A \sqrt{N} Algorithm For Mutual Exclusion in Decentralized Systems

A short story

Mamoru Maekawa

University of Tokyo

ACM Transactions on Computer Systems, Vol. 3, No. 2, May 1985 pp.145-159

Contenido

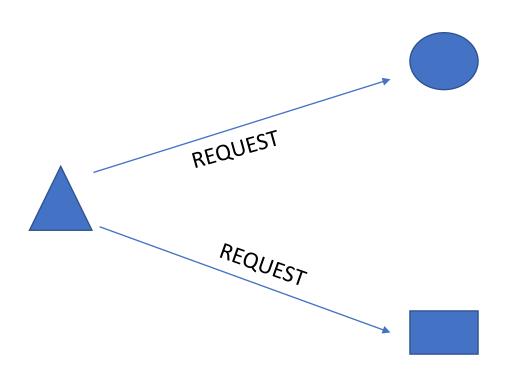
- Resumen
- Algoritmo de Maekawa
- Geometría Proyectiva
- Propiedades del Quorum
- Construcción del Quorum
- Ejemplo
- Conclusiones

Resumen

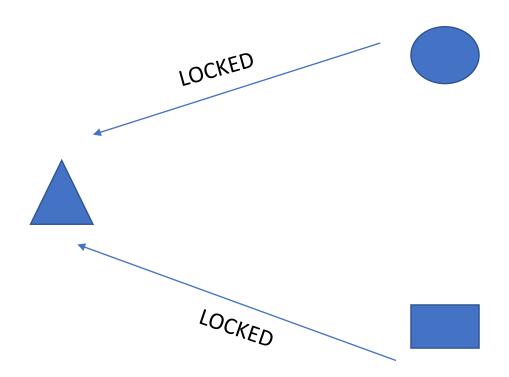
El algoritmo de Maekawa es un enfoque basado en quórum para garantizar la exclusión mutua en sistemas distribuidos.

Usa solo c \sqrt{N} mensajes para crear exclusión mutua en una red de computadoras, donde N es el número de nodos y c una constante entre 3 y 5. El algoritmo es simétrico y permite una operación totalmente paralela.

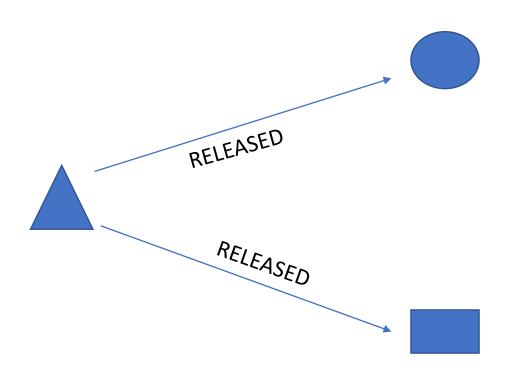
En el enfoque basado en el quórum, cada nodo solicita permiso para acceder a la sección critica desde un subconjunto de nodos (llamado quórum). Los quórums se forman de tal manera que cuando dos nodos soliciten acceso a la sección critica, un nodo recibe ambas solicitudes y es responsable de asegurarse de que solo una solicitud ejecute el acceso a la sección critica.



