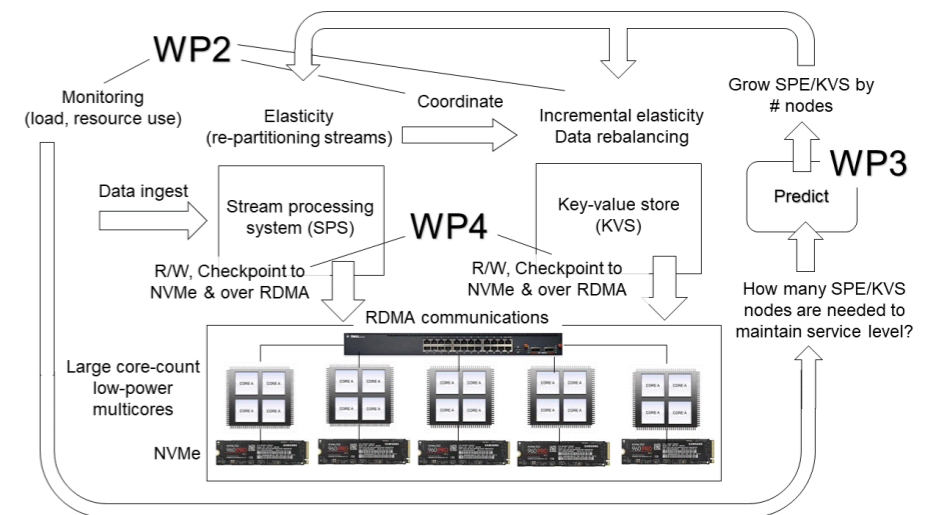




**ΕΛΙΔΕΚ.**  
Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας

**Περιγραφή Χρηματοδοτούμενου Ερευνητικού Έργου**  
**1η Προκήρυξη Ερευνητικών Έργων ΕΛ.ΙΔ.Ε.Κ. για την**  
**ενίσχυση των Μελών ΔΕΠ και Ερευνητών/τριών και την**  
**προμήθεια ερευνητικού εξοπλισμού μεγάλης αξίας**

