



Leveraging Machine Learning to Design Better Large-Scale Systems

Christina Delimitrou

Electrical and Computer Engineering

<http://sail.ece.cornell.edu>

October 18th 2019

Systems

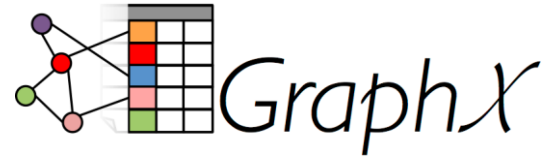


Big Data

Systems



Big Data



Systems

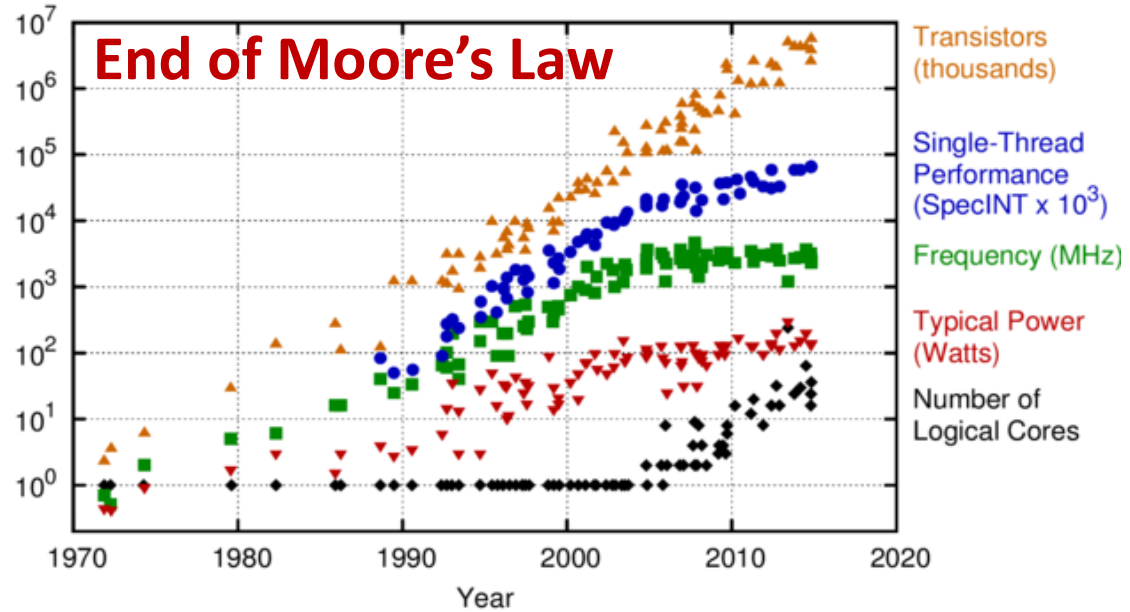


Big Data

Hardware

Hardware

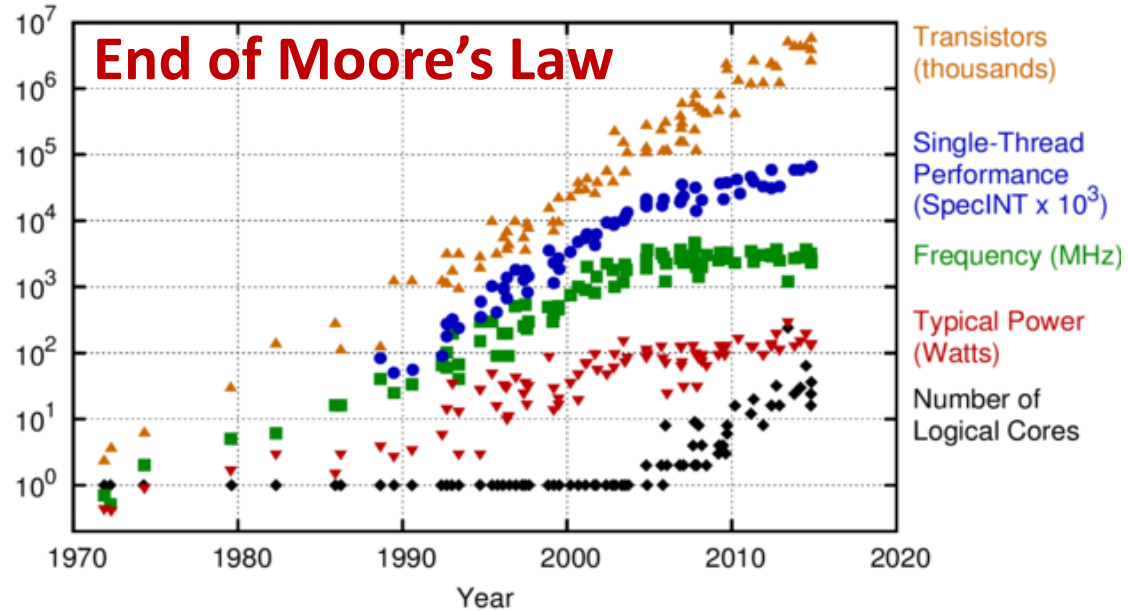
40 Years of Microprocessor Trend Data



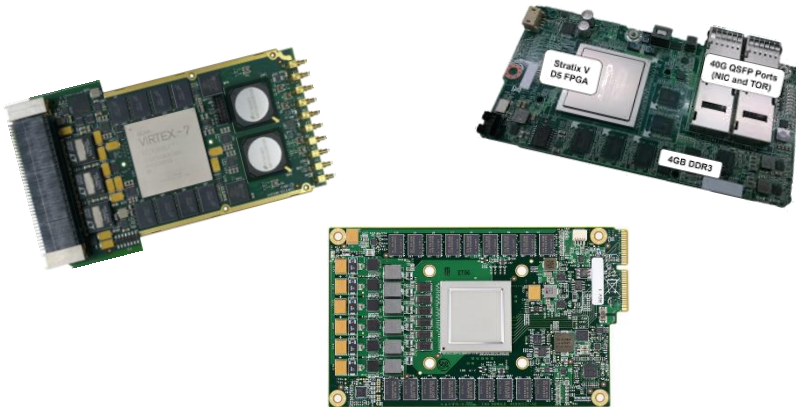
Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2015 by K. Rupp

