CAPACITY PREDICTION IN EXTERNAL SERVICES

Jerome F. Kraus, Senior Site Reliability Engineer

Alaska AIRLINES

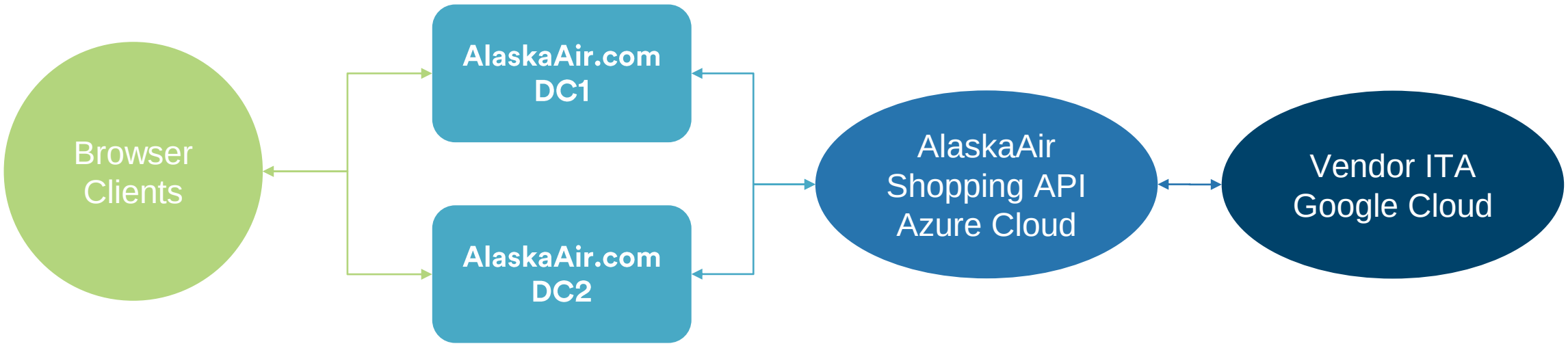
Alaska Airlines & AlaskaAir.com

- After 85 years, Alaska Airlines has grown to be the 5th largest US airline
- First airline in North America to sell tickets online
- First in the world to allow customers to check in and print boarding passes via the Internet
- Award winning Customer Service, Mileage Plan™ program and web site, AlaskaAir.com
- AlaskaAir.com averages 2.4 million page views per day, with peaks over 4 million
- AlaskaAir.com sends shopping vendor ITA an average of 10 million shopping requests per day, with peaks over 14 million

AlaskaAir.com and ITA

- Alaska Airlines files fares and pricing with ATPCO
- Alaska Airlines uses SABRE to book reservations and track inventory
- ITA combines inventory, schedules, and fare pricing data to provide shopping results to AlaskaAir.com
- AlaskaAir.com was an early adopter of ITA and their launch customer for ReShop

Data Flow


Browser
Clients
AlaskaAir.com
DC1
AlaskaAir.com
DC2
AlaskaAir
Shopping API
Azure Cloud
Vendor ITA
Google Cloud

