

Practical Federated Learning without a Server

Rishi Sharma

EPFL

In collaboration with Akash Dhasade, Anne-Marie Kermarrec,
Erick Lavoie, Johan Pouwelse, and Martijn de Vos.

March 31st, 2025

AI Tools



ChatGPT



Gemini



Claude



Copilot



Perplexity



Grok

**Thinking,
Web Search,
Deep Research,
Agents.**



~50 TBs

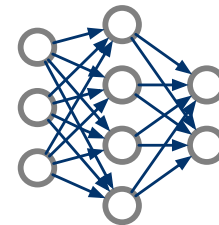
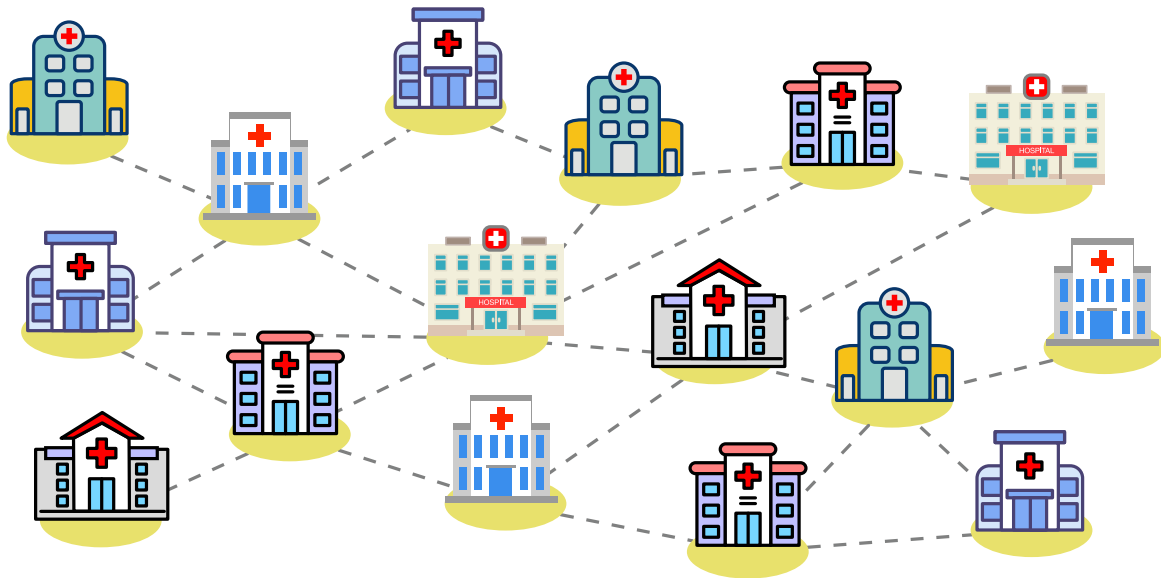
Public data

~200 PBs

Siloed data

~1 ZBs
~ 1M PBs

Collaborative Machine Learning



Shared model architecture



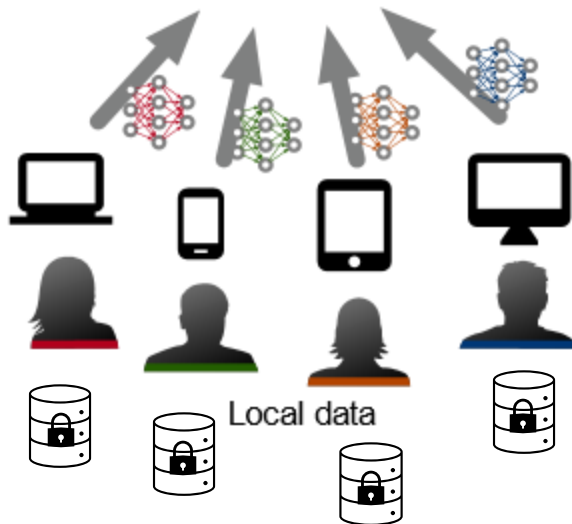
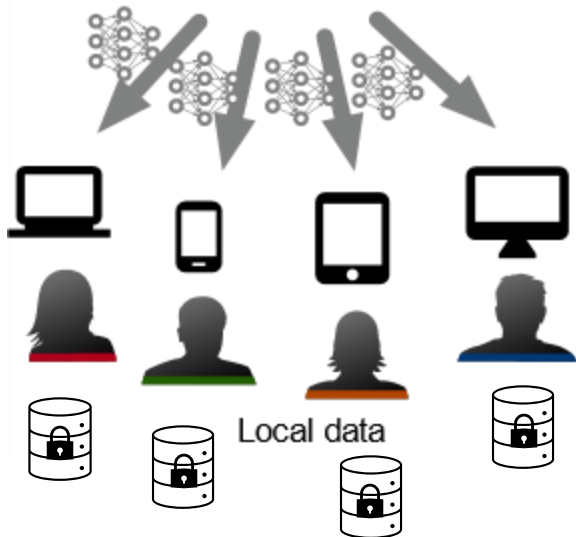
Private data