Hardware

40 Years of Microprocessor Trend Data



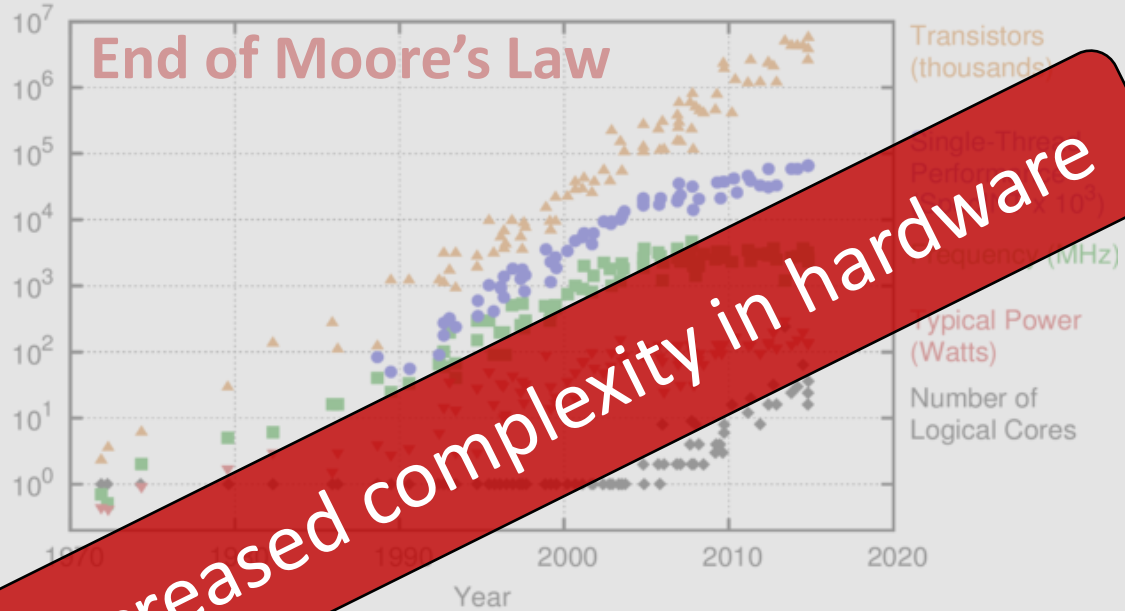
Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2015 by K. Rupp



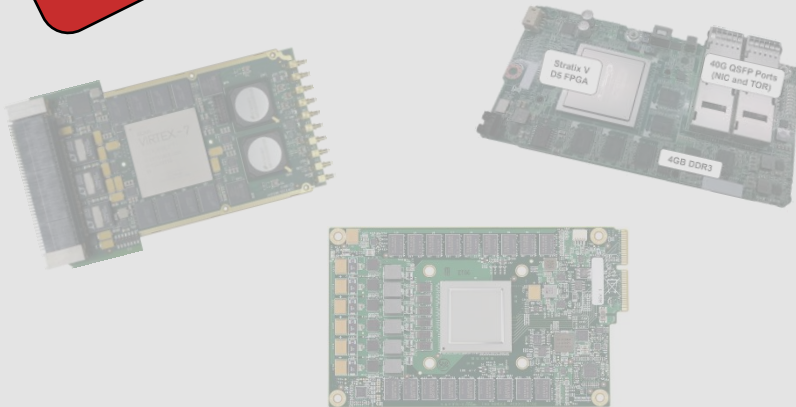
**Hardware
acceleration
in the cloud**

Hardware

40 Years of Microprocessor Trend Data



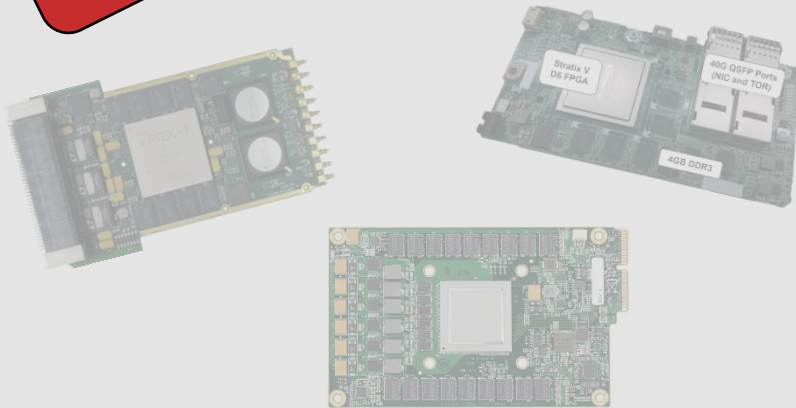
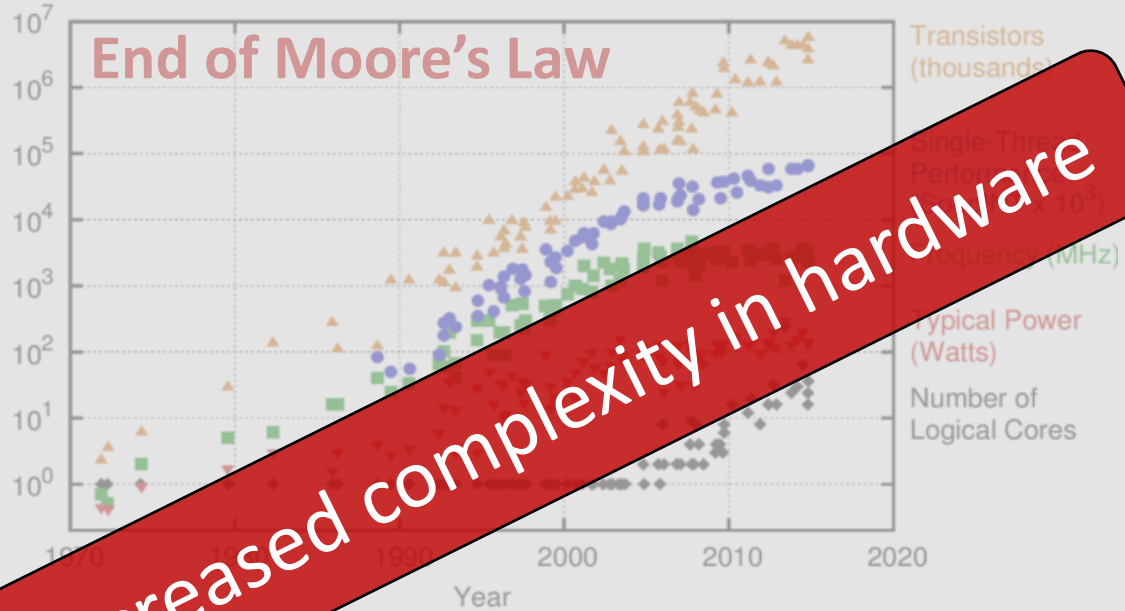
Increased complexity in hardware



