

Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχ. Και Μηχ. Υπολογιστών

Προηγμένα Θέματα Αρχιτεκτονικής Υπολογιστών

8^ο εξάμηνο, Ροή Υ

Ακαδημαϊκή Περίοδος : 2011-2012



1^η Εργασία

Γερακάρης Βασίλης
<vgerak@gmail.com>
Α.Μ. :03108092

Εισαγωγικά:

Σκοπός της άσκησης αυτής ήταν η εξοικίωση με το λογισμικό προσομοίωσης συστημάτων Simics της Wind River, και η μελέτη της επίδρασης διαφόρων παραμέτρων των μνημών L1 και L2 cache στην απόδοση συγκεκριμένων προγραμμάτων αναφοράς (benchmarks)

Πιο συγκεκριμένα, σε εικονικό περιβάλλον Fedora Linux που λειτουργούσε μέσω simics, εξετάσαμε τις μεταβολές στην επίδοση των προγραμμάτων art, gzip, equake και mcf, χρησιμοποιώντας ως δείκτη απόδοσης το IPC (Instructions Per Cycle) και Miss Rate.

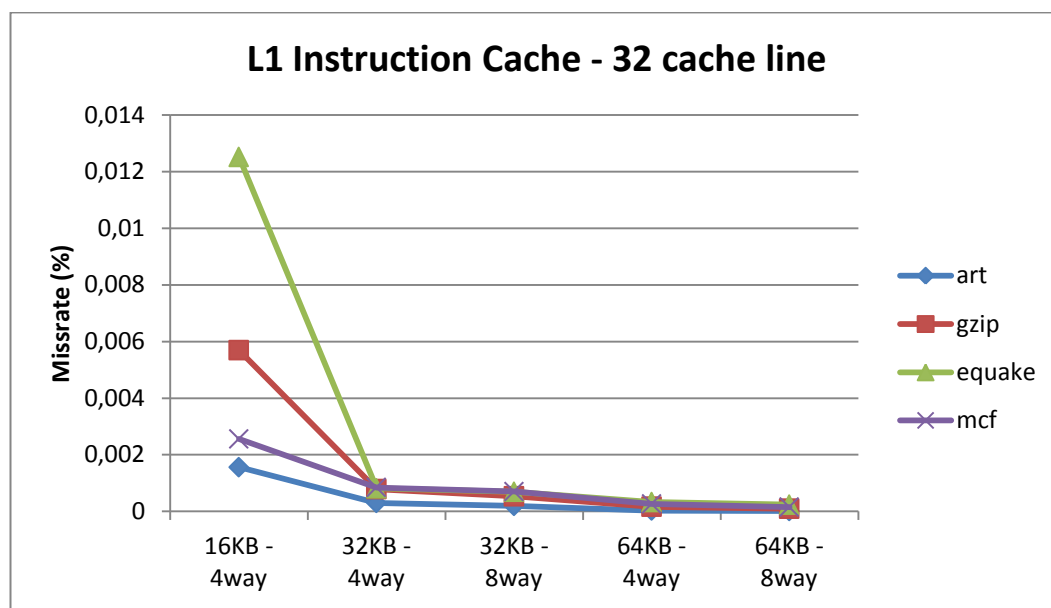
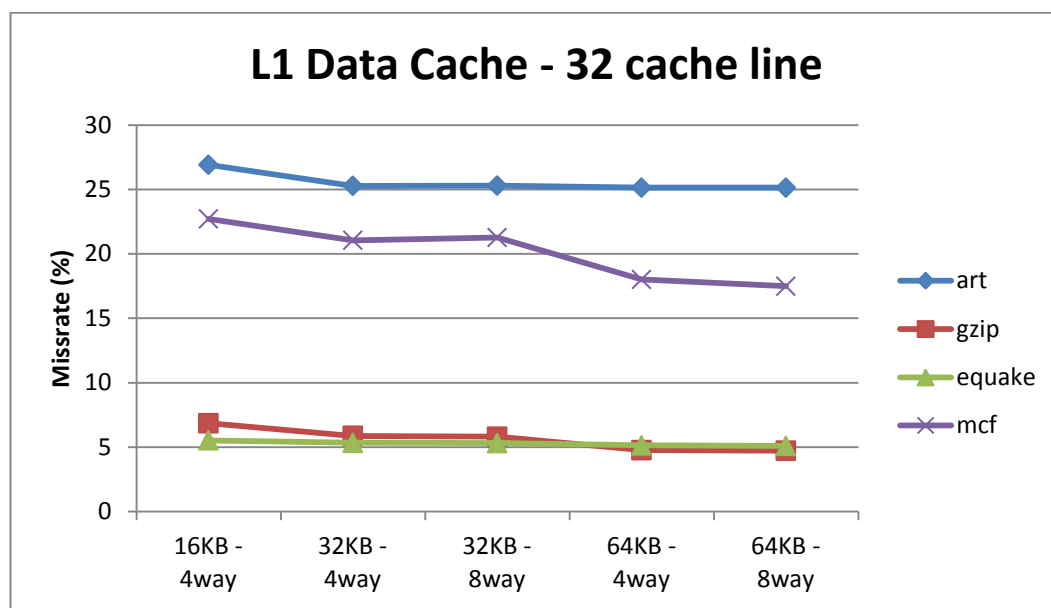
Μέρος 1^ο – L1 Cache

Στο τμήμα αυτό της άσκησης, κρατώντας σταθερά τα μεγέθη της L2 cache, παρατηρήσαμε την επίδραση που έχουν το μέγεθος γραμμής (line size), ο αριθμός γραμμών (line number) και η συσχέτιση (associativity) στο ποσοστό των Misses της L1 cache, καθώς και στο συνολικό IPC.

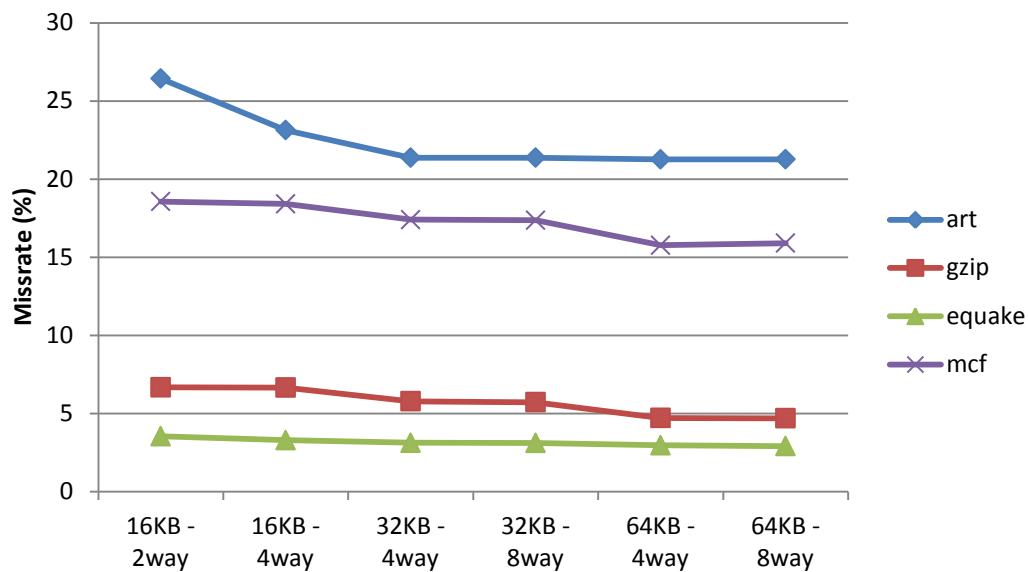
Από τα αποτελέσματα των προσομοιώσεων συμπεραίνουμε εύκολα ότι η επίδραση της L1 Instr Cache είναι αμελητέα, τόσο στο ποσοστιαίο Miss Rate, όσο και στο IPC. Παρ'όλα αυτά, αφού ζητείται από την εκφώνηση της άσκησης, παρατίθενται μερικά, ώστε να φανεί αυτό το συμπέρασμα.

Ακολουθούν τα διαγράμματα για L1 Miss rate και IPC, όπως προκύπτουν από τα αποτελέσματα των προσομοιώσεων:

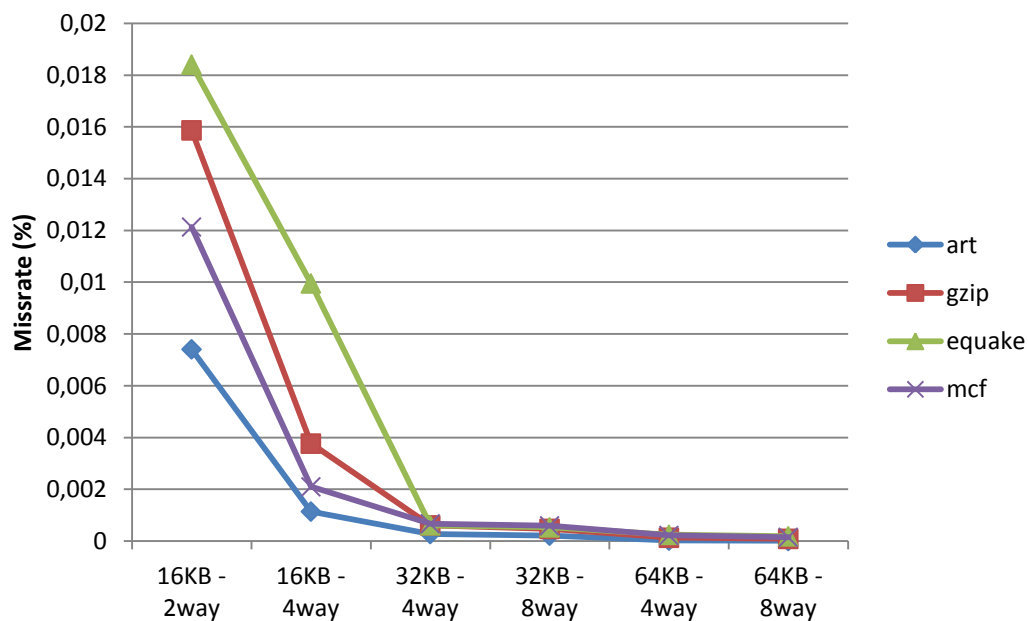
4.1.1 α) Επίδραση στο Miss Rate %:

