



Dr. Roberto Yus  
<https://robertoyus.com/>

# Data Privacy

## CMSC 463/663

L08 – Secure Multi-Party Computation



# Previously on...

- Differential Privacy is current state of the art for privacy protection
- Privacy parameter ( $\epsilon$ ) to adjust the tradeoff between the level of privacy loss and data quality
- Not for all privacy problems!
  - Statistical releases
  - Works well with large amounts of data

**Your personal data all over the web - is there a better way?**

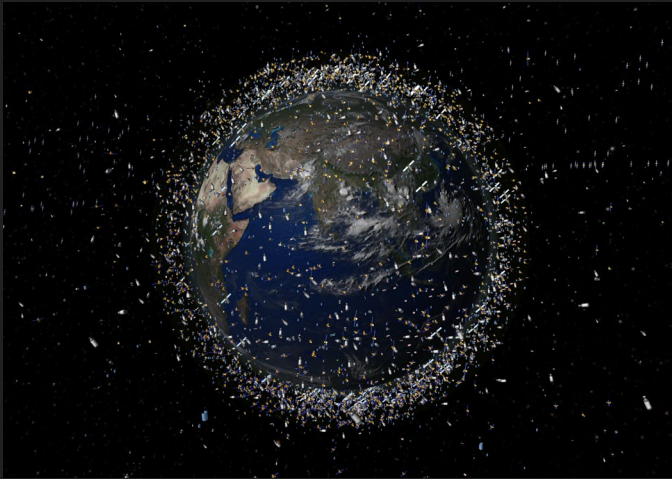
7 March 2024

By Sean McManus, Technology Reporter



***In the news!***

# Privacy and Collaboration



**Parties:** International satellite operators

**Desired output:** potential collisions

**Private information:** location, maneuver  
schedule



# Other Examples

- Elections

- N parties, each vote “yes” or “no”
- Goal: determine whether the majority voted “Yes”, but no voter should learn how other people voted

- Auctions

- Each bidder makes an offer
  - Offer should be committing! (can’t change it later)
- Goal: determine whose offer won without revealing losing offers

- Distributed data mining

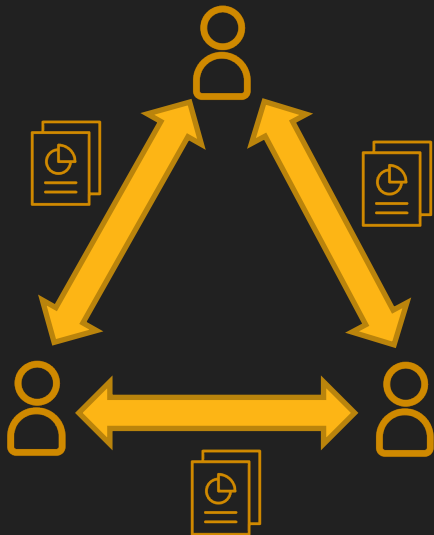
- Two companies want to compare their datasets without revealing them
  - For example, compute the intersection of two lists of names

- Database privacy

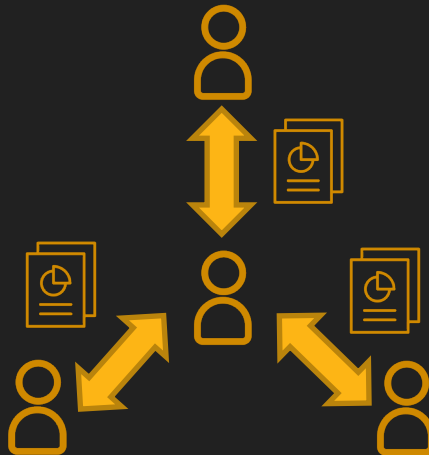
- Evaluate a query on the database without revealing the query to the database owner
- Evaluate a statistical query on the database without revealing the values of individual entries
- Many variations

# Current Approaches to Collaboration

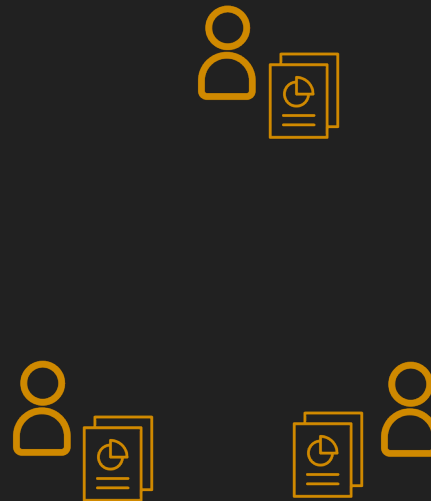
Share with each other



Share with external party



Don't share



***Collaboration requires giving data to trusted parties, accepting security and privacy risks***

# Secure Multi-Party Computation (MPC)

- Goal: replace trusted party with technology
- Requirements
  - **Correctness:** everyone learns correct result of computation
  - **Privacy/security:** no one learns anything beyond result
- **MPC provides correctness and security without trusted party**
  - For any computation
  - For any number of parties

# MPC: The First 40 Years

Shamir  
secret  
sharing    GMW    BGW

1980s: *Existence*

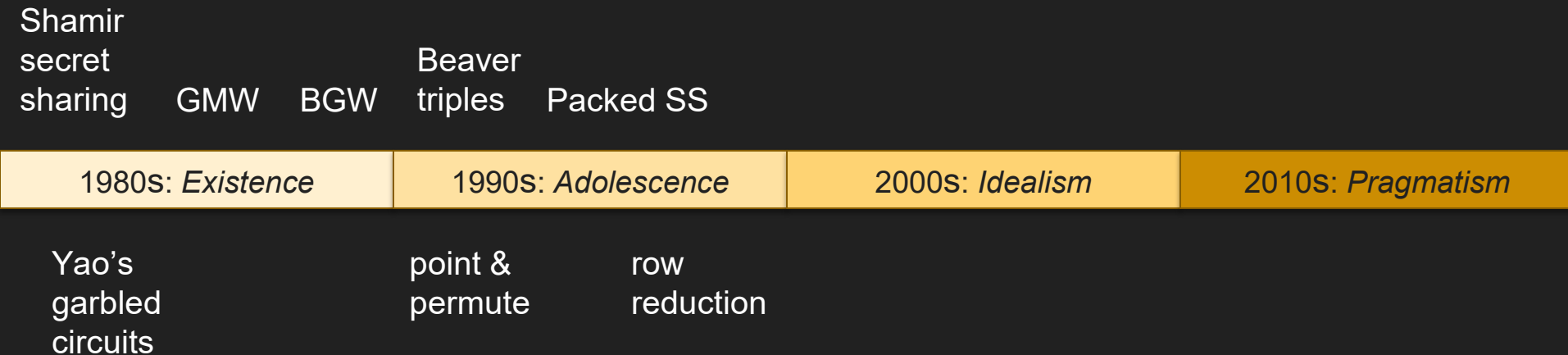
1990s: *Adolescence*

2000s: *Idealism*

2010s: *Pragmatism*

Yao's  
garbled  
circuits

# MPC: The First 40 Years



# MPC: The First 40 Years

Shamir secret sharing	GMW	BGW	Beaver triples	Packed SS	Homomorphic secret sharing	
1980s: <i>Existence</i>		1990s: <i>Adolescence</i>		2000s: <i>Idealism</i>		2010s: <i>Pragmatism</i>
Yao's garbled circuits		point & permute		row reduction	OT extension	free XOR
				Fairplay		