Hardware acceleration in the cloud

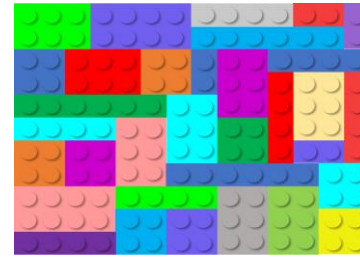
Hardware

40 Years of Microprocessor Trend Data



Hardware acceleration in the cloud

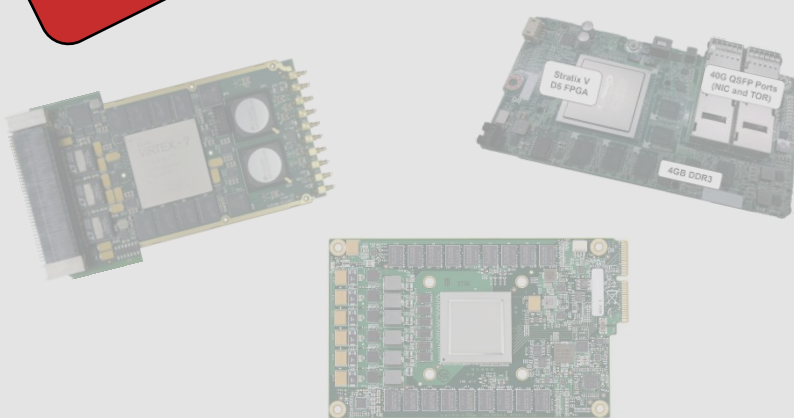
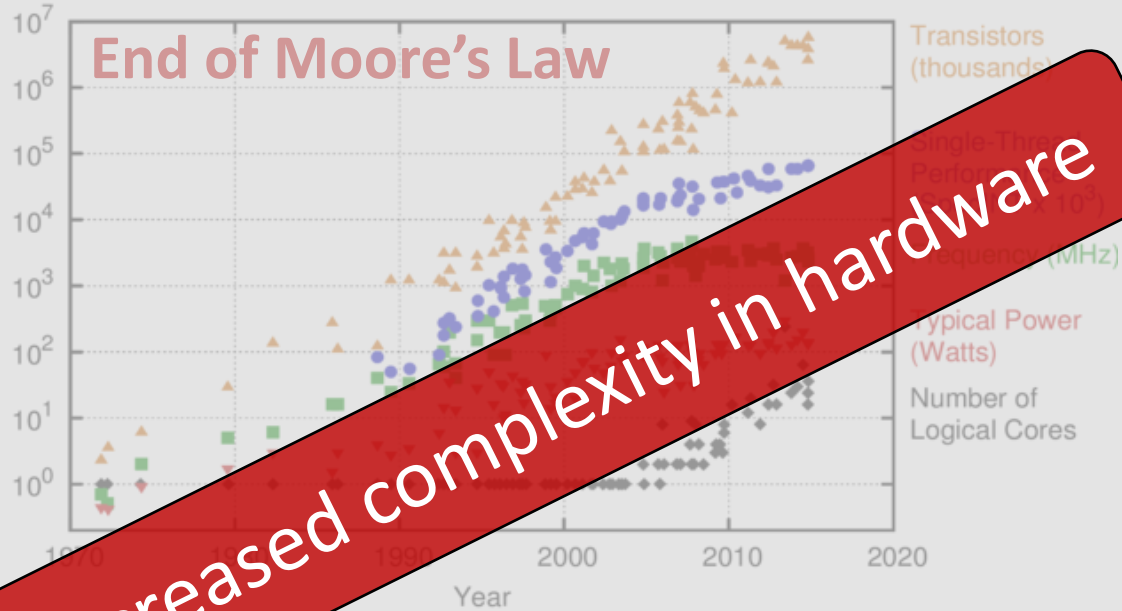
Software



