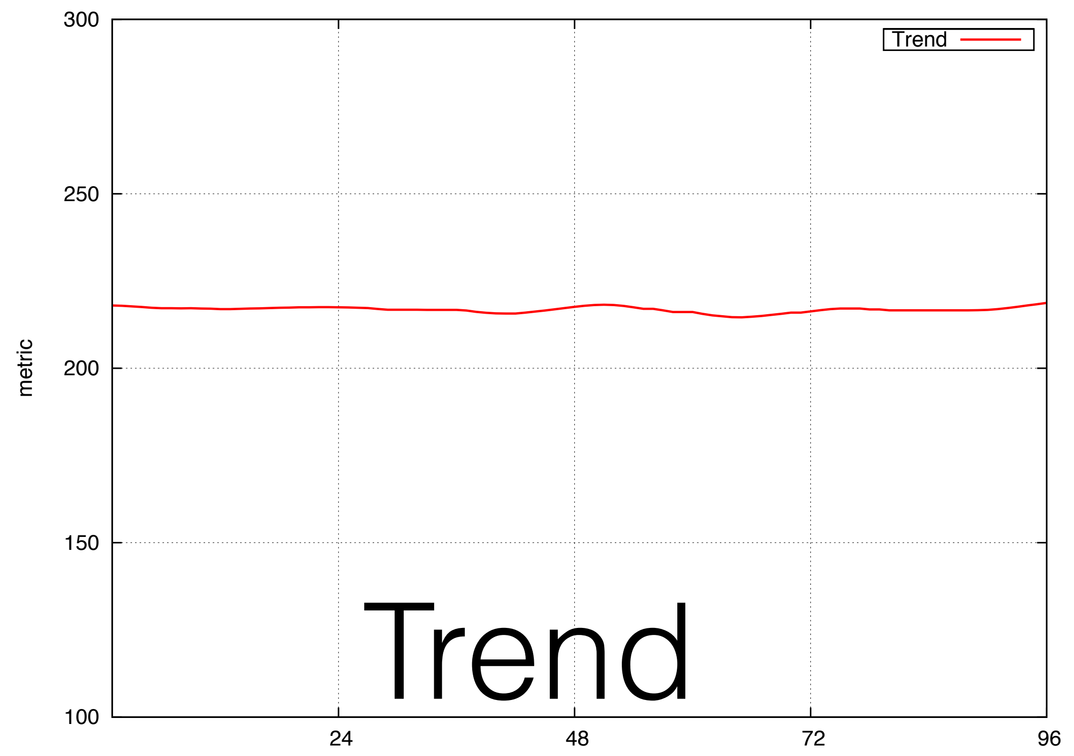


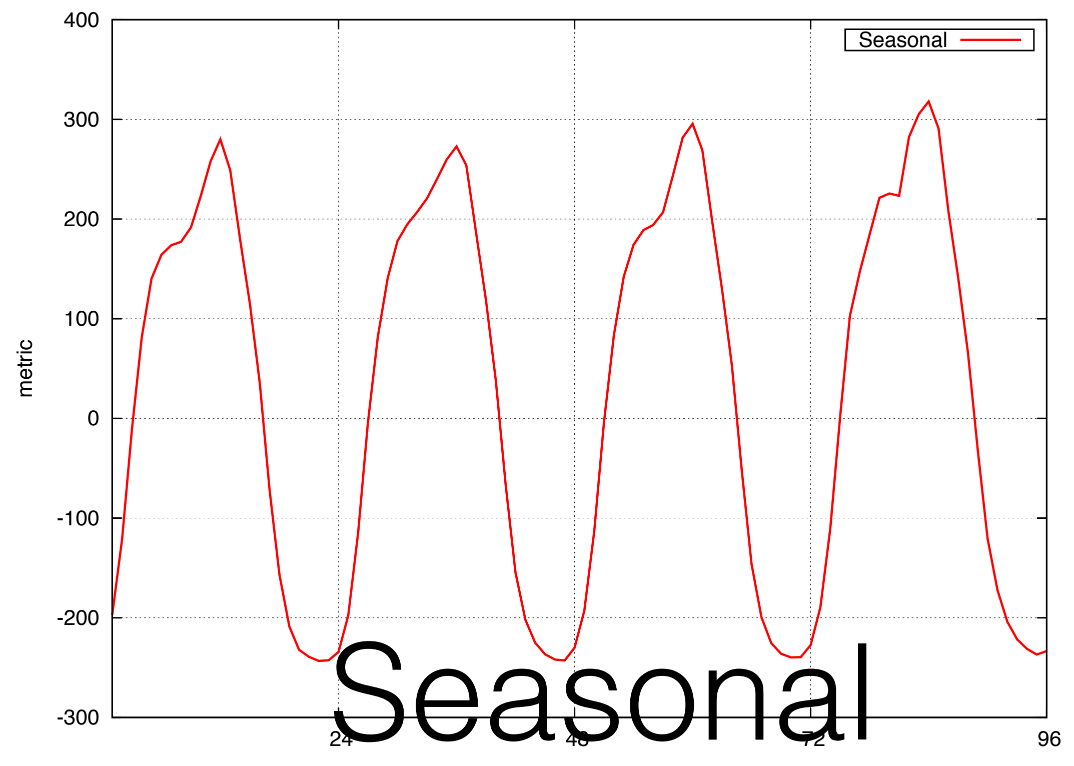
Usage's time series

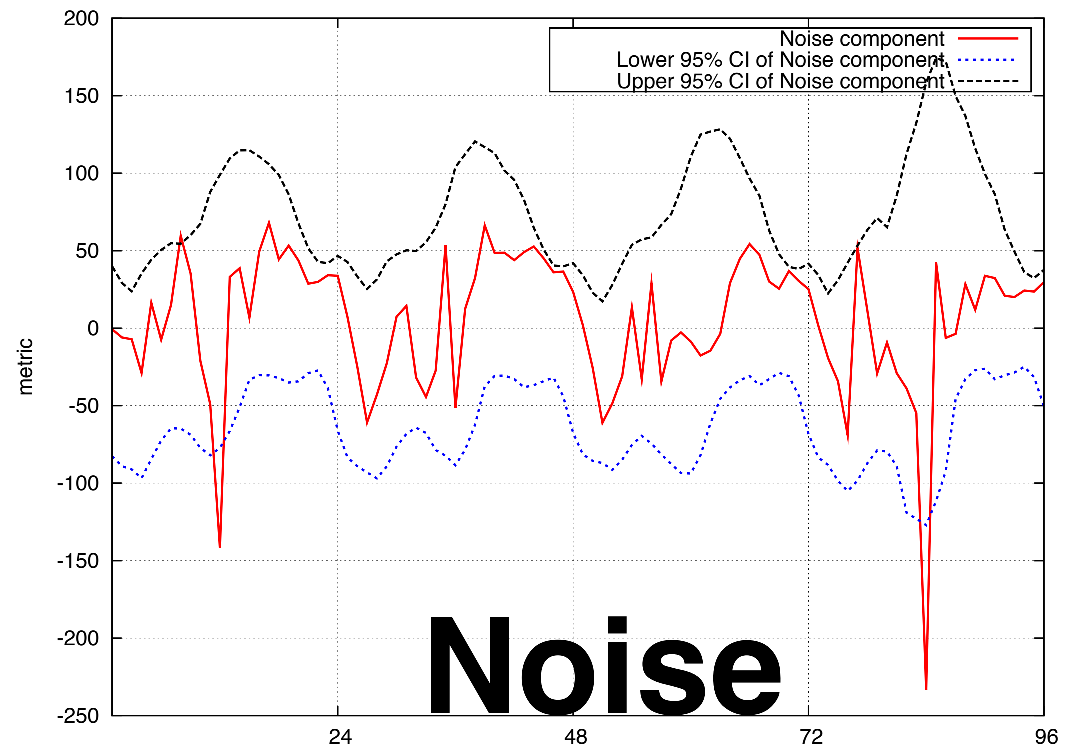
Event detection algorithm



Usage’s time series

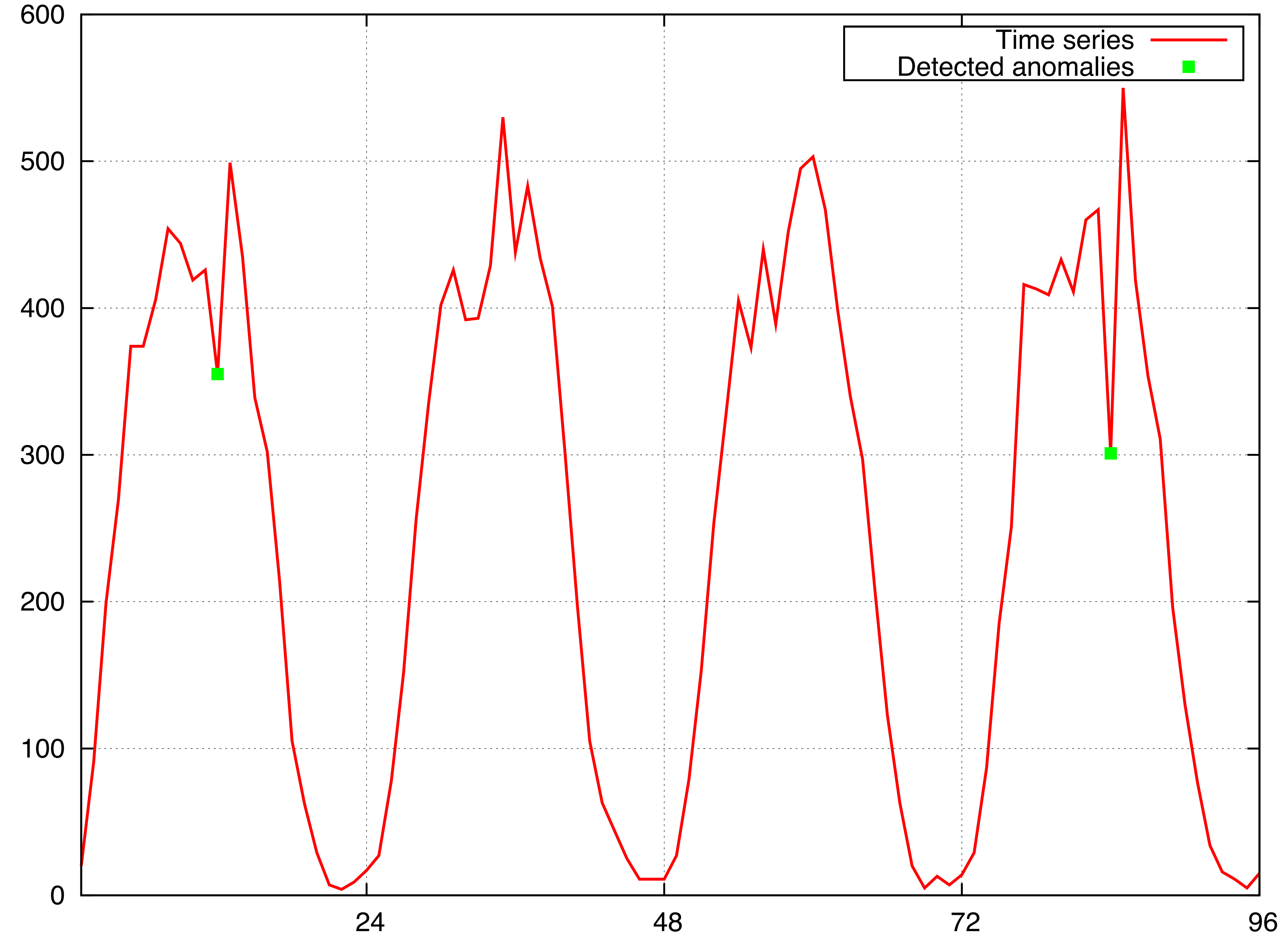




