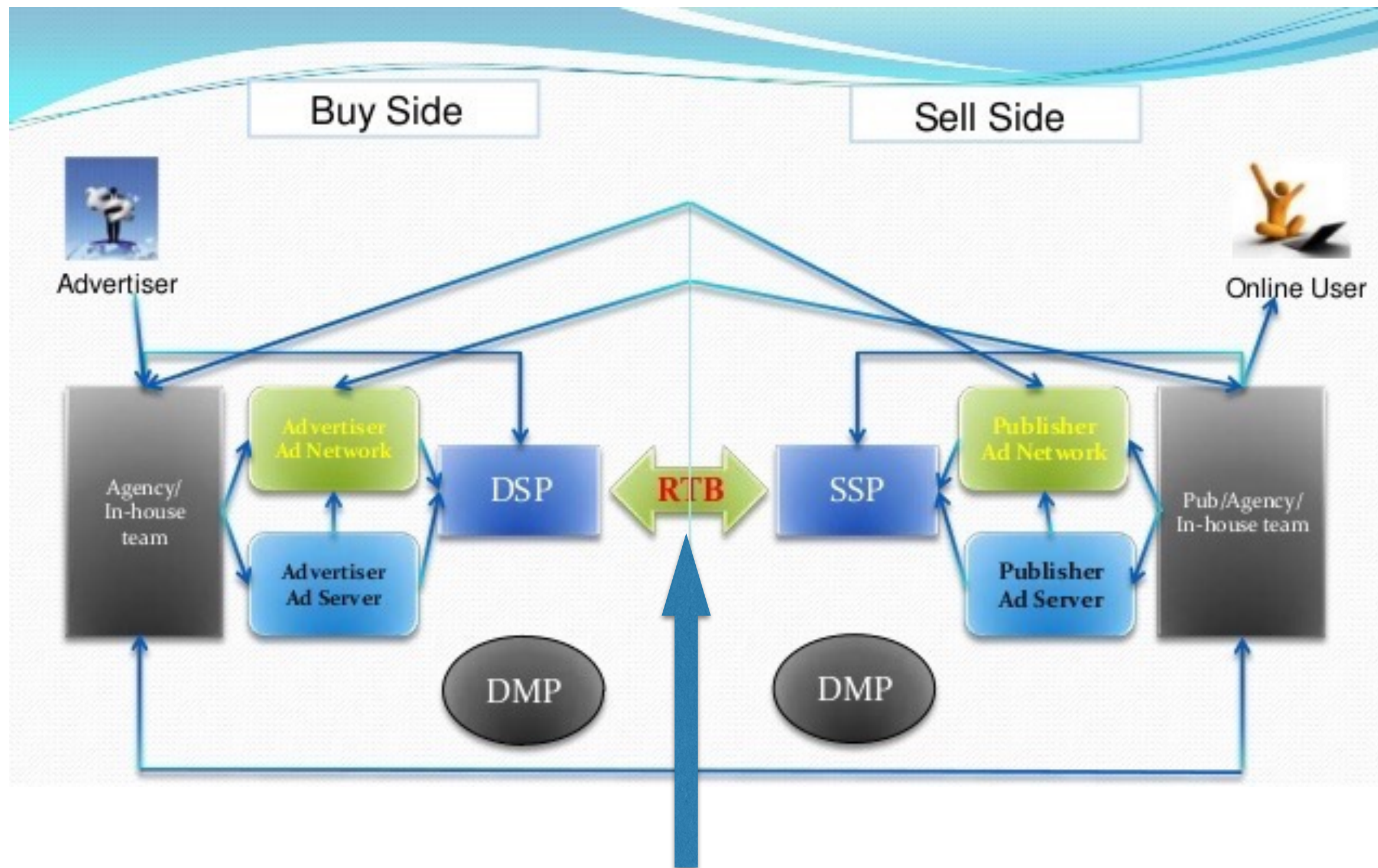
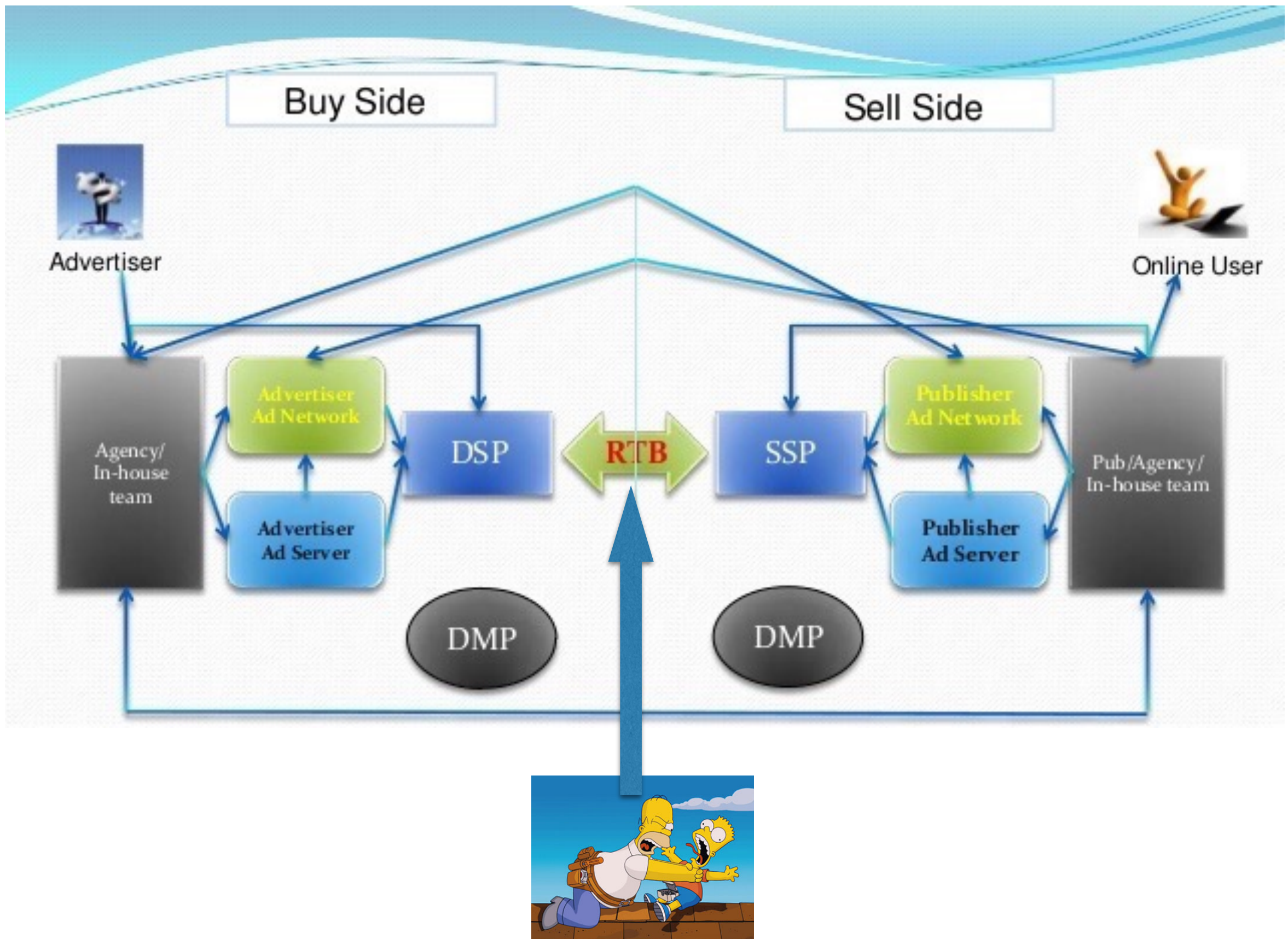


