

Stable and Accurate Health-Checking of Horizontally-Scaled Services

Lorenzo Saino



@lorenzosaino

fastly[®]



Pinterest



DAILY NEWS



imgur



FAST COMPANY



CONDÉ NAST

BuzzFeed

Rolling Stone

BUSINESS INSIDER

KICKSTARTER



FOURSQUARE



GOV.UK

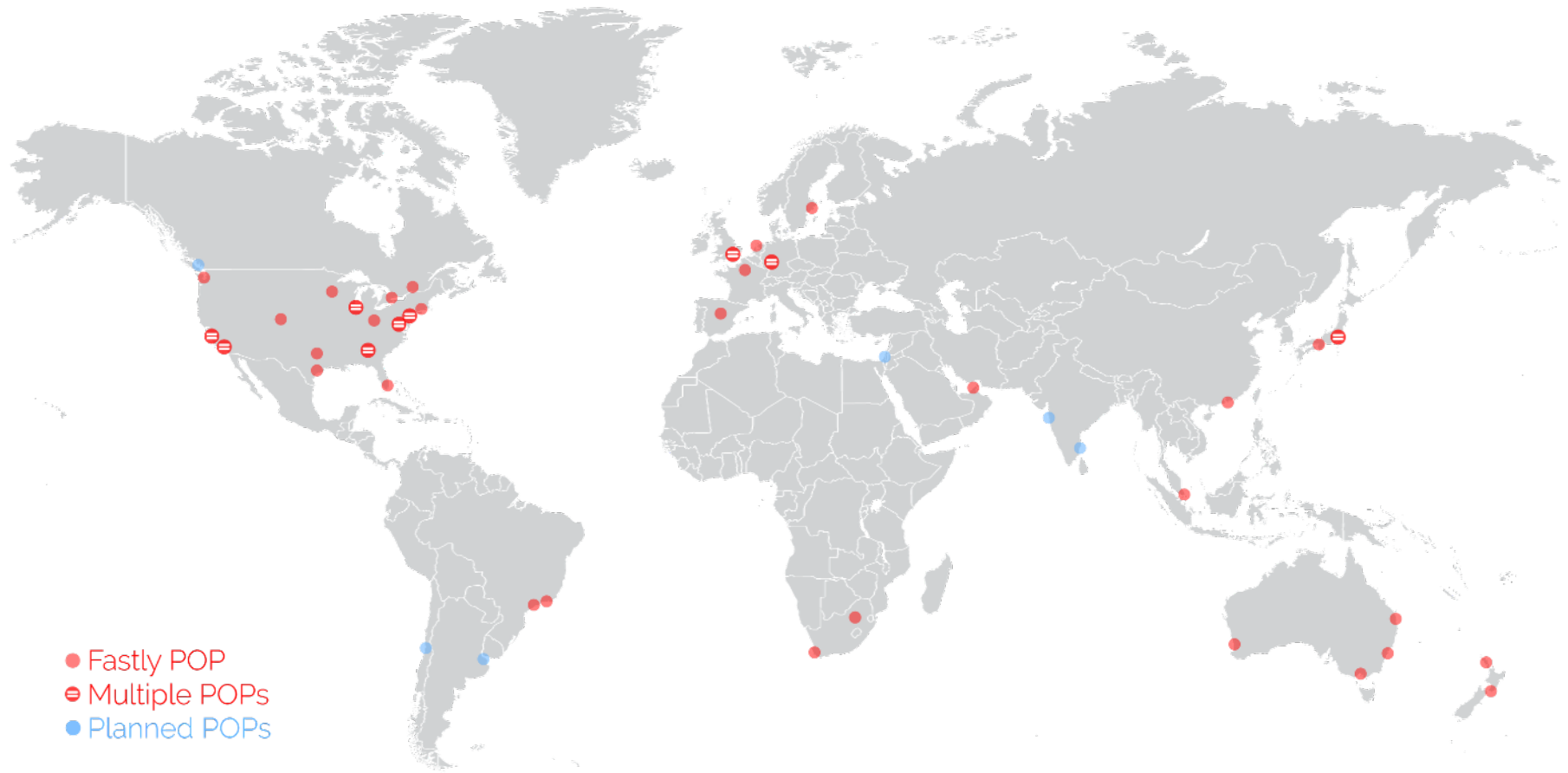
vimeo



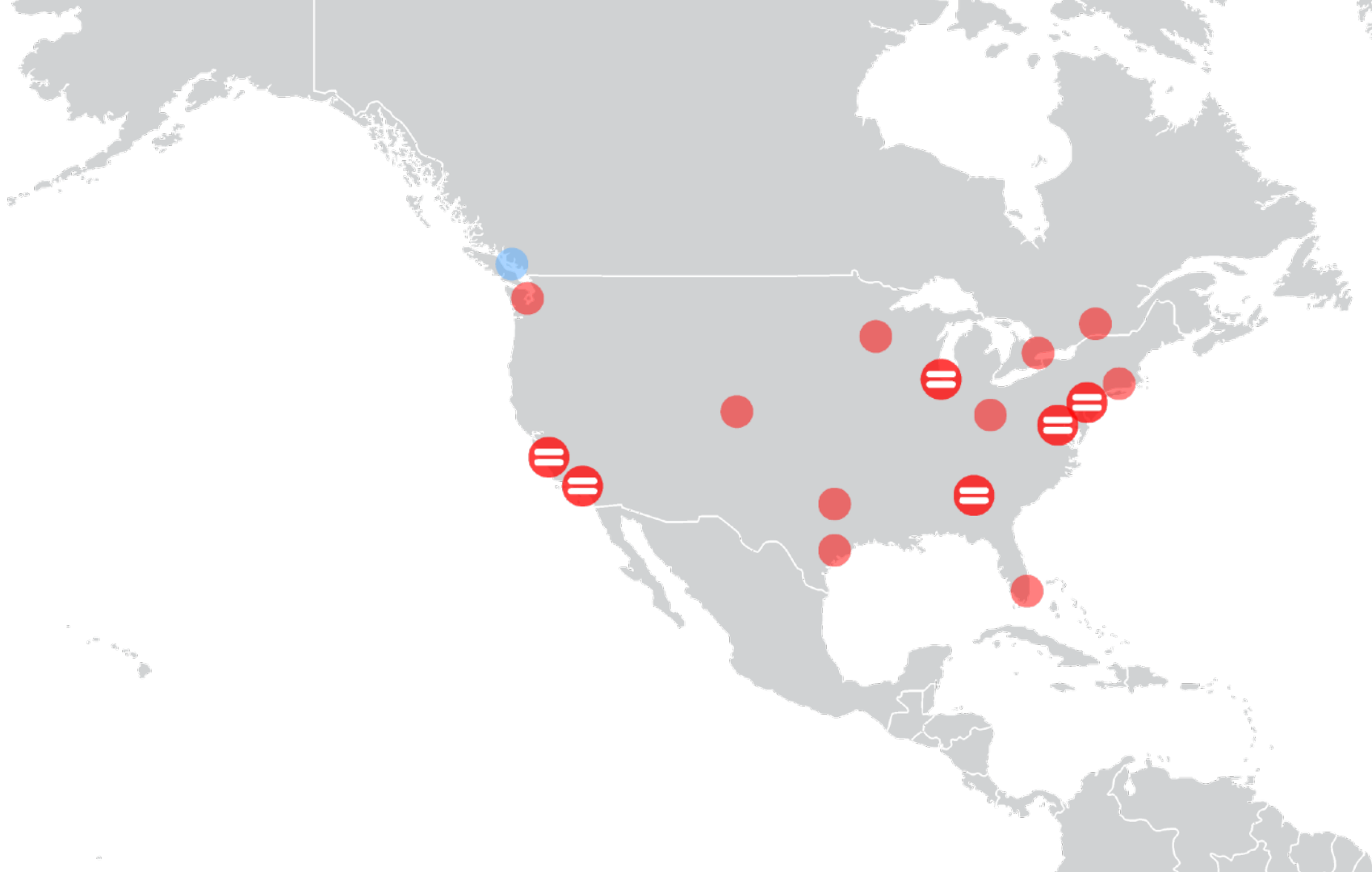
Istdibs.
THE MOST BEAUTIFUL THINGS ON EARTH

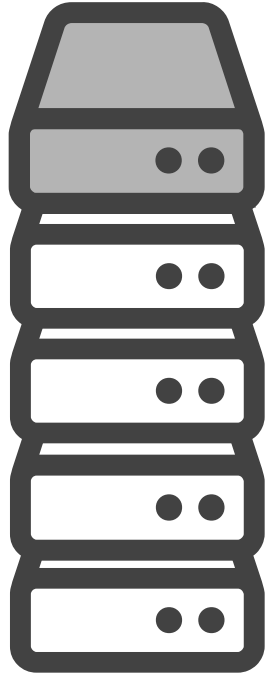


Etsy



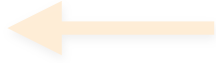
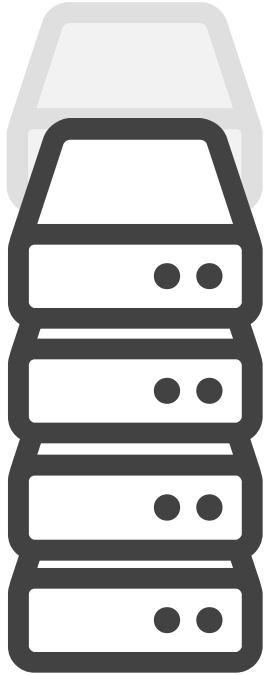
- Fastly POP
- ⊖ Multiple POPs
- Planned POPs







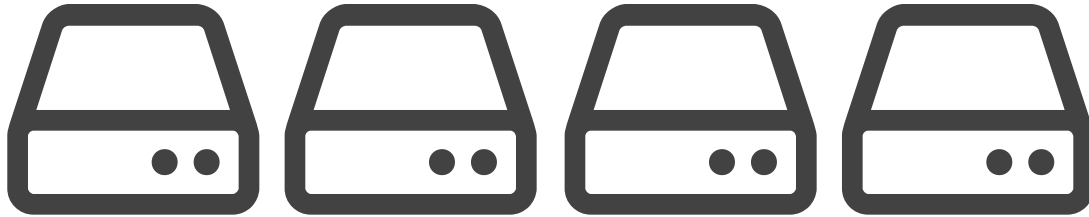
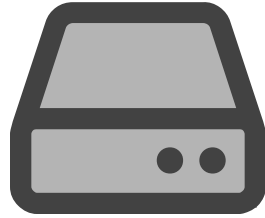
load balancer

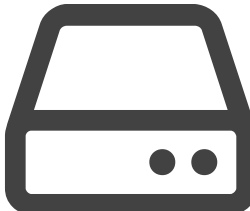
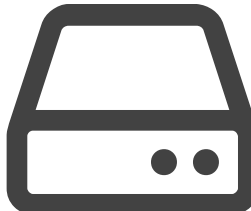
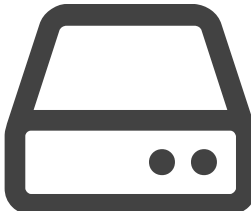
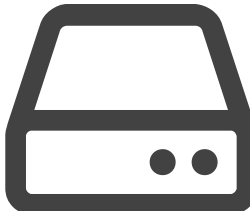
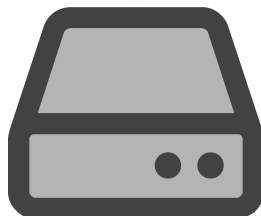


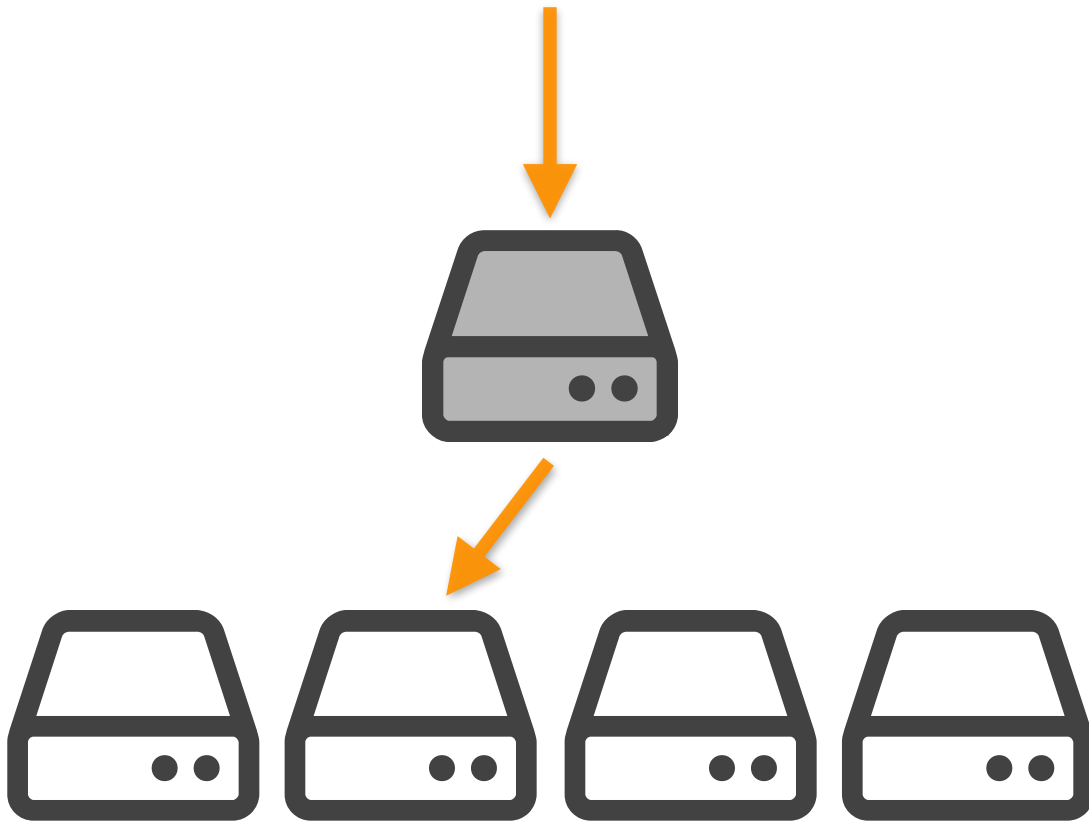
load balancer

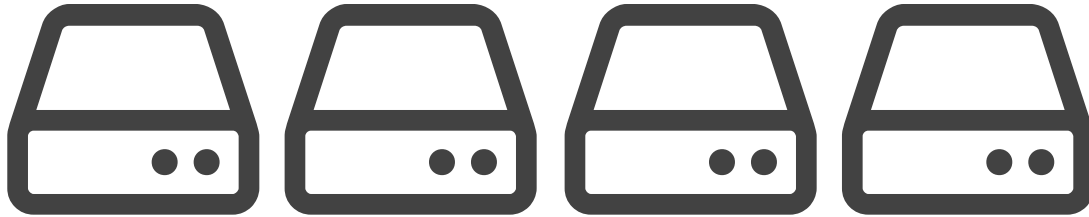
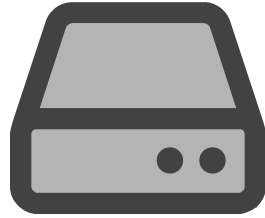


hosts









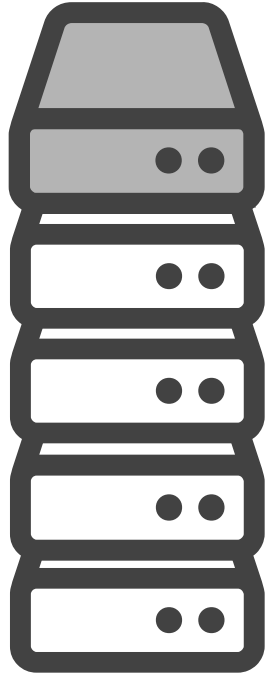


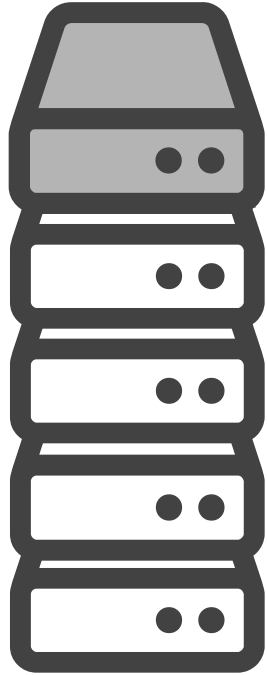
J. Taveira Araújo, L. Saino, L. Buytenhek and R. Landa

Balancing on the edge: transport affinity without network state

to appear in *USENIX NSDI '18*

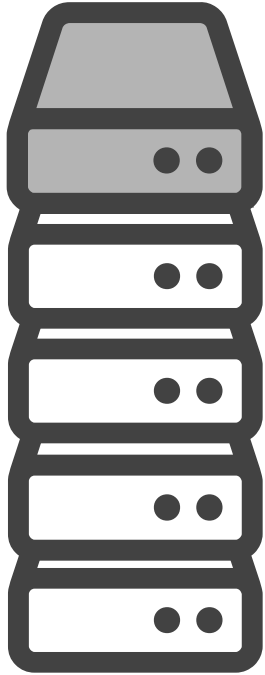






Space and power are at a premium

Careful considerations in POP design and HW selection



Space and power are at a premium

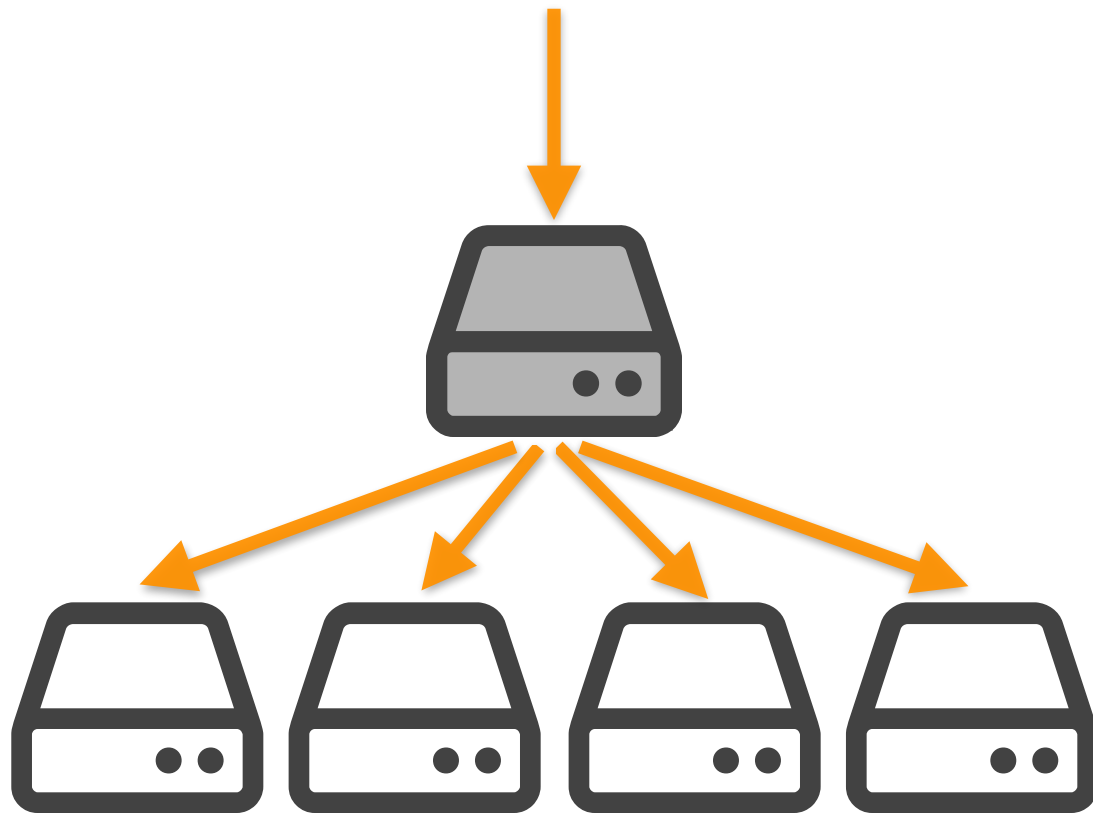
Careful considerations in POP design and HW selection

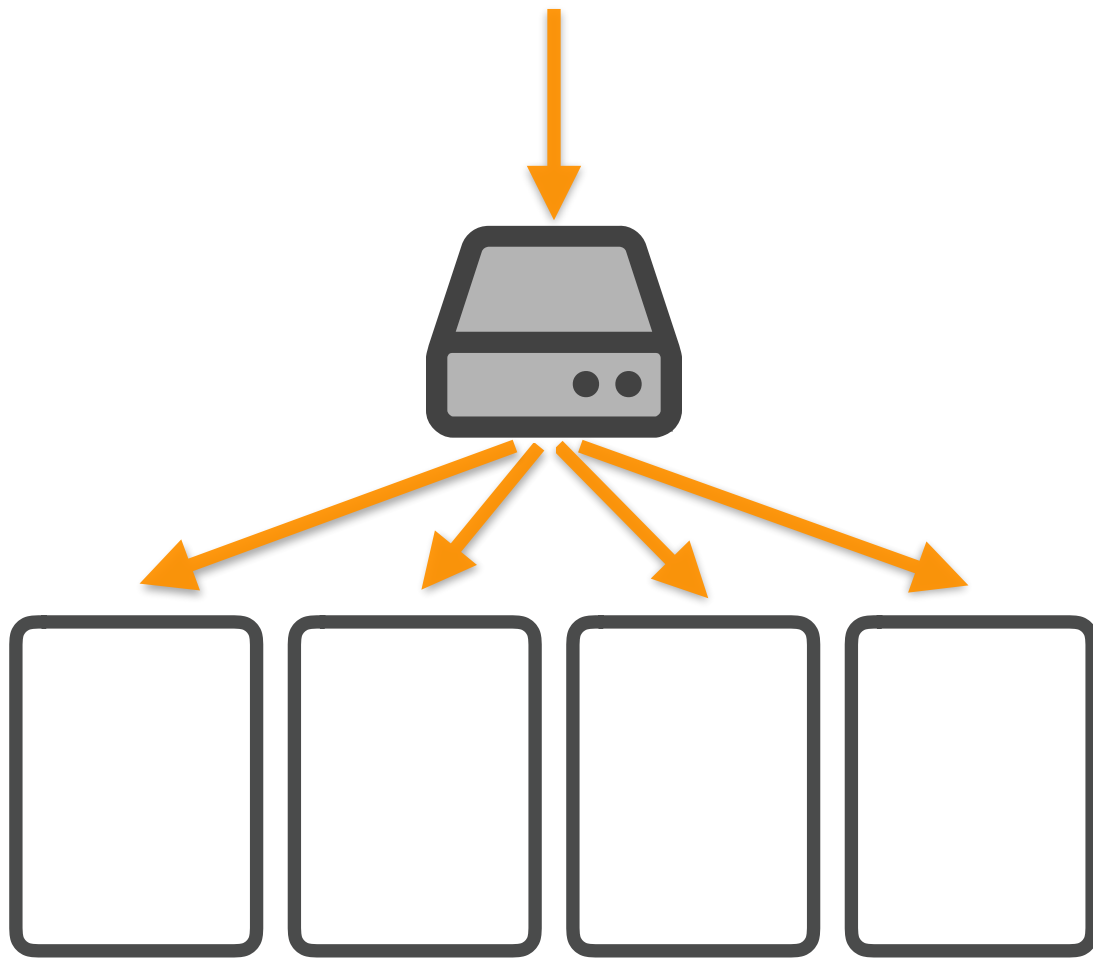
Limited scalability

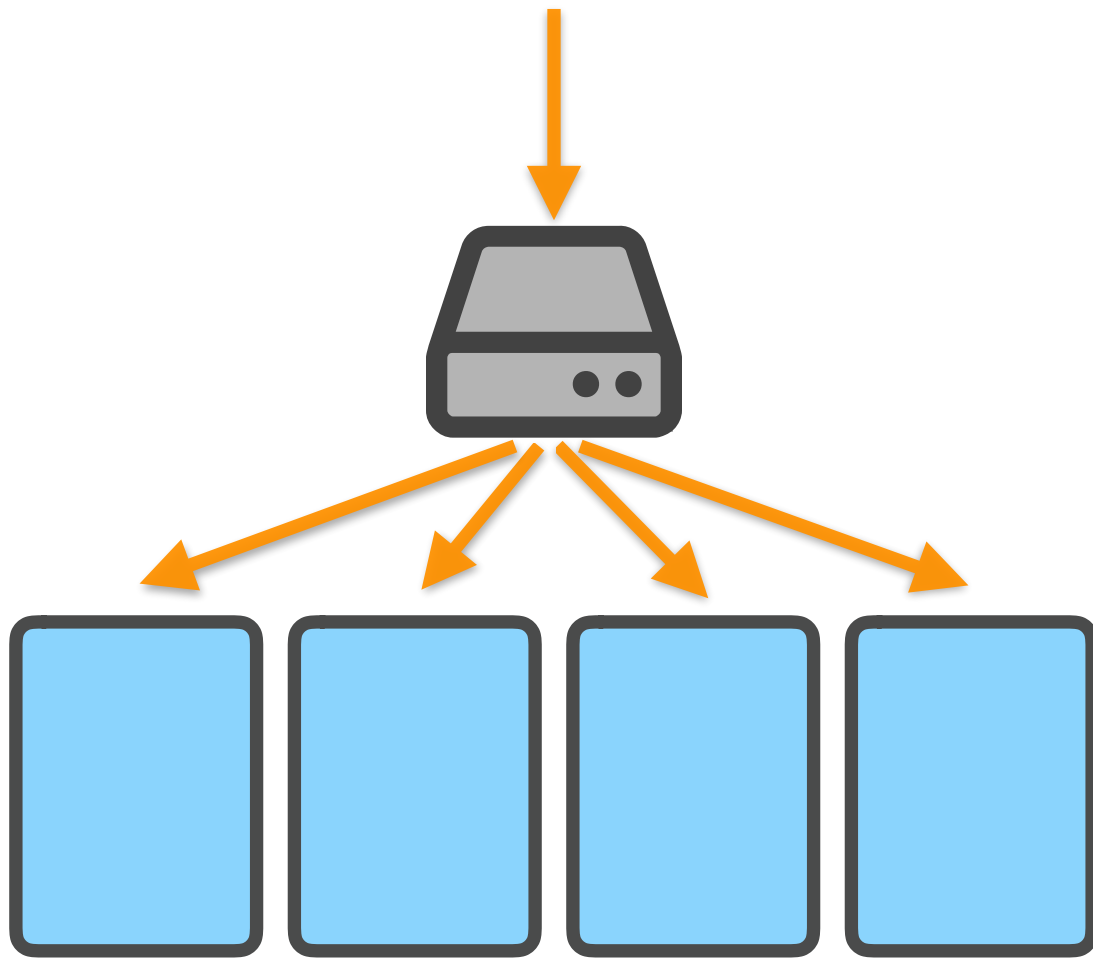
Scaling POP capacity requires racking new servers