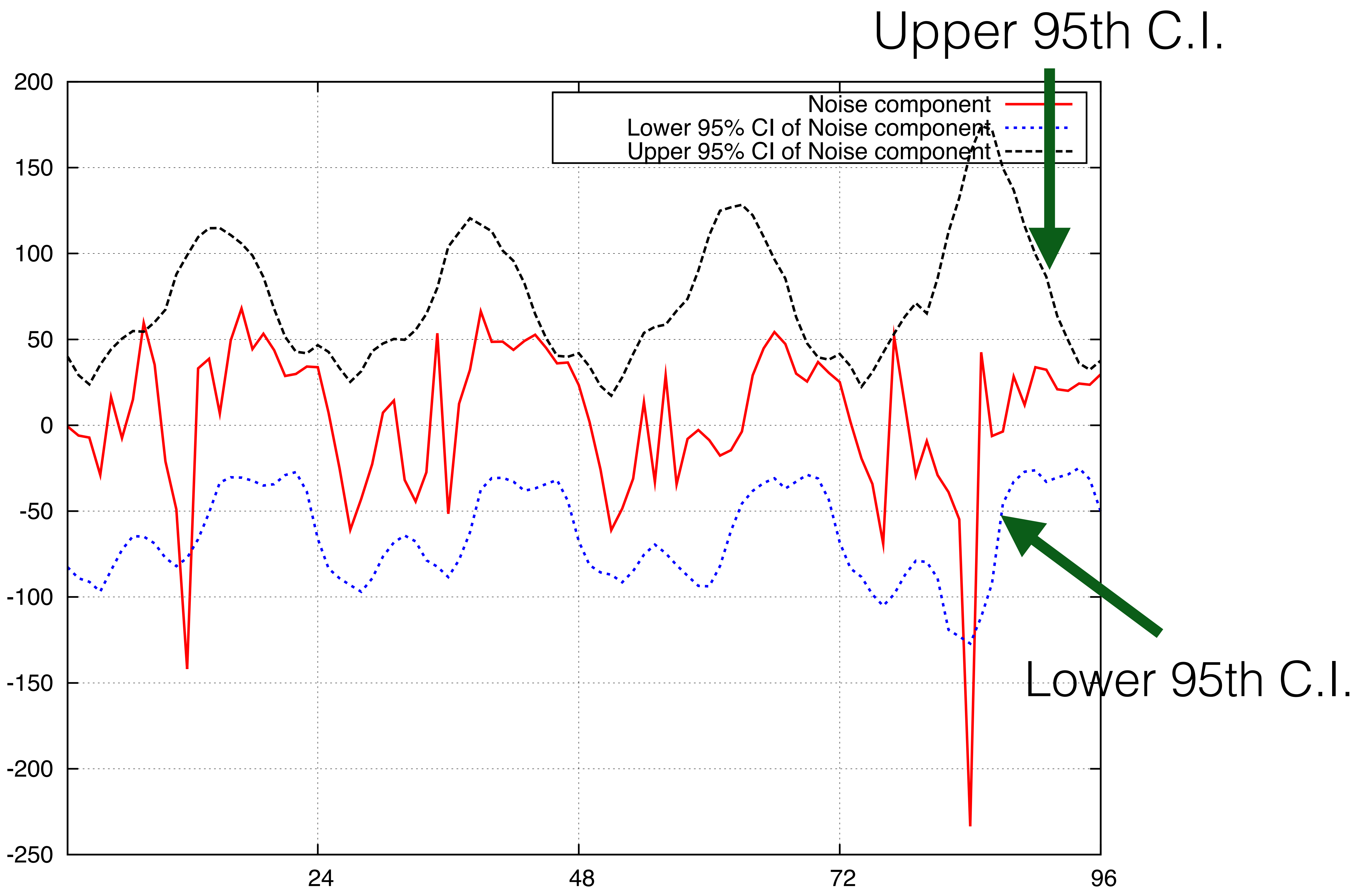


- Decompose time series: trend, seasonal, noise
- *Trend*: moving average.
- *Seasonal*: average of phasing values.
- ***Noise*** = *Time series - Trend - Seasonal*

