

Zookeeper Leader Election

الخوارزمية المطبقة:

- تقوم كل عقدة بالاتصال مع مخدّم zookeeper.
- ثم تقوم هذه العقدة بإضافة طلب ترشّح لتصبح القائد، وذلك من خلال إضافة ZNode خاصة بها تحت election ZNode ويكون لكل ZNode اسم من نمط leader_[SEQUENTIAL_NUMBER]، ويكون نمط انشاء العقد هو EPHEMERAL_SEQUENTIAL أي أن العقدة التي تصبح غير متصلة يتم حذف ZNode الخاصة بها بشكل نهائي.
- تقوم العقدة بعدها باختبار ما إذا كانت القائد ام لا.
- يتم تحديد القائد من خلال إنه كل عقدة تقوم بالحصول على جميع ZNodes الموجودة ضمن election/ وتقوم بترتيب هذه العقد تصاعدياً، والعقدة التي تكون لها أصغر ترتيب أي لها أصغر رقم تسلسلي (أي تم إضافتها قبل جميع العقد) تصبح هي القائد.
- في حال لم تكن العقدة هي القائد، تنتظر أن تصلها إشارة (حدث) بأن العقدة التي تسبقها بالرقم التسلسلي (أول عقدة لها رقم تسلسلي أقل منها) تم حذفها.
- في حال تم حذف العقدة التي تسبقها مباشرة تعيد هذه العقدة تطبيق خوارزمية انتخاب قائد وتختبر ما إذا كانت ستصبح القائد ام لا وتعيد نفس الخطوات.

مثال:

في حال كان لدينا ثلاث عقد، ستصل هذه العقد بمخدّم zookeeper وتقدّم طلب ترشّح لتصبح القائد، وسيتم إنشاء تحت election znode ثلاثة znodes كما يلي على فرض إنه سجلت Node2 أولاً ثم Node1 ثم Node3:

Node1 = leader_2

Node2 = leader_1

Node3 = leader_3

وبالتالي وبحسب الخوارزمية المطبقة فإن Node2 ستصبح هي القائد، وستنتظر Node1

حتى يتم حذف Node2 لتختبر ما اذا كانت ستصبح القائد ام لا، وستنتظر Node3 حتى

يتم حذف Node1 لتختبر ما إذا كانت القائد ام لا.

نميز مجموعة من الحالات:

- ❖ Node2 أصبحت غير متصلة: هنا سيتم حذف znode الخاصة ب Node2 وهي leader_1، وستختبر Node1 ما اذا كانت ستصبح القائد ام لا وستصبح هي القائد، دون أن تتأثر Node3 بأي شيء.
- ❖ Node1 أصبحت غير متصلة: هنا سيتم حذف znode الخاصة ب Node1 وهي leader_2، وستختبر Node3 ما اذا كانت ستصبح القائد ام لا ولكن لن تكون هي القائد لأن Node2 لا تزال متصلة، دون أن تتأثر Node2 بأي شيء.
- ❖ Node3 أصبحت غير متصلة: هنا سيتم حذف znode الخاصة ب Node3 وهي leader_3، دون أن تتأثر Node2 و Node1 بأي شيء.

• بهذا تكون كل عقدة ستعرف نفسها ما اذا كانت قائد ام لا، وتعرف متى يحق لها أن تختبر نفسها اذا ما كانت ستصبح القائد ام لا، ولكن يتبقى لدينا مشكلة واحدة وهي كيف ستعرف باقي العقد من هو القائد الجديد في حال أصبح القائد الحالي غير متصل لأنه إلى الآن فقط العقدة التي أصبحت قائد تعرف بأنها أصبحت القائد ولكن باقي العقد لم تعرف بأن القائد قد تغير ولحل هذه المشكلة نقوم بتخزين معلومة من هو القائد الحالي ضمن معطيات كل عقدة، وبنفس الطريقة السابقة بدل أن تستمع كل عقدة ما اذا كانت العقدة التي تسبقها قد حُذفت ستستمع ما اذا كانت معطيات العقدة التي تسبقها قد تغيرت، بحيث أن العقدة التي تصبح القائد ستعدل معطياتها بأنها أصبحت القائد، ستتلقى العقدة التي تليها حدث بأن العقدة السابقة قد غيرت معطياتها وبالتالي ستقوم العقدة بتعديل معطياتها أيضاً بمعلومة القائد الجديد وهكذا حتى تعلم جميع العقد من هو القائد الجديد.

وهكذا نلاحظ بأننا لن نقع في مشكلة Herd Effect وذلك لأنه عندما يصبح القائد غير متصل لن تقوم جميع العقد الباقية باستلام إشعارات (أحداث) بأن القائد أصبح غير متصل وبالتالي تتسابق العقد الباقية لاختبار ما إذا كانت ستصبح القائد ام لا، وإنما عقدة واحدة فقط ستستلم إشعار بأن القائد أصبح غير متصل وستختبر هذه العقدة ما إذا كانت ستكون هي القائد أم لا. في حال كانت القائد ستعدل معطياتها بأنها أصبحت القائد الجديد، أما باقي العقد كل عقدة ستستمع لحدث أن العقدة السابقة قد غيرت معطياتها أو إنها حُذفت.

الطالب: علي حسان سعيد.