



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

---

# ΠΡΟΗΓΜΕΝΑ ΘΕΜΑΤΑ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

---

2η ΑΣΚΗΣΗ

*Γεωργίου Δημήτριος (03115106)*

*<el15106@central.ntua.gr>*

Απρίλιος 2019

# 1 Σκοπός της Άσκησης

- Η συγκεκριμένη εργαστηριακή άσκηση έχει ως σκοπό την μελέτη της επίδρασης διαφορετικών συστημάτων πρόβλεψης εντολών άλματος και την αξιολόγηση τους με αντικειμενικό κριτήριο τον διαθέσιμο χώρο του τσιπ. **Πιο συγκεκριμένα** γίνεται εξέταση της επίδρασης των βασικών predictors στην απόδοση 12 μετροπρογραμμάτων (**benchmarks**).
- Μετά το πέρας της εξέτασης, θα γίνει επιλογή του καταλληλότερου predictor για παραπάνω 12 μετροπρογράμματα.

# 2 Εργαλεία

## PIN TOOL

- Έγινε χρήση του εργαλείου Pin, ένα εργαλείο ανάλυσης εφαρμογών που αναπτύσσεται και συντηρείται από την εταιρεία Intel
- Προσφέρει δυνατότητες για dynamic binary instrumentation, ή εναλλακτικά επιτρέπει την δυναμική εισαγωγή και τροποποίηση κώδικα της εφαρμογής κατά την διάρκεια εκτέλεσης των αυτών των μετροπρογραμμάτων. Απώτερος σκοπός είναι η συλλογή πληροφοριών που αφορούν την εκτέλεση, συγκεκριμένα **αρχιθμός εντολών, αριθμός misses και hits στην cache**
- Σημειώνεται ότι στα πλαίσια της παρούσας εργασίας έγινε χρήση της **PIN 97554 Ubuntu Linux 16.04** έκδοσης του PIN με **πυρήνα 4.4**

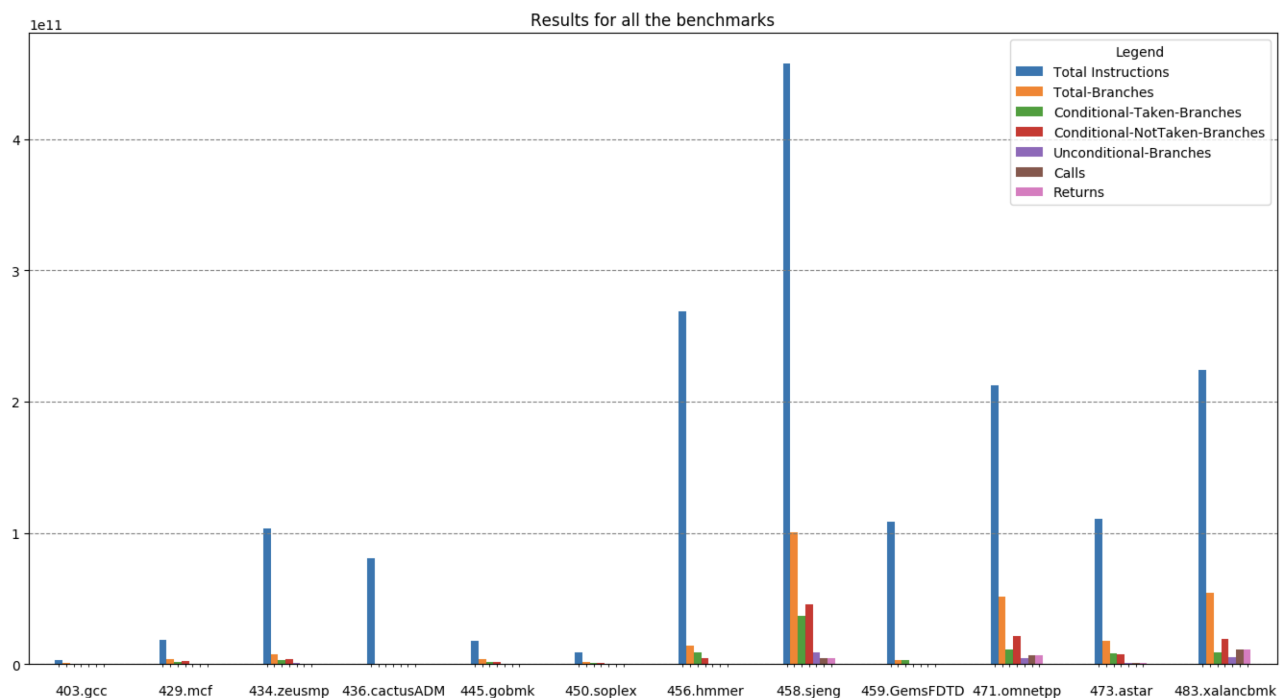
## Benchmarks

- Τα 12 μετροπρογράμματα (**benchmarks**) που χρησιμοποιήθηκαν ήταν **SPEC\_CPU2006**
- Στα πλαίσια των προσομοιώσεων μας χρησιμοποιήθηκαν 12 από αυτά τα benchmarks. Συγκεκριμένα:
  1. 403.gcc
  2. 429.mcf
  3. 434.zeusmp
  4. 436.cactusADM
  5. 445.gobmk
  6. 450.soplex
  7. 456.hmmmer
  8. 458.sjeng
  9. 459.GemsFDTD
  10. 471.omnetpp
  11. 473.astar
  12. 483.xalancbmk