Event detection algorithm



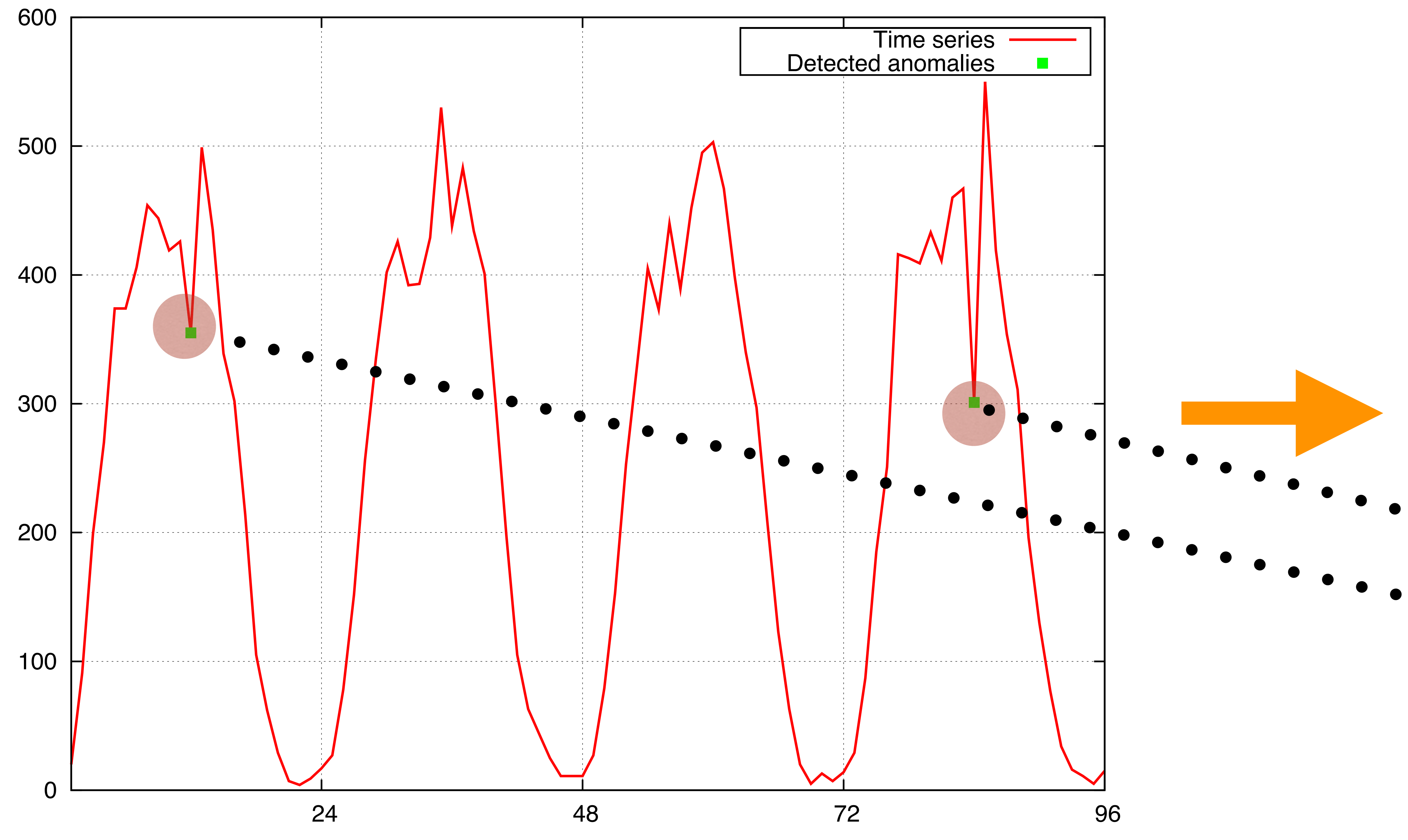
Usage's time series



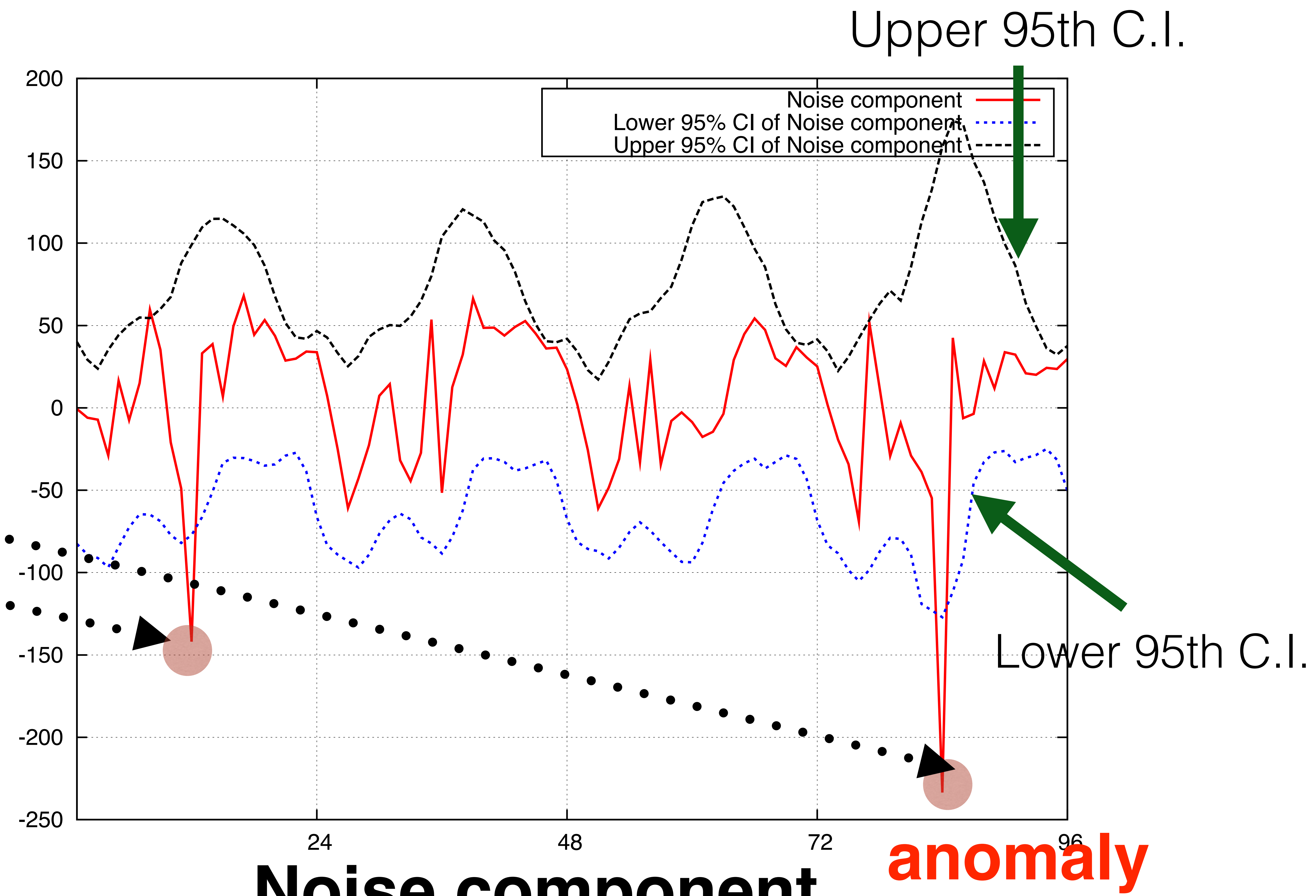
Noise component

Event detection algorithm

- If noise is out of the 95th percent Confidence Interval (CI) of noise component => anomaly.



Usage's time series



Noise component

Outline

- Motivation.
- ABSENCE overview.
- Is ABSENCE feasible?
- ABSENCE's challenges.
- ABSENCE's event detection.
- **Synthetic workload evaluation.**
- Operational validation.

Synthetic workload evaluation

- 6 months of **real CDR** from an U.S operator.
- Synthetically introduce failures:
  - Network failures: remove usage on base stations.
  - Device failures: remove usage on devices.

Parameters and metrics

Parameters

- **11,000 failures generated.**
- 100 ZIPs, 10 cities.
- Two popular device types.
- LTE/Voice.
- Duration: 1,2,3,6,12 hours.
- Quiet and busy hours.
- Impact degree: 0 - 55%.

