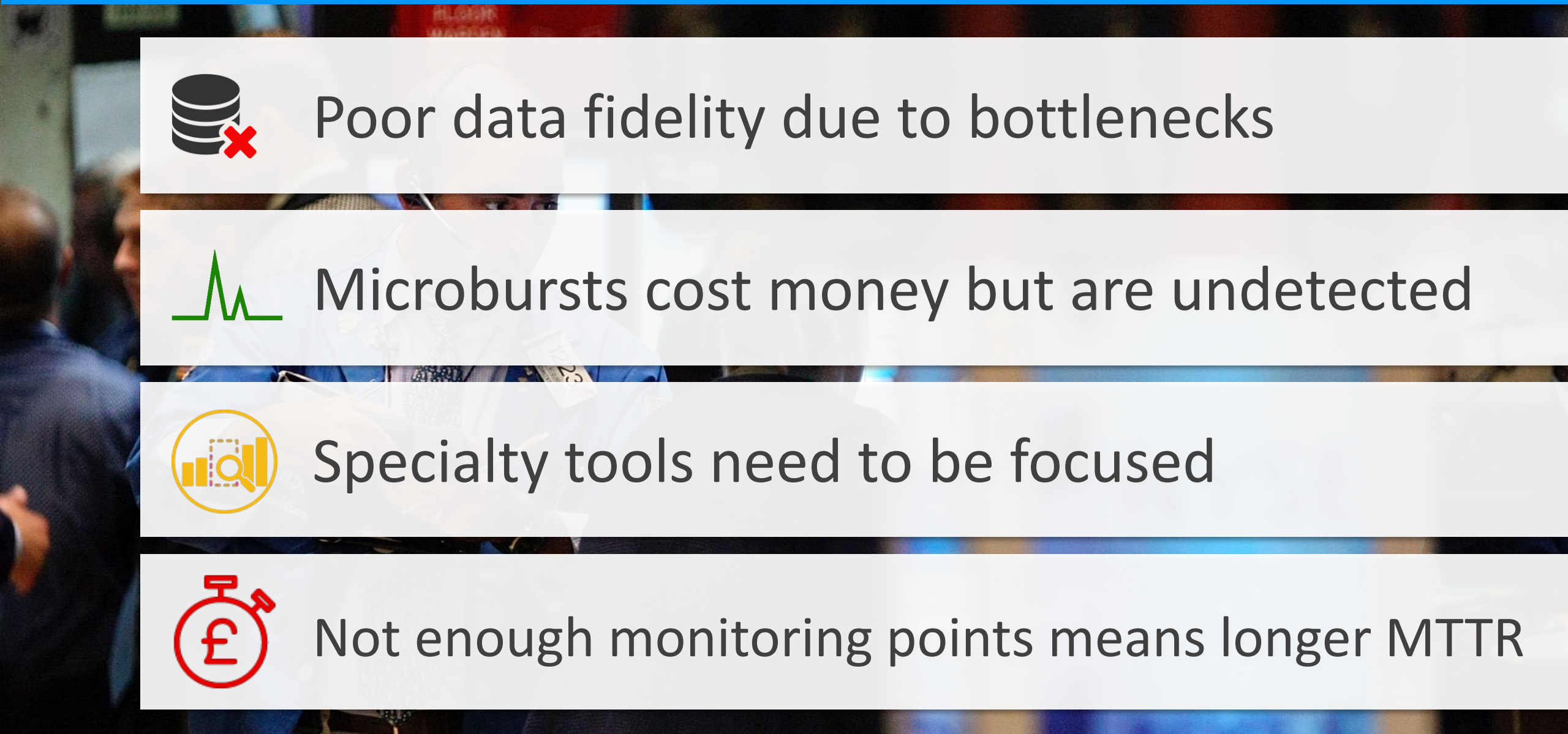
Monitoring at High Speed: Accurate Data=Better Decisions