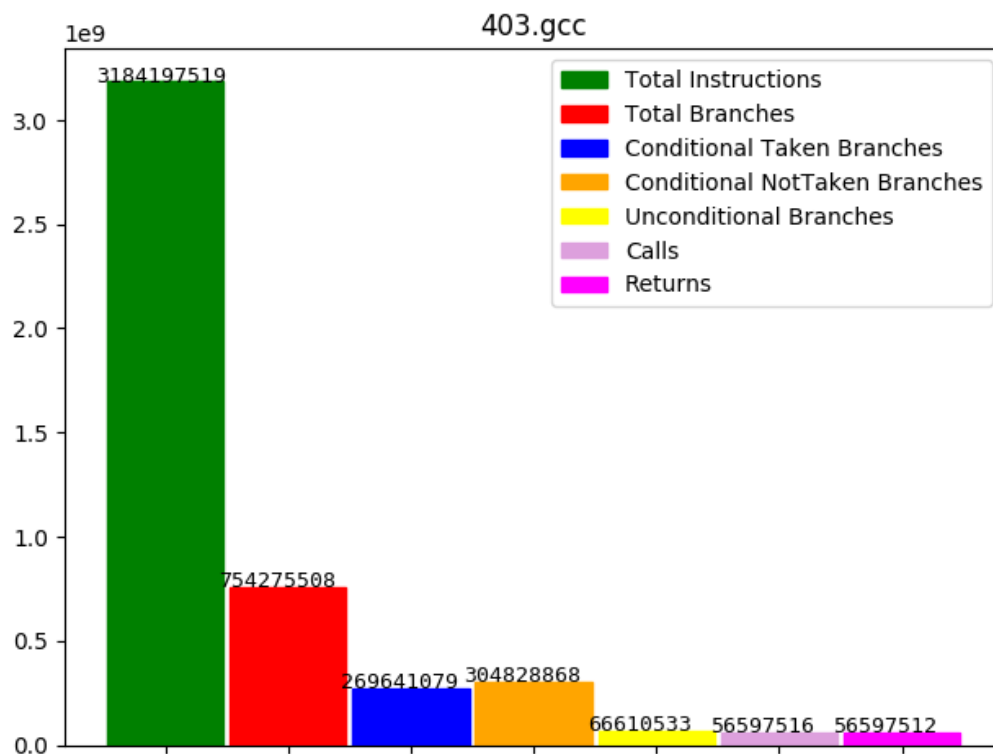
# 3 Πειραματική Αξιολόγηση

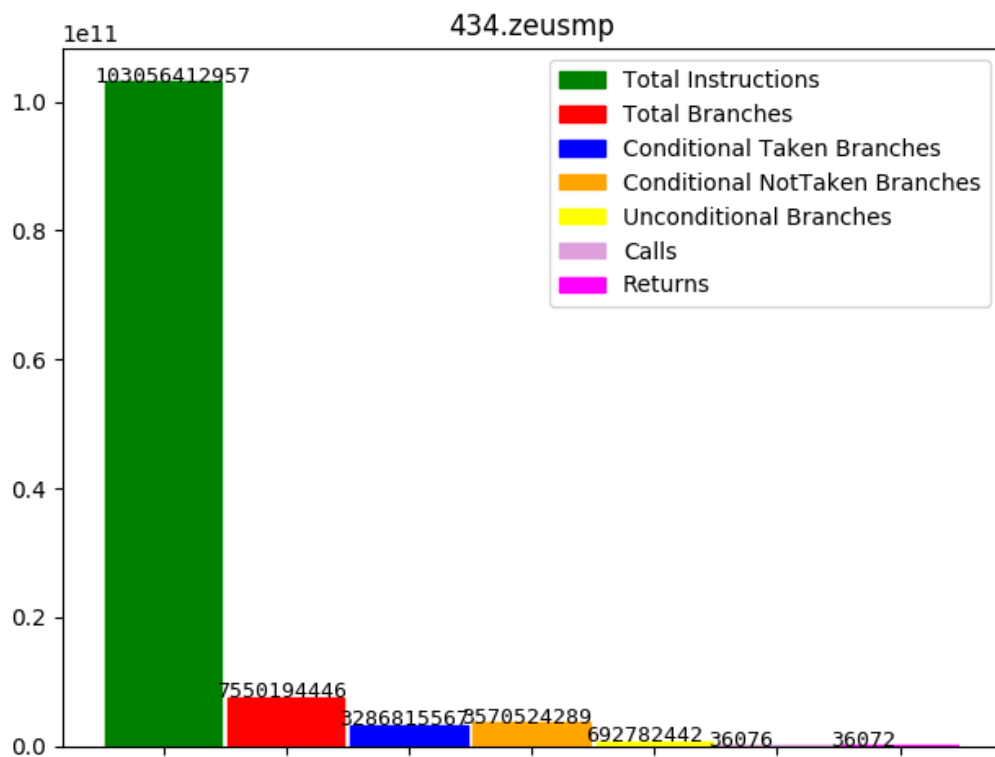
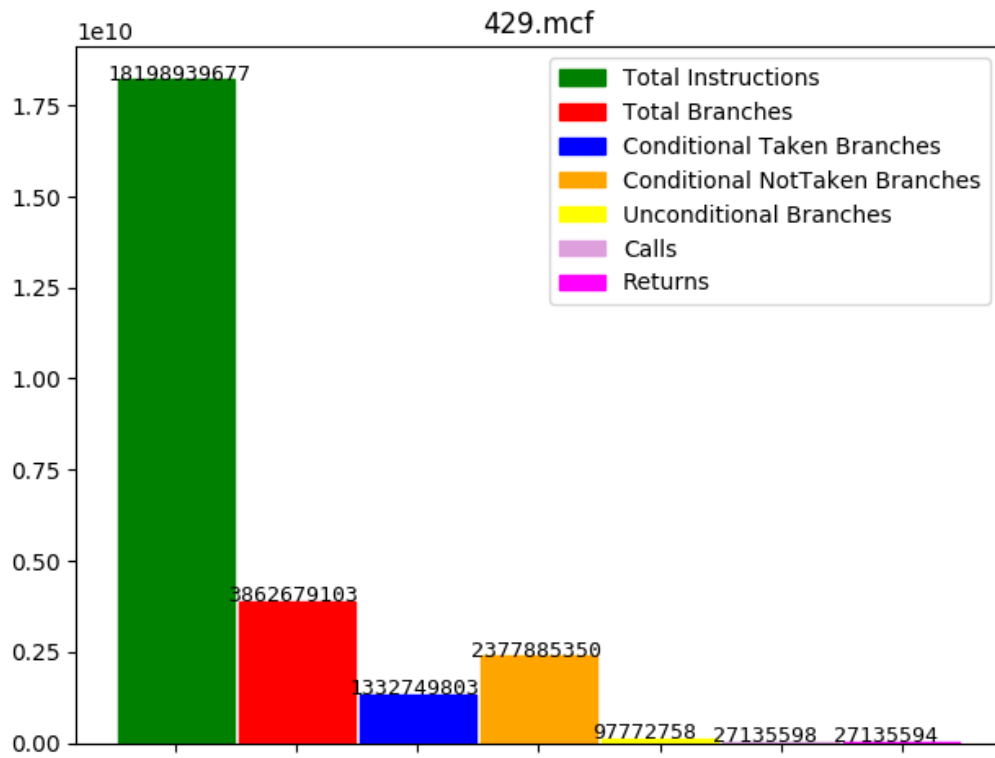
## 3.1 Μελέτη εντολών άλματος

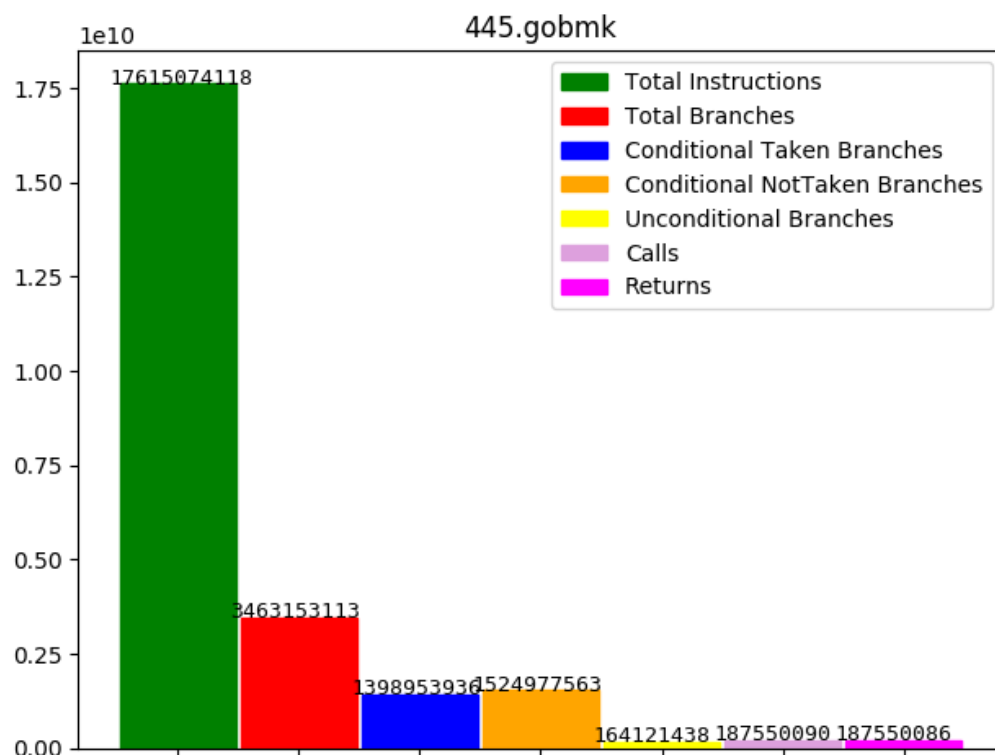
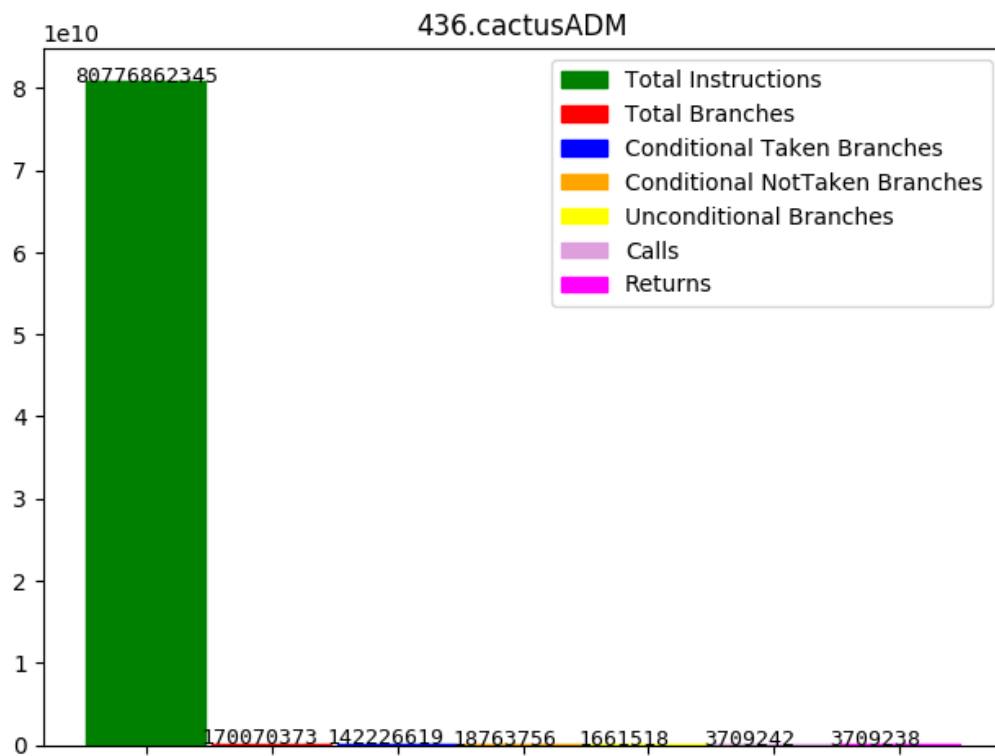
- Στο **Πρώτο Μέρος** της πειραματικής αξιολόγησης για την παρούσα άσκηση, εξετάζουμε τις εντολές άλματος που εκτελούνται για κάθε ένα από τα 12 μετροπρογράμματα. Εν συνεχεία, παρουσιάζουμε ένα διάγραμμα με το πλήθος και τα είδη των εντολών άλματος.
- **Παρατηρούμε:**
  1. Η πλειοψηφία των εντολών είναι conditional-taken και conditional-notTaken
  2. Μια γενικότερη εκτίμηση που προκύπτει από το διάγραμμα είναι ότι οι εντολές άλματος αποτελούν κατά μέσο όρο το **15%** των συνολικών εντολών που αποτελούν το μετροπρόγραμμα
  3. Έχουμε να κάνουμε με μετροπρογράμματα που αποτελούνται γενικότερα από **δισεκατομμύρια** εντολές και μάλιστα, παρατηρούμε ένα range της τάξης από **3 δισ. έως 400 δισ. εντολές**