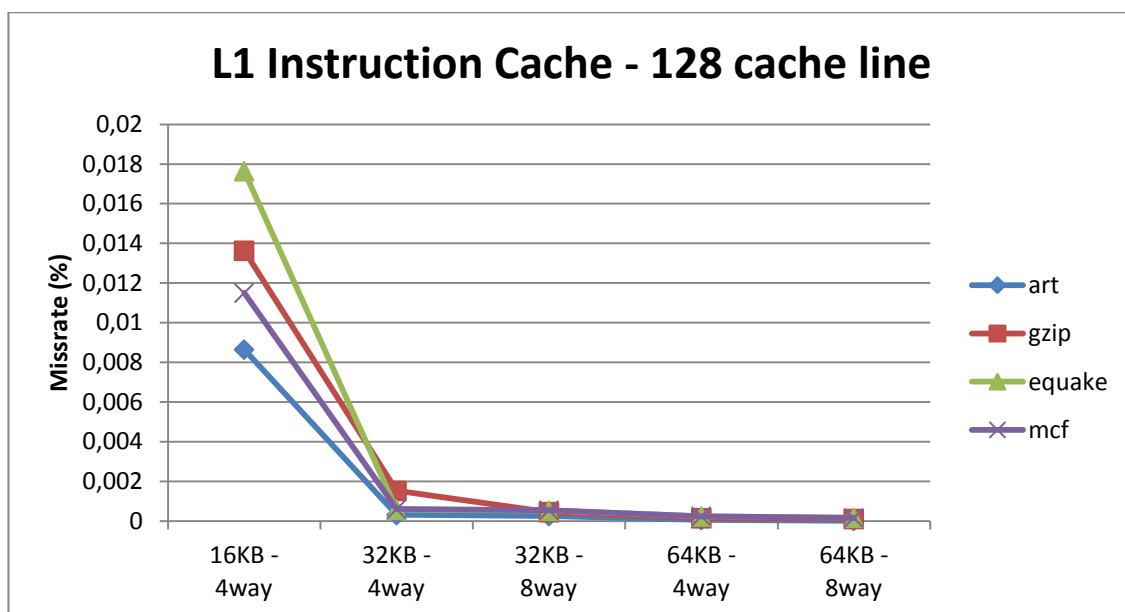
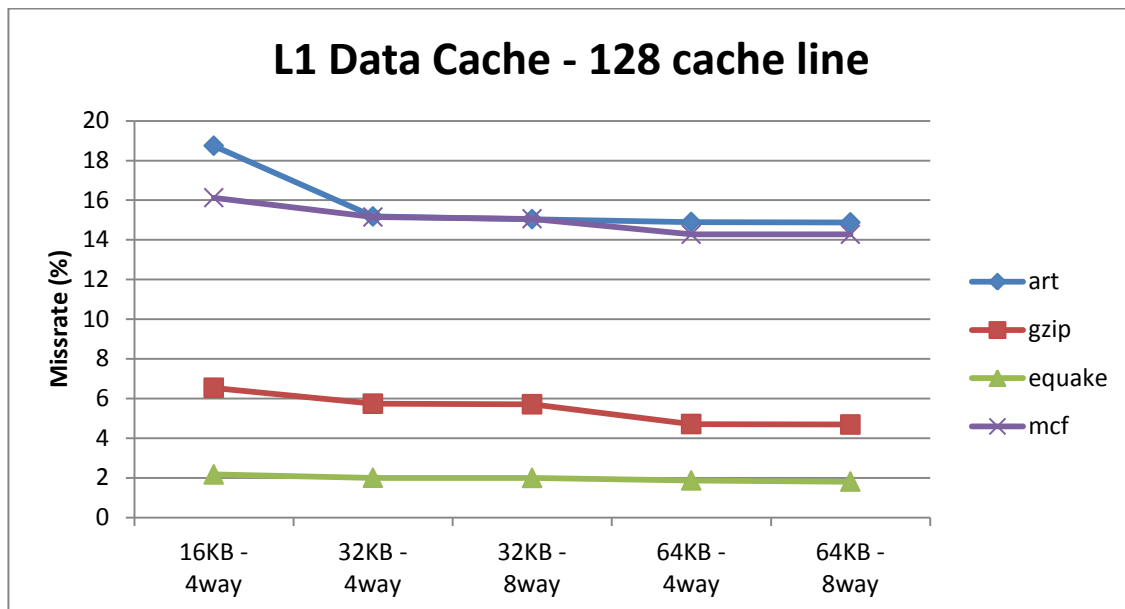


L1 Data Cache - 64 cache line



L1 Instruction Cache - 64 cache line





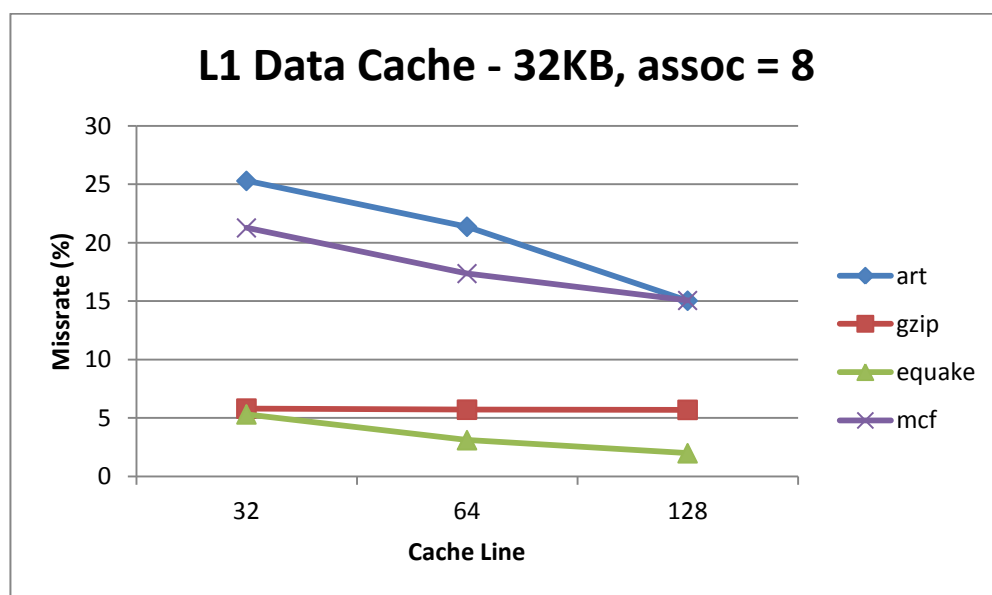
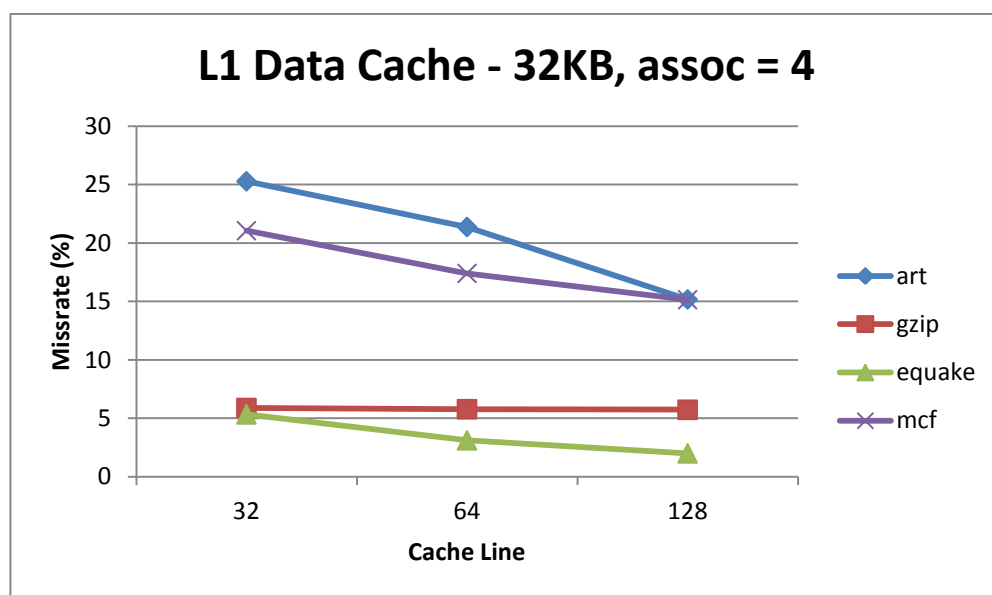
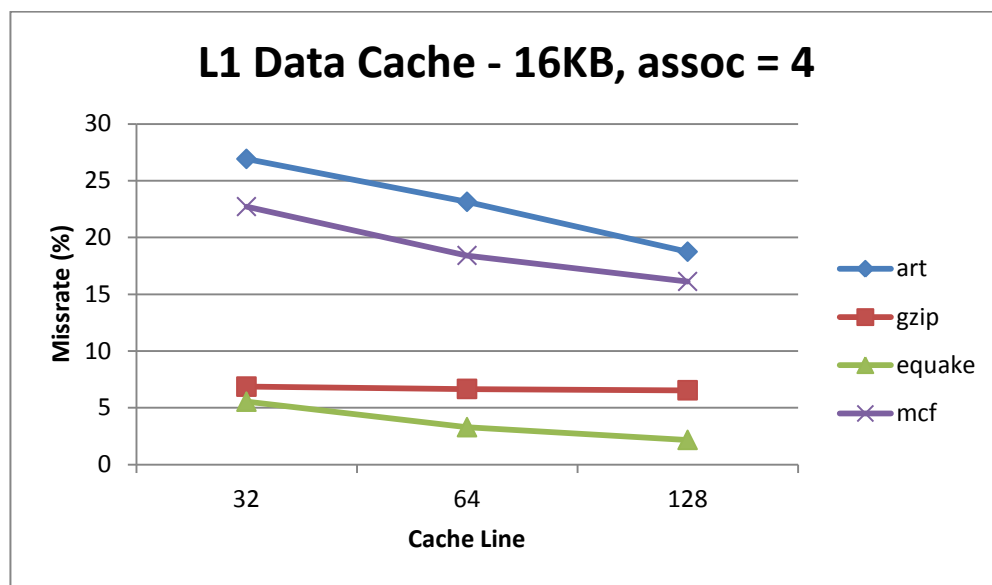
Από τα παραπάνω διαγράμματα, συμπεραίνουμε ότι αυξάνοντας το μέγεθος της L1 cache, μειώνεται το ποσοστό των misses, όπως ήταν αναμενόμενο.

