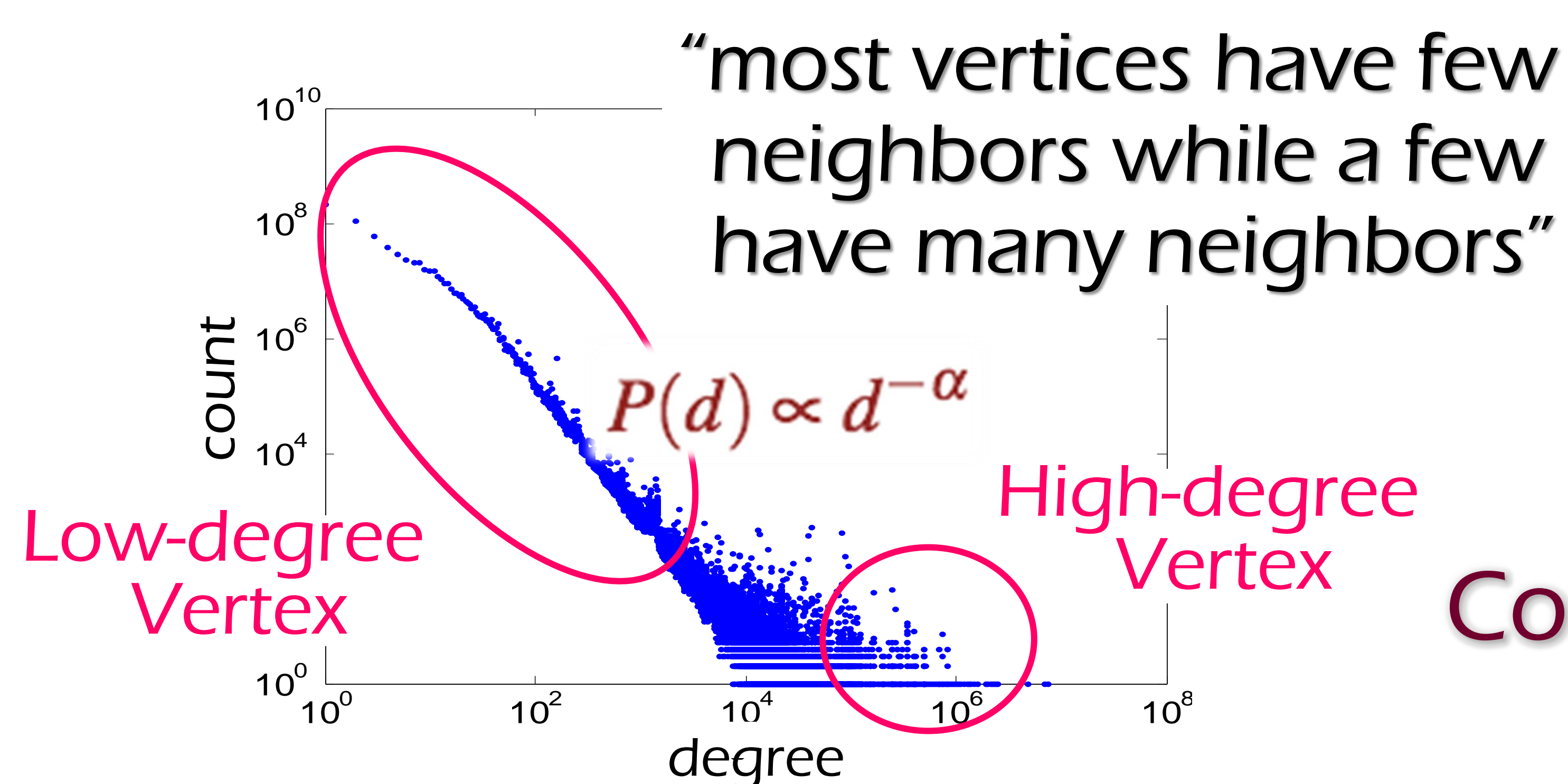




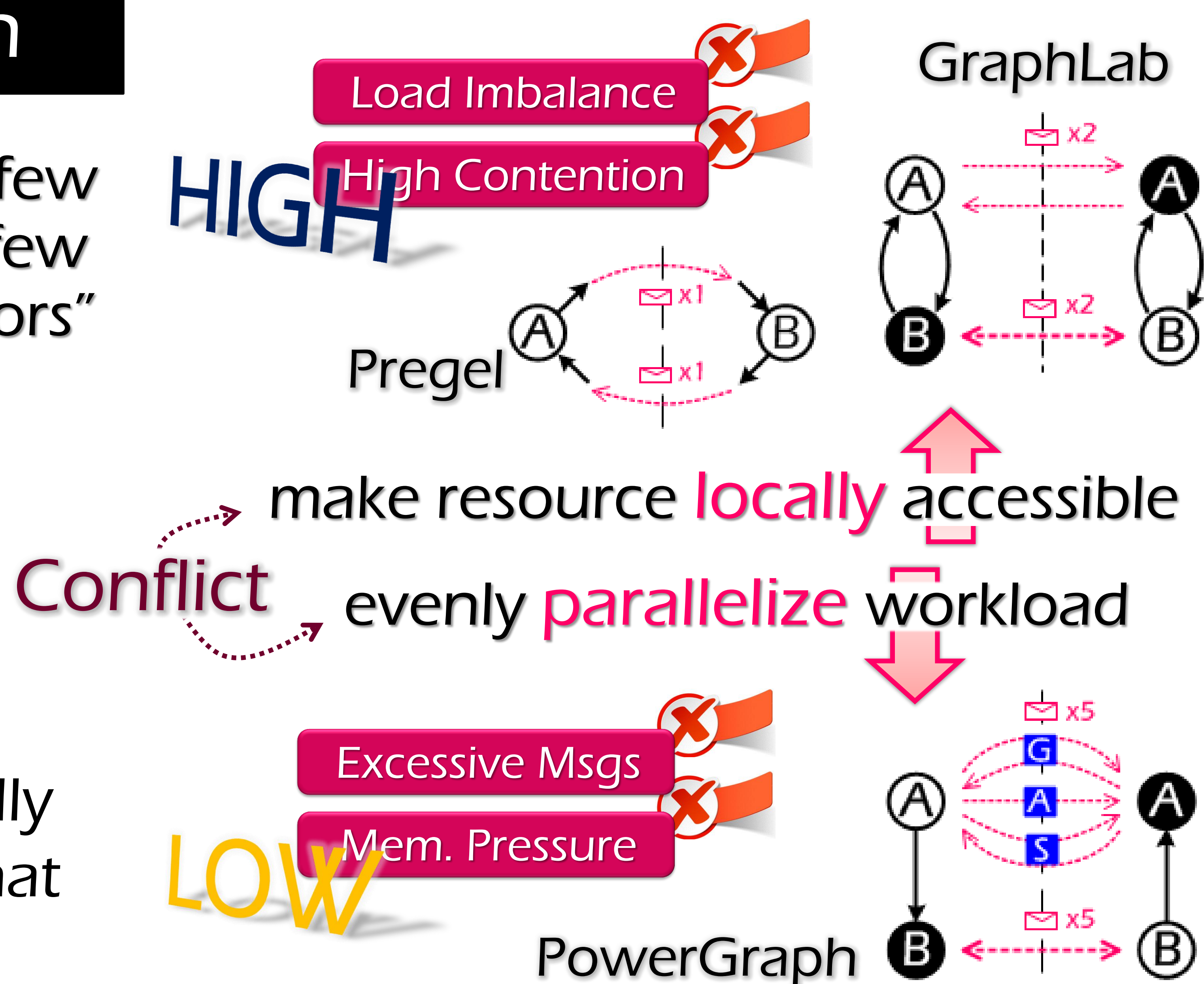
PowerLyra: Differentiated Graph Computation and Partitioning on Skewed Graphs

Rong Chen, Jiaxin Shi, Yanzhe Chen and Haibo Chen

Background and Motivation



Existing graph-parallel systems usually use a “**ONE size fits ALL**” design that uniformly processes all vertices