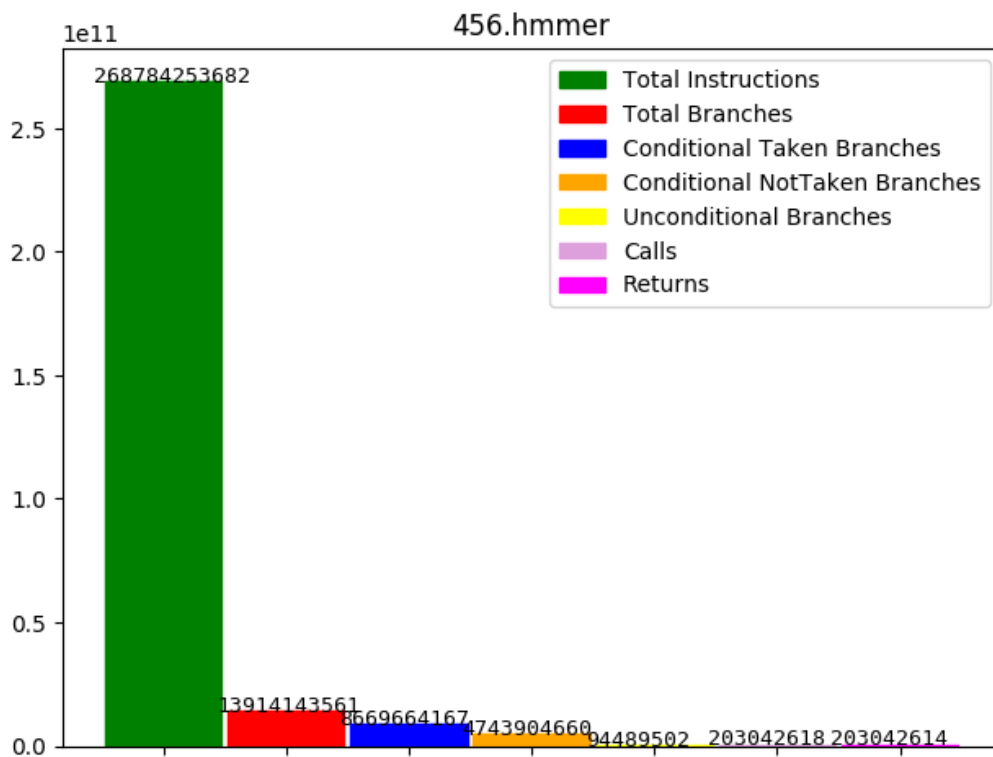
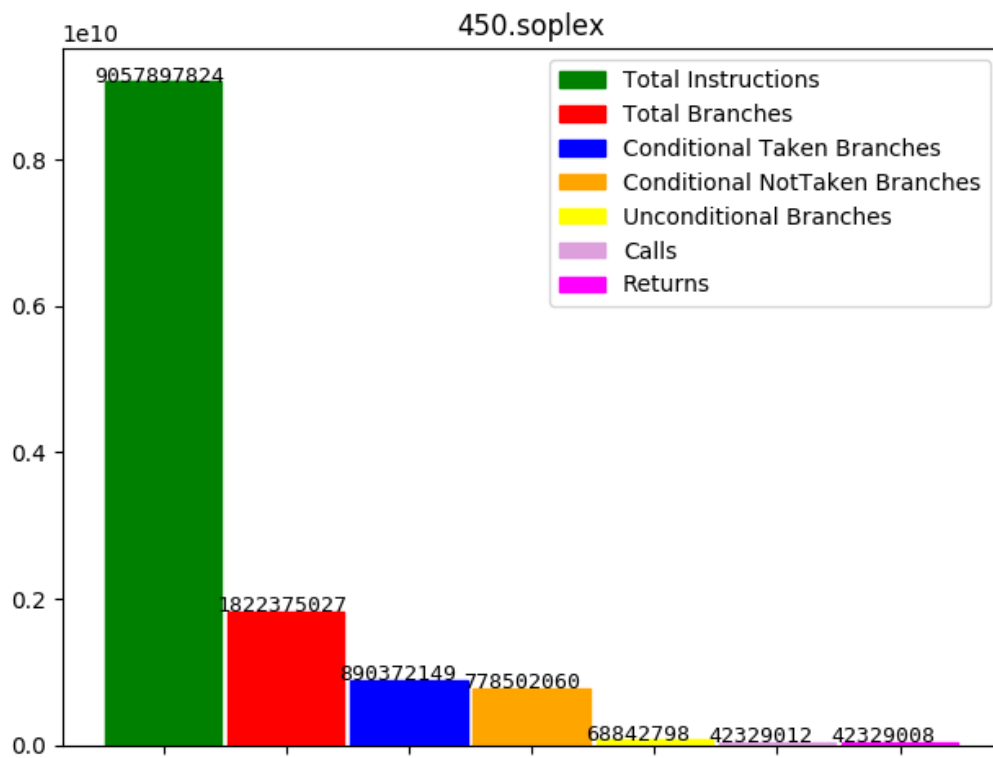


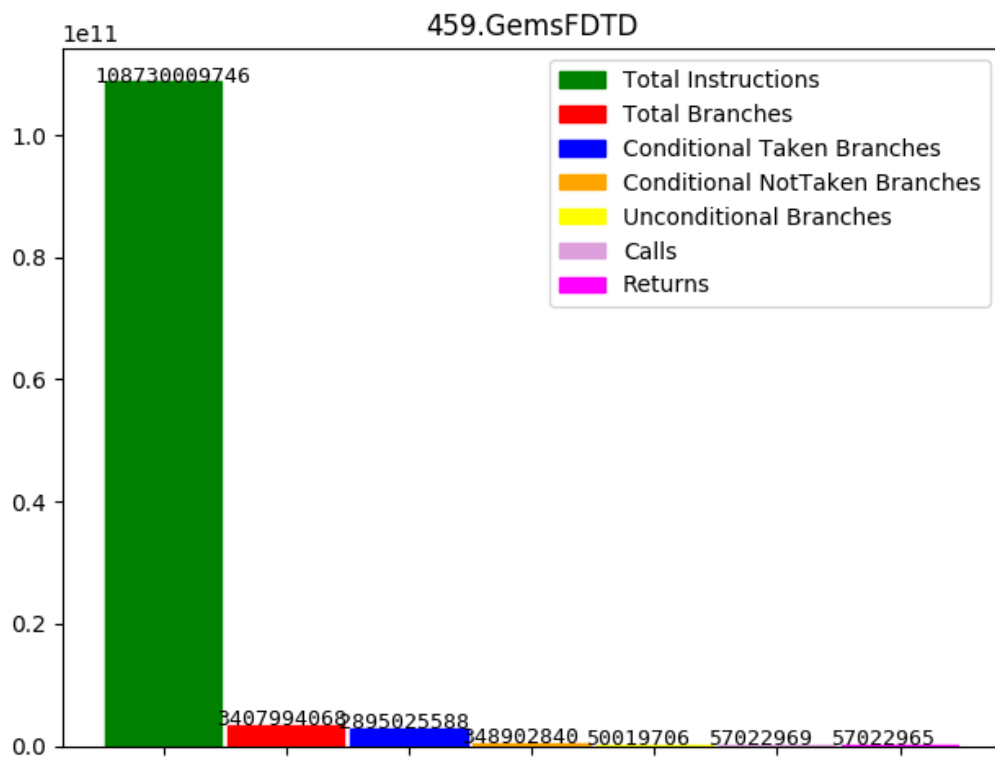
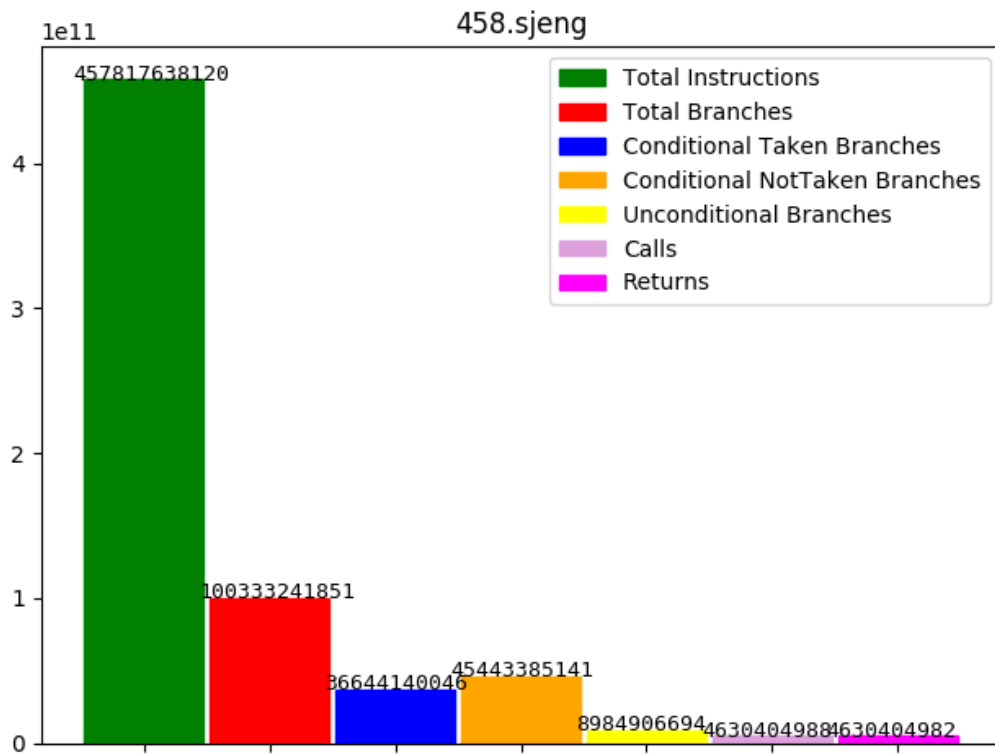
Παρακάτω παραθέτουμε ξεχωριστά για κάθε benchmark το αντίστοιχο διάγραμμα για τις εντολές άλματος, με μεγαλύτερη λεπτομέρεια, για καλύτερη επισκόπηση και εξαγωγή συμπεσμάτων