Federated Learning



Global model

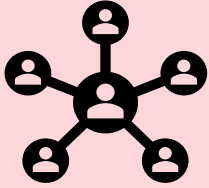
Local updates



Train locally

Upload

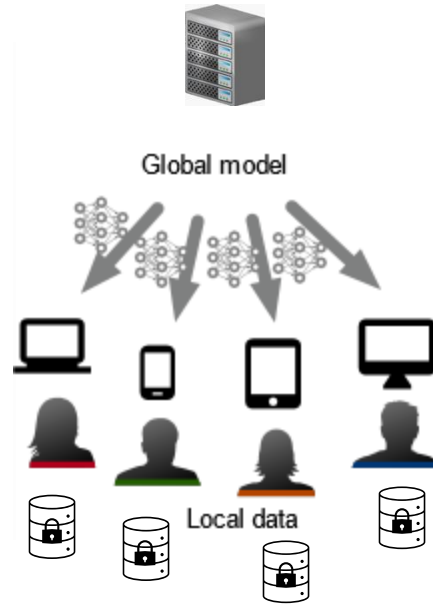
Aggregate
&
Send back



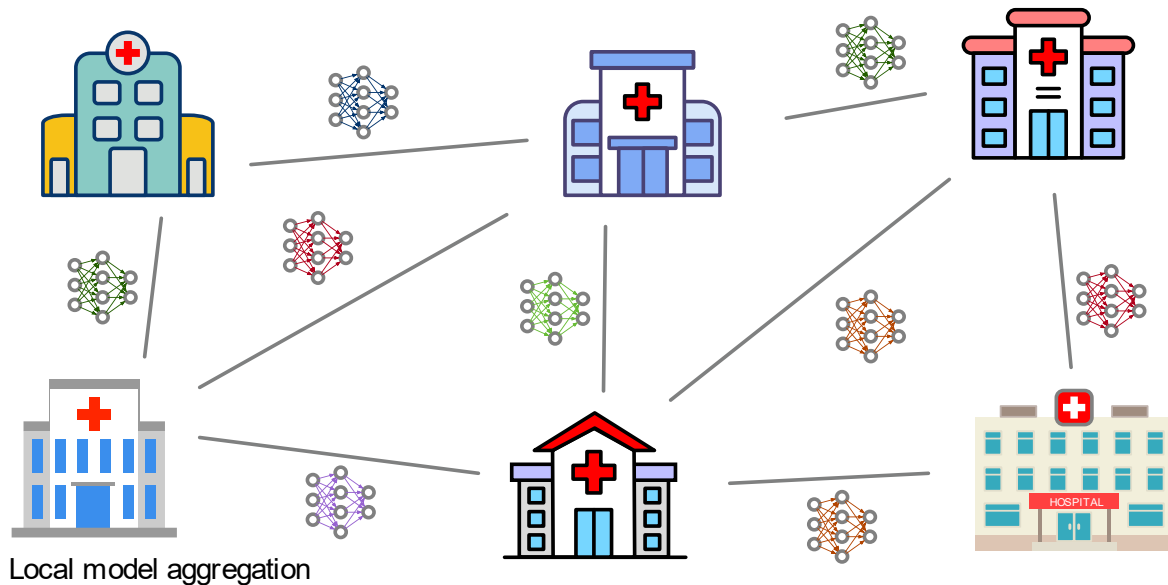
Global view



Infra. cost



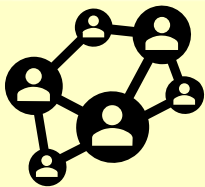
Decentralized Learning