Throttling bid requests





Wires, network, \$\$\$



Model

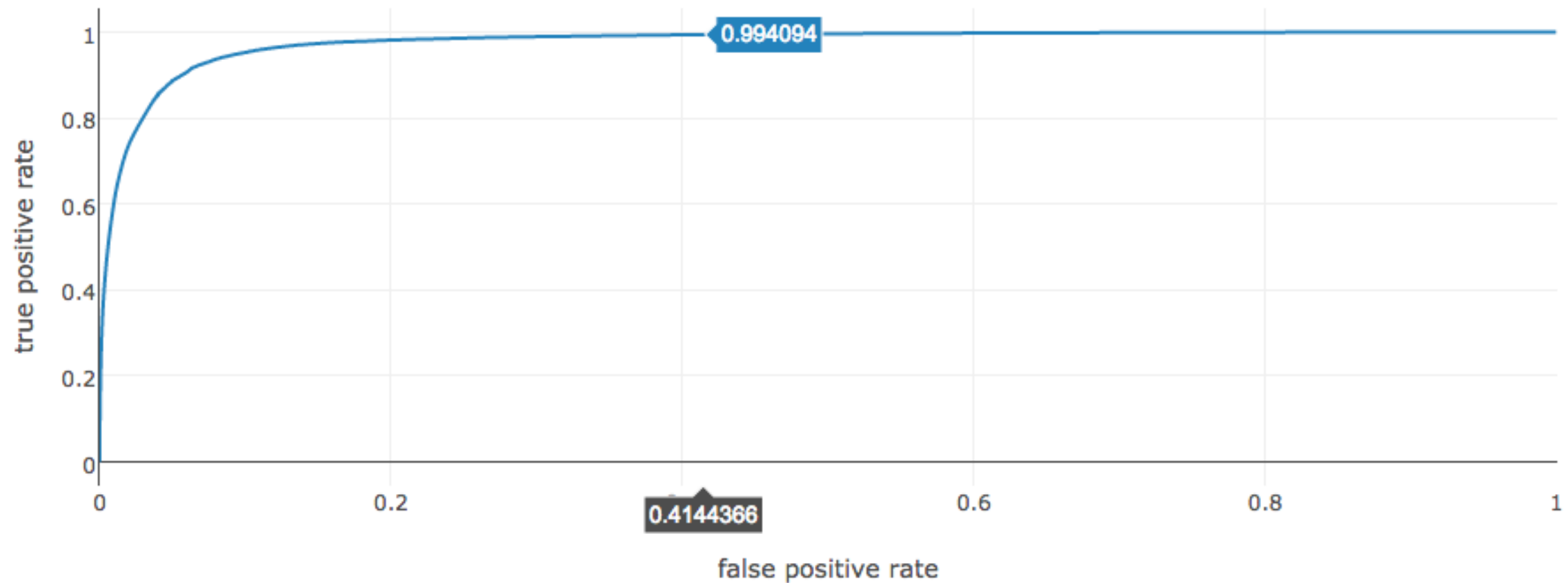
- Logistic regression
 - Not very sexy but...
 - Fast and reliable in online use, easy to optimize offline
 - Features related to: the user, the DSP, the publisher