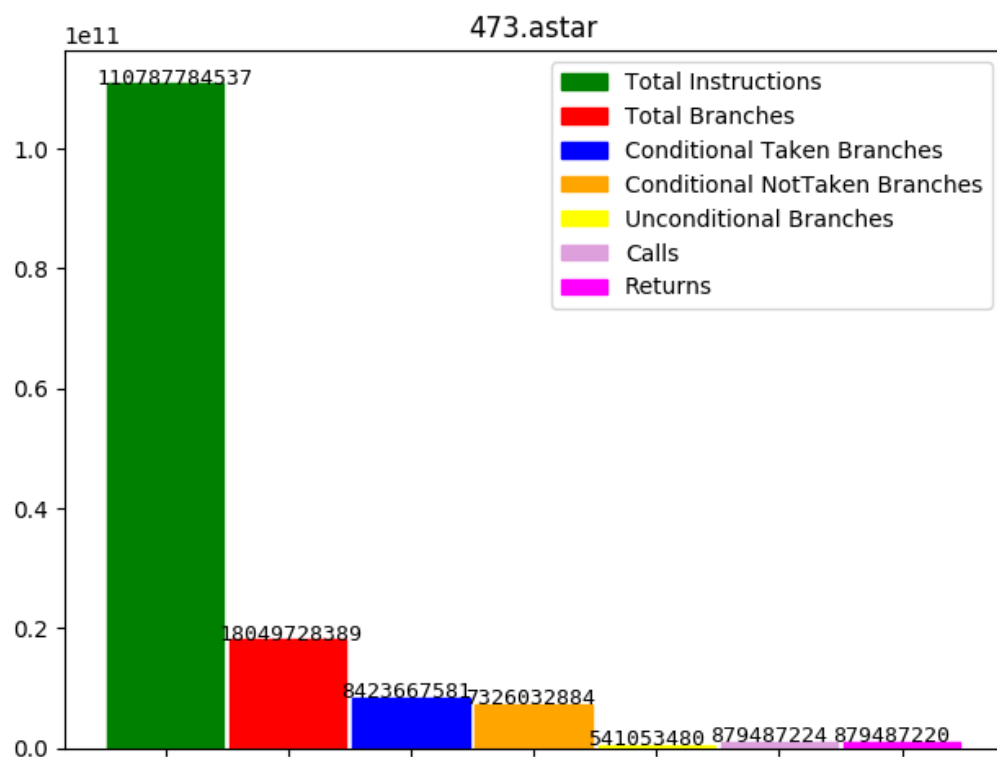
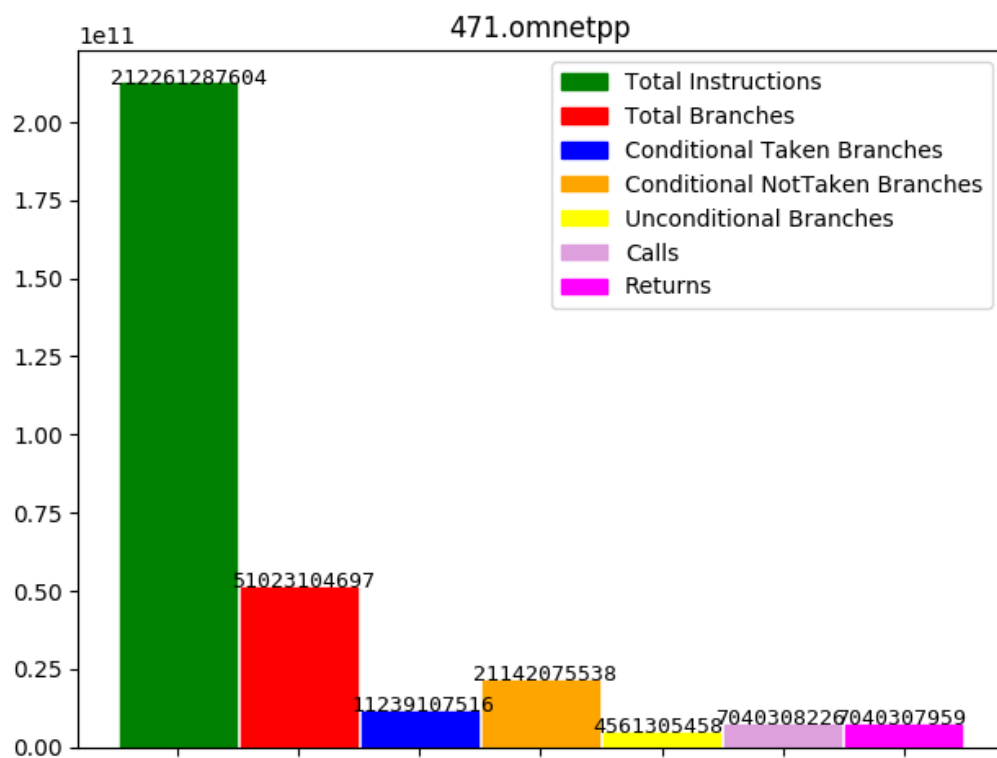




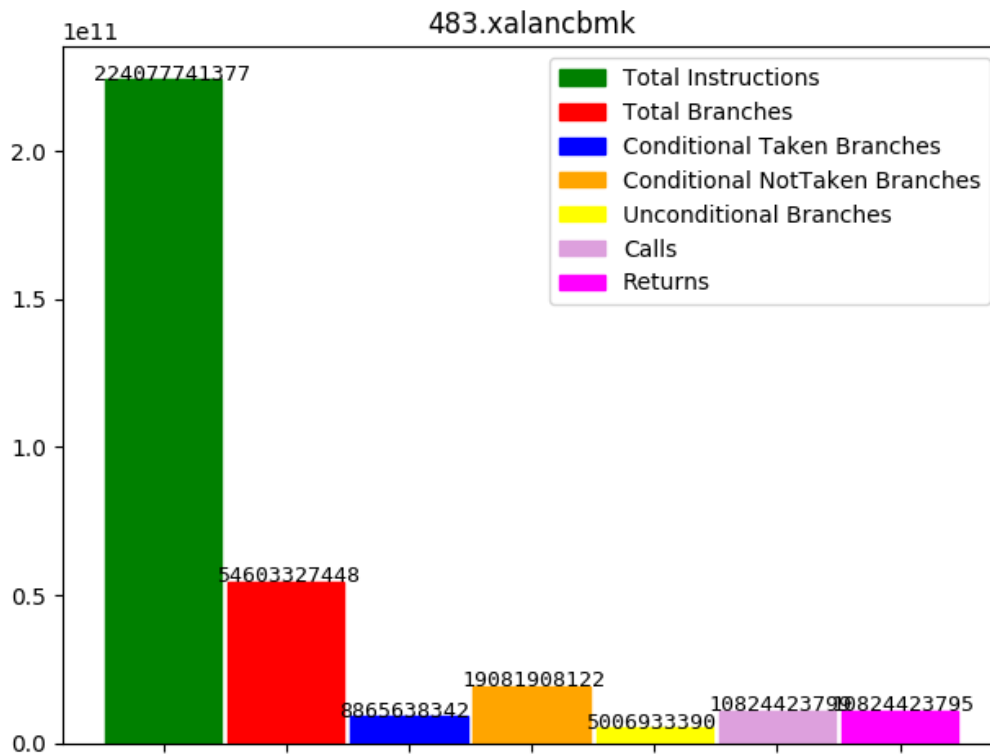












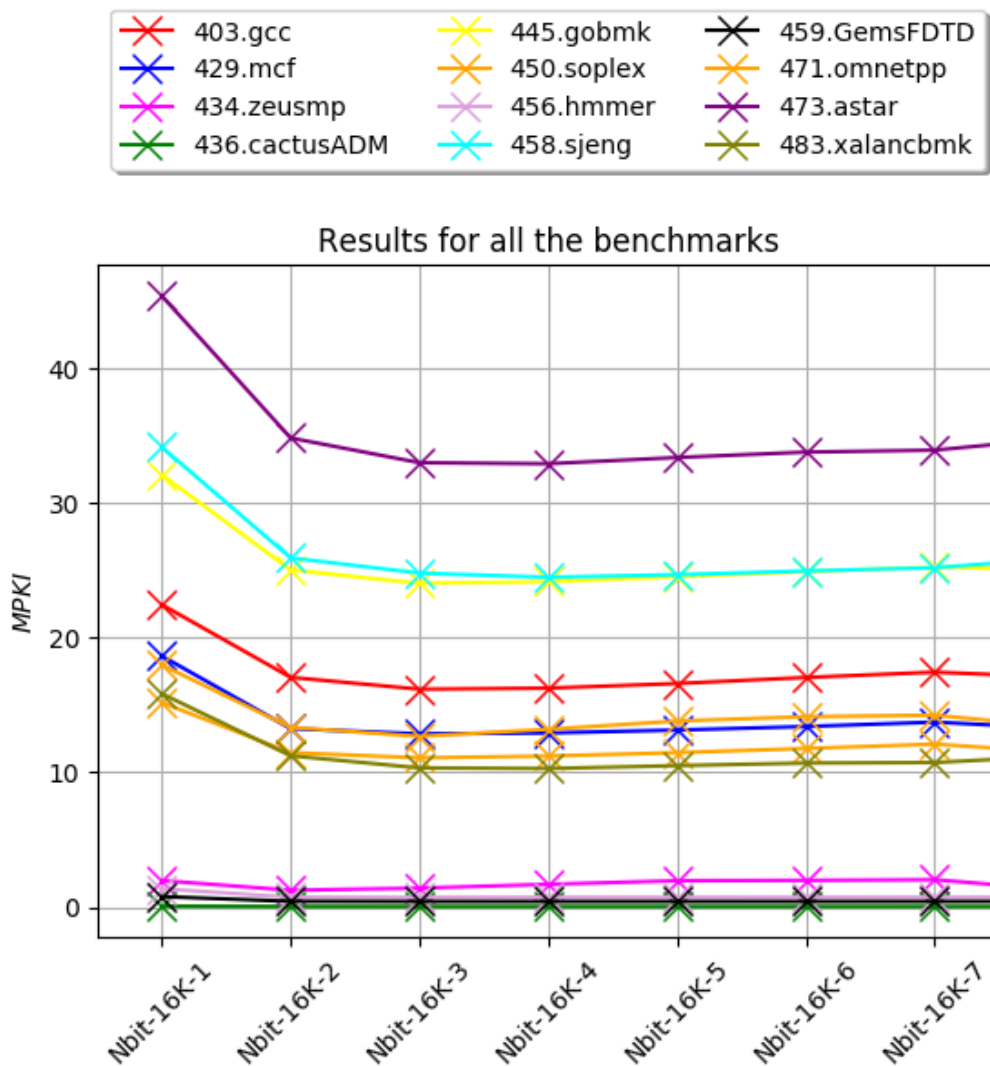
- **Συμπεράσματα:**

1. Πολύ λίγα είναι τα unconditional branches, τα calls και returns όπως αναμέναμε.
2. Γενικότερα το μεγαλύτερο ποσοστό εντολών δεν είναι τα branches, κατά μέσο όρο αποτελούν το 15% των συνολικών εντολών του μετροπρογράμματος. Εντολές **load** και **store** σημειώνουν υψηλά ποσοστά! Πάντως συγκεκριμένα για τα benchmarks μας, το *hammer* έχει υψηλά ποσοστά εντολών μνήμης, ενώ τα *gcc*, *mcfm xalancbmk*, *sjeng* έχουν υψηλότερα ποσοστά εντολών άλματος.
3. Παρ' όλα αυτά, τα υψηλά αυτά ποσοστά **δεν αποτελούν** ένδειξη για παρουσία περισσότερων **bottlenecks**. Υπάρχουν benchmarks που σημειώνουν υψηλά ποσοστά και στις δύο κατηγορίες εντολών και εντούτοις διατηρούν υψηλό IPC