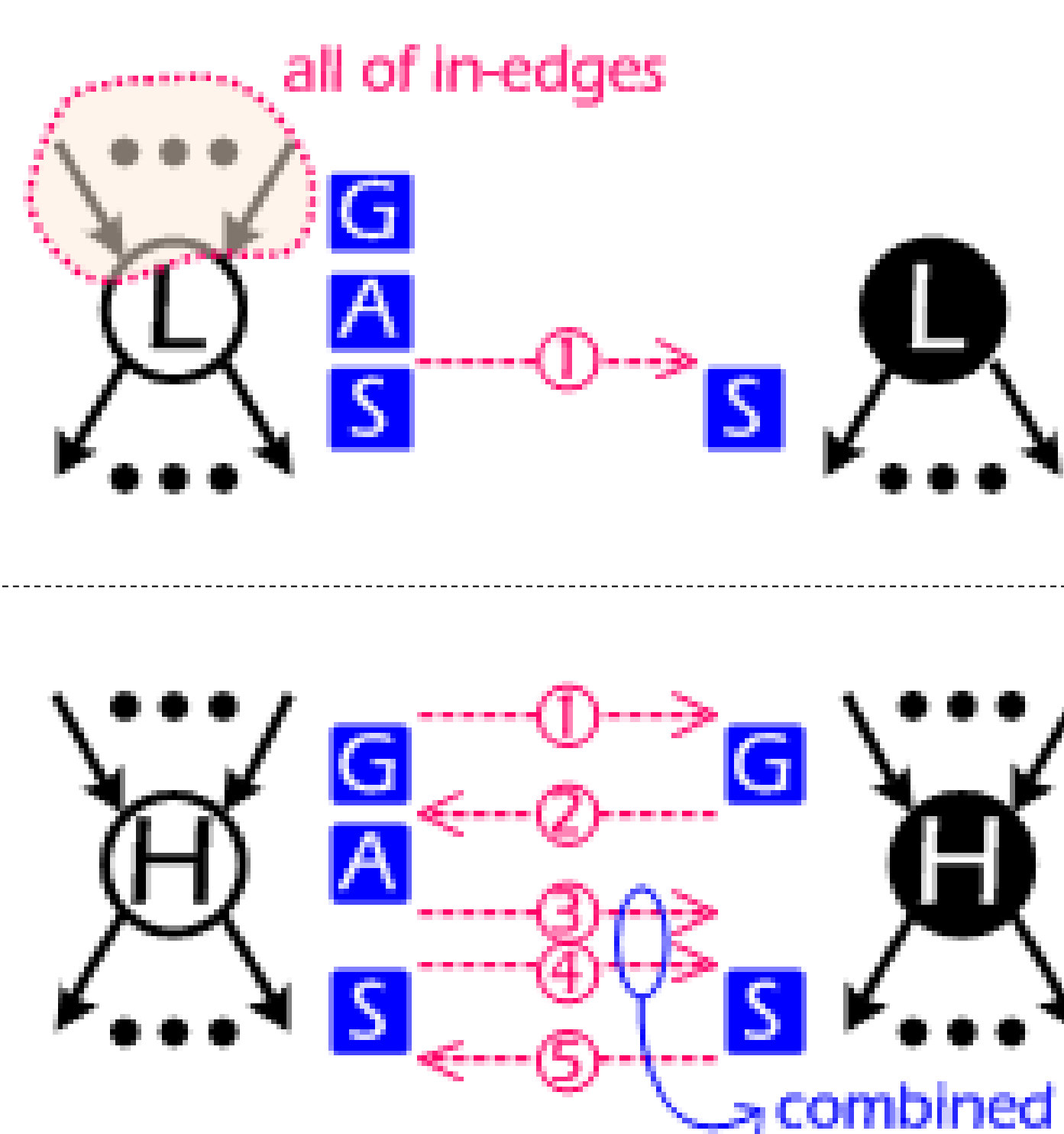


Hybrid

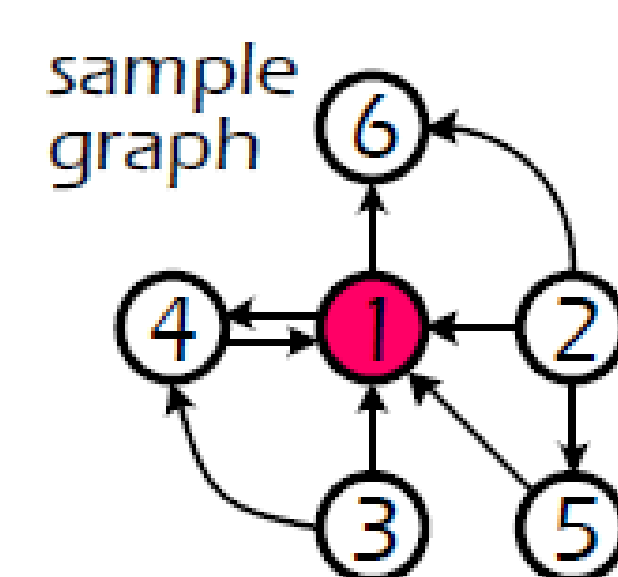
Low-degree Vertex
Locality