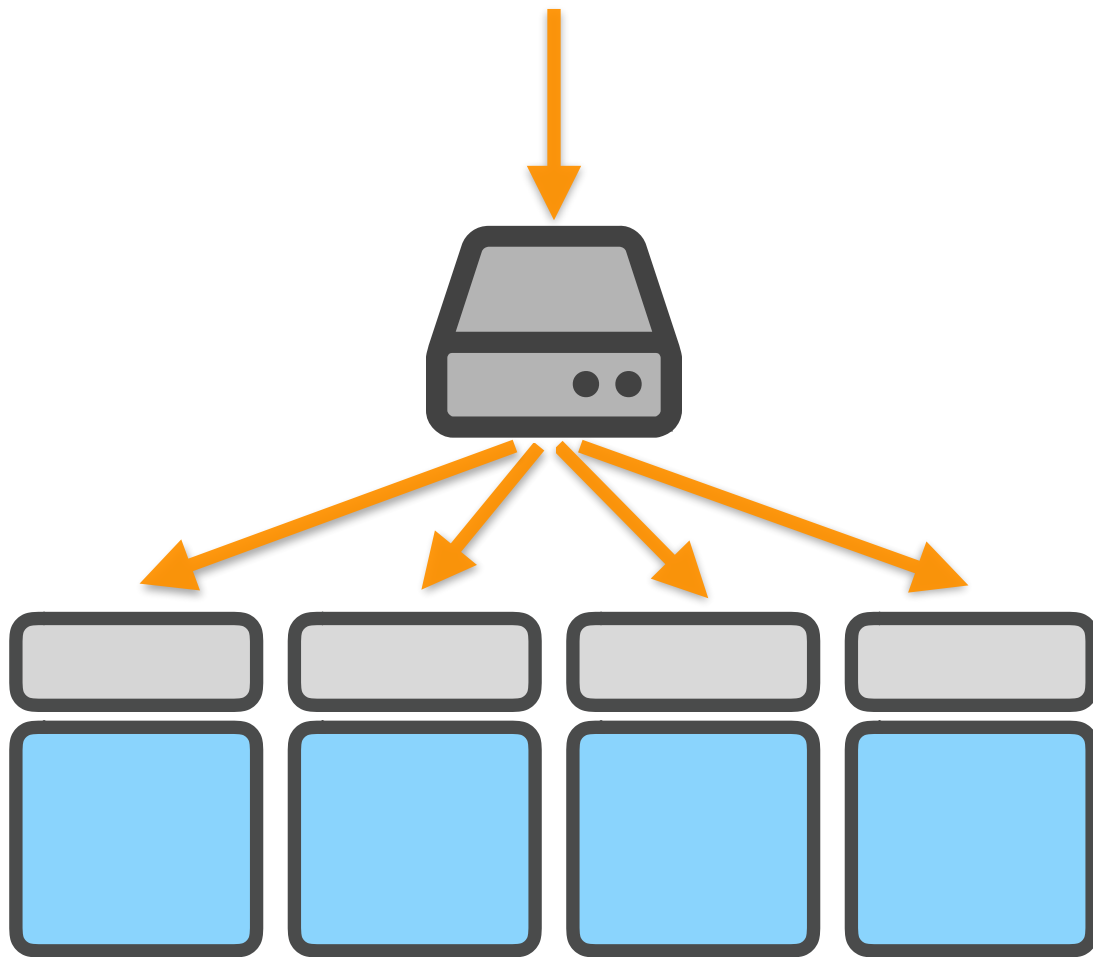


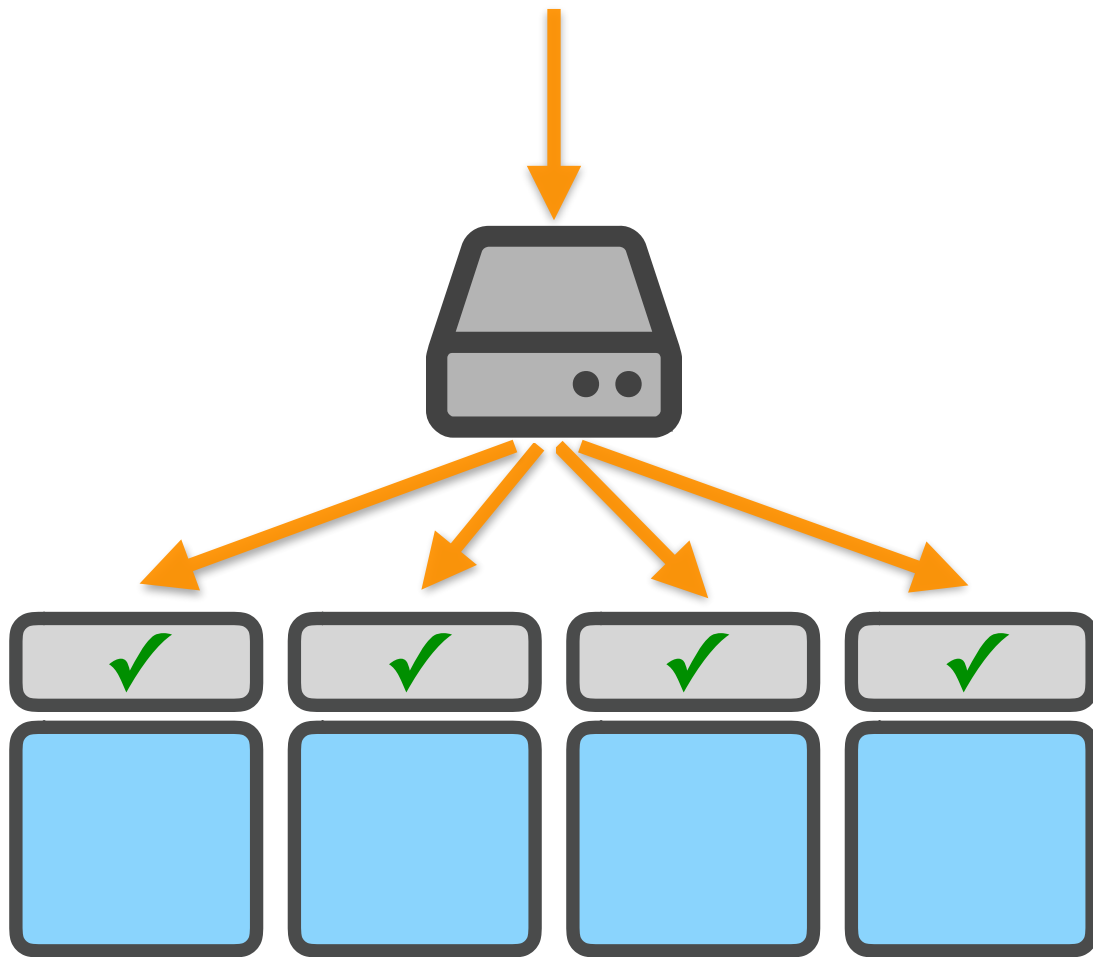
Health-checking

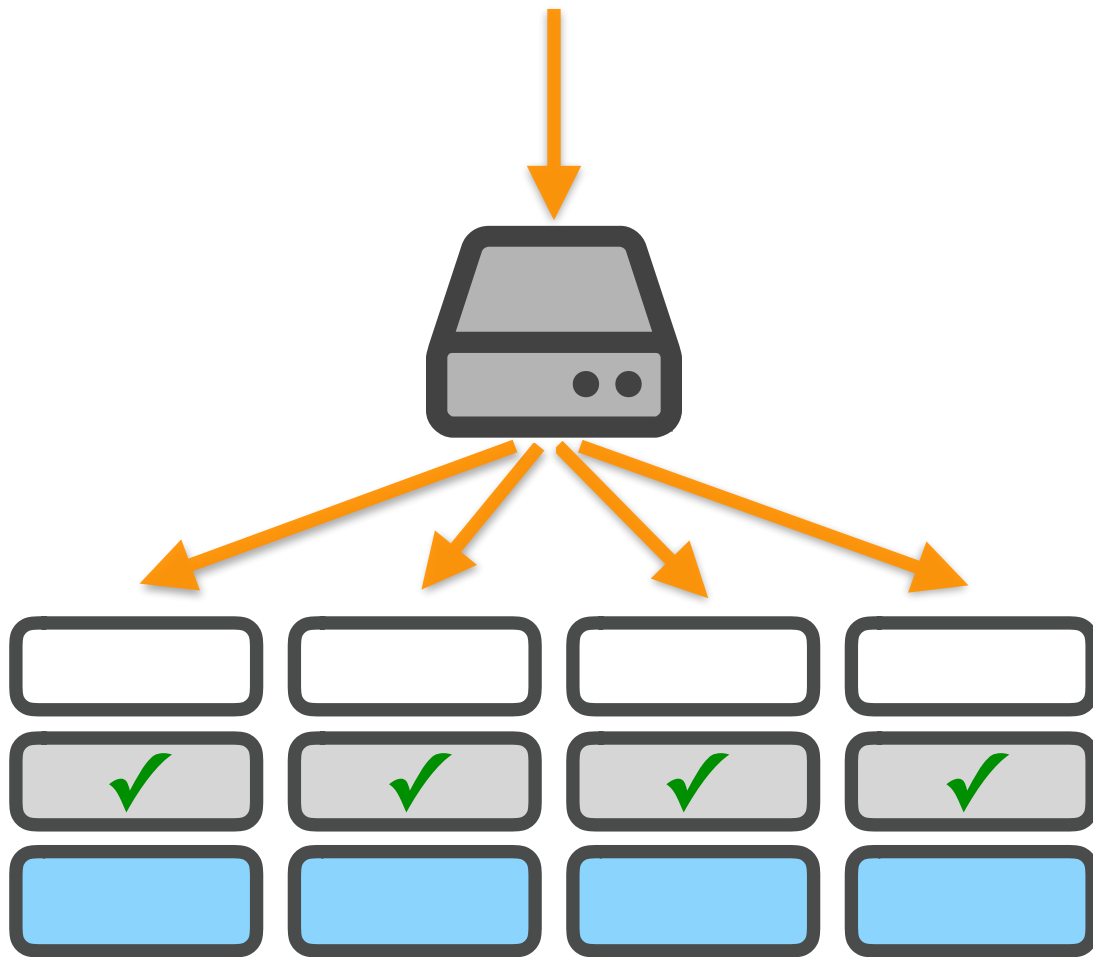


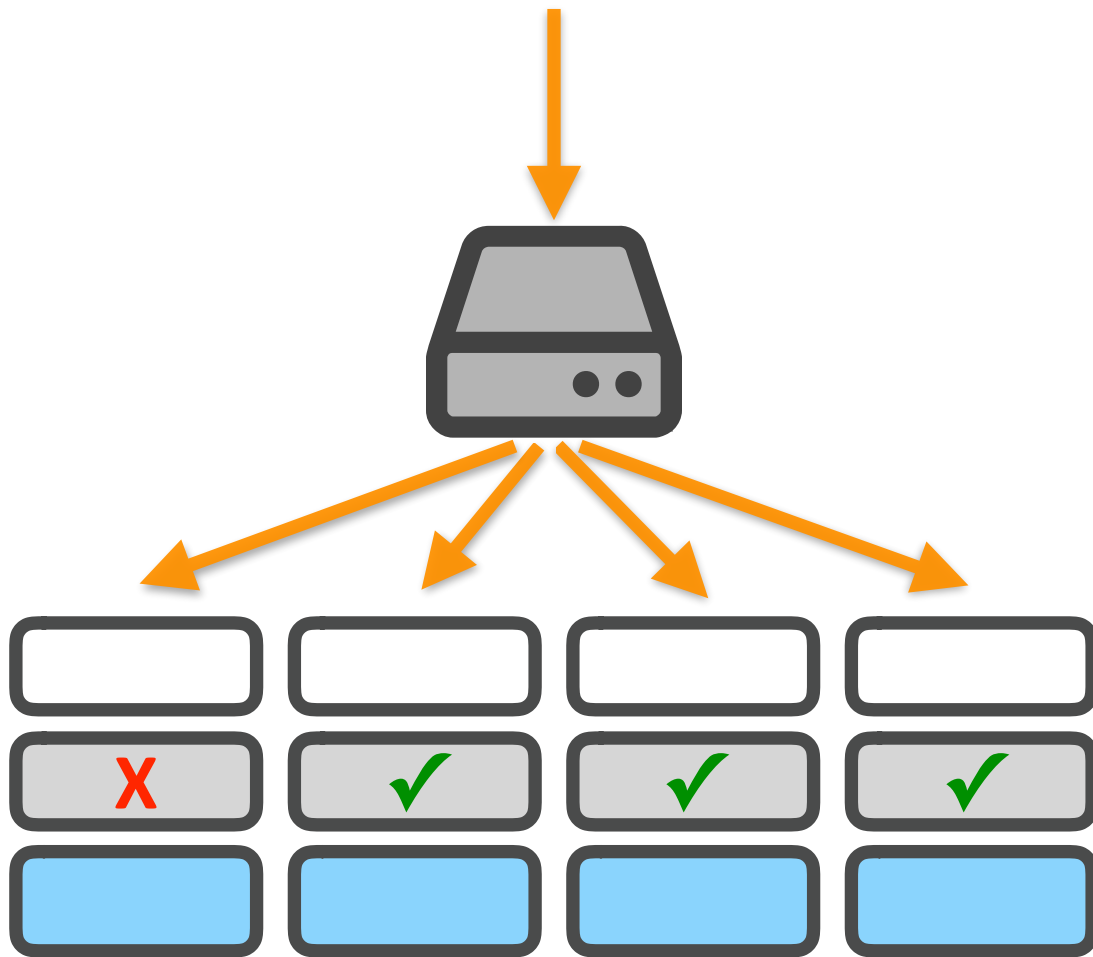


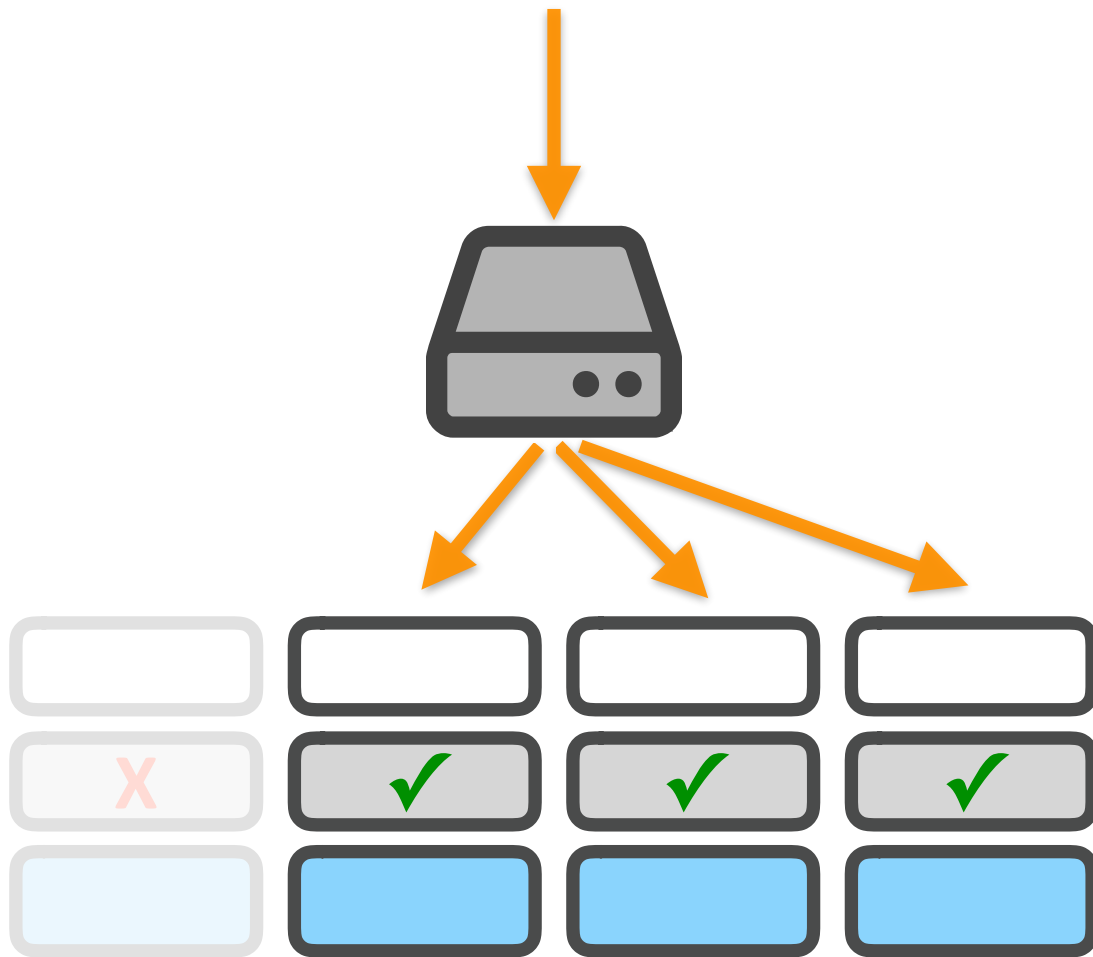


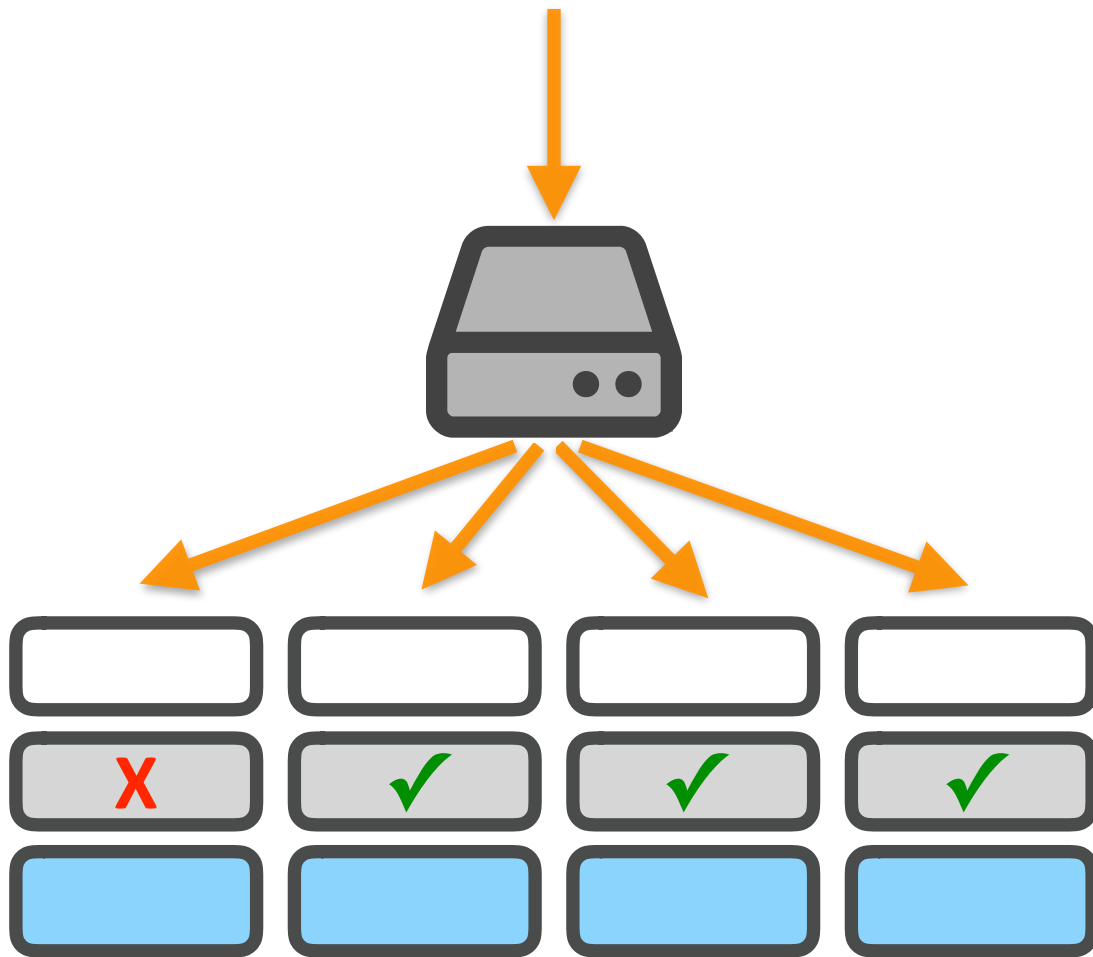


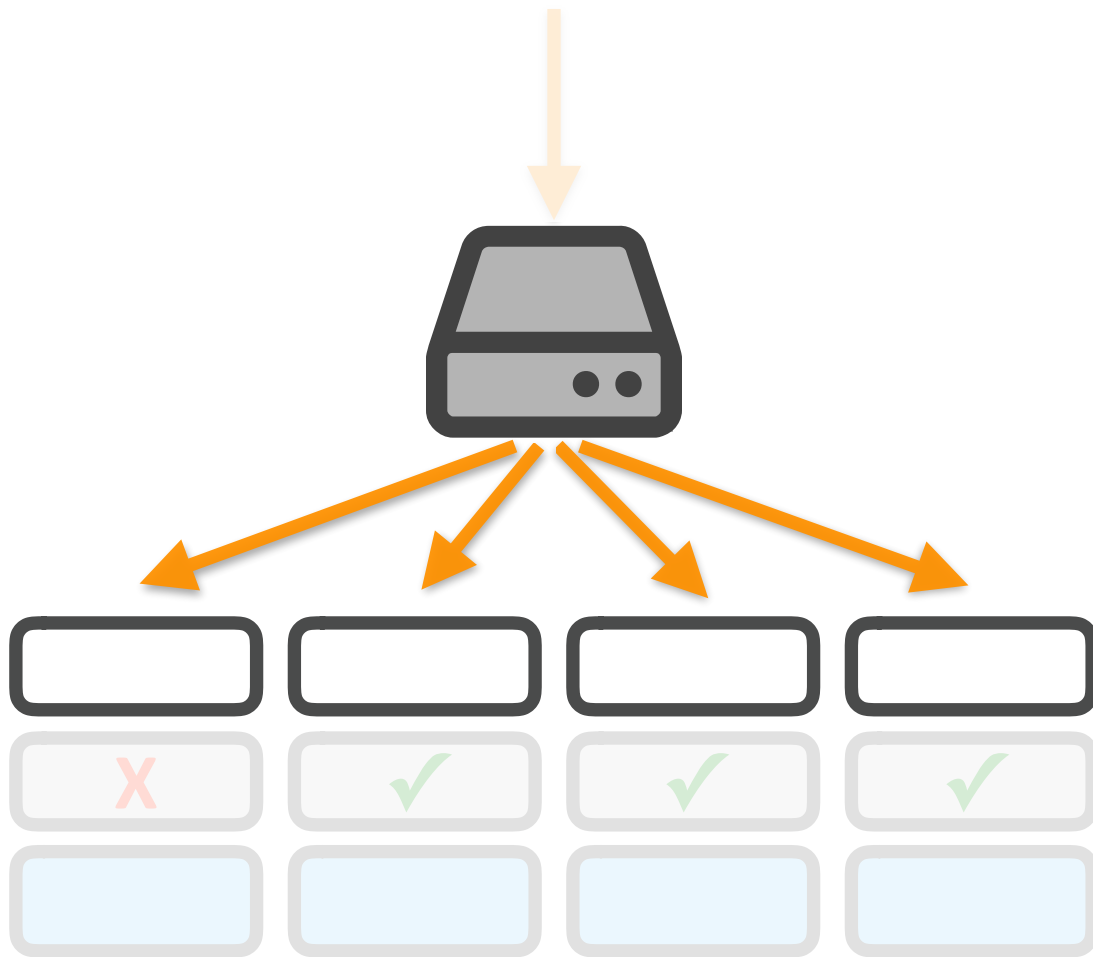


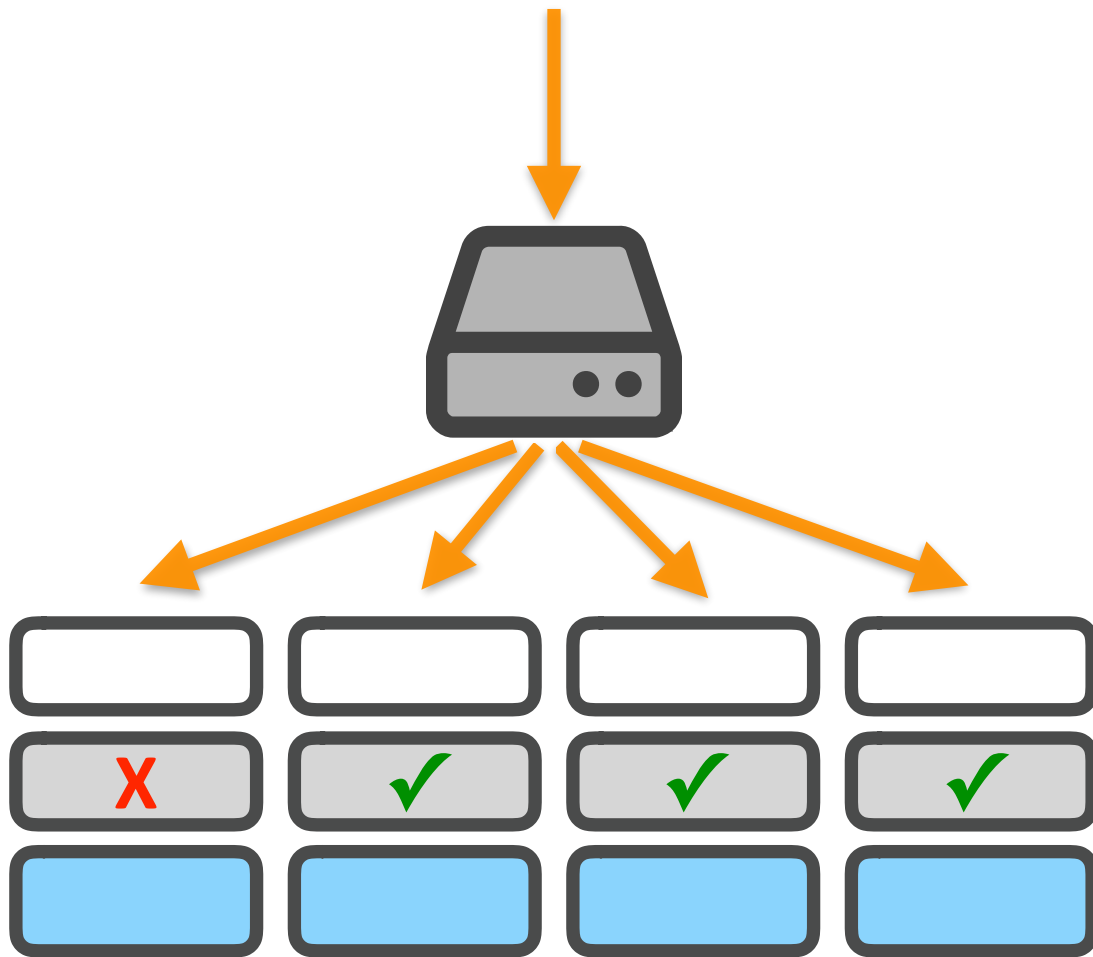


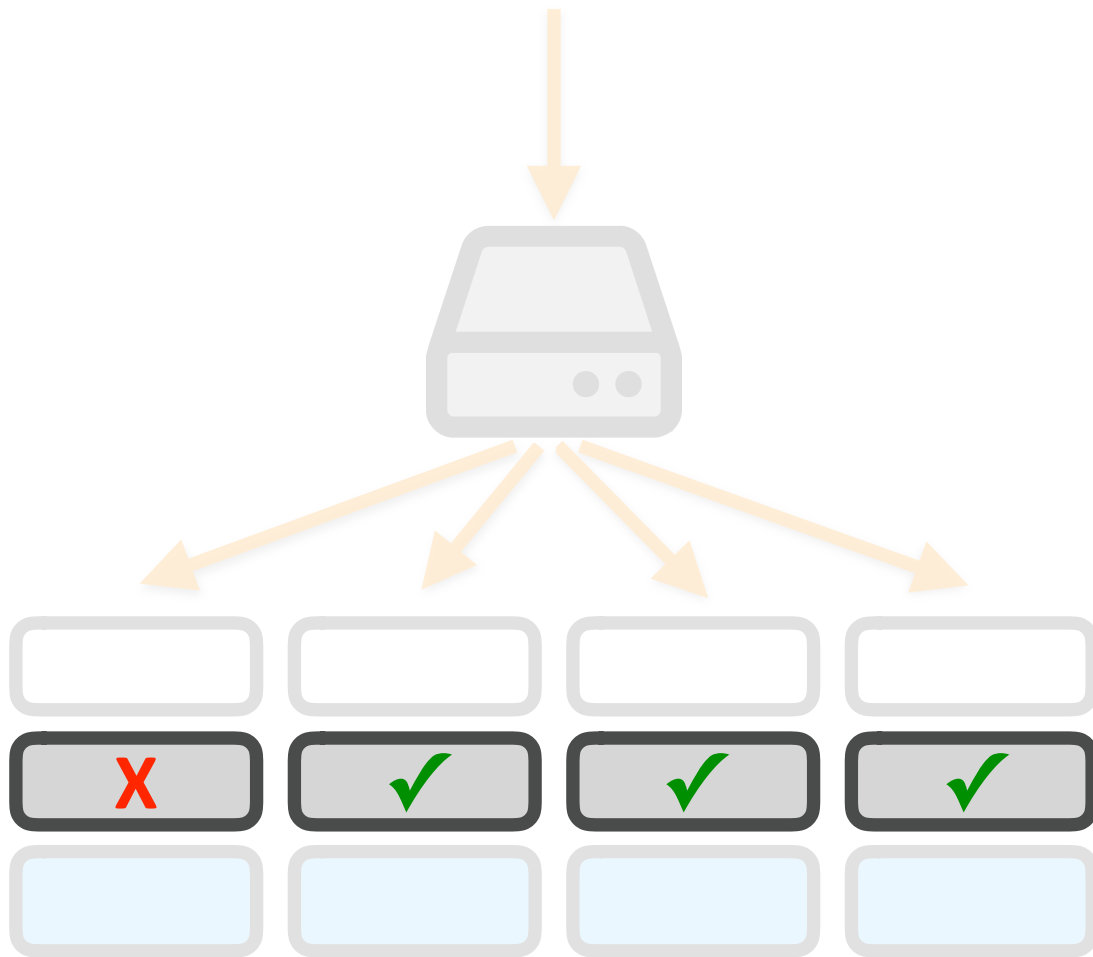


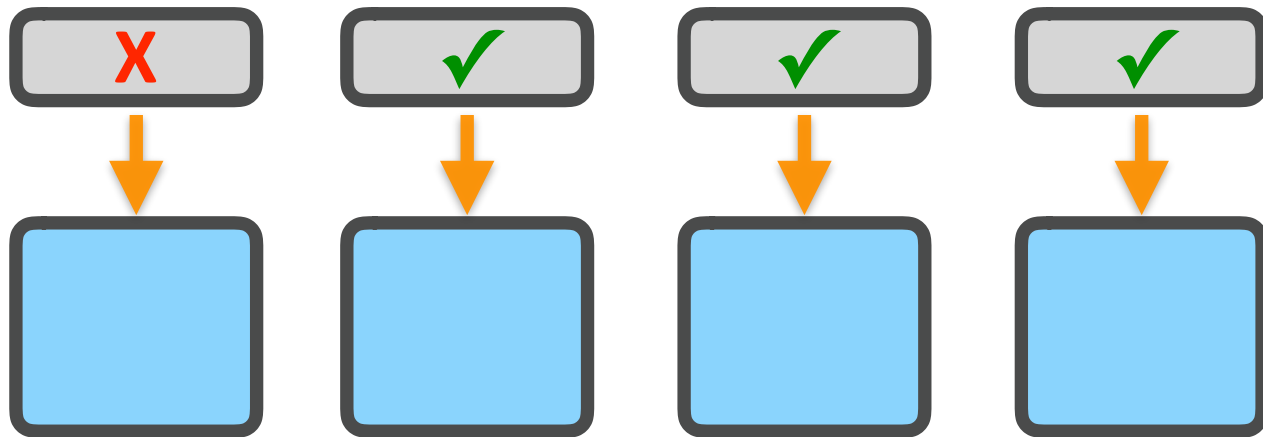


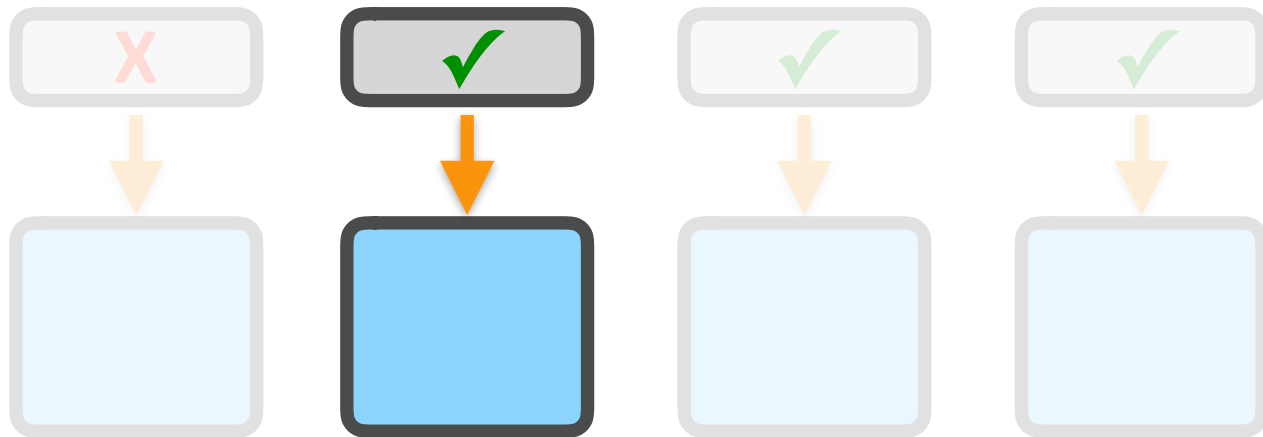


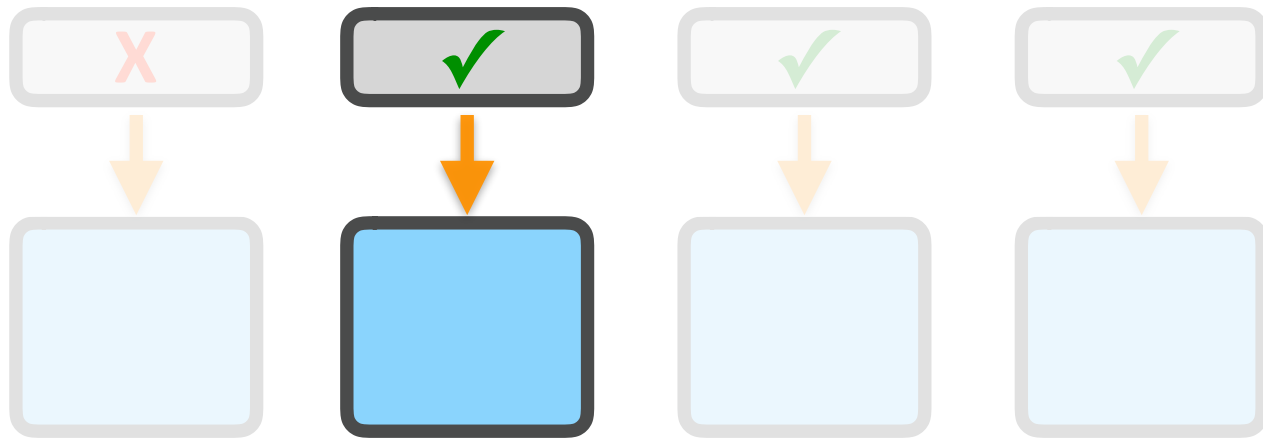


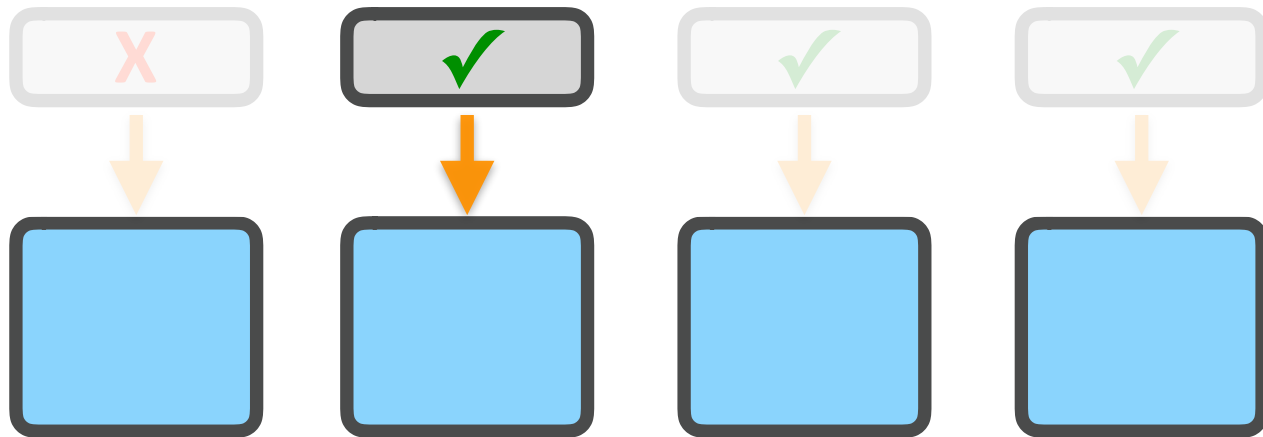


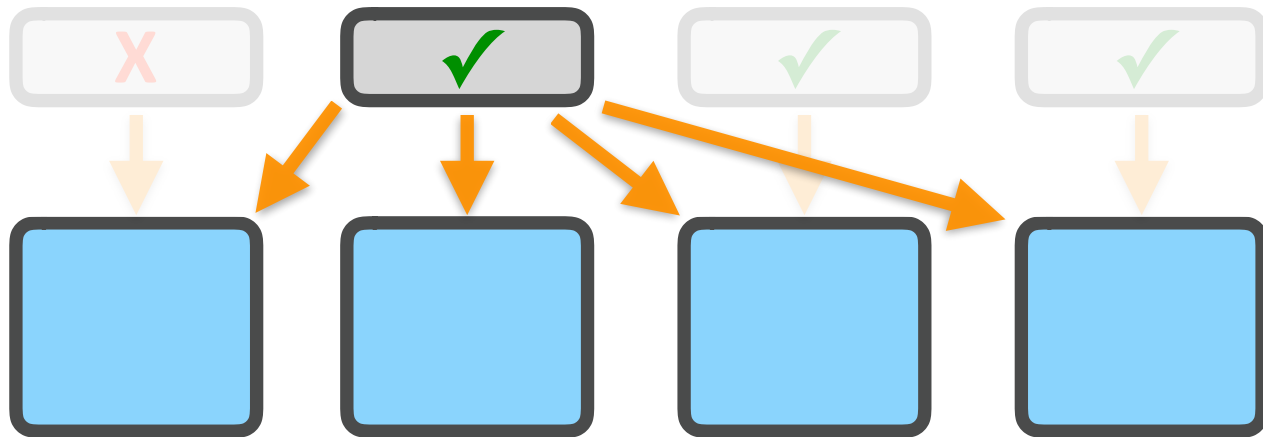


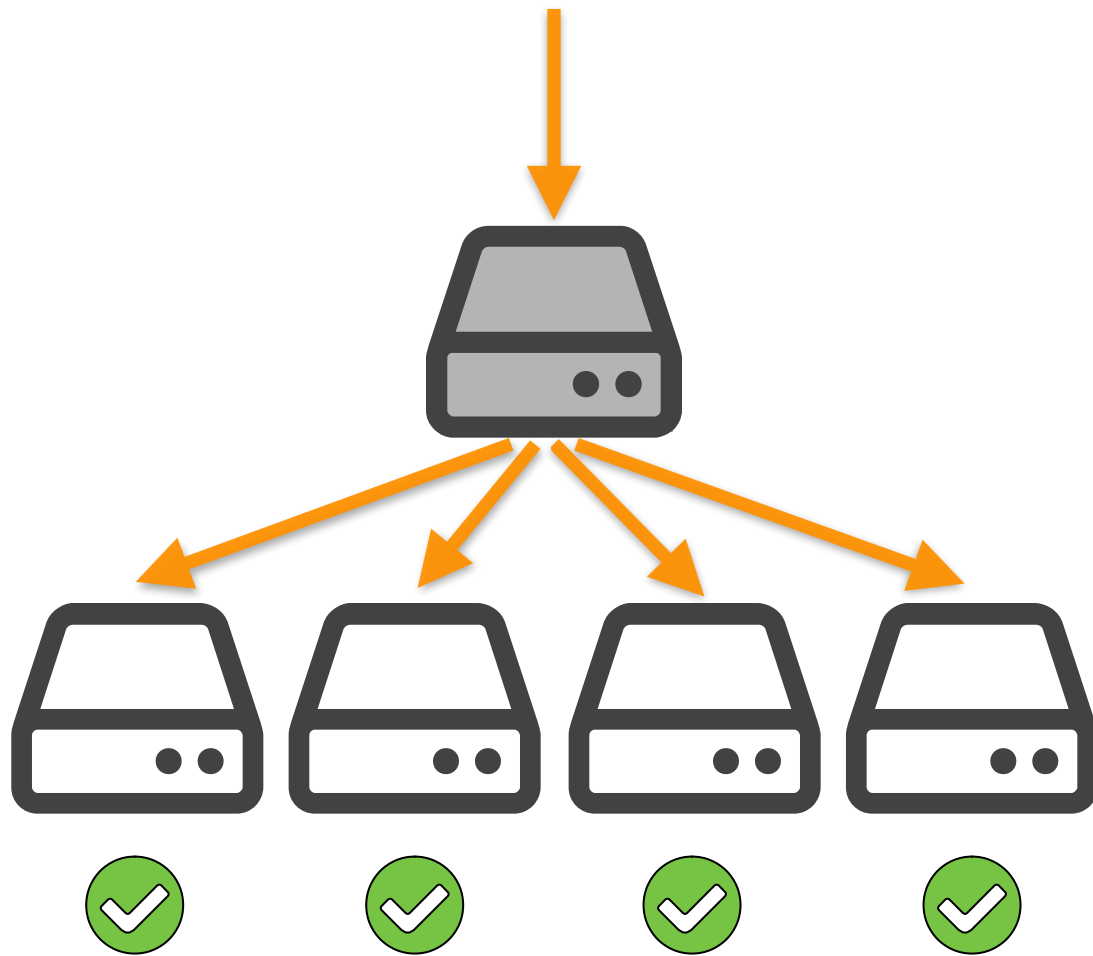


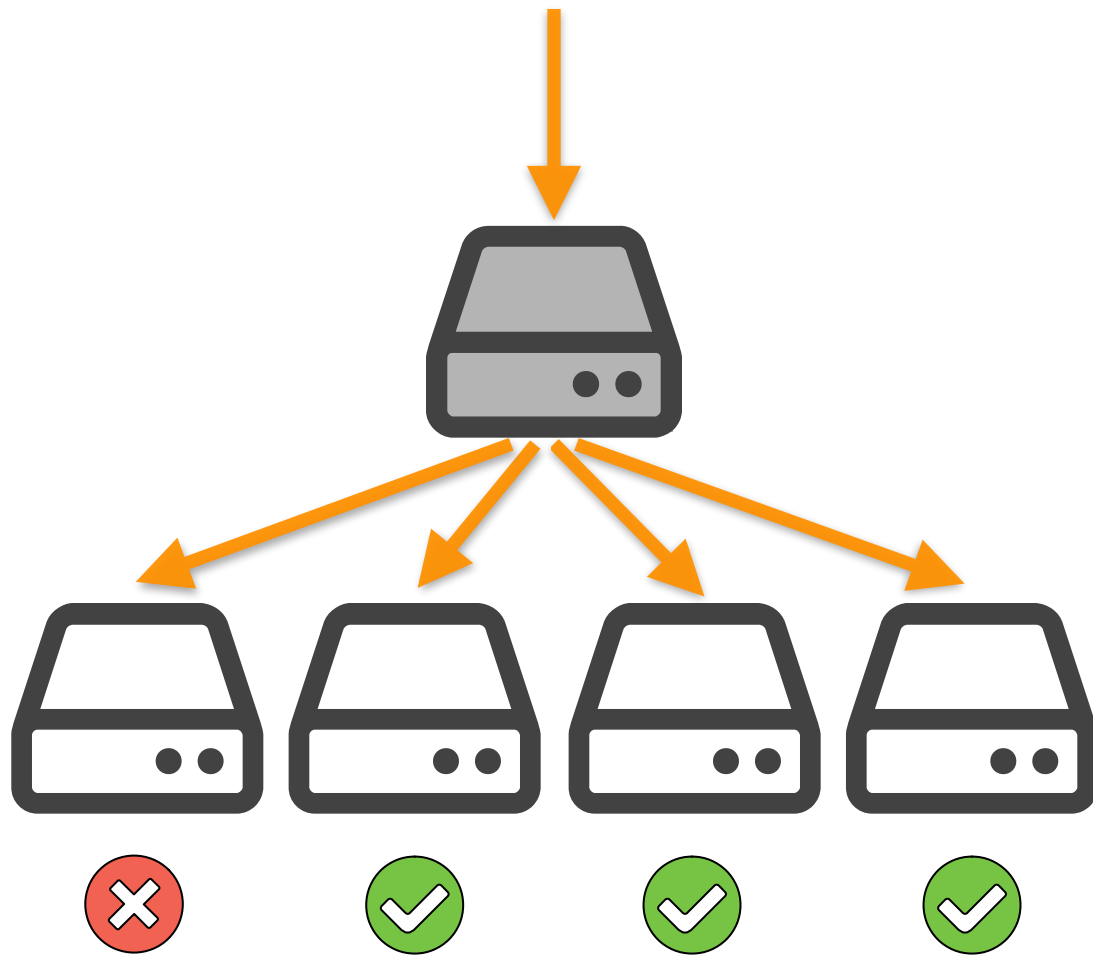


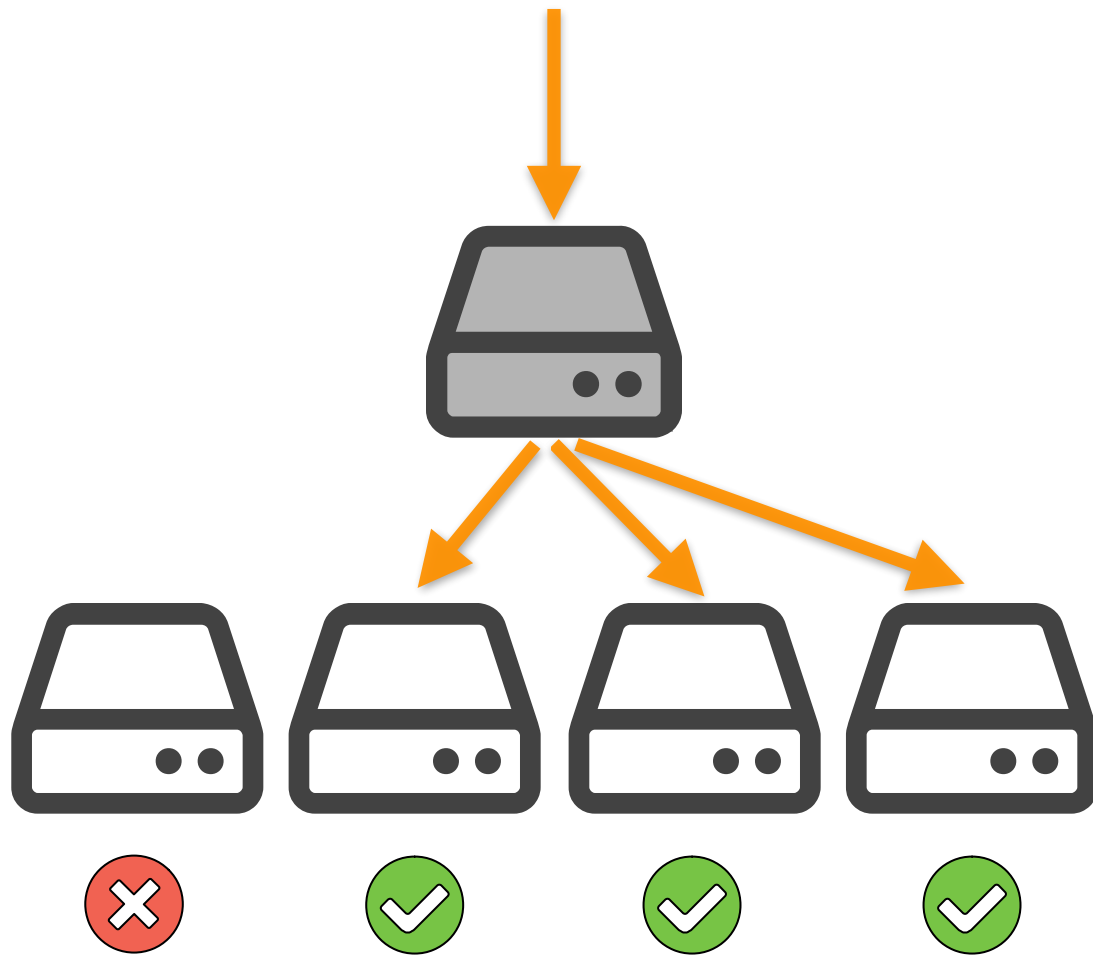


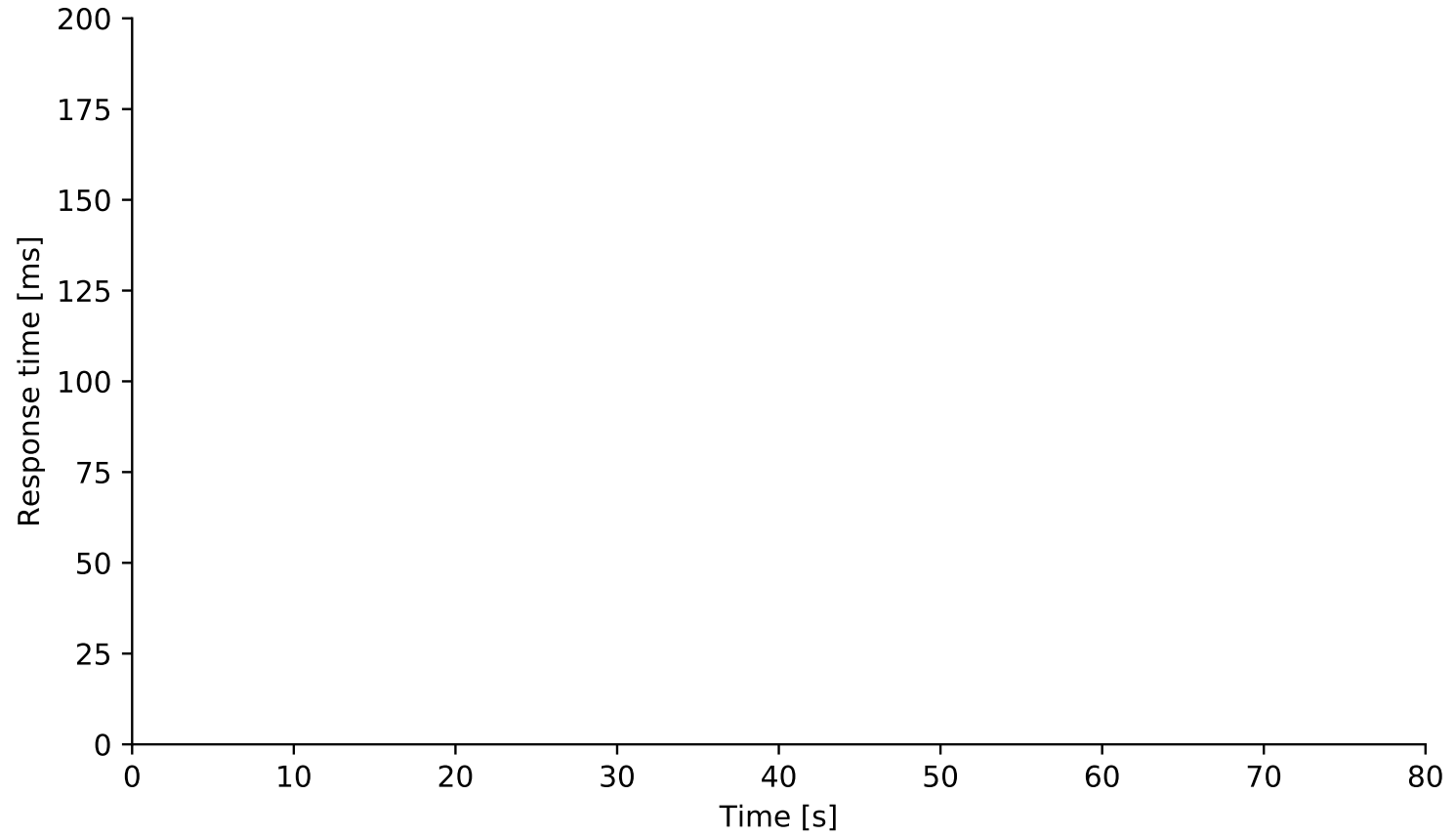


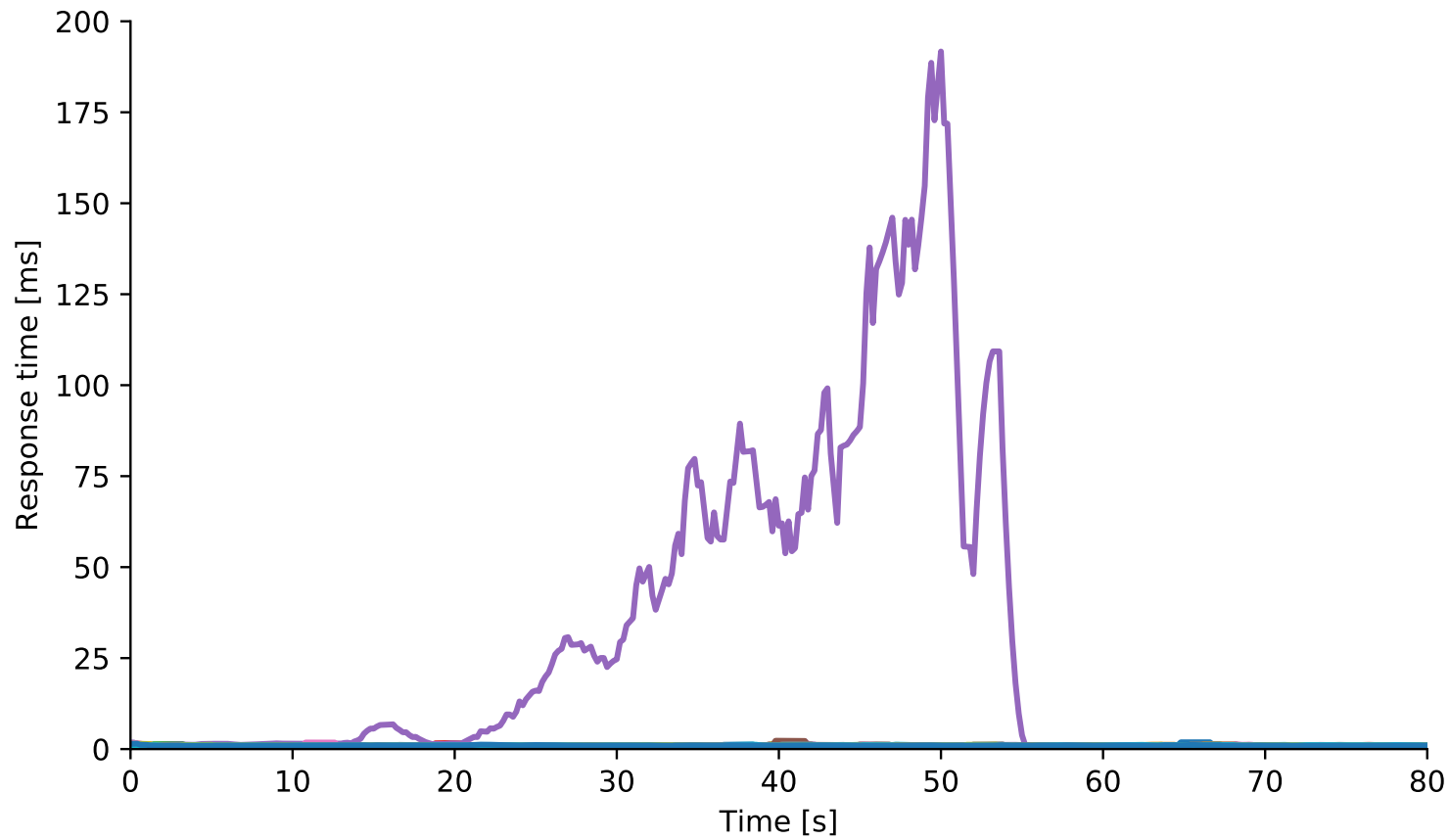


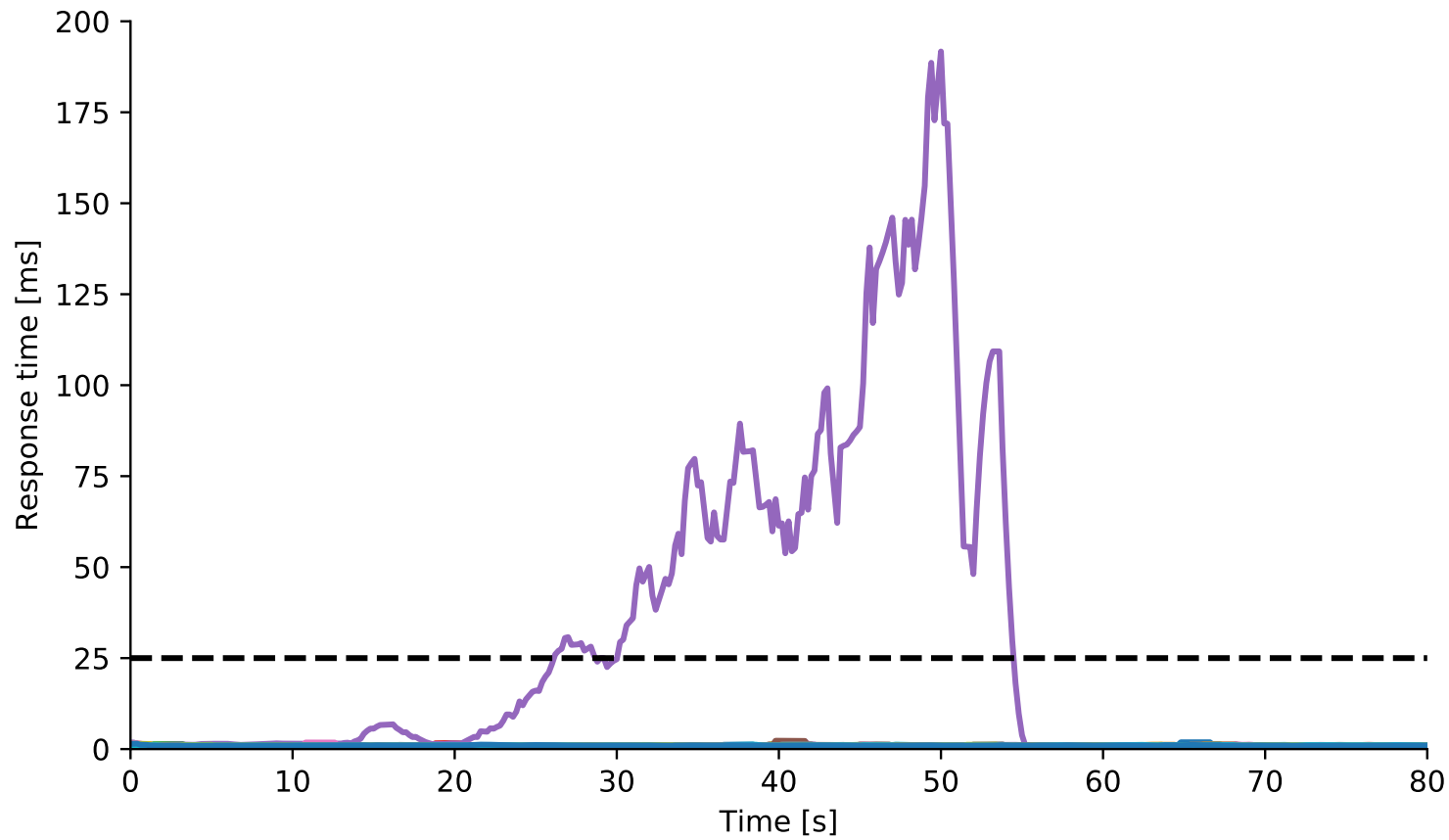


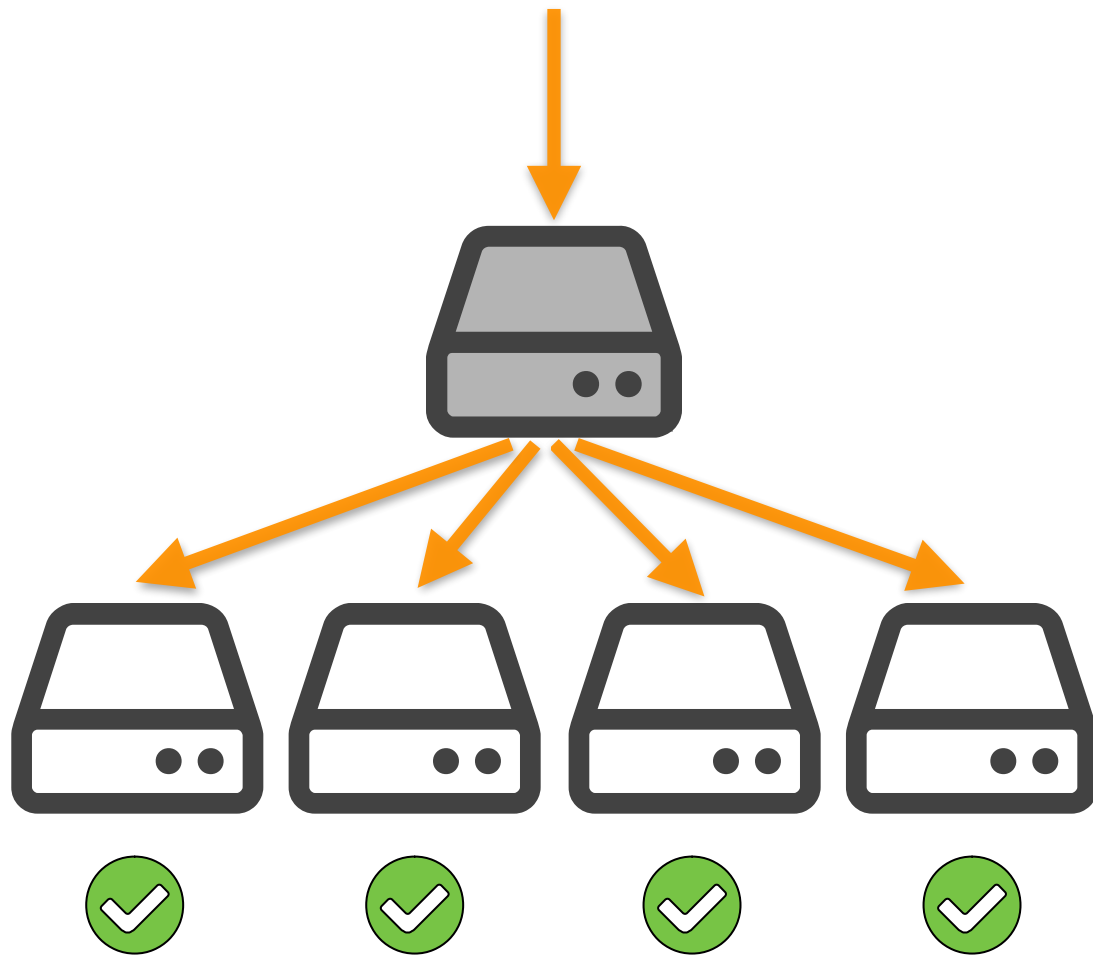


