

# UNA NUEVA FAMILIA DE TOPOLOGÍAS INDIRECTAS, EFICIENTES Y TOLERANTES A FALLOS

## -RESUM-

Actualment, els grans sistemes de còmput paral·lel estan adoptant l'arquitectura *cluster* com a base per la seua construcció (Llista Top500). Aquests *clusters* estan sent construïts amb milers de nodes de processament, els quals es connecten mitjançant una xarxa d'interconnexió d'altres prestacions.

En aquests sistemes, el rendiment, la tolerància a fallades i el cost de la xarxa són un factor clau en el disseny de tot el sistema. Per altra banda, els nivells de còmput requerits només poden ser aconseguits incrementant el nombre de nodes que componen el *cluster*. Per tant, a mesura que el sistema creix també ho fa la quantitat de components de la xarxa, i amb això la probabilitat d'una fallada en la mateixa. Atès que la disponibilitat d'aquests sistemes és una gran preocupació, és habitual que les xarxes d'interconnexió implementen mecanismes de tolerància a fallades, que solen consistir en l'augment i replicació de components, incrementant el cost total de la xarxa.

En aquest camp, les topologies indirectes sovint són triades en el disseny de *clusters* d'alt rendiment. Entre elles, la més utilitzada és el fat-tree, una topologia bidireccional multietapa que presenta un bon rendiment de xarxa i un bon nivell de tolerància a fallades, però a un alt cost. Per reduir aquest cost, es va proposar RUFT, una topologia unidireccional multietapa que obté un rendiment de xarxa similar al fat-tree utilitzant menys recursos hardware (aproximadament la meitat). No obstant això, el punt feble de RUFT és que no ofereix cap tipus de tolerància a fallades.

En aquest treball, ens centrem en dissenyar una topologia indirecta que, per una banda, aconseguisca un alt rendiment de xarxa i siga tolerant a fallades i, per altra banda, tinga un baix cost.

Concretament, proposem una nova família de topologies indirectes amb diferents propietats pel que fa a cost, rendiment i tolerància a fallades. Aquestes noves topologies obtenen un rendiment similar o millor al que ofereix el fat-tree, a més d'oferir un bon nivell de tolerància a fallades. A més, a diferència de la majoria de topologies disponibles, toleren fallades en els enllaços que connecten amb els nodes de processament.

La nostra primera contribució és RUFT-PL, una topologia que duplica els enllaços d'injecció, xarxa i ejecció, seguint el mateix patró de connexió utilitzat per RUFT per interconnectar tots els elements de la xarxa. Aquesta topologia obté un alt rendiment de xarxa i un lleuger grau de tolerància a fallades, emprant els mateixos recursos de hardware que el fat-tree.

Com a segona contribució, proposem la topologia FT-RUFT-212. Aquesta topologia incrementa el rendiment de xarxa respecte al fat-tree, oferint a més a més un bon nivell de tolerància a fallades amb un baix cost de disseny, gràcies al sistema d'injecció/ejecció proposat que implementen els nodes de processament.

La tercera contribució, FT-RUFT-222, és una topologia que aprofita les millors propietats de les dues propostes anteriors. En particular, aquesta topologia implementa la injecció/ejecció utilitzada per FT-RUFT-212 i els dobles enllaços de xarxa de RUFT-PL per a connectar els commutadors. Aquesta proposta ofereix un alt rendiment de xarxa i de tolerància a fallades, utilitzant els mateixos recursos hardware requerits pel fat-tree.

La nostra última contribució és FT-RUFT-XL, una topologia que redissenya tant la injecció / ejecció com la connexió entre els commutadors. Aquesta topologia incrementa notablement el nivell de tolerància a fallades oferit per les altres propostes, presentant alhora un elevat rendiment de xarxa. A més a més, a diferència de moltes topologies unidireccionals, aquesta permet que els paquets prenguin rutes diferents en cada etapa de la xarxa, acostant-se sempre al seu destí en cada salt.