High-degree Vertex
Parallelism

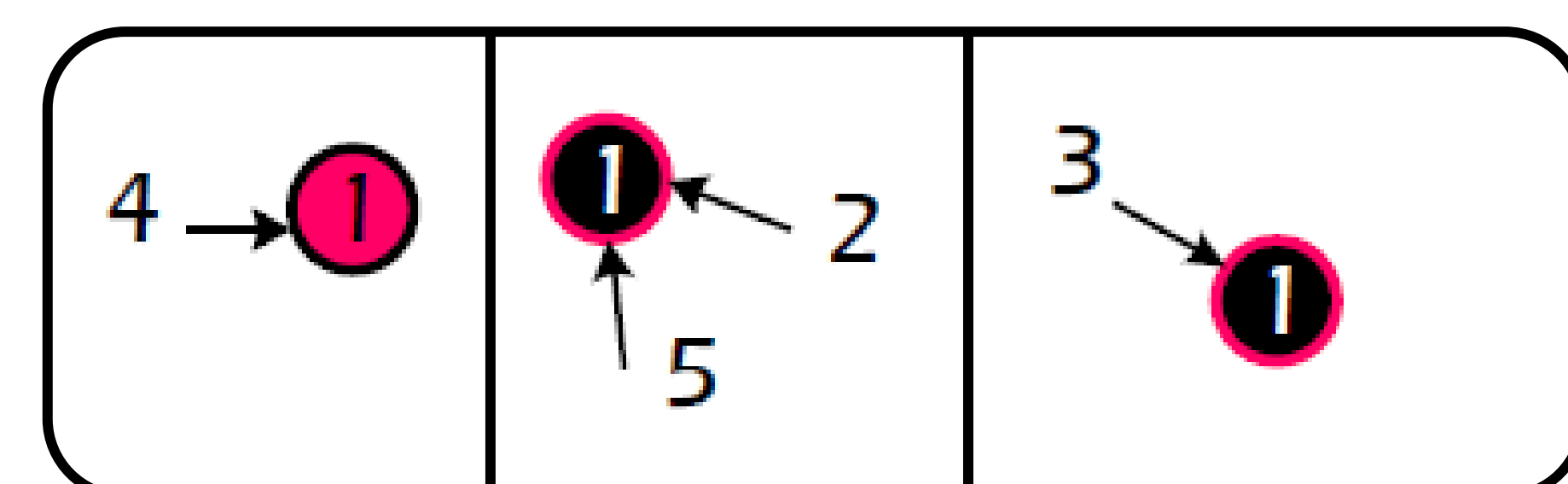
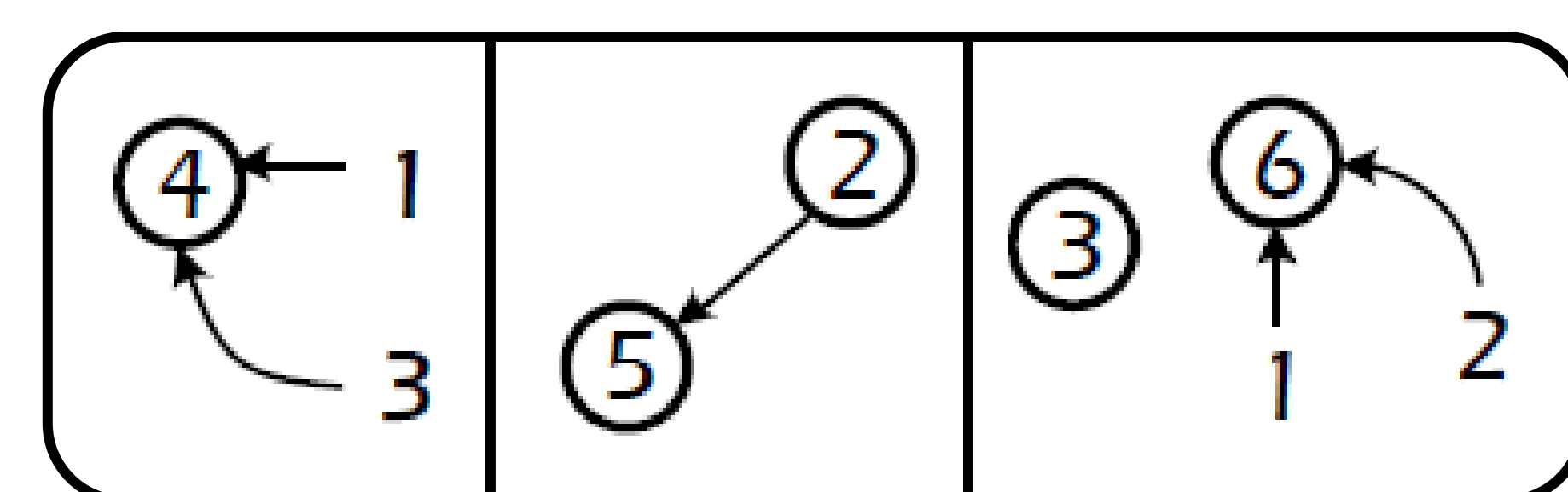
Computation