Monolith

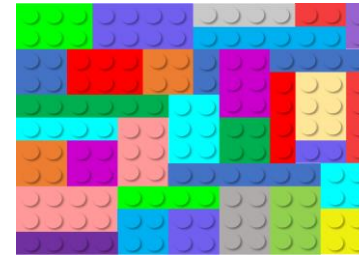
Hardware

40 Years of Microprocessor Trend Data

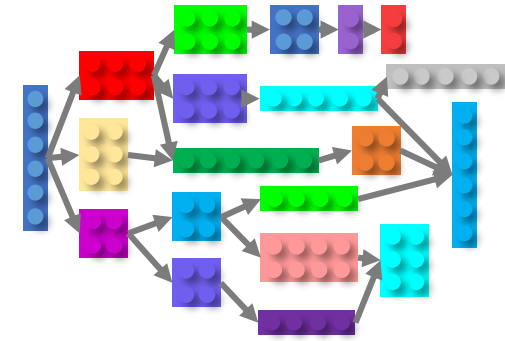


Hardware acceleration in the cloud

Software



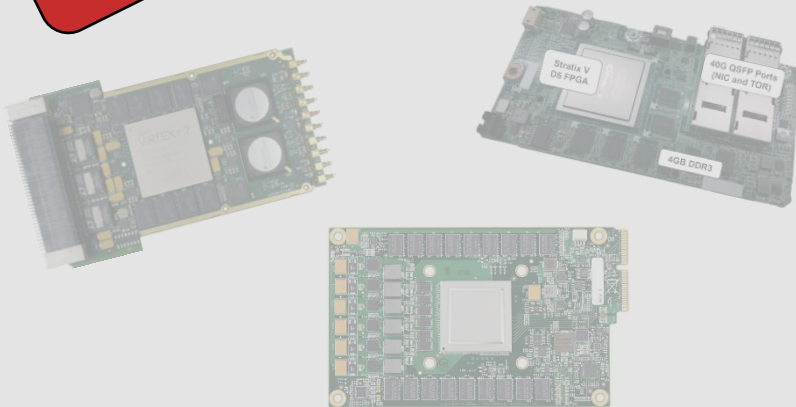
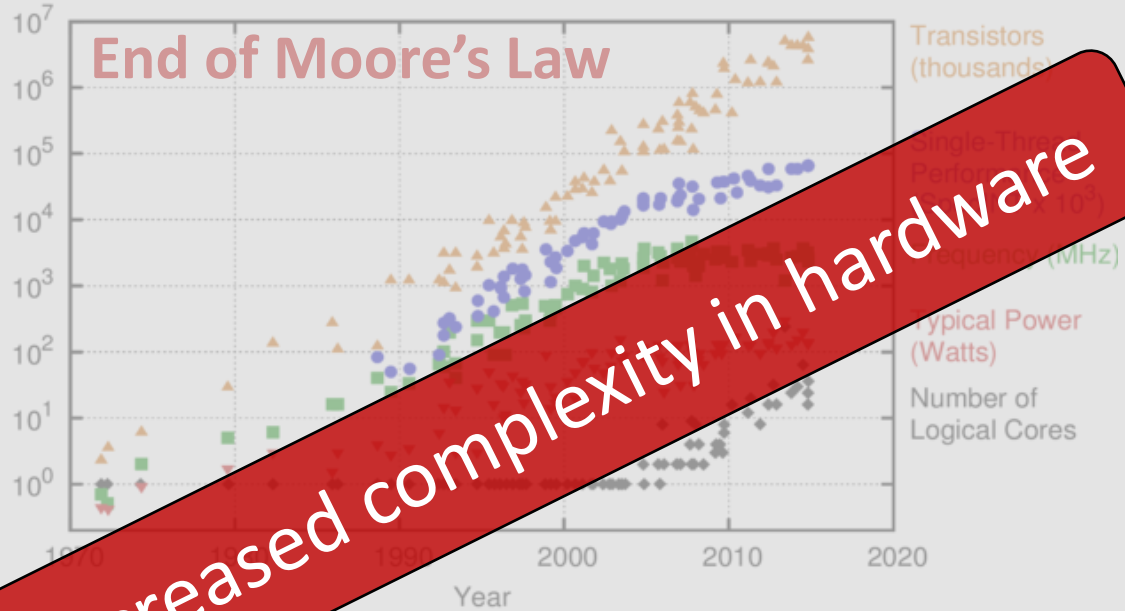
Monolith