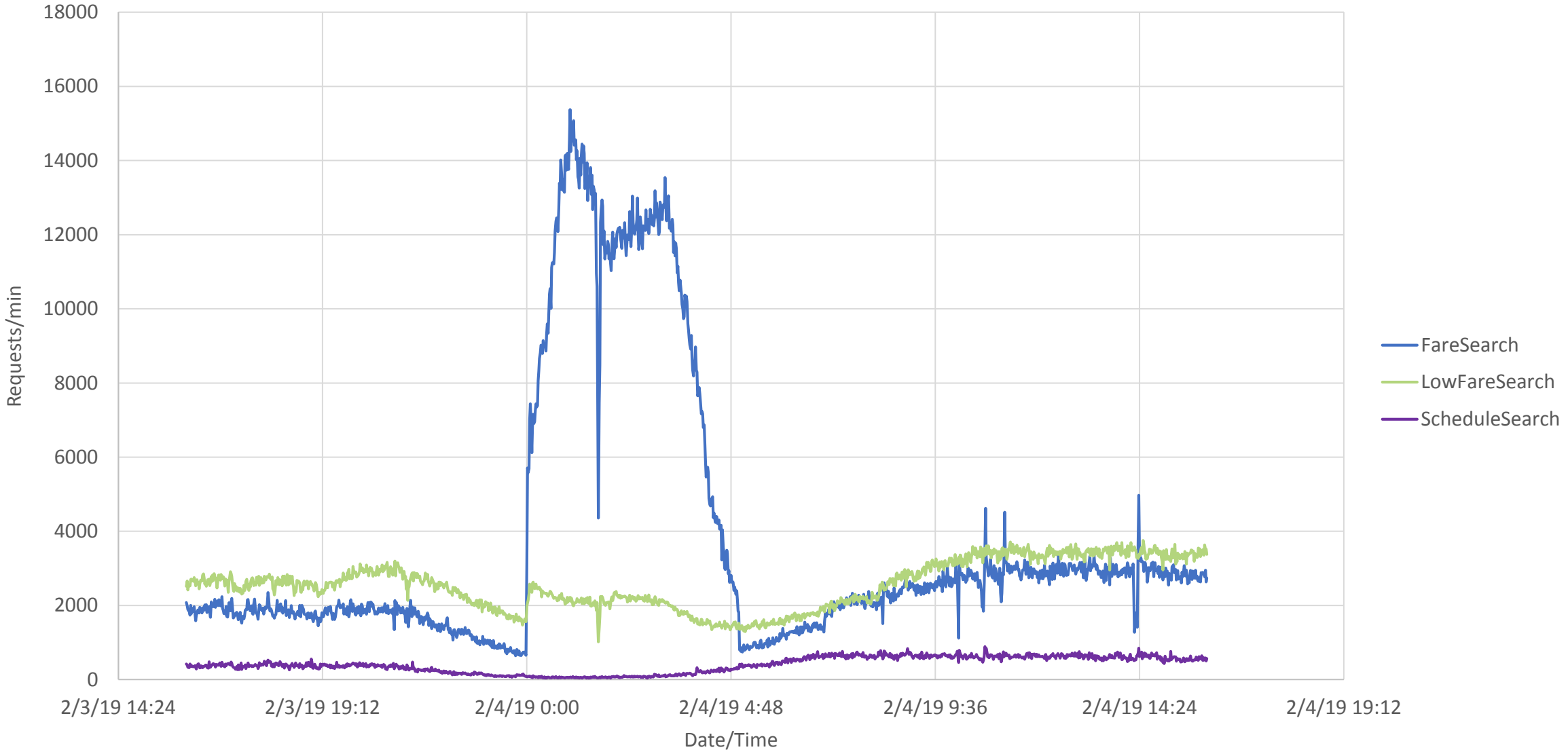
AlaskaAir.com Shopping Engine API

- In Azure cloud Service Fabric
- Receives shopping requests from AlaskaAir.com
- Orchestrates requests to vendor ITA in the Google cloud
- Builds responses and returns to AlaskaAir.com