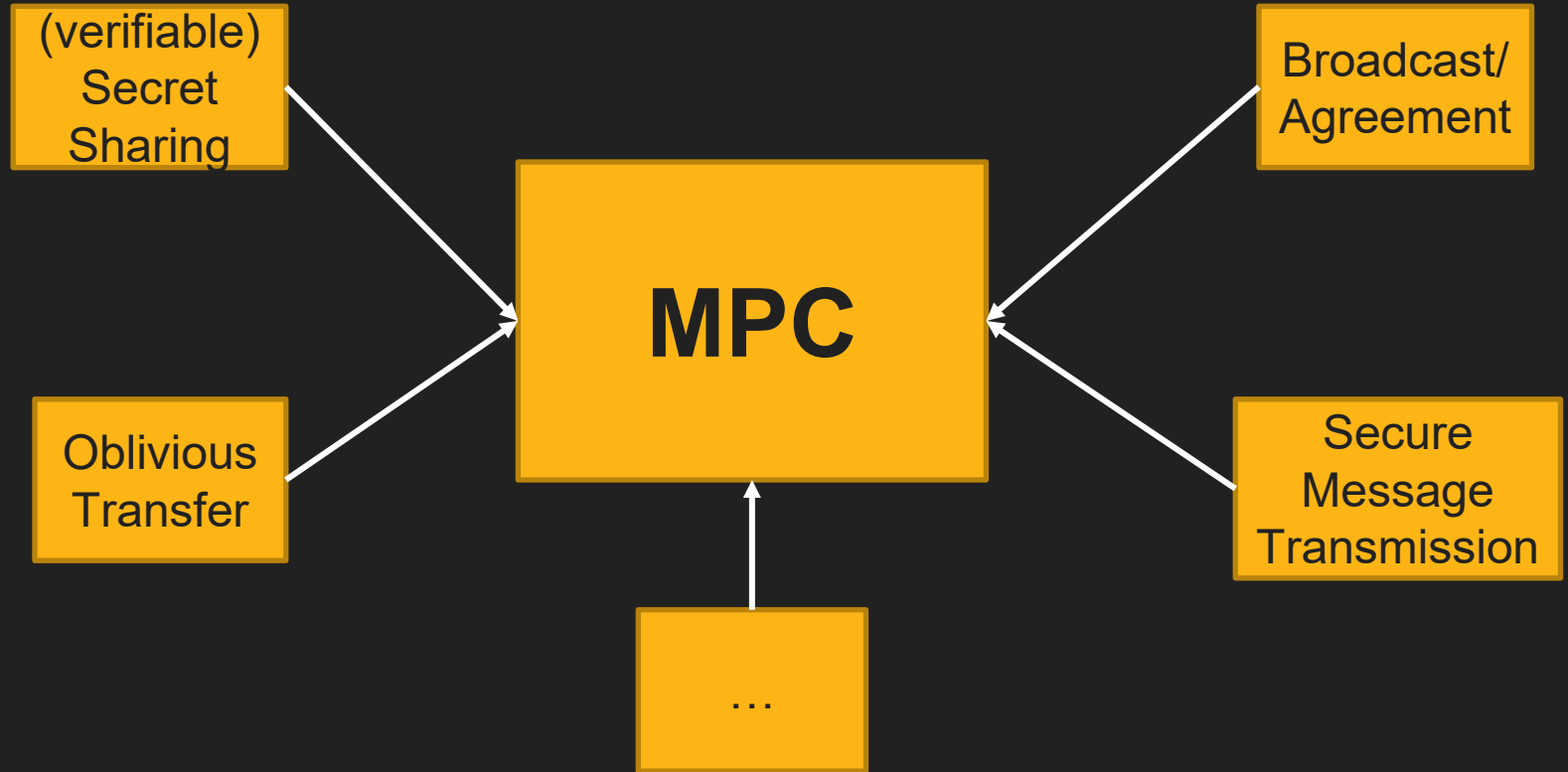
# MPC: The First 40 Years

Shamir secret sharing	GMW	BGW	Beaver triples	Packed SS	Homomorphic secret sharing	Homomorphic Enc and MACs	× via OT
1980s: <i>Existence</i>			1990s: <i>Adolescence</i>		2000s: <i>Idealism</i>		2010s: <i>Pragmatism</i>
Yao's garbled circuits			point & permute	row reduction	OT extension	free XOR	fleXOR half gates
Fairplay							

# Adversarial Models

- Some participants may be **dishonest** (corrupt)
  - If all were honest, we would not need secure multi-party computation
- **Semi-honest** (aka passive; honest-but-curious)
  - Follows protocol, but tries to learn more from received messages than they would learn in the ideal model
- **Malicious**
  - Deviates from the protocol in arbitrary ways, lies about his inputs, may quit at any point

# Building Blocks of MPC



# How MPC Works



Jane \$6200



John \$5800



Bobby \$7300



Jack \$5100

*What's the average salary?*

$$(6200 + 5800 + 7300 + 5100) / 4 \\ = \$6100 \text{ per person}$$

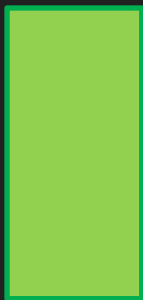
We want to know the avg salary,  
but **we don't want anybody to know our salary**

# How MPC Works

Jane \$6200



John \$5800



Bobby \$7300

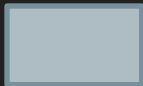
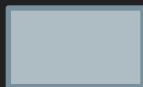


Jack \$5100



# How MPC Works

Jane \$6200



John \$5800



Bobby \$7300



Jack \$5100



# How MPC Works

Jane \$6200

-123

-3478

867

8934

John \$5800

1668

3456

-756

1432

Bobby \$7300

456

4128

-45

2761

Jack \$5100

-1789

664

9606

-3381

# How MPC Works

Jane	John	Bobby	Jack
664	-123	-1789	-3478
867	456	8934	1668
4128	-3381	3456	9606
1432	-756	2761	-45
<b>7091</b>	<b>-3804</b>	<b>13362</b>	<b>7751</b>

# How MPC Works



Jane \$6200

7081



John \$5800

-3804



Bobby \$7300

13362



Jack \$5100

7751

Raw data amount:  $6200 + 5800 + 7300 + 5100$   
 $= 24400$

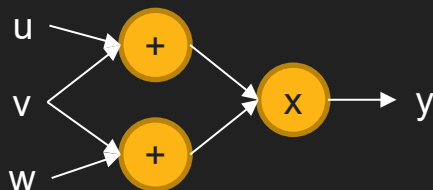
Random data amount:  $7081 - 3804 + 13362 + 7751$   
 $= 24400$

Average:  $24400 / 4 = 6100$  per person

***No one learns anything beyond  
the result of computation***

# MPC for Any Function

## MPC for Arithmetic Computation



MPC for add, multiply primitives over integers can securely compute any function!

## MPC for Boolean Computation



MPC for XOR, AND primitives over bits can securely compute any function!