Model

- Trained every hour on 3 hours of data with Spark on EMR
- Fully separated pipelines EU/US/APAC
- 10% exploration

Offline results

ROC curve

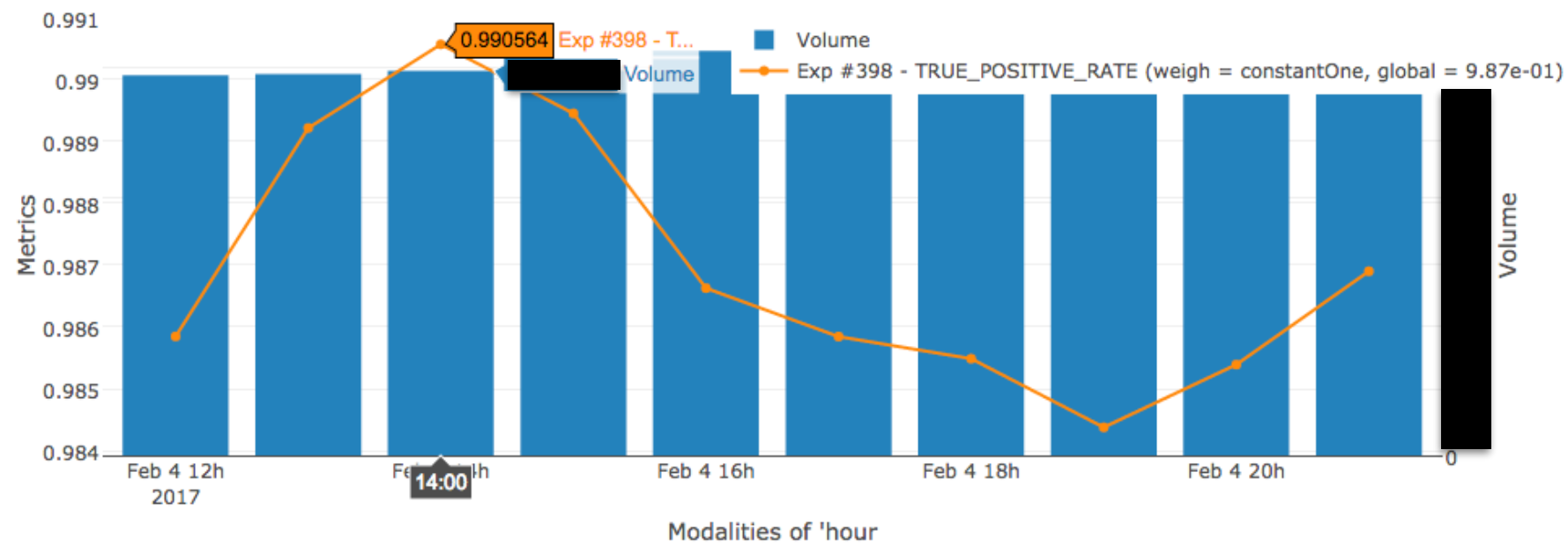


True positive rate

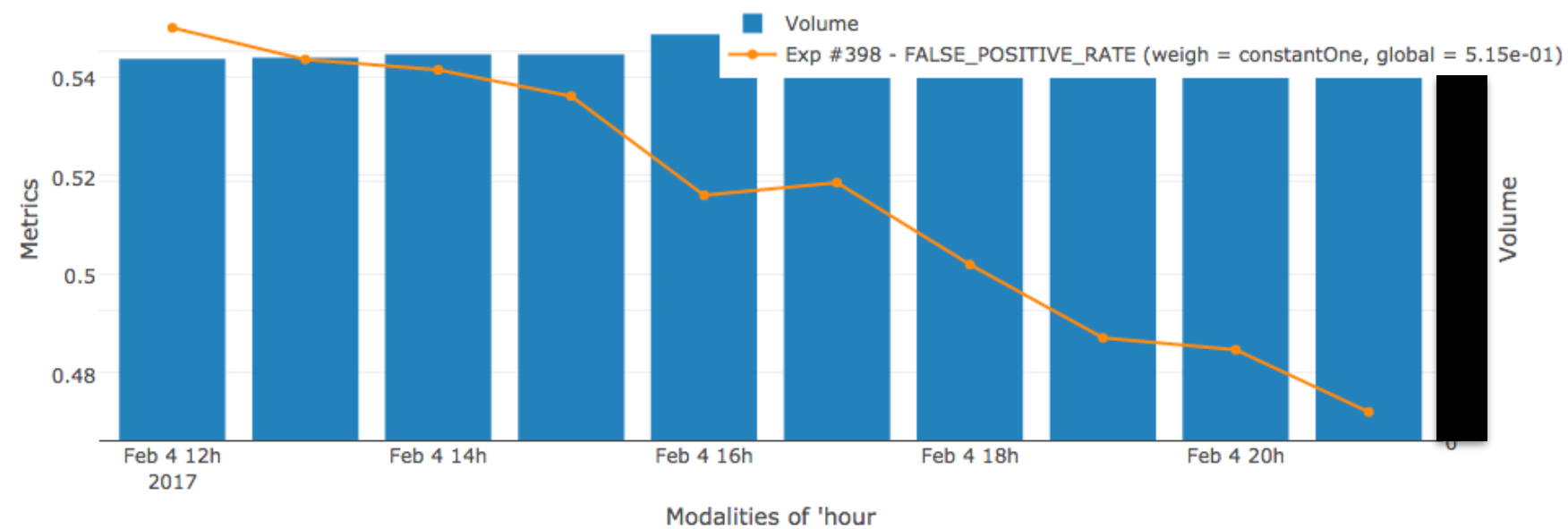
