

Links: RDD [(CURL, seq[URL])]   
 src dest

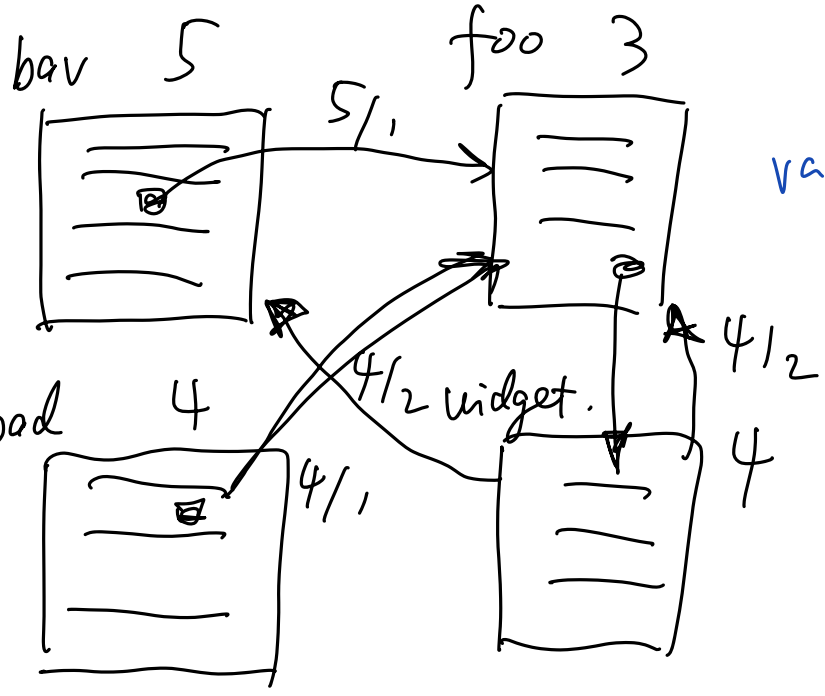
bar	foo
bad	foo
foo	widget
widget	bar, foo

RANKS: RDD [(CURL, RANK)]

bar	5
bad	4
foo	3
widget	4

RDD [(CURL, (seq[URL], RANK))]

bar	foo	5
bad	foo	4
foo	widget	3
widget	bar, foo	4



flatMap

foo	5/1
foo	4/1
widget	3/1
bar	4/2
foo	4/2

reduceBy Key

RANKS.

foo	$5 + 4 + 4/2 = 11$
widget	3
bar	2
bad	0