Train locally

Multicast

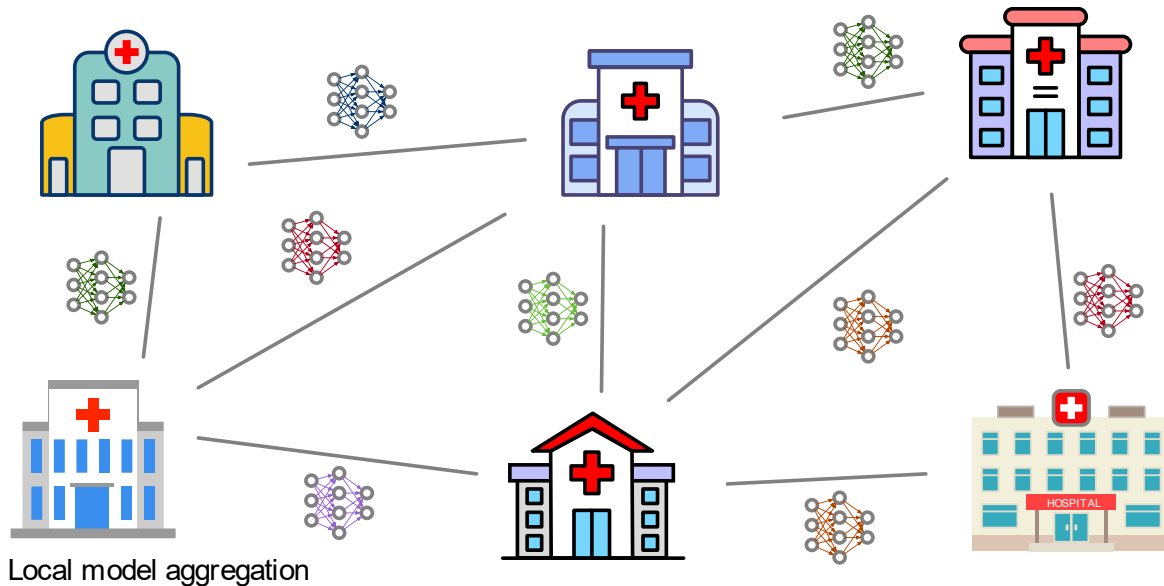
Receive & Aggregate



Scalable

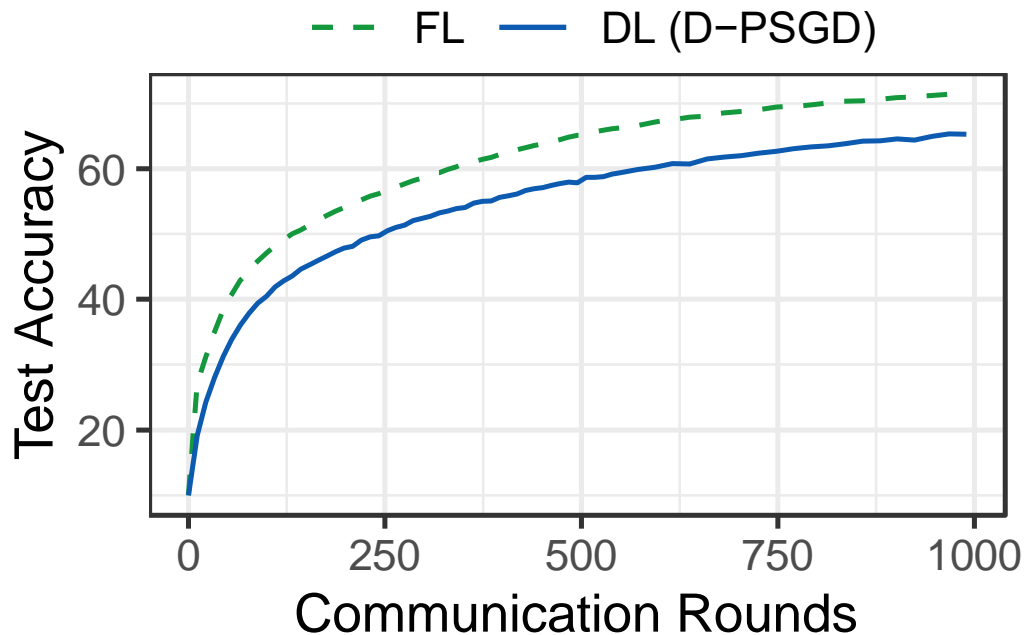


Low Infra. Cost



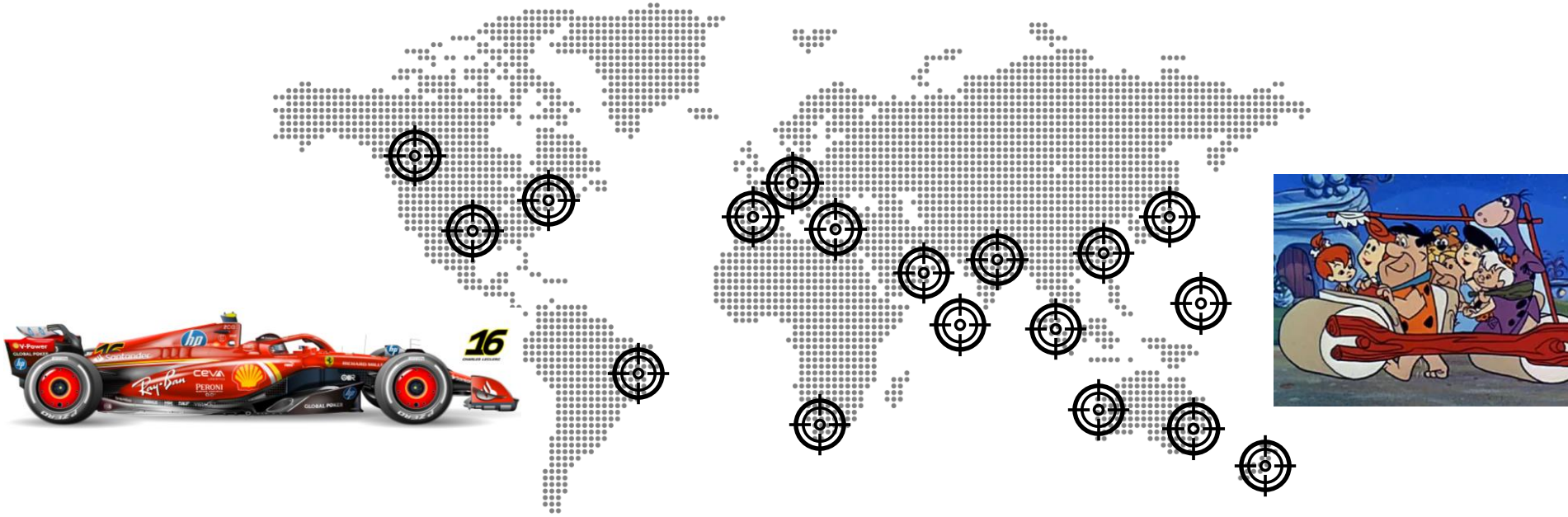
What FL Brings?

1000 nodes; CIFAR-10; non-IID;



Global aggregation = Better convergence

Best of Both Worlds?