(total usage reduction)/(total normal usage) for a given aggregation

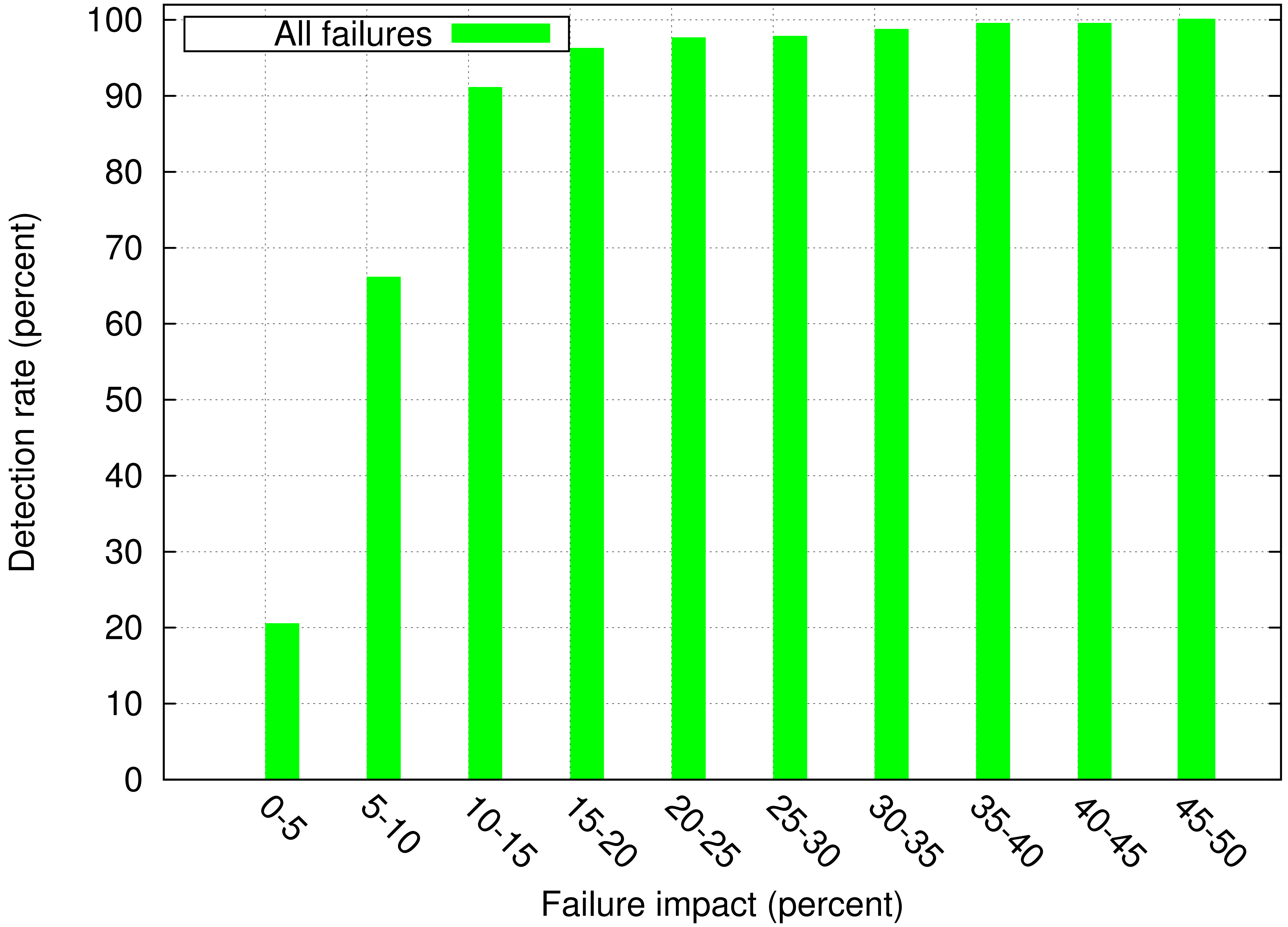
Metrics

- Detection rate = detected events/introduced events.
- Loss ratio = loss until detected/normal usage.

Example of failure scenarios:

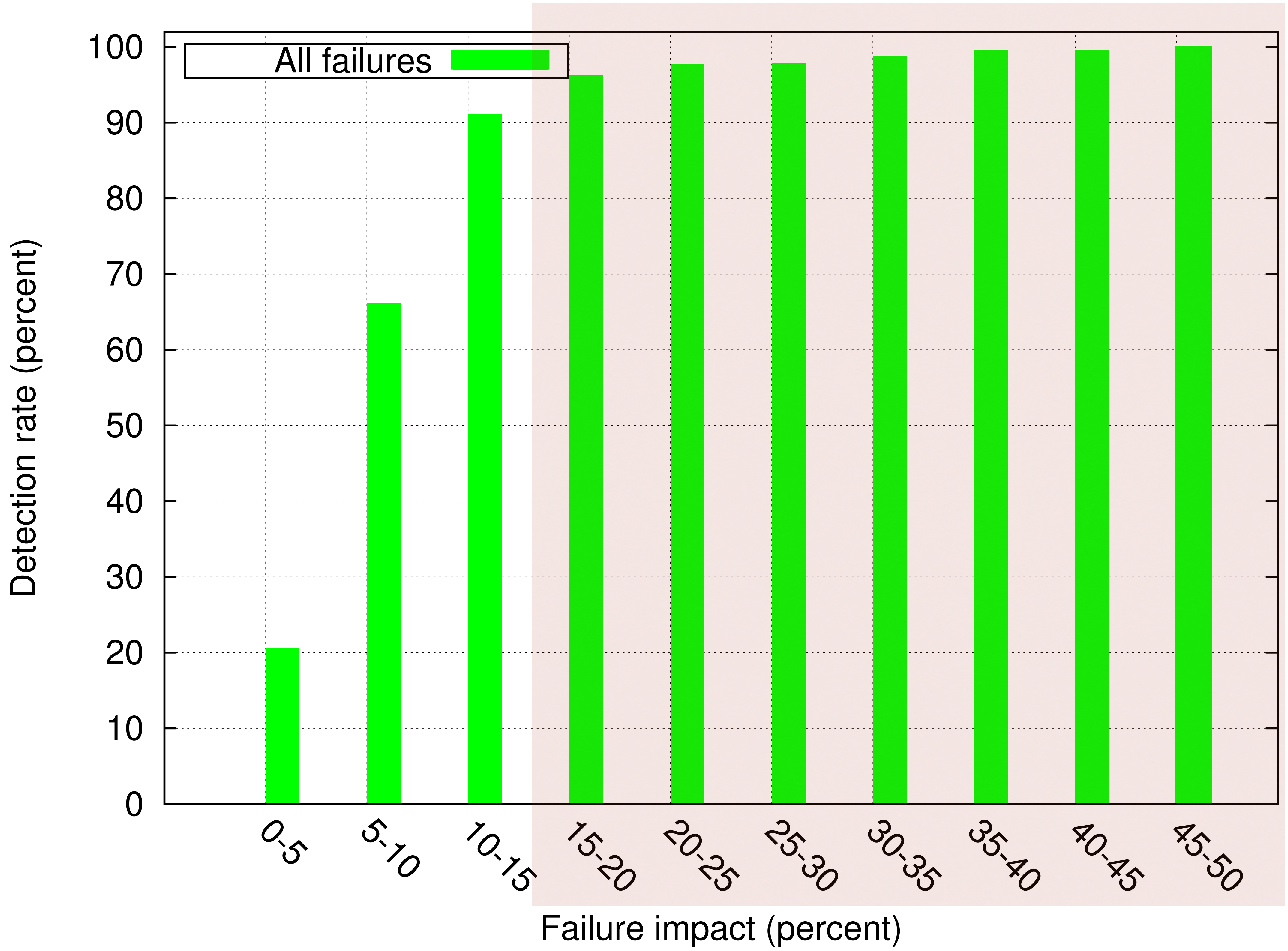
- All Android devices in Los Angeles fail.
- All Iphone5 devices in Downtown Los Angeles fail.