- TRÁFICO DE MENSAJES LIGERO
- TODOS LOS NODOS SE ENCUENTRAN LIBRES



- TRÁFICO DE MENSAJES LIGERO
- TODOS LOS NODOS SE ENCUENTRAN LIBRES



- TRÁFICO DE MENSAJES LIGERO
- TODOS LOS NODOS SE ENCUENTRAN LIBRES

TRÁFICO DE MENSAJES ALTO

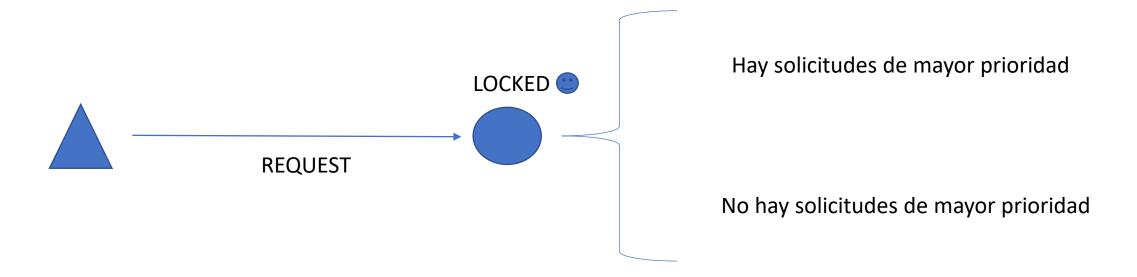
EL NODO SE ENCUENTRA BLOQUEADO



TRÁFICO DE MENSAJES ALTO

EL NODO SE ENCUENTRA BLOQUEADO

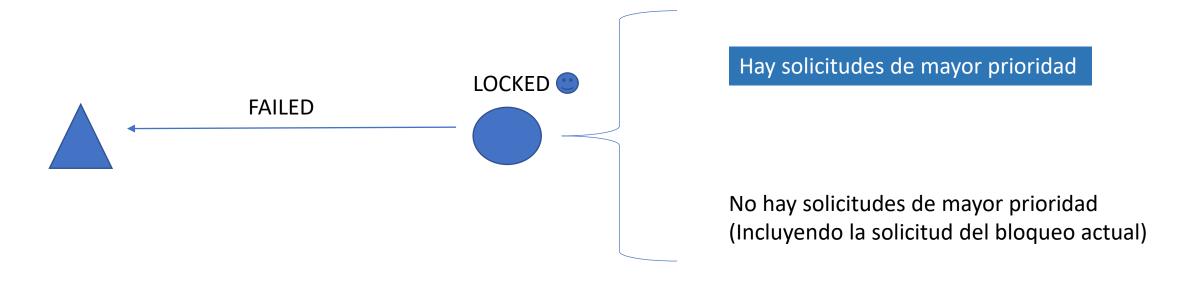
SOLICITUD SE MANDA A LA COLA DE ESPERA BUSCA PRIORIDADES



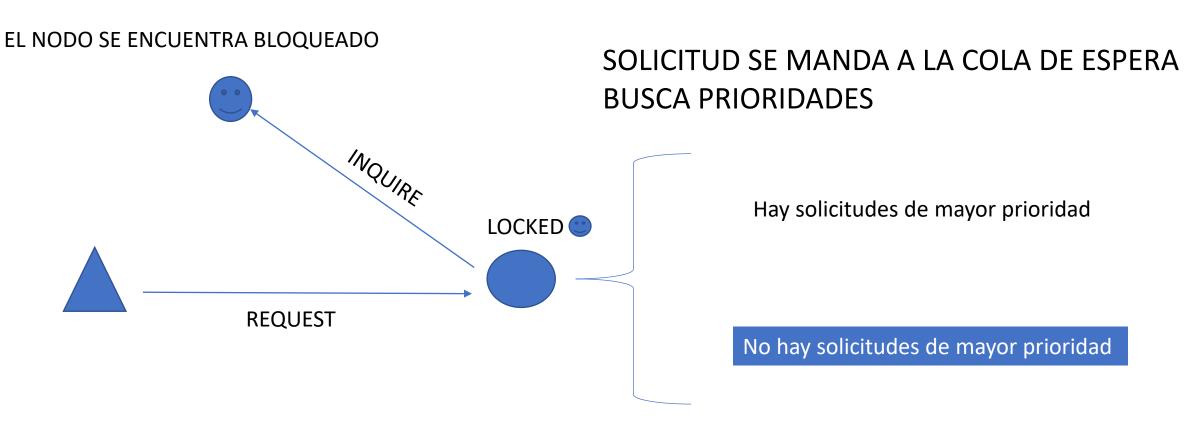
TRÁFICO DE MENSAJES ALTO

EL NODO SE ENCUENTRA BLOQUEADO

SOLICITUD SE MANDA A LA COLA DE ESPERA BUSCA PRIORIDADES



TRÁFICO DE MENSAJES ALTO

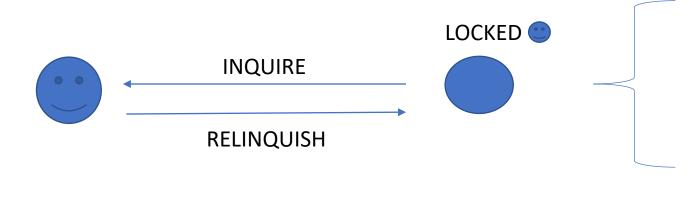


Pregunta si se han logrado bloquear los demás nodos

Supongamos que el nodo © ya tiene respuesta de todo su quorum

TRÁFICO DE MENSAJES ALTO

EL NODO SE ENCUENTRA BLOQUEADO



Se libera del estado bloqueado y se vuelve a bloquear pero esta vez con la petición de mayor prioridad que tenga en su cola de espera()