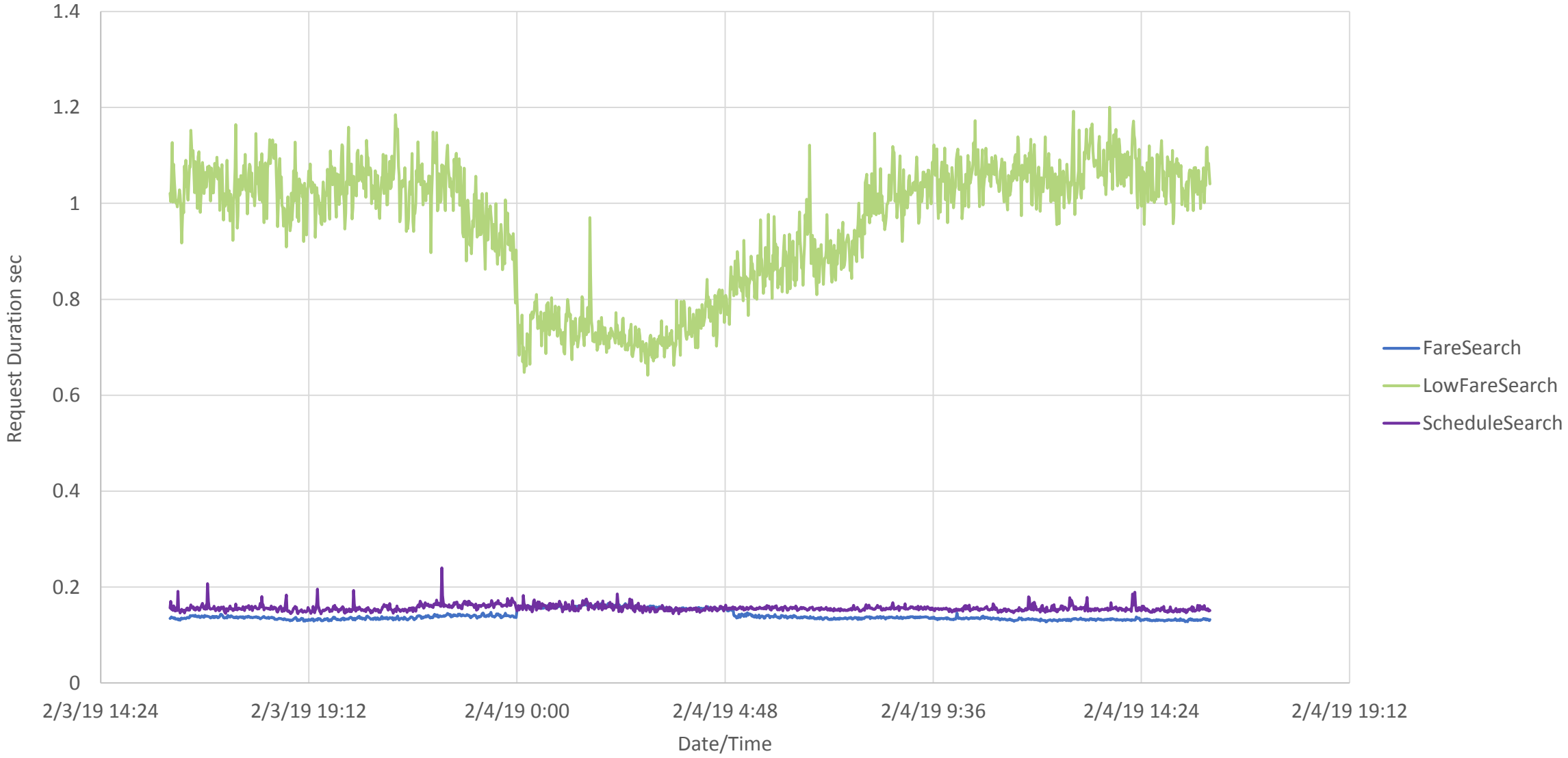
ITA Shopping Data Source Service

- Receives multiple types of Shopping Requests
  - Fare Search
  - Low Fare Search
  - Schedule Search
- Utilization of resources is controlled by Request Rate and Duration
- Different request types have different Rates and Durations