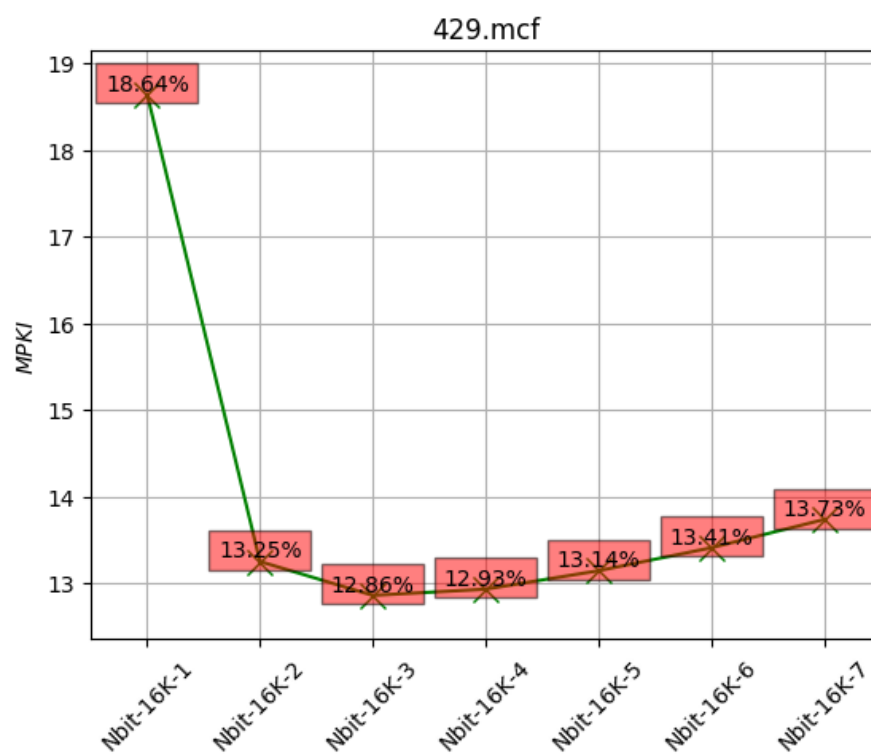
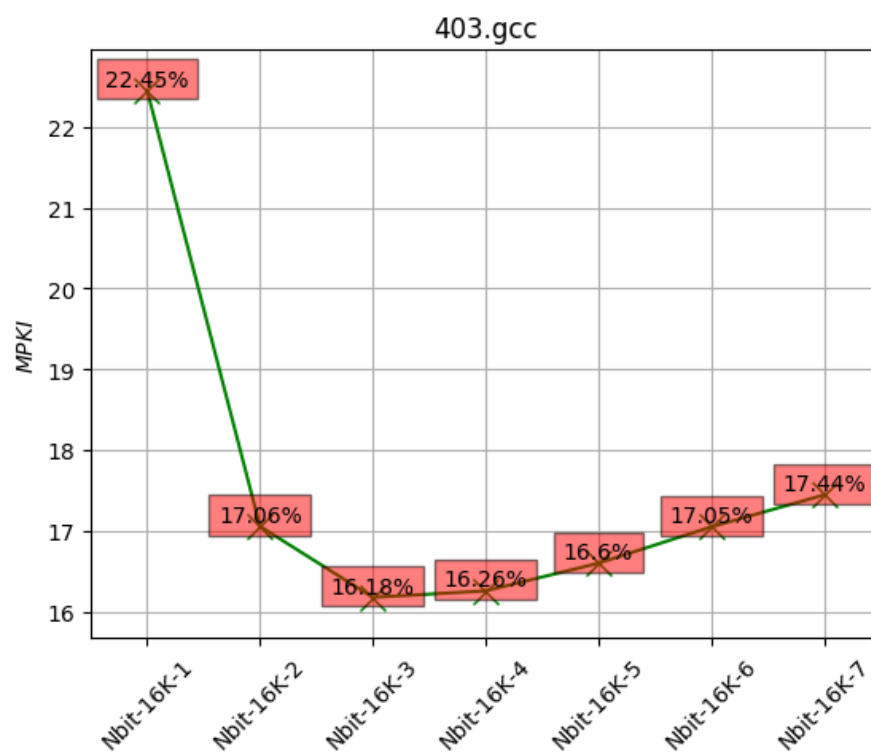
### 3.2 Μελέτη N-bit predictors

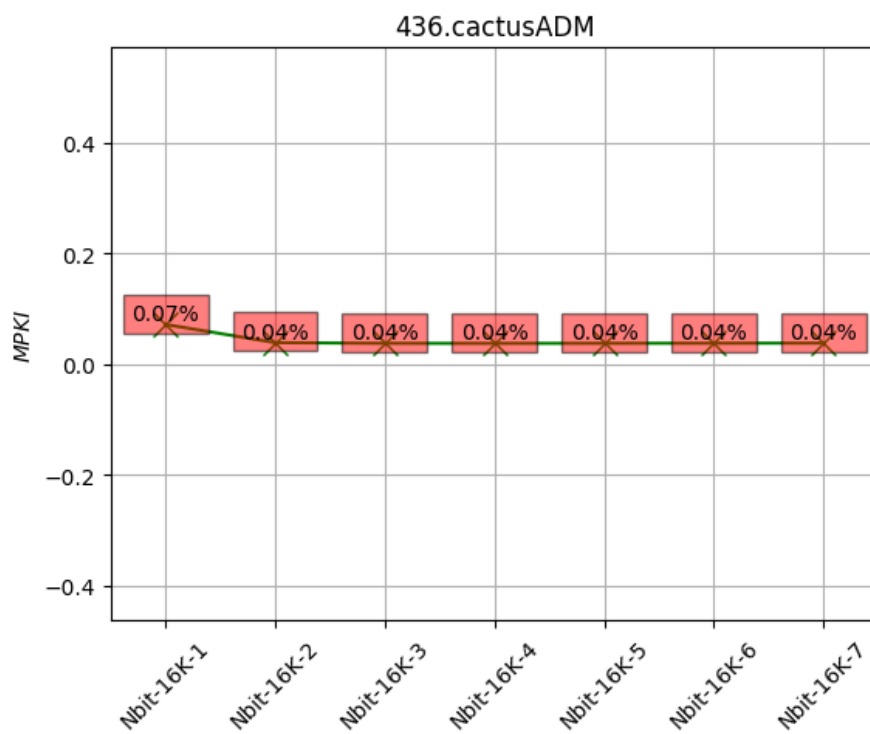
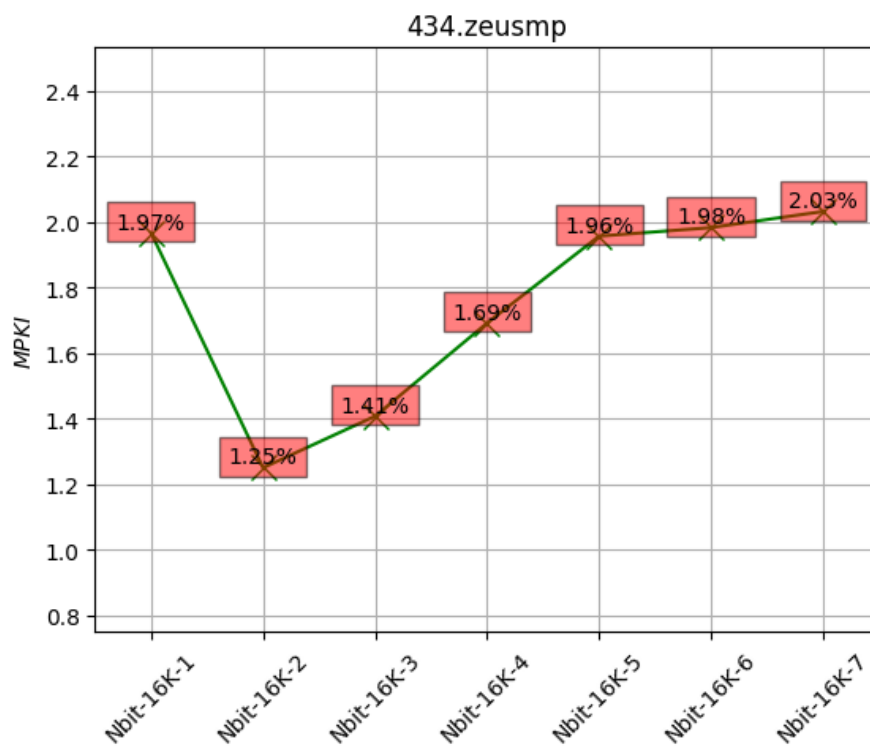
- Στο **Δεύτερο Μέρος** της πειραματικής αξιολόγησης για την παρούσα άσκηση, εξετάζουμε την απόδοση των **n-bits predictors** κάνοντας χρήση της υλοποίησης που μας δόθηκε *cslab\_branch.cpp*.
- **Συγκεκριμένα** το Δεύτερο Μέρος αποτελείται από δύο επιμέρους στάδια.
  - Στο **πρώτο στάδιο** διατηρούμε σταθερό τον αριθμό των BHT entries και ίσο με 16K, και γίνεται προσομοίωση των n-bits predictors για  $n = 1, \dots, 7$  κάνοντας χρήση των direction MPKI (Mispredictions Per Thousands Instructions)
  - Στο **δεύτερο στάδιο** διατηρούμε σταθερό το hardware και ίσο με 32K bits, και γίνεται εκτέλεση των προσομοιώσεων των n-bits predictors για τα 12 benchmarks, για για  $n = 1, 2, 4$