Recibió un mensaje FAILED

TRÁFICO DE MENSAJES ALTO

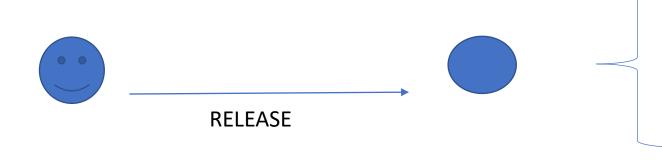
EL NODO SE ENCUENTRA BLOQUEADO



El nodo va a entrar o ya encuentra en sección crítica

TRÁFICO DE MENSAJES ALTO

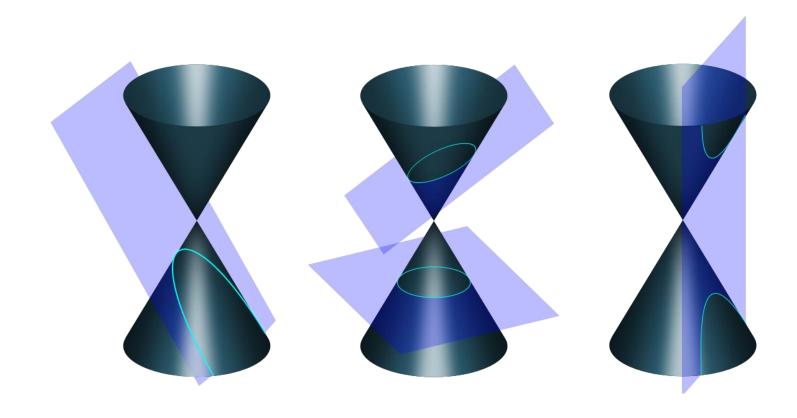
EL NODO SE ENCUENTRA BLOQUEADO



Se libera del estado bloqueado y se vuelve a bloquear pero esta vez con la petición de mayor prioridad que tenga en su cola de espera

¿Que es la Geometría Proyectiva?

La geometría proyectiva estudia las *propiedades de incidencia* de las figuras geométricas, parte de los siguientes principios:



¿Que es la Geometría Proyectiva?

La geometría proyectiva estudia las *propiedades de incidencia* de las figuras geométricas, parte de los siguientes principios:

- Dos puntos determinan una única recta.
- Dos rectas cualesquiera se cortan en un punto.

Planos Proyectivos Finitos

Es un conjunto de N puntos. Junto a una colección $S \neq \emptyset$ de subconjuntos de N, cuyos elementos son llamados rectas, tales que:

- Si *i* y *j* son dos elementos distintos de *N* entonces
- un único subconjunto S de N tal que $i \in S$ y $j \in S$.
- Si S_i y S_j son dos subconjuntos distintos de N entonces existe un único elemento i de N tal que $i \in S_i$ y $i \in S_j$.
- Existe cuatro puntos $i, j, k, l \in N$, de los cuales tres no pertenecen a un mismo subconjunto S.

- Todas las rectas tienen el mismo numero de puntos.
- Todos los puntos están en el mismo numero de rectas
- Hay el mismo numero de puntos y de rectas

Planos Proyectivos Finitos

Un plano finito de orden k es aquel en el cual cada recta tiene k+1 puntos

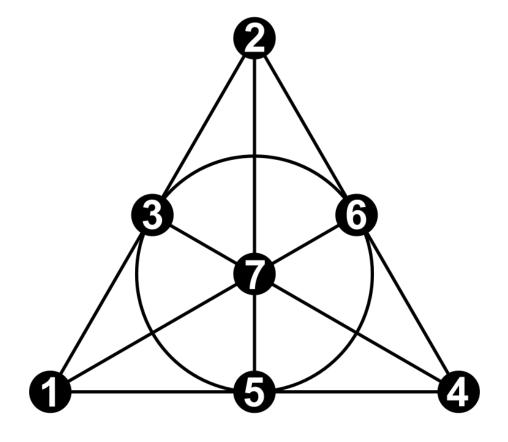
- Un plano proyectivo finito de orden k existe si k es una potencia p^m , de un primo p
- Este plano proyectivo posee k(k + 1) + 1 puntos