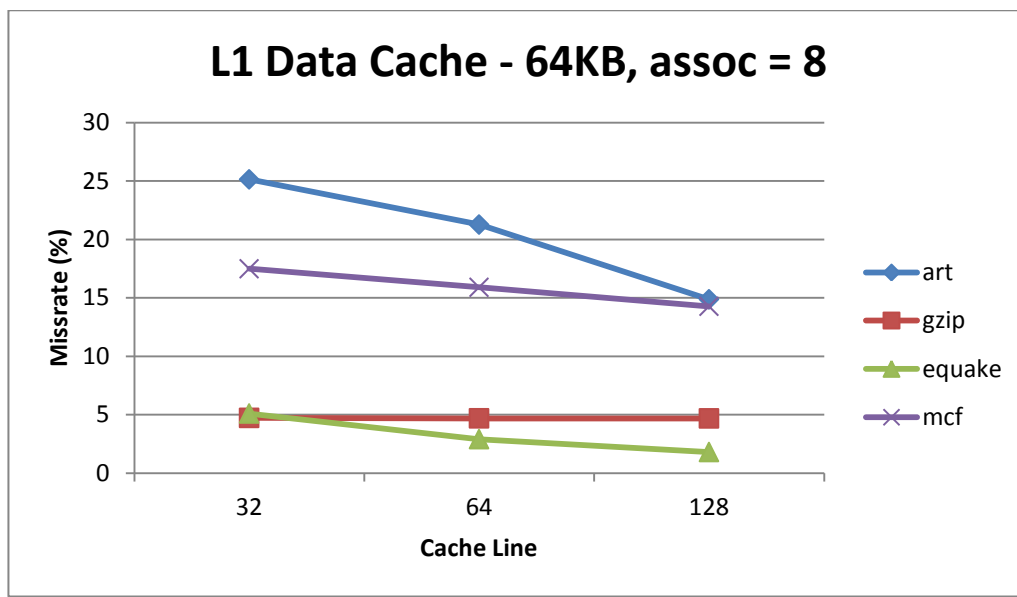
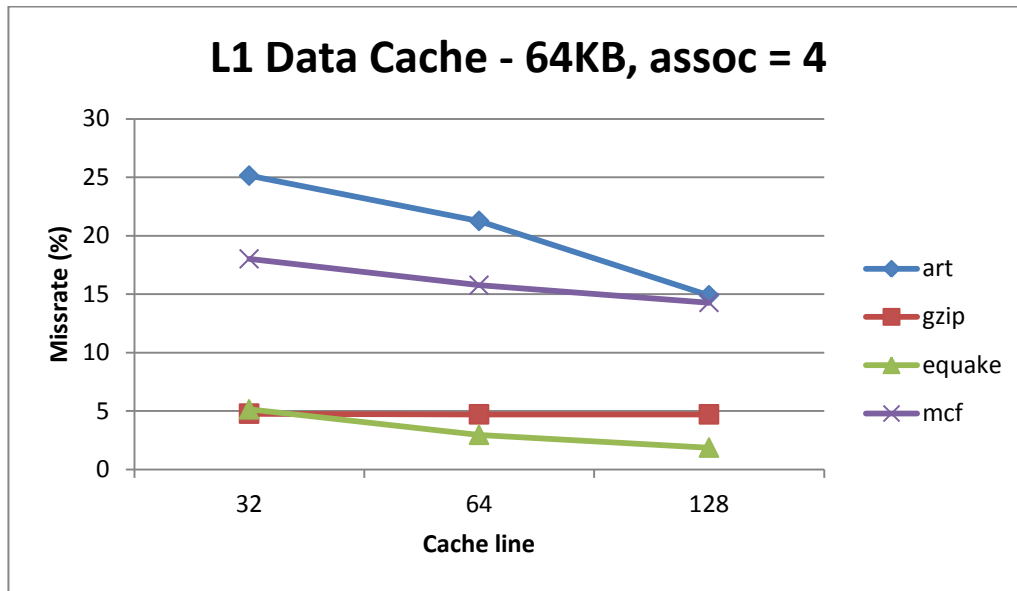
Πιο συγκεκριμένα, το art δέχεται τη μεγαλύτερη βελτίωση απόδοσης καθώς περνάει σε μέγεθος L1 cache μεγαλύτερο των 16k, και μπορούμε να υποθέσουμε ότι το working set του βρίσκεται περίπου στα 32k.

Αντίθετα, το quake φαίνεται να μην επηρεάζεται σημαντικά από τις μεταβολές της cache size.

Και για τα 4 benchmarks, παρατηρούμε ότι οι αλλαγές στο associativity δεν επιφέρουν σημαντικές βελτιώσεις στην απόδοση.

Επίδραση του Line Size:

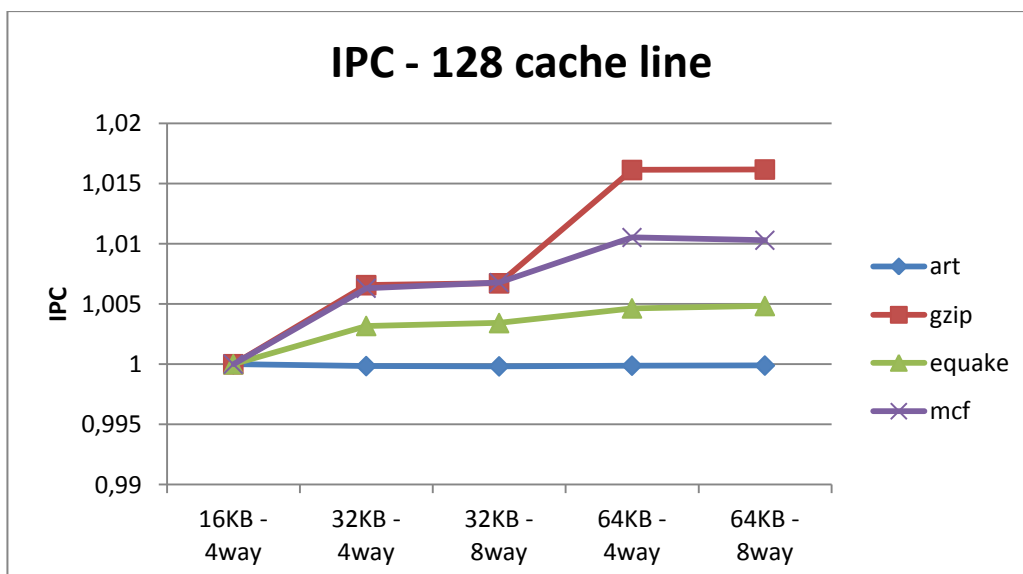
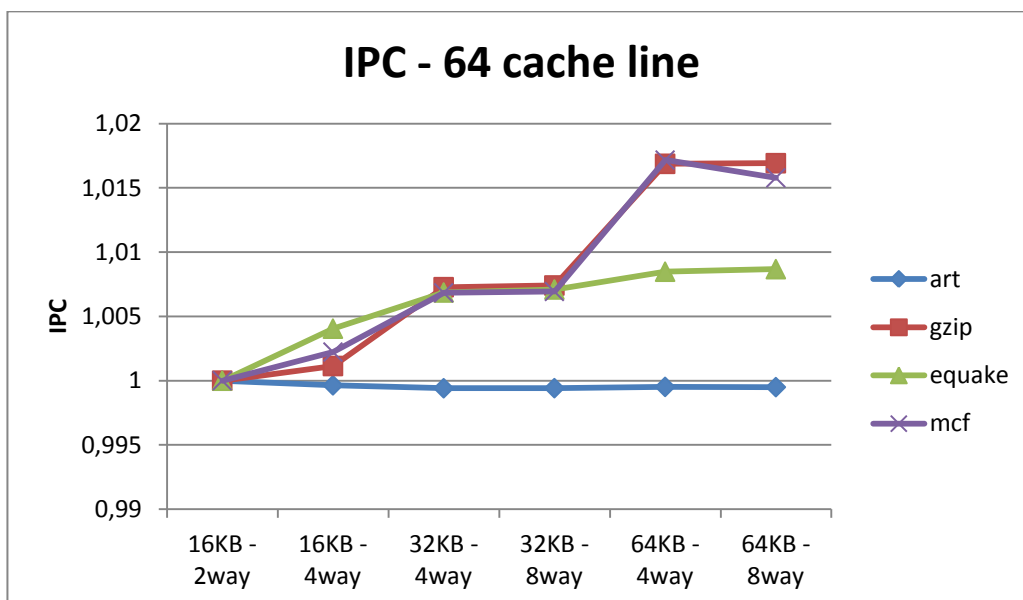
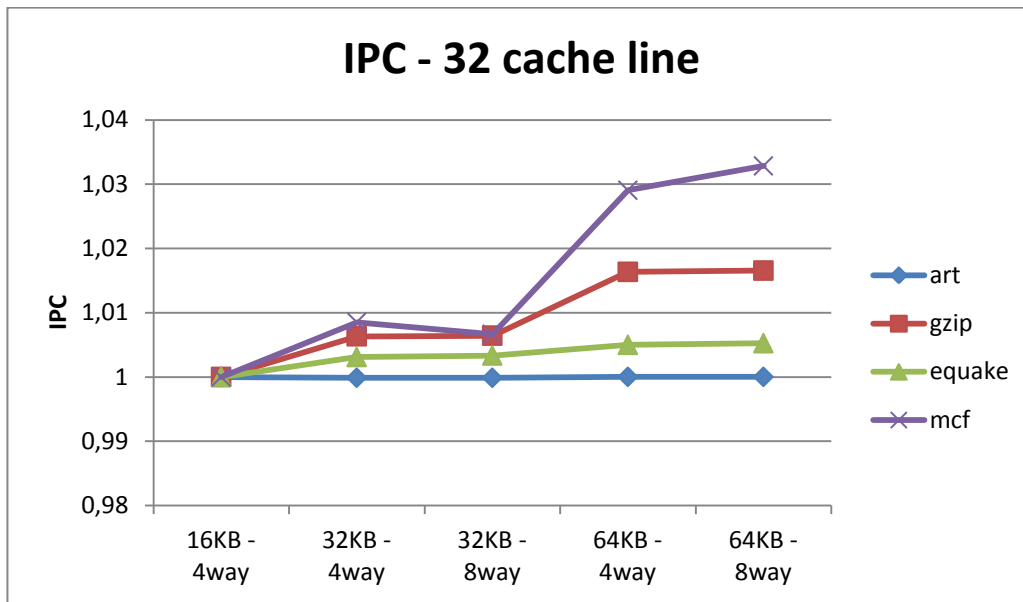




Από τα παραπάνω διαγράμματα προκύπτει ότι για όλα τα benchmarks εκτός του gzip ο αριθμός των misses μειώνεται αρκετά με την αύξηση του line size.

Αυτό σημαίνει ότι τα προγράμματα αυτά παρουσιάζουν μεγάλη τοπικότητα αναφορών, έτσι εμεταλλεύονται τη μεταφορά μεγάλου τμήματος δεδομένων προς την L1 cache, σε κάθε αναφορά στη μνήμη.

4.1.1 β) Επίδραση στο IPC:



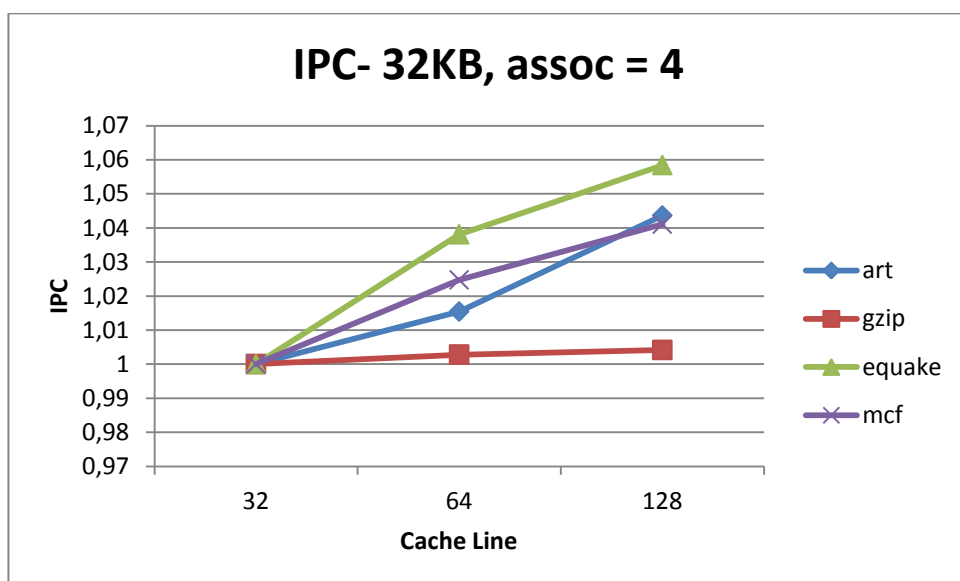
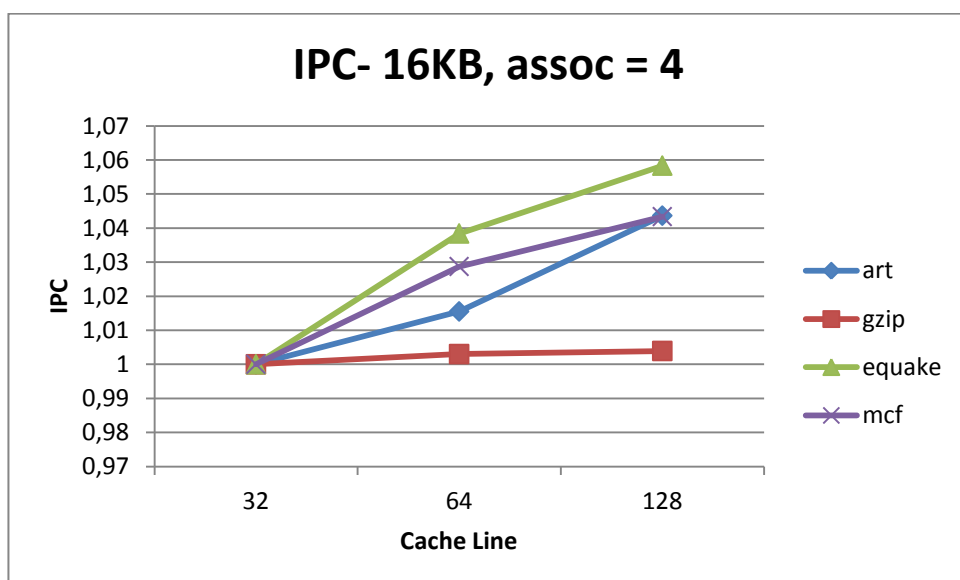
(Σε όλες τις γραφικές παραστάσεις έχει γίνει κανονικοποίηση των αποτελεσμάτων με βάση την 1^η τιμή του μεταβλητού μεγέθους)