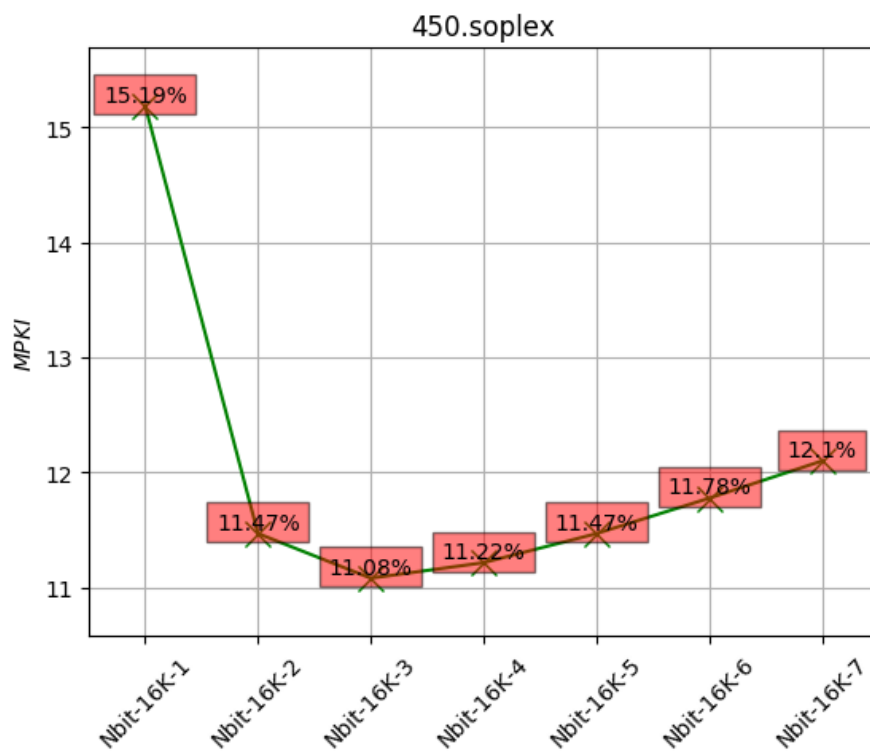
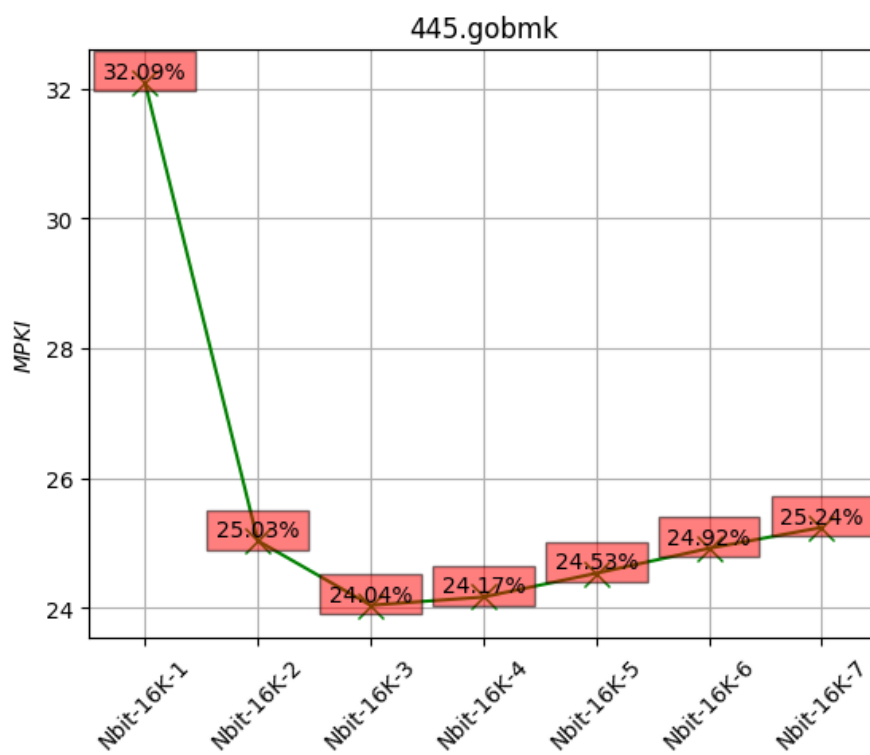
Αποτελέσματα των n-bits predictors για **BHT entries = 16K**,  $n = \{1, \dots, 7\}$  για 12 benchmarks  
ΣΤΥΝΟΛΙΚΑ

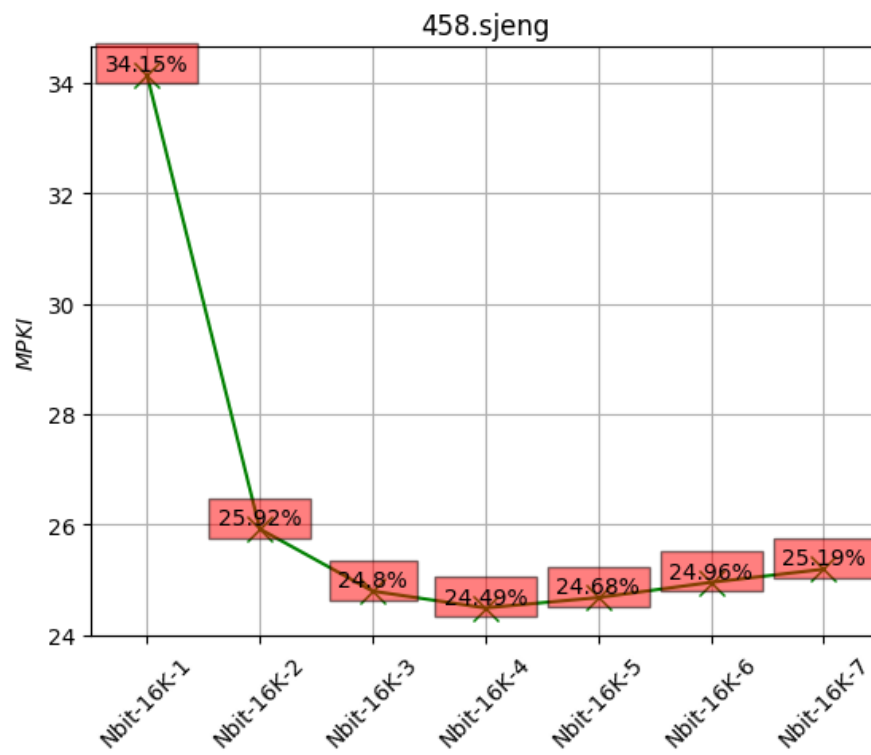
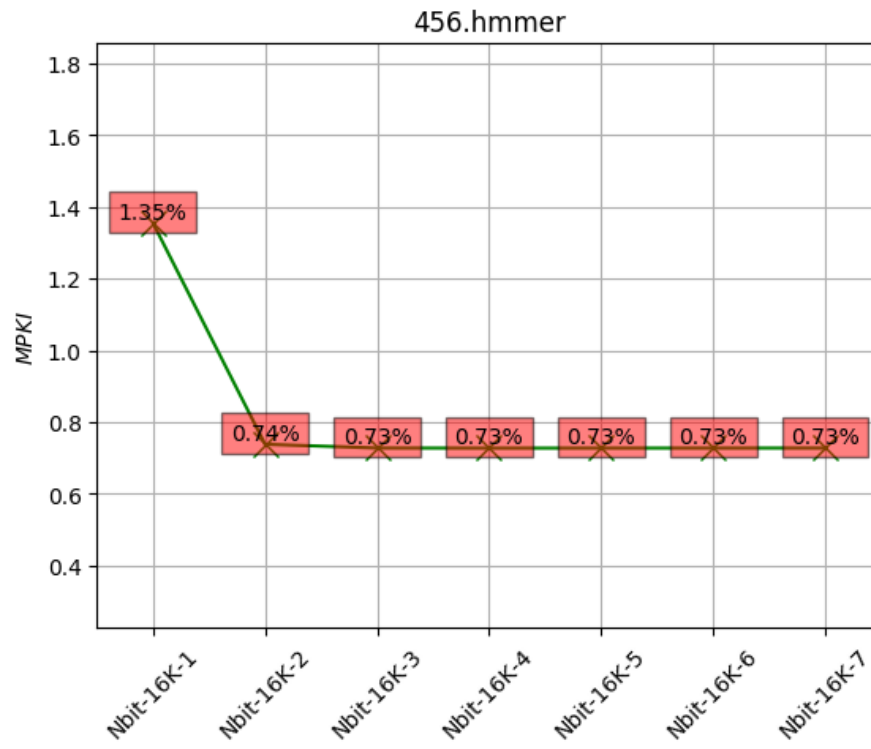