***MPC can securely compute any function using arithmetic or Boolean primitives***

# Why Secret-Sharing?

- Encryption techniques are **computationally secure**
  - A powerful adversary can break the encryption technique
    - Google, with sufficient computational capabilities, broke SHA-1 (<https://shattered.io/>)
- **Information-theoretical security**
  - Secure regardless of the computational power of an adversary
  - Quantum secure

# The Concept of Secret Sharing

$(n, t)$  LOCKED BOX REPRESENTATION

A secret  $s$



  $P_1$

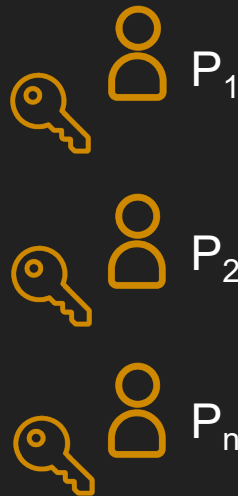
  $P_2$

  $P_n$

# The Concept of Secret Sharing

## $(n, t)$ LOCKED BOX REPRESENTATION

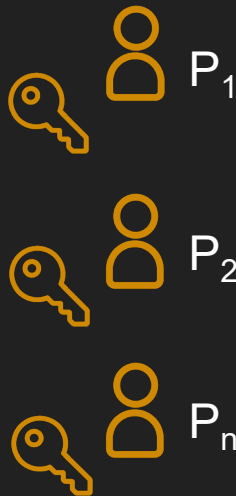
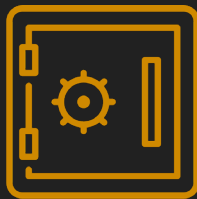
A secret  $s$  is locked in a box



# Secret Sharing: Properties

## $(n, t)$ LOCKED BOX REPRESENTATION

A **secret**  $s$  **locked** in a box



Any  $t$  parties **cannot open** the box

# Secret Sharing: Properties

## $(n, t)$ LOCKED BOX REPRESENTATION

A **secret**  $s$  **locked** in a box

Ex:  $t = 1$



Any  $t$  parties **cannot open** the box

# Secret Sharing: Properties

## $(n, t)$ LOCKED BOX REPRESENTATION

A **secret**  $s$  **locked** in a box



?



Ex:  $t = 1$



Any  $t$  parties **cannot open** the box

# Secret Sharing: Properties

## $(n, t)$ LOCKED BOX REPRESENTATION

A **secret**  $s$  **locked** in a box



?



Ex:  $t = 1$

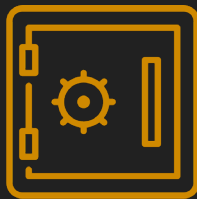


Any  $t$  parties **cannot open** the box

# Secret Sharing: Properties

## $(n, t)$ LOCKED BOX REPRESENTATION

A **secret**  $s$  **locked** in a box



Ex:  $t = 1$



Any  $t$  parties **cannot open** the box

Any  $(t + 1)$  parties **can open** the box

# Secret Sharing: Properties

## $(n, t)$ LOCKED BOX REPRESENTATION

A **secret**  $s$  **locked** in a box



Ex:  $t = 1$

Any  $t$  parties **cannot open** the box

Any  $(t + 1)$  parties **can open** the box

# Secret Sharing: Properties

## $(n, t)$ LOCKED BOX REPRESENTATION

A **secret**  $s$  **locked** in a box



Ex:  $t = 1$



Any  $t$  parties **cannot open** the box

Any  $(t + 1)$  parties **can open** the box

# Shamir's Secret-Sharing (SSS)

## Secret-Share Creation

*e.g., under the assumption that no server will collude*

**Secret  
S**

**Secret Owner**

Share 1 ( $s_1$ )

Share 2 ( $s_2$ )

Share 3 ( $s_3$ )

Share 4 ( $s_4$ )



S1



S2



S3



S4

Each server **cannot** learn the secret **S**

**Non-Communicating Public Servers**

# Shamir's Secret-Sharing (SSS)

**Secret**

**Reconstruction**

*e.g., under the assumption that no server will collude*

**Secret  
S**

**Secret Owner**

Share 1 ( $s_1$ )

Share 2 ( $s_2$ )



S1



S2



S3



S4

**Non-Communicating Public Servers**

# Shamir's Secret-Sharing (SSS)

**Secret**

**Reconstruction**

*e.g., under the assumption that no server will collude*

**Secret  
S**

**Secret Owner**

Share 2 ( $s_2$ )

Share 4 ( $s_4$ )