False positive rate (= 1 - True negative rate)

Offline results

TRUE_POSITIVE_RATE (weighted by constantOne)



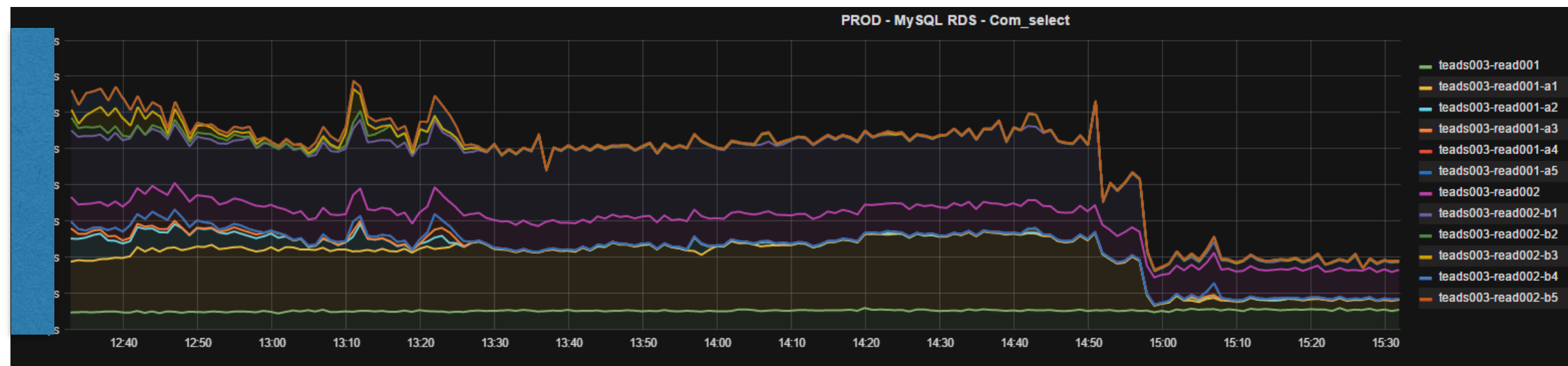
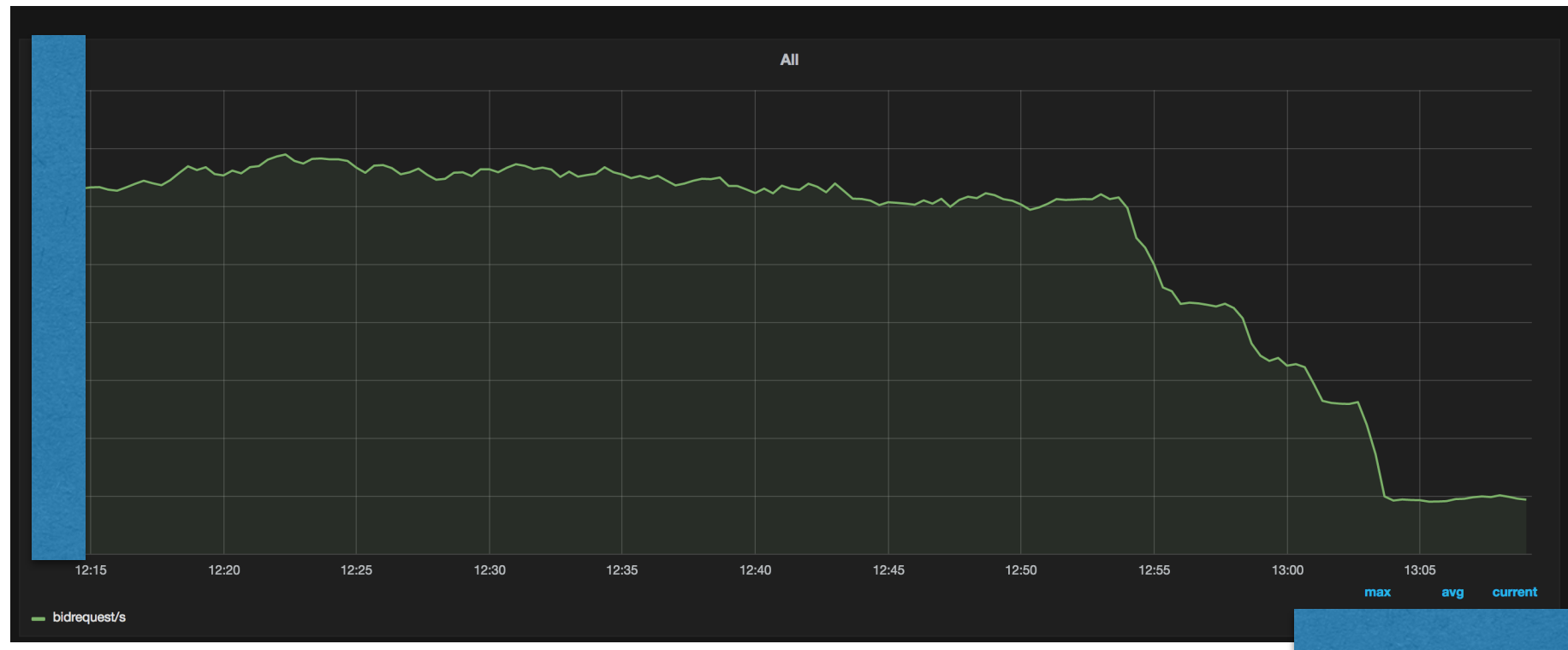
FALSE_POSITIVE_RATE (weighted by constantOne)