Microservices

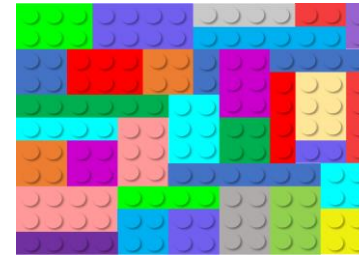
Hardware

40 Years of Microprocessor Trend Data

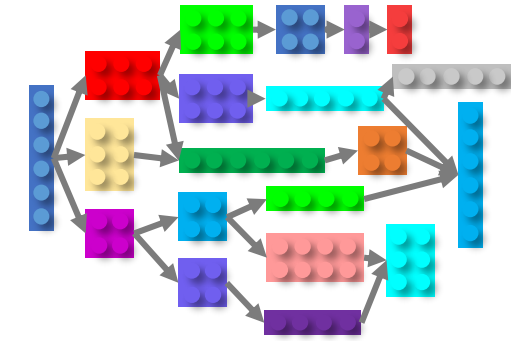


Hardware acceleration in the cloud

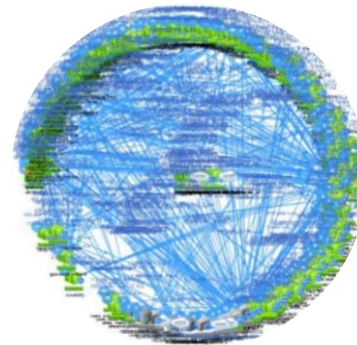
Software



Monolith



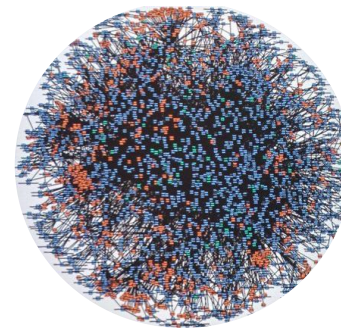
Microservices



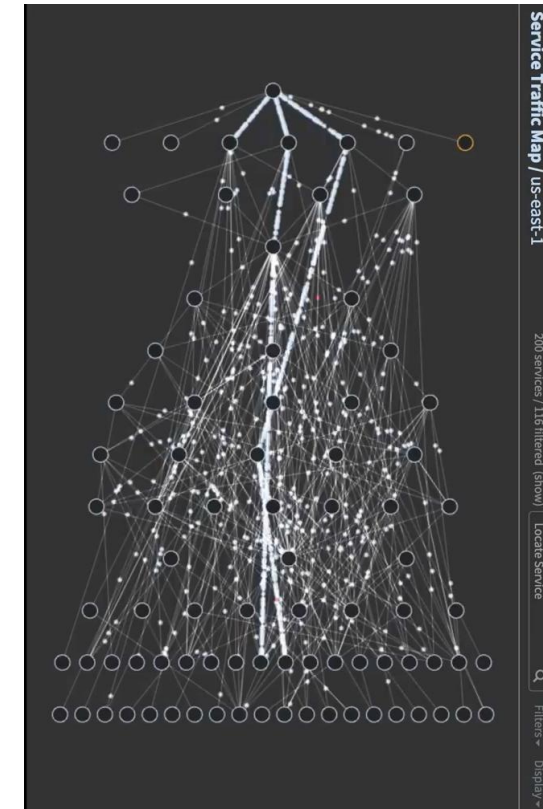
NETFLIX



twitter



amazon.com



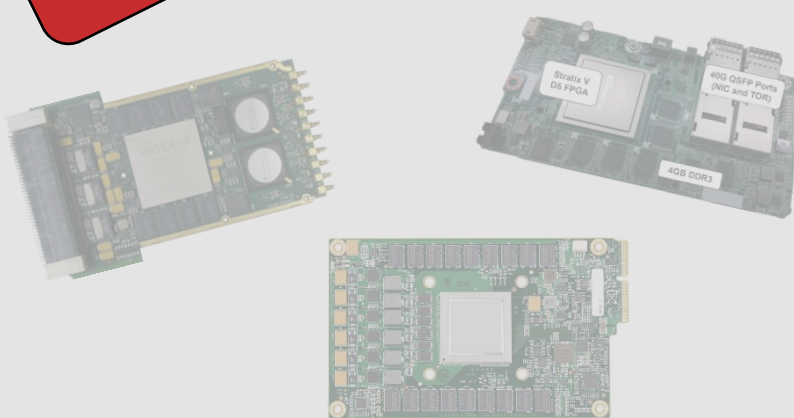
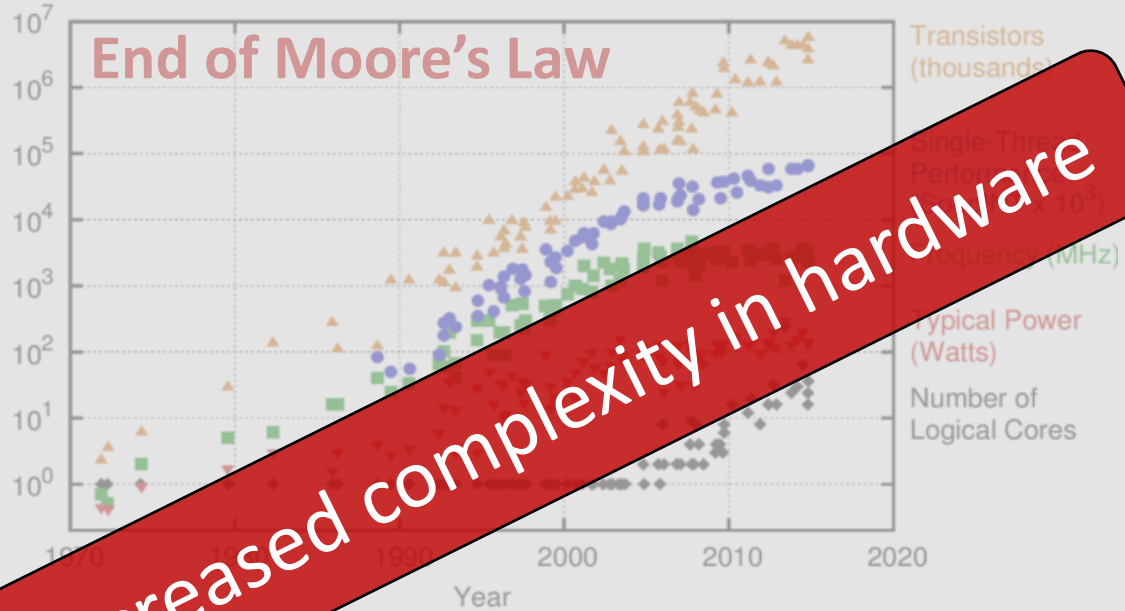
Service Traffic Map / us-east-1

200 services / 116 filtered (show) | Locate Service

Filters | Display

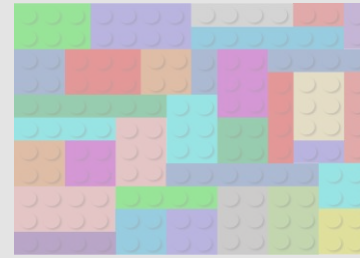
Hardware

40 Years of Microprocessor Trend Data

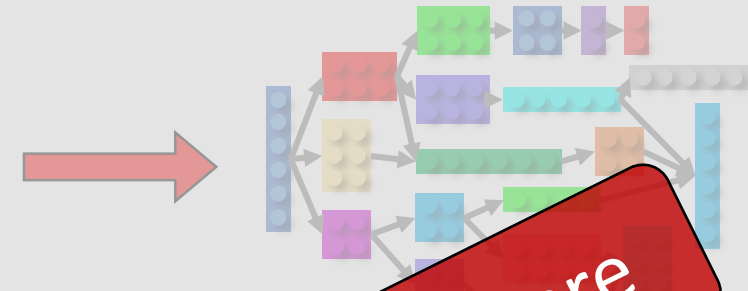


Hardware acceleration in the cloud

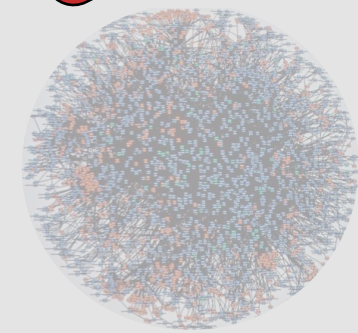
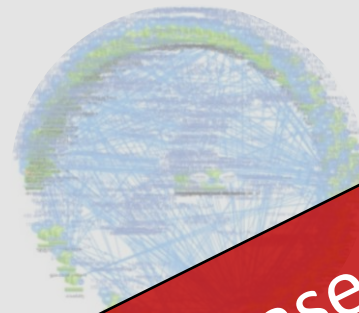
Software