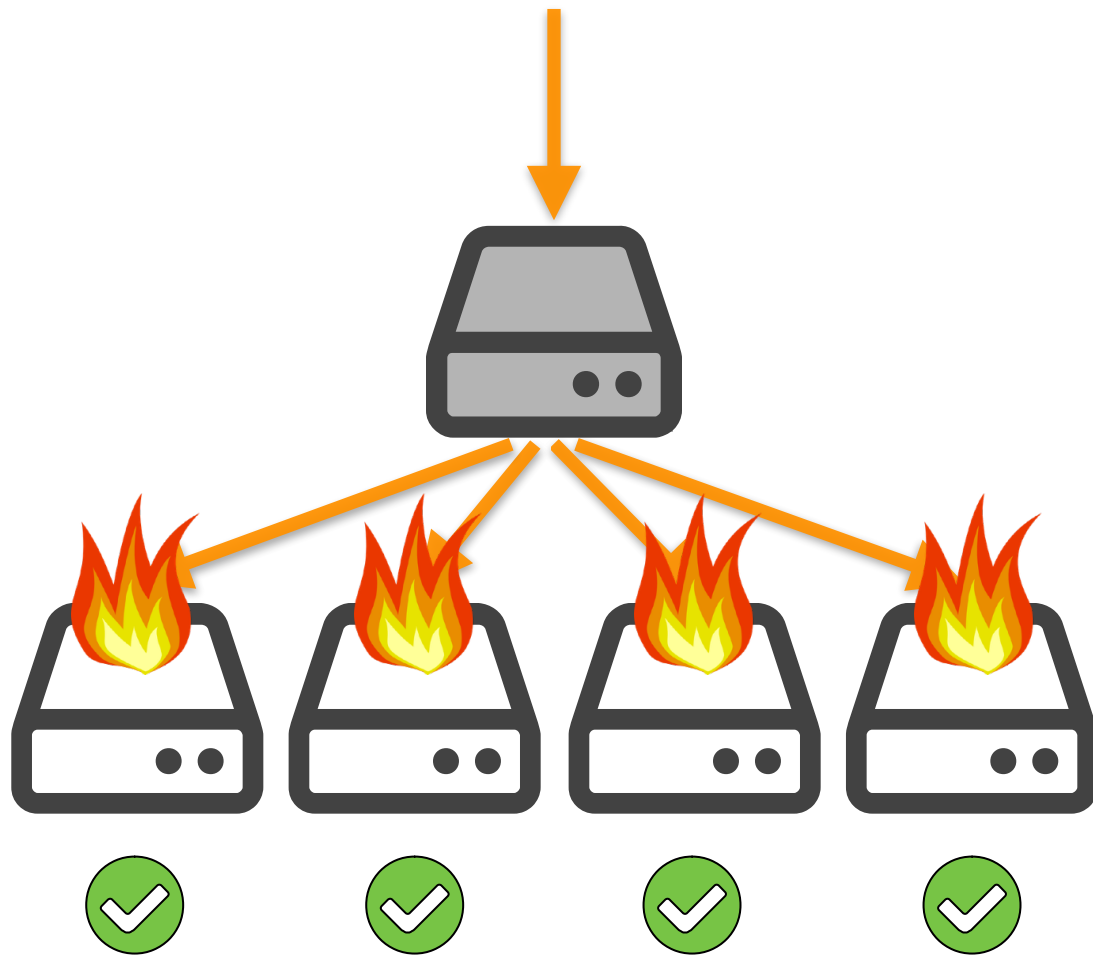


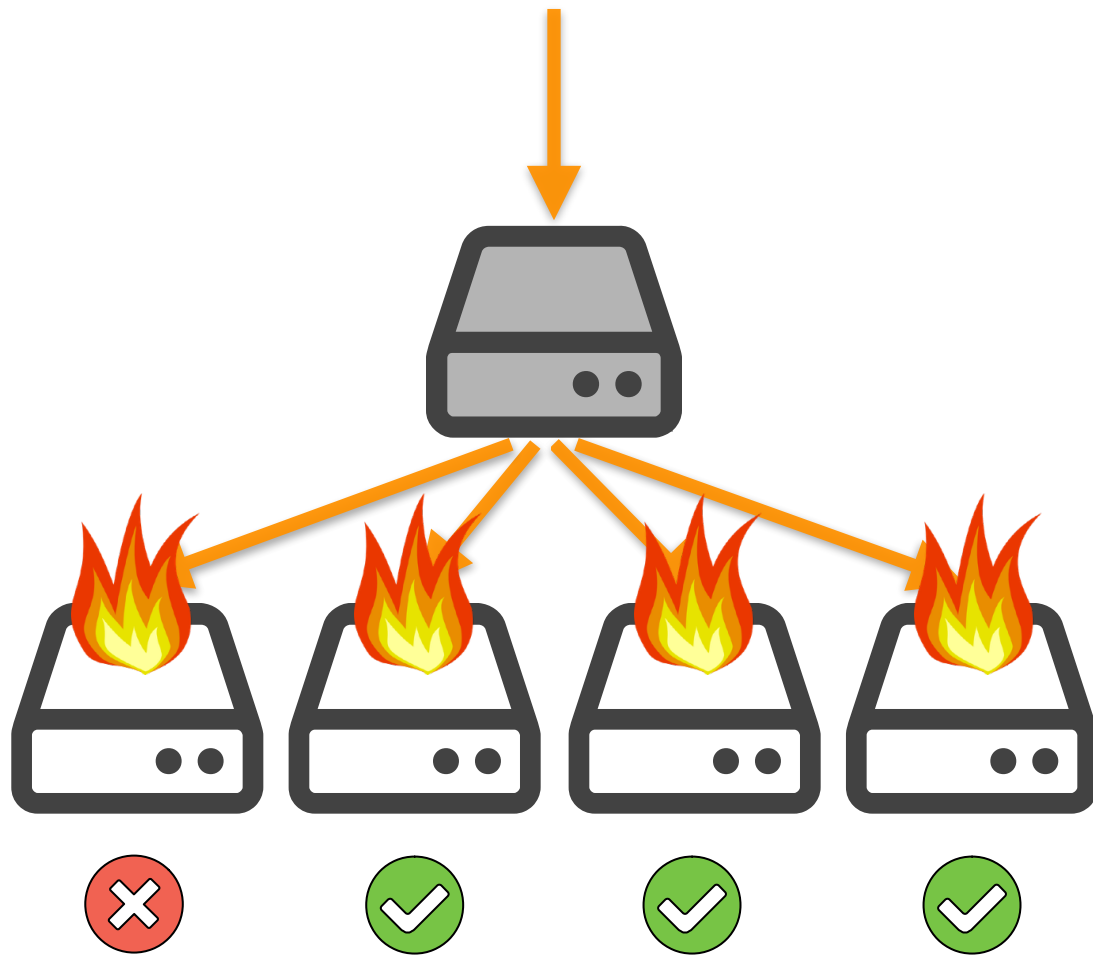


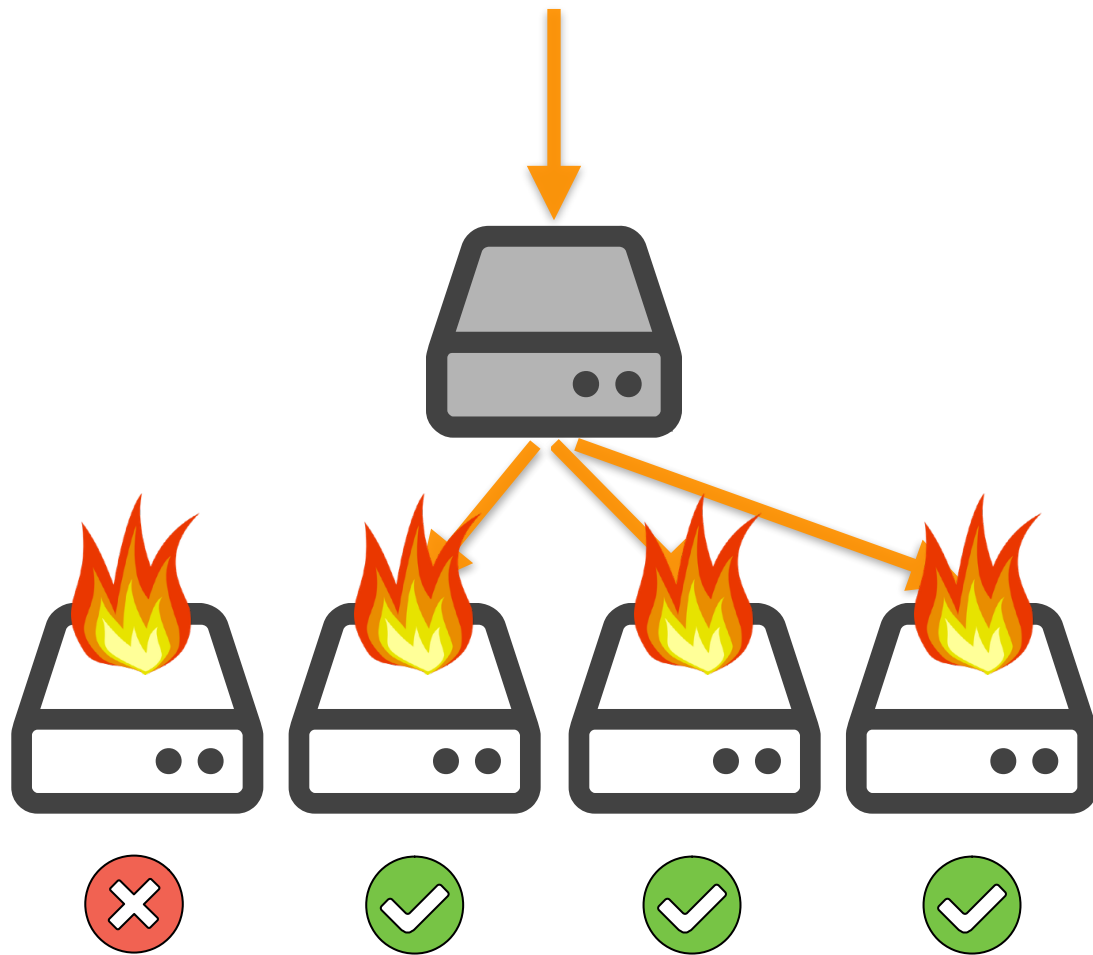


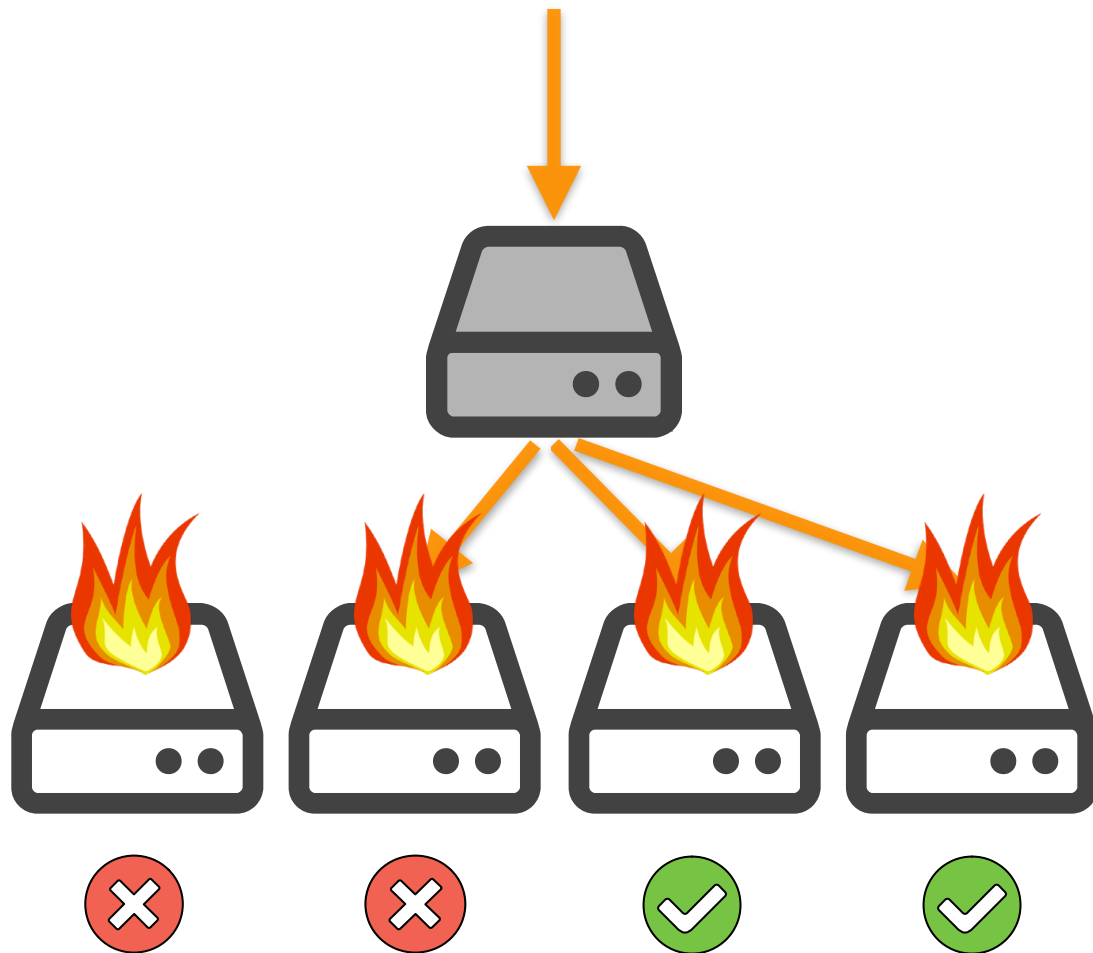


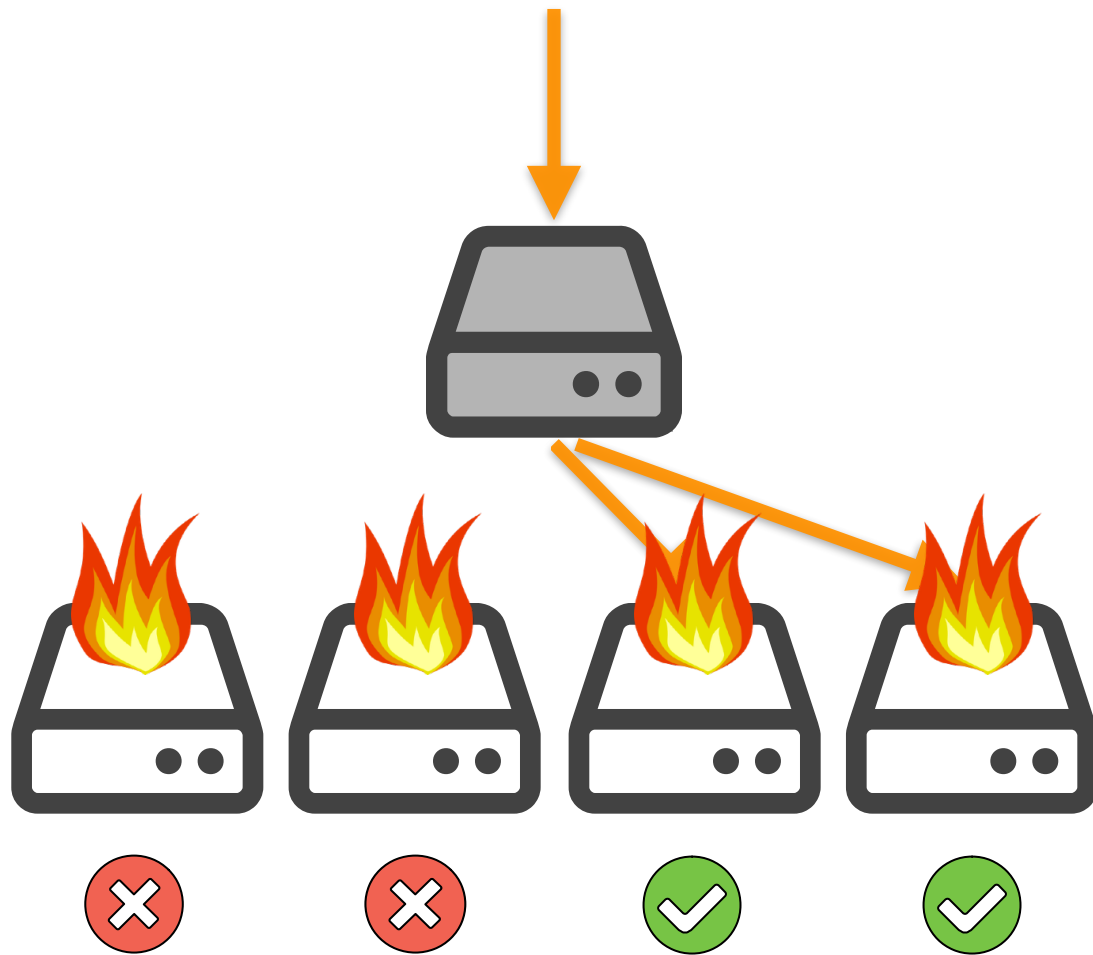


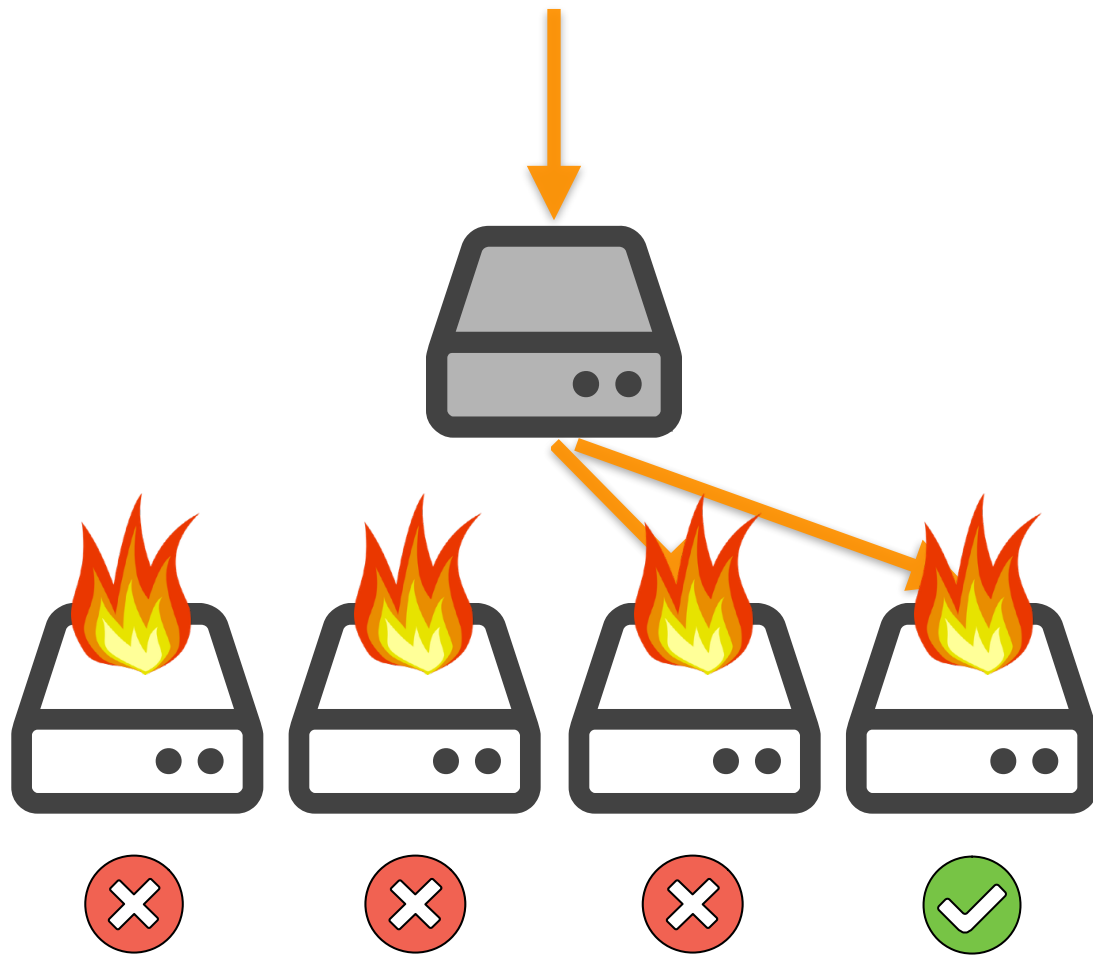


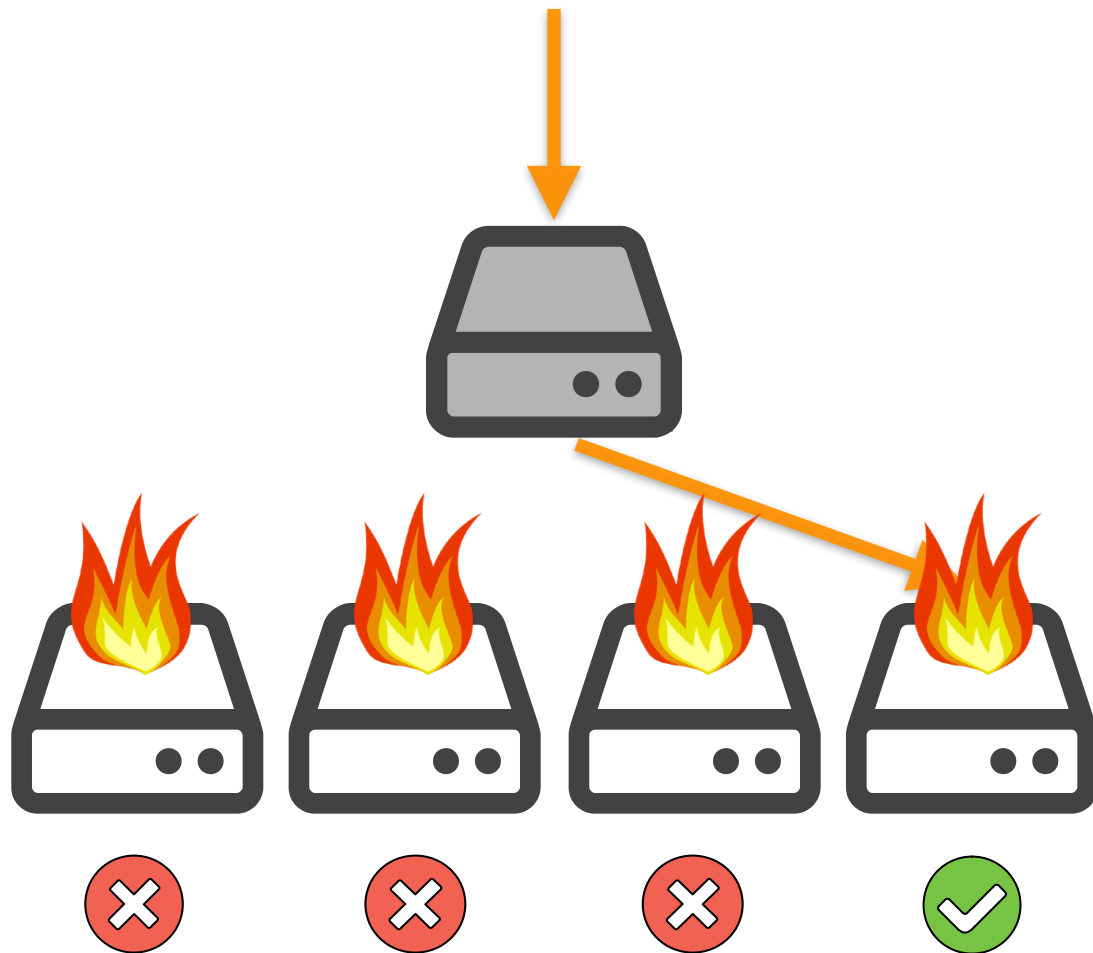


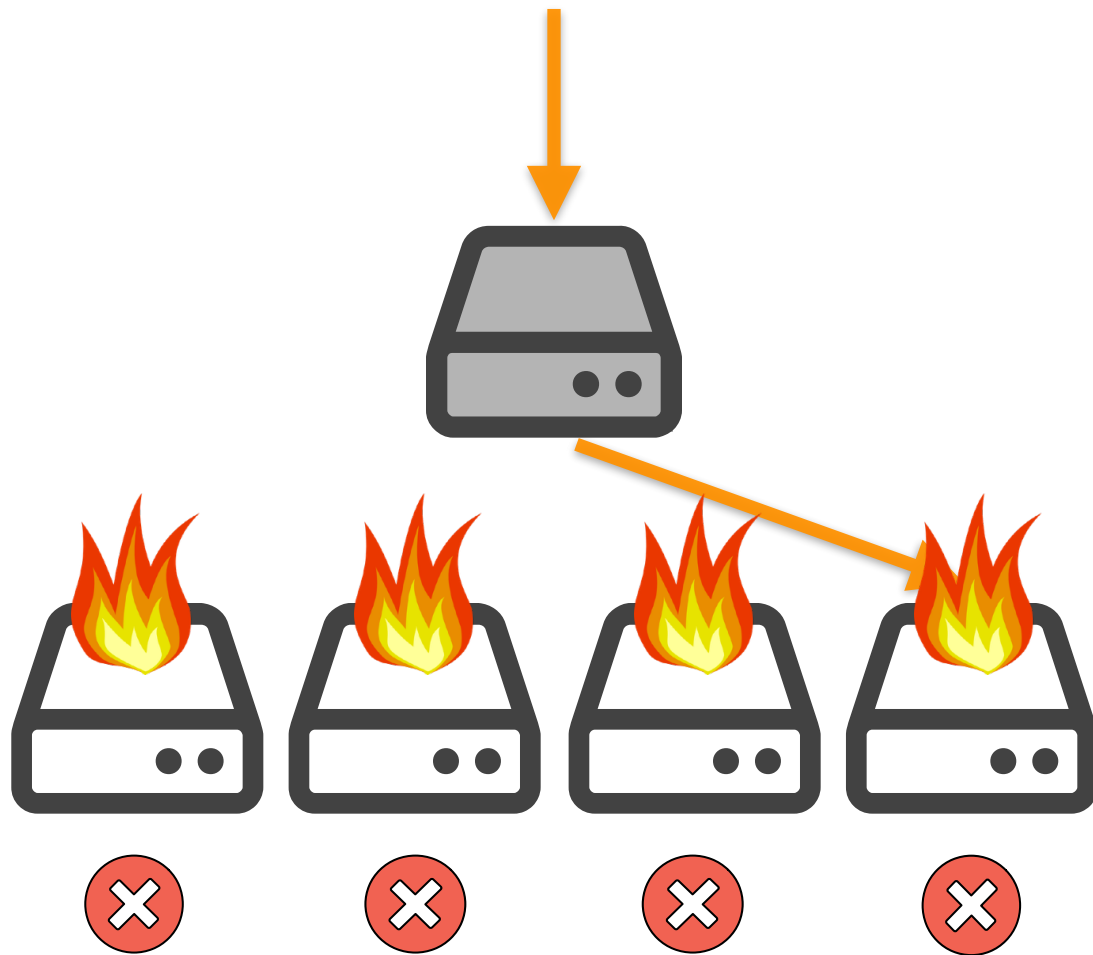


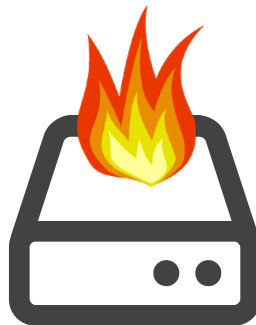
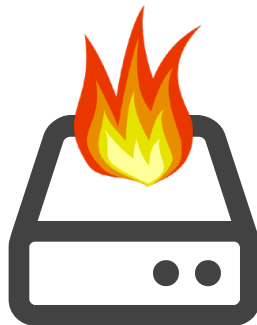
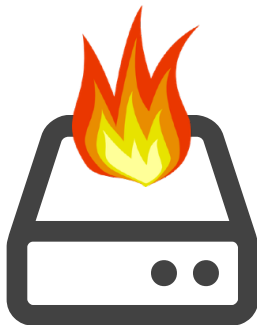
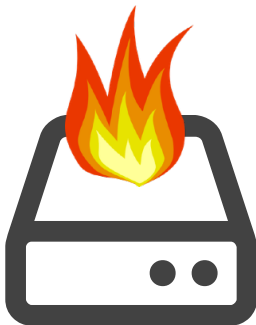
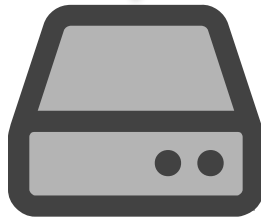


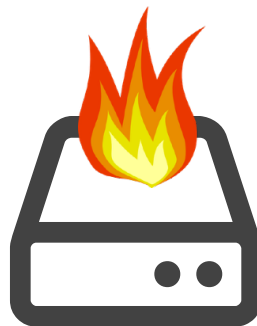
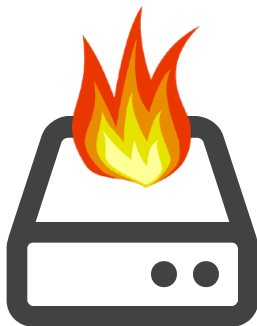
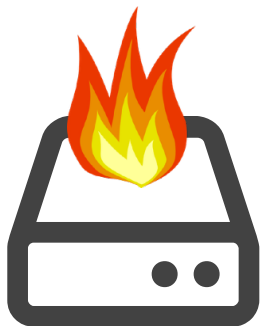


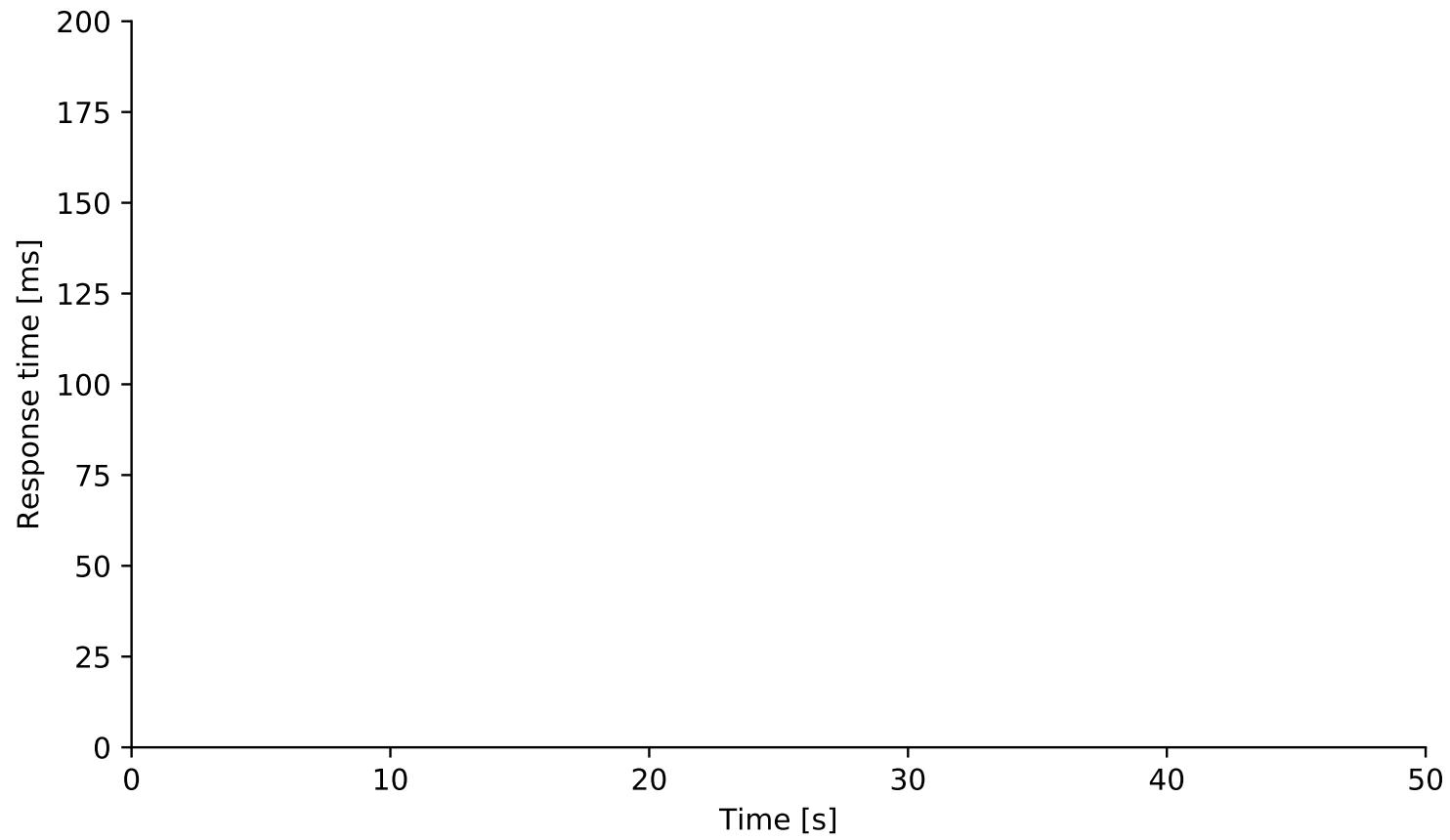


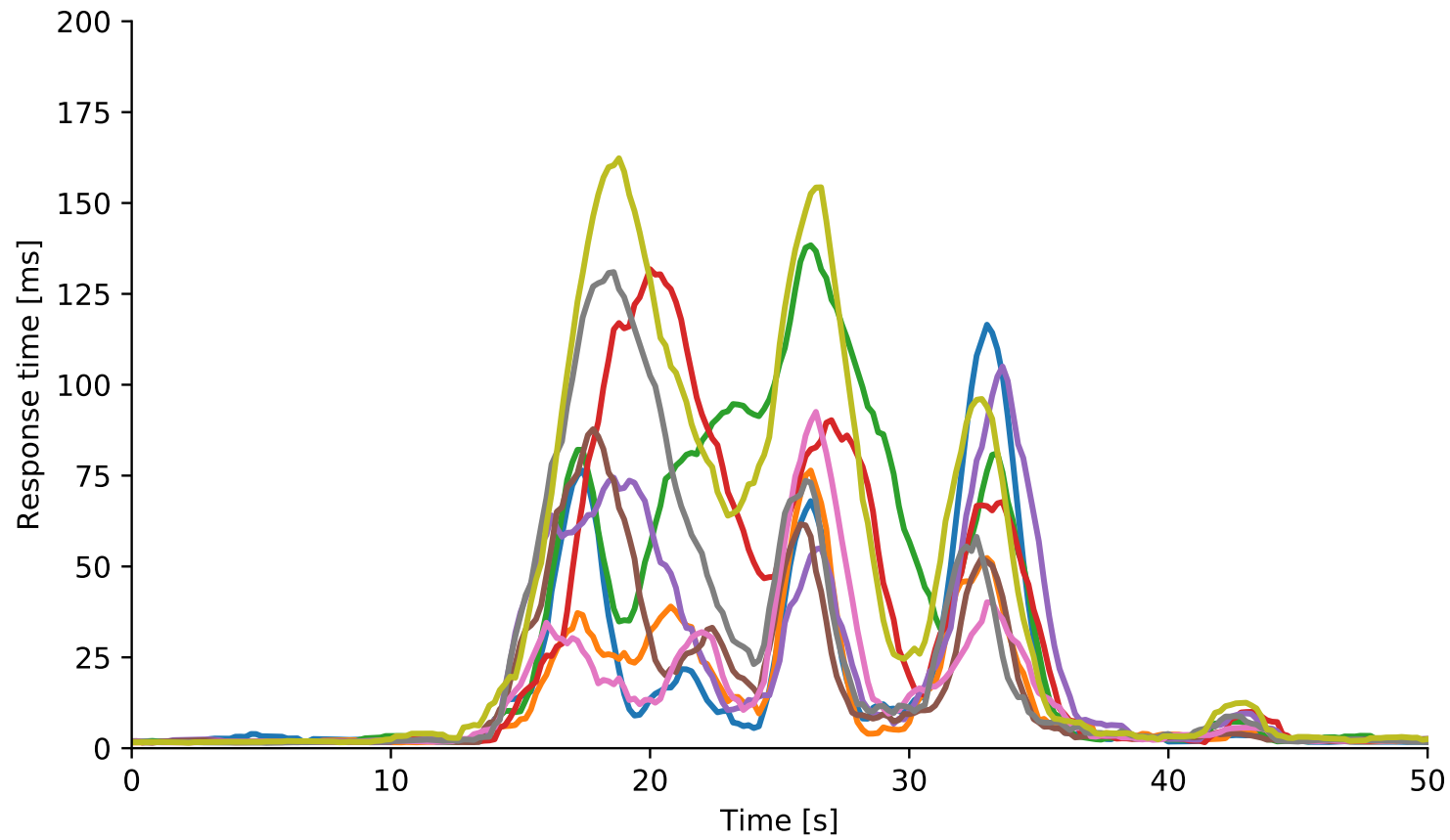


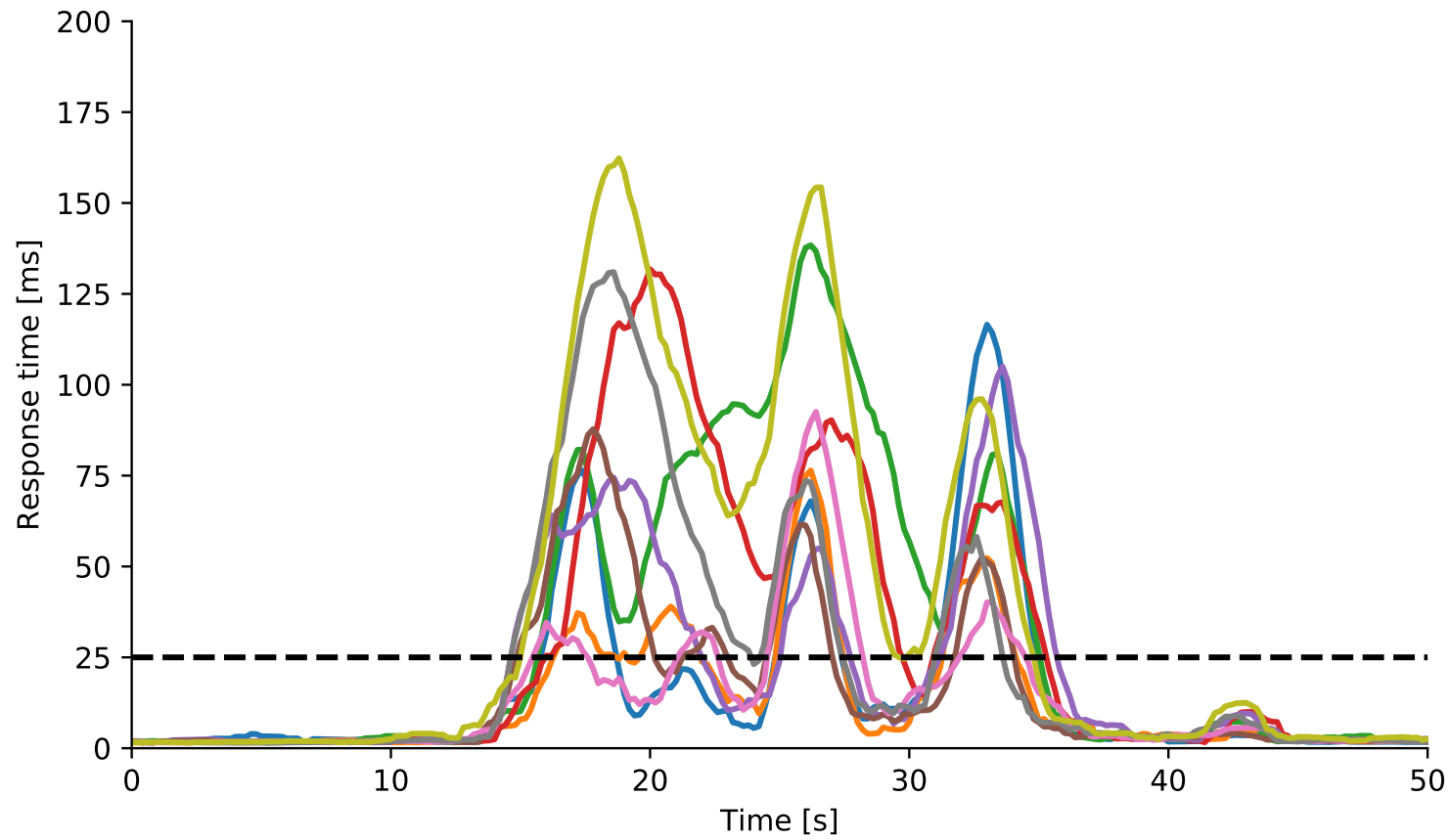


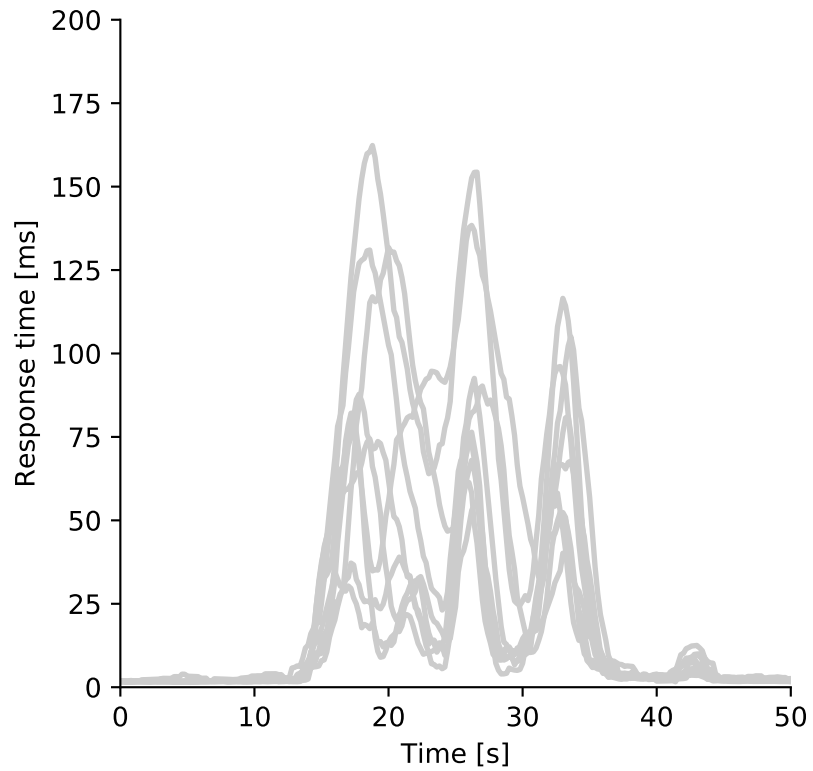
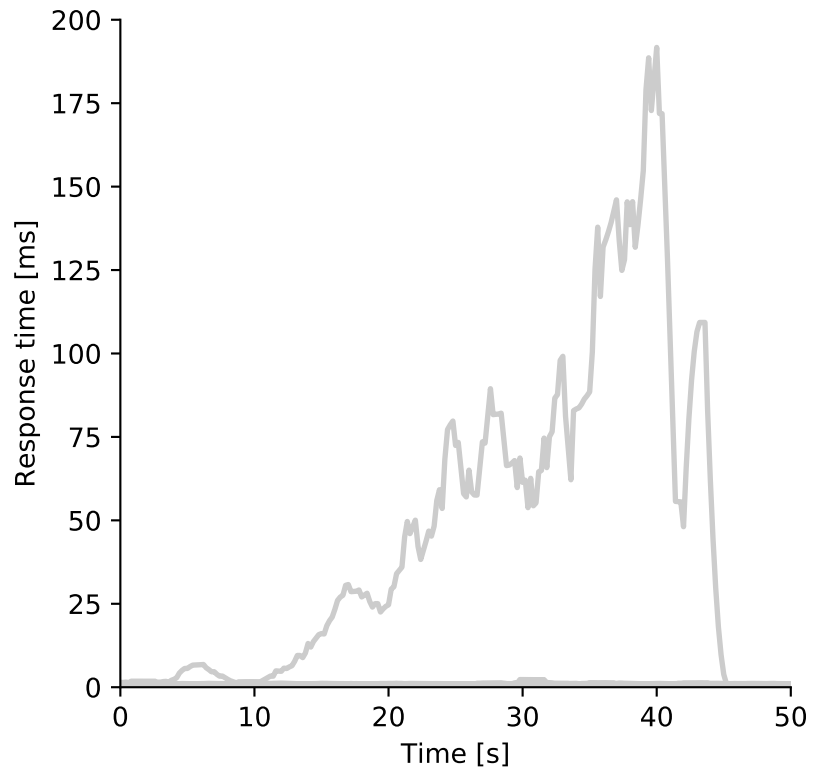


