

Θέματα Εξαμηνιαίας Εργασίας

Δ. ΤΣΟΥΜΑΚΟΣ

Elasticity - Cloud Computing

Είναι η ικανότητα ενός συστήματος να αυξάνει και να μειώνει τους πόρους που χρησιμοποιεί προκειμένου να λειτουργεί σύμφωνα με μια πολιτική του χρήστη. Το elasticity αναφέρεται κυρίως σε Cloud computing platforms/apps στις μέρες μας.

Θέματα σχετικά με elasticity:

α) Μελέτη ελαστικότητας μιας NoSQL DB (nuodb, riak, voldemort, couchbase, titan, ...), δείτε π.χ. εδώ: <http://www.vldb.org/pvldb/vol7/p1219-klems.pdf>

β) Υλοποίηση συστήματος αυτόματης εφαρμογής ελαστικότητας
δείτε π.χ. εδώ: http://www.cslab.ece.ntua.gr/~dtsouma/index_files/tiramola_CCGrid13.pdf

γ) μέτρηση της διαφοράς ανάμεσα σε load balancing NoSQL vs. elasticity (Hbase compaction,...)

δ) Επέκταση του TIRAMOLA σε RWU workloads: <http://code.google.com/p/tiramola/>

ε) μελέτη της επίδρασης του migration VM/containers vs elasticity

Big Data

Σχετικά θέματα:

α) Υλοποίηση μοντέλου κόστους (Cost model) για το πόσο κοστίζουν σε I/O, χρόνο, κλπ, διάφορες λειτουργίες (π.χ. joins) σε Big Data Engines (Spark, MonetDB, etc) όπως π.χ. στην εργασία εδώ: <http://dl.acm.org/citation.cfm?id=2742792>

β) Σύγκριση απόδοσης απλών τελεστών σε δεδομένα γράφου (φόρτωση δεδομένων, BFS, DFS, Shortest paths, etc) ανάμεσα σε hama - titan - neo4j - hadoop (ανά 2 ή ανά 3)

γ) σχεδιασμός και υλοποίηση caching αλγορίθμου σε δεδομένα γράφων

δ) Σύγκριση απόδοσης σύνθετων τελεστών (π.χ. pagerank, clustering, regression, etc) πάνω από spark και hadoop (yarn)

ε) Επέκταση του H₂RDF: <http://code.google.com/p/h2rdf/>

στ) Σύγκριση απόδοσης και δυνατοτήτων των SQL queries σε Big Data Engines (Spark, Hive, Hadoop)

ζ) Εισαγωγή μεγάλου όγκου κώδικα (π.χ. όλα τα github java projects) με μορφή γράφου (συντακτικούς γράφους) σε NoSQL βάση δεδομένων και επερωτήσεις επί αυτής

P2P

Υλοποίηση εφαρμογών

Σχετικά θέματα:

- a) Εύρεση συσχετίσεων ανάμεσα σε τυχαία datasets (available online) και αναπαράστασή τους.
- b) ελαστική λειτουργία μιας συγκεκριμένης εφαρμογής πάνω από υποδομή cloud
- γ) εύρεση της συχνότητας αλλαγής δεδομένων κλίμακας (π.χ. web pages, tweets, facebook profiles etc) και εξαγωγή γνώσης