Josh Joiner, Director of Solutions Engineering

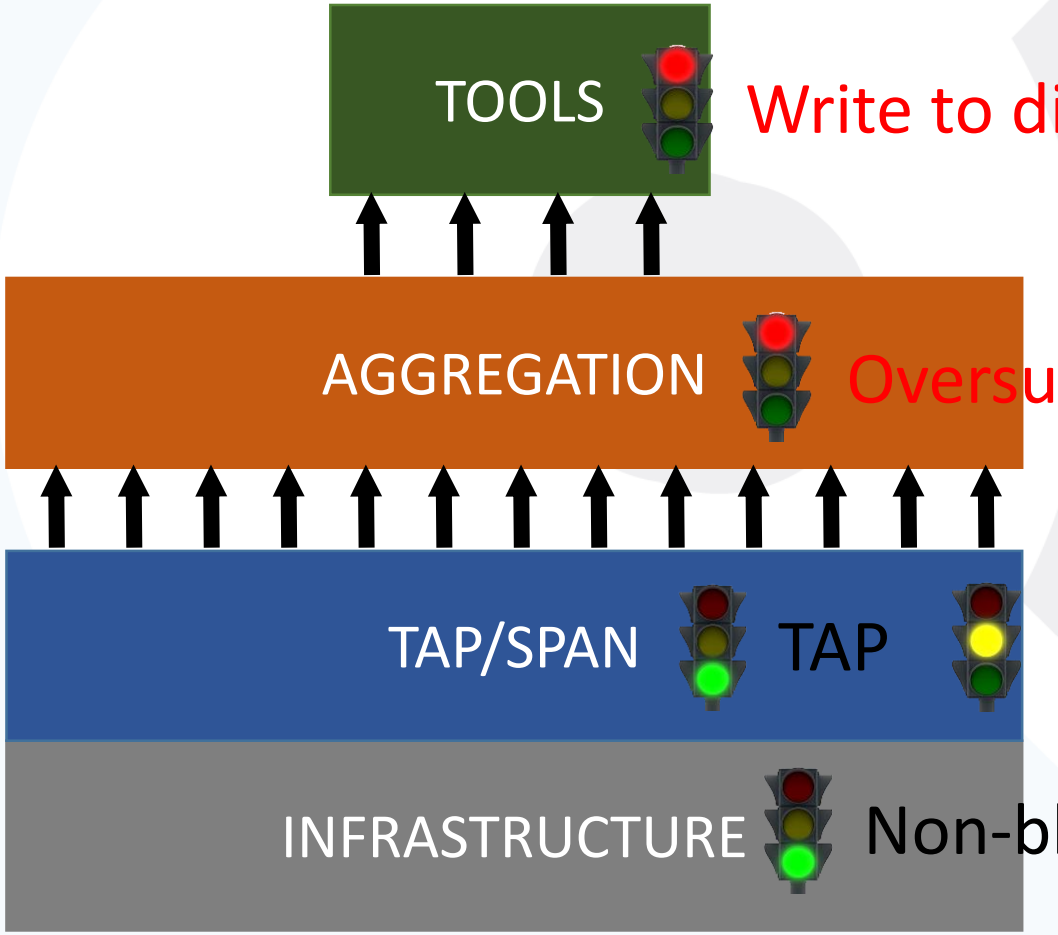
Nov 2017

Key Issues Affecting Data Accuracy in Today's Networks



- Poor data fidelity due to bottlenecks
- Microbursts cost money but are undetected
- Specialty tools need to be focused
- Not enough monitoring points means longer MTTR