Στα παραπάνω διαγράμματα φαίνεται ότι η αύξηση του cache size επιφέρει βελτίωση του IPC στα benchmarks gzip, equake και mcf. Το αποτέλεσμα αυτό ήταν αναμενόμενο, αφού η αύξηση του μεγέθους της μνήμης L1 cache αυξάνει την απόδοση των benchmarks.

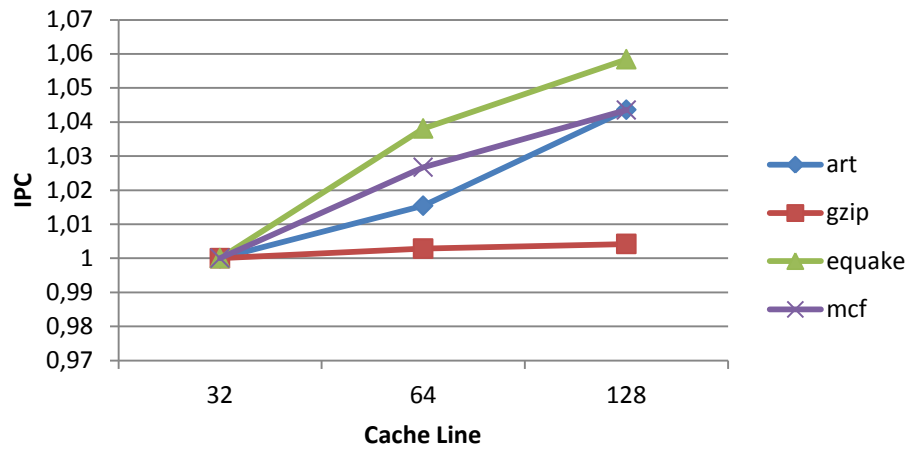
Στα benchmarks gzip και mcf, οι μεταβολές στο associativity φαίνεται να μην προκαλούν σημαντικές αλλαγές στην απόδοση, και μάλιστα παρατηρούμε ότι η αύξηση του associativity πιθανώς να επιφέρει μείωση του IPC.

Το art μένει σχεδόν ανεπηρρέαστο από τις μεταβολές που πραγματοποιήσαμε.

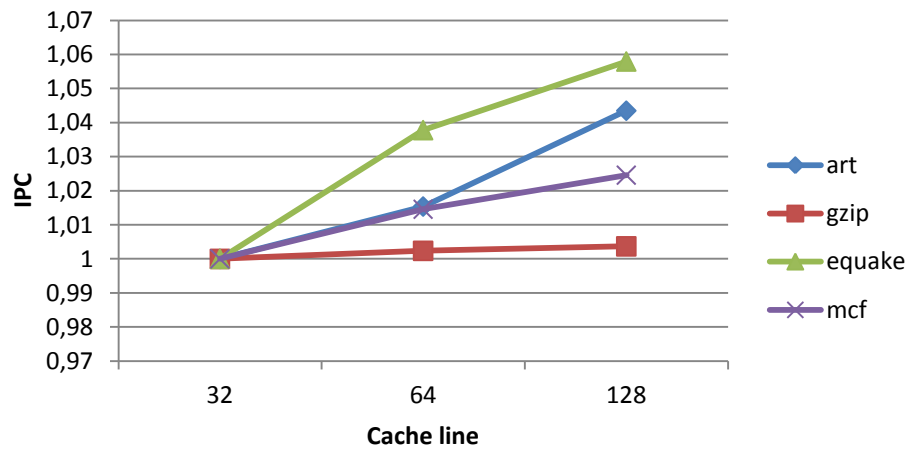
Επίδραση του Line Size:



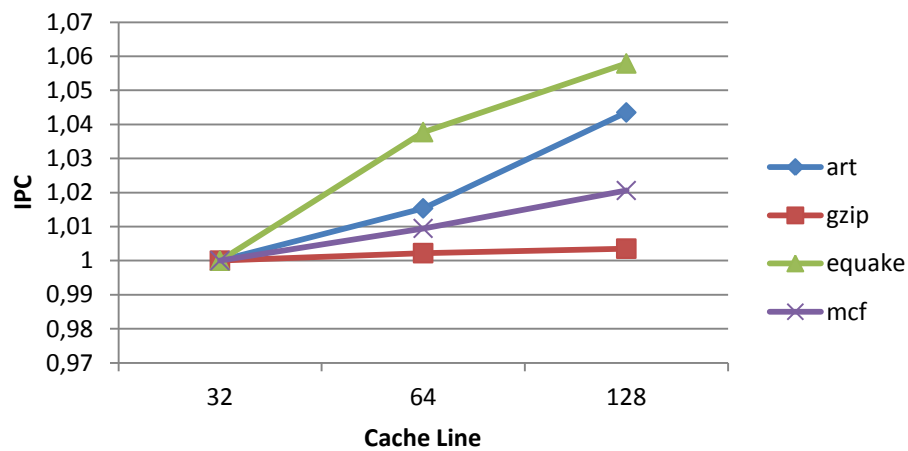
IPC- 32KB, assoc = 8



IPC- 64KB, assoc = 4



IPC- 64KB, assoc = 8



Από τα παραπάνω διαγράμματα προκύπτει ότι για όλα τα benchmarks εκτός του gzip, τα misses μειώνονται σημαντικά με αύξηση του line size, γεγονός που υποδηλώνει αρκετά μεγάλη τοπικότητα αναφορών, όπως αναφέραμε και προηγουμένως.

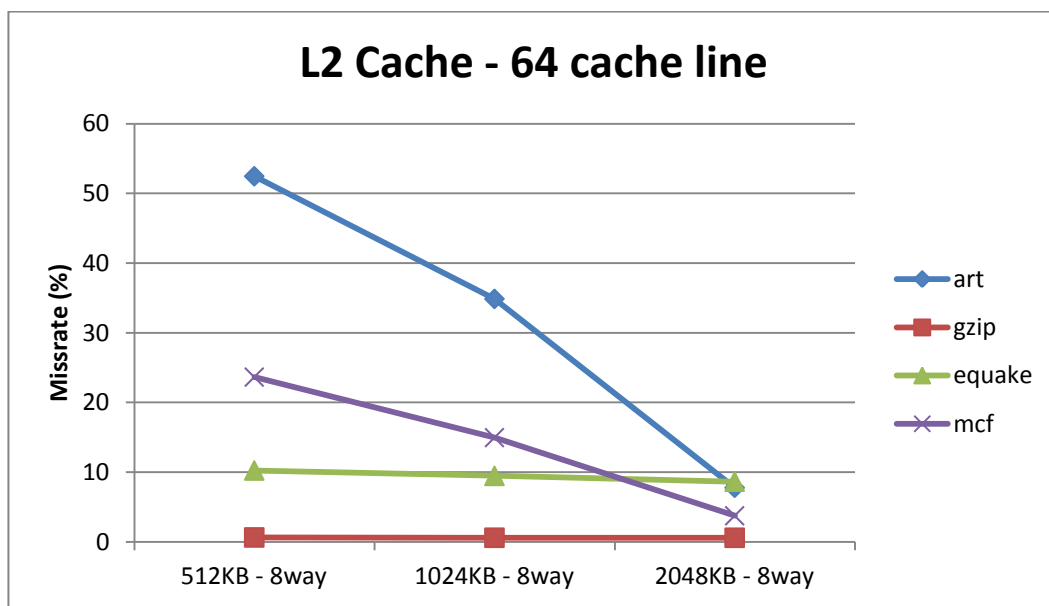
Αντίθετα, στο gzip δεν παρατηρείται σημαντική βελτίωση της απόδοσής με την αύξηση του line size.