Online results

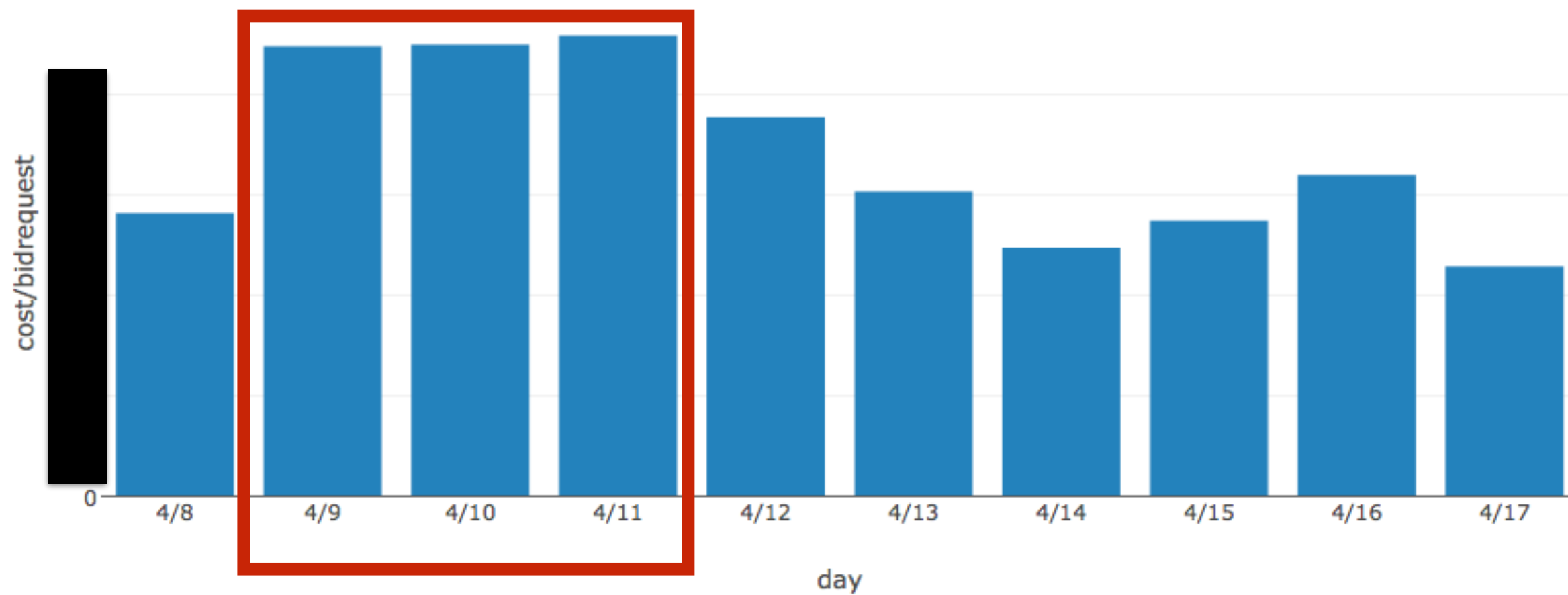


Impact on infrastructure



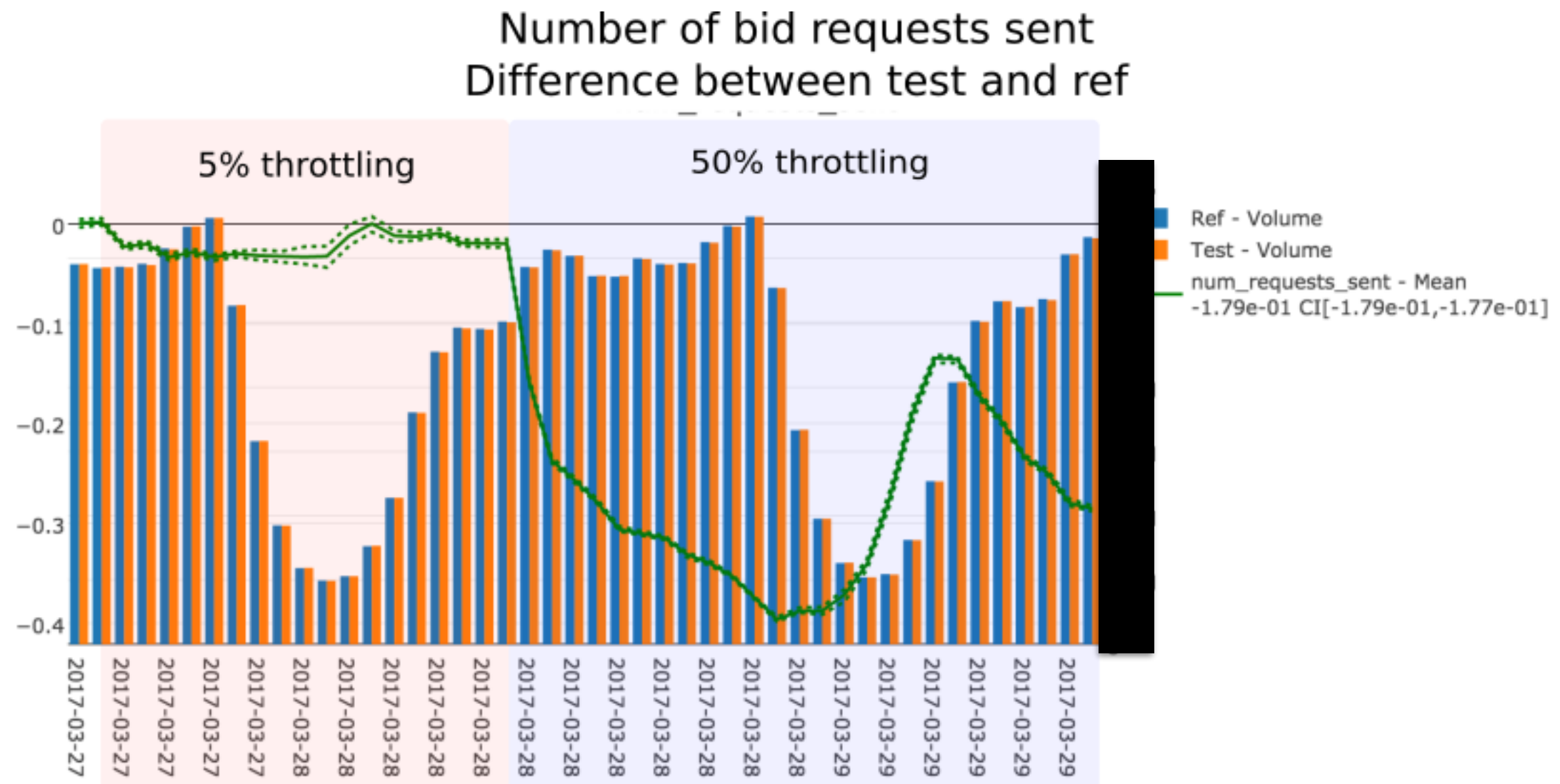
\$\$\$

cost/million bidrequest US