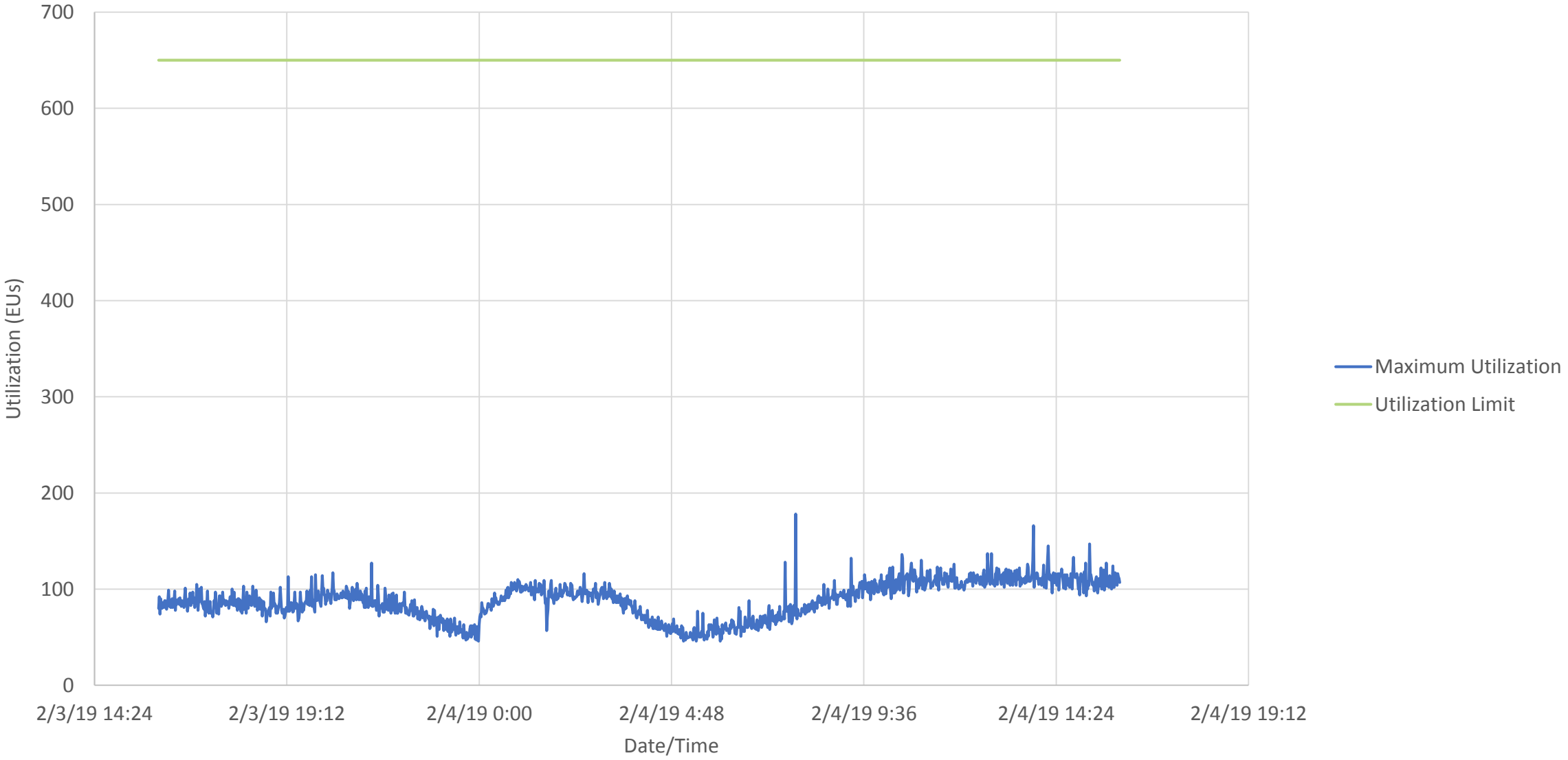
Shopping Request Rates



Shopping Request Duration



Utilization Capacity



Linear Regression Analysis

- Maximum utilization is a function of Request Rate and Duration
- There is a linear relationship between Utilization and Request Rate multiplied by Duration (Load Factor)
- The linear relationship can be modelled mathematically using linear regression analysis
- The mathematic model can be used to predict utilization based on expected query rates and durations
- Regression analysis was performed on actual Utilization and total Load Factor
- Total Load Factor is the summation of Request Rate multiplied by Duration for the three request types