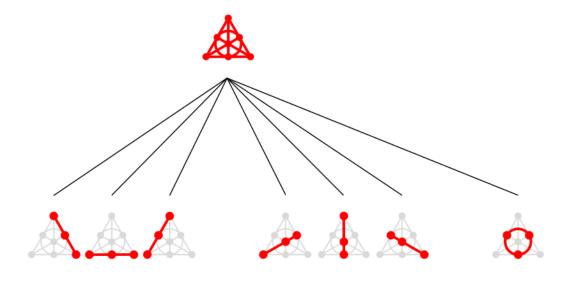
Planos Proyectivos Finitos

Ejemplo: Sea $N=\{1,2,3,4,5,6,7\} \leftarrow Plano\ proyectivo\ finito\ de\ orden\ 2$, la geometría más pequeña que satisface los tres axiomas

$$S_1 = \{1,2,3\}$$

 $S_2 = \{2,4,6\}$
 $S_3 = \{3,5,6\}$
 $S_4 = \{1,4,5\}$
 $S_5 = \{2,5,7\}$
 $S_6 = \{1,6,7\}$
 $S_7 = \{3,4,7\}$





- Todas las rectas poseen el mismo numero de puntos
- Todos los puntos están en el mismo numero de rectas
- Hay el mismo numero de puntos y de rectas

Propiedades de Quorum

- Para cualquier $i, j, 1 \le i, j \le N, S_i \cap S_j \ne \emptyset$
- → Cada uno de los nodos son un arbitro, la exclusión mutua se lleva a cabo

• El nodo i siempre pertenece a S_i

→ Minimiza la cantidad de mensajes

• Para cualquier i, $|S_i| = K$, es decir

→ Garantiza que cada nodo mande y reciba la misma la misma cantidad de mensajes

$$|S_1| = S_2 = \dots = |S_N| = K$$

- → Cada nodo es arbitro de la misma cantidad de nodos
- Todo nodo j, $1 \le j \le N$, pertenece a D S_i 's

Construcción del Quorum

- Para cualquier $i, j, 1 \le i, j \le N, S_i \cap S_j \ne \emptyset$
- El nodo i siempre pertenece a S_i
- Para cualquier i, $|S_i| = K$, es decir

$$|S_1| = S_2 = \dots = |S_N| = K$$

• Todo nodo j, $1 \le j \le N$, pertenece a D S_i 's

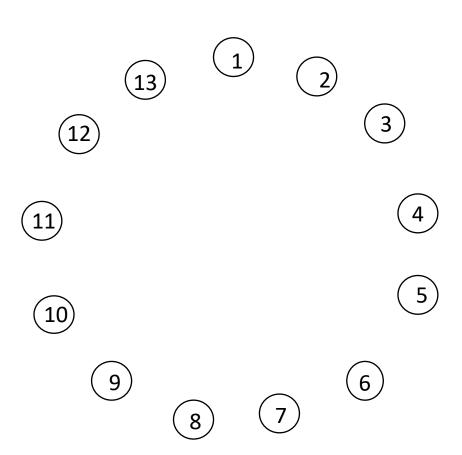
Cual es el número máximo de subconjuntos que satisfacen $S_i \cap S_j \neq \emptyset$?

$$(D-1)K+1$$
 $N=(D-1)K+1, K=D$
 $N=K(K-1)+1 \rightarrow K^2-K+1-N=0$

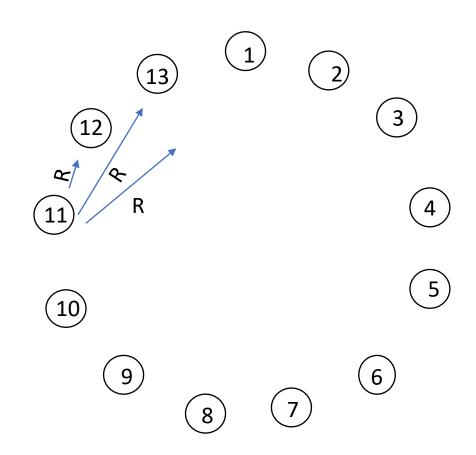
$$K = \frac{\left[-(-1) \pm \sqrt{(-1)^2 - 4 * (1)(1 - N)}\right]}{2}$$

$$K = \frac{\left[1 \pm \sqrt{4N - 3}\right]}{2} \rightarrow K_{+} = \frac{\left[1 + \sqrt{4N - 3}\right]}{2} \approx \sqrt{N}$$

