

ΠΡΟΣ

- 1) Όλα τα μέλη ΔΕΠ του Τμήματος Επιστήμης Υπολογιστών
- 2) Τους εκπροσώπους των Μεταπτυχιακών φοιτητών του Τμήματος Επιστήμης Υπολογιστών
- 3) Την Επταμελή Εξεταστική Επιτροπή
- 4) Όλα τα μέλη της Πανεπιστημιακής Κοινότητας

Πρόσκληση σε Δημόσια Παρουσίαση της Διδακτορικής Διατριβής του

κ. Παπαϊωάννου Αντωνίου

Doctoral Dissertation Defense

Mr. Antonios Papaioannou

Την Παρασκευή, 22/10/2021 και ώρα 10:00 π.μ. μέσω Τηλεδιάσκεψης (e-presence) και με ζωντανή μετάδοση μέσω της υπηρεσίας Δίαυλος του ΕΔΥΤΕ, <https://diavlos.grnet.gr/epresence-conference-9951> και από το κανάλι του Τμήματος στο Youtube, https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB_Gnt6Q/live, θα γίνει η δημόσια παρουσίαση και υποστήριξη της Διδακτορικής Διατριβής του υποψηφίου διδάκτορα του Τμήματος Επιστήμης Υπολογιστών κ. Παπαϊωάννου Αντωνίου με θέμα:

“ Αποδοτικοί μηχανισμοί προσαρμογής για την βελτίωση της απόδοσης κατά τη διάρκεια εσωτερικών ή εξωτερικών μεταβολών σε καταναμημένα συστήματα αποθήκευσης δεδομένων”

“Efficient adaptation mechanisms for improving performance during internal or external changes in distributed data stores”

ΠΕΡΙΛΗΨΗ

Η εξέλιξη στον χώρο των καταναμημένων συστημάτων διαχείρισης δεδομένων, ιδιαίτερα μιας κατηγορίας τέτοιων συστημάτων που αναπτύχθηκαν τα τελευταία 15-20 χρόνια γνωστά και ως συστήματα NoSQL, είναι ραγδαία. Μια πληθώρα τέτοιων συστημάτων βασίζεται σε ποικίλες σχεδιαστικές αποφάσεις που εστιάζουν στην βελτιστοποίηση για συγκεκριμένους τύπους εφαρμογών, μορφών δεδομένων και χαρακτηριστικών φόρτου εργασίας. Τα συστήματα αυτά αποτελούν συνήθως μέρος πολυεπίπεδων εφαρμογών ικανών να εξυπηρετούν φόρτο μεγάλης κλίμακας (συνήθως προερχόμενο από το Internet) κάτω από πολύπλοκες συνθήκες. Ένας σημαντικός στόχος για τους παρόχους υπηρεσιών είναι να διασφαλίσουν το επίπεδο

απόδοσης των συστημάτων αυτών παρά τις εσωτερικές ή εξωτερικές αλλαγές και προκλήσεις που αντιμετωπίζουν. Ένας τέτοιος στόχος μπορεί να επιτευχθεί μέσω μηχανισμών προσαρμογής που επιτρέπουν στα συστήματα διαχείρισης δεδομένων να προσαρμόζονται στις αλλαγές αυτόματα και με διαφάνεια, διατηρώντας παράλληλα τους στόχους και την αποδοτικότητα τους.

Σε αυτή τη διατριβή διερευνούμε μηχανισμούς προσαρμογής σε κατανεμημένα συστήματα διαχείρισης δεδομένων που έρχονται αντιμέτωπα με αλλαγές που πηγάζουν είτε από το εσωτερικό του ίδιου του συστήματος ή από εξωγενείς παράγοντες. Μελετάμε ιδιαίτερα αλλαγές που αντιμετωπίζει το σύστημα λόγω αύξησης του φόρτου εργασίας, την εκτέλεση έκτακτων/περιοδικών δραστηριοτήτων στο παρασκήνιο ή την αλληλεπίδραση και συνέργεια με εξωτερικά συστήματα τα οποία εξελίσσονται ανεξάρτητα και παράλληλα με τα συστήματα διαχείρισης δεδομένων. Προτείνουμε νέους μηχανισμούς προσαρμογής των συστημάτων καθώς και βελτιώσεις σε υφιστάμενους μηχανισμούς σε τρία διαφορετικά πλαίσια (ελαστικότητα του συστήματος, διαχείριση της επίπτωσης δραστηριοτήτων που εκτελούνται στο παρασκήνιο και αποδοτικότερη συνέργεια και αλληλεπίδραση με εξωτερικά συστήματα επεξεργασίας δεδομένων). Μελετάμε προκλήσεις που δεν είχαν αντιμετωπιστεί μέχρι σήμερα με στόχο τη βελτίωση της συνολικής απόδοσης των συστημάτων.