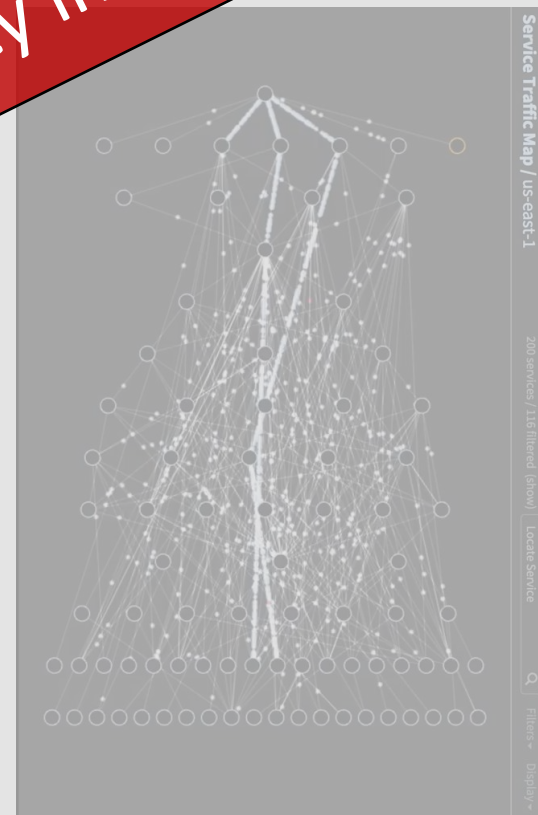
Monolith



Increased complexity in software



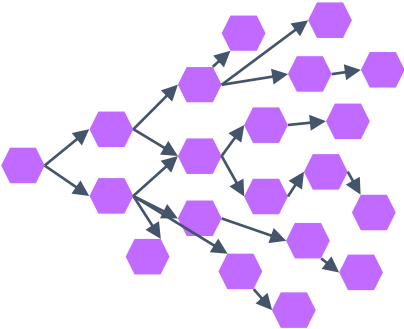
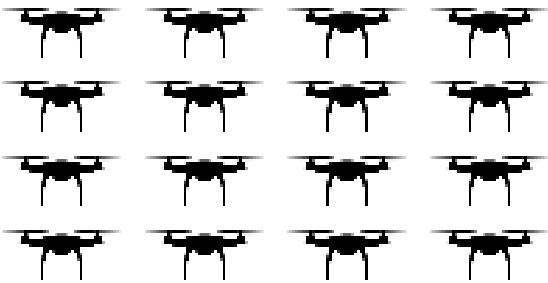
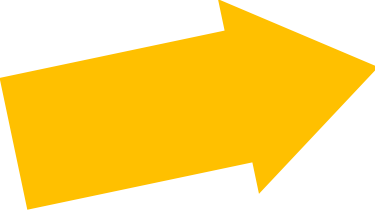
amazon.com



ML in Systems

1

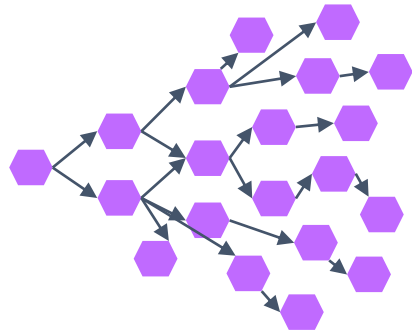
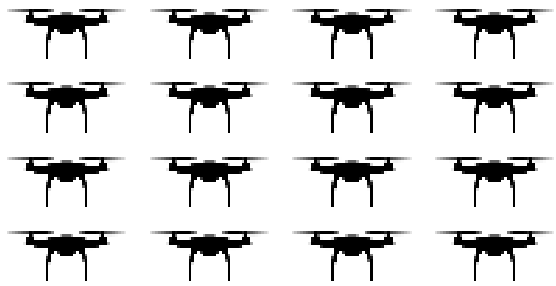
Massive amounts of data



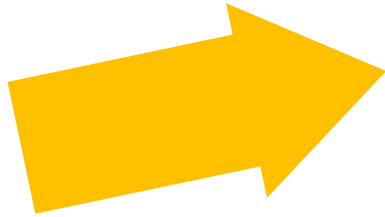
ML in Systems

1

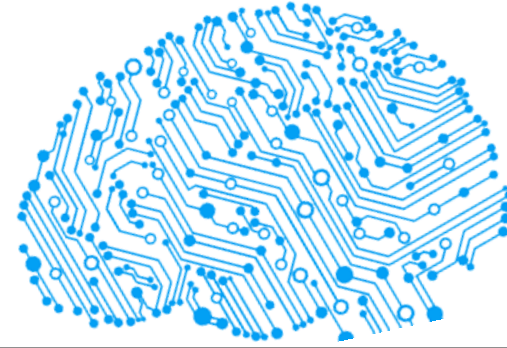
Massive
amounts of data



TraceDB



2



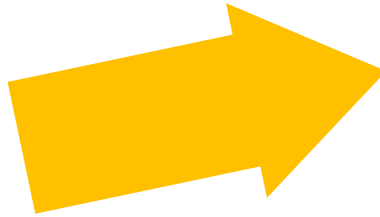
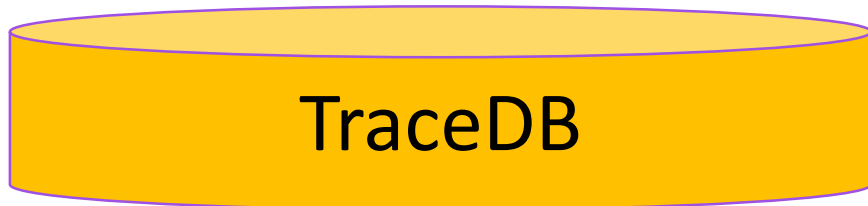
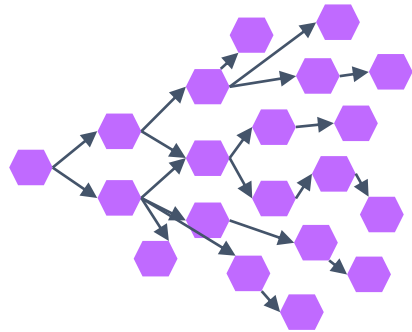
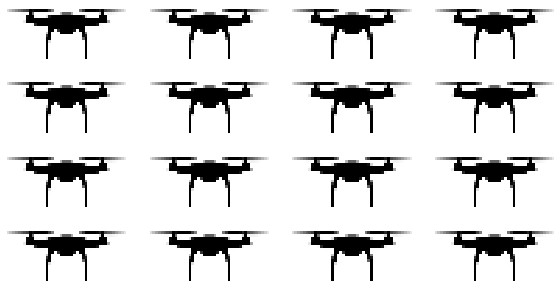
ML algorithmic
advancements