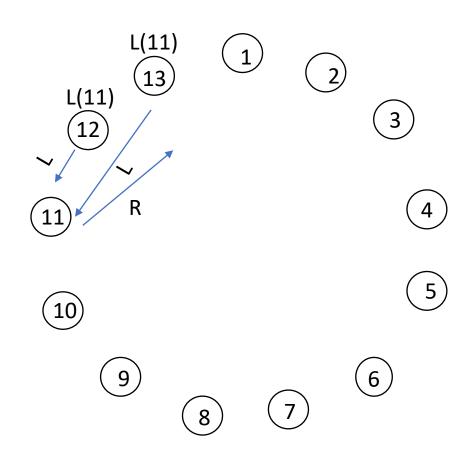
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



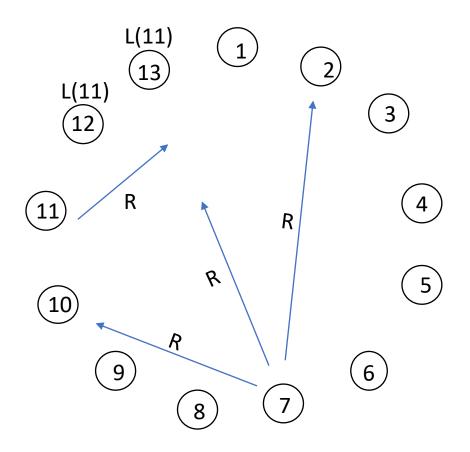
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



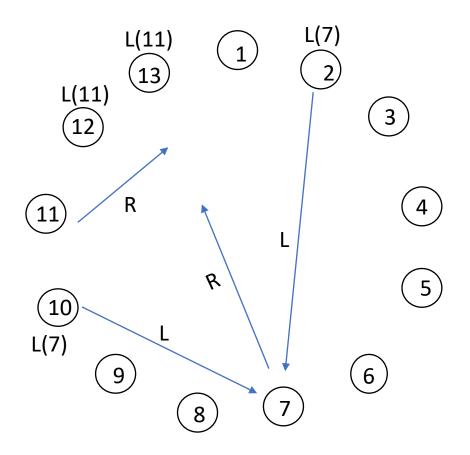
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



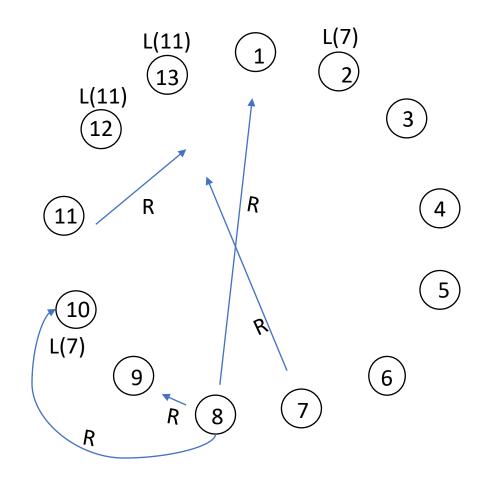
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



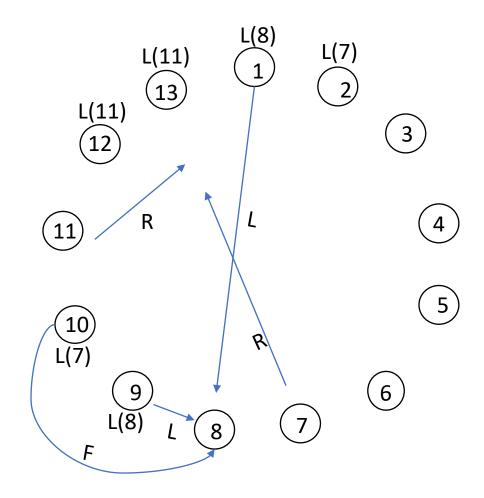
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