Overall detection rate



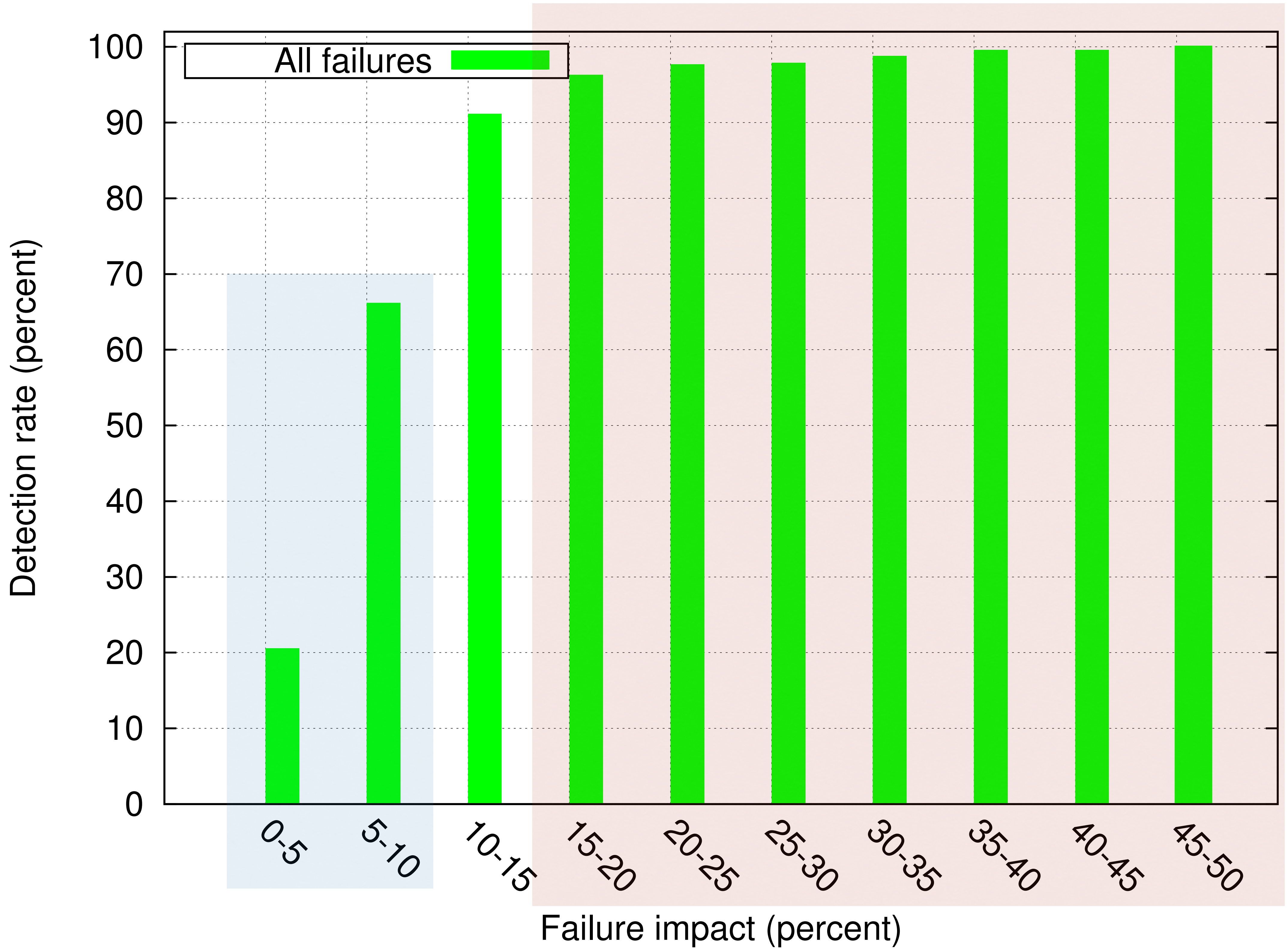
- With the 11,000 introduced failures:
  - ABSENCE detected >96% of failures that have more than 15% of impact.
  - ABSENCE tends to miss events that are <10% of impact.

Overall detection rate



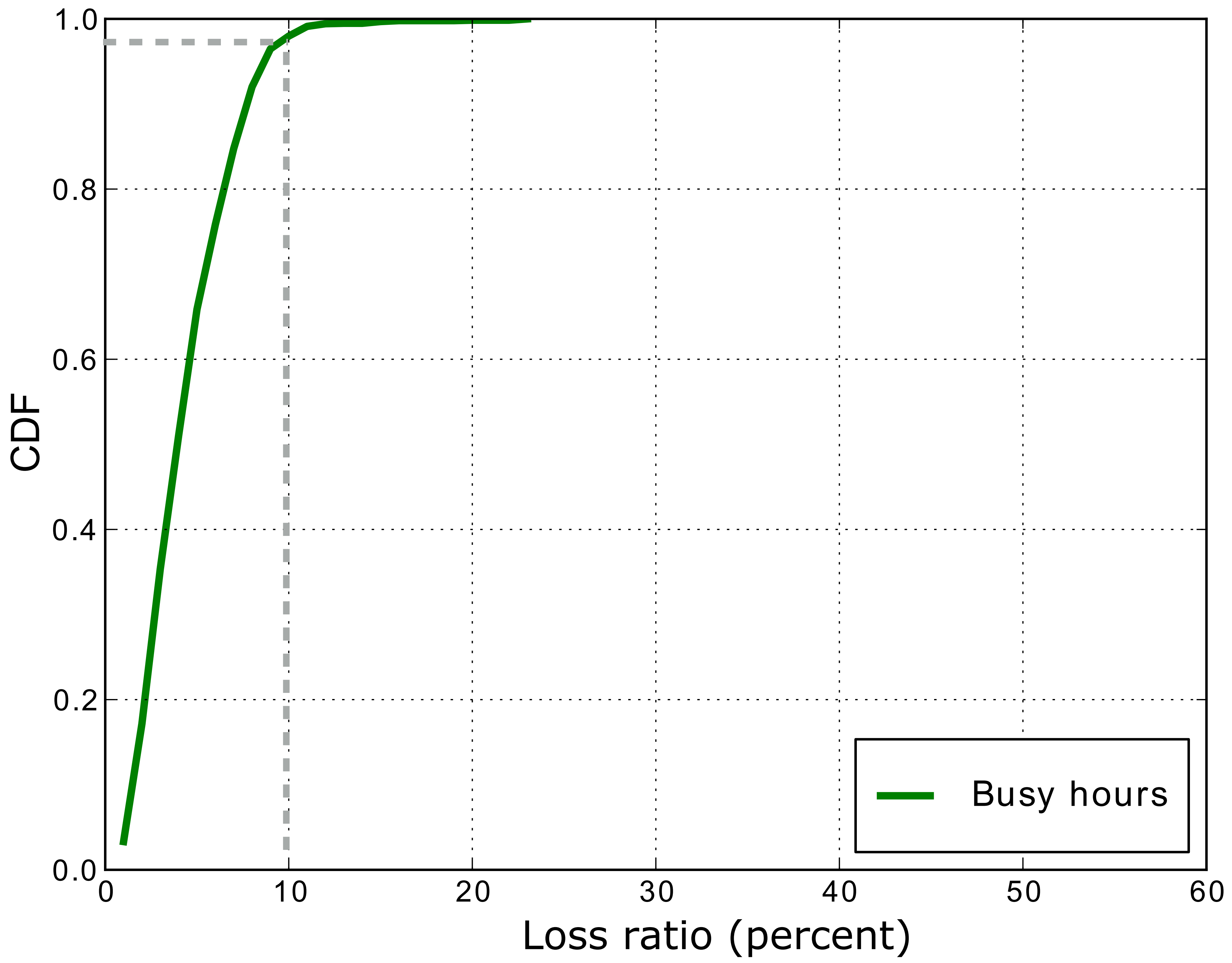
- With the 11,000 introduced failures:
  - ABSENCE detected >96% of failures that have more than 15% of impact.
  - ABSENCE tends to miss events that are <10% of impact.

Overall detection rate



- With the 11,000 introduced failures:
  - ABSENCE detected >96% of failures that have more than 15% of impact.
  - ABSENCE tends to miss events that are <10% of impact.

Loss ratio of detected failures



Loss Ratio= (usage loss until detection)/(normal usage during the failure period)

- All detected failures:
    - ~97% of them are detected when <10% of usage is lost (during busy hours).