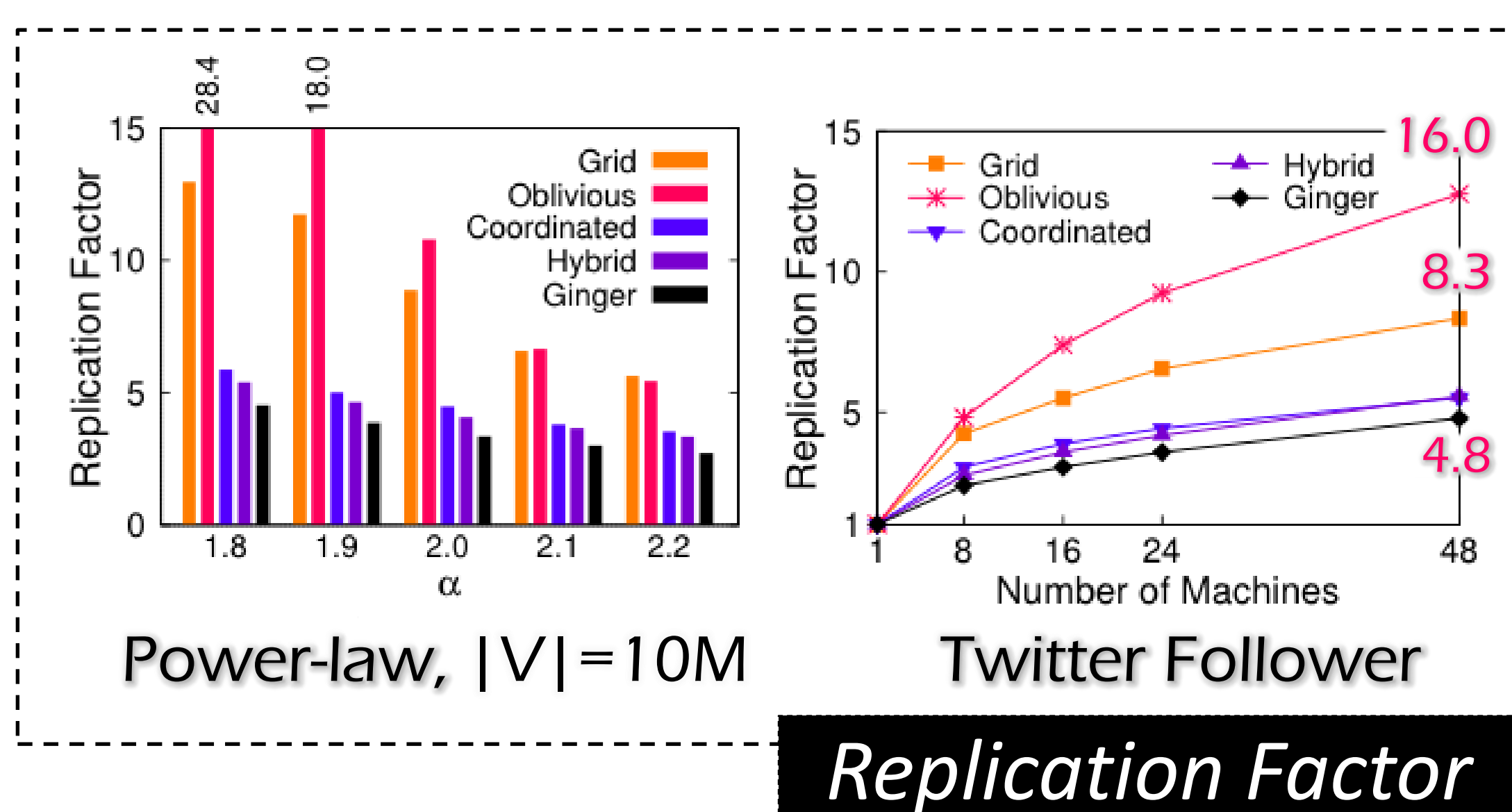
Graph Partitioning



- ② Low-master
- ② Low-mirror
- ① High-master
- ① High-mirror



Evaluation (Code and Instruction: <http://ipads.se.sjtu.edu.cn/projects/powerlyra.html>)