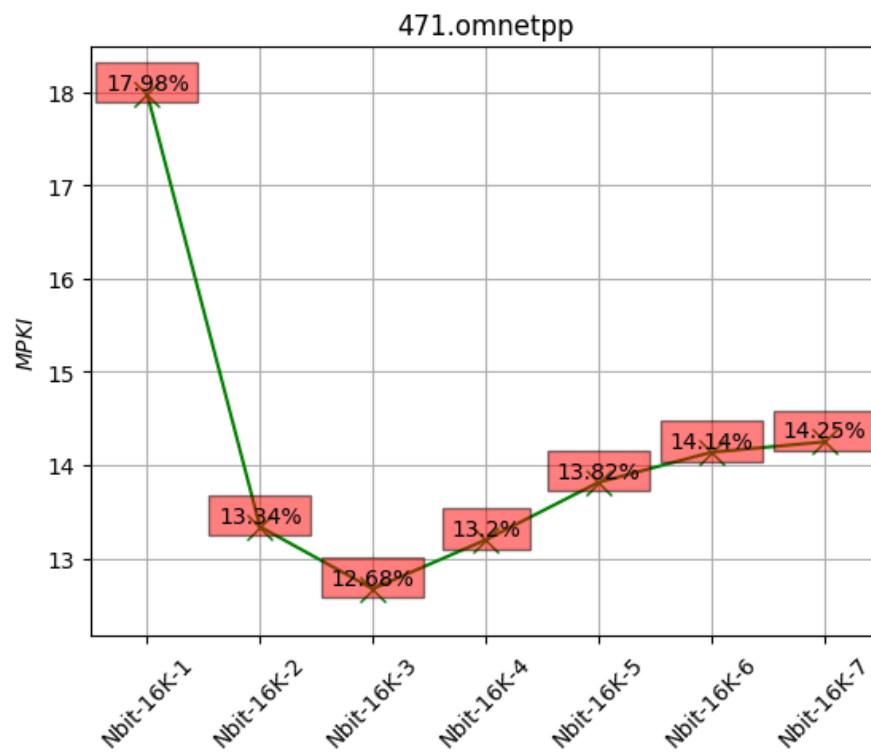
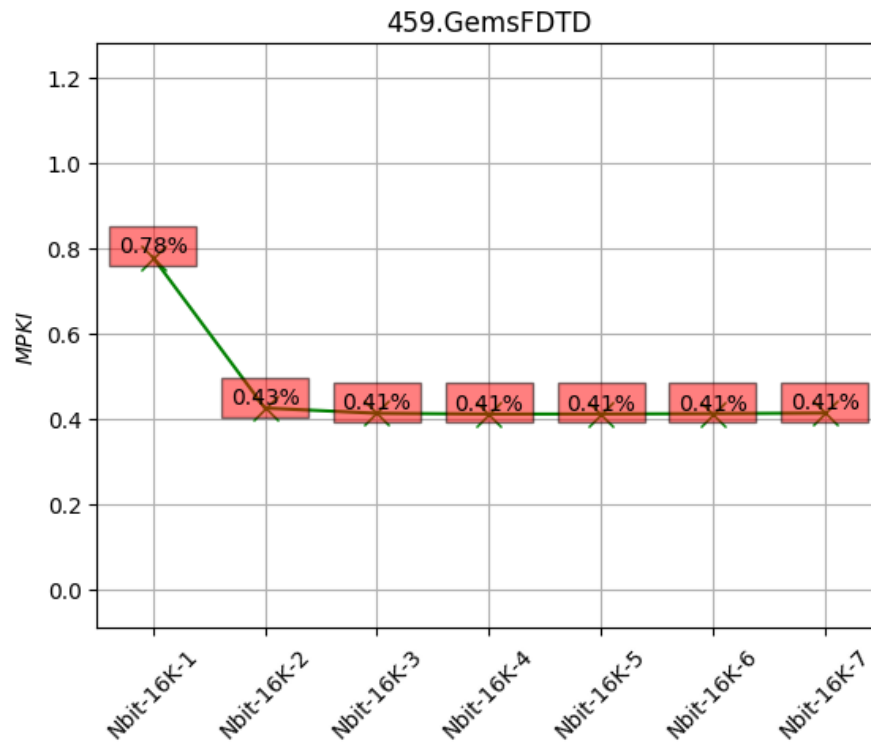
Παρακάτω παραθέτουμε ξεχωριστά για κάθε benchmark το αντίστοιχο διάγραμμα για τους  $n$ -bit predictors για  $BHT\ entries = 16K$ ,  $n = \{1, \dots, 7\}$  με μεγαλύτερη λεπτομέρεια, για καλύτερη επισκόπηση και εξαγωγή συμπεσμάτων

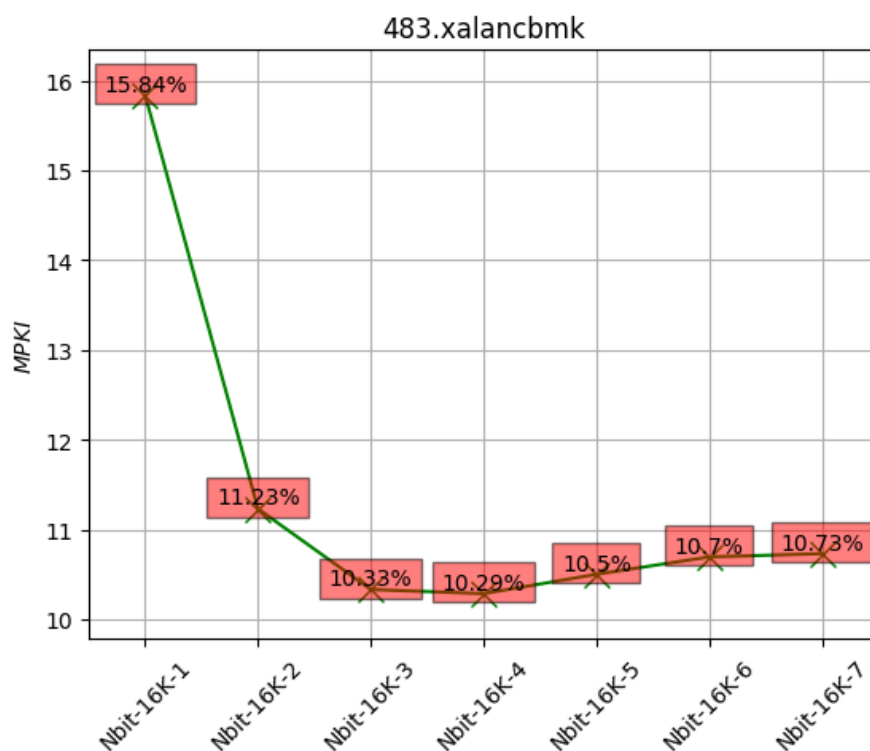
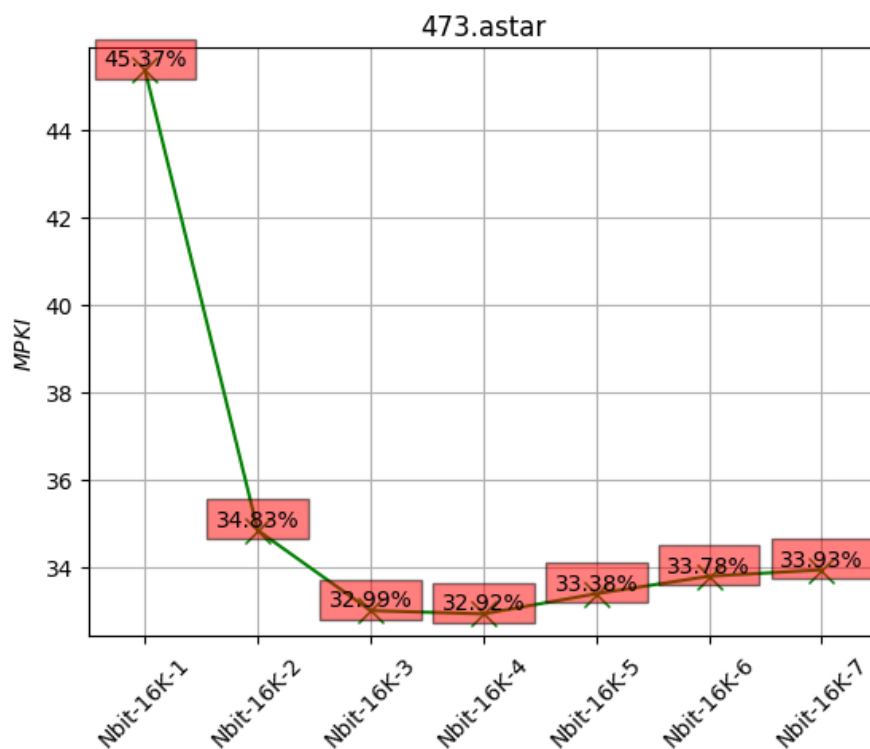












• Συμπεράσματα:

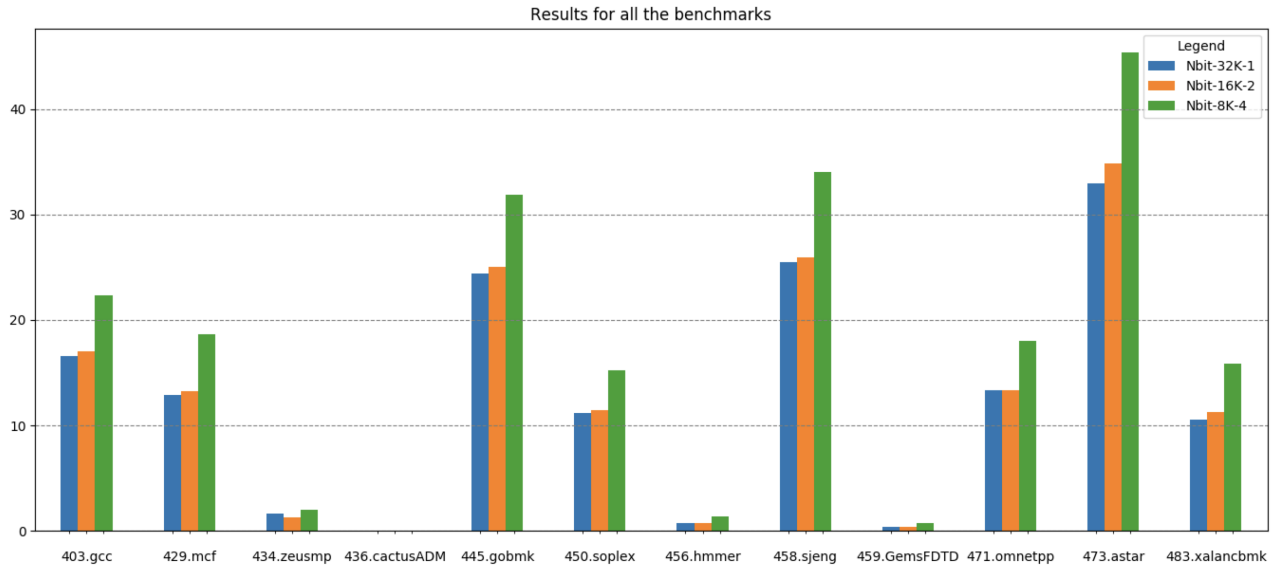
1. Για bit  $n = 1$  bit, παρατηρούμε ότι σχεδόν όλα τα benchmarks έχουν την χειρότερη δυνατή απόδοση, καθώς το MPKI αγγίζει μεγαλύτερες τιμές για  $n = 1$  συγκριτικά με οποιοδήποτε άλλο αριθμό bits. Αναλυτικά, η μετάβαση σε  $n = 2$  bit οδηγεί σε **μεγάλη βελτίωση της απόδοσης**, ενώ σε  $n = 3, \dots, 7$  bits δεν παρατηρούμε σημαντική βελτίωση αντιθέτως η **απόδοση διατηρείται σταθερή** (δηλαδή και ο δείκτης MPKI) **ή και χειροτερεύει** για τα περισσότερα benchmarks (αρκετά μικρή χειροτέρευση)
2. Τα χαμηλότερα ποσοστά mispredictions σημειώνονται για τα benchmarks (**Κάτω από 2%**):
  - 434.zeusmp
  - 436.cactusADM