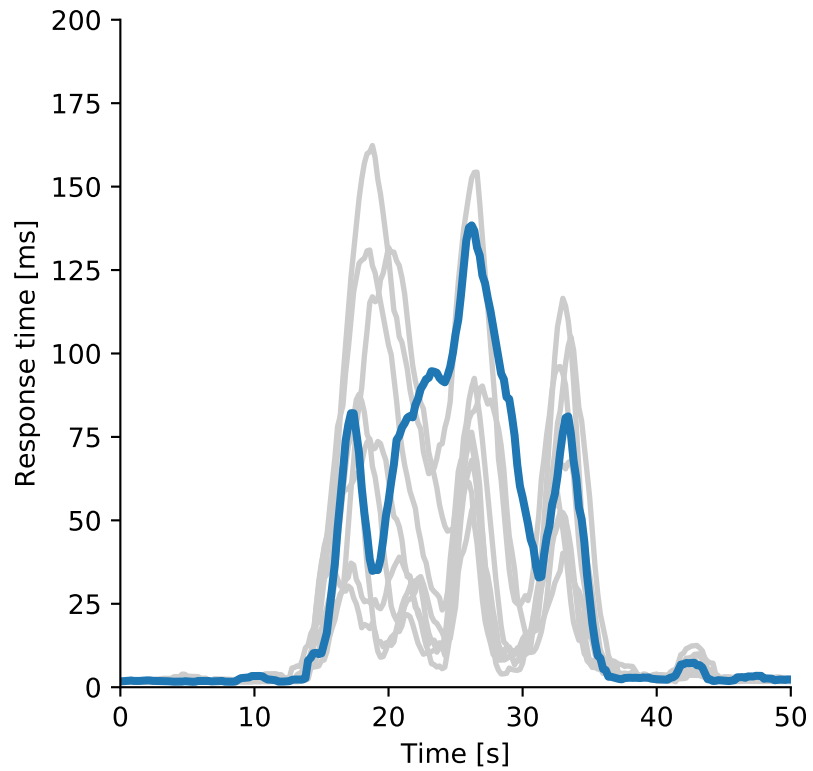
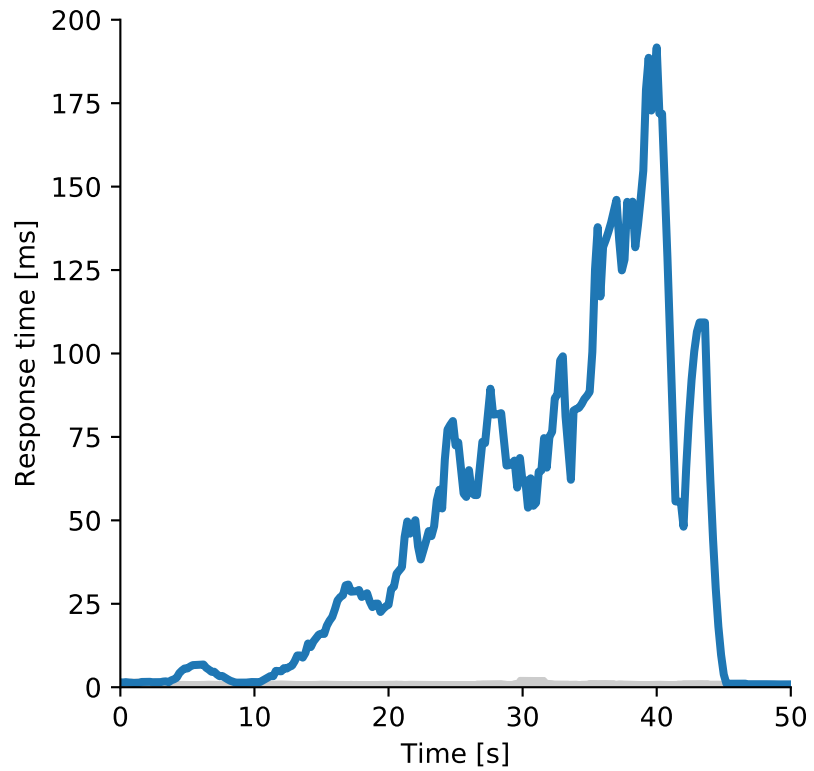






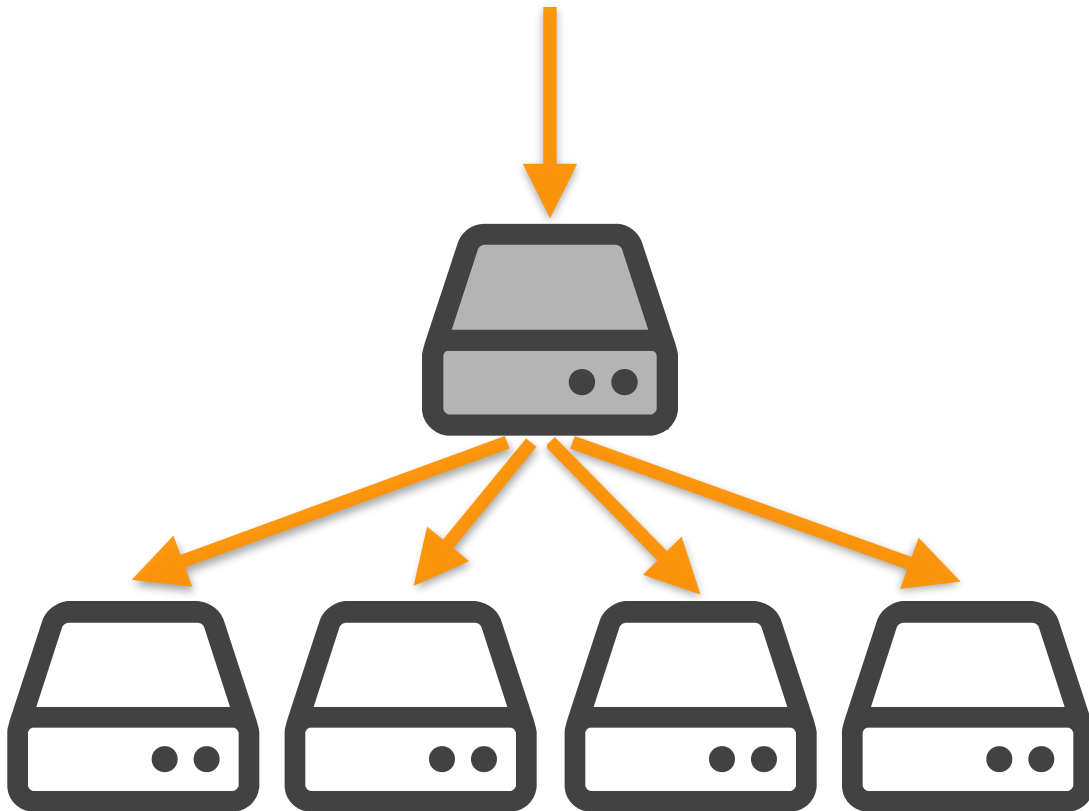


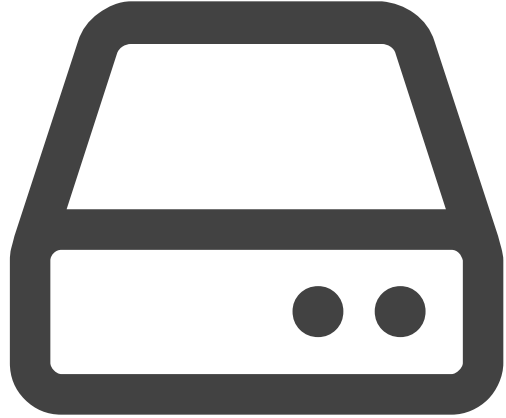


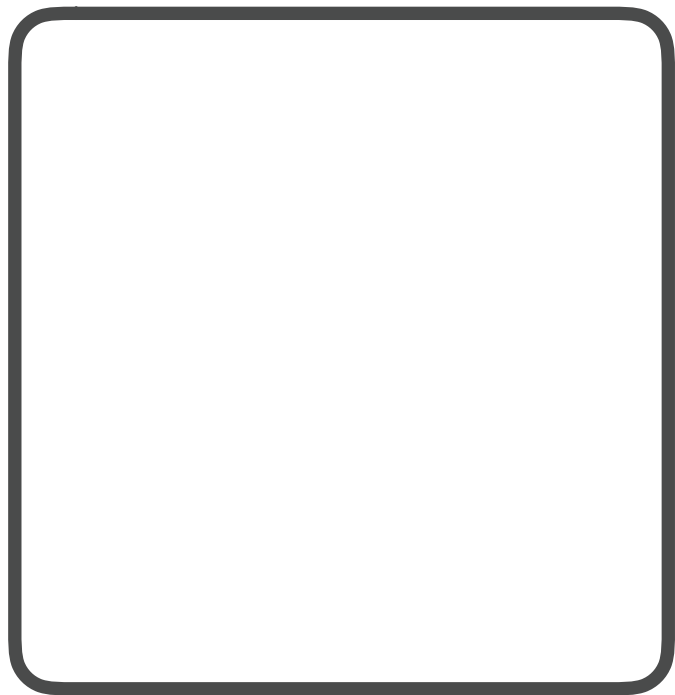


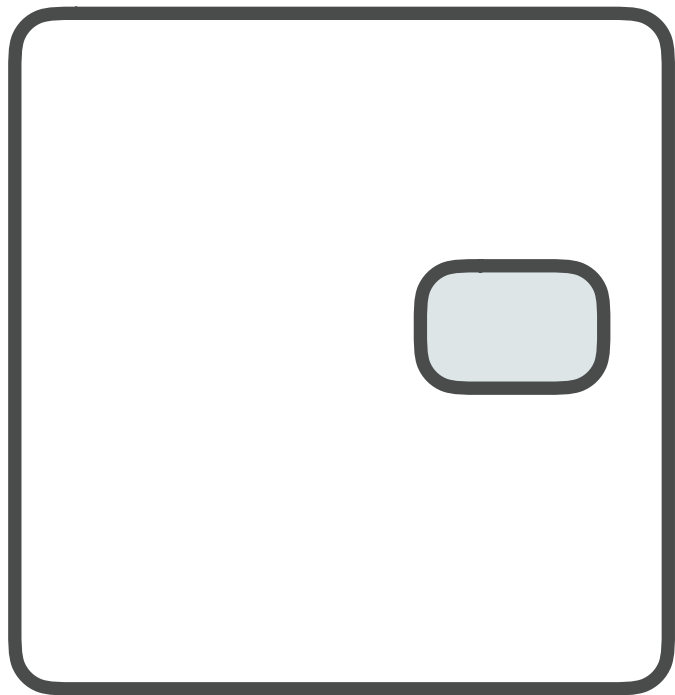


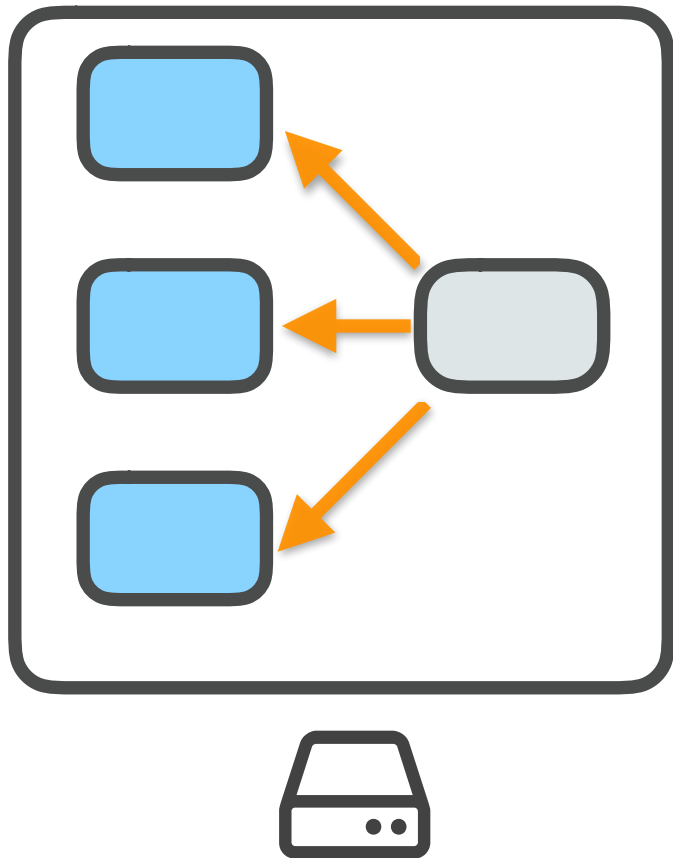
Architecture





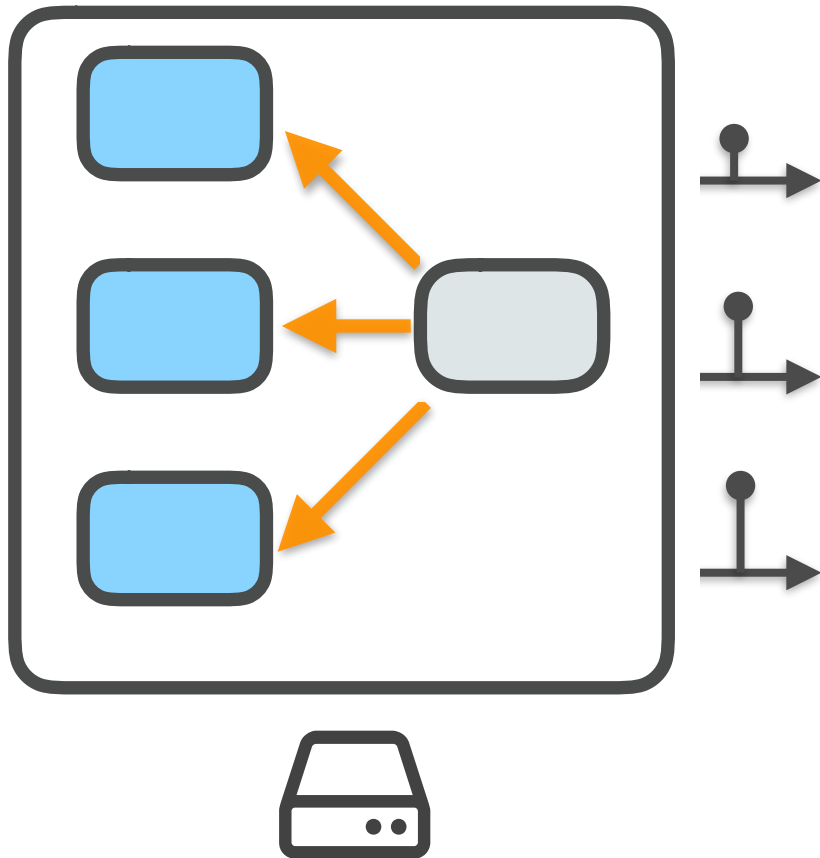






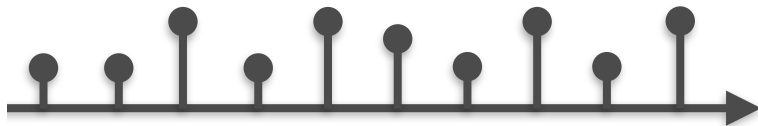
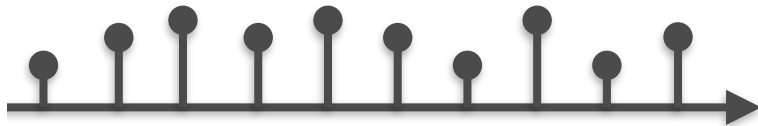
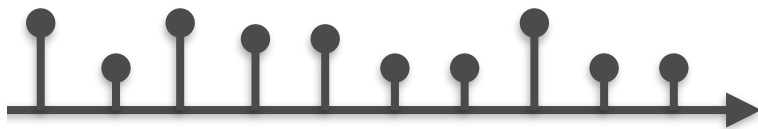
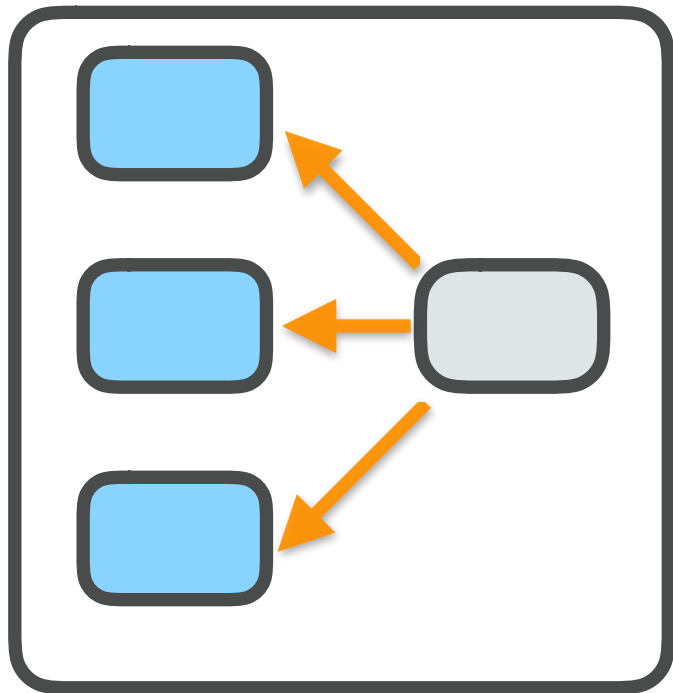
Observable signals:

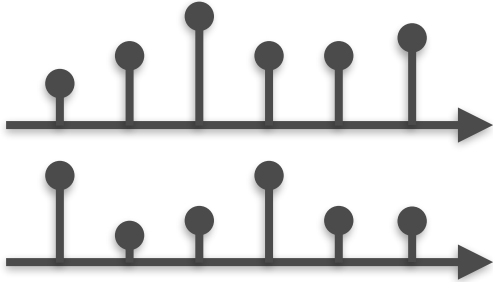
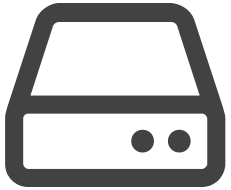
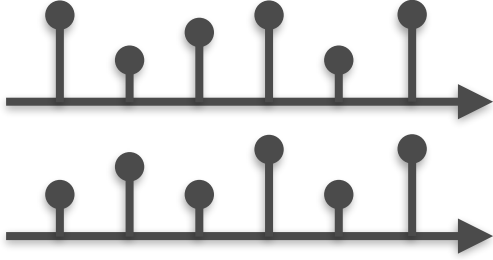
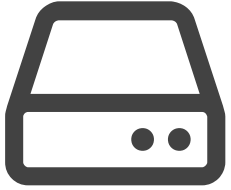
- Service error rate
- Service response time
- Process uptime
- CPU load
- Available memory
- Rate of requests served
- ...

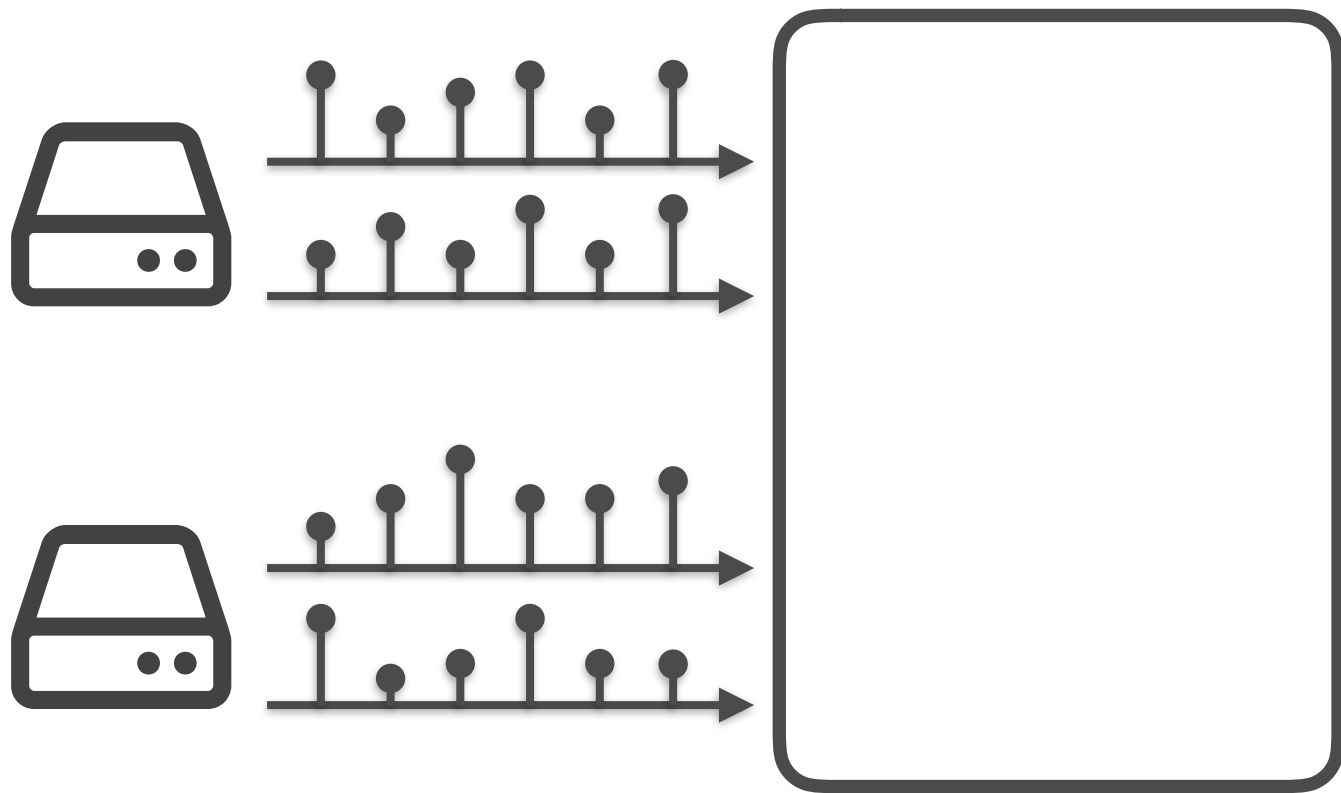


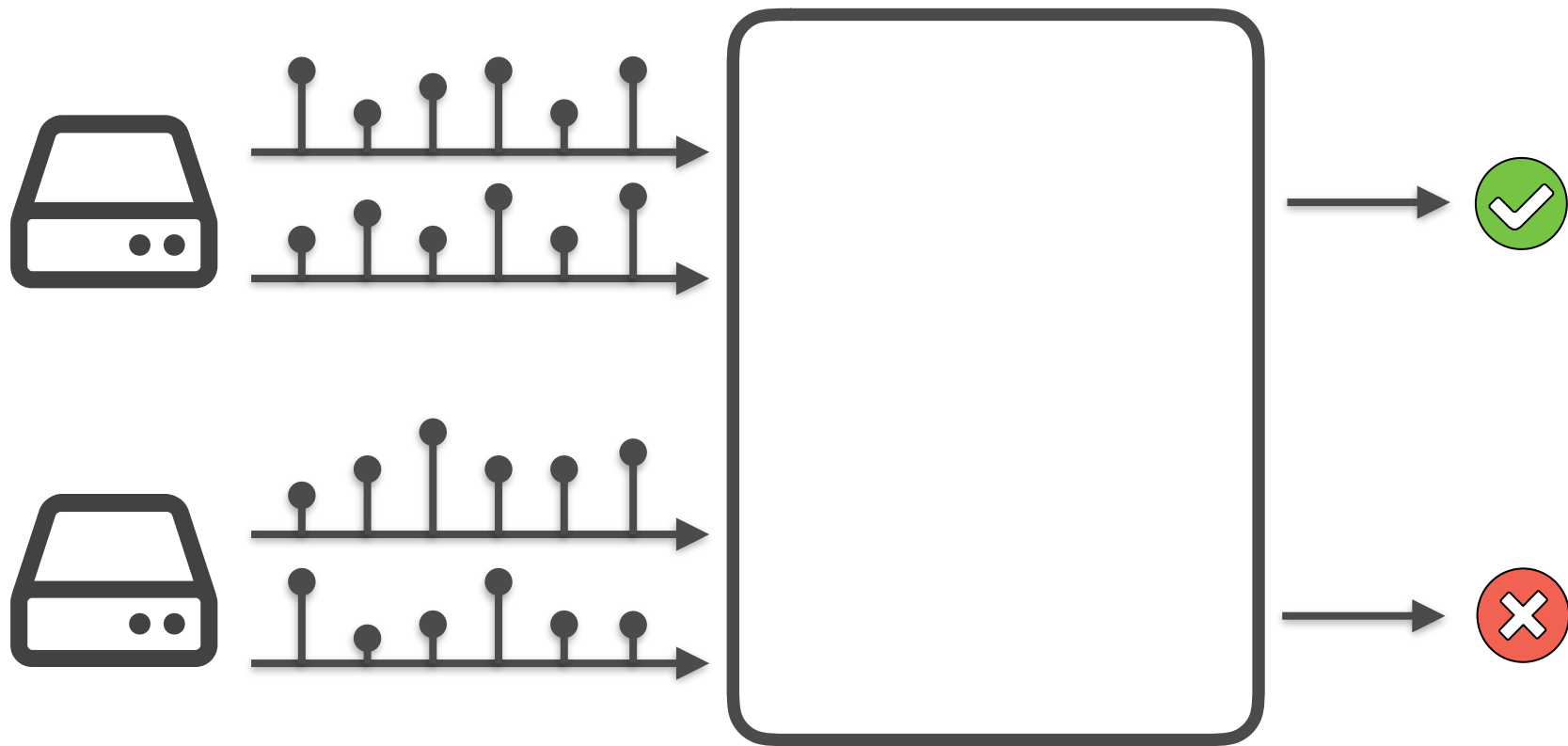
Observable signals:

- Service error rate
- Service response time
- Process uptime
- CPU load
- Available memory
- Rate of requests served
- ...













Field	Concept
Machine learning	Classifier
Signal processing	Filter
Control theory	Controller



Field	Concept
Machine learning	Classifier
Signal processing	Filter
Control theory	Controller



Design



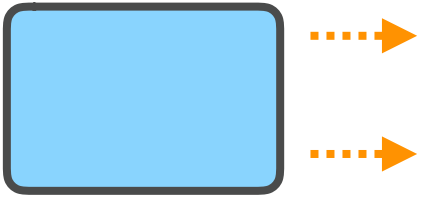
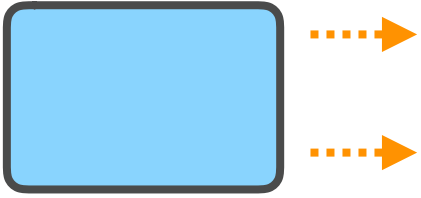
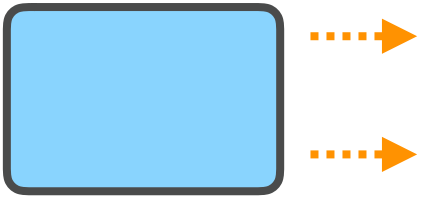
HTTP reverse proxy