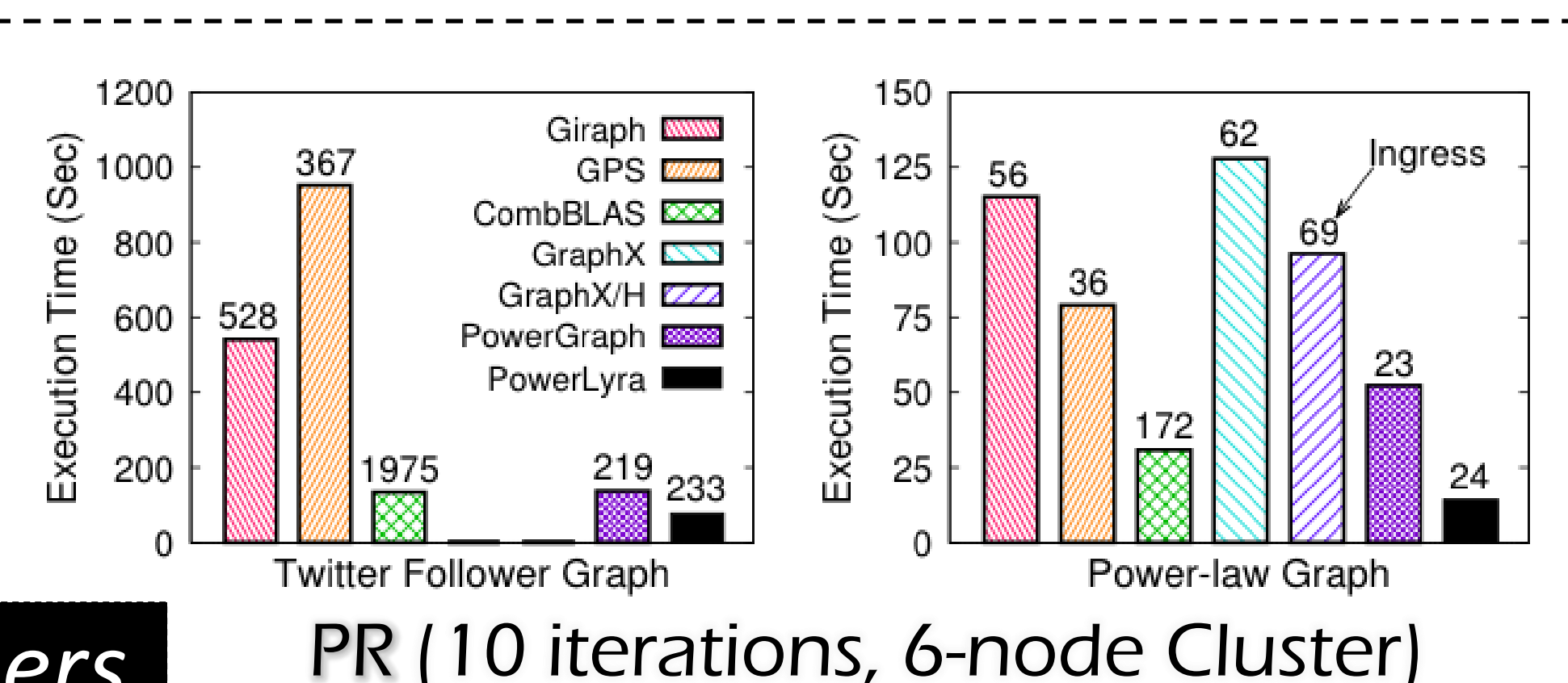
Netflix Movie Recommendation [52]				Replication Factor	
$ V $	$ E $	Vertex Data	Edge Data	Grid	Hybrid
0.5M	99M	$8d + 13$	16	12.3	2.6

ALS [52]	$d=5$	$d=20$	$d=50$	$d=100$
PowerGraph	10 / 33	11 / 144	16 / 732	Failed
PowerLyra	13 / 23	13 / 51	14 / 177	15 / 614