# STREAMSTORE



## Ελαστική και Προβλέψιμη Αποθήκευση Δεδομένων Τύπου NOSQL για Κλιμακώσιμες Εφαρμογές Επεξεργασίας Ροών

## ΕΥ και Ερευνητική Ομάδα

Επιστημονικός Υπεύθυνος: Αν. καθ. Κων/νος Μαγκούτης, Πανεπιστήμιο Κρήτης



Φιλικός προς τον αναγνώστη τίτλος: STREAMSTORE

Επιστημονική Περιοχή: Μαθηματικά και Επιστήμες της Πληροφορίας

Ποσό Χρηματοδότησης: 169.422 Ευρώ

Φορέας Υποδοχής: Ίδρυμα Τεχνολογίας και Έρευνας

Διάρκεια Χρηματοδότησης: 36 Μήνες

Ιστοσελίδα προβολής του Έργου: <https://streamstore-project.eu>

## Σύνοψη Ερευνητικού Έργου

Η επεξεργασία μεγάλου όγκου δεδομένων ροής αποτελεί σημαντική επιχειρησιακή λειτουργικότητα για δημοφιλείς υπηρεσίες Διαδικτύου, που έχει οδηγήσει στην δημιουργία μιας νέας γενιάς συστημάτων επεξεργασίας ροών (ΣΕΡ). Η επεξεργασία ροών συχνά απαιτεί την χρήση μεγάλων δομών δεδομένων για την αποθήκευση γνώσης (ground-truth data) όπως κοινωνικοί γράφοι, προφίλ χρηστών, κλπ. Τέτοιες μόνιμες δομές τυπικά αποθηκεύονται σε καταναμημένα συστήματα κλειδιού-τιμής (ΣΚΤ) τύπου NoSQL. Προηγμένα ΣΕΡ επιπρόσθετα πρέπει να υποστηρίζουν ελαστική προσαρμογή σε ταχείες διακυμάνσεις επιπέδου φόρτου ροών, συχνά της τάξης 2x-3x σε σύντομο χρονικό διάστημα. Μια σημαντική πρόκληση είναι ότι το ΣΚΤ σε αυτές τις περιπτώσεις πρέπει να προσαρμόζεται τουλάχιστον όσο γρήγορα όσο το ΣΕΡ ώστε να διατηρείται το επίπεδο απόδοσης των εφαρμογών. Η προσαρμογή των ΣΚΤ απαιτεί ανακατανομή δεδομένων μέσω μεταφοράς στους συμμετέχοντες κόμβους, μια χρονοβόρα και απαιτητική σε πόρους διαδικασία. Επιπρόσθετα, δράσεις ελαστικότητας στο ΣΚΤ σήμερα είναι αποσυνδεδεμένες και δεν συντονίζονται με δράσεις ελαστικότητας στο ΣΕΡ. Το ερευνητικό όραμα του STREAMSTORE είναι να υποστηρίξει πλήρως ελαστικά, κλιμακώσιμα ΣΕΡ τα οποία ταυτόχρονα αποθηκεύουν δεδομένα μεγάλου όγκου σε ΣΚΤ. Κατά τη διάρκεια περιόδων μεγάλης διακύμανσης φόρτου, η διατήρηση καλής απόκρισης απαιτεί ικανότητα ελαστικότητας με χαμηλό κόστος σε όλα τα μέρη του συστήματος (ΣΕΡ και ΣΚΤ). Οι ερευνητικοί στόχοι του STREAMSTORE απαιτούν την εξέλιξη της ελαστικότητας, προβλεψιμότητας, και διαθεσιμότητας των ΣΕΡ και ΣΚΤ πέραν της τρέχουσας τεχνολογικής στάθμησης. Το STREAMSTORE στοχεύει στη δημιουργία ενός ελαστικού (ταχέως κλιμακώσιμου) ΣΚΤ που αξιοποιεί τις αναδυόμενες αρχιτεκτονικές διακομιστών και είναι πλήρως συντονισμένο με τρέχοντα ελαστικά, κλιμακώσιμα ΣΕΡ. Η υλοποίηση αυτών των στόχων απαιτεί την επέκταση μοντέρνων τεχνικών όπως προοδευτική ελαστικότητα (incremental elasticity) και πρόβλεψης απόδοσης σε ΣΕΡ και ΣΚΤ και ευθυγράμμισής τους με αναδυόμενες τεχνολογίες συστοιχιών διακομιστών.

## Πρωτοτυπία του Ερευνητικού Έργου

Το έργο STREAMSTORE προάγει την τεχνολογική στάθμιση στοχεύοντας στον συντονισμό των δράσεων ΣΕΡ και ΣΚΤ κατά τη διάρκεια φάσεων ελαστικότητας και στην δημιουργία νέων τεχνικών για την ελαστική διαχείριση κατανεμημένων ΣΚΤ τύπου NoSQL. Αρκετά ΣΚΤ υποστηρίζουν υποτυπώδεις δράσεις αναδιοργάνωσης (ελαστικότητας) οι οποίες ωστόσο είναι χρονοβόρες και απαιτητικές σε πόρους όταν αλλάζει ο αριθμός των διακομιστών που αποθηκεύουν δεδομένα. ΣΕΡ σε επιχειρησιακά περιβάλλοντα υπόκεινται σε δραστικές αλλαγές φόρτου της τάξης 2x-3x. Ενώ τα ΣΕΡ μπορούν να αναδιοργανωθούν ταχέως με νέους πόρους επεξεργασίας, η αναδιοργάνωση των ΣΚΤ διαρκεί περισσότερο και έχει αρνητικό αντίκτυπο στην απόδοση των εφαρμογών, δημιουργώντας την ανάγκη για καλύτερη υποστήριξη ελαστικότητας στα ΣΚΤ και στην καλύτερη ενσωμάτωση και συντονισμό μεταξύ ΣΚΤ και ΣΕΡ. Το STREAMSTORE προτείνει την χρήση τεχνολογίας προοδευτικής ελαστικότητας (incremental elasticity) για την προαγωγή της ελαστικότητας των ΣΕΡ, ένα ερευνητικό πεδίο στο οποίο η επίδραση περιορισμών των ΣΚΤ δεν έχει μελετηθεί σε βάθος. Το STREAMSTORE θα μελετήσει την μόχλευση καινοτόμων δυνατοτήτων των πολυπύρηνων διακομιστών (υψηλό παραλληλισμό, δίκτυα τύπου RDMA, solid-state τεχνολογίες αποθήκευσης), τα οποία αναμένεται να υποστηρίξουν ιδανικά τις απαιτήσεις της προοδευτικής ελαστικότητας. Το STREAMSTORE θα προάγει επίσης την τεχνολογική στάθμιση στον χώρο της πρόβλεψης απόδοσης κλιμακώσιμων ΣΕΡ και ΣΚΤ, εστιάζοντας την εφαρμογή τεχνικών μηχανικής μάθησης όπως support vector regression (SVR), artificial neural network (ANN) regression, multivariate adaptive regression splines (MARS), οι οποίες μελετώνται από την ομάδα έργου για την πρόβλεψη απόδοσης κατανεμημένων ΣΚΤ. Το STREAMSTORE θα μοχλεύσει επίσης τεχνολογίες αιχμής όπως RDMA για να βελτιώσει την απόδοση της ενοποιημένης ΣΕΡ/ΣΚΤ πλατφόρμας.