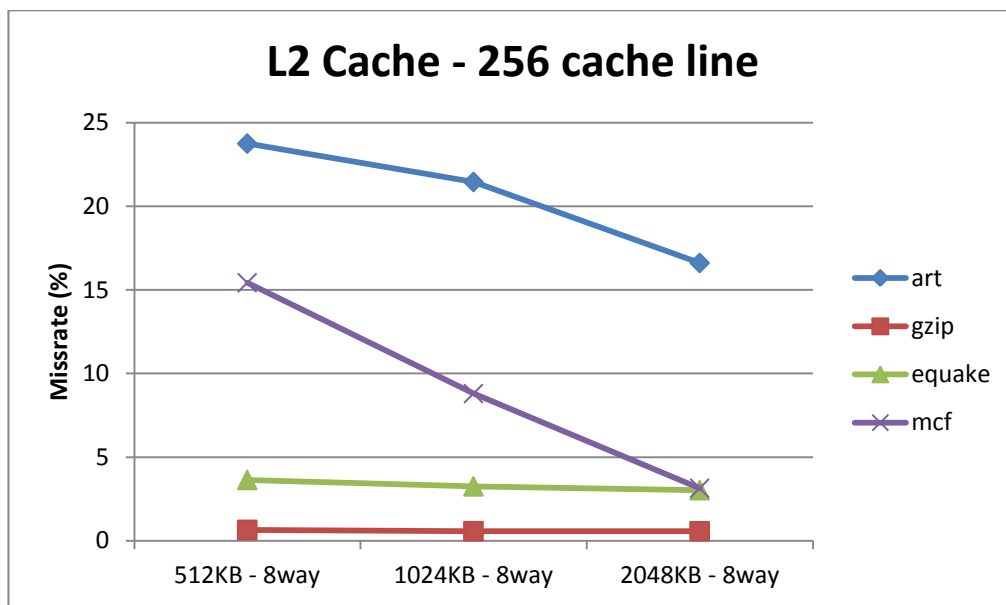
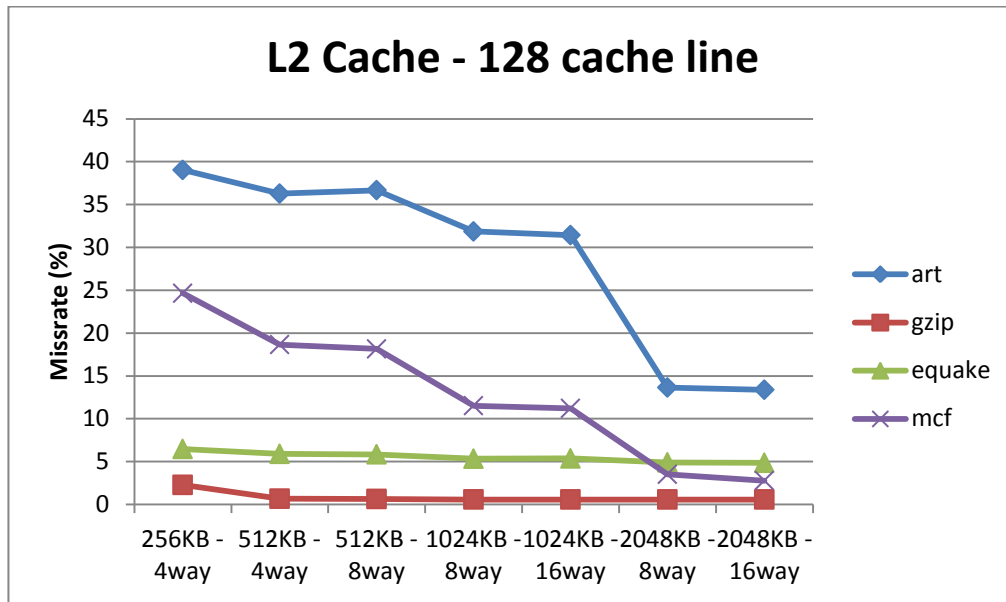
Μέρος 2^ο – L2 Cache

Στο τμήμα αυτό της άσκησης, κρατώντας σταθερά τα μεγέθη της L1 cache, παρατηρήσαμε την επίδραση που έχουν το μέγεθος γραμμής (line size), ο αριθμός γραμμών (line number) και η συσχέτιση (associativity) στο ποσοστό των Misses της L2 cache, καθώς και στο συνολικό IPC.

Ακολουθούν τα διαγράμματα για L2 Miss Rate % και IPC, όπως προκύπτουν από τα αποτελέσματα των προσομοιώσεων:

4.1.2 α) Επίδραση στο Miss Rate %:



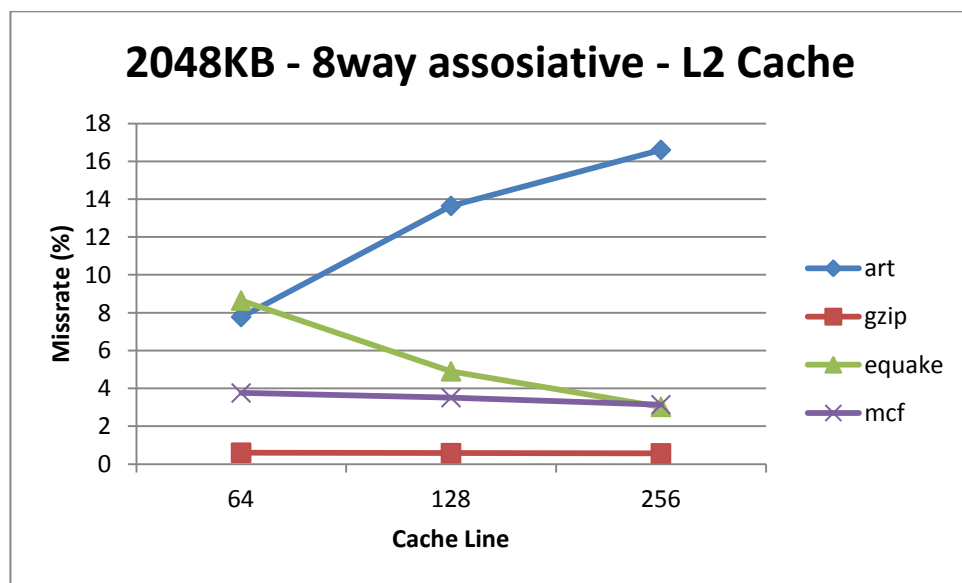
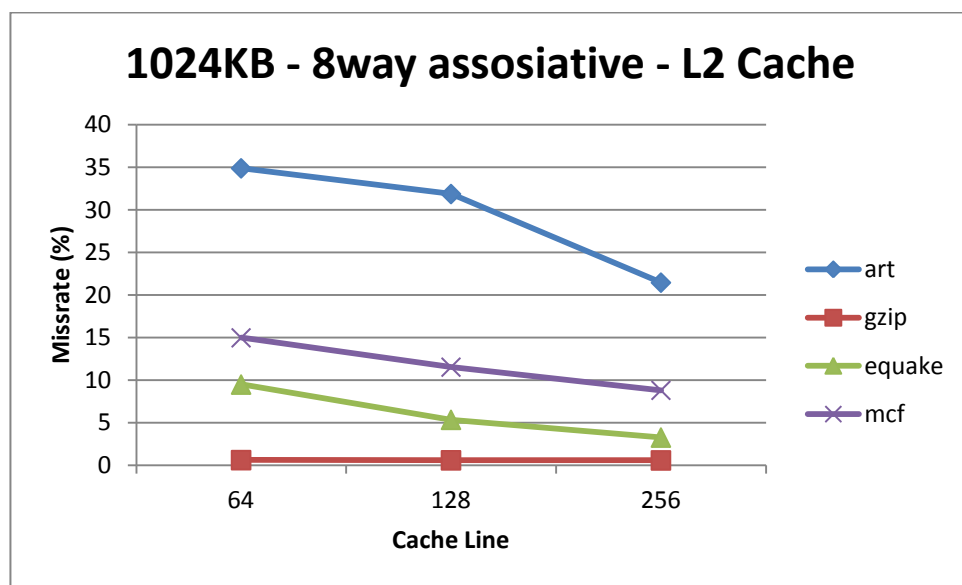
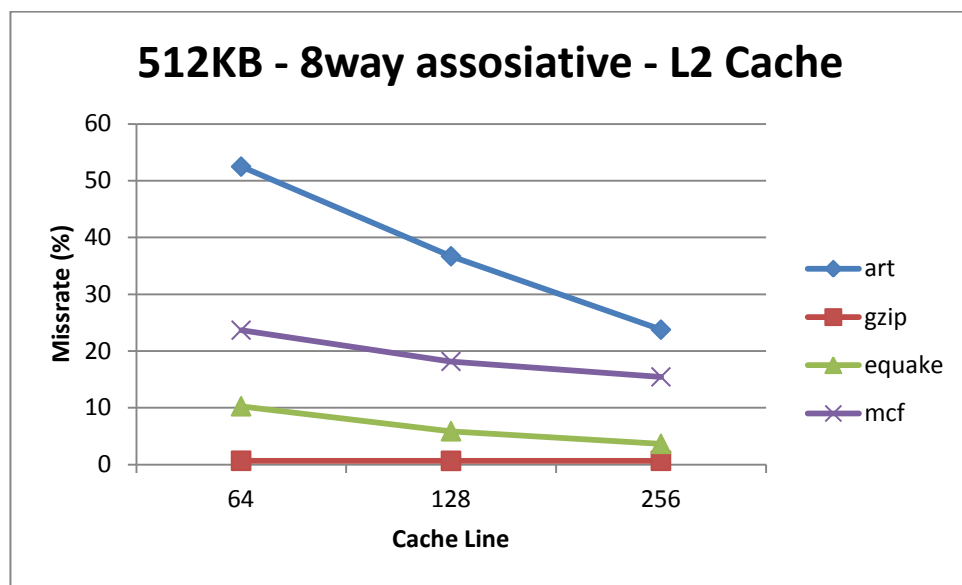


Παρατηρούμε ότι σε όλα τα benchmarks, η αύξηση του cache size συνεπάγεται μείωση του miss rate, όπως θα περιμέναμε.

Στα gzip και equake η μείωση αυτή είναι λιγότερο εμφανής, αλλά αυτό οφείλεται εν μέρει και στο ήδη αρκετά χαμηλό ποσοστό misses που παρουσιάζουν.

Παρατηρούμε επίσης ότι η μεταβολή της συσχέτισης δεν προκαλεί σημαντικές αλλαγές σε κανένα benchmarks, και ότι η μεγαλύτερη μείωση του miss rate εμφανίζεται κατά την αύξηση της L2 cache από 1024k σε 2048k.

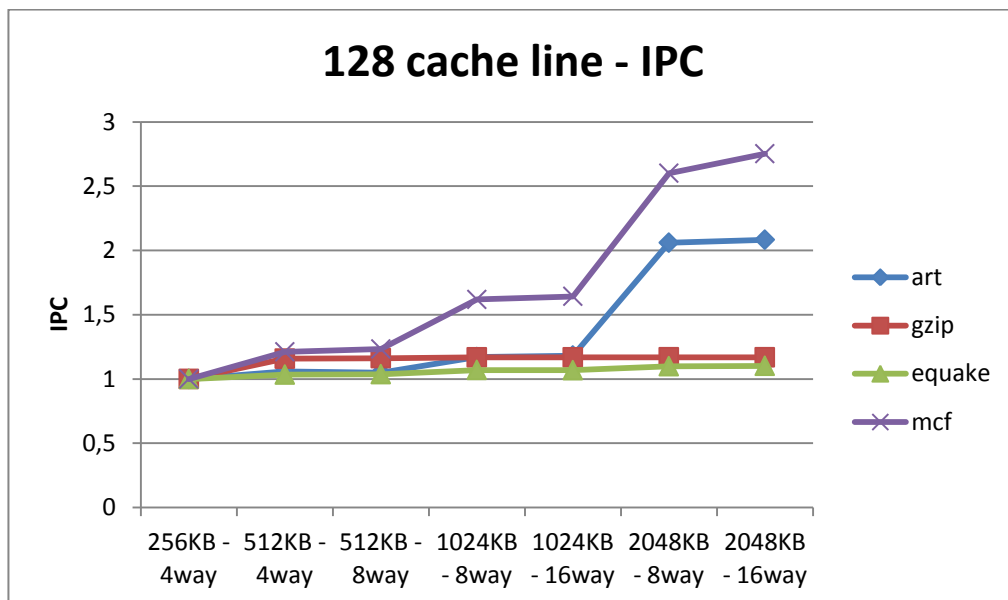
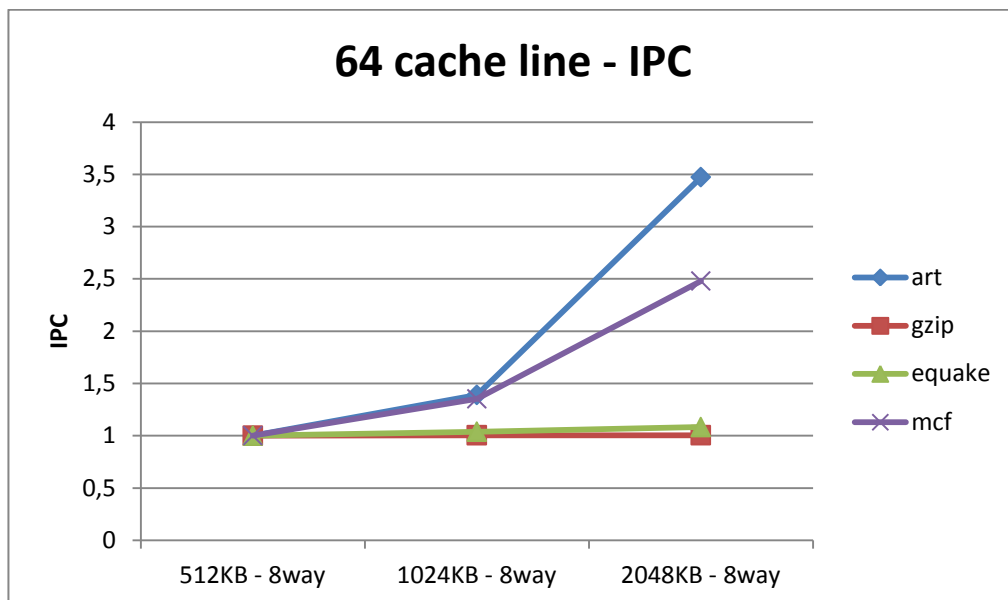
Επίδραση του Line Size:

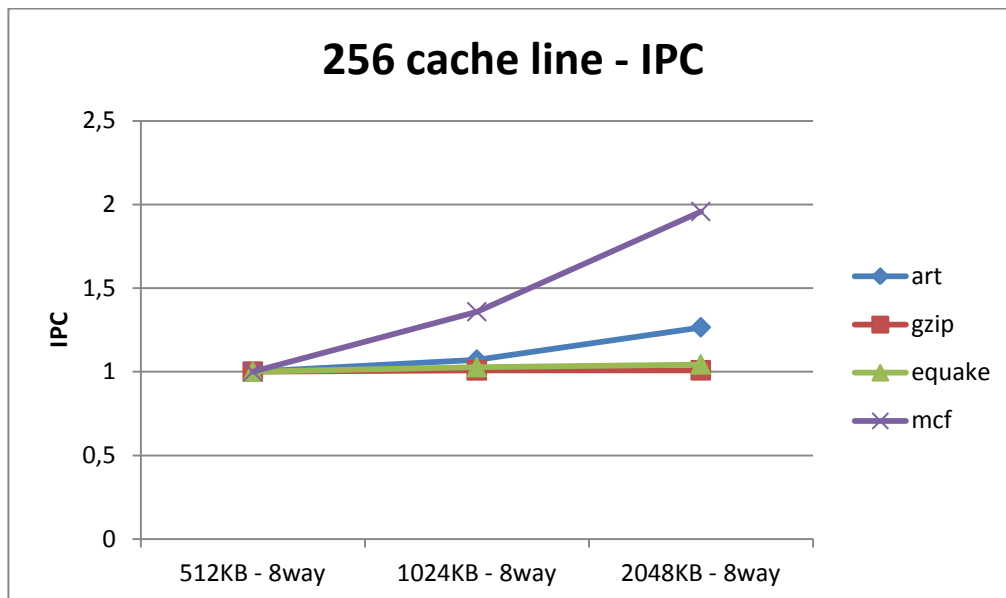


Στα παραπάνω διαγράμματα φαίνεται ότι η αύξηση του line size οδηγεί σε βελτίωση της απόδοσης των benchmarks (εκτός του gzip). Αυτό το αποτέλεσμα επιβεβαιώνει για μια ακόμη φορά την τοπικότητα που παρουσιάζουν οι αναφορές.

Στο art παρατηρείται μια μεγάλη αύξηση του miss rate για line size = 256 και cache size = 2048k.

4.1.2 β) Επίδραση στο IPC:

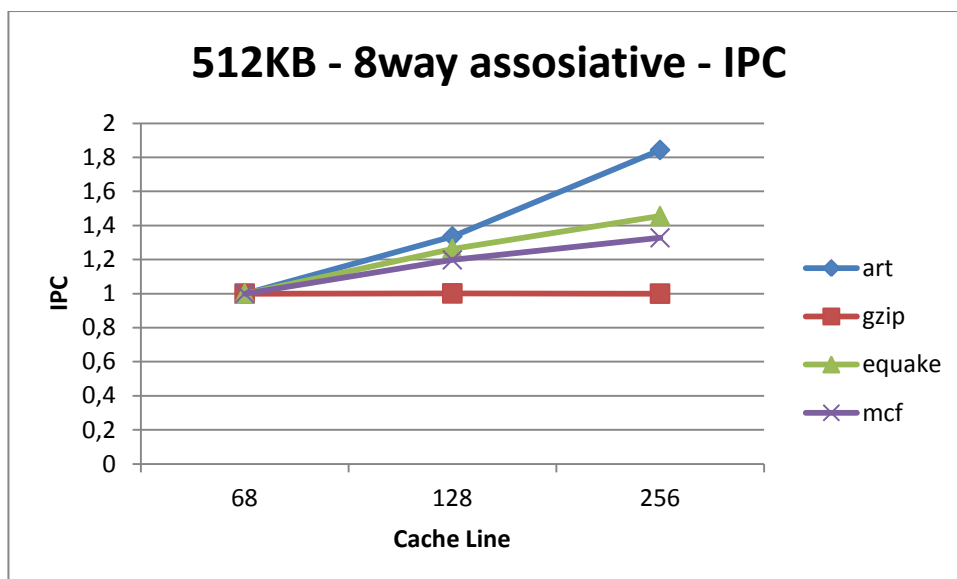


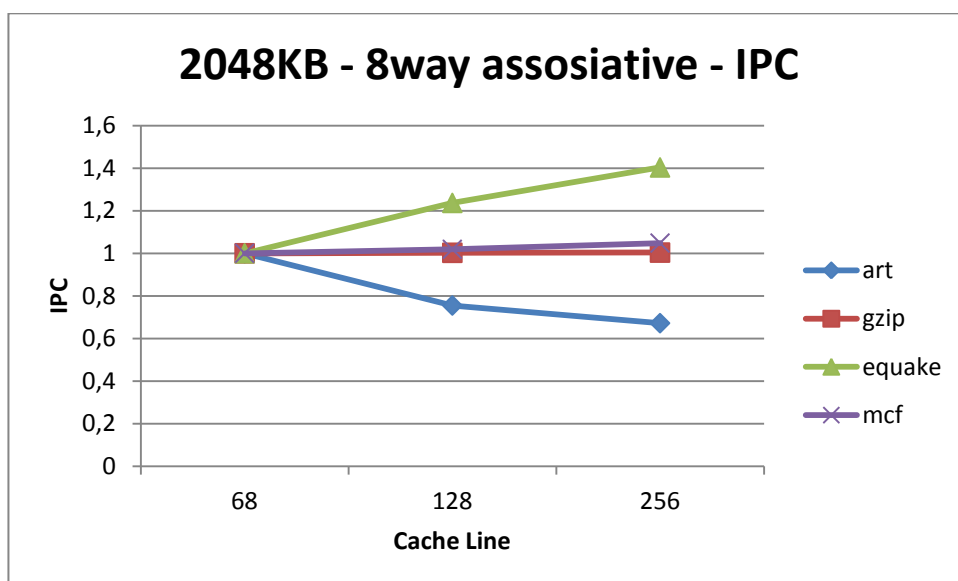
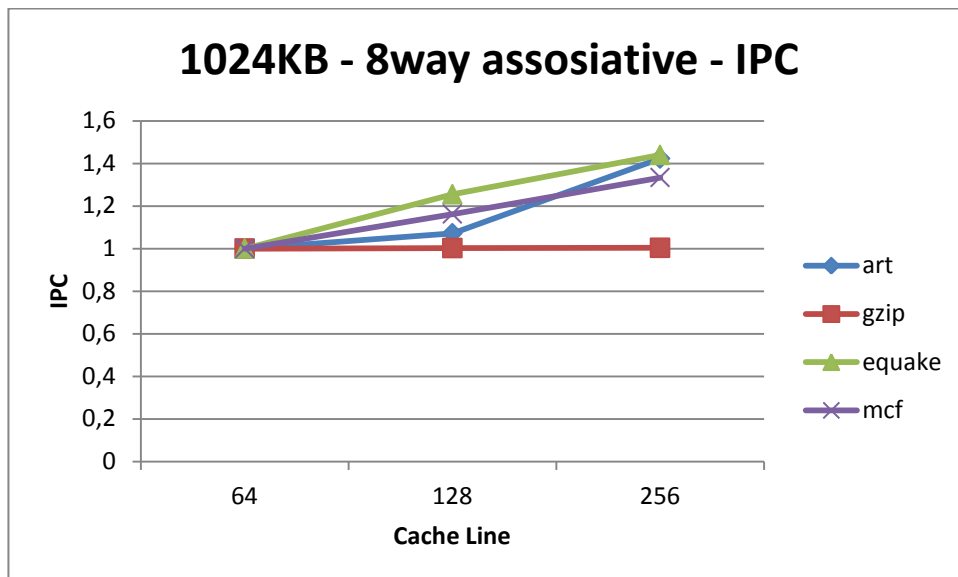


Εδώ φαίνεται ότι το IPC αυξάνεται με την αύξηση του cache size.

Όπως και πριν, παρατηρούμε ότι στα benchmarks gzip, equake δεν παρατηρείται μεγάλη βελτίωση της απόδοσης, παρα μόνο στην άνοδο από τα 256k στα 512k.

Τα art, mcf επηρεάζονται άμεσα από τη μεταβολή του μεγέθους της cache, όμως ελάχιστα από αλλαγές στη συσχέτιση.