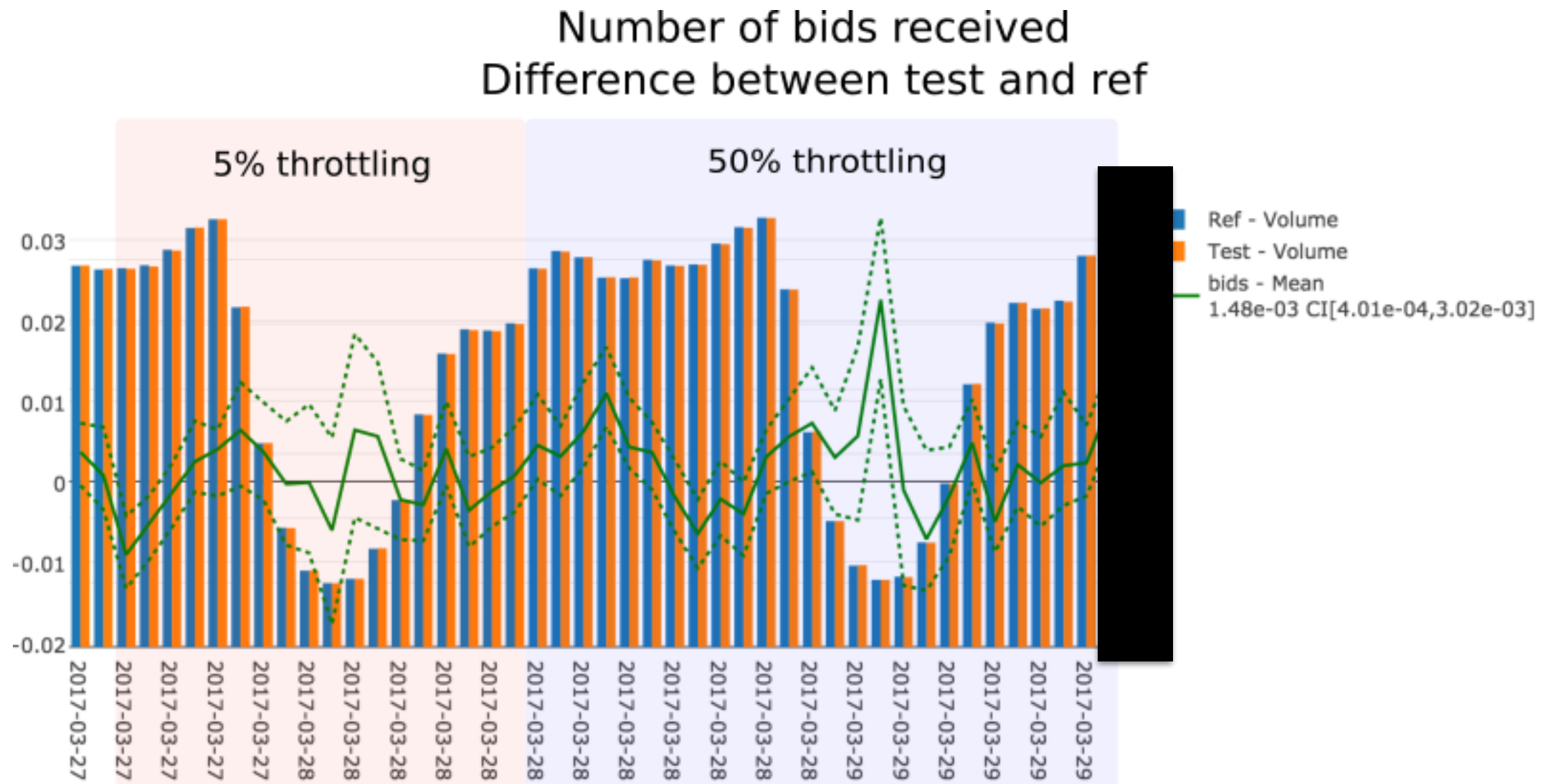
Impact on business

AB test



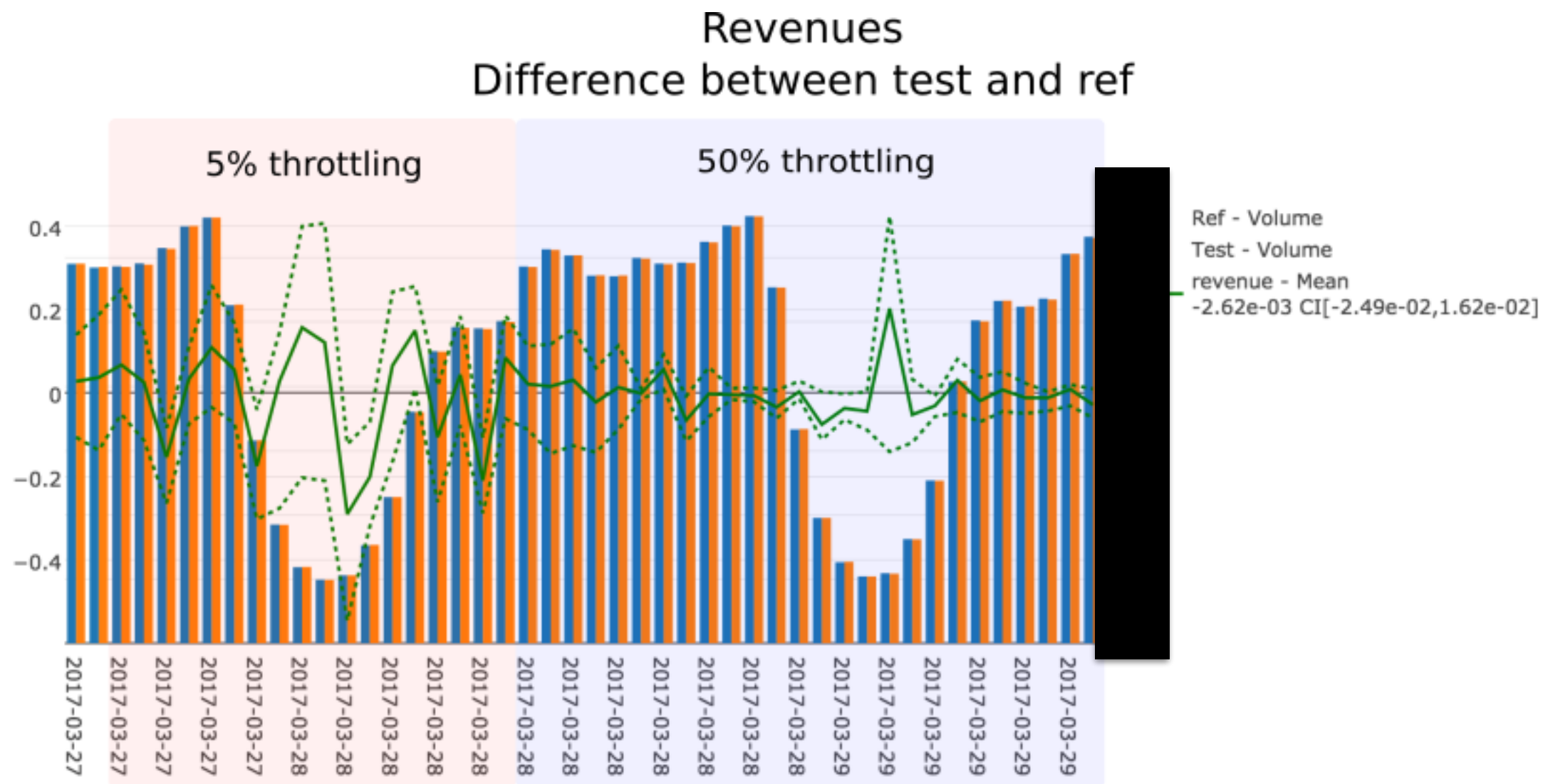
Impact on business

AB test