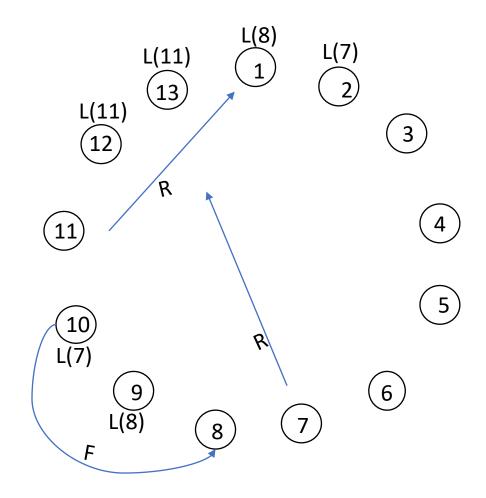
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$

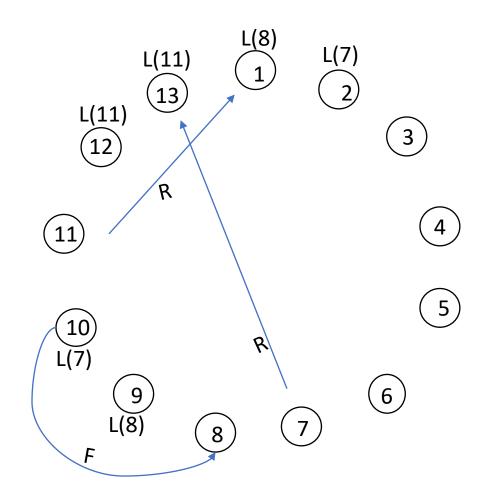


$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



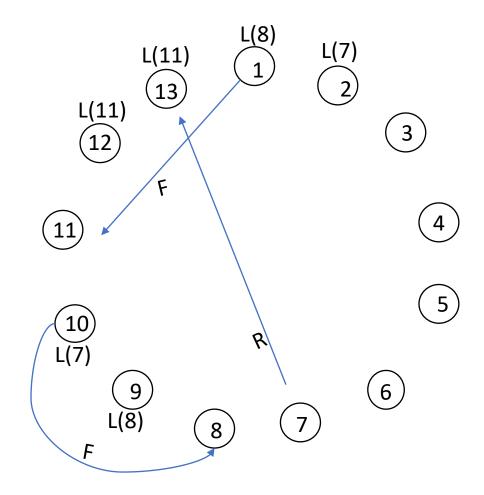
$$S_1 = \{1,2,3,4\}$$

 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$

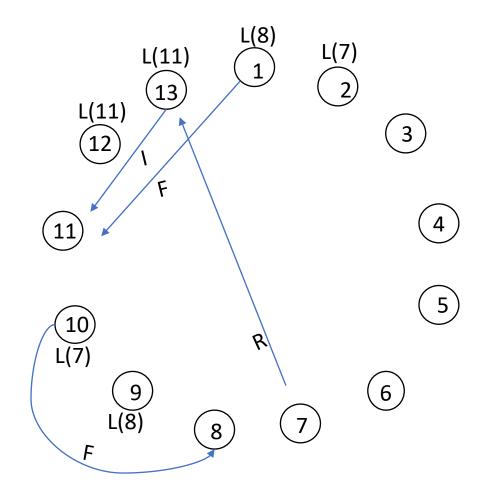


$$S_1 = \{1,2,3,4\}$$

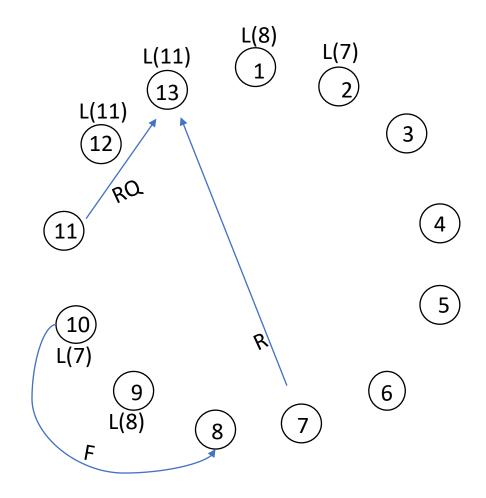
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 \cap S_{11} = \{13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