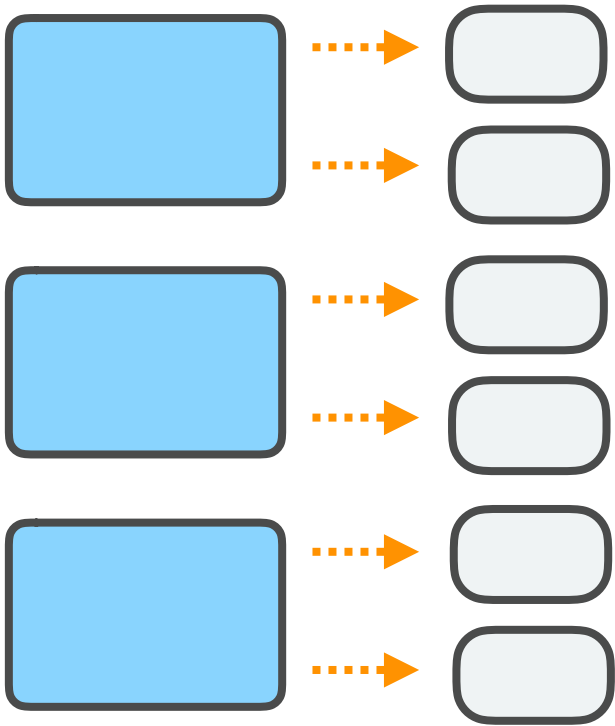




Request error rate

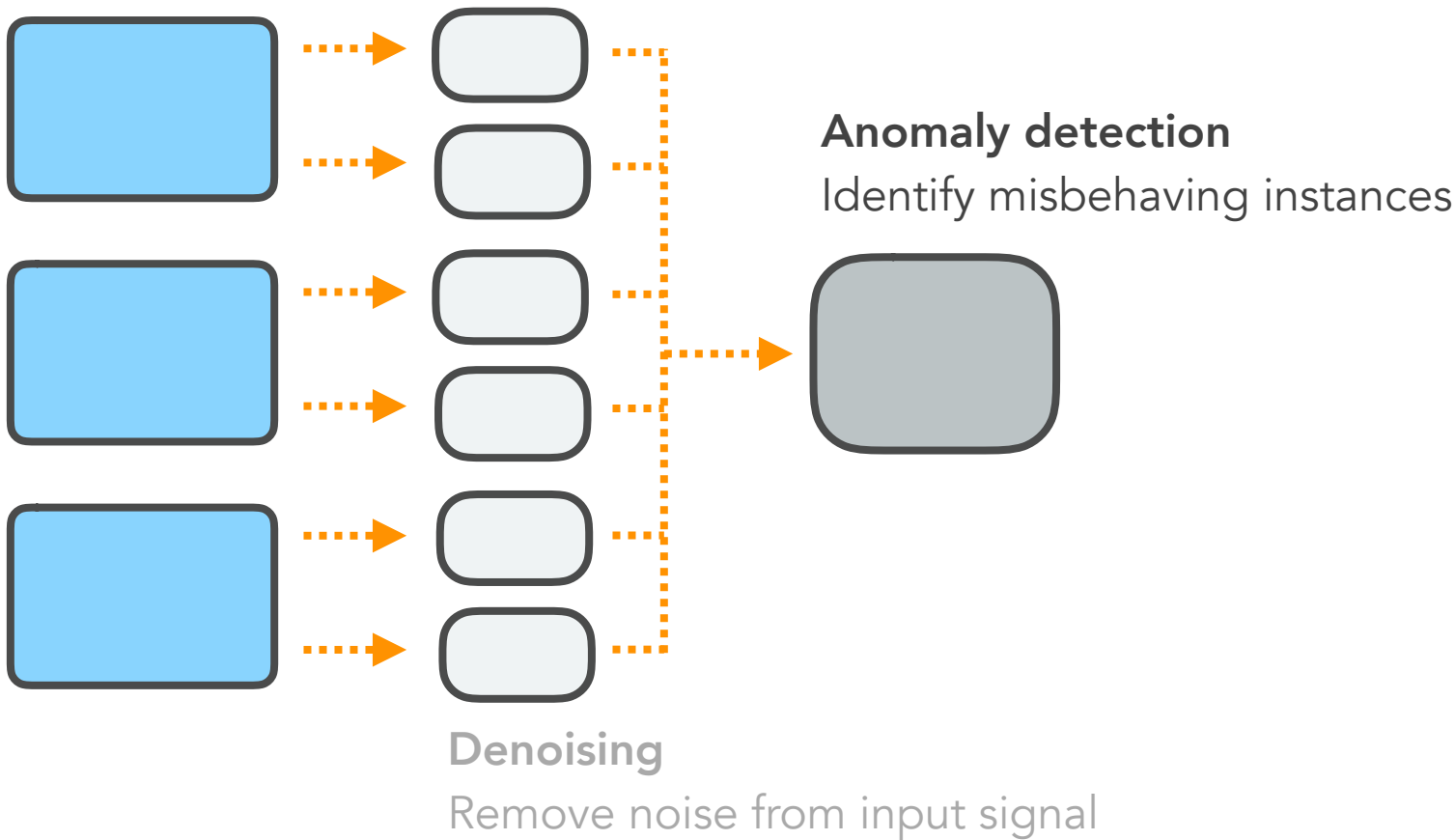
Response time

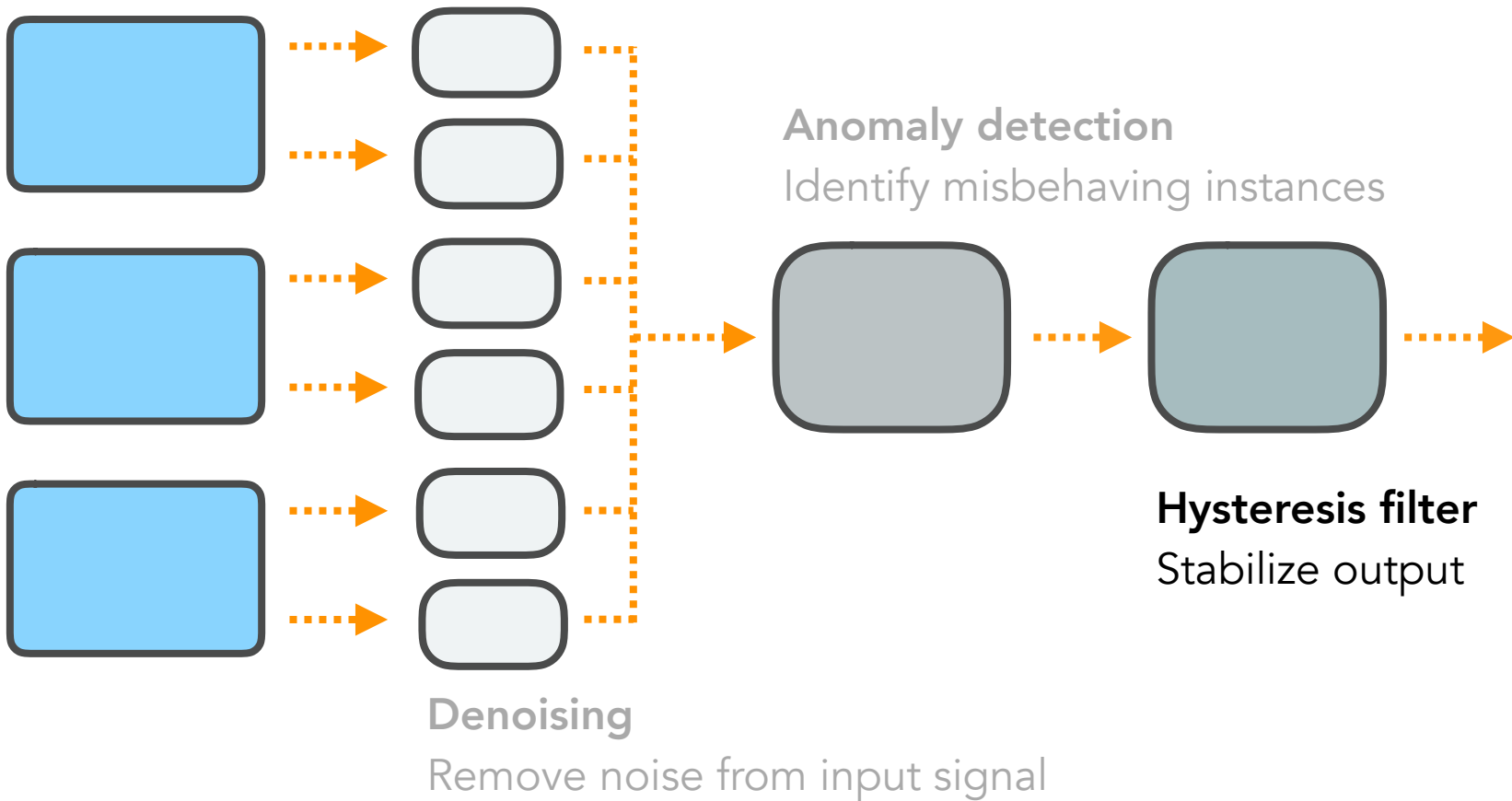


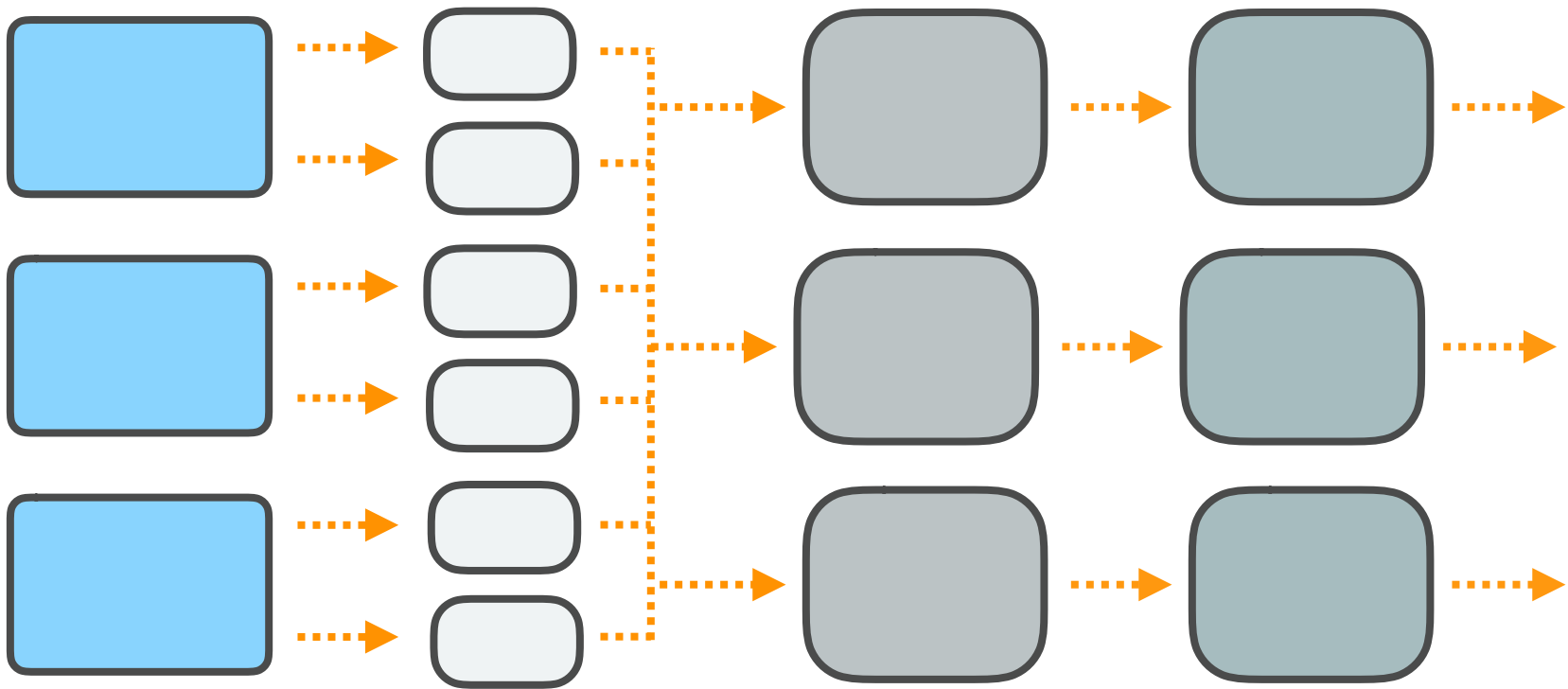


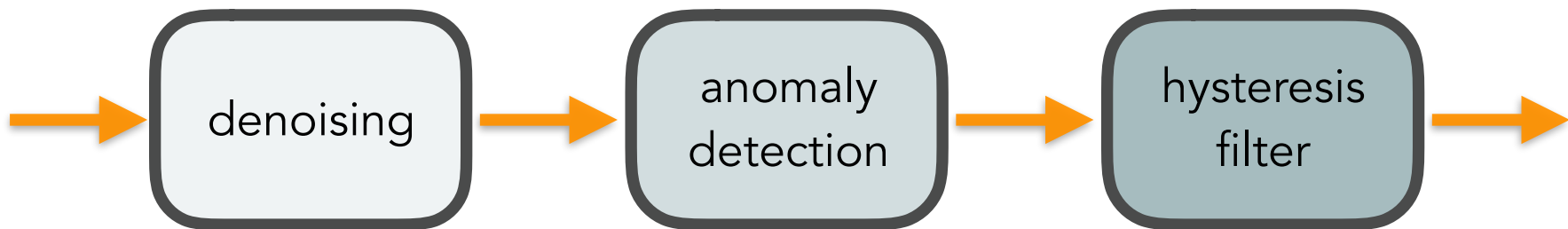
Denoising

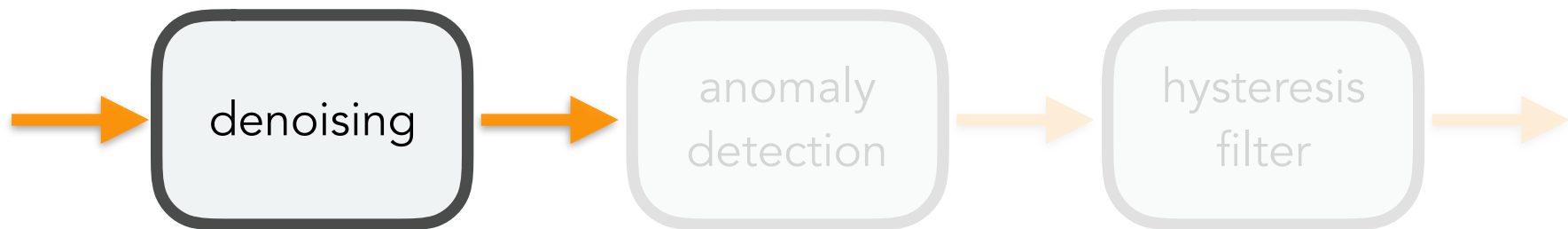
Remove noise from input signal

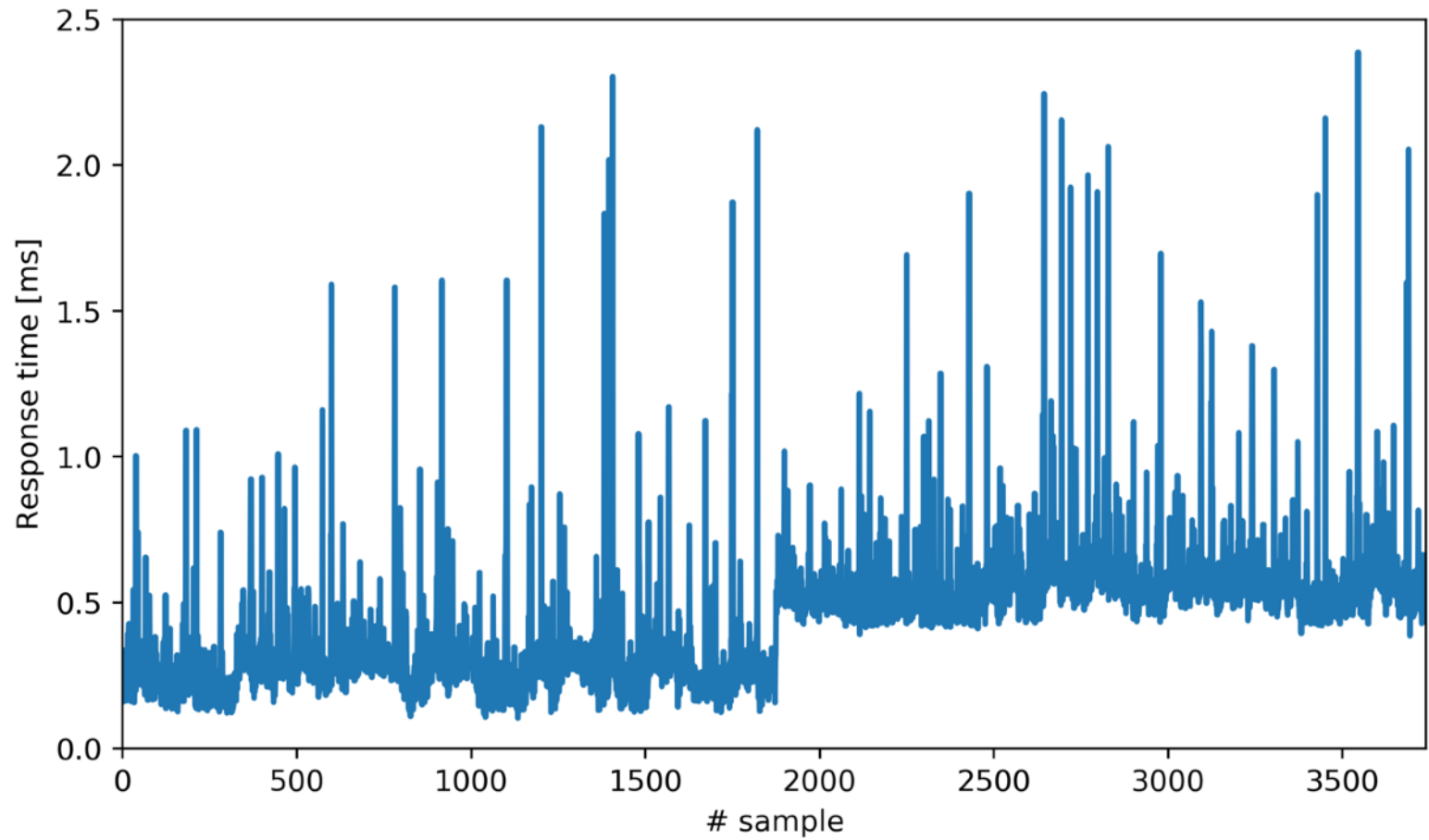


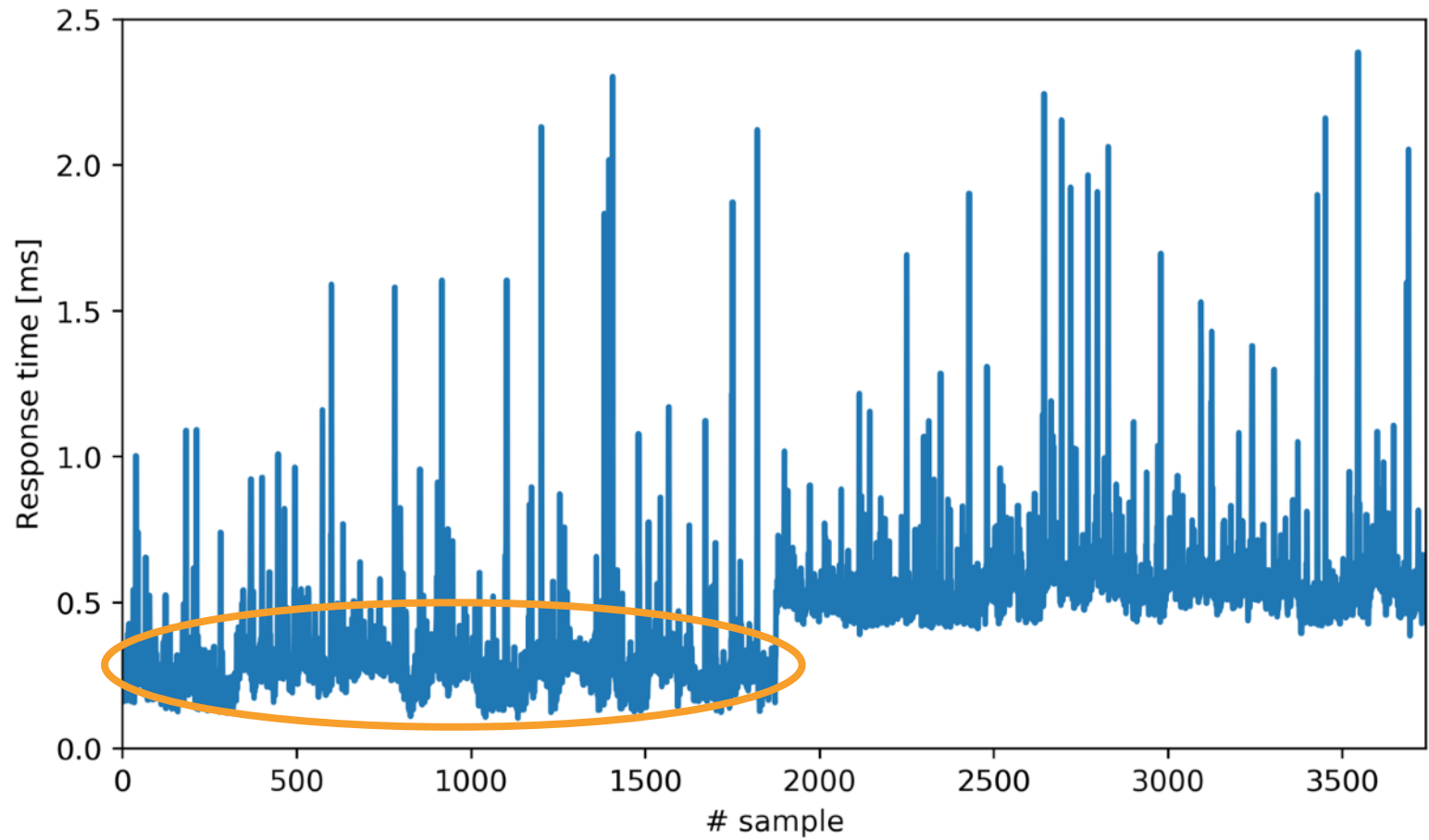


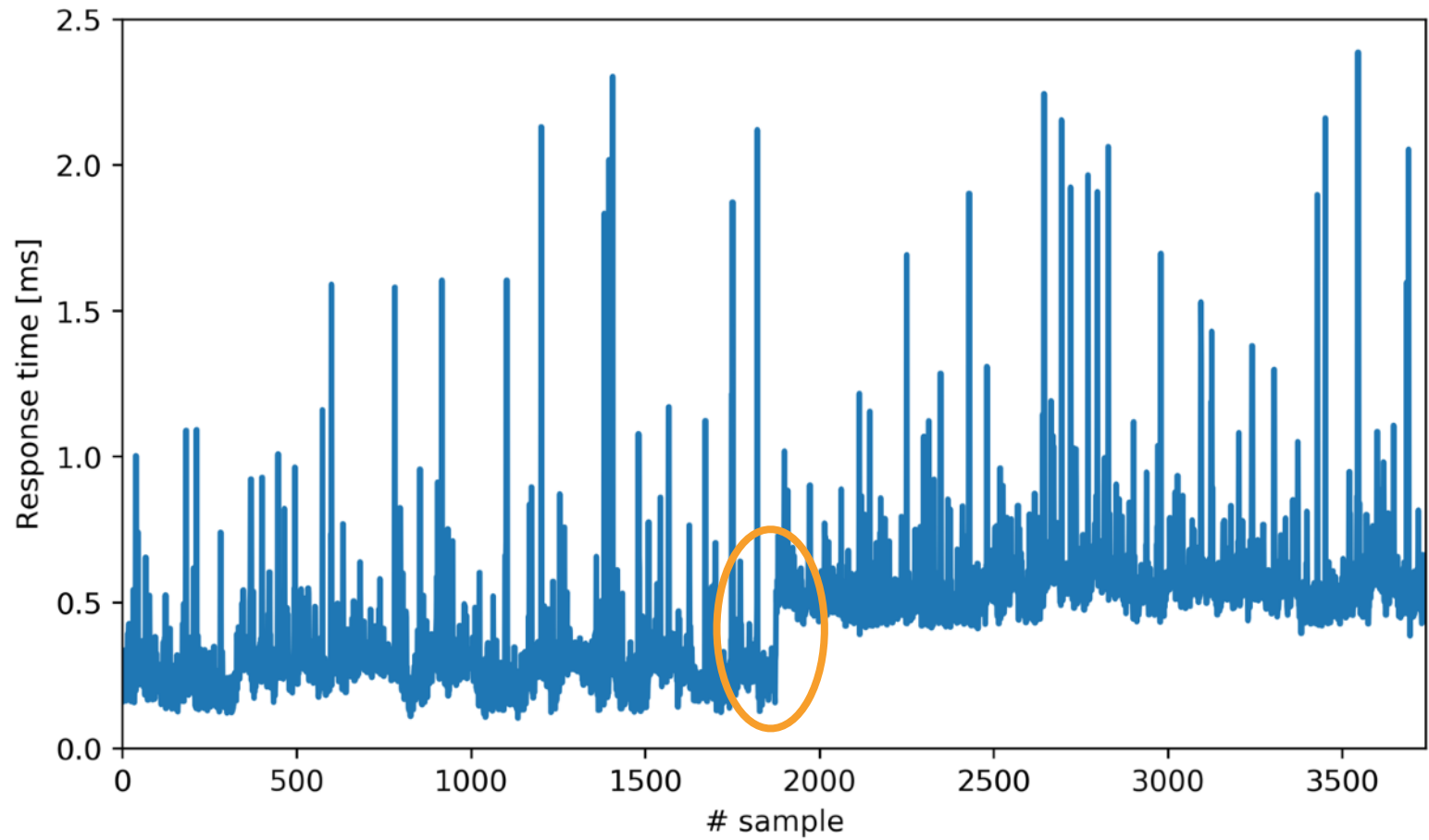


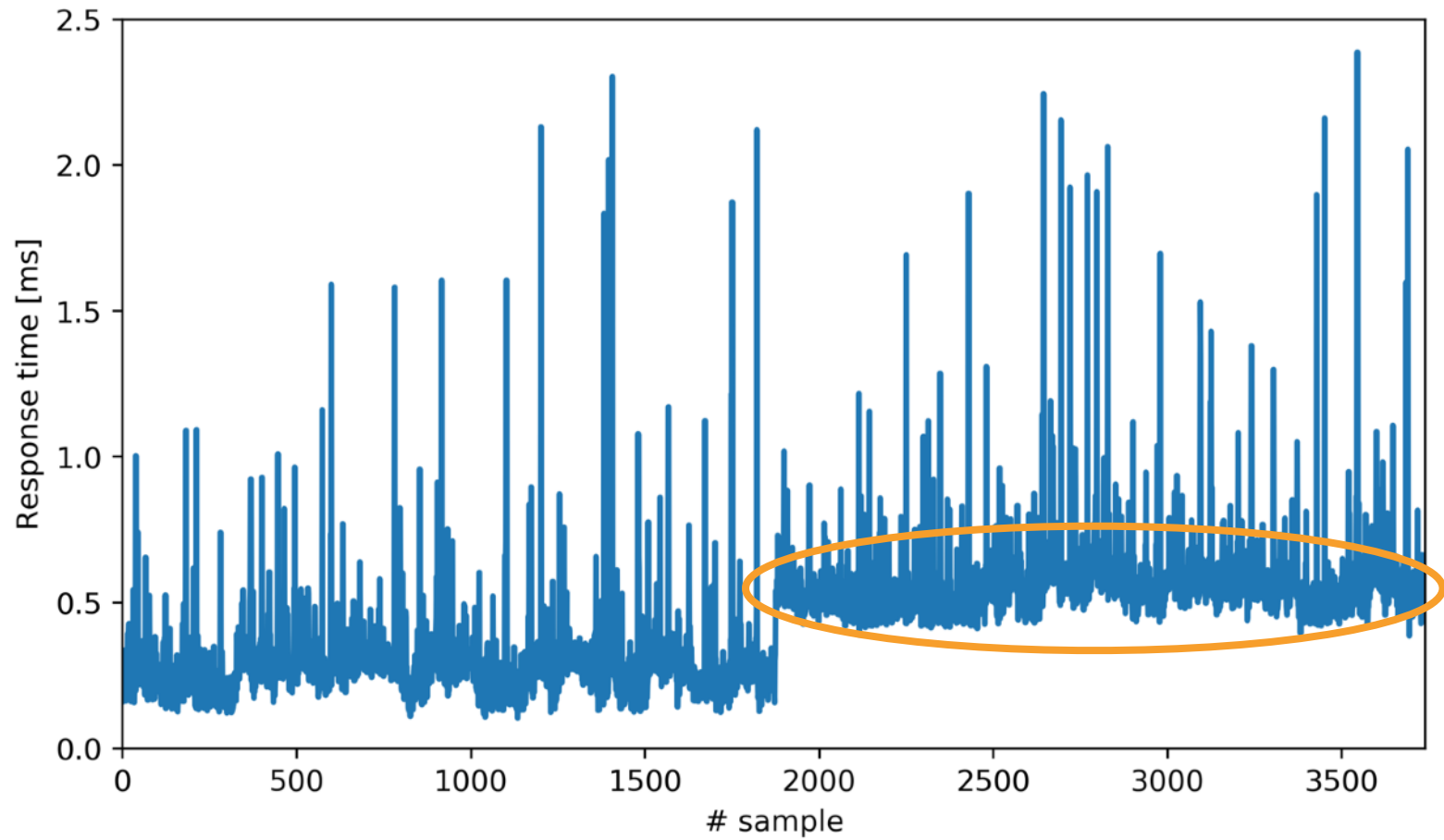


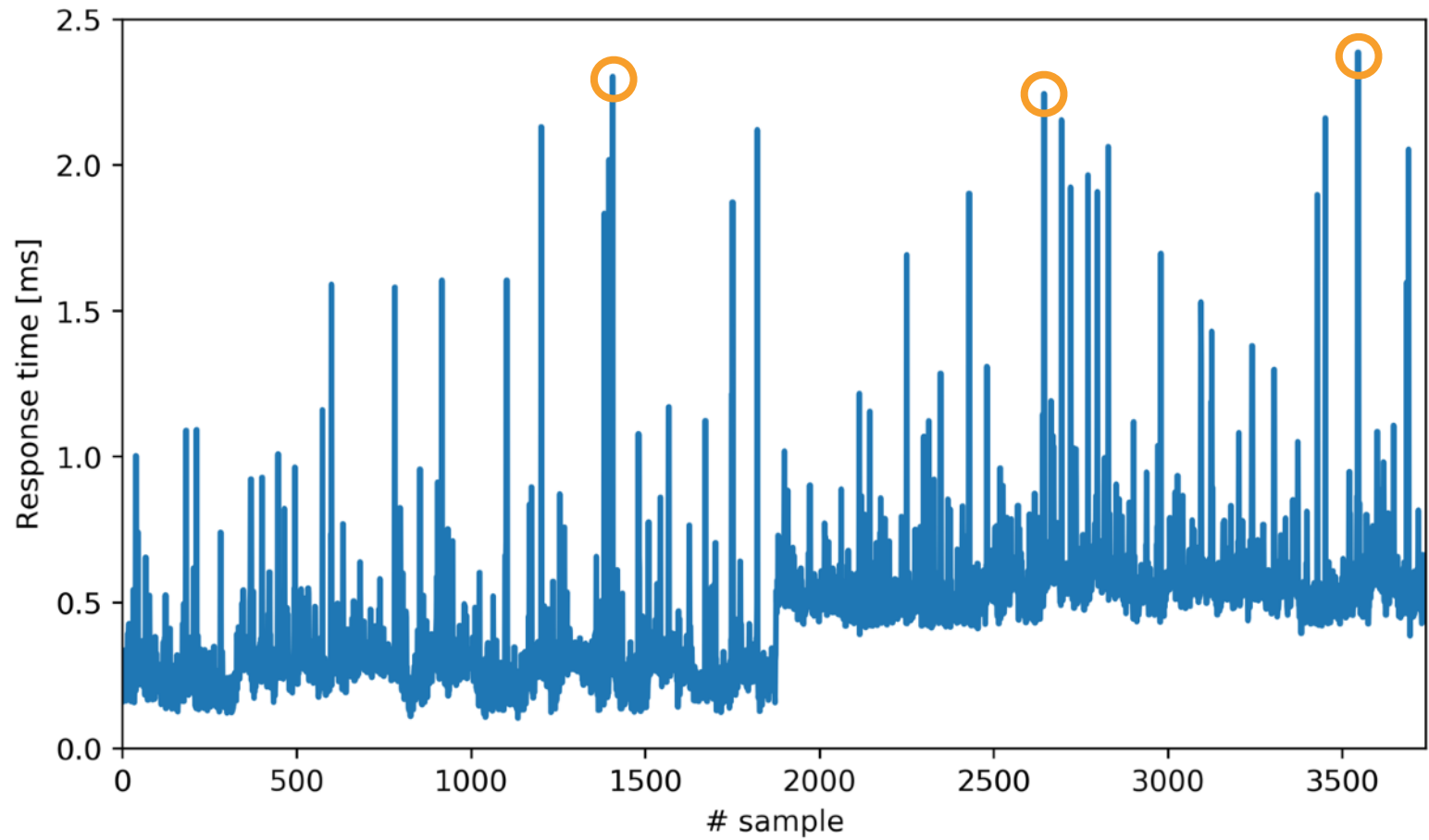










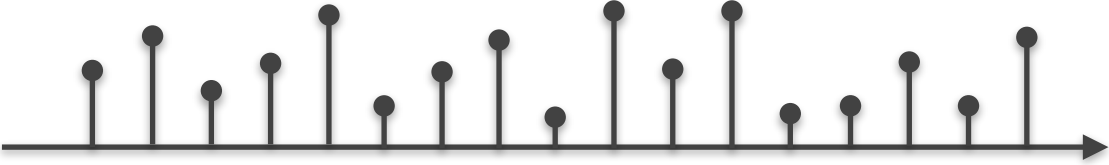


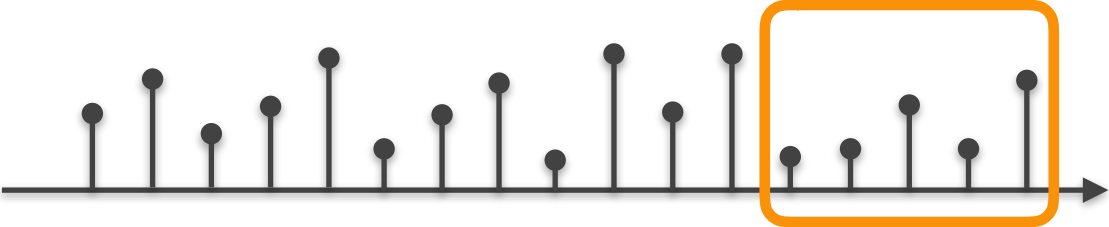


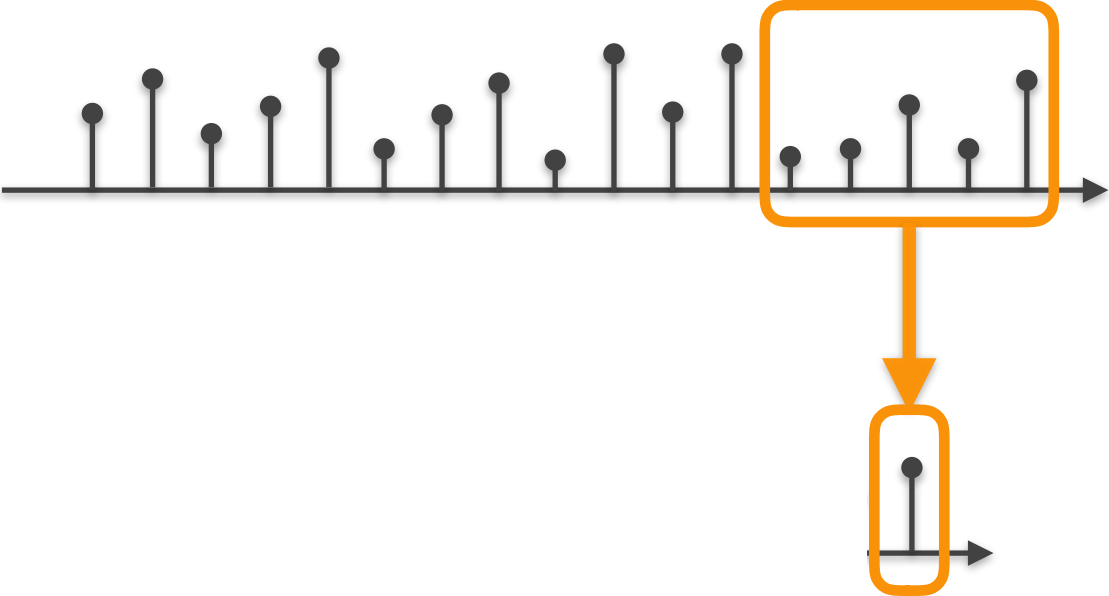
- Moving Average (MA)
- Weighted MA
- Low-pass filtering (FIR, IIR) / Convolution
- Rolling quantile
- Karhunen-Loève transform
- Subspace projection (wavelets, splines)

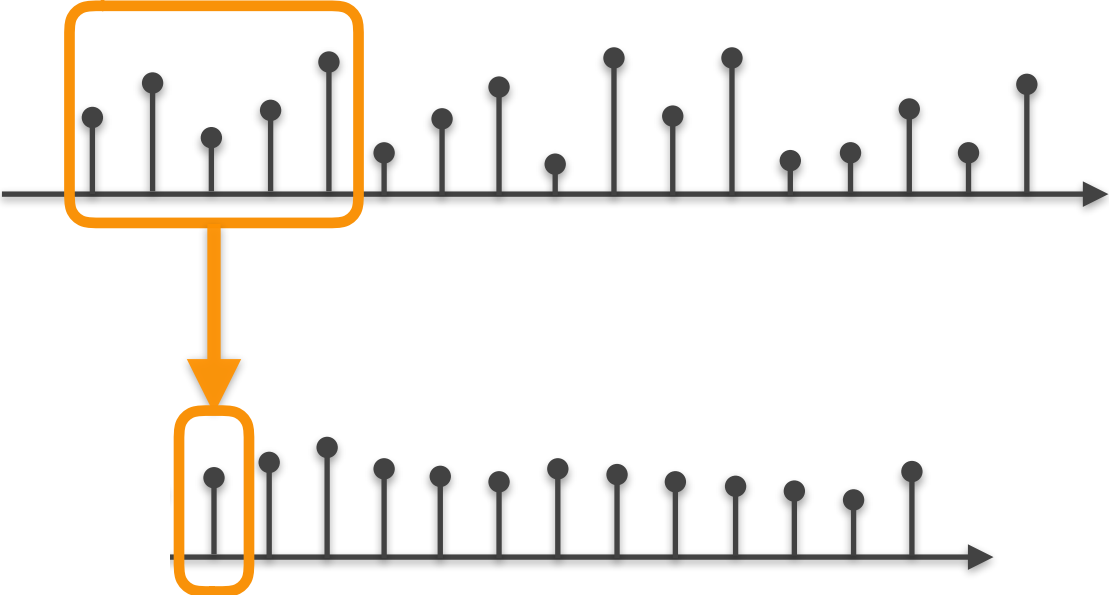


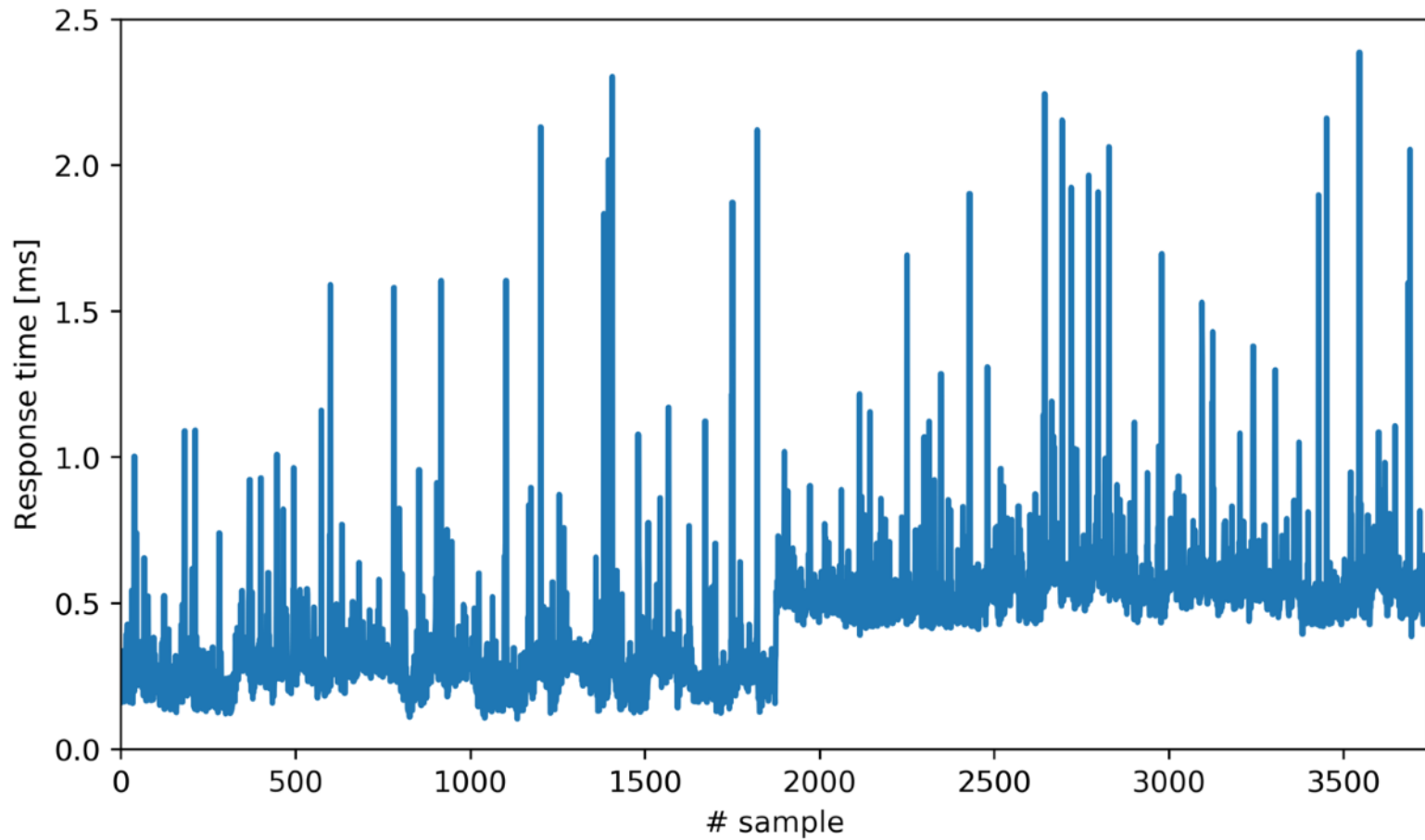
- **Moving Average (MA)**
- Weighted MA
- Low-pass filtering (FIR, IIR) / Convolution
- Rolling quantile
- Karhunen-Loève transform
- Subspace projection (wavelets, splines)