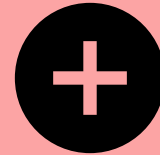
Select Team

**Decentralized
peer sampling**



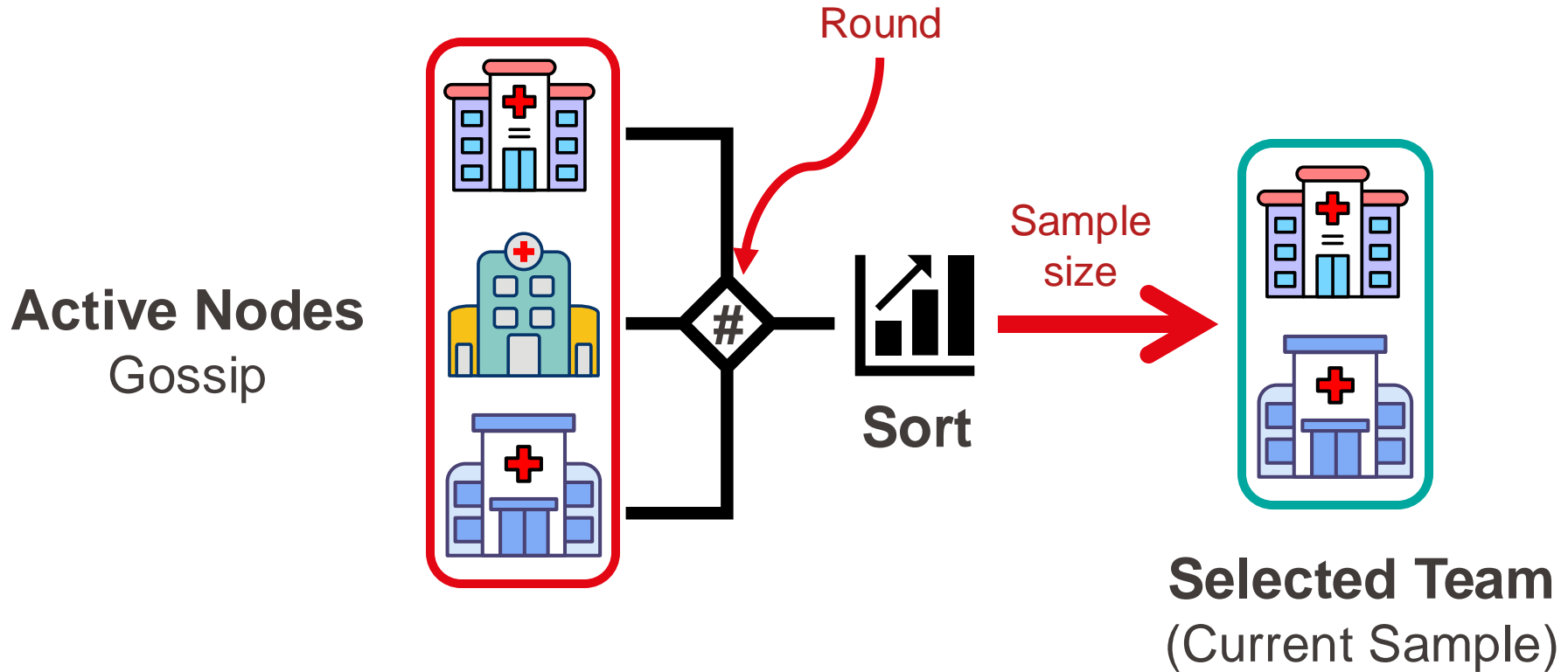
Aggregator

**Aggregator
selection**



Aggregate

**Aggregate &
Push**

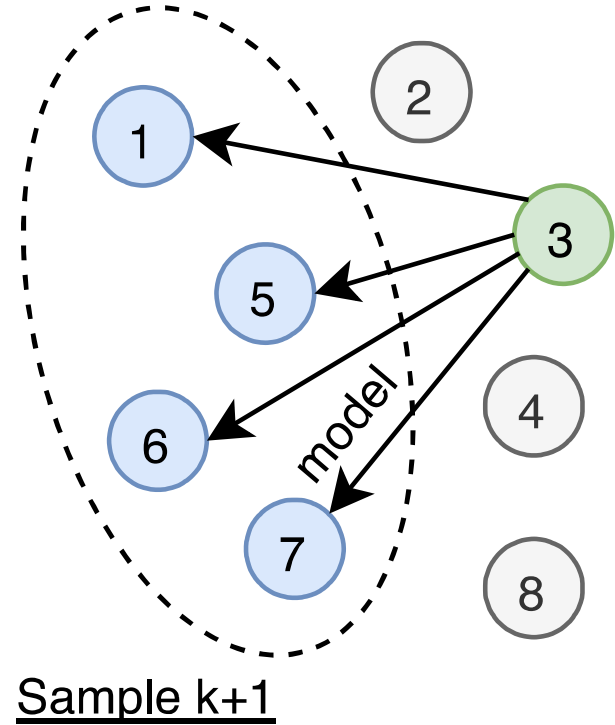
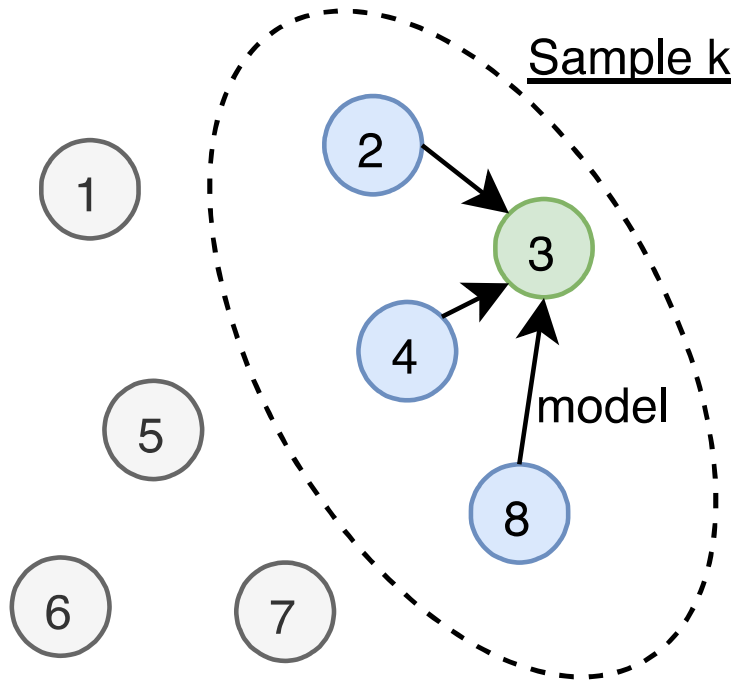