S1



S2



S3



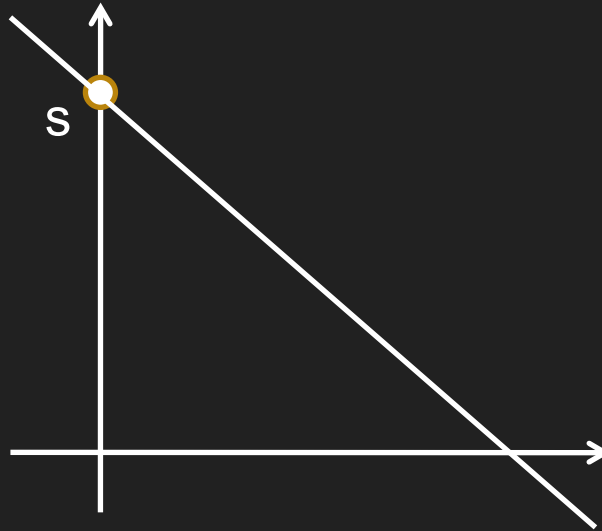
S4

**Non-Communicating Public Servers**

# SSS Intuition

$n = 3$  and  $t = 1$

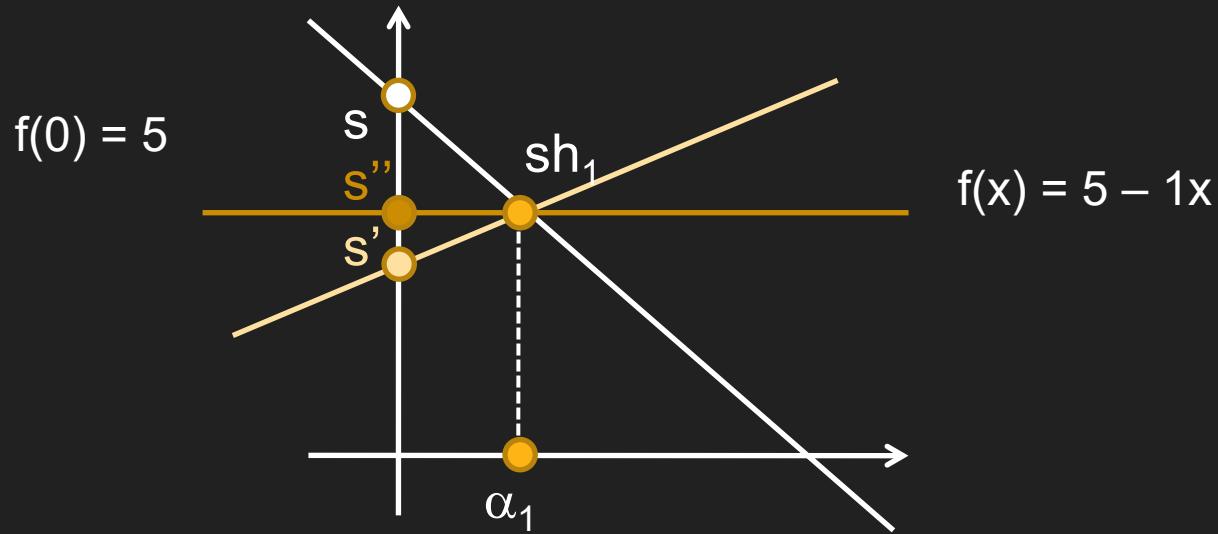
$$f(0) = 5$$



$$f(x) = 5 - 1x$$

# SSS Intuition

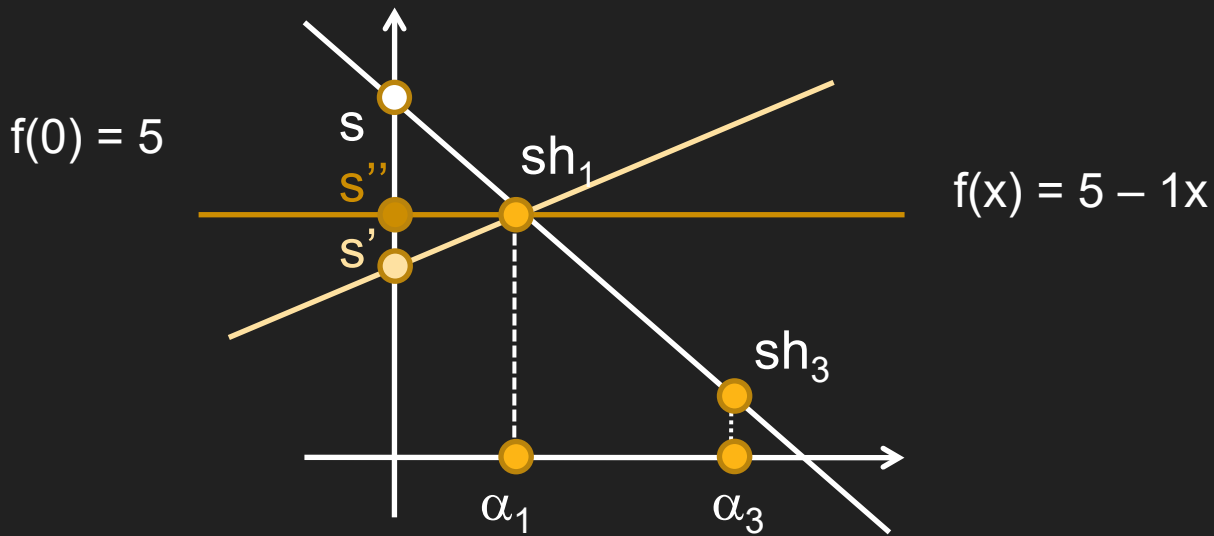
$n = 3$  and  $t = 1$



Only 1 share  $\rightarrow$  all possible straight lines over the field

# SSS Intuition

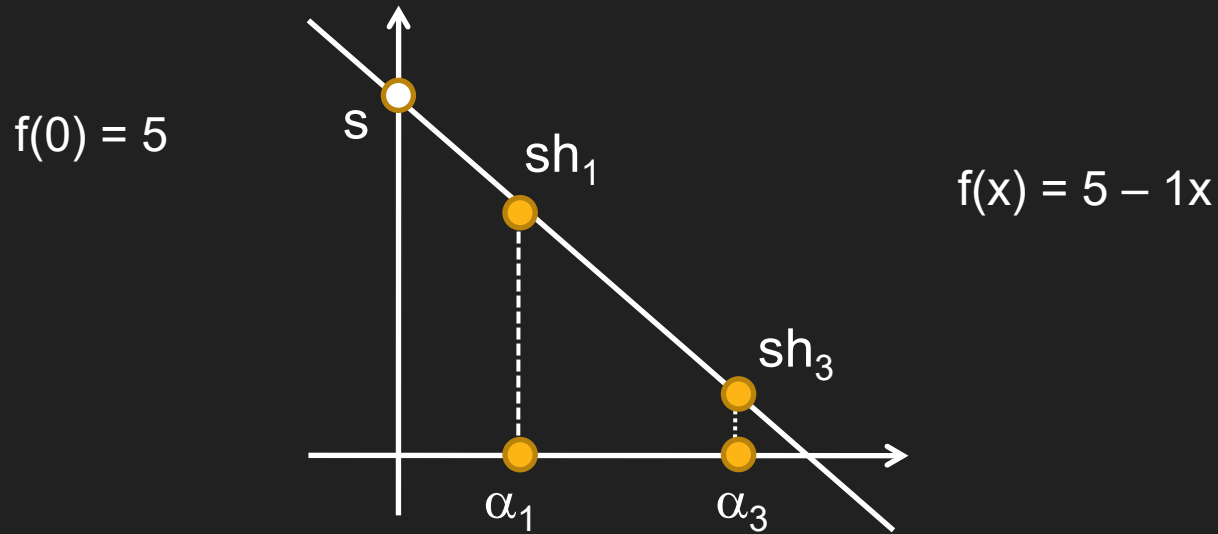
$n = 3$  and  $t = 1$



Only 1 share  $\rightarrow$  all possible straight lines over the field

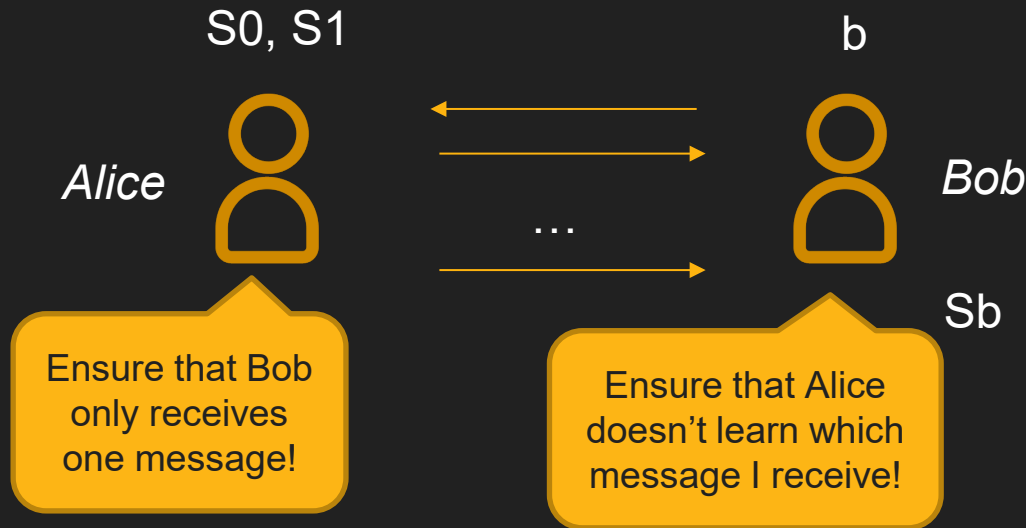
# SSS Intuition

$n = 3$  and  $t = 1$



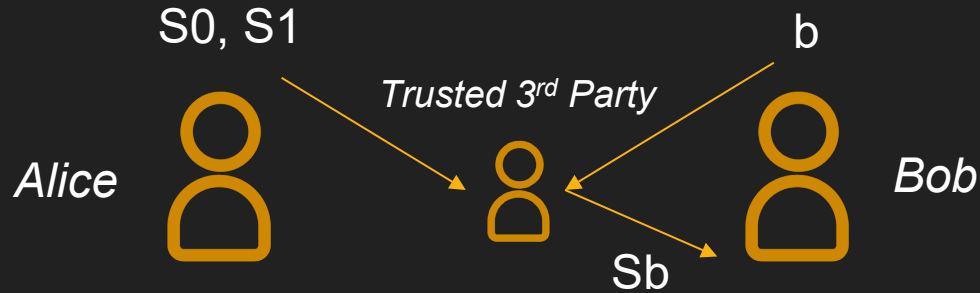
Any set of 2 shares  $\rightarrow$  original straight line and the secret