Outline

- Motivation.
- ABSENCE overview.
- Is ABSENCE feasible?
- ABSENCE's challenges.
- ABSENCE event detection.
- Synthetic workload evaluation.
- **Operational validation.**

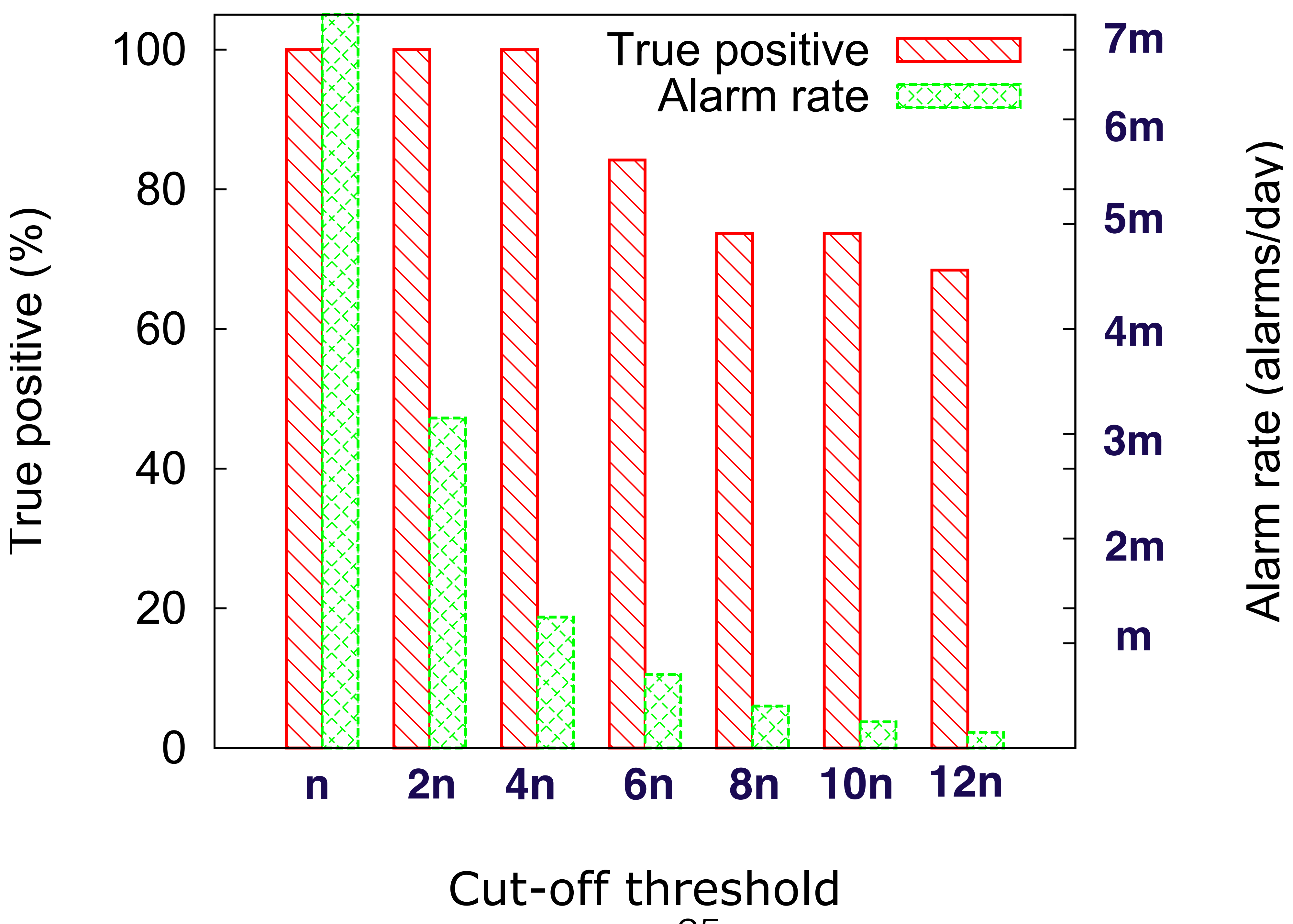
Evaluate against known silent failures from the operator

Evaluate against known silent failures from the operator

- 19 silent failure events: not known by the network operator when they happened.
- **Detected19/19, 100% true positive.**