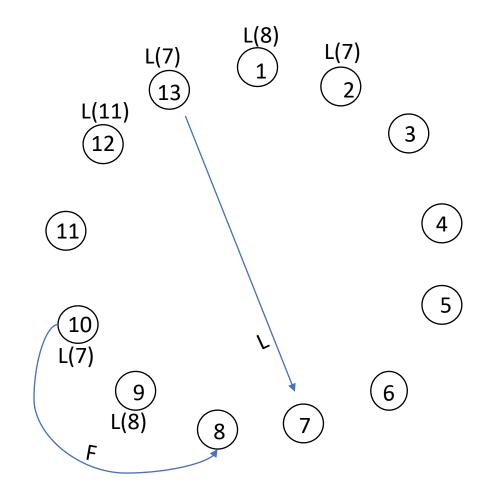
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$

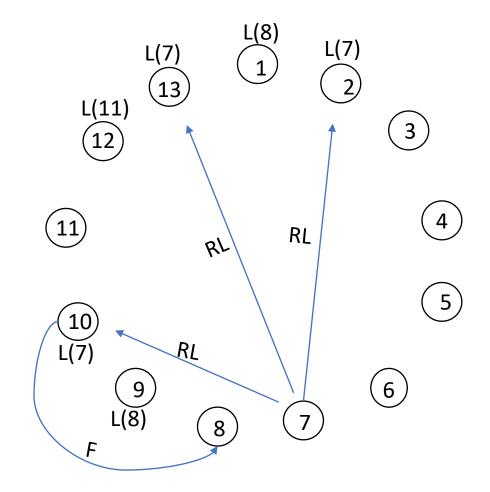


$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$

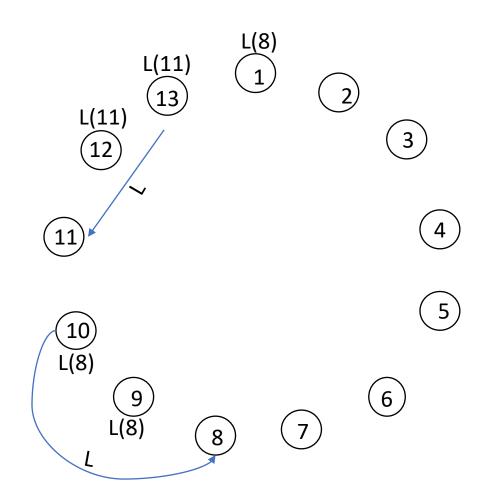


$$S_1 = \{1,2,3,4\}$$

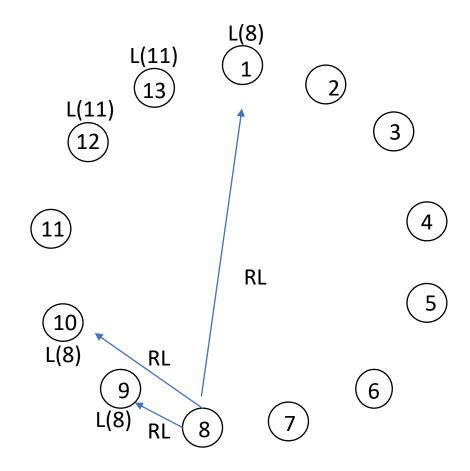
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



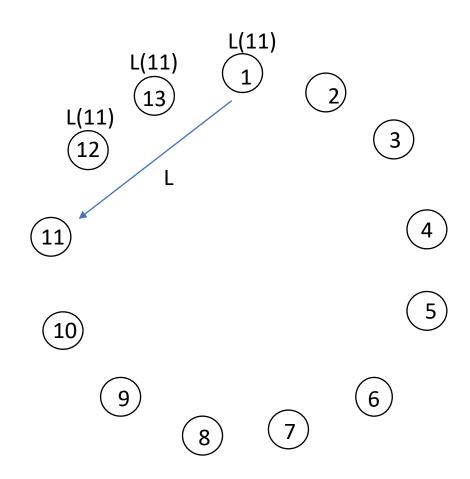
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



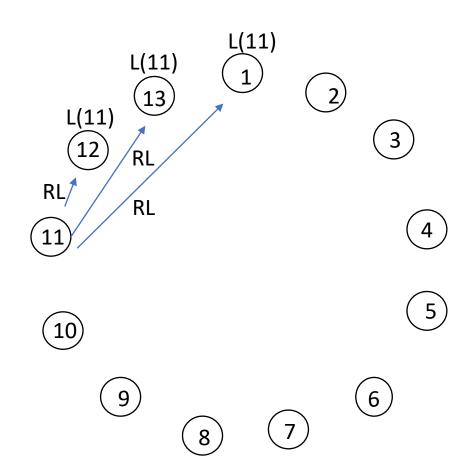
$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$