Aggregator Selection



**Highest
bandwidth
in the
sample**

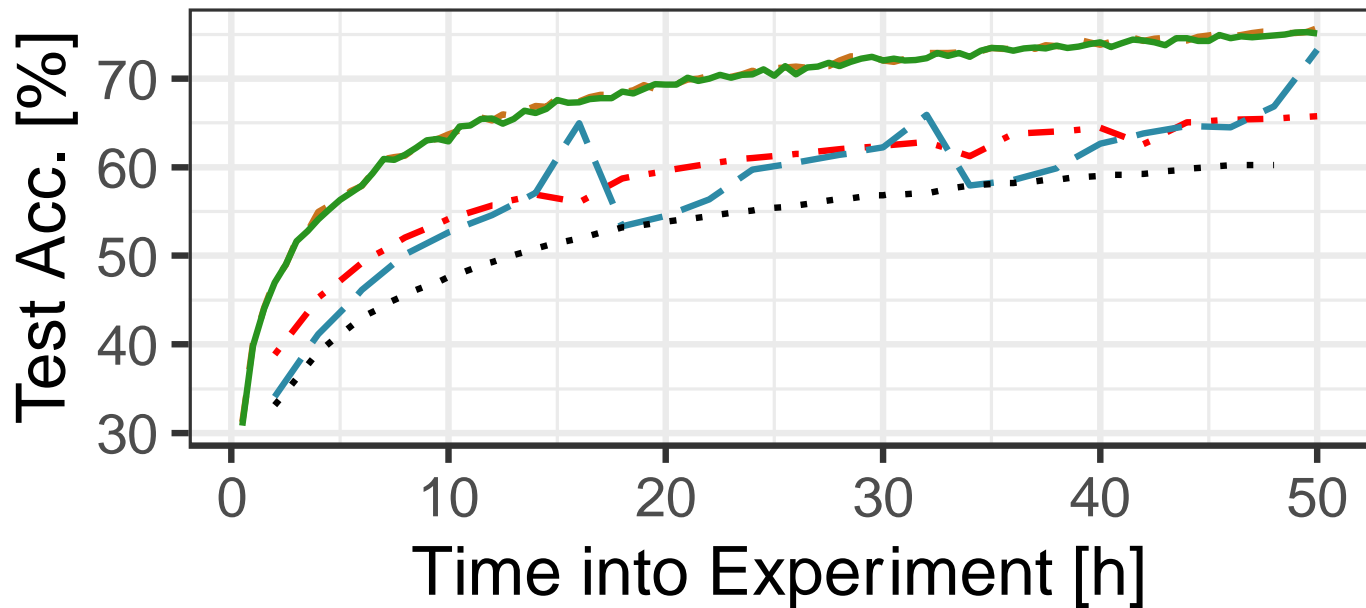
Aggregate & Push



Results: Convergence

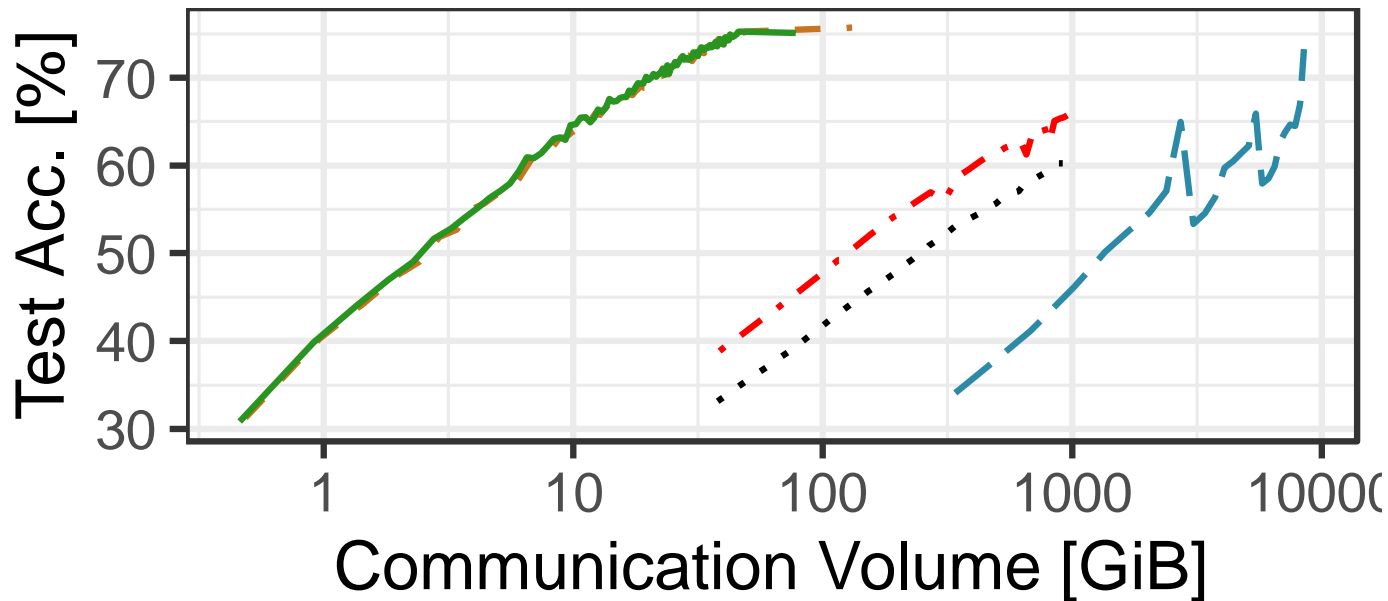
1000 nodes; CIFAR-10; IID;