Linear Regression Analysis Results

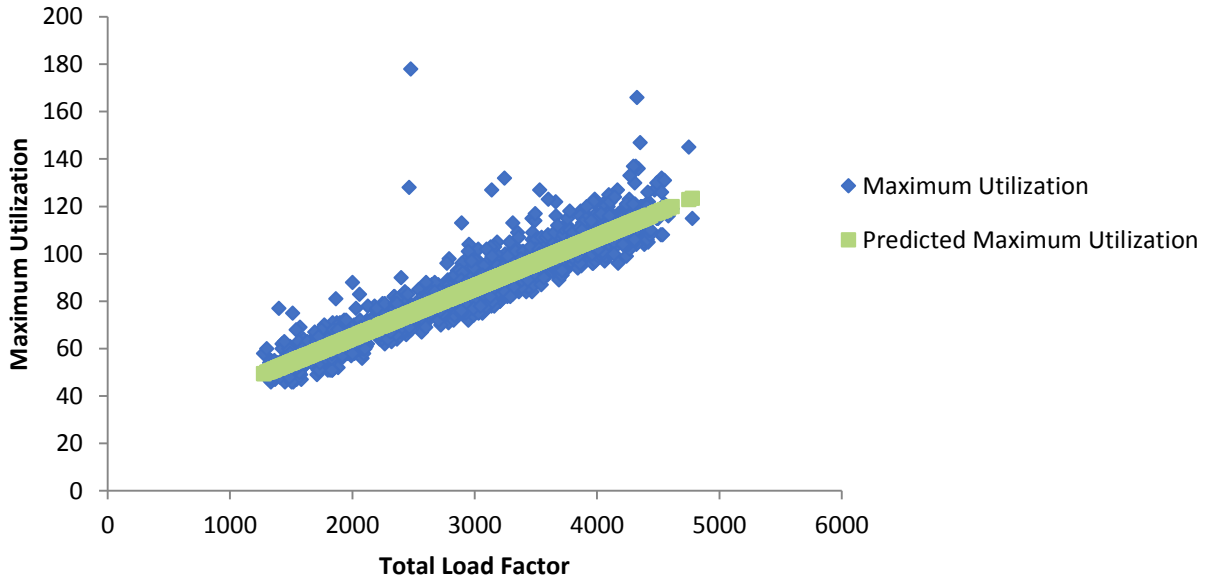
Regression Statistics

Multiple R	0.9262
R Square	0.8578
Adjusted R Square	0.8577
Standard Error	7.1923
Observations	1440