Sources of Data Loss: The Monitoring Stack



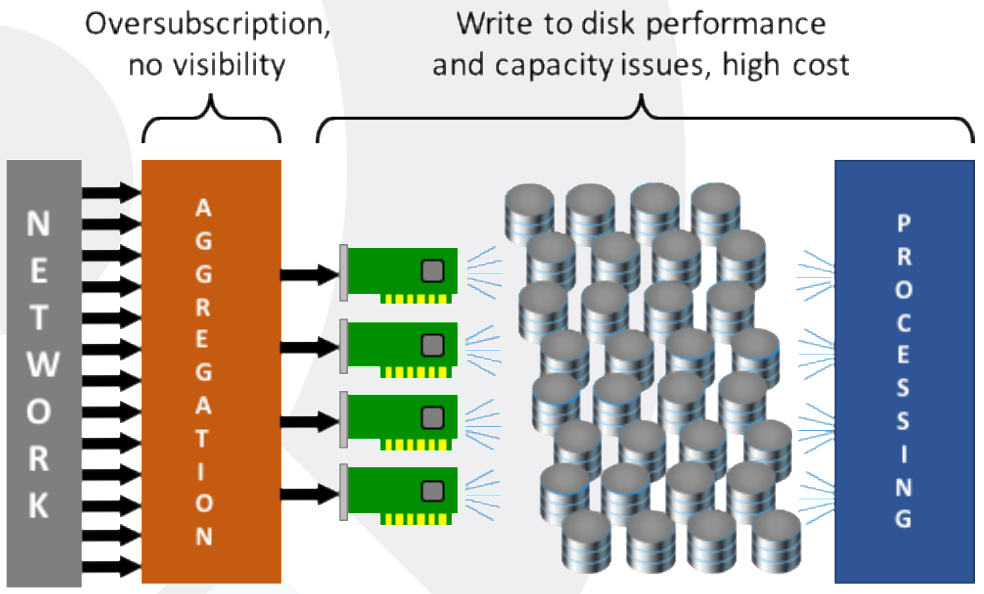
Critical Infrastructure Bottlenecks

Challenges

- Oversubscription
- Data loss at the tool
- False positives

What's Needed

More reliable monitoring stack that bypasses the bottlenecks inherent in monitoring architecture



How Much Data in One Microburst?

$$\frac{10^{11} bits}{1\ second} * \frac{1 byte}{8\ bits} = 12{,}500{,}000{,}000\ bytes/second \quad (1)$$

$$\frac{12{,}500{,}000{,}000\ bytes}{second} * \frac{1\ second}{1000\ milliseconds} = 12{,}500{,}000\ bytes/millisecond \quad (2)$$

At 100 Gbps, a one millisecond spike is 12.5 megabytes…

How much market data is dropped due to microbursts?

Always On Millisecond Accuracy

Spike detection for market data feeds

Challenges

- Markets = Event Driven
- Spikes = Lost Data
- Spikes Are Invisible

What's Needed

Predictive high-resolution monitoring that identifies problems before end-users are impacted

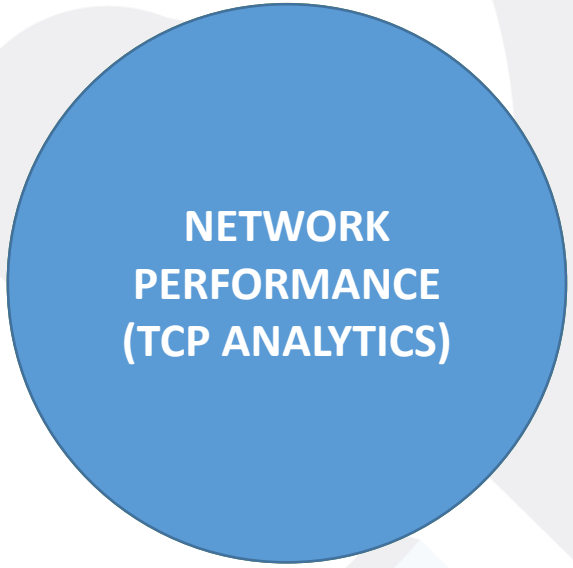
Offload Performance Intensive Operations

Challenges

- High cost of specialized tools
- Limited network issue visibility
- Performance issues