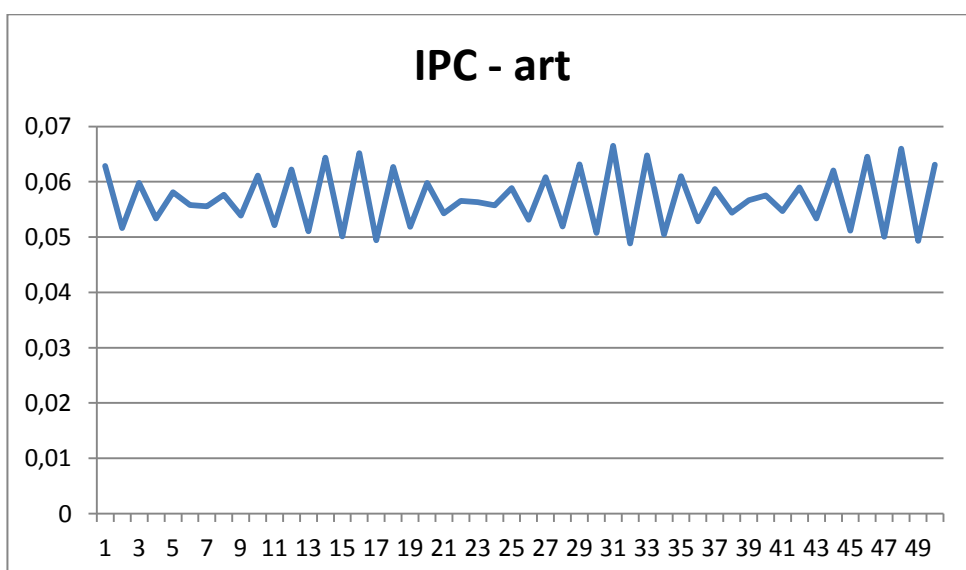
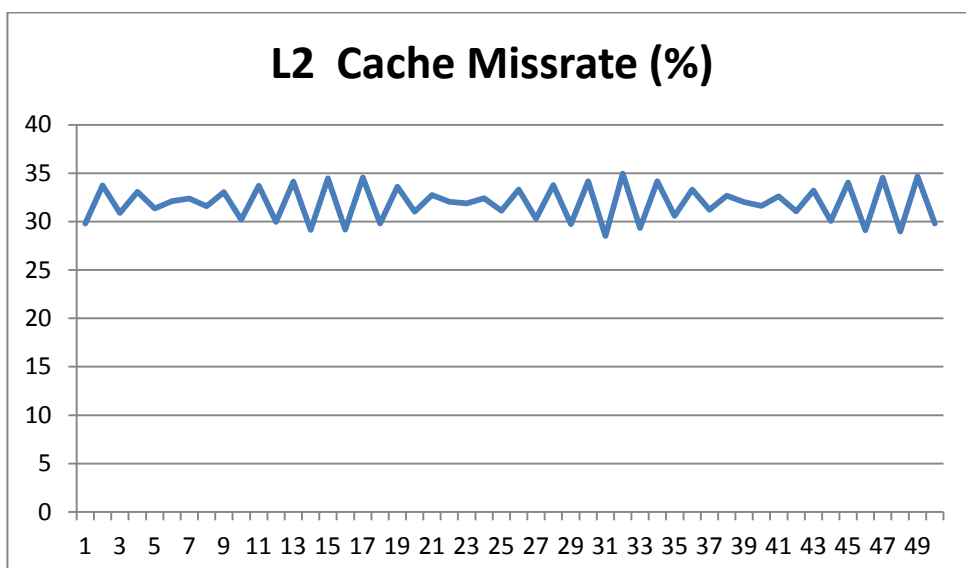
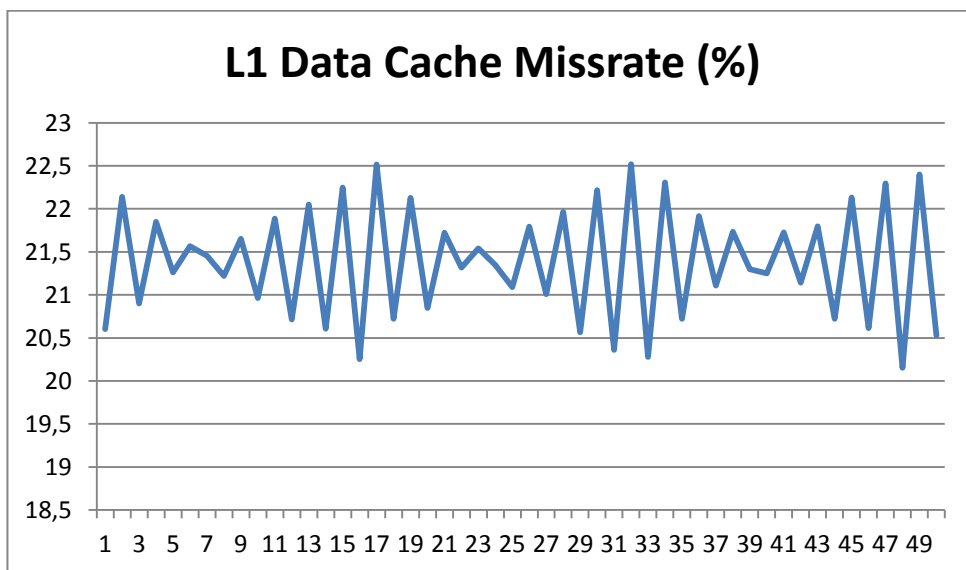
Από τα διαγράμματα αυτά παρατηρούμε ότι η αύξηση του line size επιφέρει γενικά αύξηση του IPC, γεγονός που υποδηλώνει για μία ακόμα φορά την τοπικότητα αναφορών.

Στο art, για cache size 2048k και cache line 256 παρατηρείται απότομη πτώση του IPC, όπως φάνηκε και στο προηγούμενο μέρος με την αύξηση των misses της L2, για το ίδιο configuration.

4.2 Μελέτη απόδοσης στο χρόνο

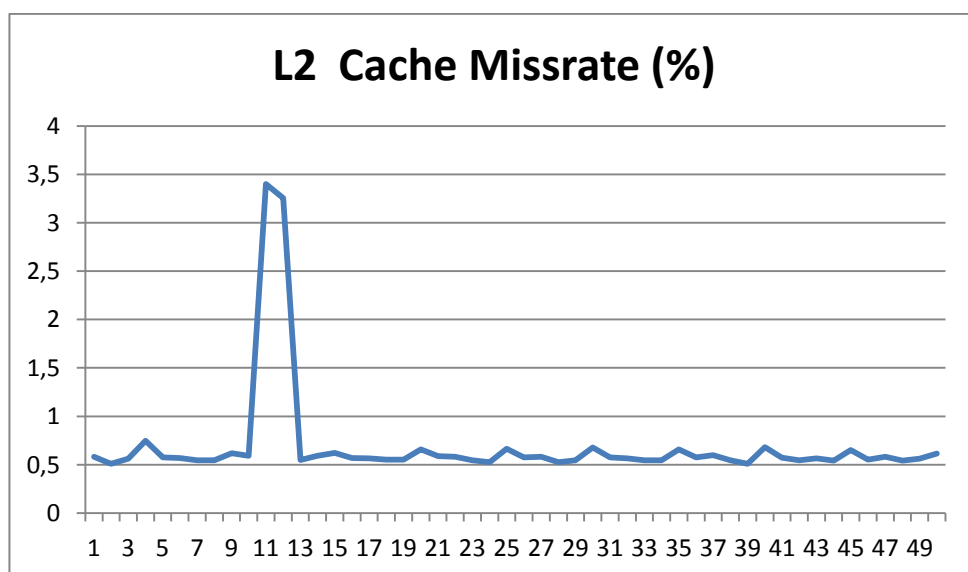
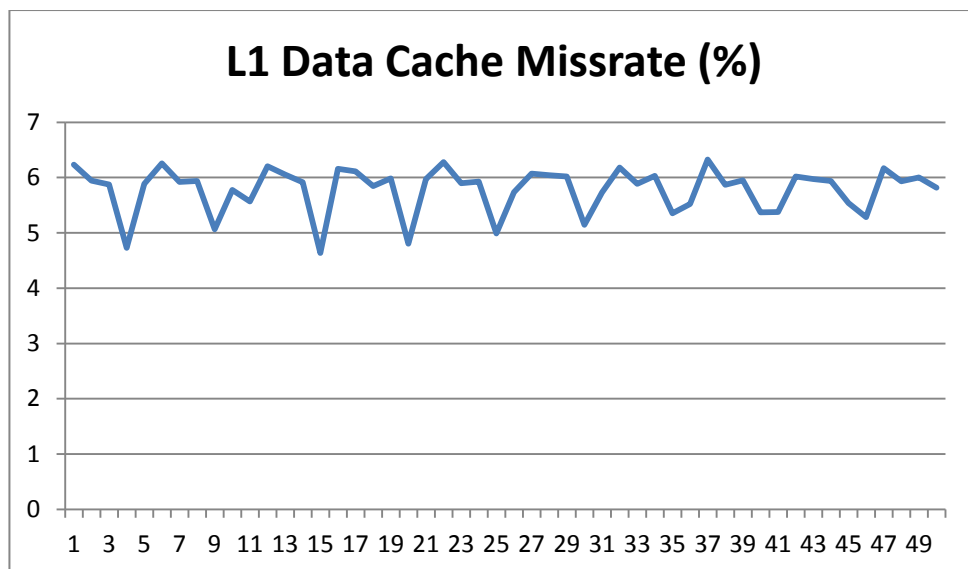
Στο τμήμα αυτό θα μελετήσουμε τη συμπεριφορά των benchmarks στον άξονα του χρόνου, παρατηρώντας τις διάφορες «φάσεις» εκτέλεσής τους. Με τον τρόπο αυτό θα έχουμε μια πιο ολοκληρωμένη και αξιόπιστη εικόνα των μεταβολών που παρατηρήσαμε προηγουμένως.

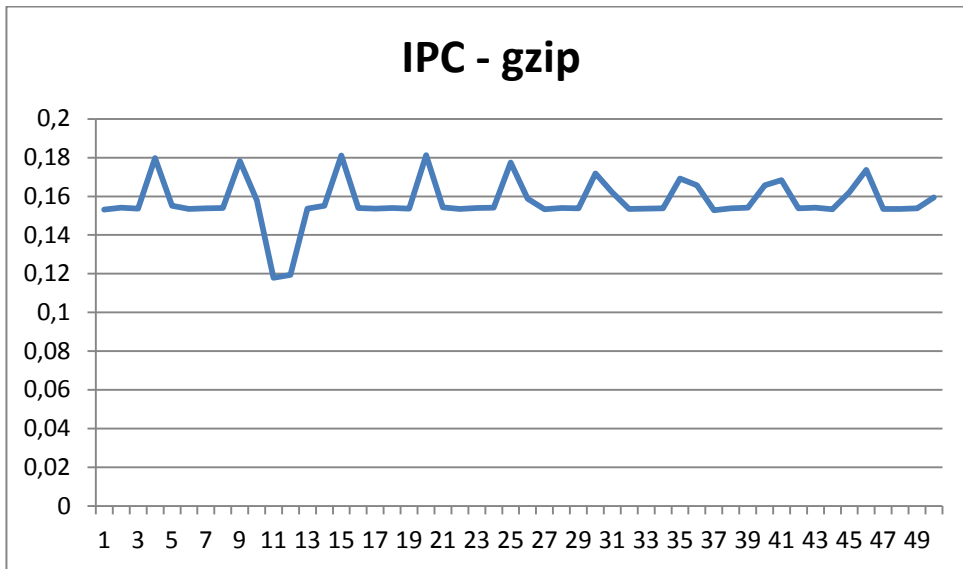
Art:



Το benchmark αυτό παρουσιάζει ομοιόμορφη συμπεριφορά στις διάφορες φάσεις του. Αυτό σημαίνει ότι τα αποτελέσματα που προέκυψαν για τις πρώτες 1 δις εντολές ήταν χαρακτηριστικά, και δίνουν μια αξιόπιστη εικόνα για την εκτέλεση του art.

