**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ**

**ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

**ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ**

**Νταλλαρής Ευστράτιος**

**Μεταπτυχιακός Φοιτητής

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης**

**Επόπτης Μεταπτυχιακής Εργασίας: Αναπλ. Καθηγητής, Κ. Μαγκούτης**

**Πέμπτη, 31 Μαρτίου 2022, ώρα 11:00 π.μ.**

**Join Zoom Meeting**

<https://zoom.us/j/93060654122>

**“ Αποδοτική Επαναστάθμιση Δεδομένων στο Κατανεμημένο Σύστημα Αποθήκευσης Κλειδιού-Τιμής Redis”**

**ΠΕΡΙΛΗΨΗ**

Η επαναστάθμιση δεδομένων είναι ένας σημαντικός μηχανισμός που χρησιμοποιείται σε κατανεμημένα συστήματα αποθήκευσης μεγάλου όγκου δεδομένων μετά από αλλαγές στους πόρους του συστήματος ή ανισορροπία στον φόρτο και απαιτεί την ανακατανομή των δεδομένων μεταξύ των διαθέσιμων κόμβων του συστήματος. Η επαναστάθμιση δεδομένων απαιτεί μεταφορά δεδομένων μεταξύ κόμβων για να εξασφαλίσει ομοιόμορφη κατανομή από άποψη χώρου και επεξεργαστικής ισχύος. H αποδοτικότητα της διαδικασίας επαναστάθμισης δεδομένων, σε όρους ταχύτητας και επίδρασης στην επίδοση της εφαρμογής (ειδικά κατά την διαδικασία της δυναμικής μεταβολής πόρων, γνωστής και ως ελαστικότητας), είναι σημαντική. Γρήγορες και με χαμηλό αντίκτυπο ενέργειες ανακατανομής δεδομένων σε κατανεμημένα συστήματα αποθήκευσης κλειδιού-τιμής είναι επιθυμητές ιδιότητες, απαιτείται ωστόσο προσεκτική σχεδίαση και υλοποίηση για να επιτευχθούν.

Στην παρούσα εργασία μελετούμε τον μηχανισμό επαναστάθμισης δεδομένων στο δημοφιλές σύστημα Redis Cluster. Αρχικά αναλύουμε τον προκαθορισμένο μηχανισμό επαναστάθμισης δεδομένων του Redis Cluster, μελετώντας τις πηγές αναποτελεσματικότητας στη σχεδίαση και την υλοποίησή του, και παρουσιάζουμε την επιρροή τους με μετρήσεις κάτω από φόρτο εργασίας παραγόμενο από την εφαρμογή στάθμισης απόδοσης YCSB. Στη συνέχεια σχεδιάζουμε και υλοποιούμε μία έκδοση του Redis Cluster με βελτιωμένο μηχανισμό επαναστάθμισης δεδομένων, εκμεταλλευόμενοι τις δυνατότητες του μηχανισμού Remote Direct Memory Access (RDMA) για μεταφορές δεδομένων με υψηλή ταχύτητα και μειωμένο κόστος σε υπολογιστικούς πόρους. Τα αποτελέσματά μας δείχνουν πως η βελτιωμένη έκδοση του Redis Cluster μπορεί να μεταφέρει δεδομένα σε έναν νέο εισαγόμενο κόμβο κοντά στην ταχύτητα γραμμής (1GB/s) σε ένα δίκτυο συστοιχίας 5 κόμβων, αφήνοντας ταυτόχρονα ελεύθερους πόρους επεξεργασίας διαθέσιμους στην εφαρμογή. Η βελτιωμένη έκδοση του Redis Cluster πετυχαίνει μία σημαντικά πιο ταχεία επαναστάθμιση δεδομένων σε σύγκριση με τον μηχανισμό επαναστάθμισης δεδομένων της δημόσια διαθέσιμης έκδοσης του Redis Cluster υπό φόρτο του YCSB κατά την ανάγνωση κλειδιών-τιμών, με χαμηλή χρήση πόρων επεξεργασίας και μνήμης στην μεταφορά δεδομένων.

**University of Crete**

**Computer Science Department**

**M.Sc. Thesis**

**Ntallaris Efstratios**

**Master's Thesis Supervisor: Associate Professor, K. Magoutis**

**Thursday, 31 March 2022, 11:00 a.m.**

**Join Zoom Meeting**

<https://zoom.us/j/93060654122>

**“Increasing the Efficiency of Data Rebalancing in the Redis Distributed Key-Value Store”**

**ABSTRACT**

Data rebalancing is an important mechanism used in data-intensive distributed systems when a change in the underlying resources or imbalance in the workload (e.g., a skew in the access pattern) necessitates re-distributing data across available storage nodes. Data rebalancing requires transferring (migrating) data between nodes to ensure a more equitable allocation in terms of space and/or processing capacity. The efficiency of the data rebalancing process, in terms of speed as well as impact on application performance is important, especially when used in an on-line fashion; in this case, the rebalancing process is also known as elasticity. Fast, low-impact data rebalancing actions are desirable; however, they require careful design and implementation to achieve in a distributed key-value store.

In this thesis, we study data rebalancing in a widely-deployed real-world system, Redis Cluster. Our study starts by analyzing the default data rebalancing mechanism (data path) in Redis Cluster, exposing sources of inefficiency in its design and implementation, and demonstrating their impact in measured performance in experiments under an I/O-intensive workload, the Yahoo Cloud Serving Benchmark (YCSB). We then design and implement a version of Redis Cluster that features an improved data-rebalancing path, leveraging remote direct memory access (RDMA) for low-overhead data transfers. Our results show that our improved version of Redis Cluster can transfer data to a new (joining) node in nearly wire speed (1GB/s) over a RDMA-capable network connecting a 5-node cluster, while allowing sufficient spare resources for application (YCSB) processing. Our improved version of Redis Cluster achieves significantly faster data rebalancing compared to standard (and tuned) Redis in YCSB read-only workloads while also maintaining lower CPU and memory footprint during data transfers.