Αρχικά εστιάζουμε στην φάση της επέκτασης του συστήματος κατά τη διάρκεια της οποίας ενσωματώνονται νέοι πόροι ώστε να βελτιωθεί η απόδοσή του και να ανταποκριθεί στην αύξηση του φόρτου και των προκλήσεων που αυτή συνεπάγεται. Τα δεδομένα ανακατανέμονται εσωτερικά ώστε οι νέοι πόροι να λάβουν το αναλογούν τους μερίδιο φόρτου. Μελετάμε τον αντίκτυπο στην απόδοση του συστήματος εξαιτίας της μεταφοράς δεδομένων πάνω από το δίκτυο κατά αυτή την ανακατανομή των δεδομένων. Προτείνουμε έναν μηχανισμό ο οποίος δημιουργεί και εφαρμόζει ένα πλάνο μεταφοράς των δεδομένων σε μικρότερα αυτόνομα τμήματα. Κάθε φορά που η μεταφορά ενός τμήματος των δεδομένων ολοκληρώνεται, η εξυπηρέτηση του φόρτου που αντιστοιχεί στα δεδομένα αυτά γίνεται από τους νέους πόρους. Με αυτόν τον τρόπο το συνολικό σύστημα ενεργοποιεί οφέλη των νέων πόρων καθώς η διαδικασία επέκτασης βρίσκεται ακόμα υπό εξέλιξη (πριν ολοκληρωθεί η πλήρης μεταφορά όλων των δεδομένων) με αποτέλεσμα να αυξάνεται προοδευτικά η συνολική χωρητικότητα του συστήματος.

Στη συνέχεια, μελετάμε δραστηριότητες που εκτελούνται στο παρασκήνιο (λειτουργίες για την εσωτερική αναδιοργάνωση των δεδομένων ή ενέργειες δημιουργίας αντιγράφων ασφάλειας) οι οποίες καταναλώνουν πόρους του συστήματος με αποτέλεσμα να έχουν αρνητική επίδραση στην απόδοση του συστήματος. Οι κόμβοι των συστημάτων αυτών είναι συνήθως οργανωμένοι σε ομάδες αντιγράφων για λόγους διαθεσιμότητας των δεδομένων. Συστήματα τα οποία υποστηρίζουν ισχυρή συνέπεια των δεδομένων περιορίζουν την ανάγνωση και εγγραφή των δεδομένων σε ένα υποσύνολο κόμβων. Προτείνουμε ένα μηχανισμό προσαρμογής της ομάδας αντιγράφων ώστε οι κόμβοι που εξυπηρετούν κυρίως τον φόρτο των χρηστών να είναι αυτοί οι οποίοι δεν έχουν ενεργές διεργασίες παρασκηνίου. Επιπλέον παρατηρούμε ότι ο μηχανισμός αλλαγής των κόμβων που

εξυπηρετούν τον φόρτο που παράγουν οι χρήστες εμφανίζει μια αδυναμία η οποία σχετίζεται με την κρυφή μνήμη (cache). Αυτή οφείλεται στο ότι τα δευτερεύοντα αντίγραφα δεν ενημερώνουν πλήρως την cache τους με αποτέλεσμα να μην είναι επαρκώς προετοιμασμένα όταν καλούνται να ξεκινήσουν να εξυπηρετούν τους χρήστες. Αυτή η παρατήρηση μας οδήγησε να προτείνουμε έναν νέο μηχανισμό που επιτρέπει σε δευτερεύοντα αντίγραφα να διατηρούν ενημερωμένες caches. Ο μηχανισμός που προτείνουμε βασίζεται στην διάδοση των εντολών ανάγνωσης σε όλους τους κόμβους που φιλοξενούν αντίγραφα των δεδομένων χωρίς να αλλοιώνεται το μοντέλο συνέπειας και διαθεσιμότητας των δεδομένων του συστήματος.