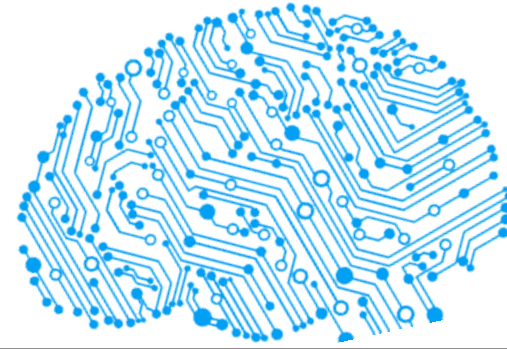
ML in Systems

1

Massive
amounts of data



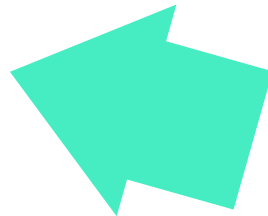
2



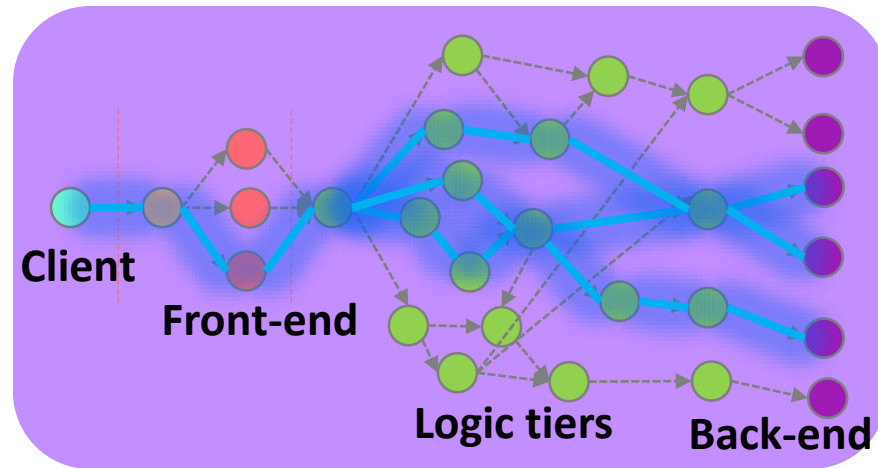
ML algorithmic
advancements



Better system
design & mgmt

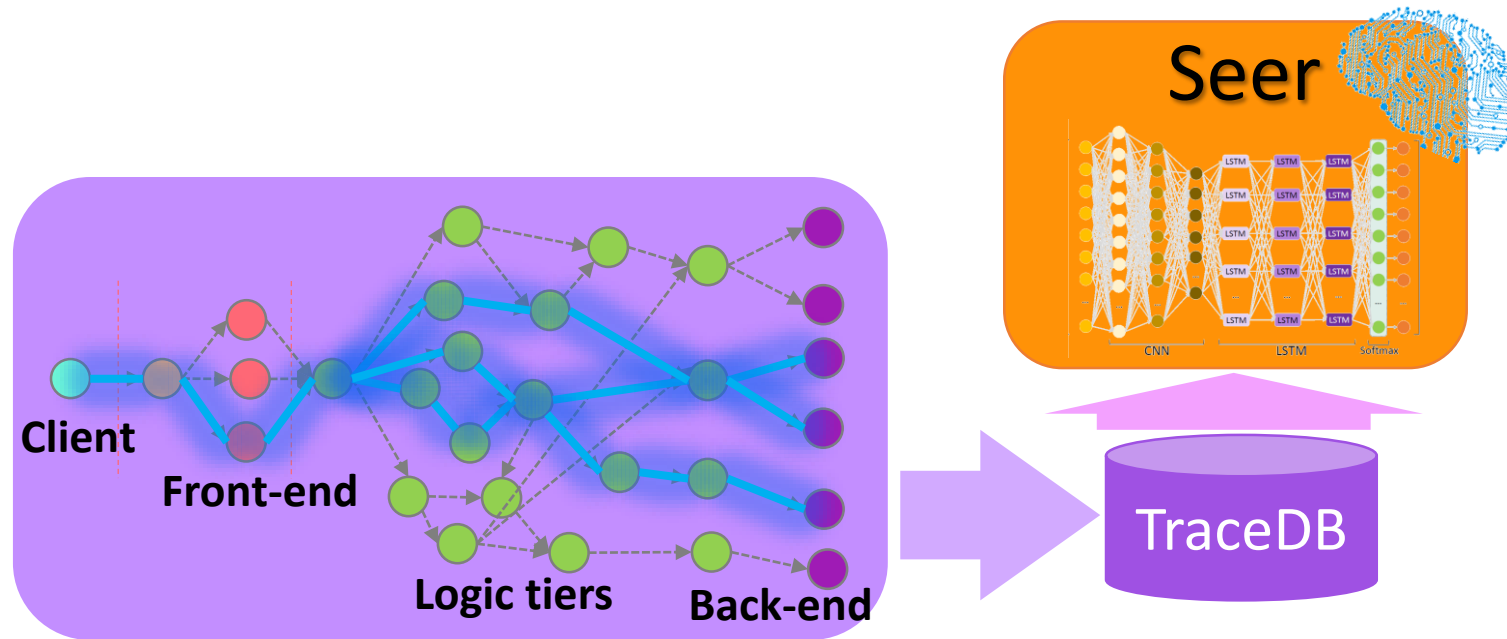


Seer: Using Deep Learning to Debug Cloud Performance



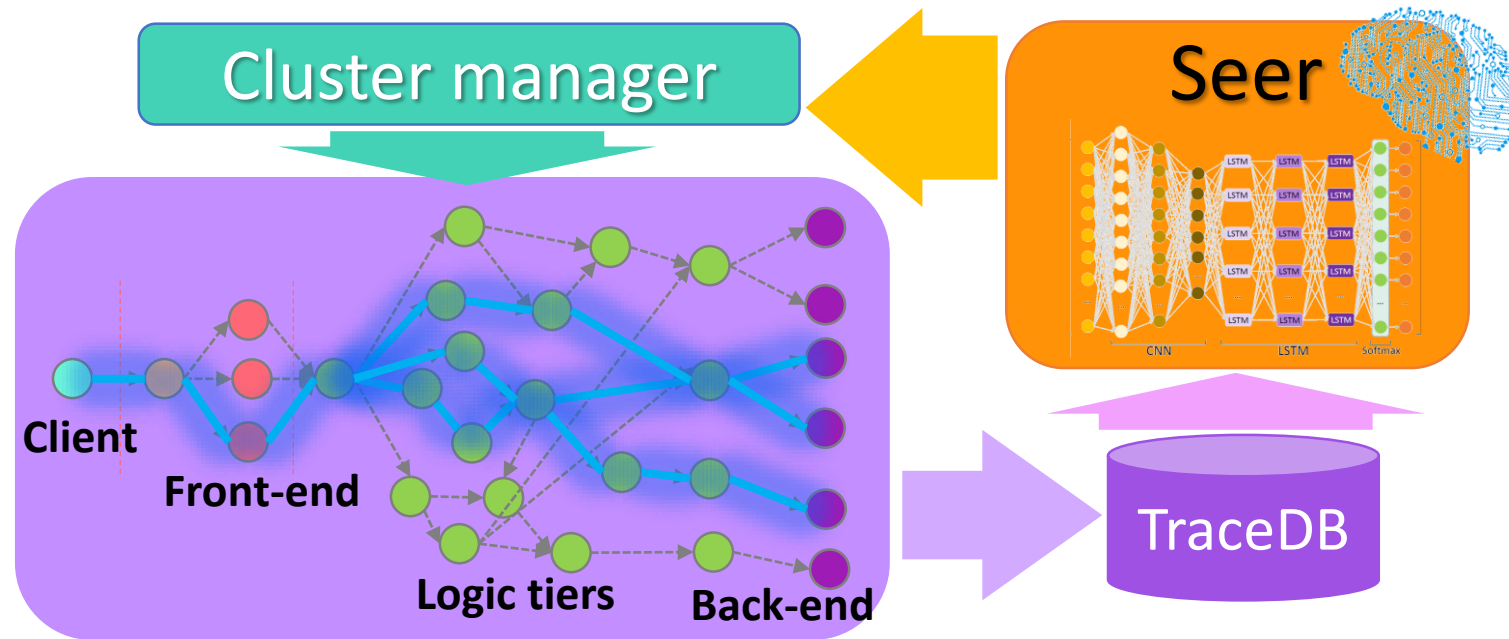
- Use deep learning to anticipate failures and act proactively
- Improves performance & reliability in Cornell internal social network
- Collaborations with Google, Facebook & Microsoft to evaluate in production

Seer: Using Deep Learning to Debug Cloud Performance



- Use deep learning to anticipate failures and act proactively
- Improves performance & reliability in Cornell internal social network
- Collaborations with Google, Facebook & Microsoft to evaluate in production

Seer: Using Deep Learning to Debug Cloud Performance



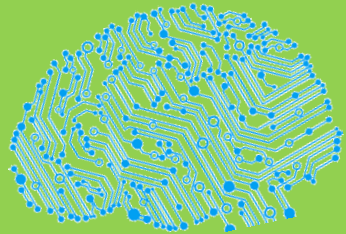
- Use deep learning to anticipate failures and act proactively