— Plexus — FL ···· GL - - - D-PSGD (OP) - - - D-PSGD (k-reg)



FL in Decentralized!

— Plexus — FL ···· GL -.- D-PSGD (OP) - - - D-PSGD (k-reg)



FL in Decentralized!

More results in
the paper...

Plexus – Practical Federated Learning without a Server

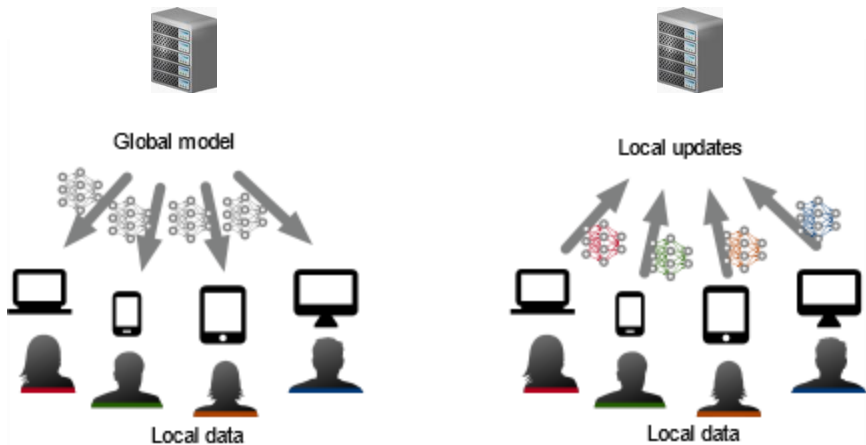
- Best of both worlds
- Contributions:
 - Emulating the server in decentralized settings
- Future work:
 - Nodes leaving and joining
 - Load balancing
 - Asynchronous learning



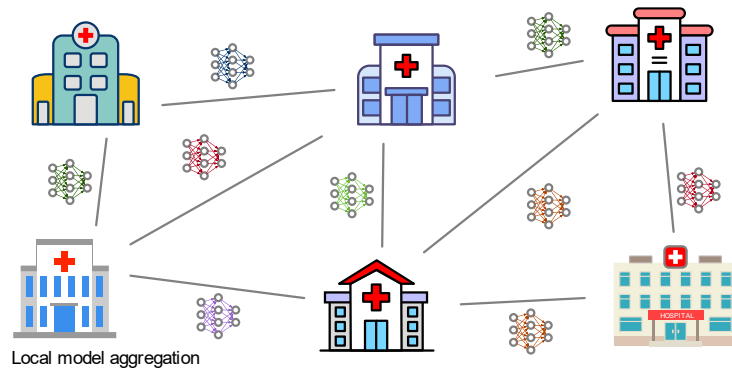
rishi.sharma@epfl.ch



rishi-s8.github.io



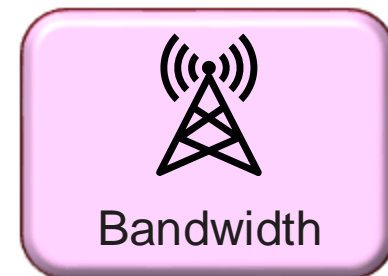
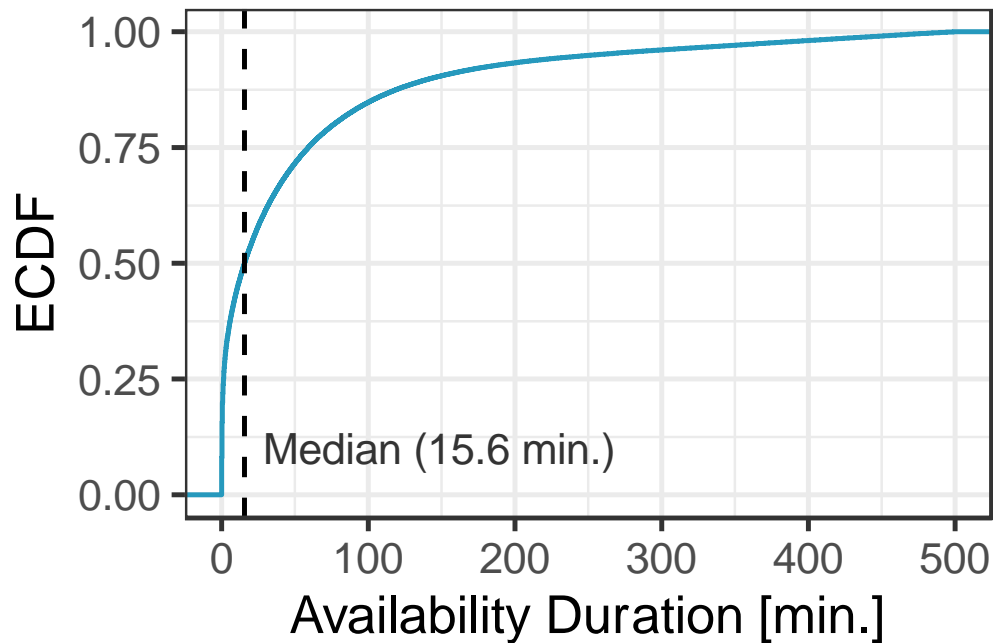
Federated Learning (FL)