Coefficients

Intercept	22.486
Total Load Factor	0.021

Total Load Factor Line Fit Plot



$$y = 0.021x + 22.486 + 7.192$$

coefficient *intercept* *standard error*

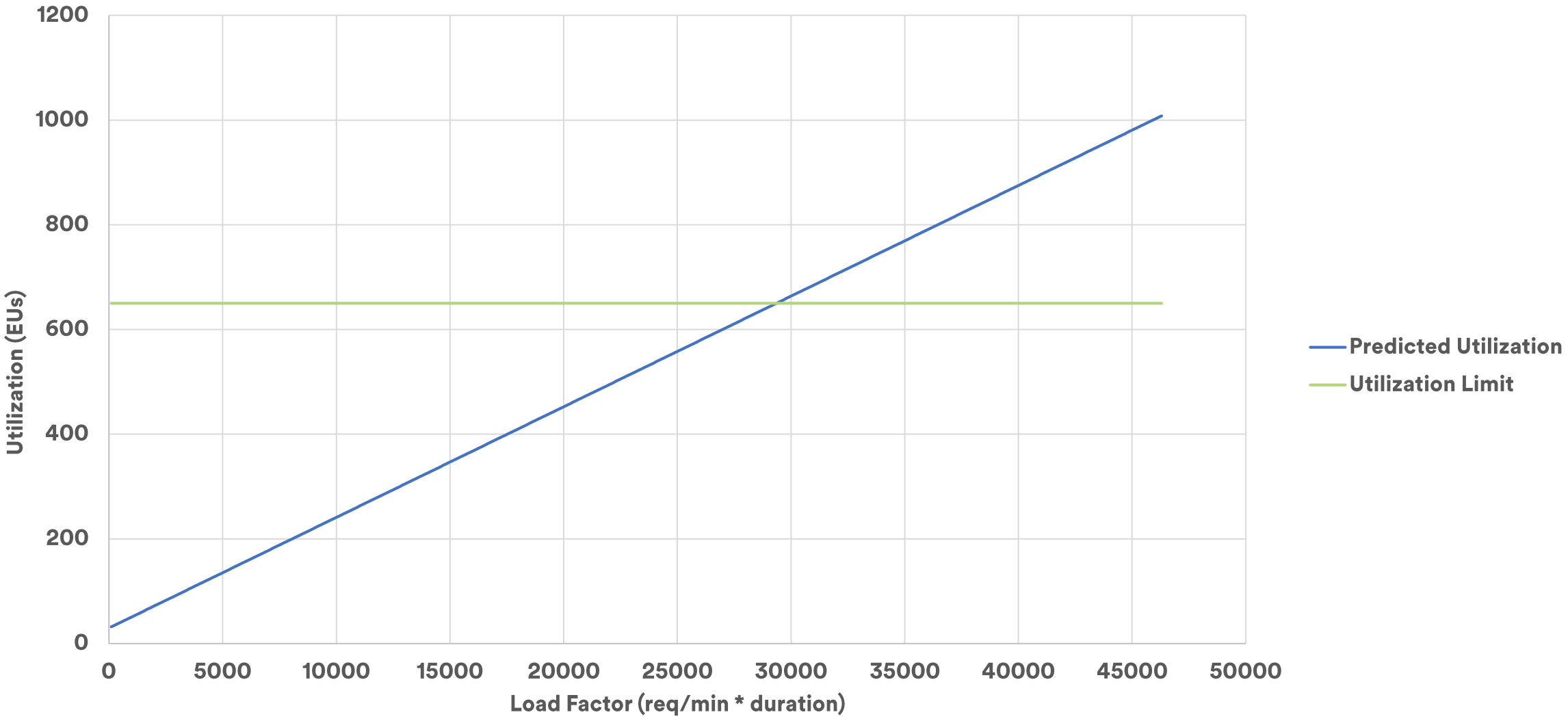