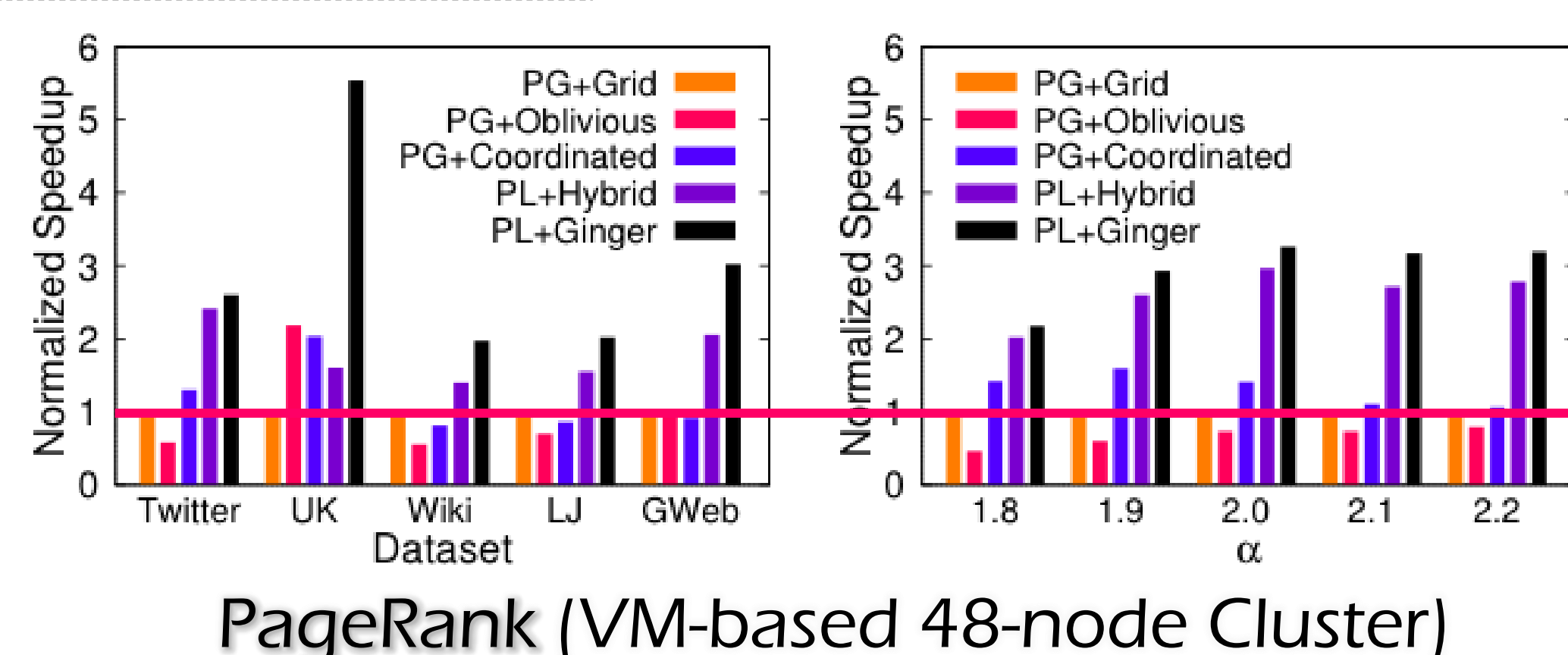
SGD [42]	$d=5$	$d=20$	$d=50$	$d=100$
PowerGraph	15 / 35	17 / 48	21 / 73	28 / 115
PowerLyra	16 / 26	19 / 33	19 / 43	20 / 59