- Improves performance & reliability in Cornell internal social network
- Collaborations with Google, Facebook & Microsoft to evaluate in production

ML in Systems



Hardware design
& acceleration

ML in Systems



Cluster management



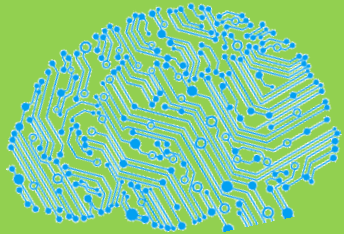
Hardware design
& acceleration

ML in Systems

Serverless



Programming frameworks

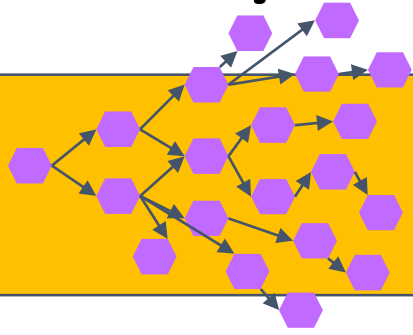


Cluster management

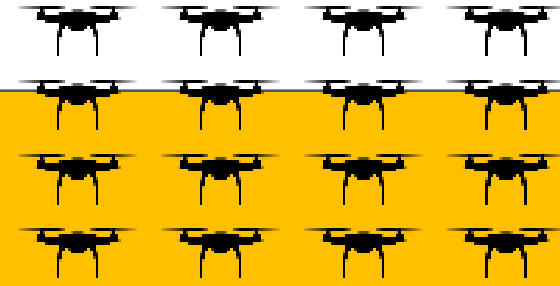


Hardware design
& acceleration

ML in Systems



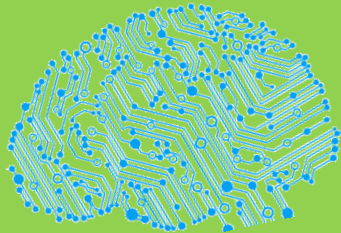
Applications



Serverless



Programming frameworks

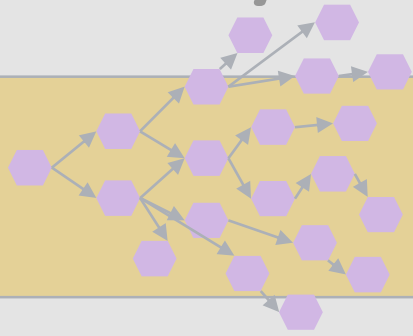


Cluster management

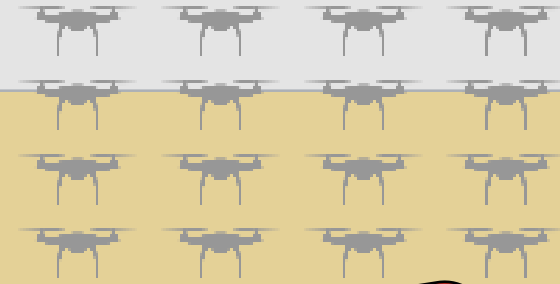


Hardware design
& acceleration

ML in Systems



Applications

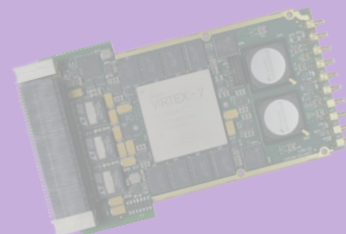


Serverless

ML → practical & scalable improvement
in design & mgmt of complex systems



Cluster management



Hardware design
& acceleration

More info:

<http://sail.ece.cornell.edu>