Predictive Model

$$y = 0.021x + 22.486 + 7.192$$

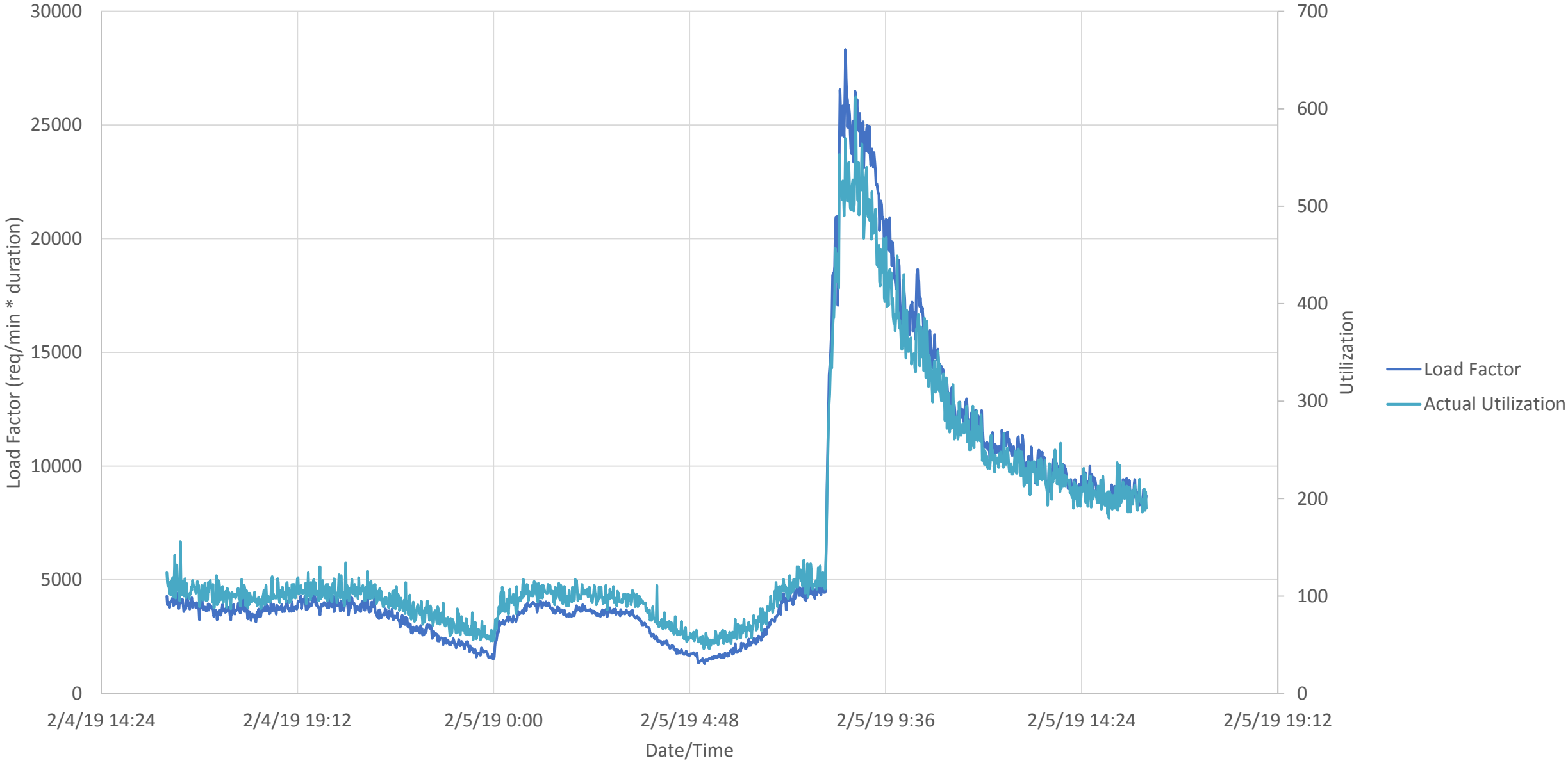
$y = Utilization$
$x = Load\ Factor = (fsr * fsd) + (lfr * lfd) + (sfr * sfd)$