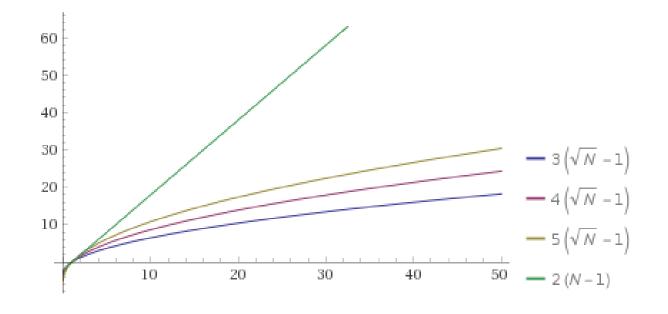


$$S_1 = \{1,2,3,4\}$$
 $S_2 = \{2,5,8,11\}$
 $S_3 = \{3,6,8,13\}$
 $S_4 = \{4,6,10,11\}$
 $S_5 = \{1,5,6,7\}$
 $S_6 = \{2,6,9,12\}$
 $S_7 \cap S_8 = \{10\}$
 $S_7 = \{2,7,10,13\}$
 $S_8 = \{1,8,9,10\}$
 $S_9 = \{3,7,9,11\}$
 $S_{10} = \{3,5,10,12\}$
 $S_{11} = \{1,11,12,13\}$
 $S_{12} = \{4,7,8,12\}$
 $S_{13} = \{4,5,9,13\}$



Para tráfico de mensajes ligero se tiene que mandar:

(K-1) REQUEST (K-1)LOCKED (K-1)RELEASE TOTAL $3(K-1)=3(\sqrt{N}-1)$



Para tráfico de mensajes alto se tiene que mandar:

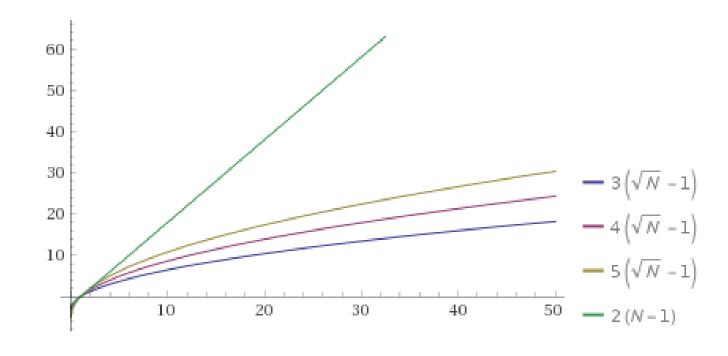
(K-1) REQUEST

(K-1) FAILED

(K-1)LOCKED

(K-1)RELEASE

TOTAL 4(K-1)= $4(\sqrt{N-1})$



Para tráfico de mensajes alto se tiene que mandar:

(K-1) REQUEST

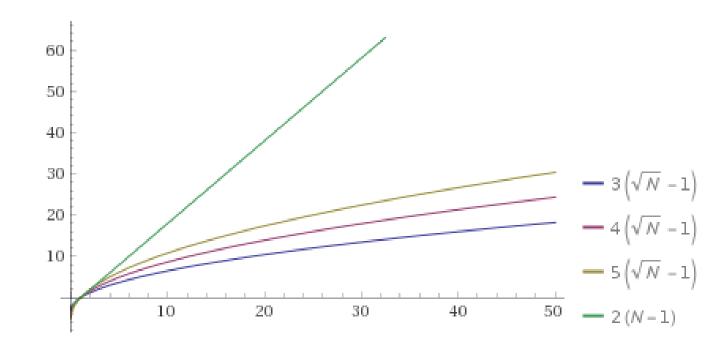
(K-1) INQUIRE

(K-1) RELINQUISH

(K-1)LOCKED

(K-1)RELEASE

TOTAL 5(K-1)= $5(\sqrt{N-1})$



El algoritmo de Maekawa reduce significativamente la complejidad de los mensaje de invocar la exclusión mutua al hacer que los nodos soliciten permiso solo de un subconjunto de nodos.

Sin embargo este algoritmo tiende a generar Deadlocks, porque un nodo está bloqueado exclusivamente por otros nodos y las solicitudes no tienen prioridad en sus marcas de tiempo(timestamps).