What's Needed

Ability to allocate resources without sacrificing data quality while saving cost



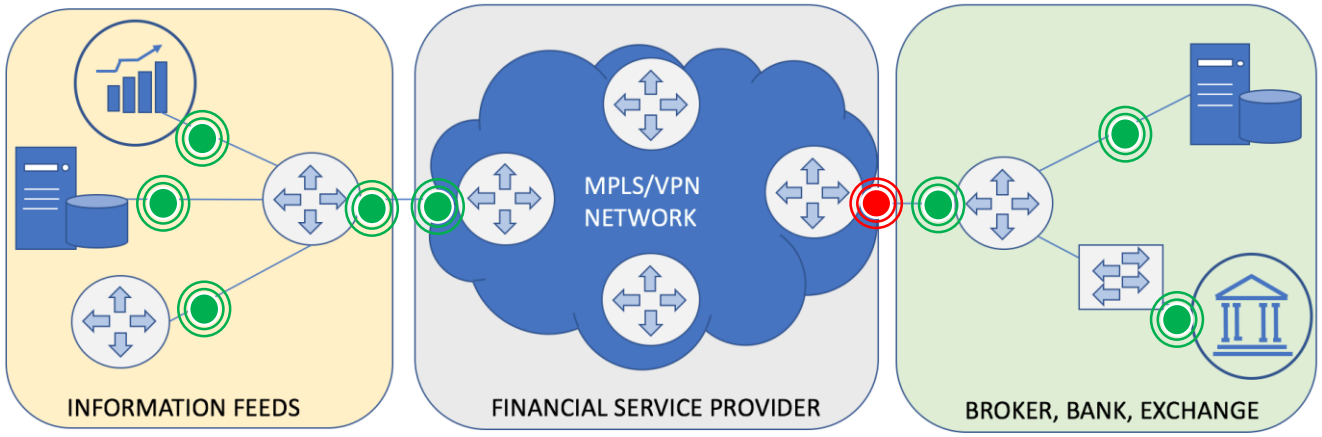
More Visibility Points to Isolate Issues

Challenges

- Lack of visibility (low resolution)
- Too many blind spots
- Long MTTR

What's Needed

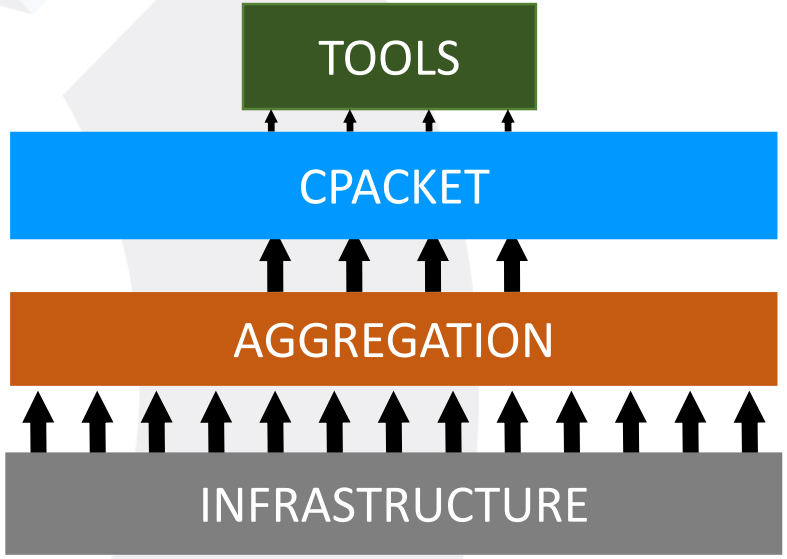
Increase the number of visibility points in order to reduce troubleshooting time