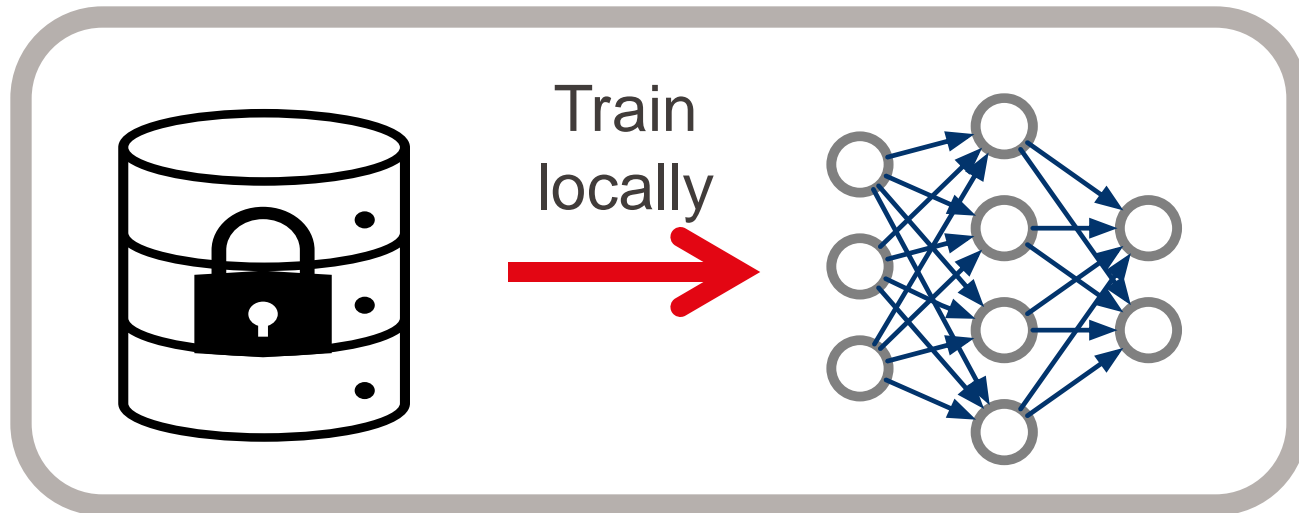
Decentralized Learning (DL)