# Oblivious Transfer (OT)



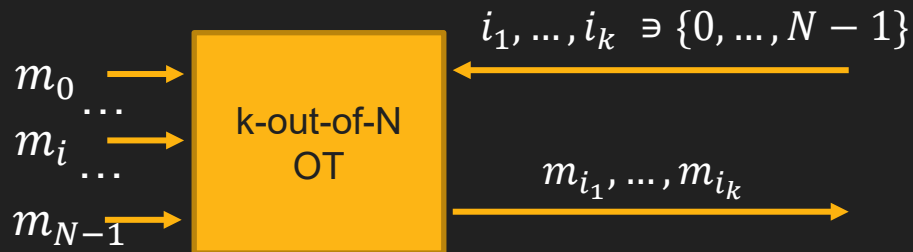
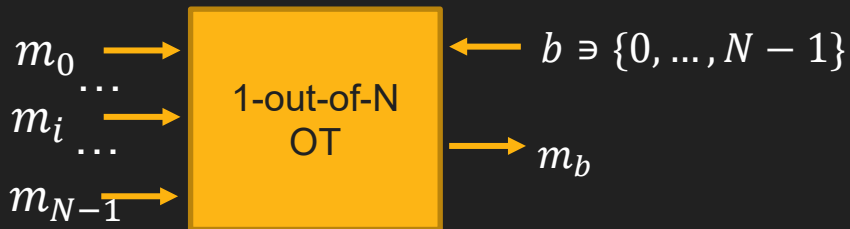
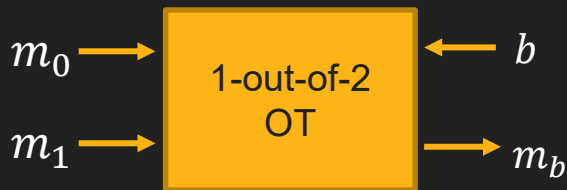
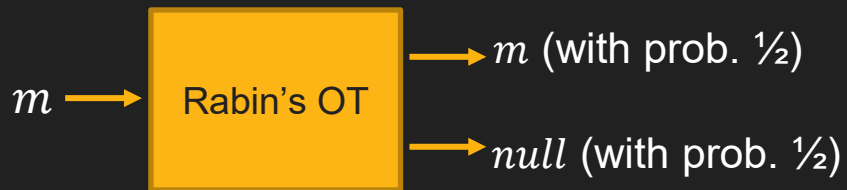
1-out-of-2 Oblivious Transfer

# Oblivious Transfer (OT)

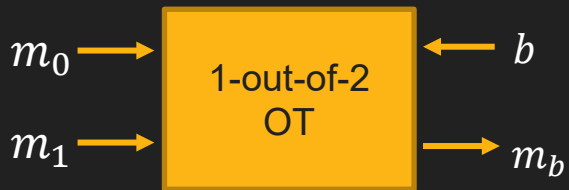


*“Learn no more than what they would if they were interacting with a trusted third party”*

# OT Variants



# Using OT to Compute Operations



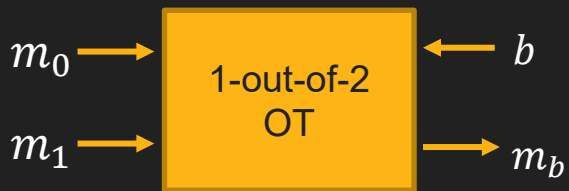
$$m_b = (1 \oplus b)m_0 \oplus bm_1$$

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

$$\text{If } b = 0 \rightarrow m_b = m_0 \oplus 0 = m_0$$

$$\text{If } b = 1 \rightarrow m_b = 0 \oplus m_1 = m_1$$

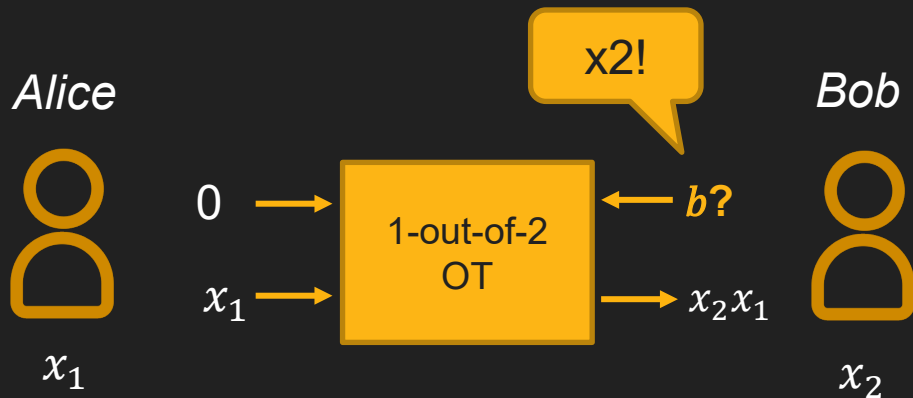
# Using OT to Compute Operations



$$m_b = (1 \oplus b)m_0 \oplus bm_1$$

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

$$m_b = (1 \oplus x_2)0 \oplus x_2x_1 = x_2x_1$$



**Compute AND operator!**

# How is MPC Deployed in Practice?

Input Parties



...



Computing Parties

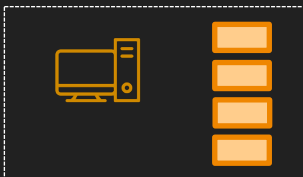


Result Parties

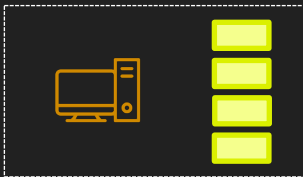


# How is MPC Deployed in Practice?

Input Parties



...