Τέλος, στη διατριβή αυτή διερευνούμε τα οφέλη του αυτόματου συντονισμού κατανεμημένων συστημάτων διαχείρισης δεδομένων με εξωτερικά συστήματα με τα οποία αυτά αλληλεπιδρούν (π.χ. κατανεμημένα συστήματα επεξεργασίας δεδομένων) μέσω του συντονισμού της ανάπτυξης των συστημάτων επί των υπολογιστικών και αποθηκευτικών πόρων. Προτείνουμε ένα σύστημα που διερευνά συνεχώς ευκαιρίες συντονισμού μεταξύ των συστημάτων και βελτιώνει την τοποθεσία των δεδομένων ώστε να βρίσκονται εγγύτερα στους πόρους επεξεργασίας τους. Ο συντονισμός των συστημάτων επιτυγχάνεται χρησιμοποιώντας μηχανισμούς προσαρμογής που σχετίζονται με την δημιουργία αντιγράφων και την μεταφορά δεδομένων. Επιπλέον μελετάμε για πρώτη φορά σε αυτό το πλαίσιο την προσαρμογή του μηχανισμού με τον οποίο τα δεδομένα διαμερίζονται και κατανέμονται στους κόμβους του συστήματος.

Η πειραματική αξιολόγηση των μηχανισμών που μελετώνται σε αυτή την διατριβή επιβεβαιώνει τα οφέλη τους στην βελτίωση της απόδοσης ευρέως διαδεδομένων συστημάτων. Η μελέτη αυτή συμβάλλει στην εξέλιξη των συστημάτων διαχείρισης δεδομένων προς την κατεύθυνση συστημάτων που μπορούν να προσαρμόζονται αποδοτικά καθώς έρχονται αντιμέτωπα με εσωτερικές ή εξωτερικές αλλαγές κατά την διάρκεια του κύκλου ζωής τους.

Επιβλέπων: Αναπλ. Καθηγητής, Κωνσταντίνος Μαγκούτης

ABSTRACT

The evolution of distributed data management systems, especially a class of systems developed during the last 15-20 years commonly known as NoSQL data stores, has led to a multitude of designs optimised for different application types, data formats, and workload characteristics. Given the complexity of the environments they operate in as parts of multi-tier software stacks driven by Internet workloads, data stores are facing significant challenges during their operation. An important objective for service

operators is to ensure that data store performance levels and guarantees are maintained despite internal or external changes that they face. Such an objective can be reached via automated adaptation mechanisms by which data stores adapt to changes automatically and transparently while maintaining efficiency and performance goals as the data store transitions to new configurations.

In this dissertation we explore adaptation mechanisms in distributed data stores facing internally or externally-induced changes, with a focus on workload variations, occasional background activities, or the evolution of an external middleware component that interoperates with a distributed data store. We propose novel adaptation mechanisms and improvements to existing mechanisms in three different contexts (data store elasticity, masking background activities, and alignment with external distributed middleware), aiming to improve the overall performance during the aforementioned contexts in the lifecycle of scalable data stores, aiming at challenges that had not been addressed so far.

First, this dissertation focuses on the expansion phase of a data store when the need arises to adapt its capacity as workload demands increase and the system tries to improve its performance by incorporating more resources. We study the performance impact of data transfers over the network during this phase and propose a mechanism that schedules data transfers in a fine-grain manner, reducing their performance impact while progressively increasing the processing capacity in an incremental fashion. The proposed method realizes early benefits from data transfers during the elasticity action as it incorporates new resources and makes data sub-sections available prior to completing the full data transfers.

Next, we study the performance overhead of background activities that often impact data store performance. We propose replica-group reconfiguration as a way to mask performance bottlenecks in replicated data stores and investigate the benefits of changing replica-group leadership prior to resource-intensive background tasks (e.g. internal data reorganization, garbage collection or data backup tasks). Our observation of an occasional performance glitch during reconfiguration actions, caused by cold-cache misses in the cache of a new leader that was not adequately prepared for the transition to the new configuration, led us to propose a new mechanism to maintain up-to-date read caches across replicas without affecting the data consistency and availability by disseminating read-hints within the replica group.

Finally, in this dissertation we investigate the benefits of automatically aligning data stores with distributed middleware systems that rely on those data stores to maintain their state. We do that by appropriately co-locating data partitions of data store with processing tasks of the distributed middleware systems. We propose a system that continuously strives to discover such alignment opportunities across systems and improve data locality. The alignment actions combine multiple data store mechanisms in common use, such as data replication and migration, as well as the adaptation of the partitioning schemes across systems, a mechanism that has not been studied before in this context.

The evaluation of the proposed mechanisms over widely deployed systems confirms their performance improvements, advancing the state of the art in distributed data stores in the direction of systems that adapt more efficiently and in new ways through internal and external changes in their lifecycle.

Supervisor: Associate Professor, Konstantinos Magoutis