Private data



Decentralized Learning



Train locally

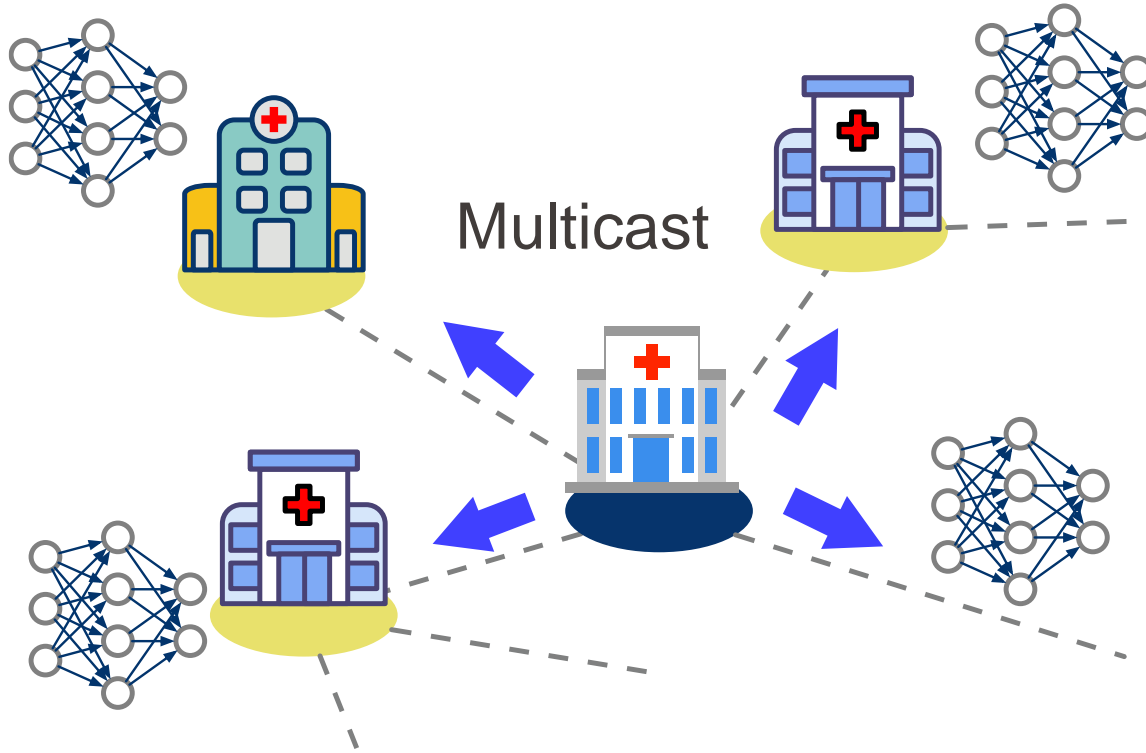
↓

Multicast

↓

Receive & Aggregate

Decentralized Learning



Train
locally

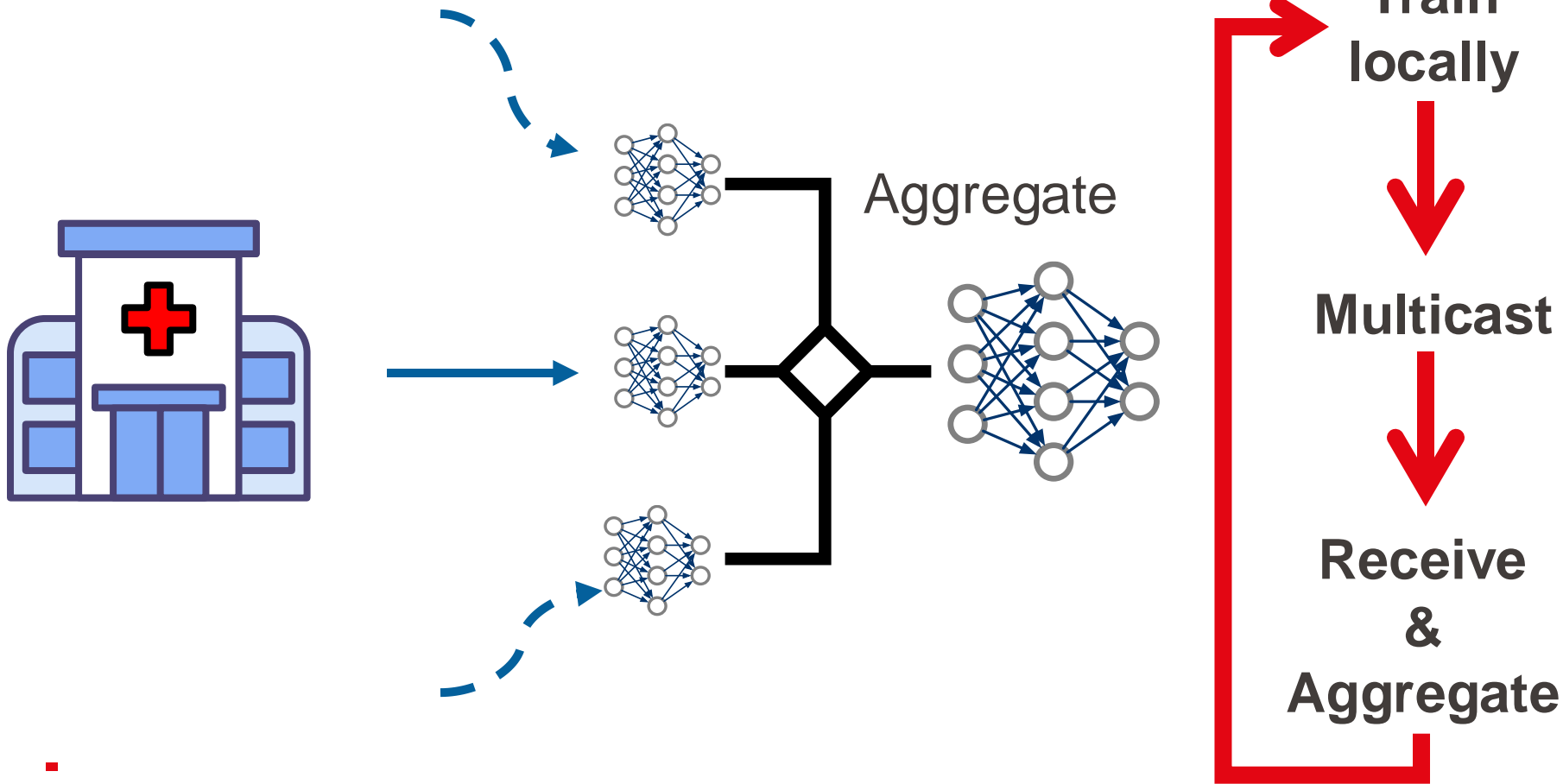


Multicast



Receive
&
Aggregate

Decentralized Learning



Communication Stragglers

	Cape Town	Tokyo	Mumbai	Sydney	Stockholm	Milan	Bahrain	Sao Paulo	Ohio	Oregon	Bandwidth (Mbps)
Cape Town		26.1	36.0	20.8	59.8	67.1	33.6	27.1	43.6	35.9	
Tokyo	354.0		89.3	112.1	42.1	48.1	66.8	39.3	85.8	108.8	
Mumbai	272.0	127.2		75.9	81.3	103.2	336.3	30.8	53.3	48.5	
Sydney	410.4	102.3	146.8		32.0	42.4	59.6	31.2	57.0	80.8	
Stockholm	179.7	241.2	138.9	295.7		404.6	81.8	48.2	94.7	67.6	
Milan	162.4	214.8	110.8	238.8	30.2		105.7	49.4	104.9	70.1	
Bahrain	287.0	164.3	36.4	179.2	137.9	108.2		29.9	49.4	38.7	
Sao Paulo	340.5	256.6	305.6	310.5	214.9	211.9	320.0		92.3	60.5	
Ohio	237.0	131.8	197.3	187.9	120.0	109.2	212.7	121.9		105.0	
Oregon	276.6	96.7	215.8	139.7	162.0	157.8	251.4	178.3	55.2		

Round trip time (ms)