Computing Parties

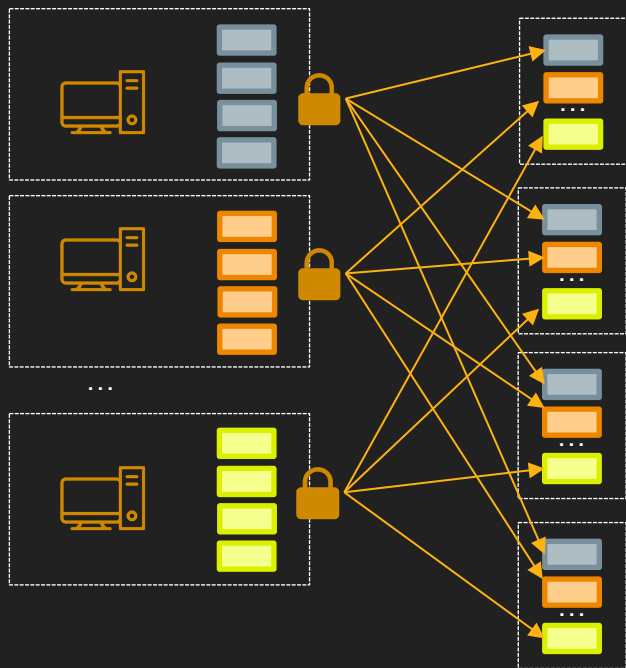


Result Parties

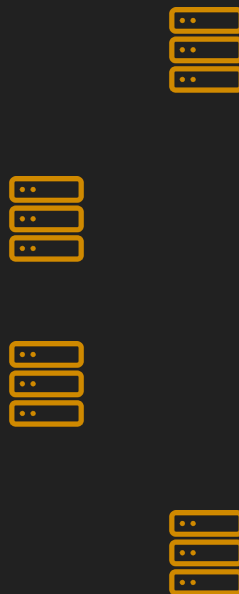


# How is MPC Deployed in Practice?

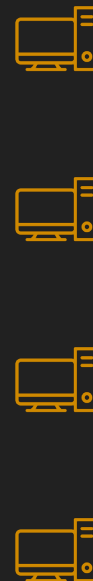
Input Parties



Computing Parties



Result Parties

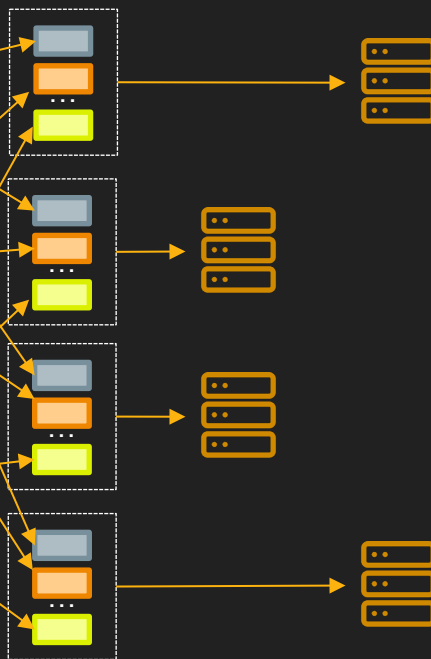


# How is MPC Deployed in Practice?

Input Parties



Computing Parties



Result Parties

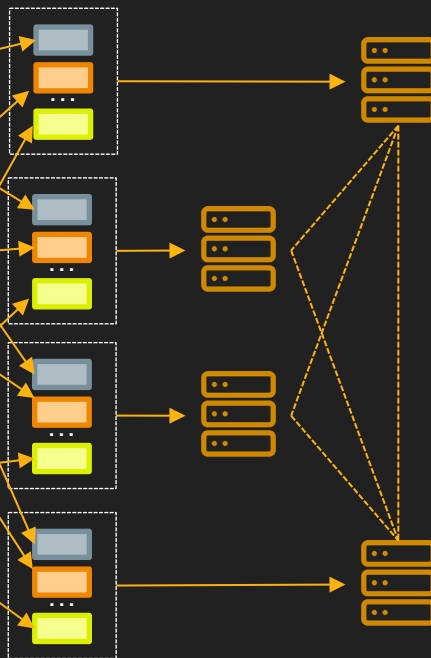


# How is MPC Deployed in Practice?

Input Parties



Computing Parties



Result Parties

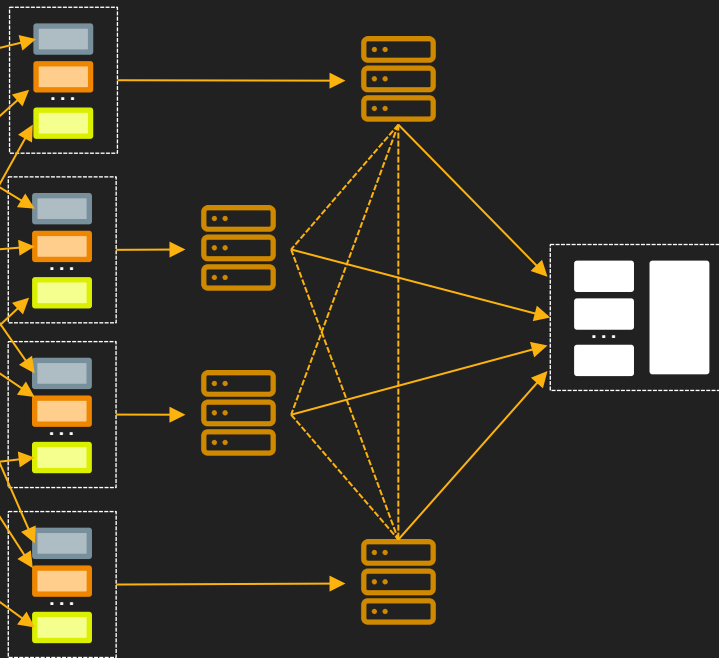


# How is MPC Deployed in Practice?

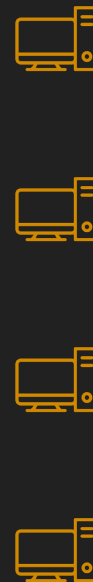
Input Parties



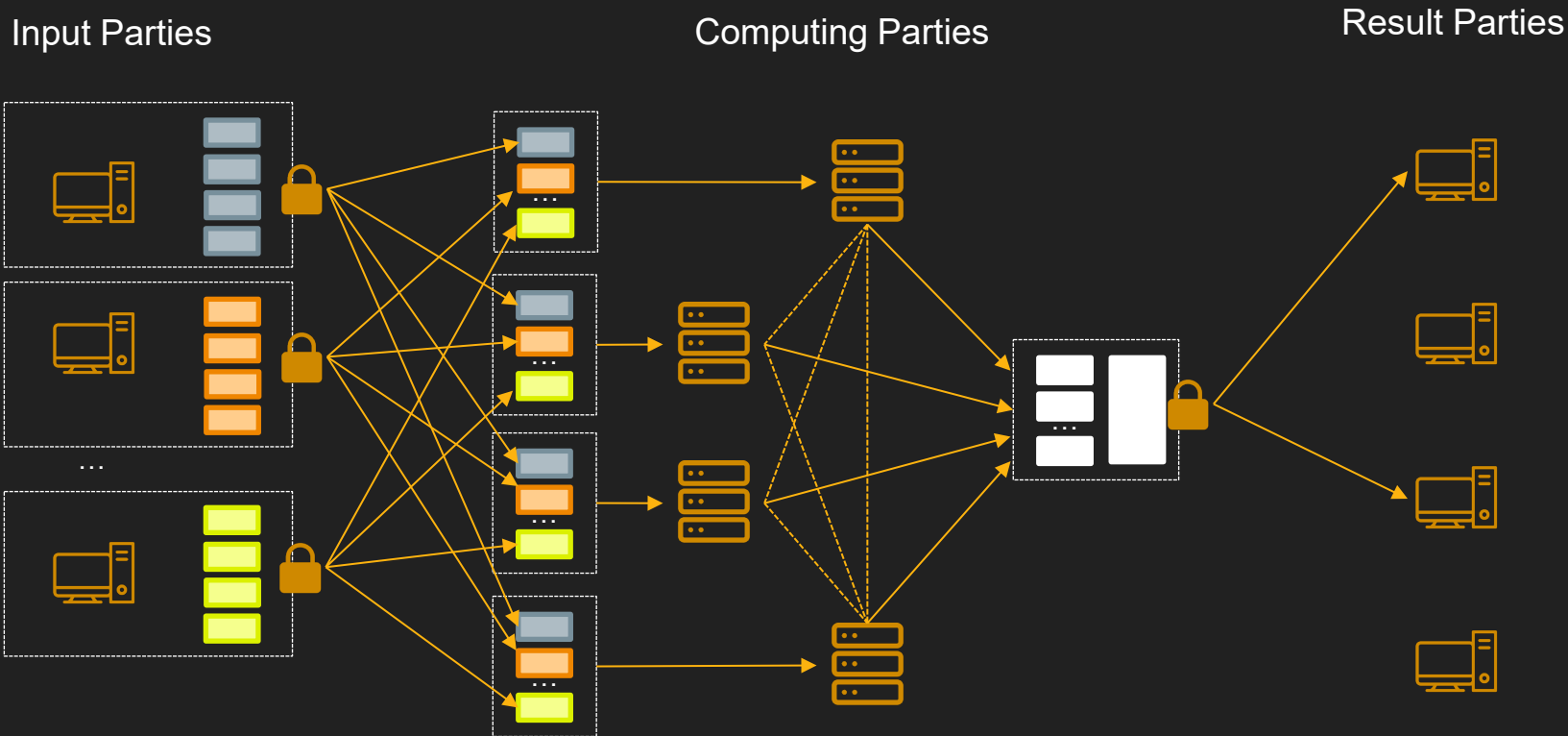