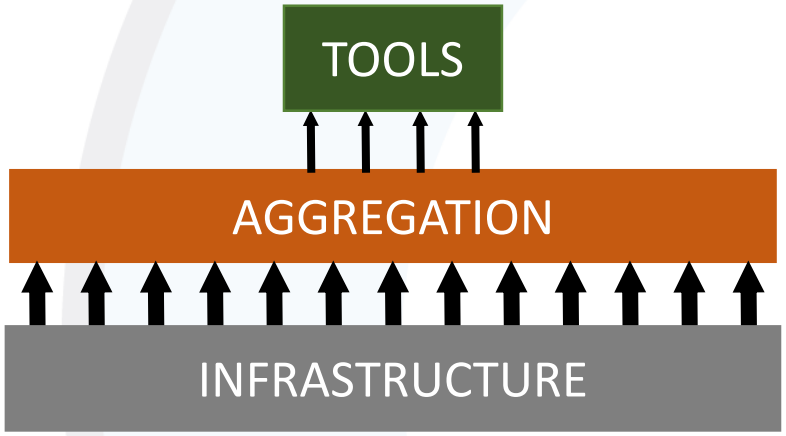
cpacket NETWORKS

cPacket Approach

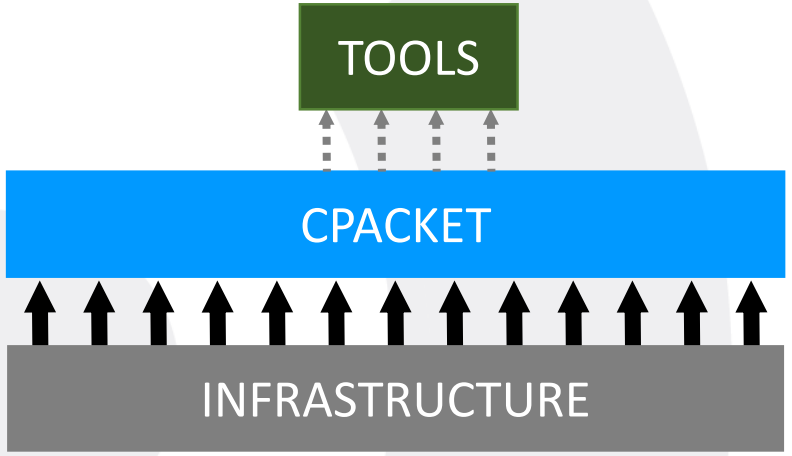
- Push processing to the wire
  - Address the oversubscription and bottlenecks in the aggregation layer
  - Reduce data storage and capture requirements at the tool layer
- Build new technology
  - Develop new hardware, build our own ASIC
- Additional benefits
  - Accuracy, Scalability, Coverage



Increased Coverage = Better Data



1. Aggregation layer provides no visibility into network behavior
2. Oversubscription at aggregation distorts tool analysis
3. Performance and cost constraints at tool layer means limited visibility points