## Αναμενόμενα αποτελέσματα & Αντίκτυπος του Ερευνητικού Έργου

Η έρευνα στο έργο STREAMSTORE προάγει την τεχνολογική στάθμιση μέσω της ενσωμάτωσης και του συντονισμού δράσεων ελαστικότητας ΣΕΡ και ΣΚΤ στα οποία βασίζεται η αποθήκευση και χρήση γνώσης και πληροφορίας. Το κύριο αποτέλεσμα του STREAMSTORE θα είναι μια κλιμακώσιμη και ελαστική πλατφόρμα ΣΕΡ και ΣΚΤ η οποία θα μοχλεύει νέες τεχνολογίες δικτύωσης και αποθήκευσης με στόχο την τεχνολογική ετοιμότητα βαθμίδας TRL 4. Ο οικονομικός και κοινωνικός αντίκτυπος του STREAMSTORE αναμένεται να είναι σημαντικός, ιδιαίτερα στους τομείς της κοινωνικής δικτύωσης (χώρος που μεγάλες εταιρίες όπως οι Google, Twitter, και LinkedIn έχουν εφαρμόσει αντίστοιχες τεχνολογίες), στην επισκόπηση γεγονότων σε πραγματικό χρόνο για διάγνωση γεγονότων ενδιαφέροντος και οικονομικής δραστηριότητας (high-speed financial trading), και άλλες εφαρμογές στις οποίες η επεξεργασία γεγονότων μεγάλου όγκου και ρυθμού, σε συνδυασμό με μεγάλες ποσότητες αποθηκευμένης πληροφορίας και περιόδους υψηλής διακύμανσης εισερχόμενου φόρτου ροών, πρέπει να είναι διαχειρίσιμη με τρόπο που να σιγουρεύει ότι η απόκριση δεν επηρεάζεται αρνητικά. Τα αποτελέσματα του STREAMSTORE θα βελτιώσουν την ικανότητα των συστημάτων να χειρίζονται διακυμάνσεις κοινωνικής ψηφιακής δραστηριότητας και να προσφέρουν ταχεία διάγνωση γεγονότων και τάσεων με χαμηλή απόκριση. Τα ερευνητικά αποτελέσματα του STREAMSTORE αναμένεται να έχουν περαιτέρω επίδραση σε περιοχές εφαρμογών πέραν της επεξεργασίας ροών όπως η υλοποίηση συμφωνιών επιπέδου υπηρεσιών (SLAs) για κλιμακώσιμα αποθηκευμένα δεδομένα.

## Η σημασία της χρηματοδότησης

Η χρηματοδότηση μέσω των ερευνητικών προγραμμάτων για Μέλη ΔΕΠ και Ερευνητές του ΕΛ.ΙΔ.Ε.Κ. για ελεύθερη έρευνα (Blue Sky research) έδωσε στον ΕΥ και την Ερευνητική του Ομάδα την δυνατότητα να εστιάσουν σε έρευνα αιχμής στο πεδίο των αναδυομένων τεχνολογιών κατανεμημένων συστημάτων, επιτρέποντάς τους να υλοποιήσουν ένα φιλόδοξο ερευνητικό πρόγραμμα. Όντας προσανατολισμένο στον ερευνητή και ακολουθώντας μια στρατηγική «από κάτω προς τα πάνω» (bottom-up), η ερευνητική χρηματοδότηση του ΕΛ.ΙΔ.Ε.Κ επιτρέπει στον ΕΥ και την Ερευνητική του Ομάδα να εστιάσουν τις προσπάθειές τους στην προαγωγή του ερευνητικού τους σχεδιασμού με μεγαλύτερο βαθμό αποδοτικότητας και ευελιξίας από ότι ήταν εφικτό στο παρελθόν.



**ΕΛΙΔΕΚ.**  
Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας

## ΕΠΙΚΟΙΝΩΝΙΑ

Λ. Συγγρού 185 & Σάρδεων 2  
ΤΚ. 17121, Νέα Σμύρνη, Ελλάδα  
210 64 12 410, 420  
communication@elidek.gr  
[www.elidek.gr](http://www.elidek.gr)