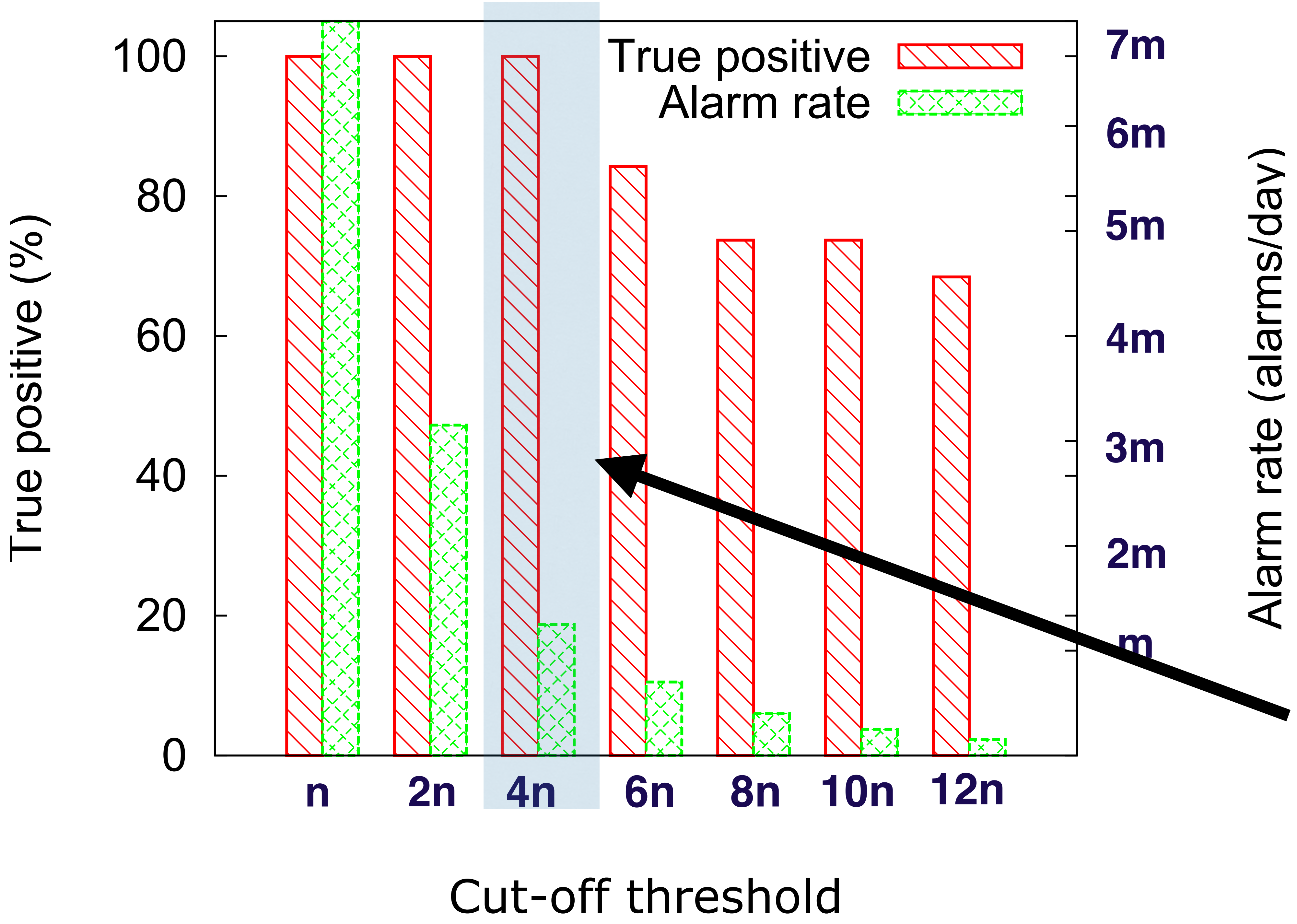
Alarm rate and true positive

- Use the 19 known events from operator.
- Alarm rate (**m**): average number of alarms per day that an operation team needs to handle.
- Cut-off threshold (**n**): filter out events that less impactful.
- Increase cut-off threshold could reduce alarm rate while maintaining true positive rate of ABSENCE



Alarm rate and true positive

- Use the 19 known events from operator.
- Alarm rate (**m**): average number of alarms per day that an operation team needs to handle.
- Cut-off threshold (**n**): filter out events that less impactful.
- Increase cut-off threshold could reduce alarm rate while maintaining true positive rate of ABSENCE

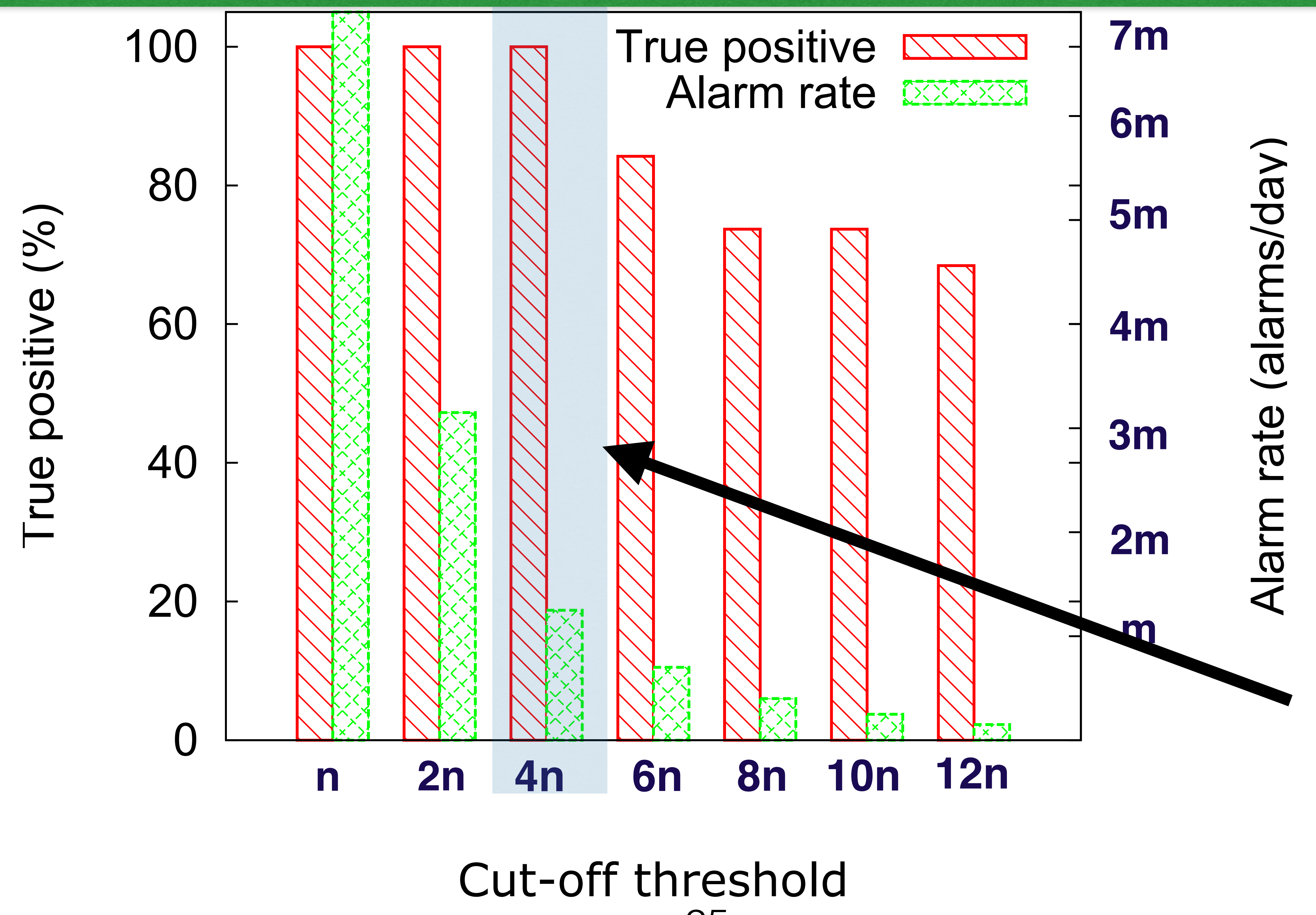


100% true positive,
m alarms per day
m is small enough

Alarm rate and true positive

- Use the 19 known events from operator.
- Alarm rate (**m**): average number of alarms per day that an operation team needs to handle.
- Cut-off threshold (**n**): filter out events that less impactful.
- Increase cut-off threshold could reduce alarm rate while maintaining true positive rate of ABS

ABSENCE's alarm rate is reasonable for practical!



100% true positive,
m alarms per day
m is small enough

Conclusions

- ***Absence of customer usage*** is a reliable indicator of service disruptions a mobile network.
- Appropriate grouping users results in predictable usage and high fidelity for anomaly detection.
- Synthetic evaluation and operational validation.
- Practical in an operational environment.

Thank you!

ABSENCE: Usage-based Failure Detection in Mobile Networks

Binh Nguyen, Zihui Ge, Jacobus Van der Merwe, He Yan, Jennifer Yates
Mobicom 2015

THE UNIVERSITY OF UTAH

AT&T Labs Research
The Science Behind The Service