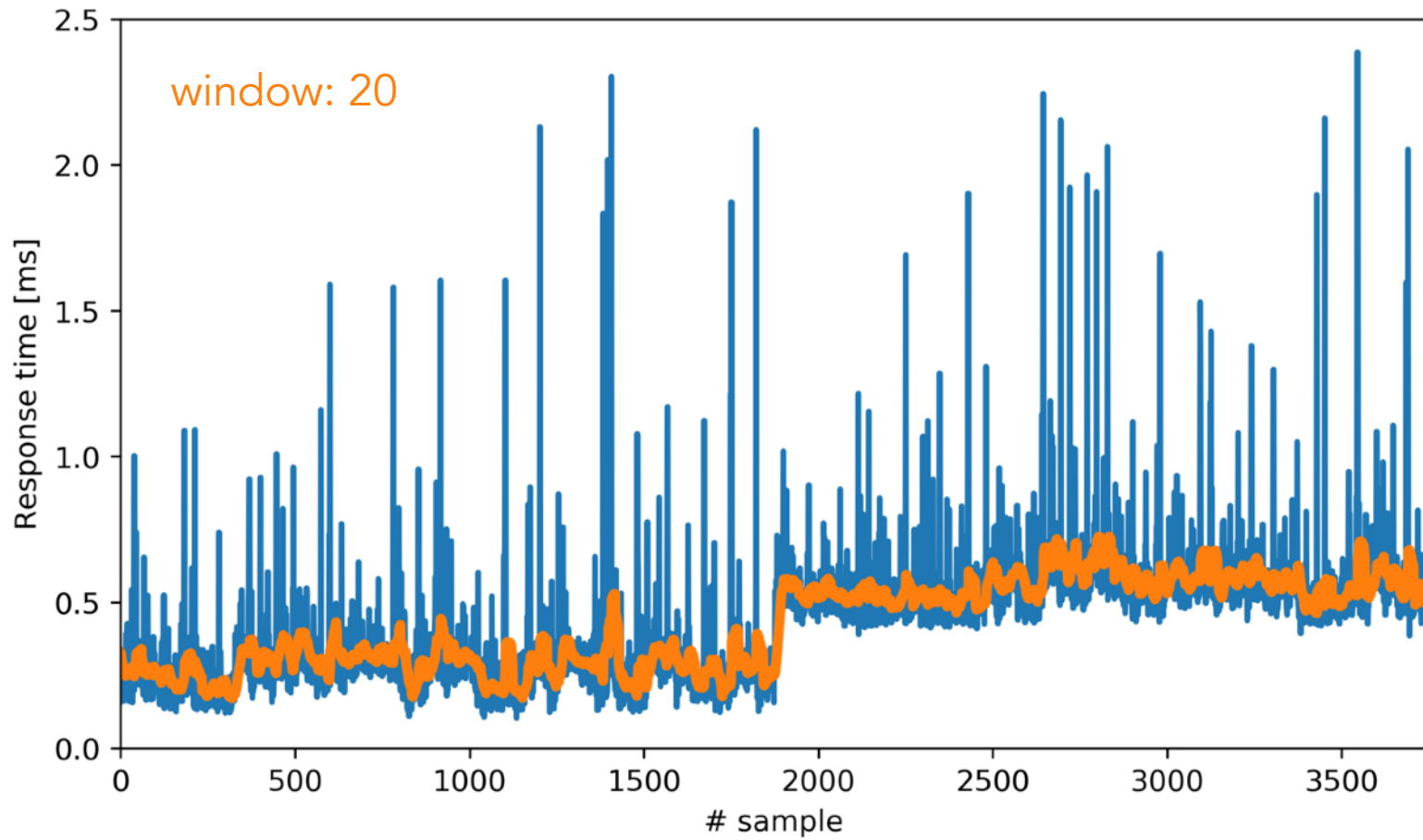


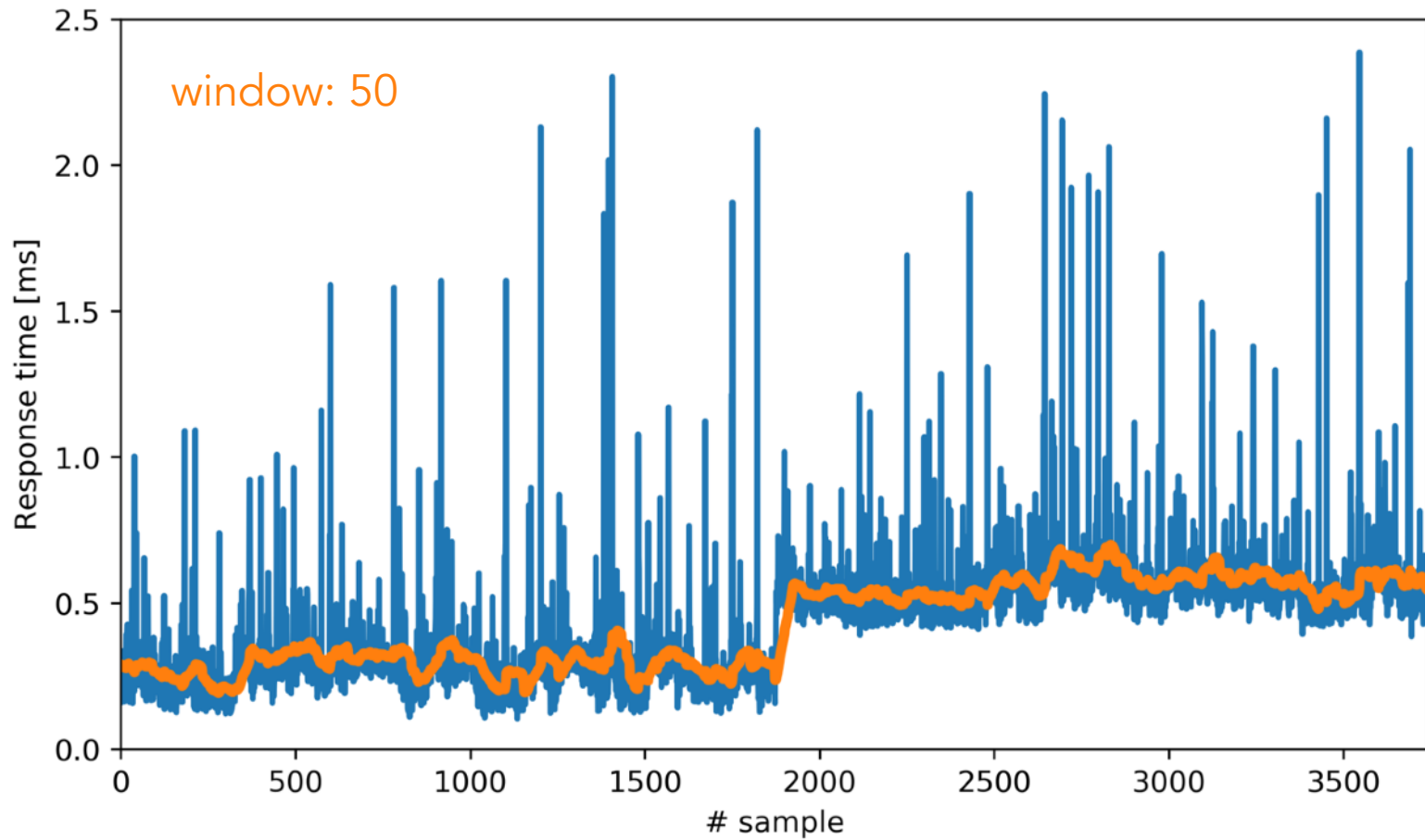


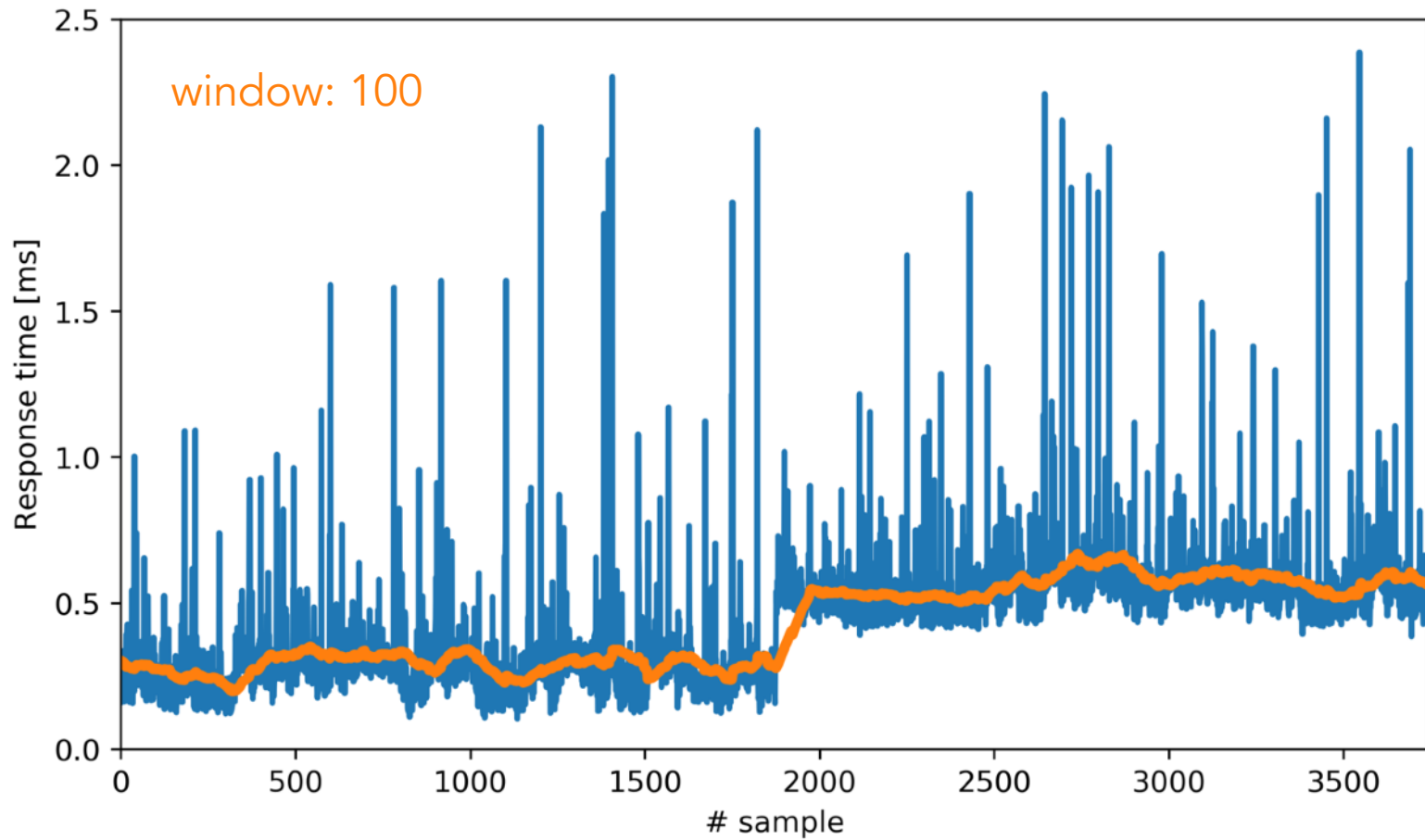


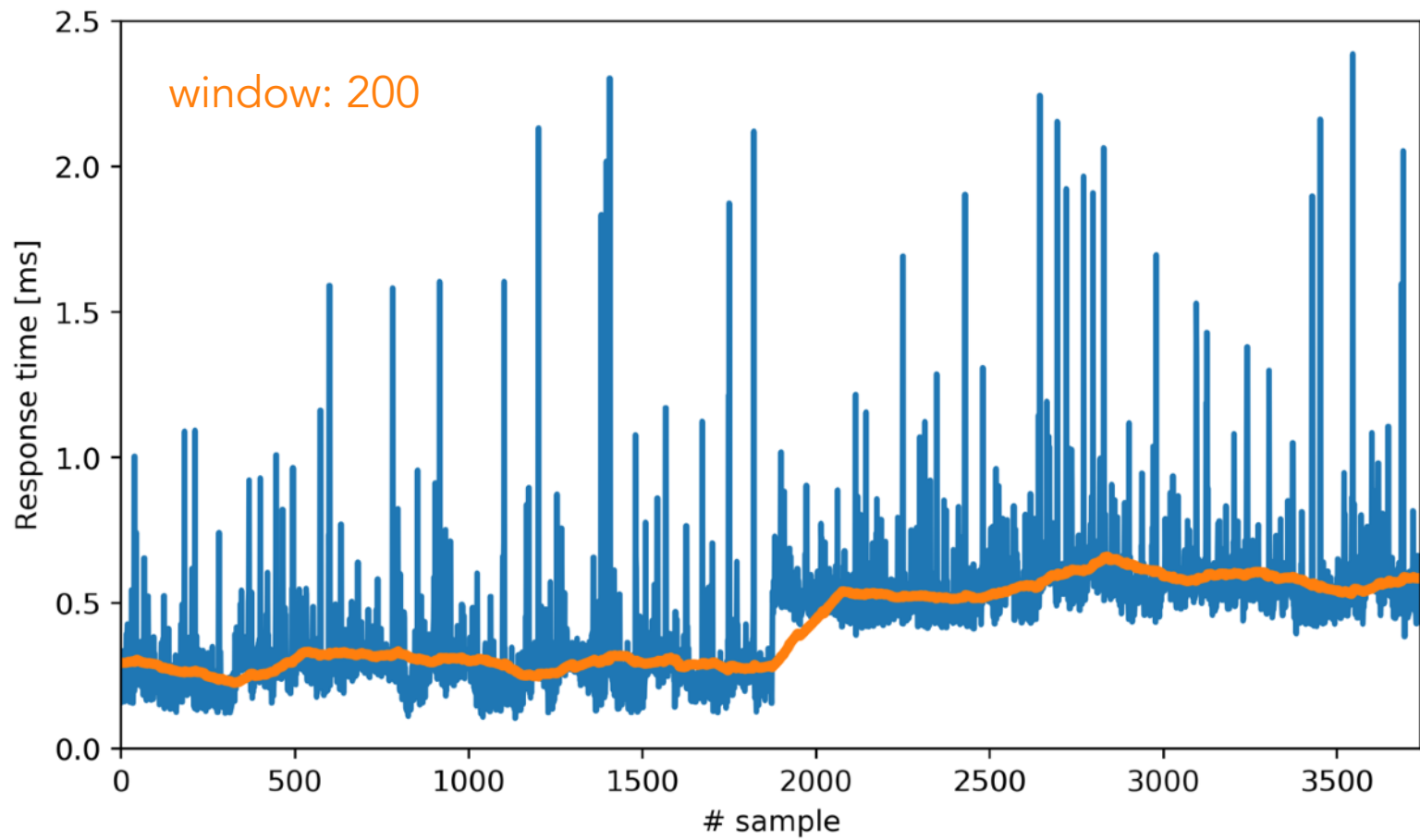


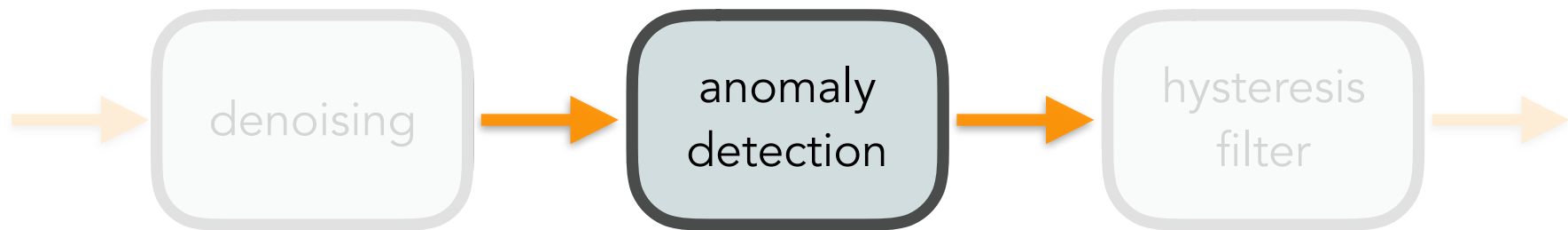


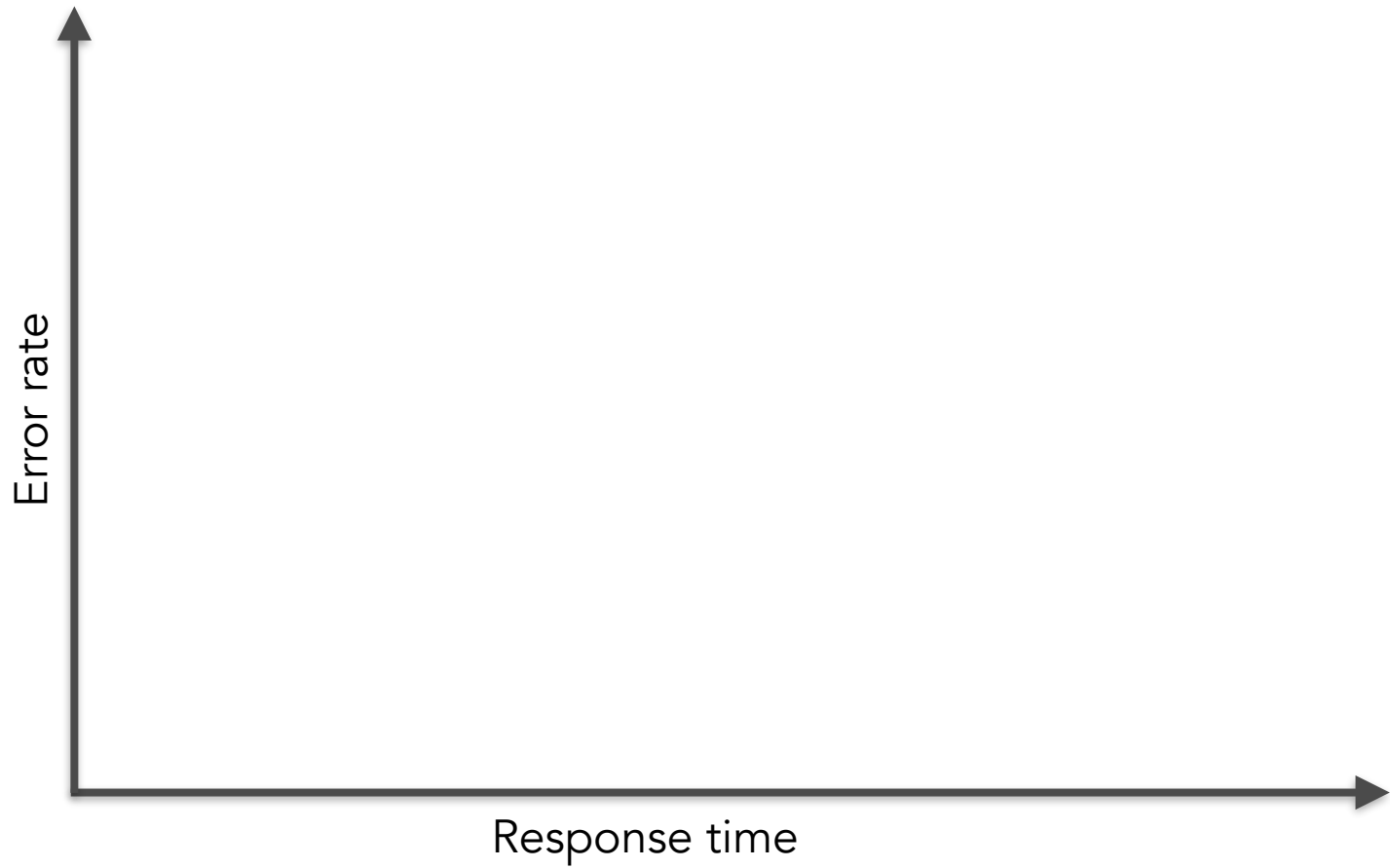


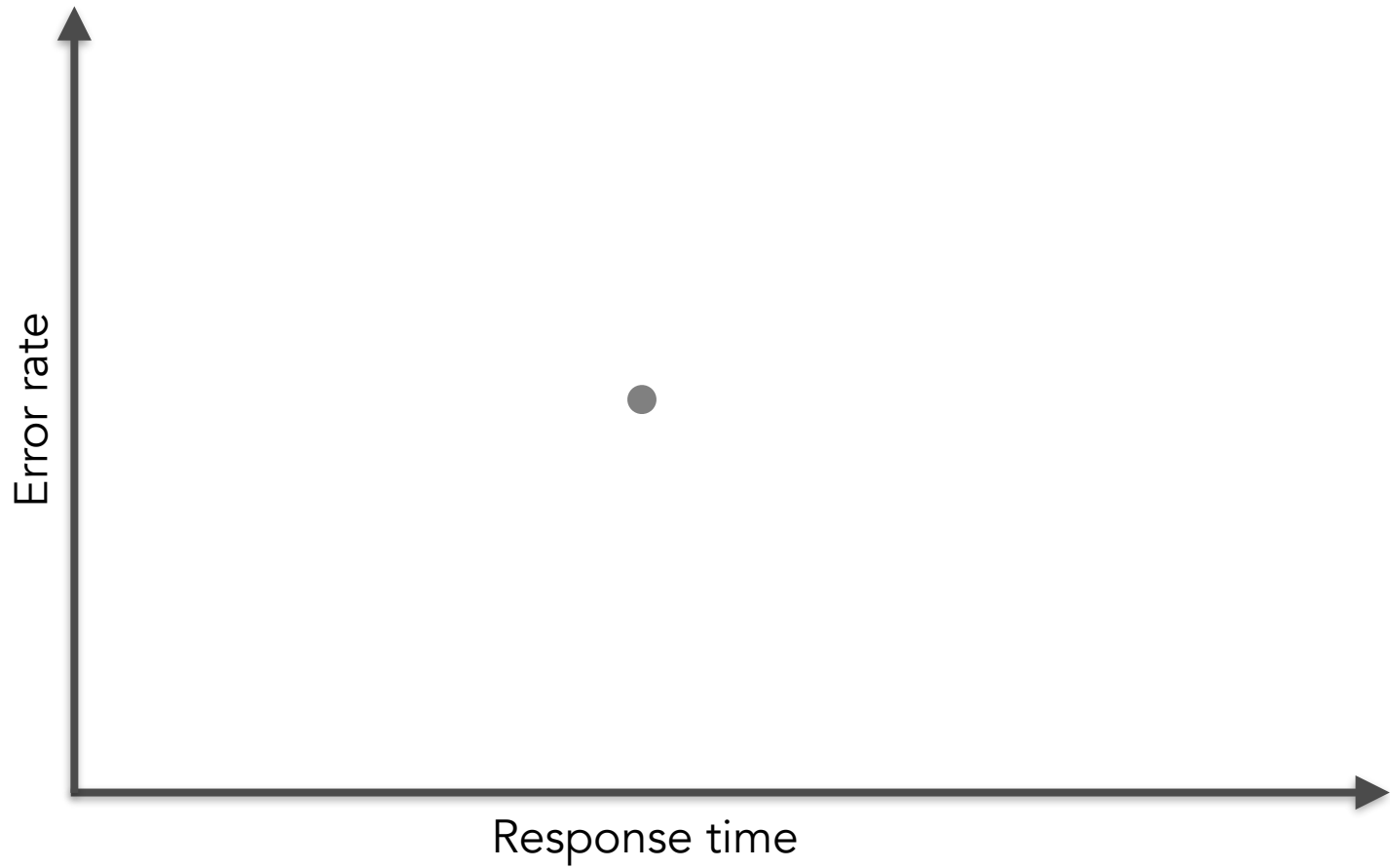


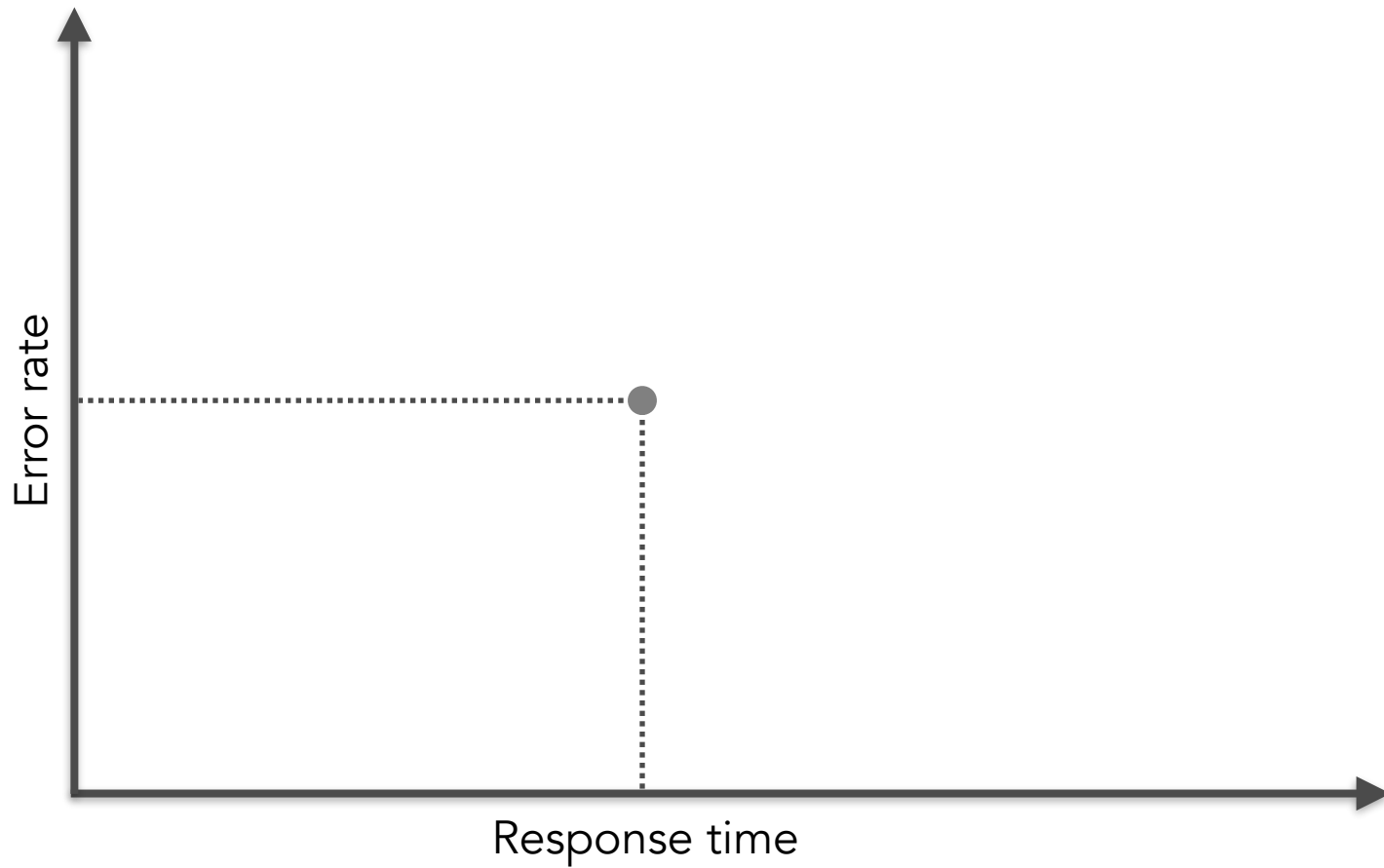


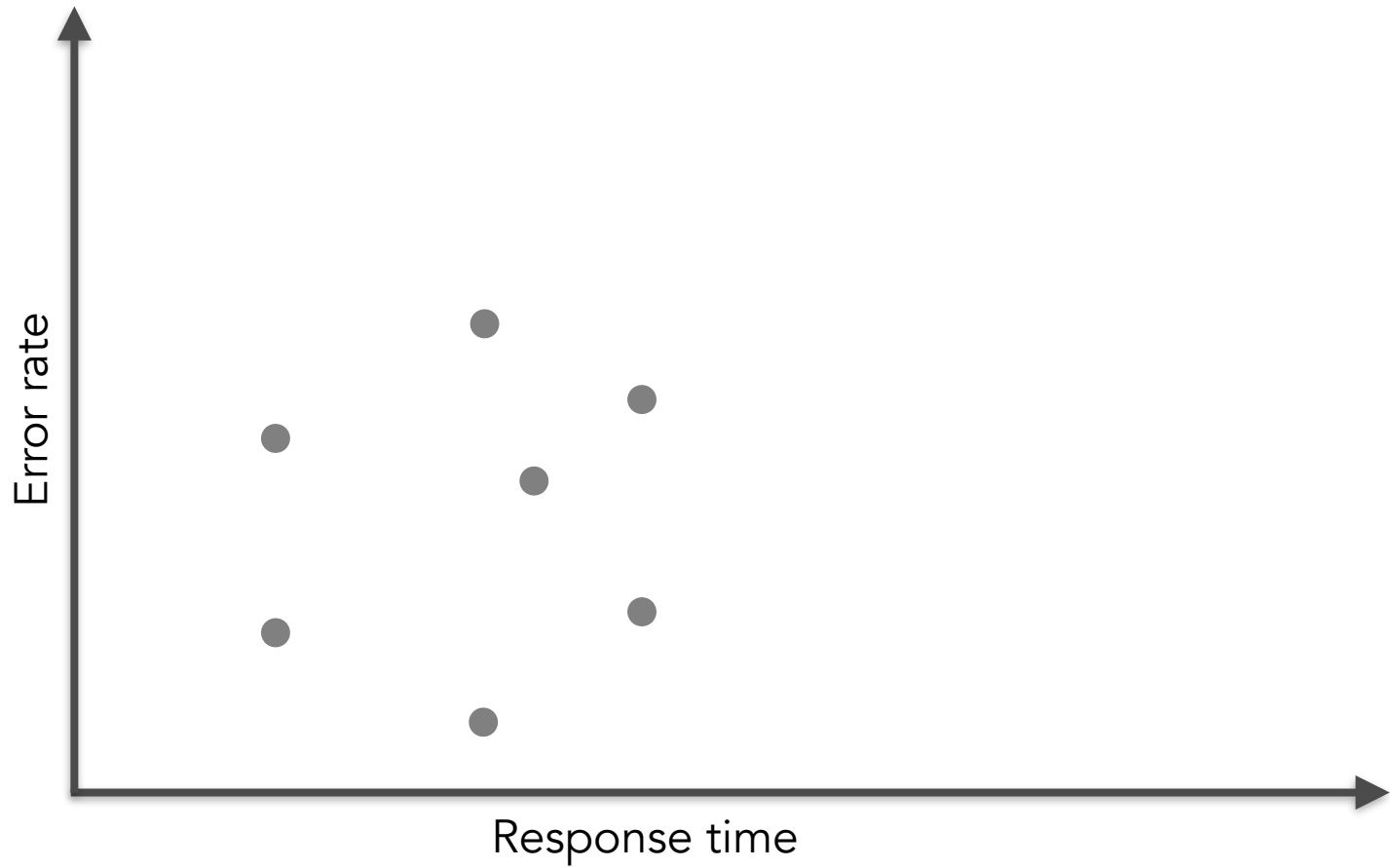


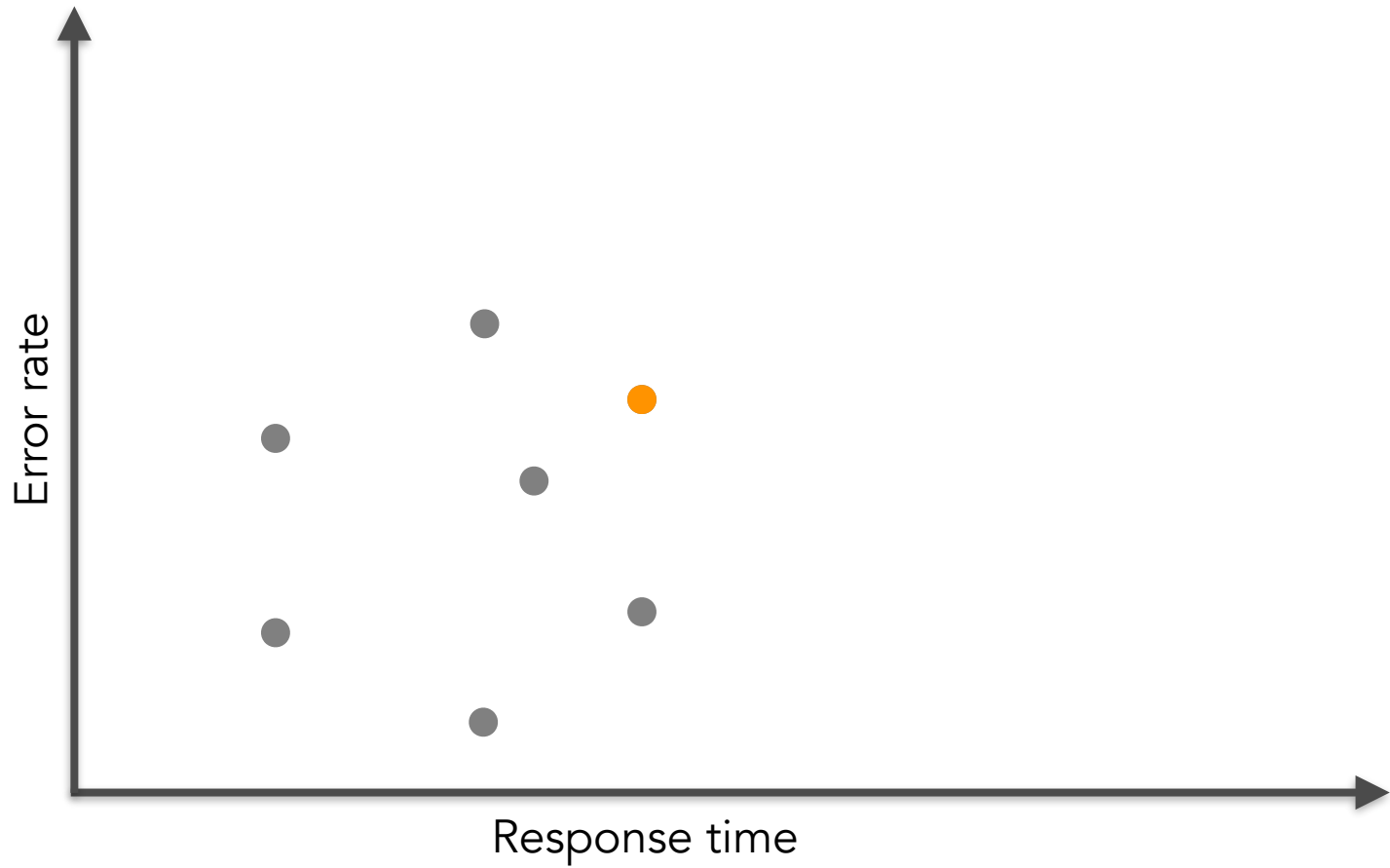


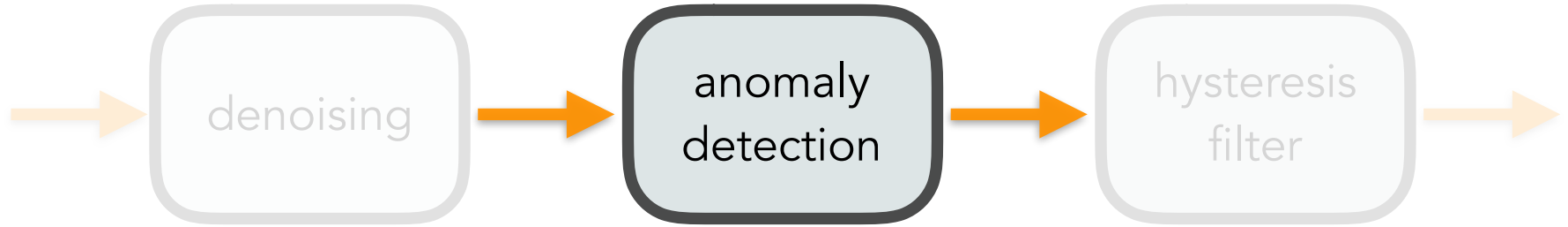




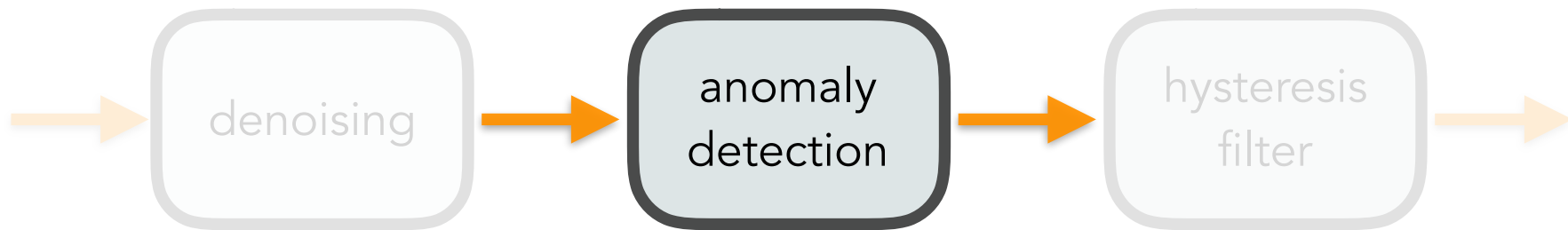




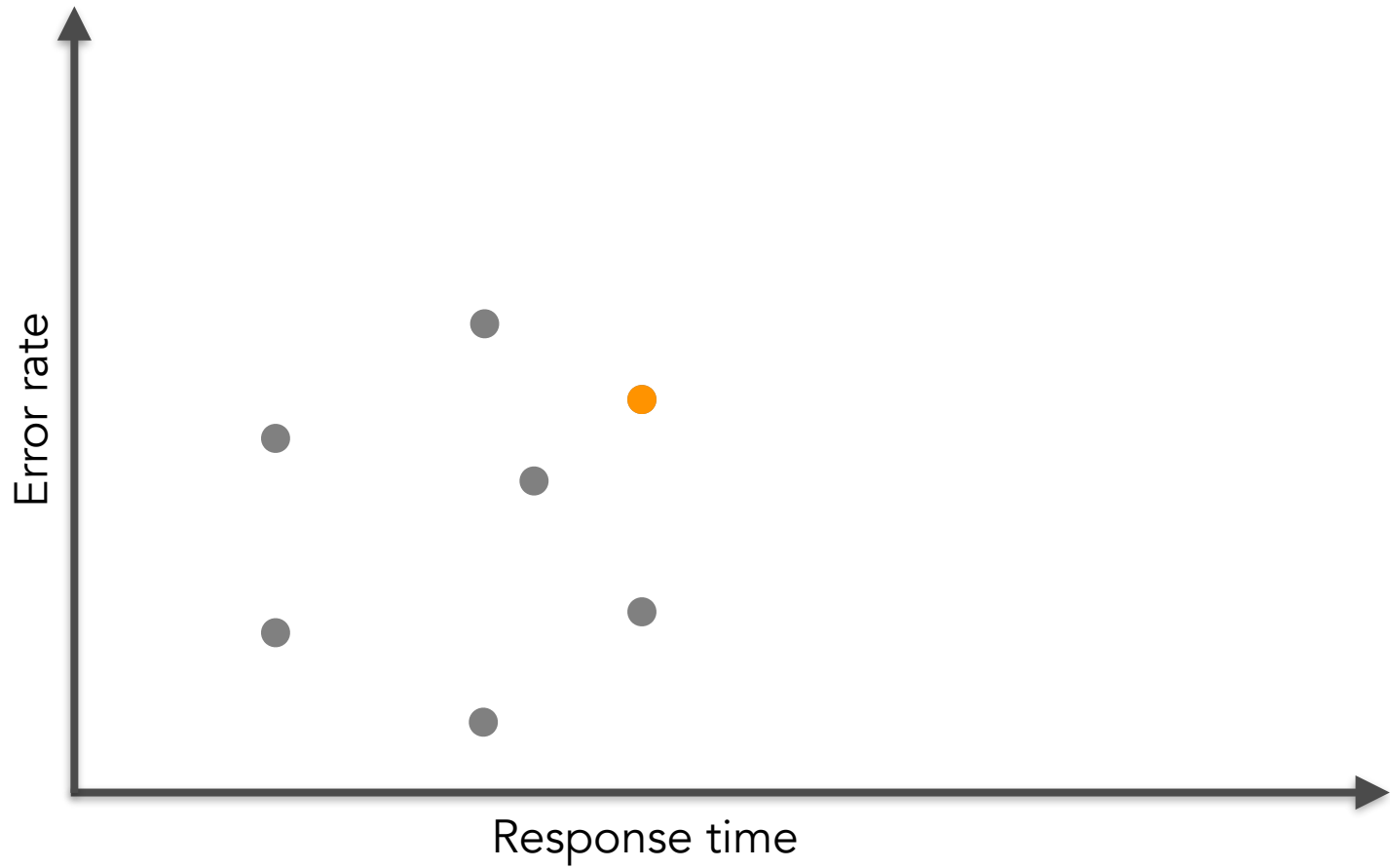


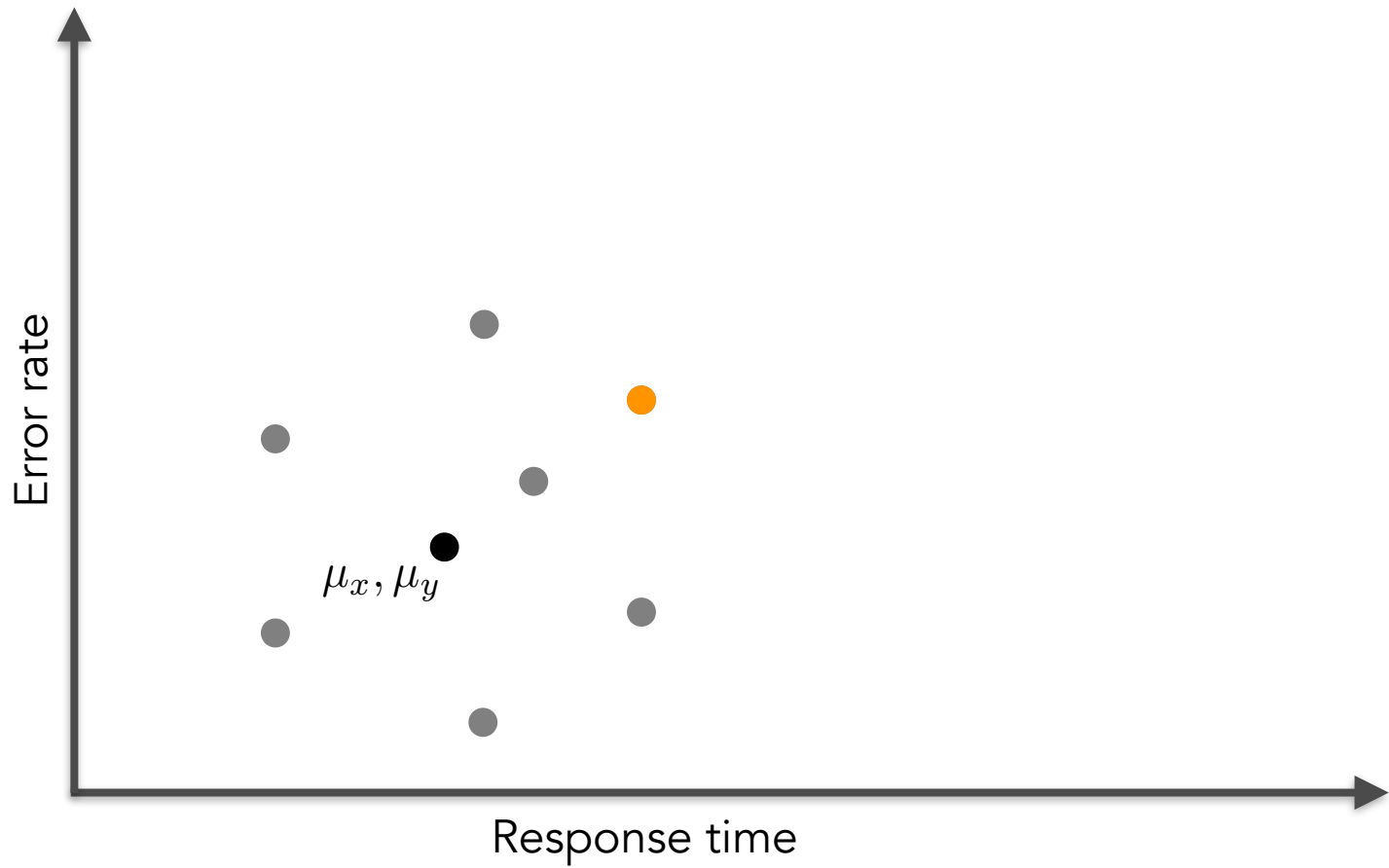


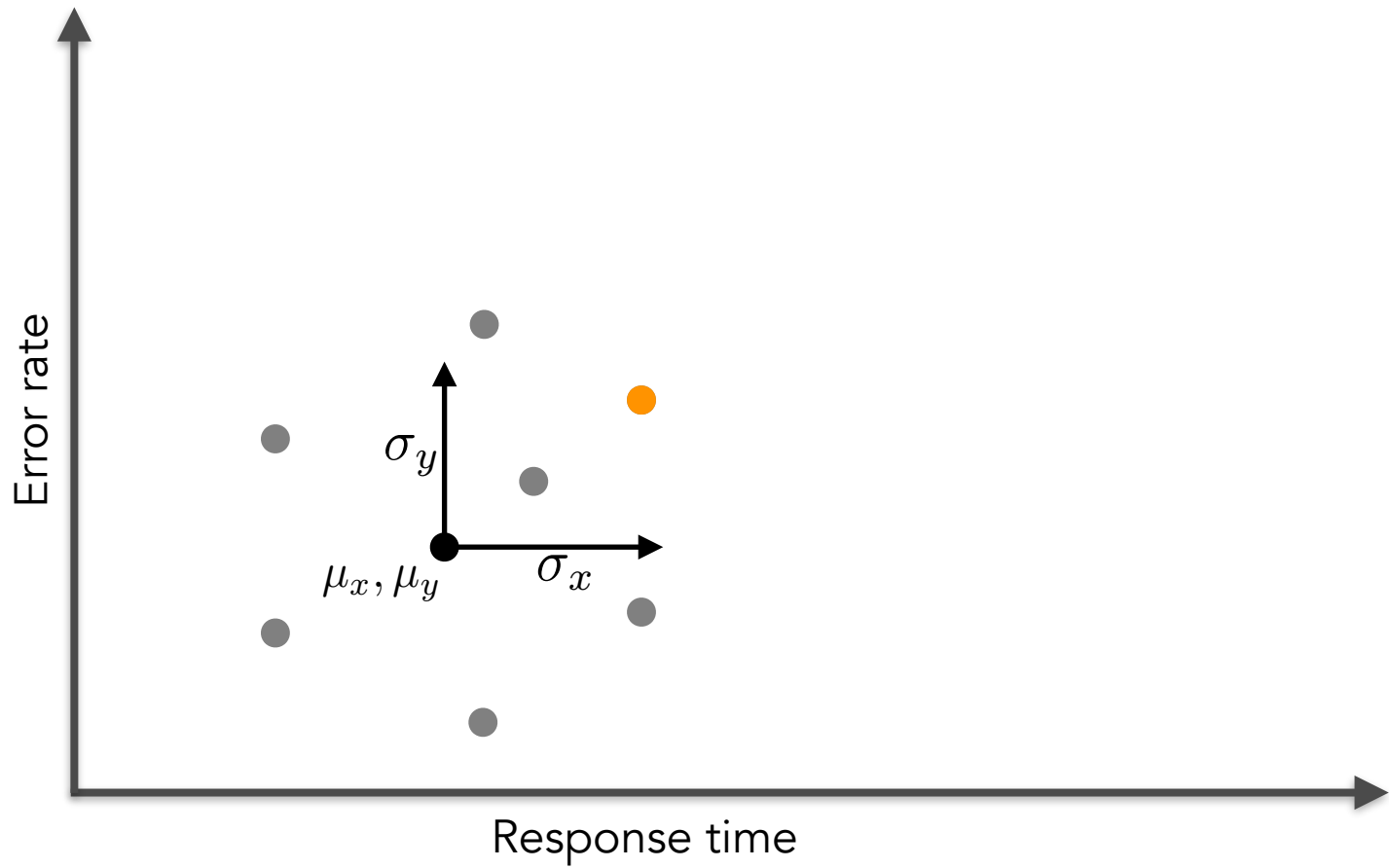
- Simple thresholding
- Hypothesis testing
- Conditional entropy
- Distributional thresholding
 - Mahalanobis distance
 - Kullback-Leibler divergence
- Pattern matching / Clustering

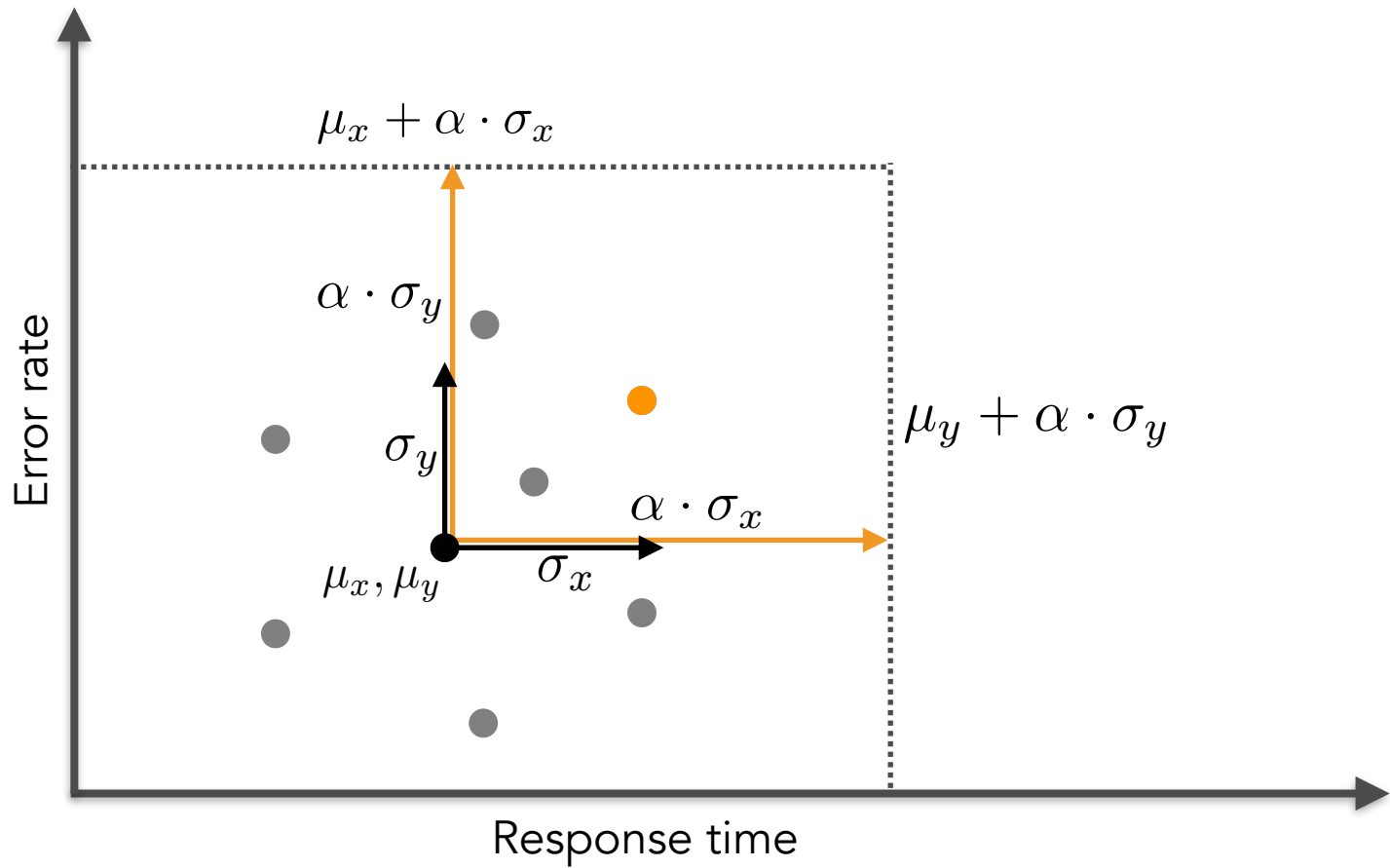


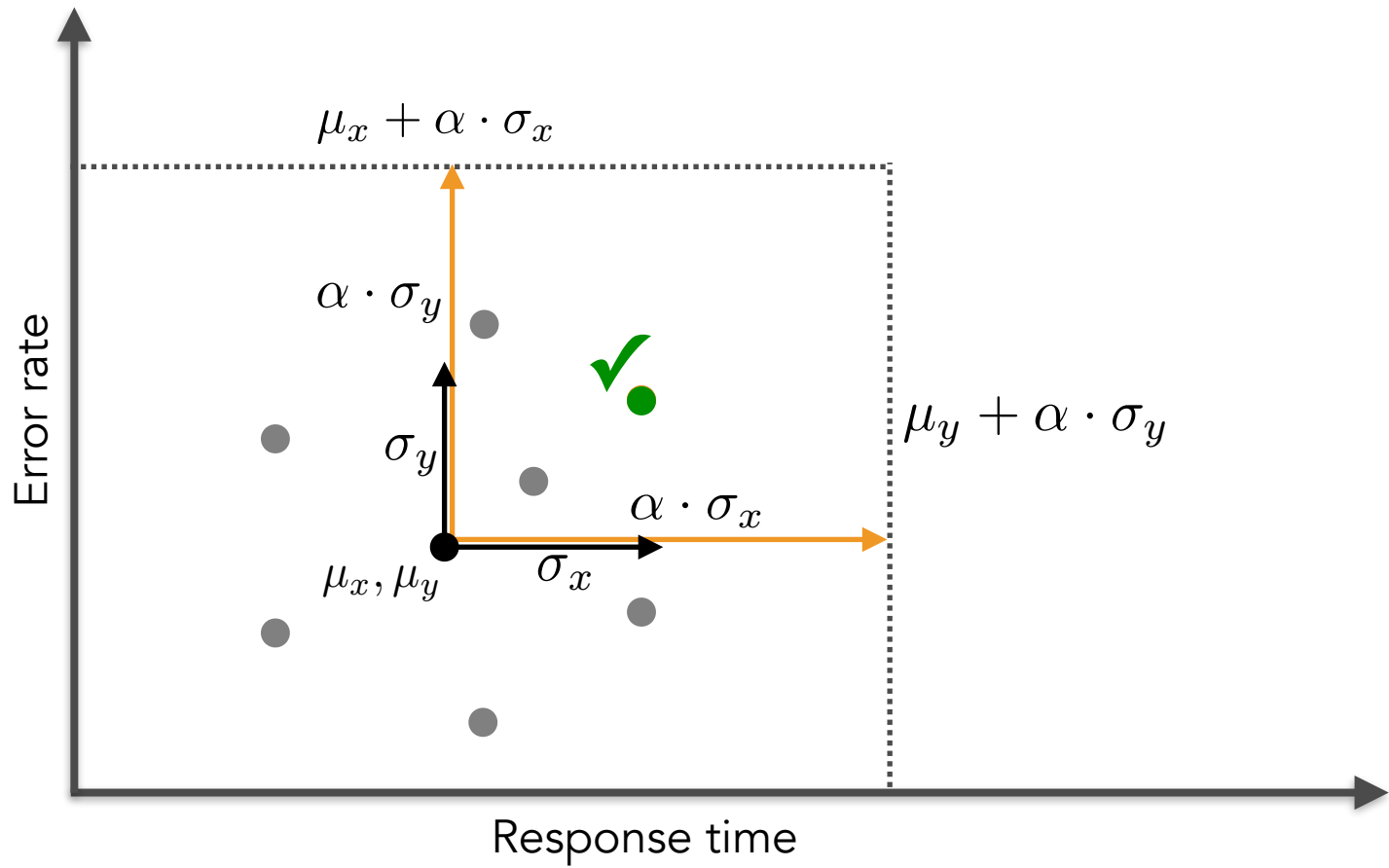
- **Simple thresholding**
- Hypothesis testing
- Conditional entropy
- Distributional thresholding
 - Mahalanobis distance
 - Kullback-Leibler divergence
- Pattern matching / Clustering