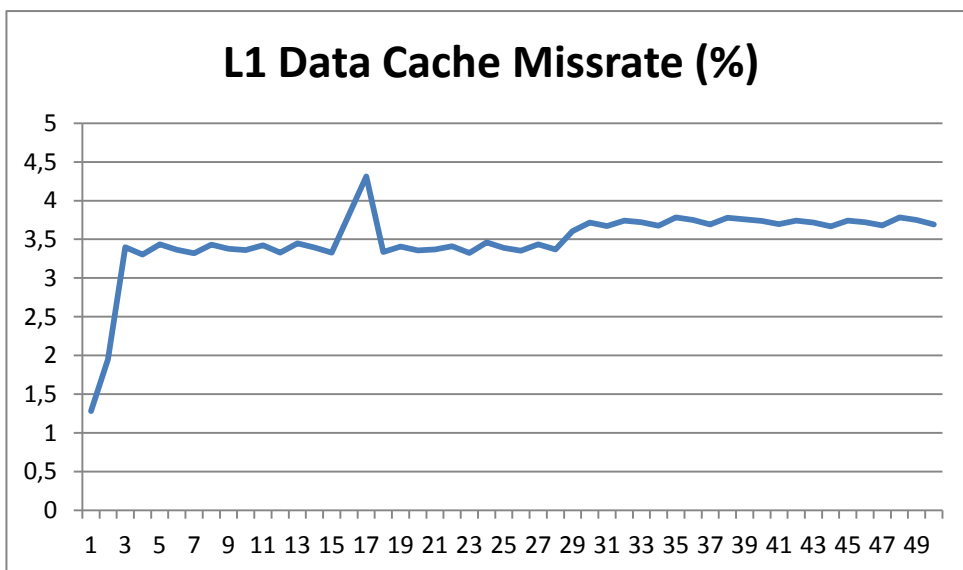
Gzip:

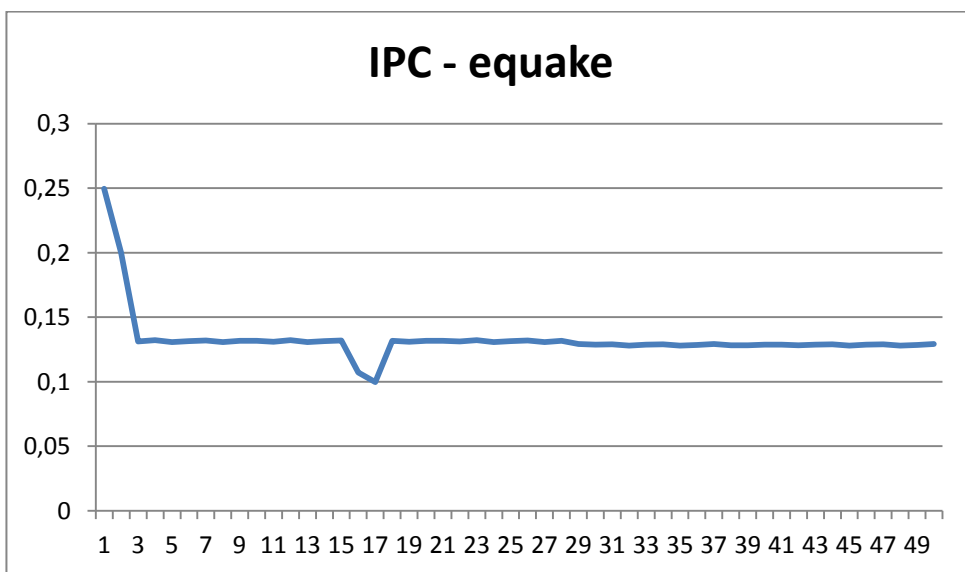
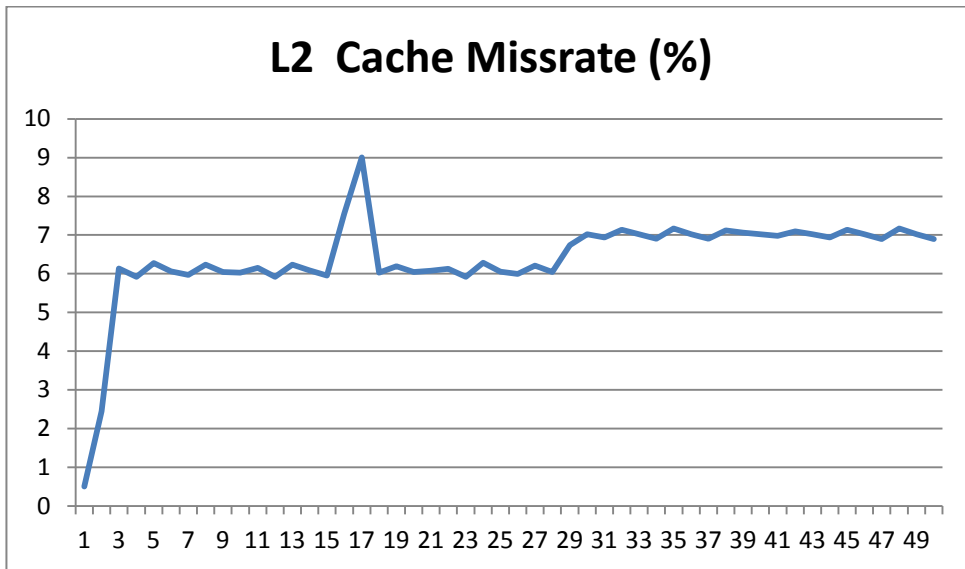




Παρατηρούμε ότι ενώ τα αποτελέσματα του gzip για την L1 cache είναι σχετικά ομοιόμορφα στον άξονα του χρόνου, δεν ισχύει το ίδιο για τα misses της L2 cache καθώς και για το IPC. Κατά συνέπεια, τα αποτελέσματα που πήραμε από τις πρώτες 1 δις εντολές δεν μπορούν να χαρακτηριστούν ως τελείως αντιπροσωπευτικά του προφίλ εκτέλεσης αυτού του benchmark.

Equake:

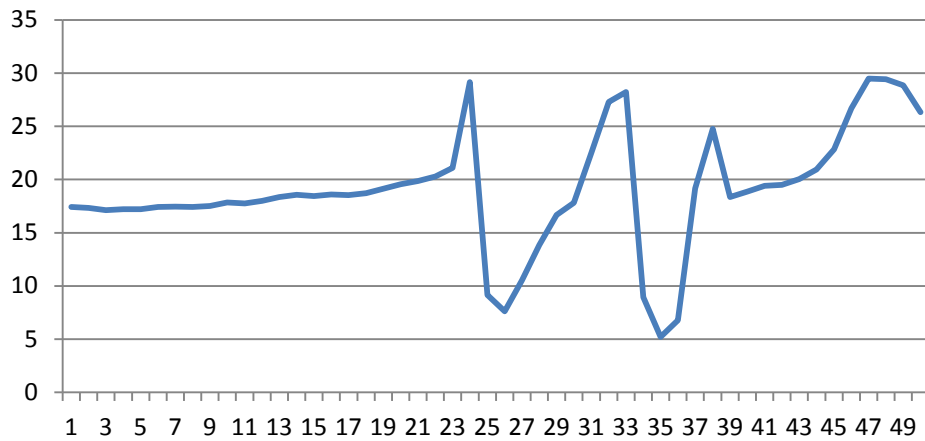




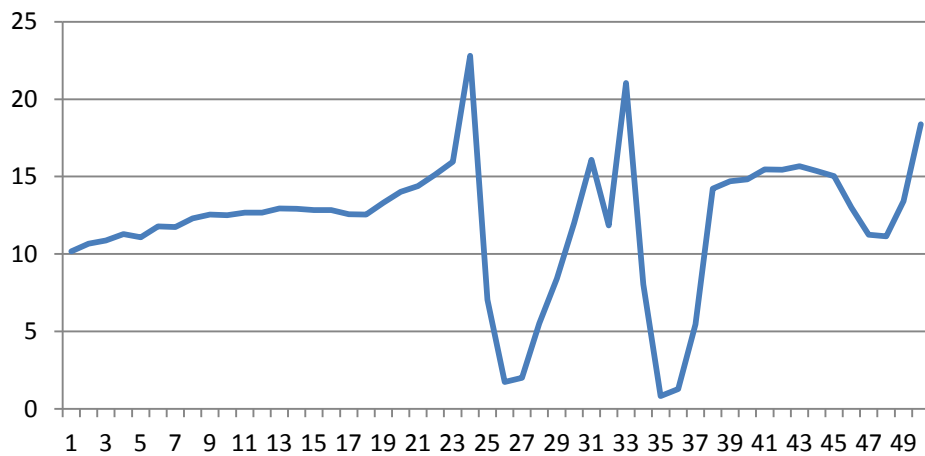
Και εδώ παρατηρείται εύκολα ότι οι μετρήσεις που έγιναν στο 1^ο μέρος της εργασίας καθίστανται αναξιόπιστες. Παρατηρείται μεγάλη απόκλιση της συμπεριφοράς του equake κατά τις πρώτες 300.000.000 εντολές με αυτή που διατηρείται για το μεγαλύτερο μέρος της εκτέλεσης.

Mcf:

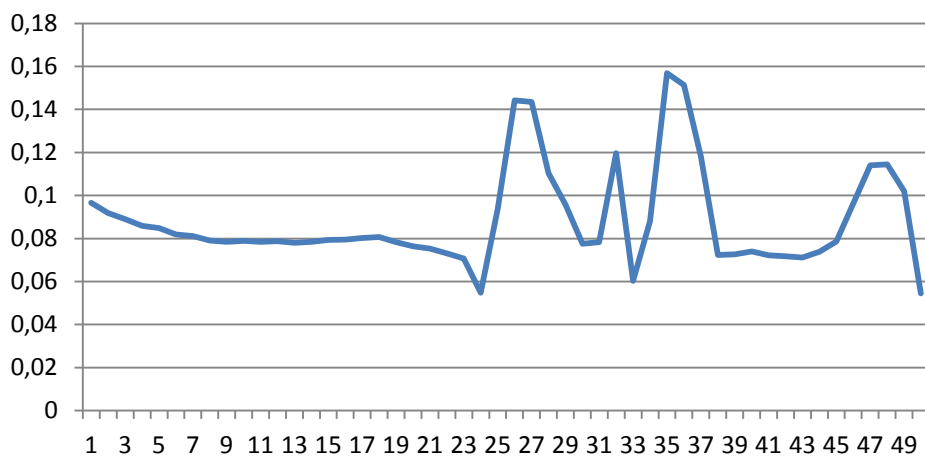
L1 Data Cache Missrate (%)



L2 Cache Missrate (%)



IPC - mcf



Τέλος, το συγκεκριμένο benchmark παρατηρούμε ότι εμφανίζει τελείως ακαθόριστη συμπεριφορά στον άξονα του χρόνου, επομένως δεν μπορούμε να χαρακτηρίσουμε τα αποτελέσματα που προέκυψαν ως αξιόπιστα.

Συνολικά συμπεράσματα:

- Στη γενική περίπτωση, ο παράγοντας που επηρεάζει περισσότερο το ποσοστό των misses της μνήμης είναι το μέγεθος της. Μεγαλύτερο line size επιφέρει επίσης βελτίωση της απόδοσης.
- Αλλαγές στο associativity φαίνεται να μην επηρεάζουν σημαντικά την απόδοση.
- Για ορισμένα benchmarks/προγράμματα (ειδικού σκοπού), οι παραπάνω γενικεύσεις μπορεί να μην ισχύουν, και για το σχεδιασμό ενός αποδοτικού μοντέλου μνήμης χρειαζόμαστε πρώτερη γνώση για την προβλεπόμενη χρήση της μνήμης από το σύστημα.
- Οι προσομοιώσεις που πραγματοποιήσαμε στο simics, ενώ είναι αρκετά λεπτομερείς και παραμετροποιήσιμες, μπορεί, αν εκτελεστούν για συγκεκριμένη φάση ενός benchmark, να μη δώσουν αντιπροσωπευτικά αποτελέσματα.