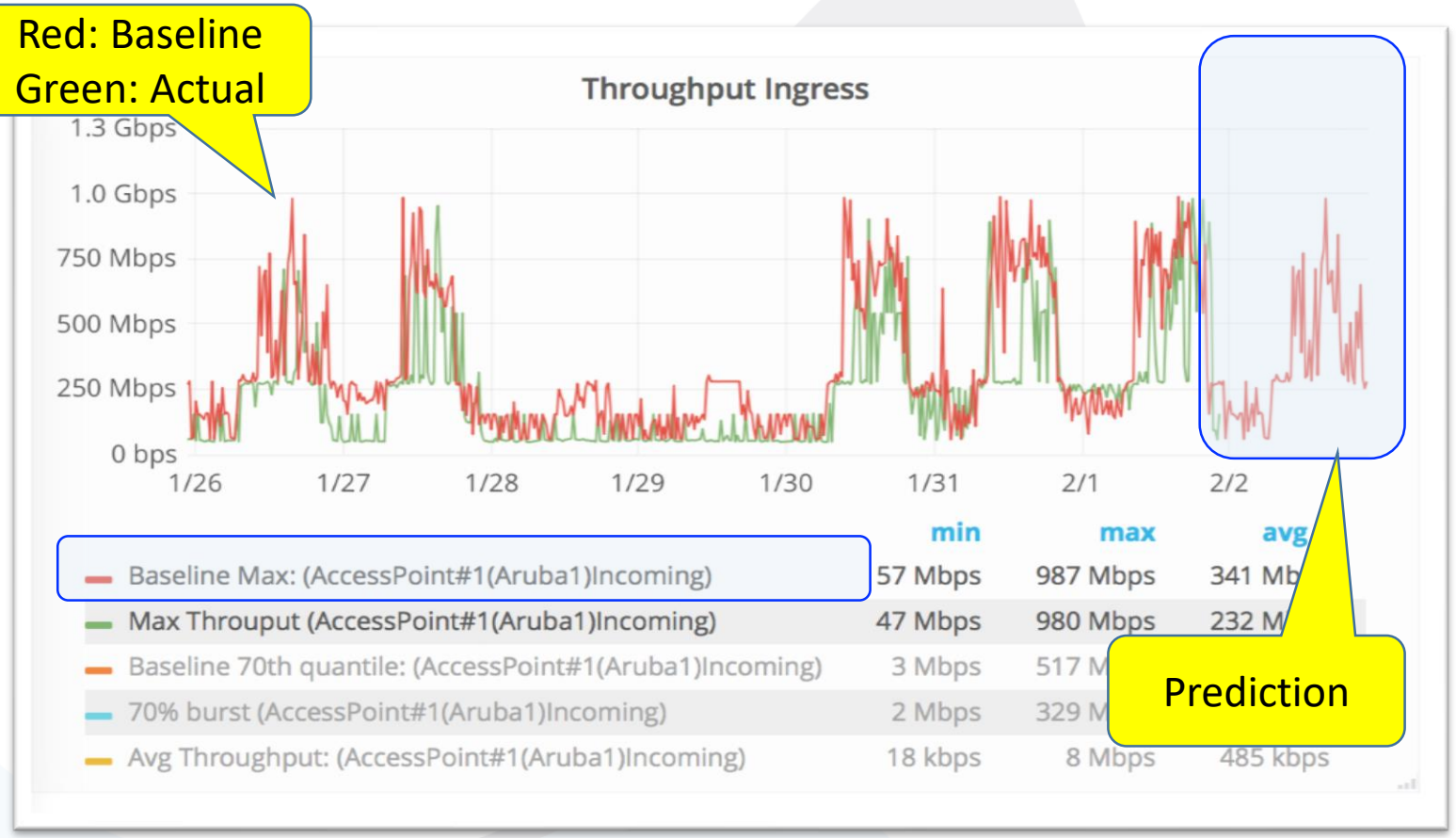


1. Smart connection to infrastructure for immediate KPI extraction
2. KPIs extracted before aggregation for reliable, high resolution data analysis
3. Hardware architecture provides more connections at same cost

Every connection to the network is a visibility point: Always On Coverage

Better Data for Better Predictions

- Analyze past network behavior
- Predict future behavior
- Set automatic thresholds and alerts
- Use trends to do ‘What if’ analysis
- Proactively alert on future trends for preventive action



Validate network performance against past behavior and predict future behavior

It’s About the Data

The secret to Google’s success: “We don’t have better algorithms. We just have more data.” P. Norvig, Chief Scientist #Zeitgeist

- Beyond real life experience in high speed networking, the value of data to machine learning performance is well established. The accuracy of predictive algorithms depends more on the amount and quality of data than the algorithms.

  - More data usually beats better algorithms (Anand Rajaraman, Stanford)
  - Scaling to Very Very Large Corpora for Natural Language Disambiguation (Michelle Banko and Eric Brill, Microsoft Research)
  - The Unreasonable Effectiveness of Data (Alon Halevy, Peter Norvig, Fernando Pereira, Google)

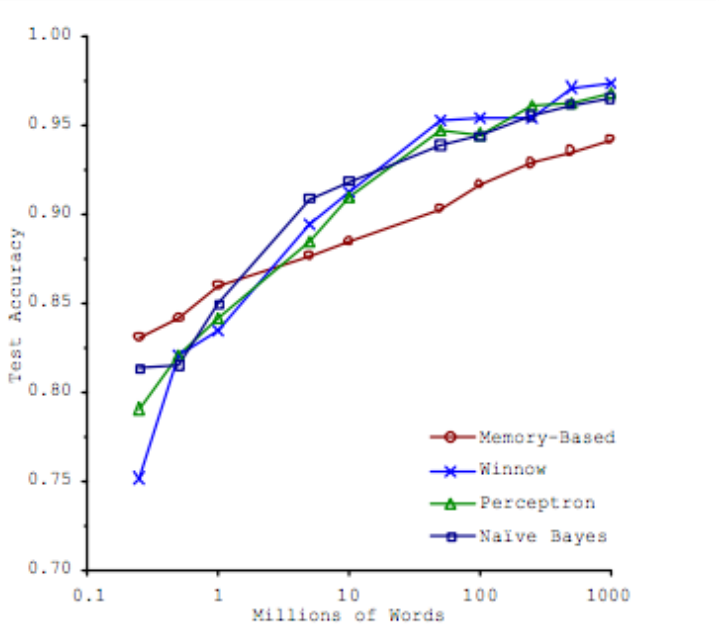


Figure 1. Learning Curves for Confusion Set Disambiguation

Thank You

Please email questions to:

josh@cpacket.com