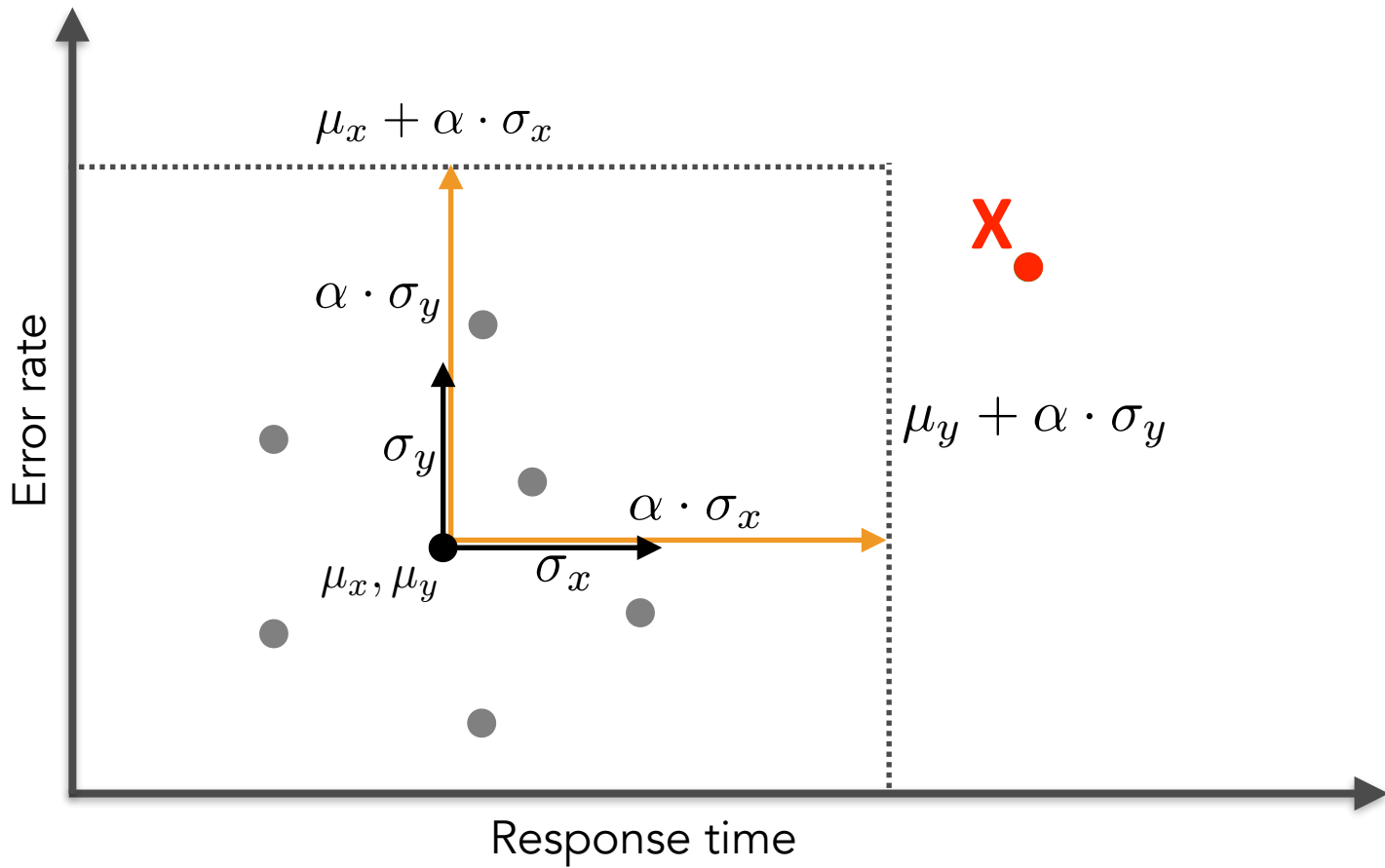


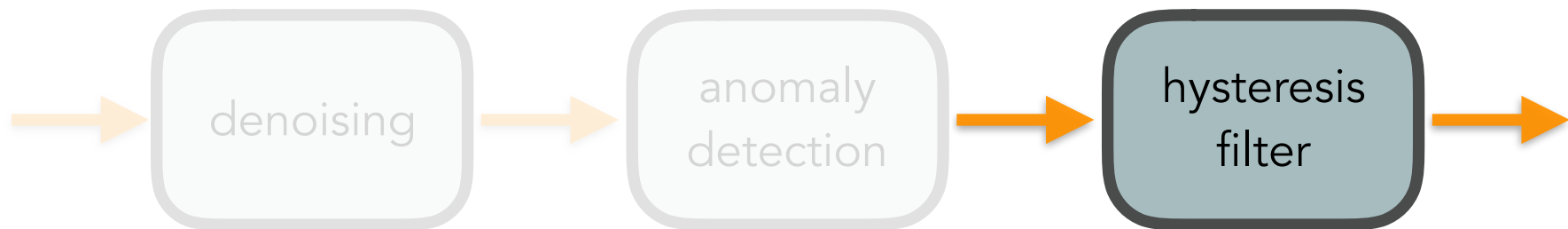


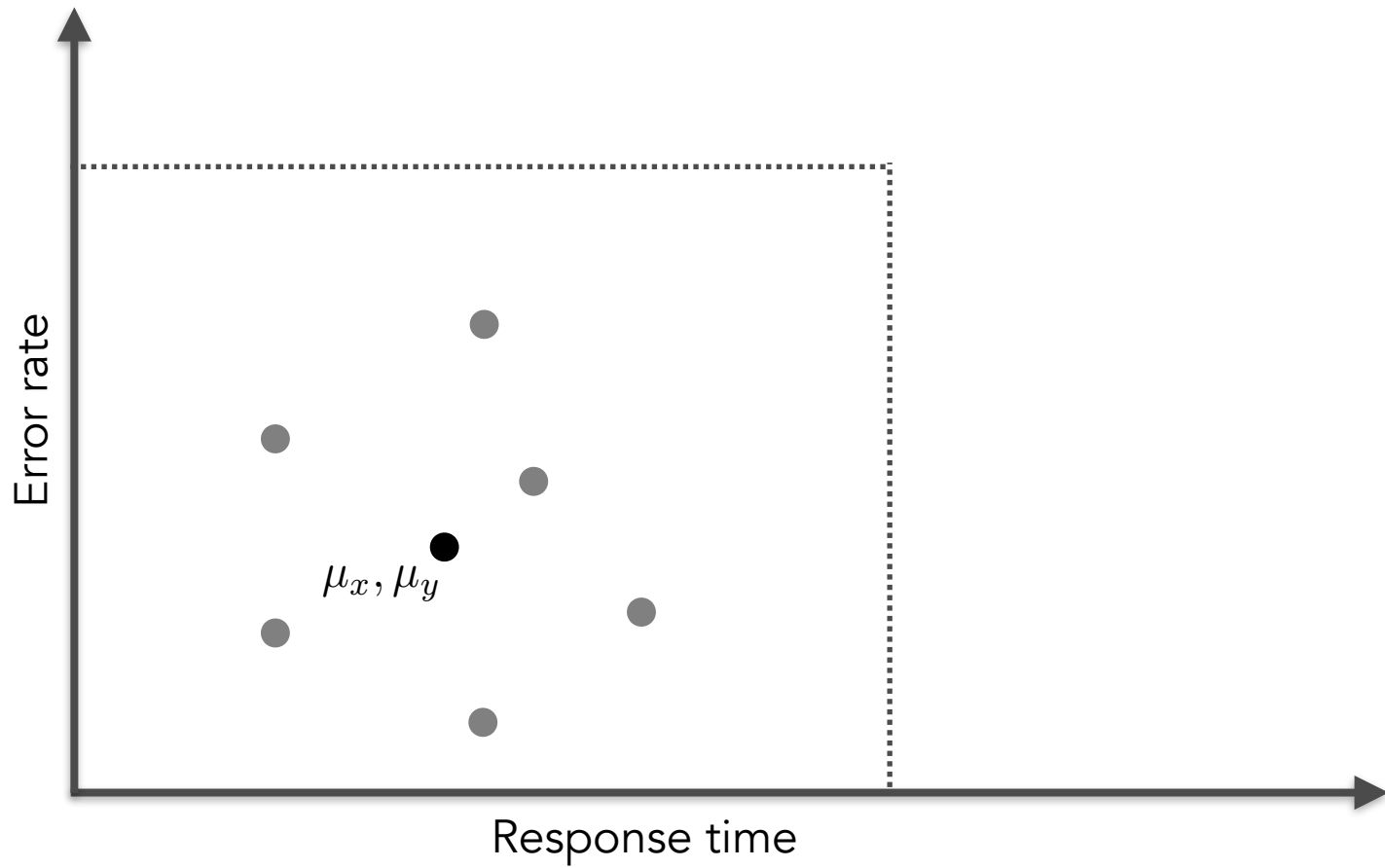


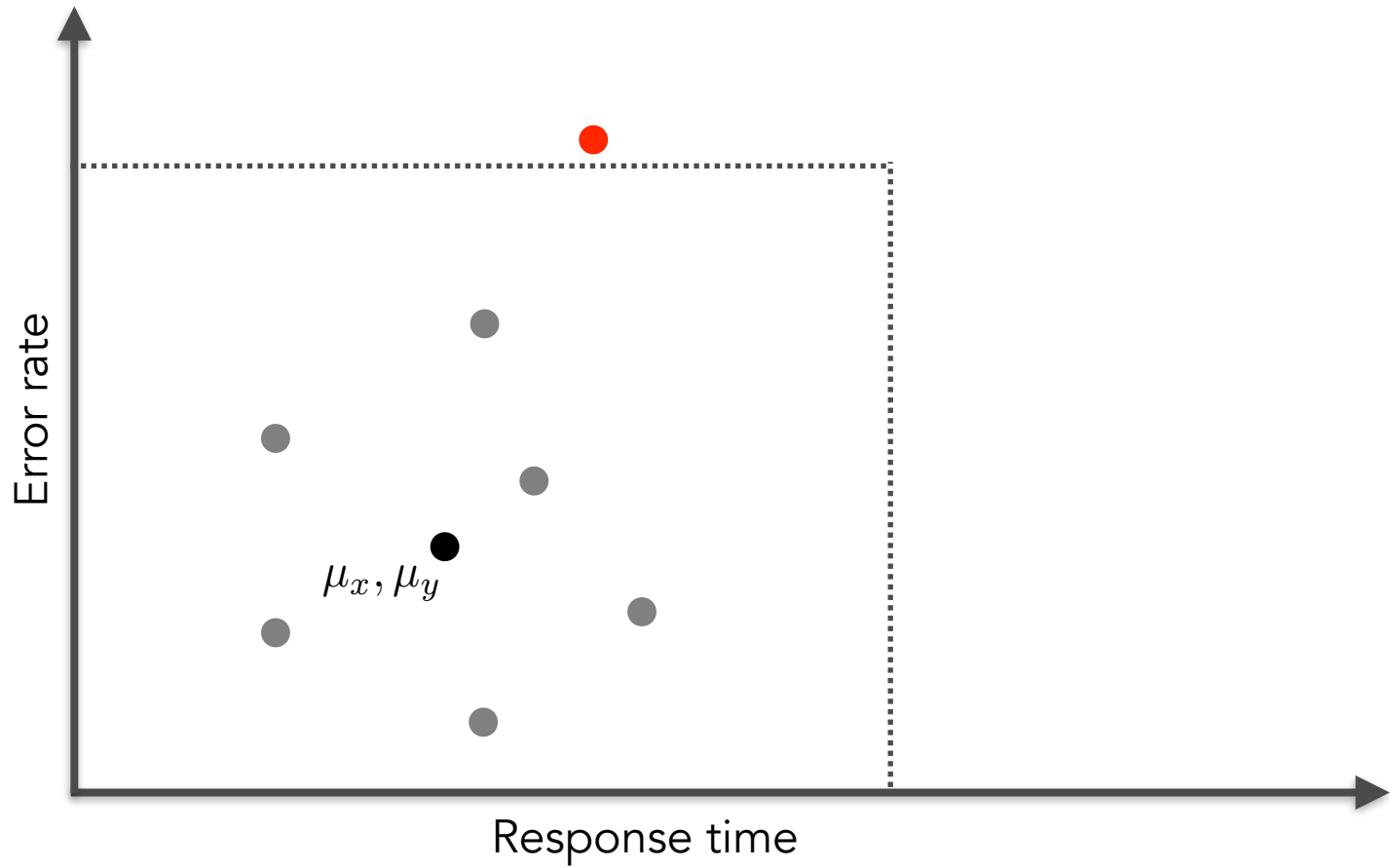


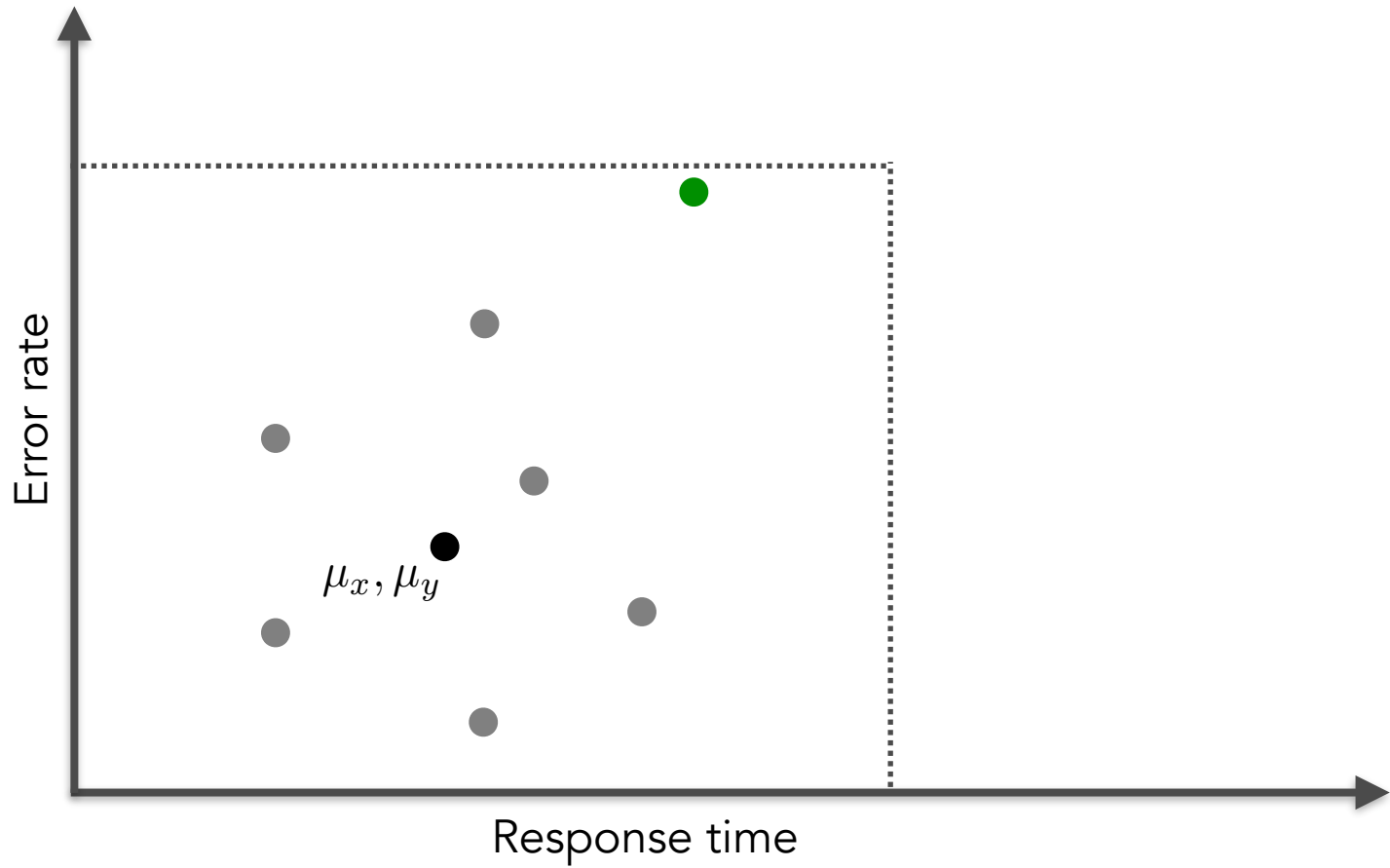


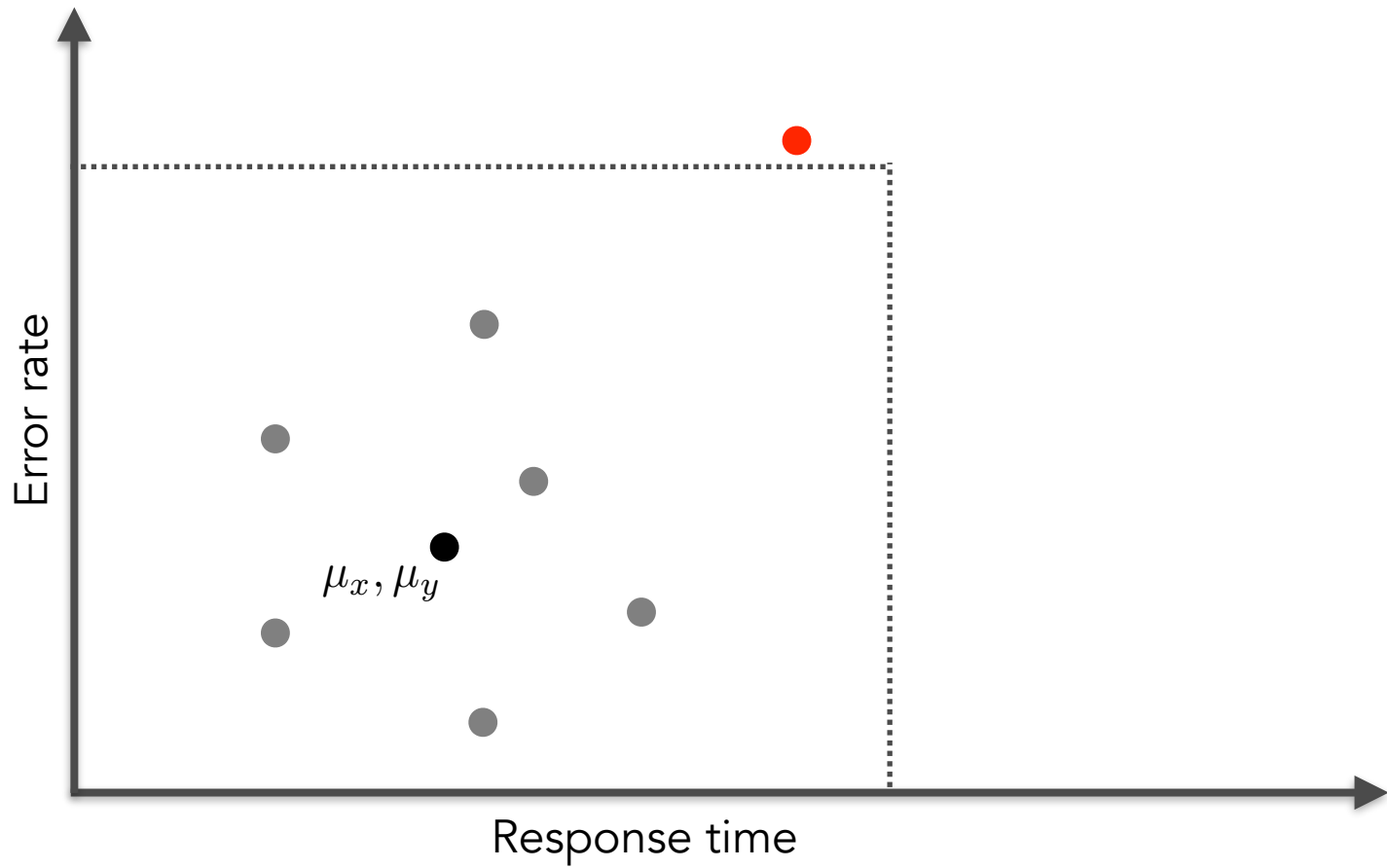






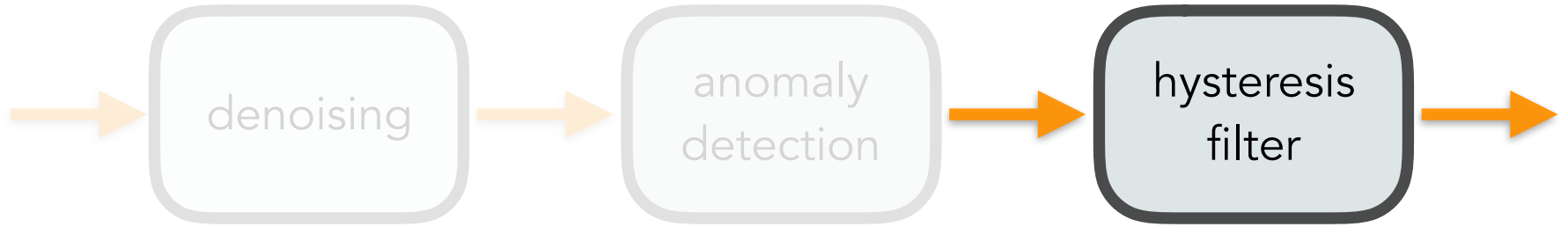




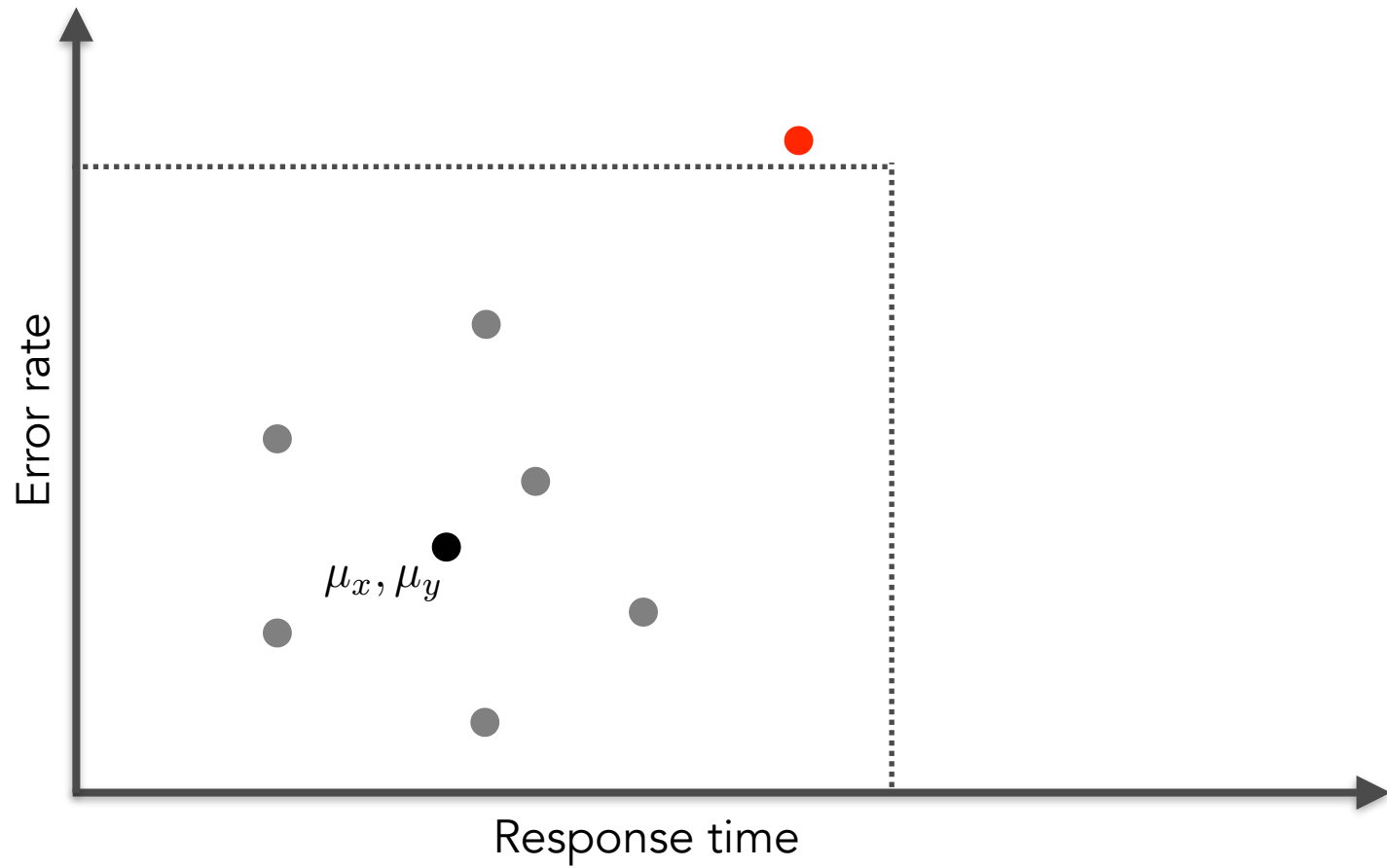


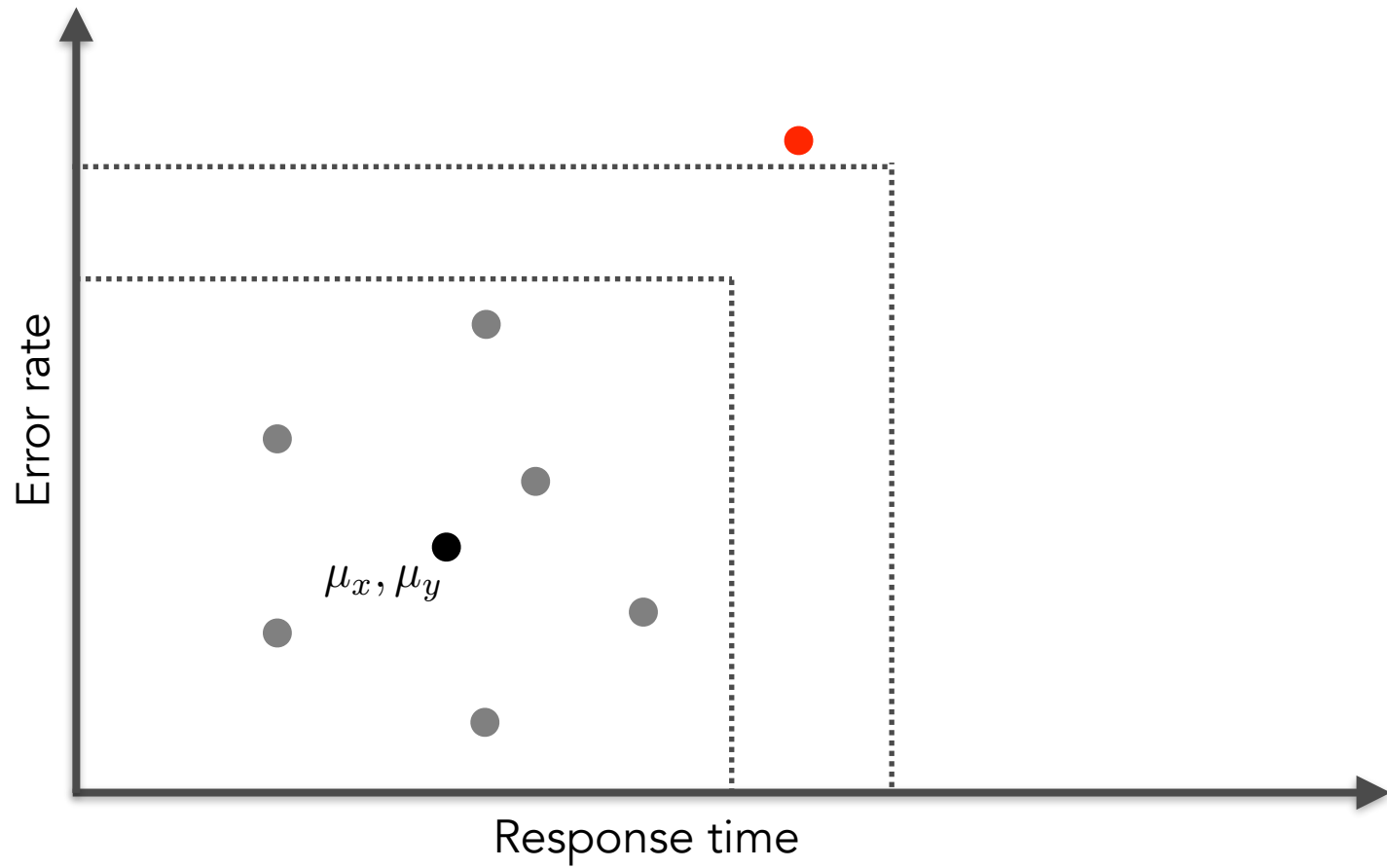


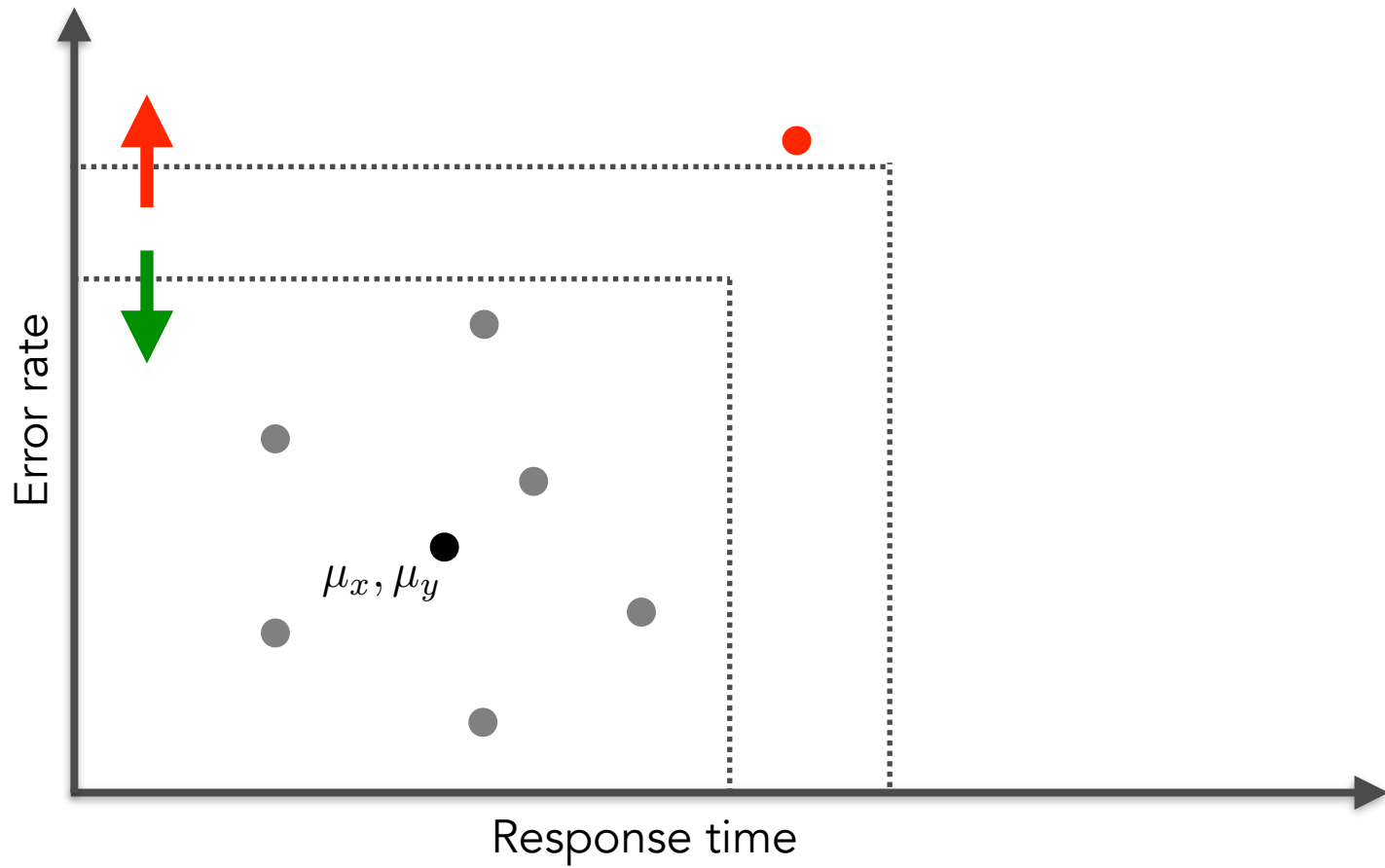
- Sharp hysteresis
- Continuous hysteresis
- Finite State Machine
- Fuzzy logic program

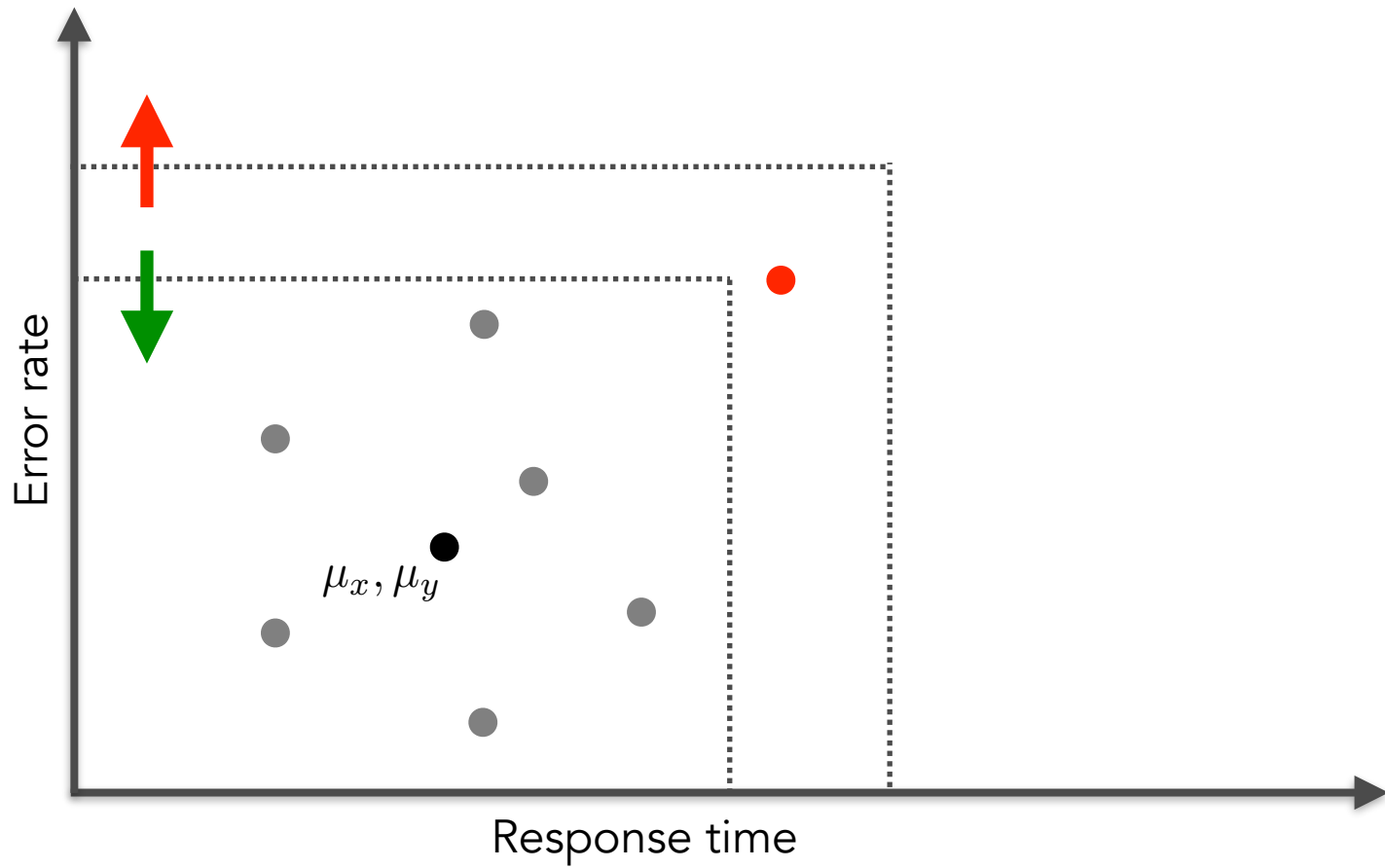


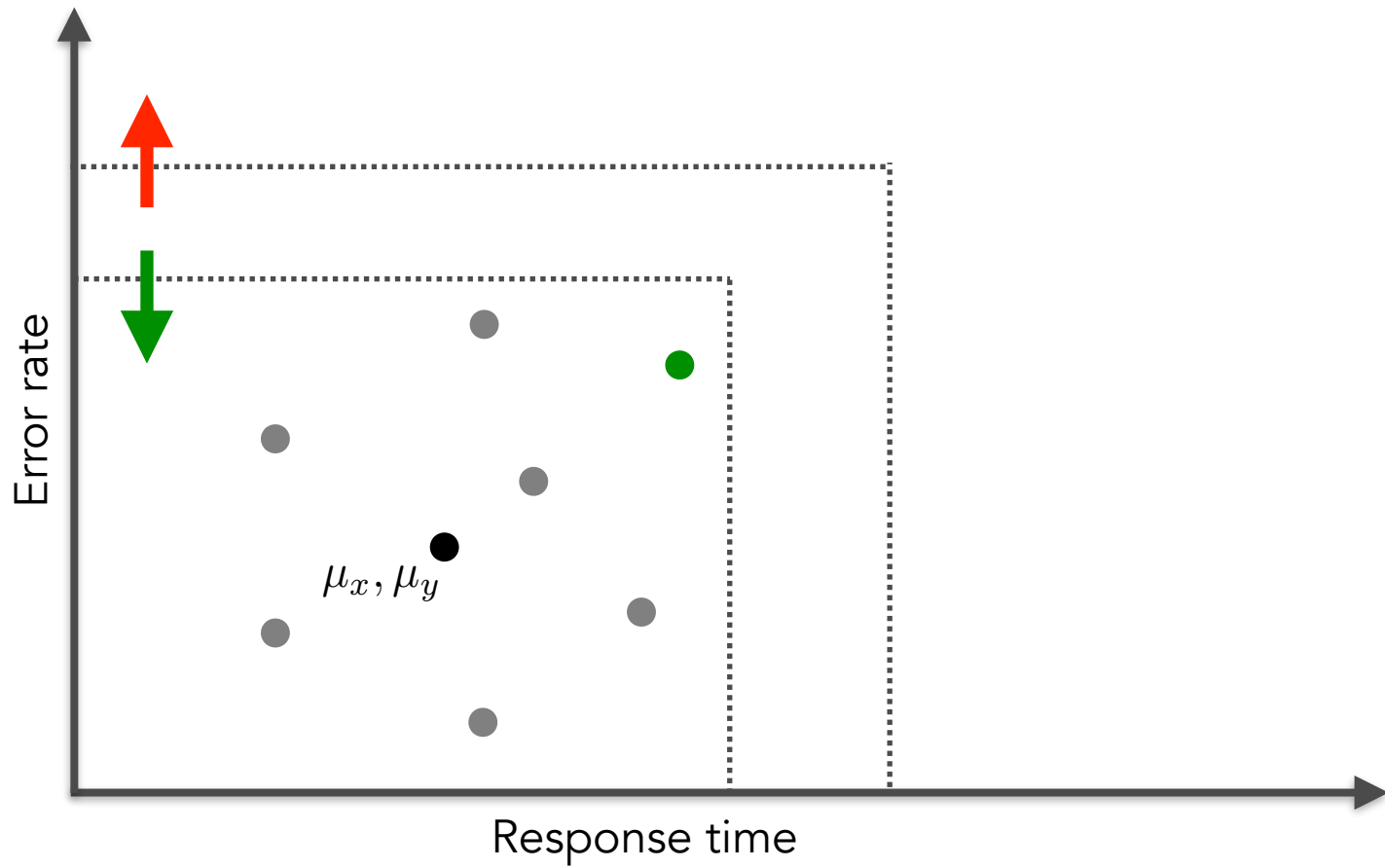
- **Sharp hysteresis**
- Continuous hysteresis
- Finite State Machine
- Fuzzy logic program