fsr = FareSearch Rate
fsd = FareSearch Duration
lfr = LowFareSearch Rate
lfd = LowFareSearch Duration
sfr = ScheduleFareSearch Rate
sfd = ScheduleFareSearch Duration

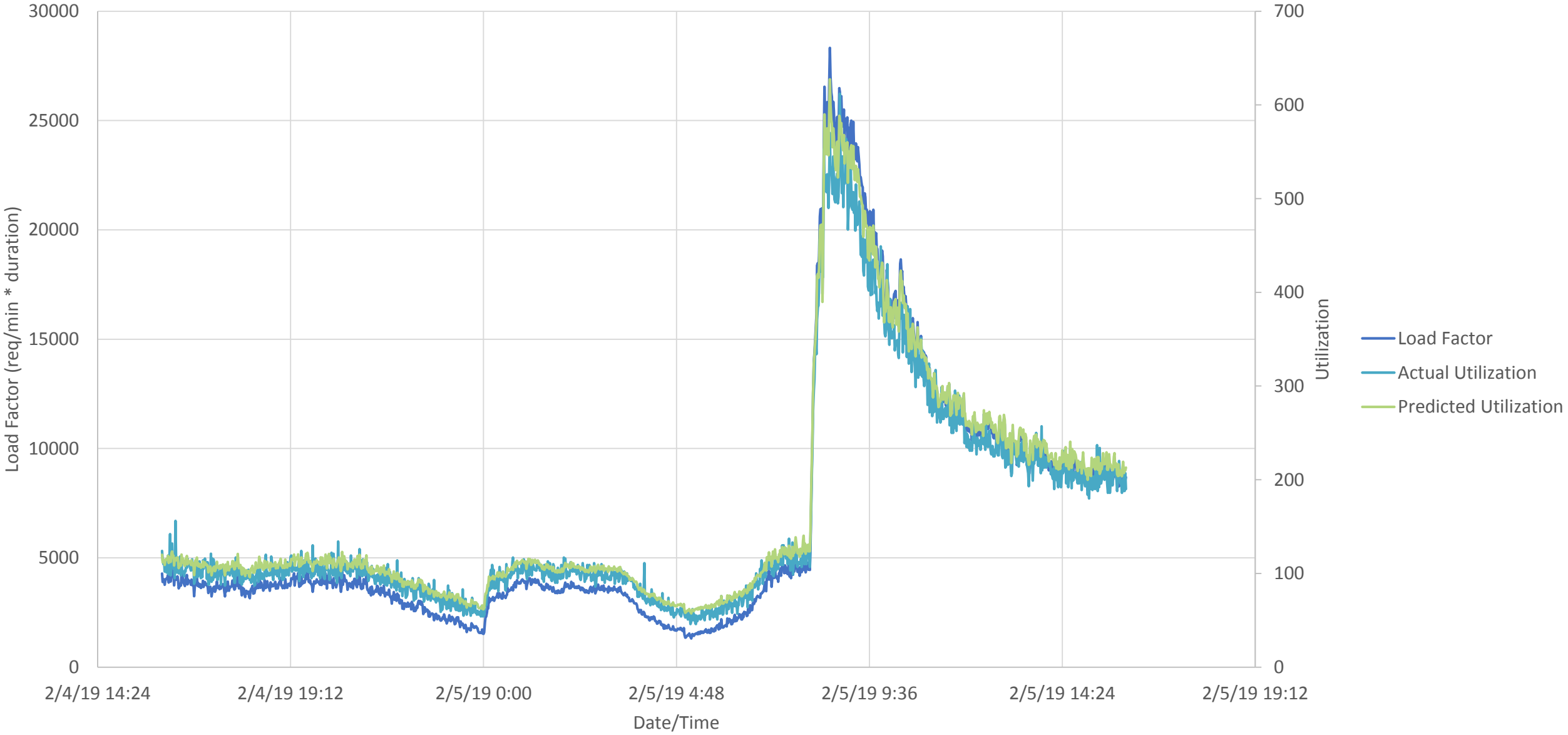
Predictive Model Plot



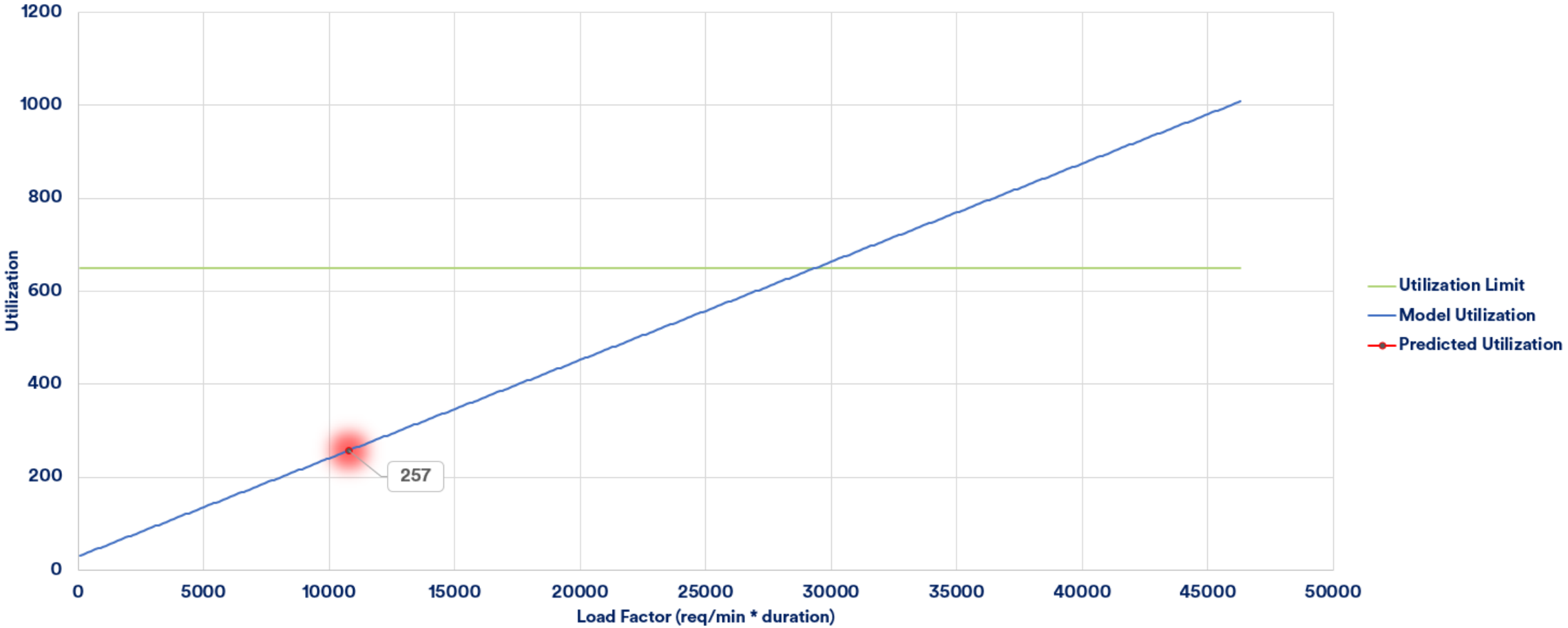
Predictive Model Validation



Predictive Model Validation



Adapting to Change: Demo



Fare Search Rate

Low Fare Search Rate

Schedule Search Rate

	Rate (req/min)	Duration (sec)	Load Factor
FareSearch	5514	0.16	883.67
LowFareSearch	8085	1.11	8981.94
ScheduleFareSearch	5417	0.17	902.94
		Total Load Factor	10768.55
		Predicted Utilization	257

What's next?

Live data feed

Machine learning

Adapt to other metrics

Questions?