Computing Parties



Result Parties

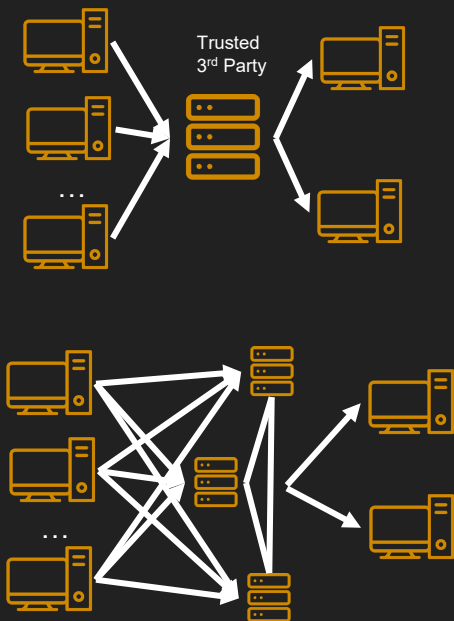


# How is MPC Deployed in Practice?

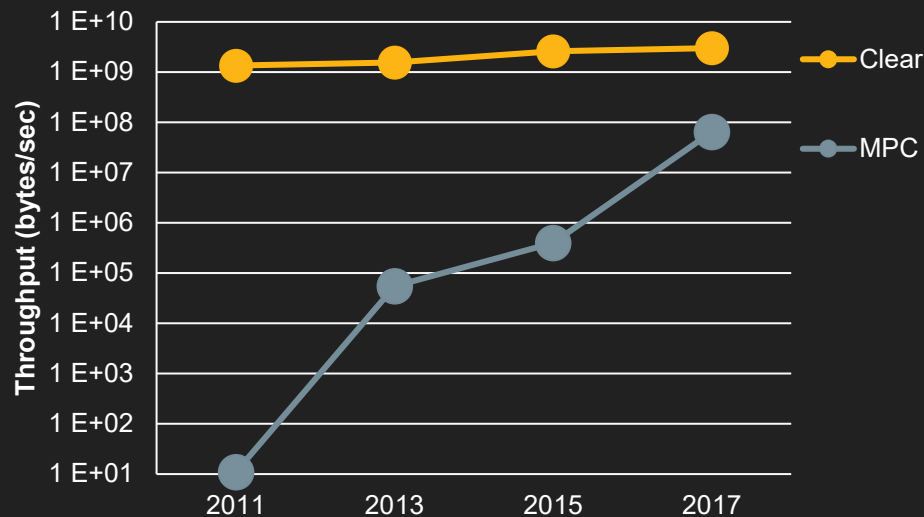


# MPC Challenges

- Communication overheads!



- High Computational Cost!

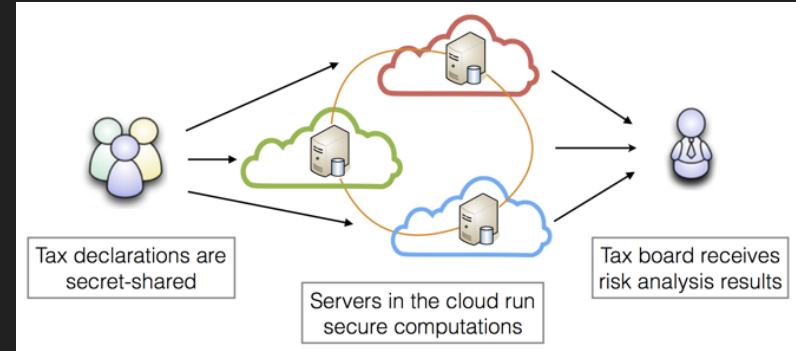


Source: Mayank Varia, Boston University, "A Survey of MPC Offerings"

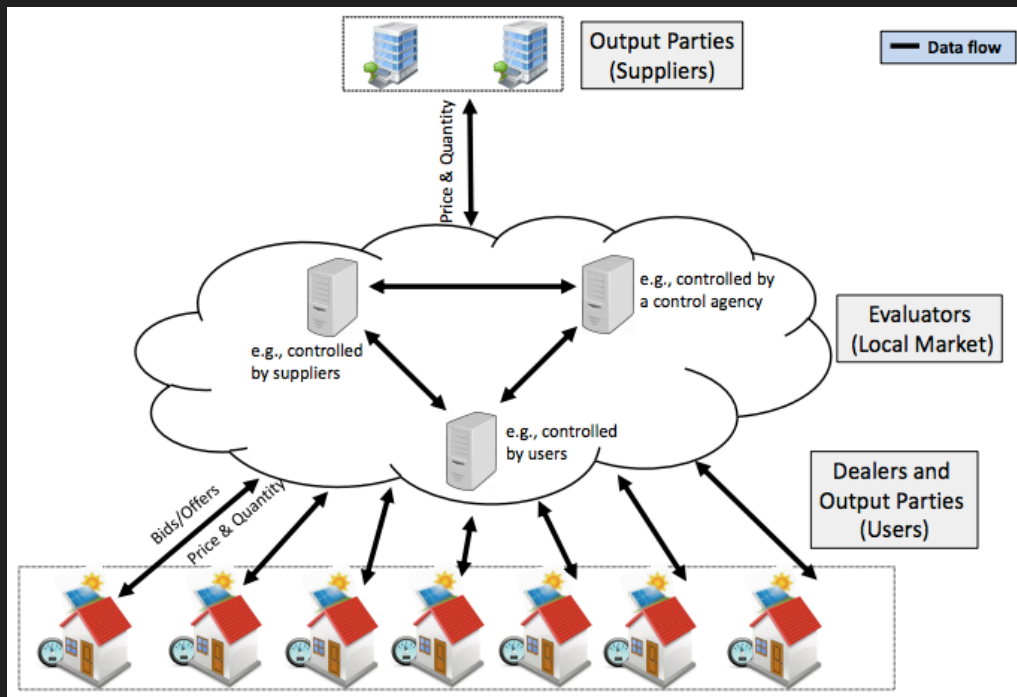
MPC in Use!

# Tax Fraud

- ITL economic benchmarks
  - Collection of Estonian companies
  - Aggregate economic indicators: profit, # employees, salaries
- VAT tax revenue
  - Worked with Estonian Tax and Customs Board
  - Test if Company A's VAT credit == Company B's VAT reported