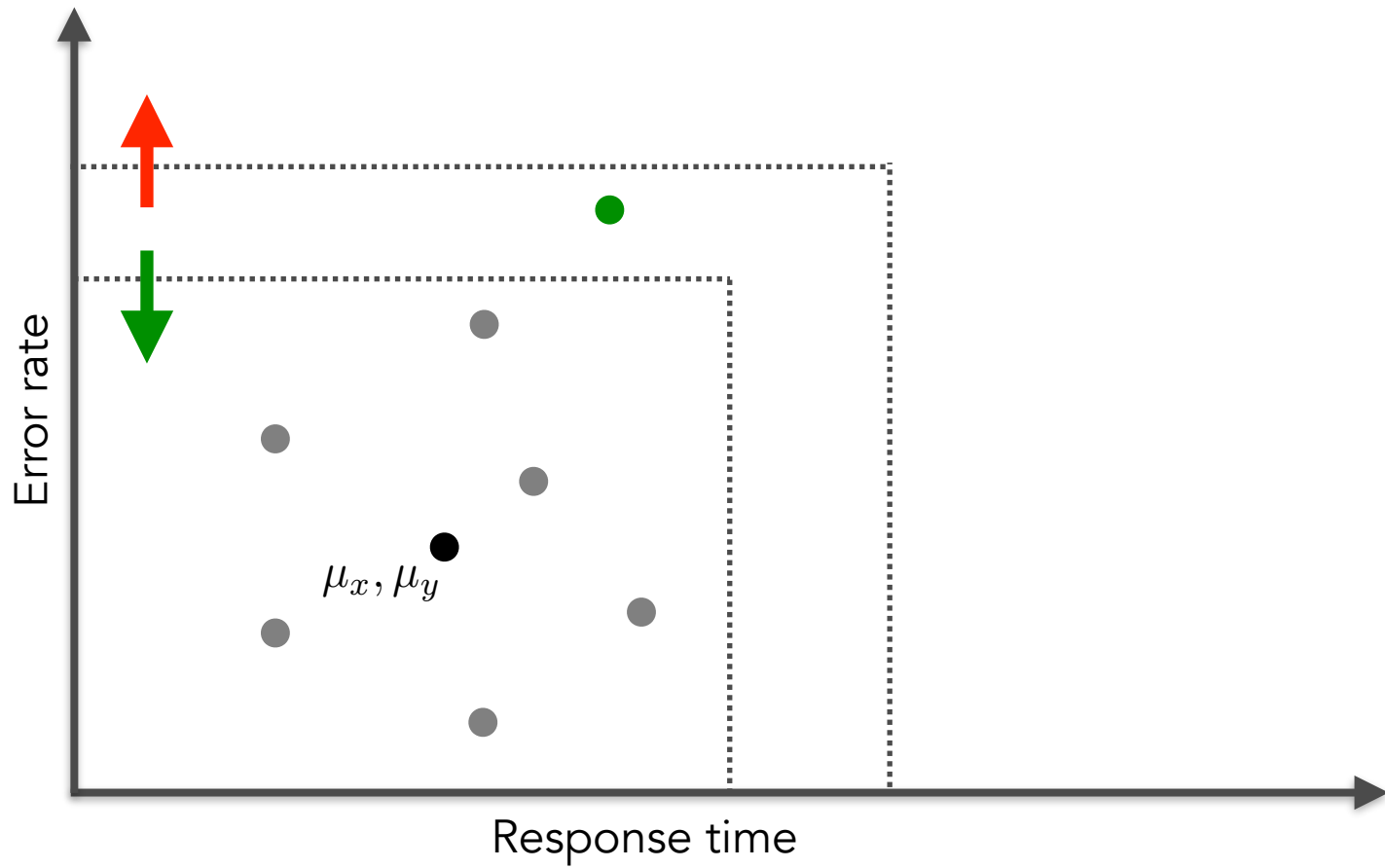


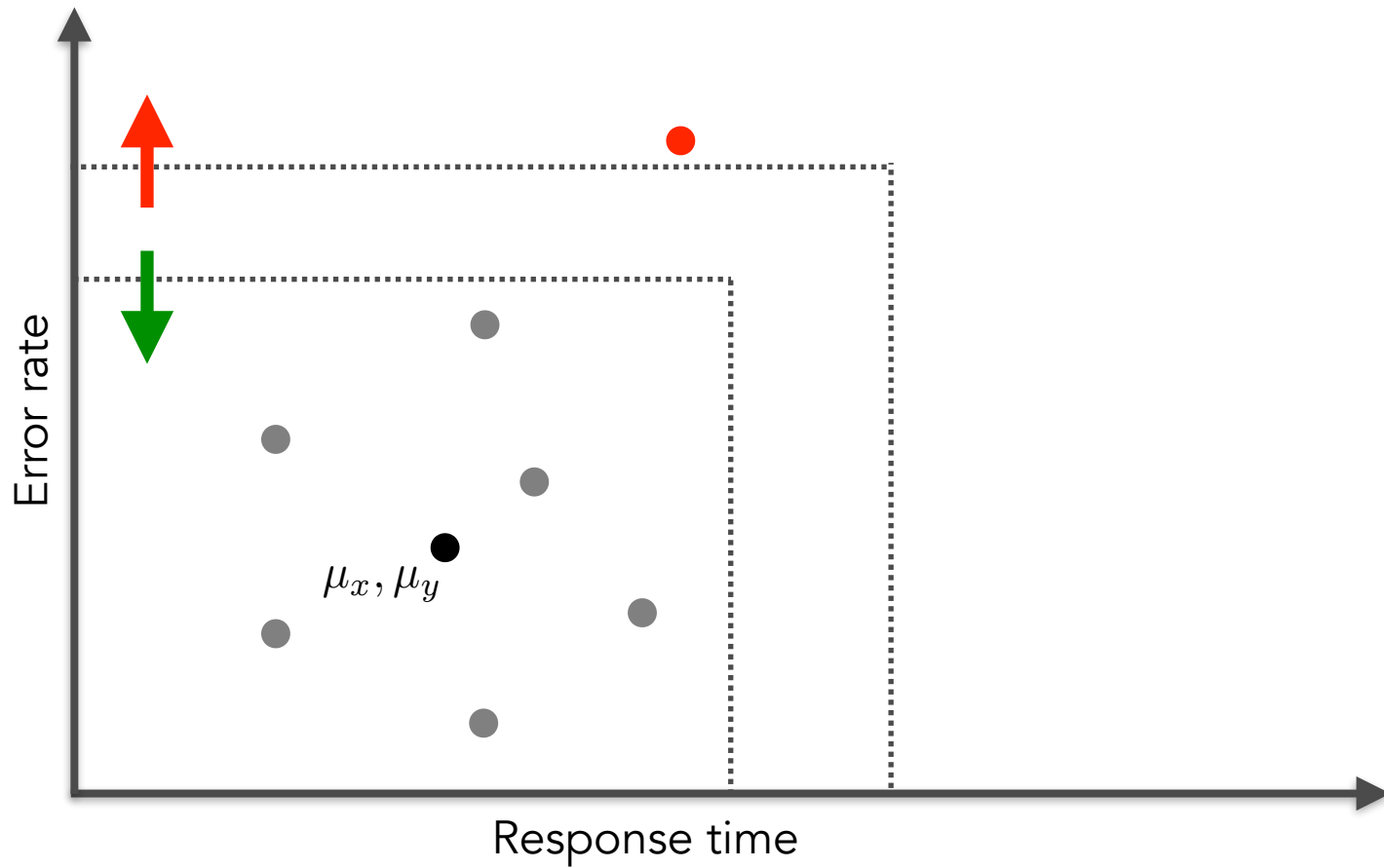


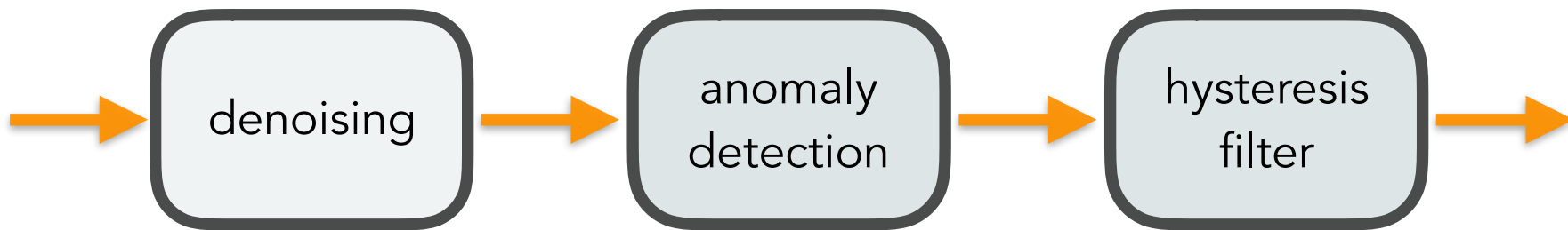












- **Moving Average (MA)**

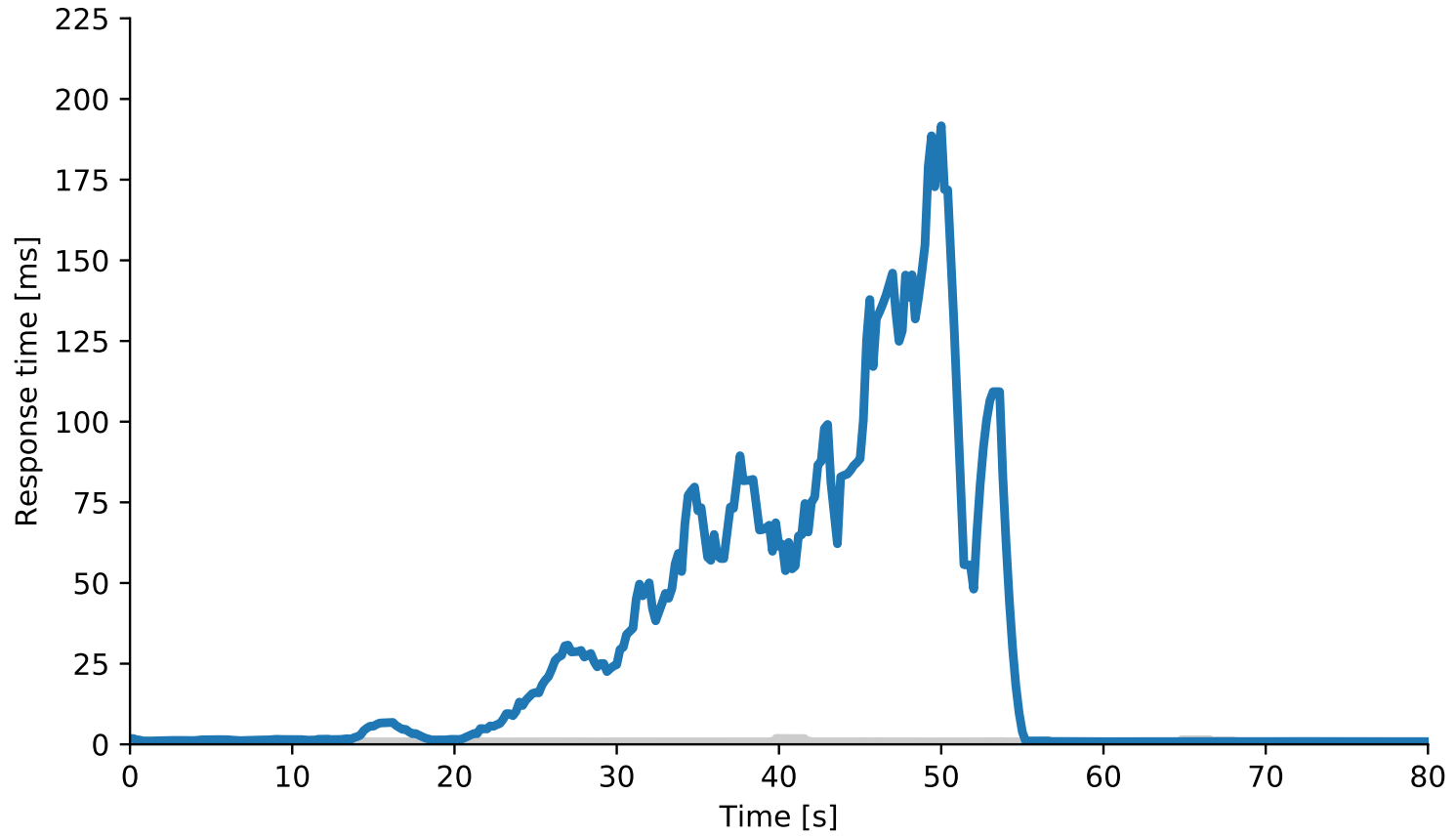
- Weighted MA
- Low-pass filtering
- Rolling quantile
- Karhunen-Loève transform
- Subspace projection

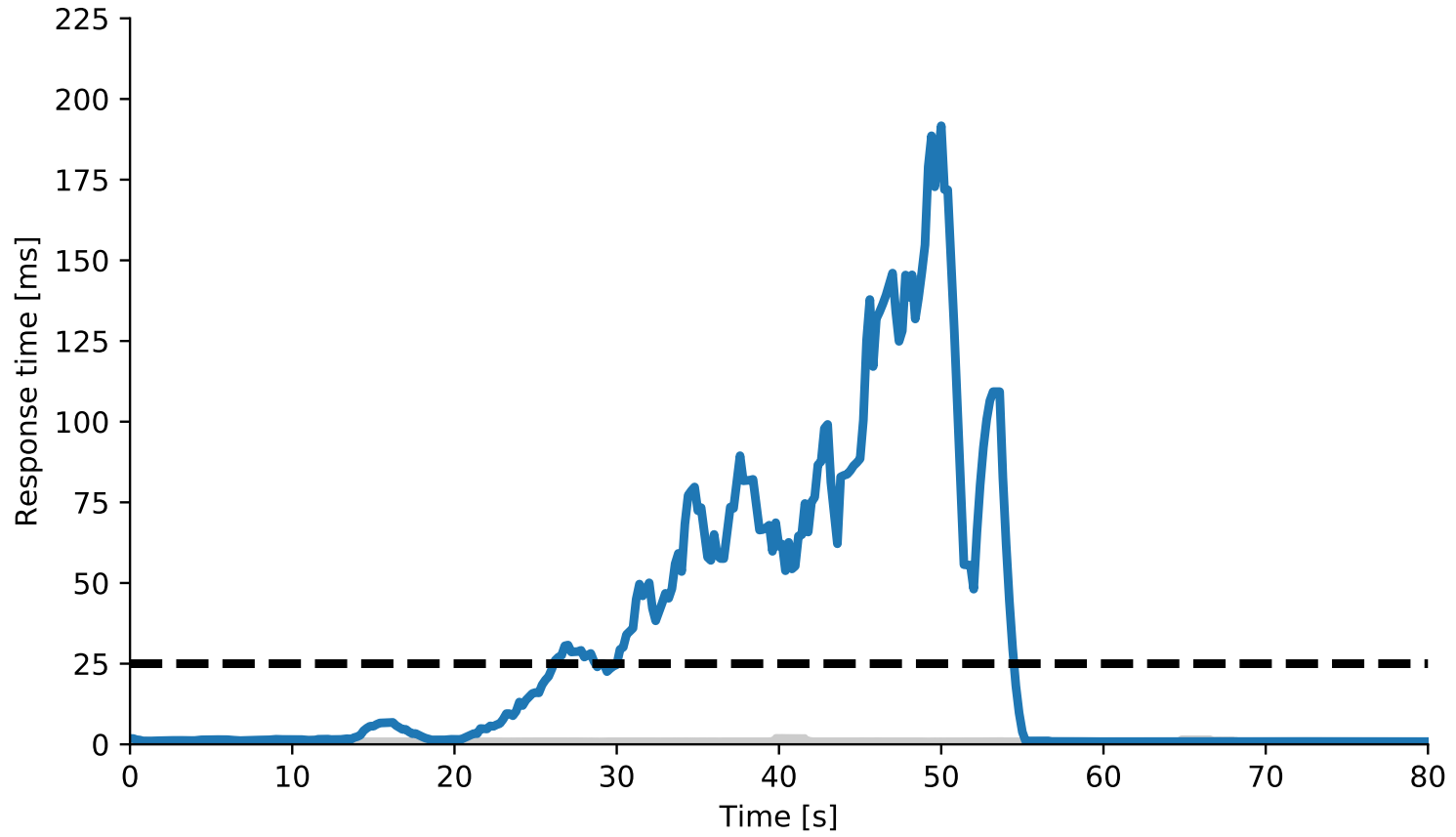
- **Simple thresholding**

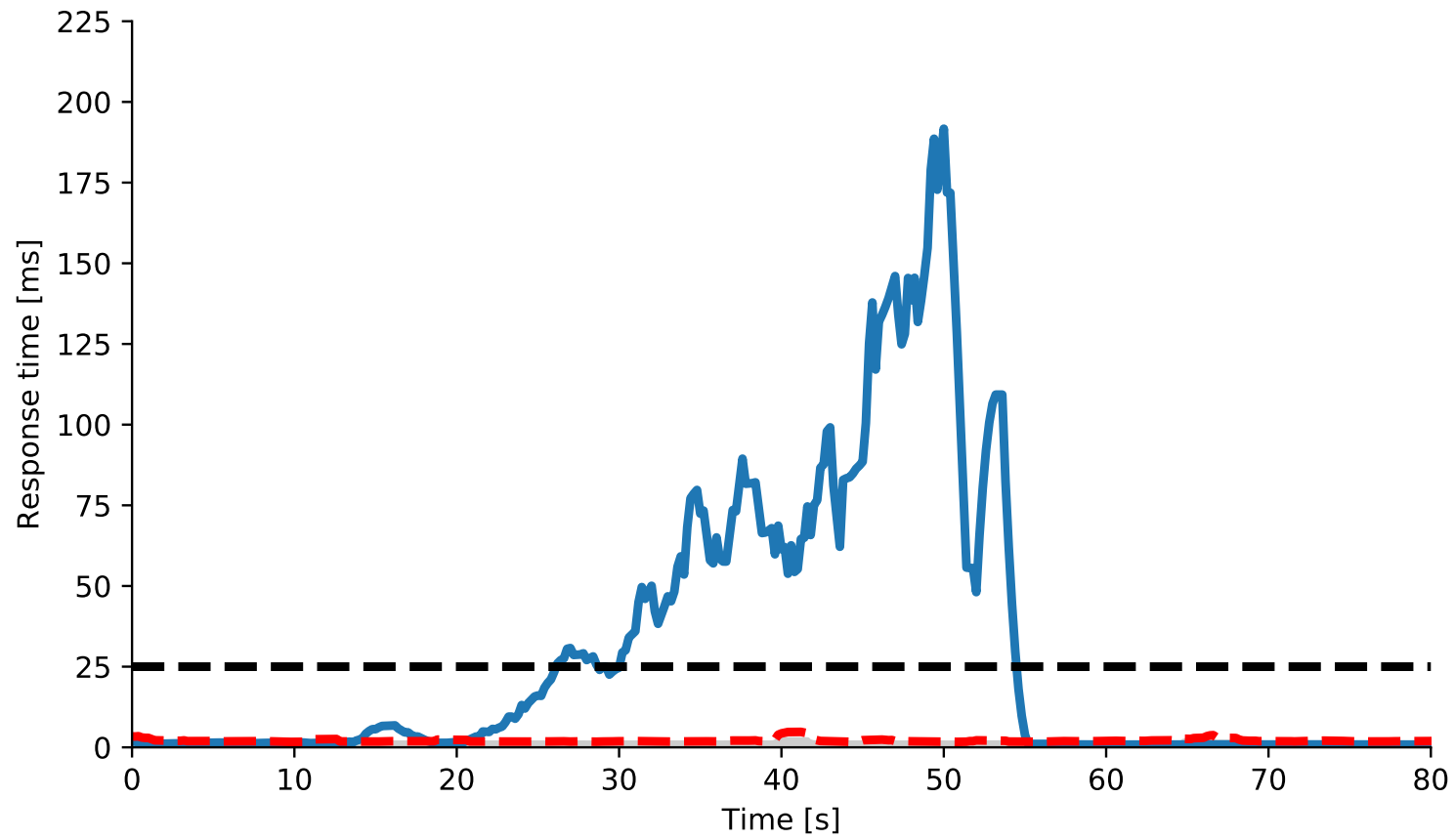
- Hypothesis testing
- Conditional entropy
- Distributional thresholding
- Pattern matching/Clustering

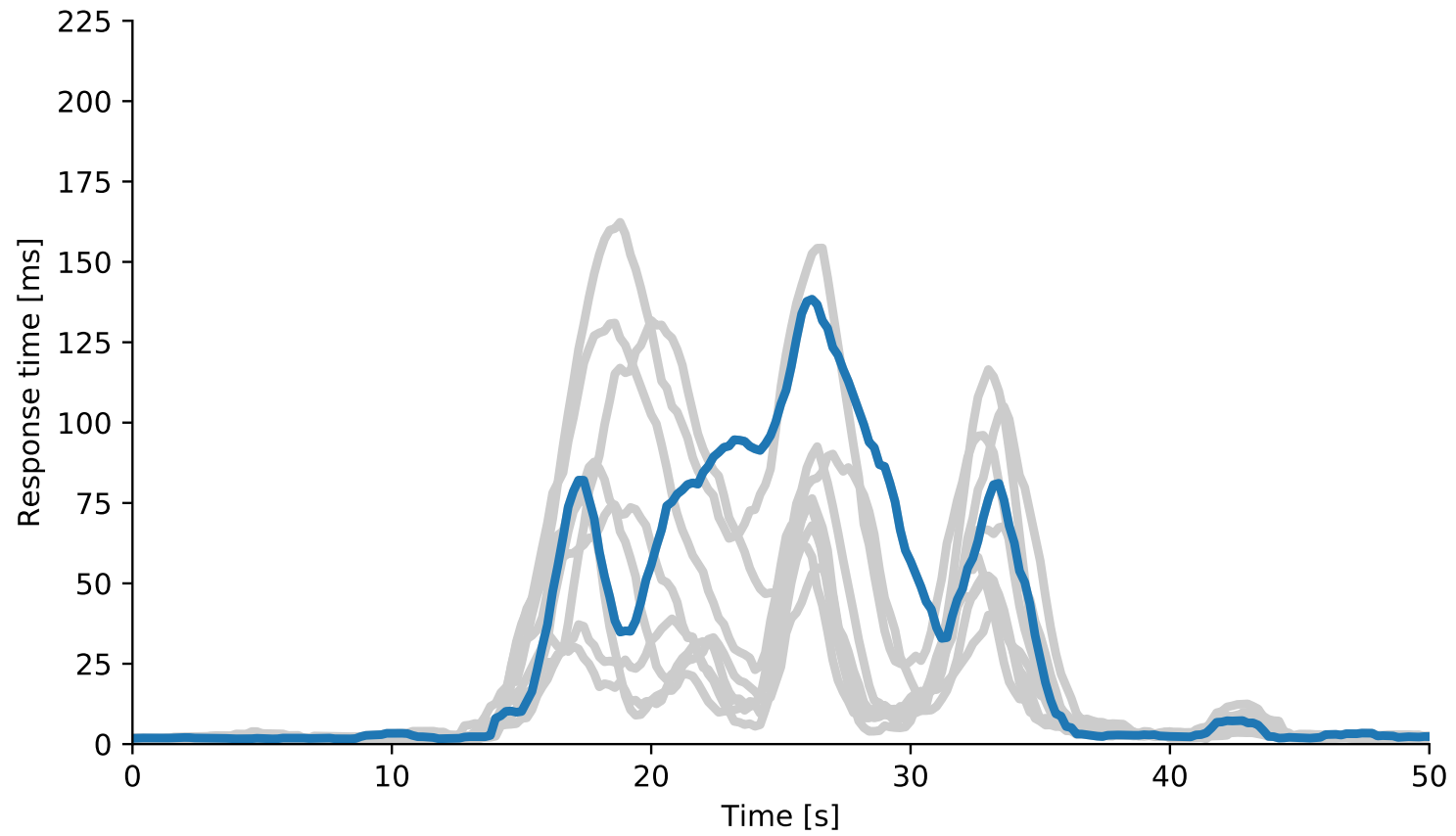
- **Sharp hysteresis**

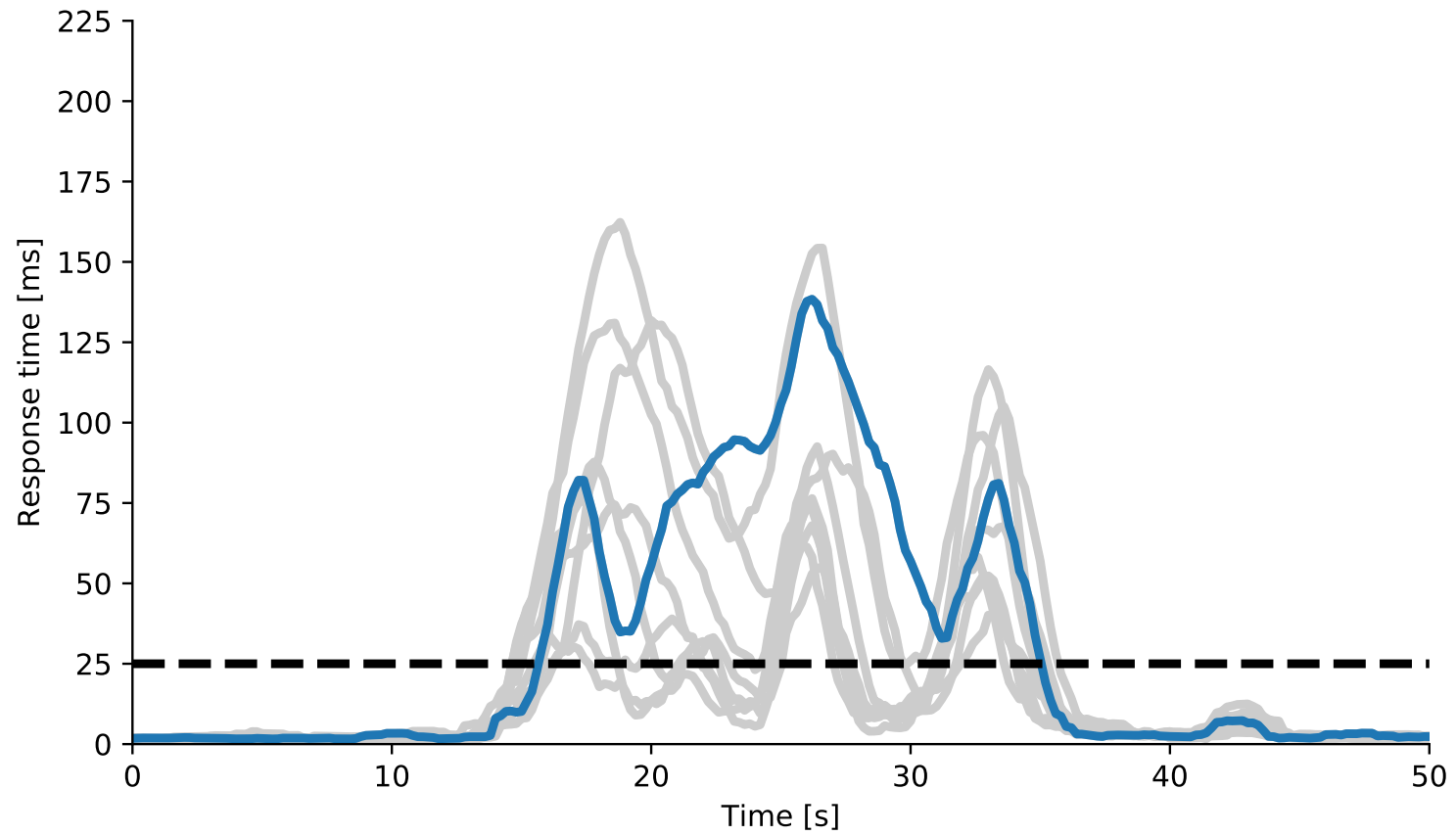
- Continuous hysteresis
- Finite State Machine
- Fuzzy logic program

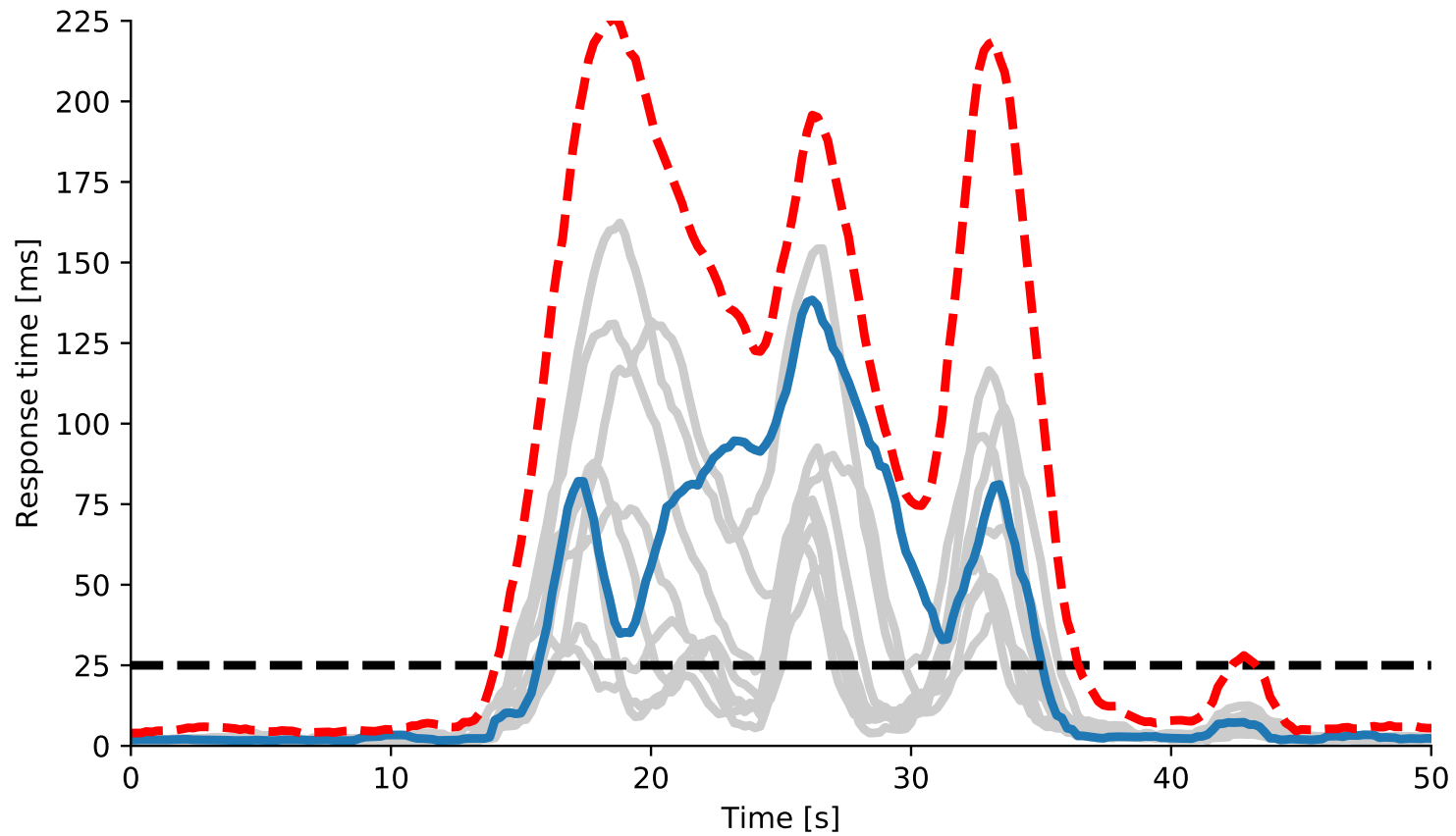








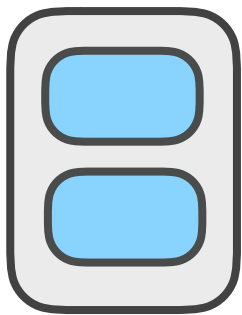




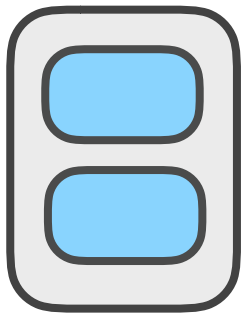


fastly

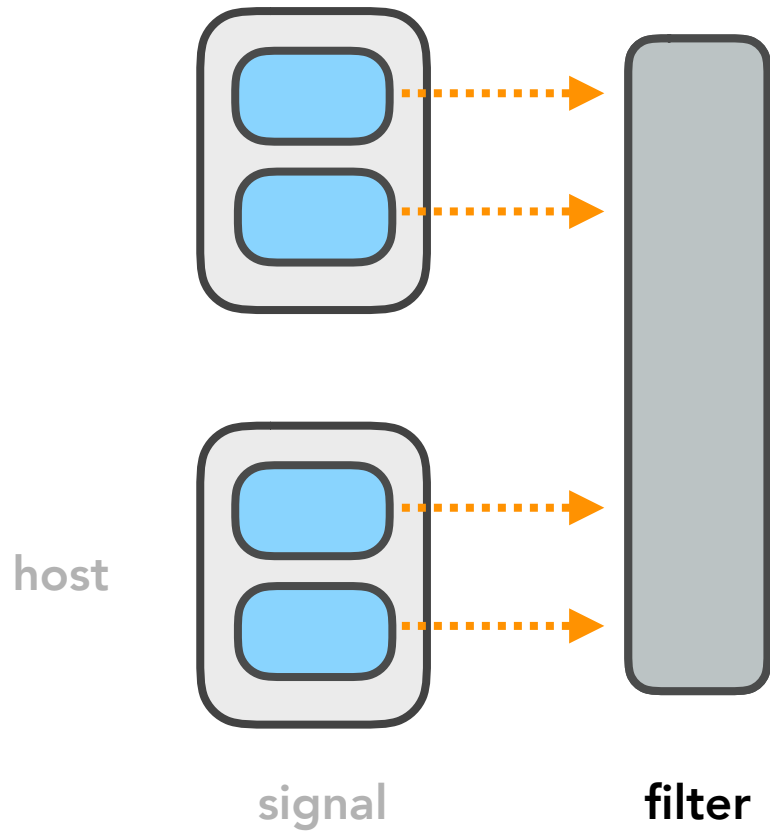
Implementation

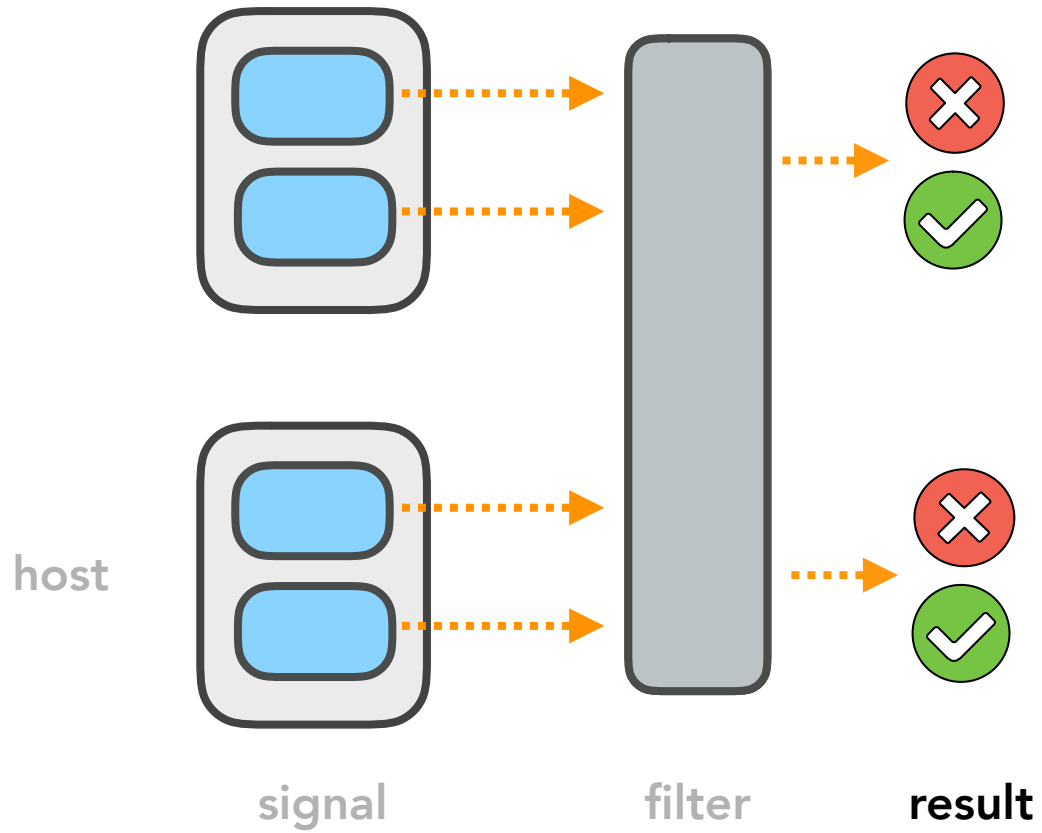


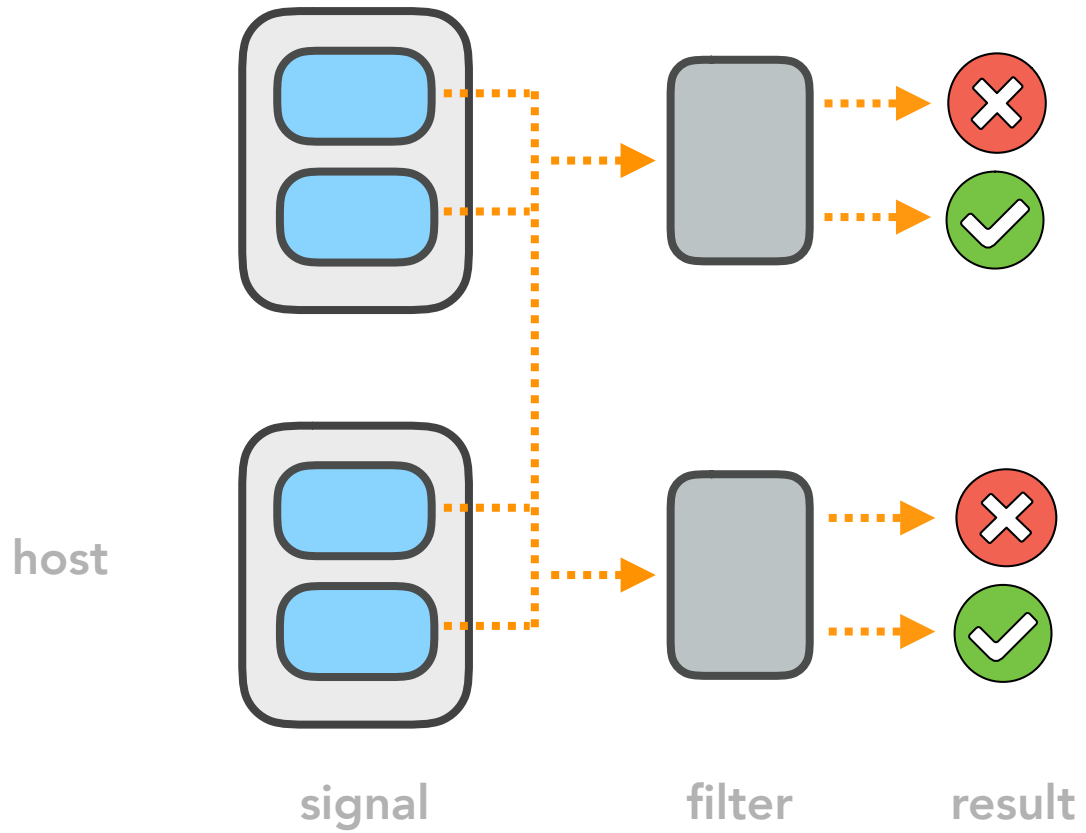
host

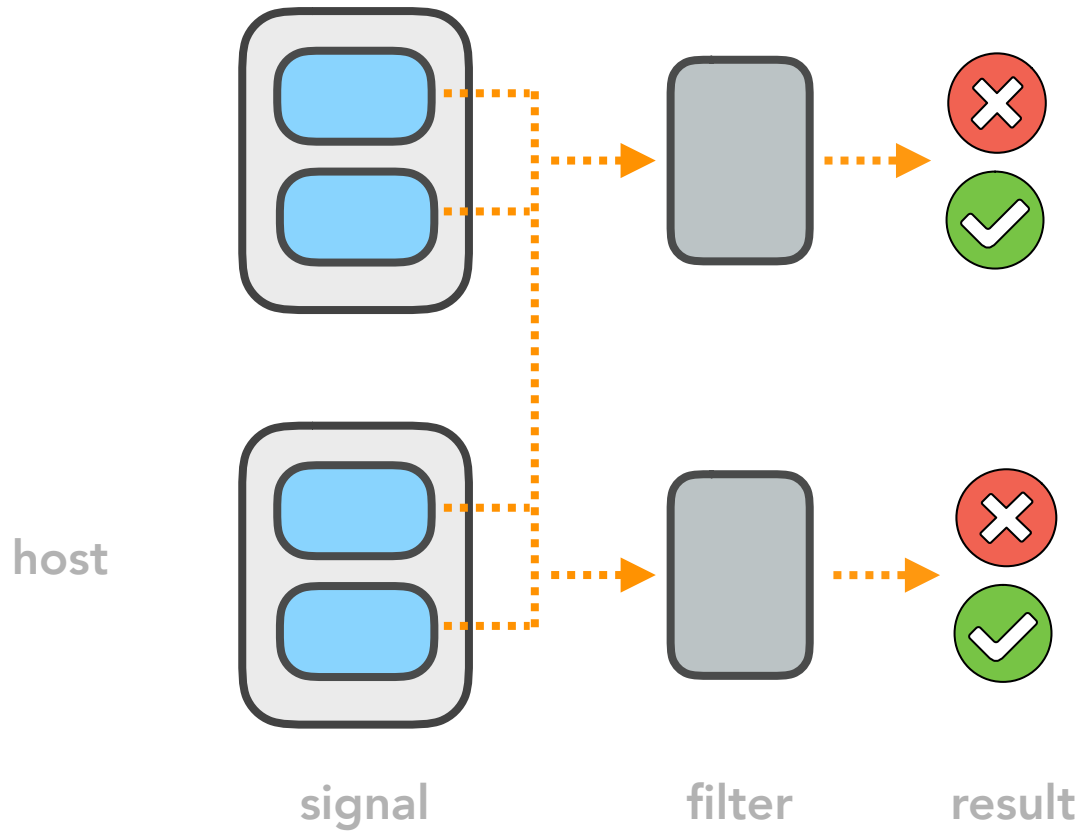


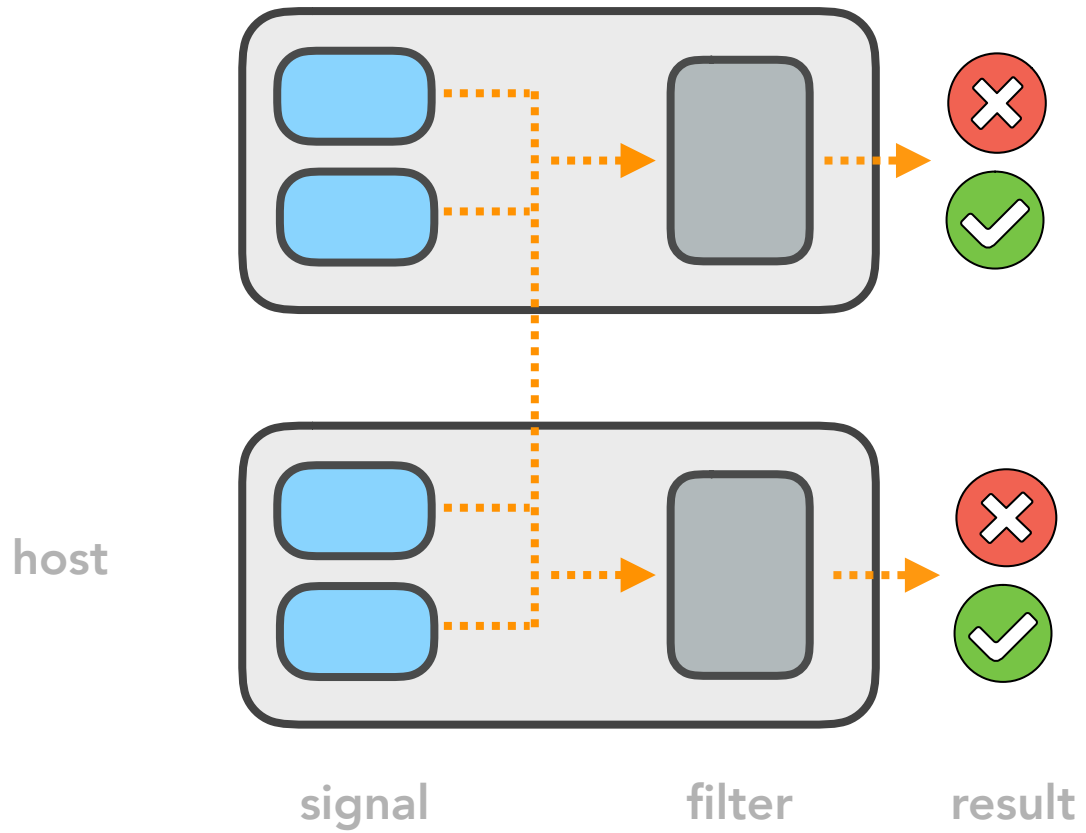
signal











Summary

Health-checking based on global service state

Local state is insufficient to timely and accurately identify all failure modes

Model health-checking as a classification/filter design/stable control problem

Makes it possible to exploit tools and techniques from other disciplines

Distribute replicas of health-check processing logic to individual hosts

Resilient design without requiring distributed consensus

References

Denoising

- S. Mitra, J. Kaiser, "Handbook for Digital Signal Processing", Wiley Interscience, 1993
- J. Proakis, D. Manolakis, "Digital Signal Processing", 4th edition, Prentice Hall, 2006

Anomaly detection

- C. Bishop, "Pattern Recognition and Machine Learning", Springer, 2013
- T. Dunning and E. Friedman, "Practical Machine Learning", O'Reilly, 2014

Stable control

- K. Ogata , "Modern Control Engineering", 5th edition, Pearson, 2009
- P. Janert, "Feedback control for Computer Systems", O'Reilly, 2014

FIN