

ملخص البحث رقم (2)

ملخص البحث باللغة العربية :

نمو البيانات في السنوات الأخيرة كان سريعاً، مما أدى إلى ظهور علم البيانات الضخمة. تستخدم أنظمة الملفات الموزعة (DFS) بشكل شائع للتعامل مع البيانات الضخمة، مثل نظام الملفات (GFS) Google، ونظام الملفات الموزعة (HDFS) Hadoop، وغيرها. يجب أن تحقق أنظمة الملفات الموزعة توفير البيانات وموثوقية النظام في حالة الفشل، حيث يقوم DFS بنسخ الملفات في مواقع مختلفة لتوفير التوفر والموثوقية.

تستهلك هذه النسخ مساحة التخزين والموارد الأخرى. تختلف أهمية هذه الملفات حسب تكرار استخدامها في النظام. لذلك، لا تستحق بعض هذه الملفات النسخ عدة مرات لأنها غير مهمة في النظام. تقدم هذه الورقة سياسة النسخ الديناميكي باستخدام التجميع التعلّم الآلي (DRMO على HDFS)، والتي تستخدم التعلّم الآلي لتجميع الملفات في مجموعات مختلفة وتطبيق سياسات النسخ الأخرى لكل مجموعة لتقليل استهلاك التخزين، وتحسين أوقات عمليات القراءة والكتابة، والحفاظ على التوفر والموثوقية من HDFS كحوسبة موزعة عالية الأداء (HPC).

وتفضلوا سيادتكم بقبول فائق الاحترام والتقدير،،