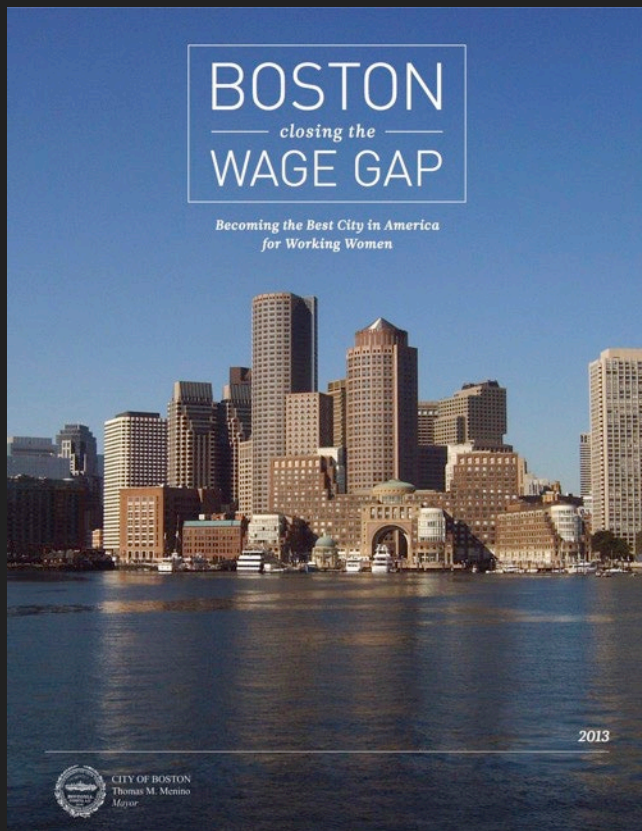
# Electricity Markets



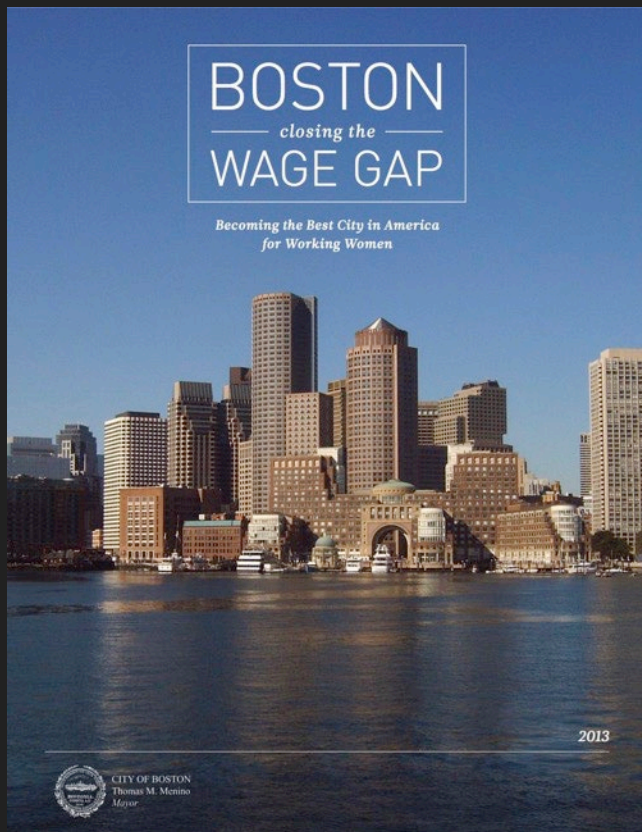
Energy trading with smart meters

- Handles 2500 bids in ~5 min
- Auction run every 30 min

# Public Good (Wage Disparity)



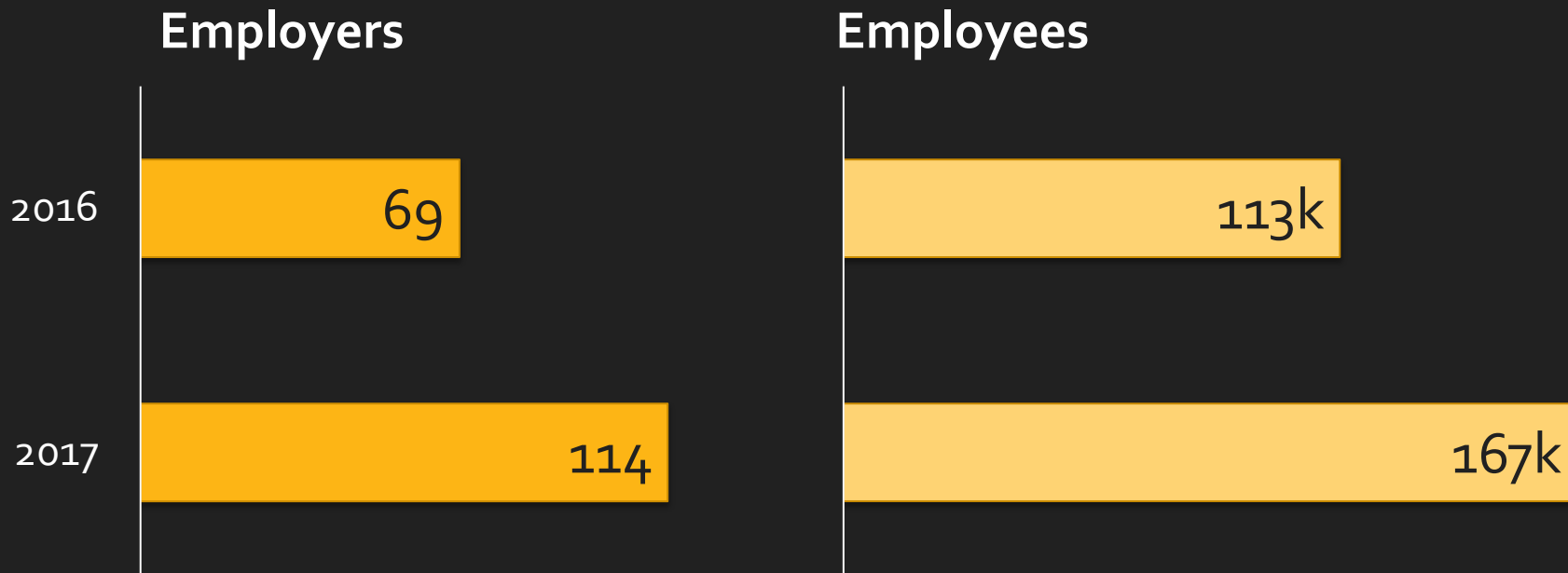
# Public Good (Wage Disparity)



## Goal 3: Evaluating Success

Employers agree to contribute data to a report *compiled by a third party* on the Compact's success to date. *Employer-level data would not be identified* in the report.

# Public Good (Wage Disparity)



# “Student Right to Know Before You Go” Bill

- Empower prospective college students to make more informed decisions
- Measure annual earnings and accumulated debt of recent graduates

115TH CONGRESS  
1ST SESSION






S. \_\_\_\_\_

IN THE SENATE OF THE UNITED STATES

Mr. WYDEN (for himself, Mr. RUBIO, and Mr. WARNER) introduced the following bill; which was read twice and referred to the Committee on

“in designing, establishing, and maintaining the higher education data system, ... the Commissioner shall **use secure multiparty computation technologies**”

# “Student Right to Know Before You Go” Bill

MAY 9, 2013		Earlier Version — Introduced This activity took place on a related bill, <a href="#">S. 915 (113th)</a> .
MAY 5, 2015		Earlier Version — Introduced This activity took place on a related bill, <a href="#">S. 1195 (114th)</a> .
NOV 29, 2017		Earlier Version — Introduced This activity took place on a related bill, <a href="#">S. 2169 (115th)</a> .
MAR 6, 2019		<b>Introduced</b> Bills and resolutions are referred to committees which debate the bill before possibly sending it on to the whole chamber. <a href="#">Read Text »</a>
MAR 30, 2022		Reintroduced Bill — Introduced This activity took place on a related bill, <a href="#">S. 3952</a> .

# Conclusions

- MPC provides a mechanism to promote collaboration
- Goal: prevent other parties from learning about shared data
- MPC maintains data usability and “privacy” (more like “confidentiality”)
  - Not the differential privacy definition of privacy! Attacks are still possible...
- High computational and communication costs!
- Assumptions about maliciousness of participants

# Group Activity

- Think about your group project
- Do you need to collaborate to learn something?
  - What data would you need to share?
  - Who would be the collaboration parties?
  - What would you want to learn?
  - What would you want to protect?