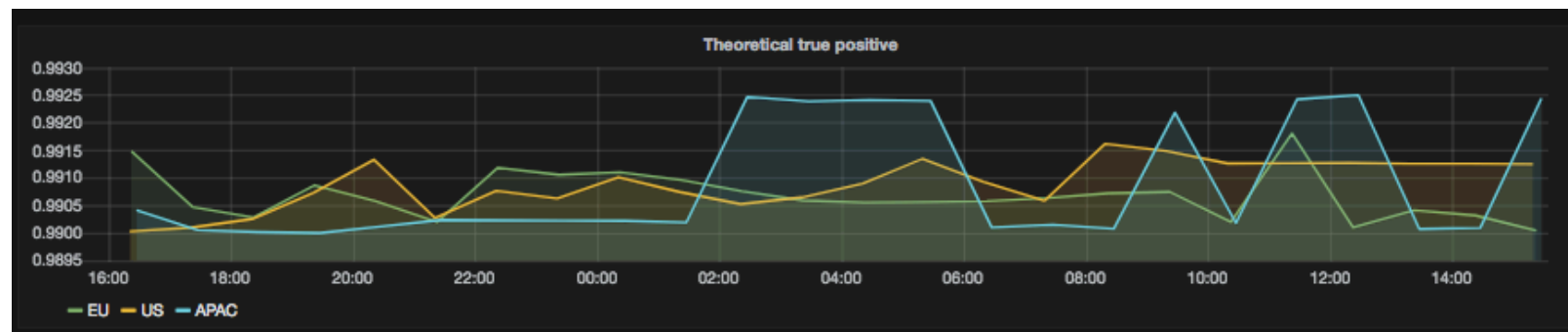
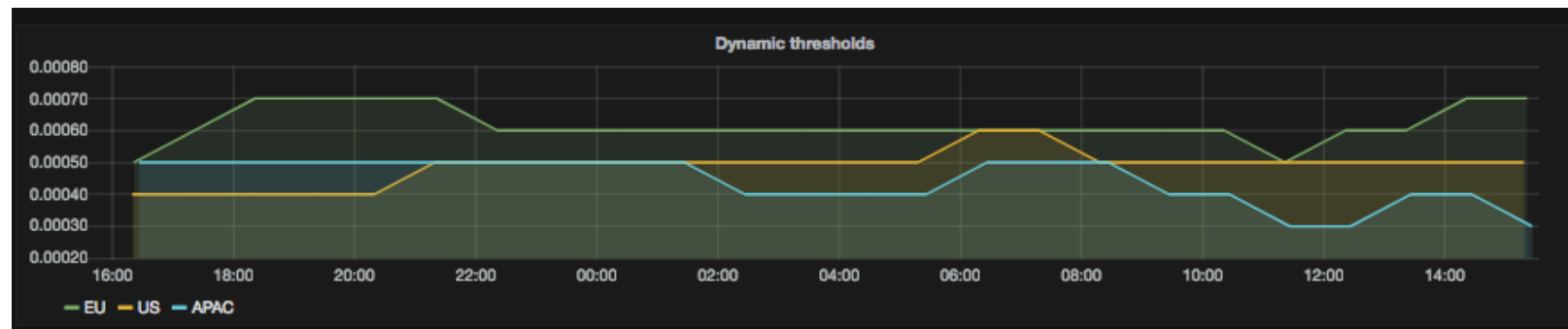
Impact on business

AB test



Dynamic threshold

- The model is good but how can we determine and adjust the threshold?
- Fix true positive rate target, set in db
- Hourly job to compute optimal threshold over the last x hours to reach target.



Thank you

Questions?