MLDM Institute of Parallel And Distributed Systems

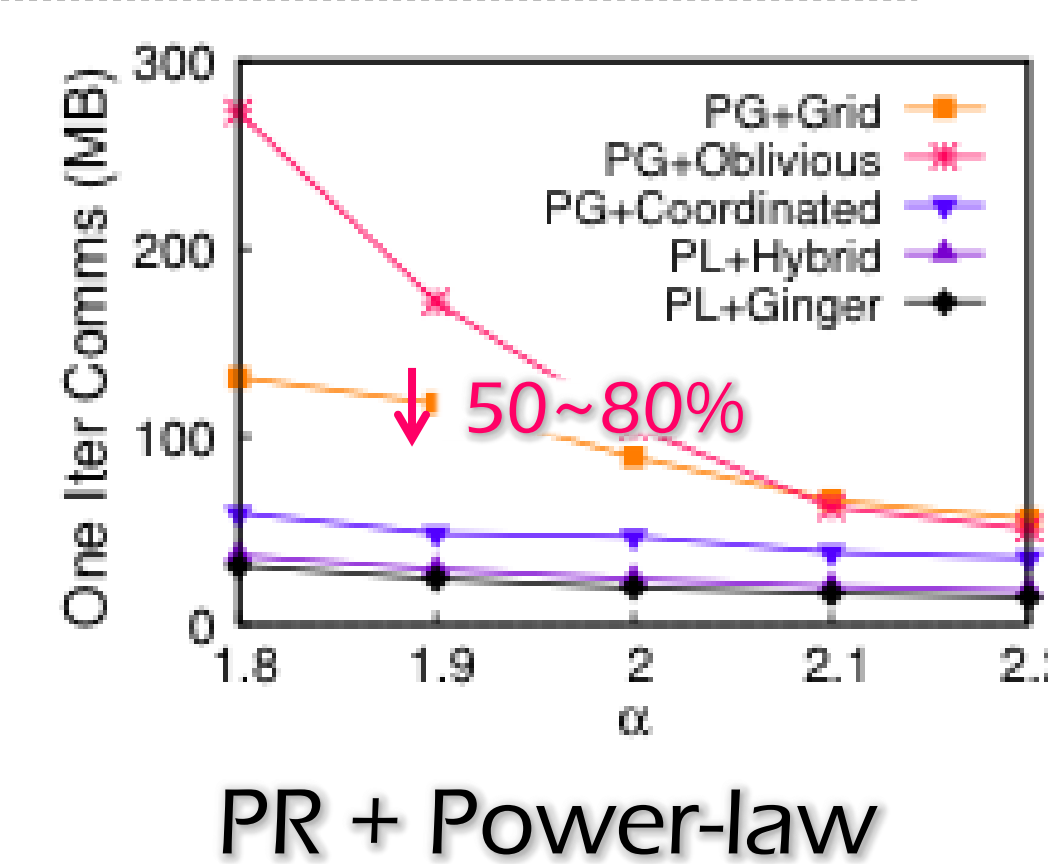
6-node Cluster (24 pCore, 64G RAM)
48-node Cluster (4 vCore, 12G RAM)



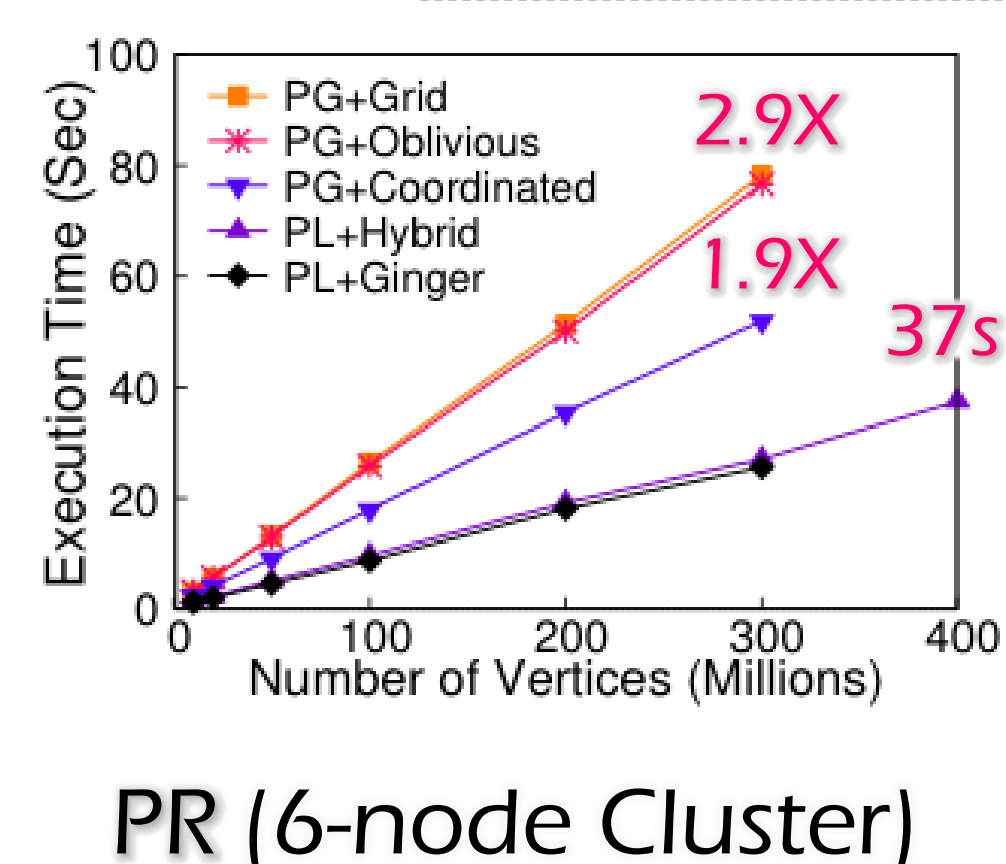
Performance



Communication



Scalability



Memory