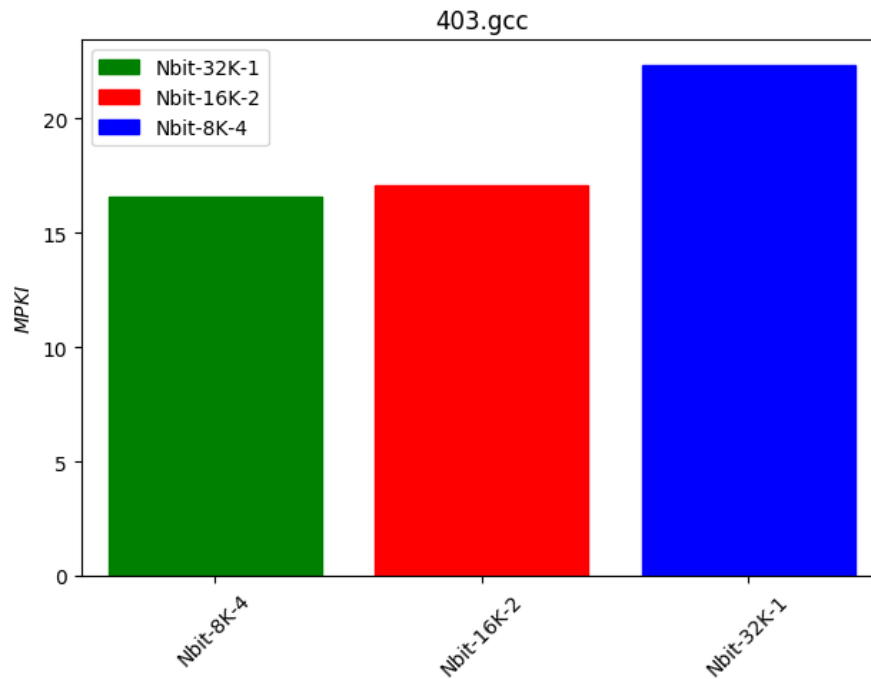
- 456.hmmmer
- 459.GemsFDTD

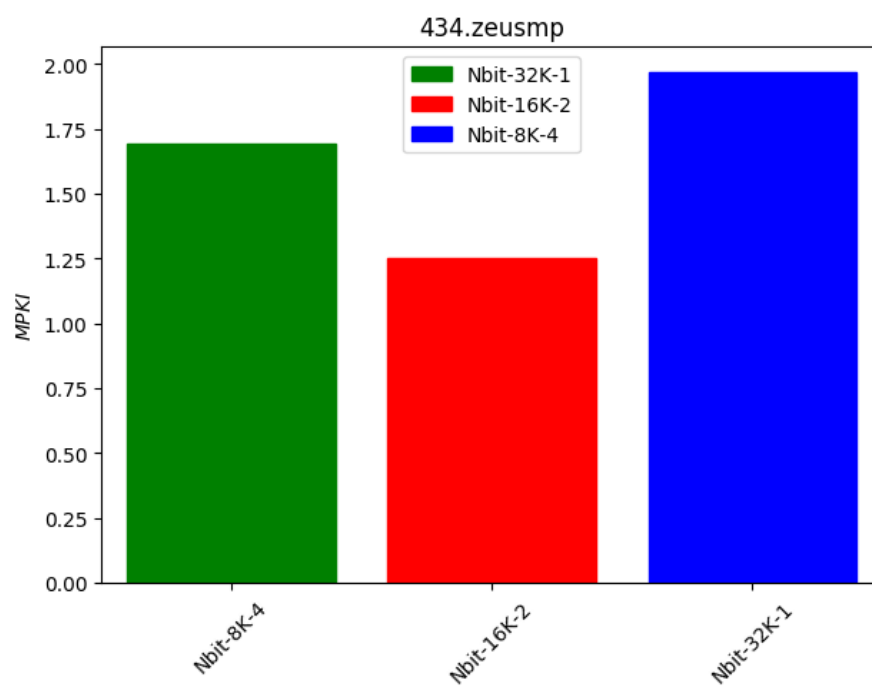
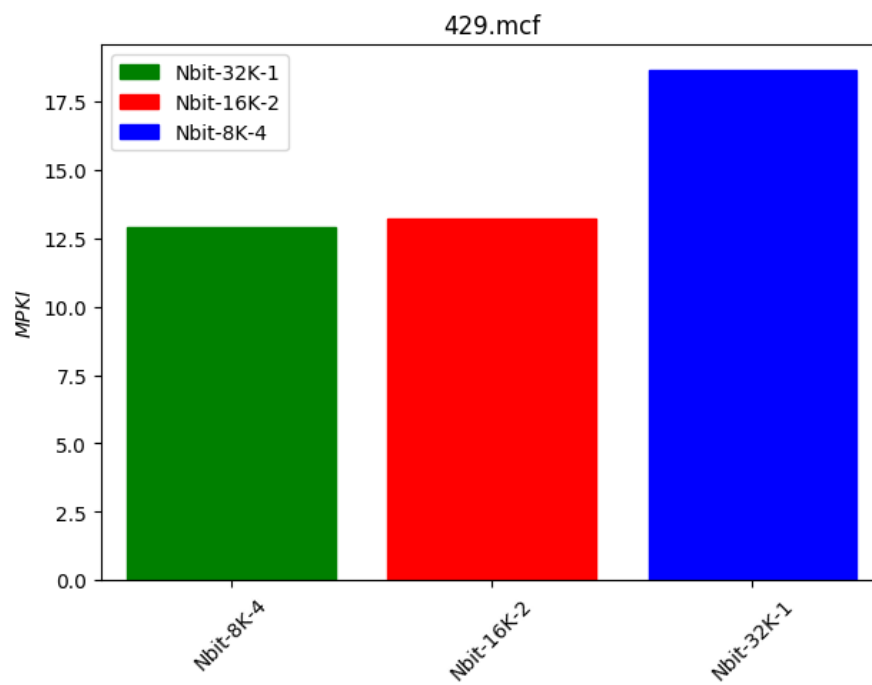
Μάλιστα, η τάξη MPKI που σημειώνουν αυτά τα 4 benchmarks είναι αρκετά μικρότερη συγκριτικά με τα υπόλοιπα 6 benchmarks.

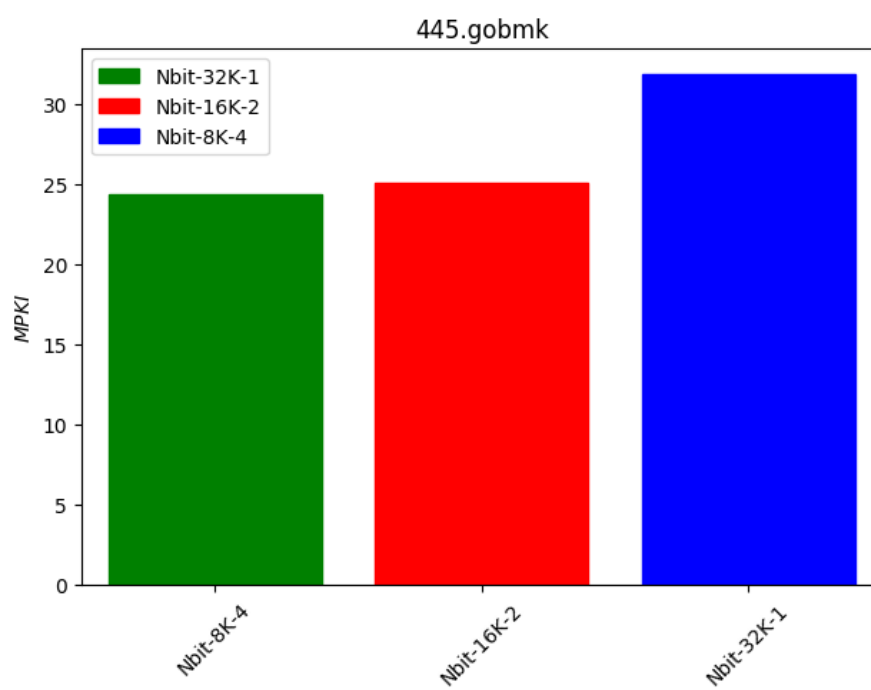
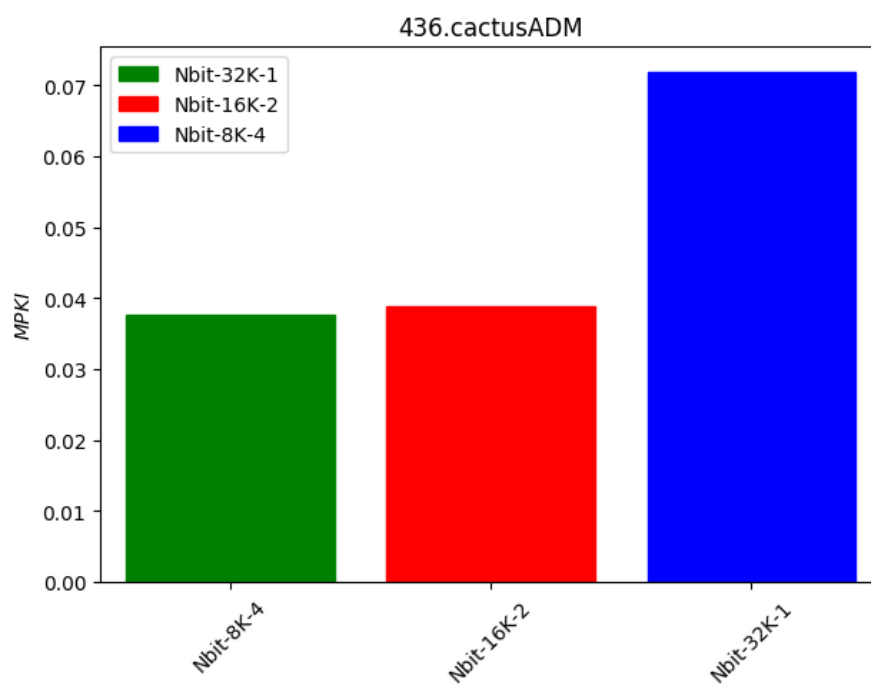
Αποτελέσματα των  $n$ -bits predictors για **Hardware** = 32K bits,  $n = \{1,2,4\}$  για 12 benchmarks ΣΥΝΟΛΙΚΑ

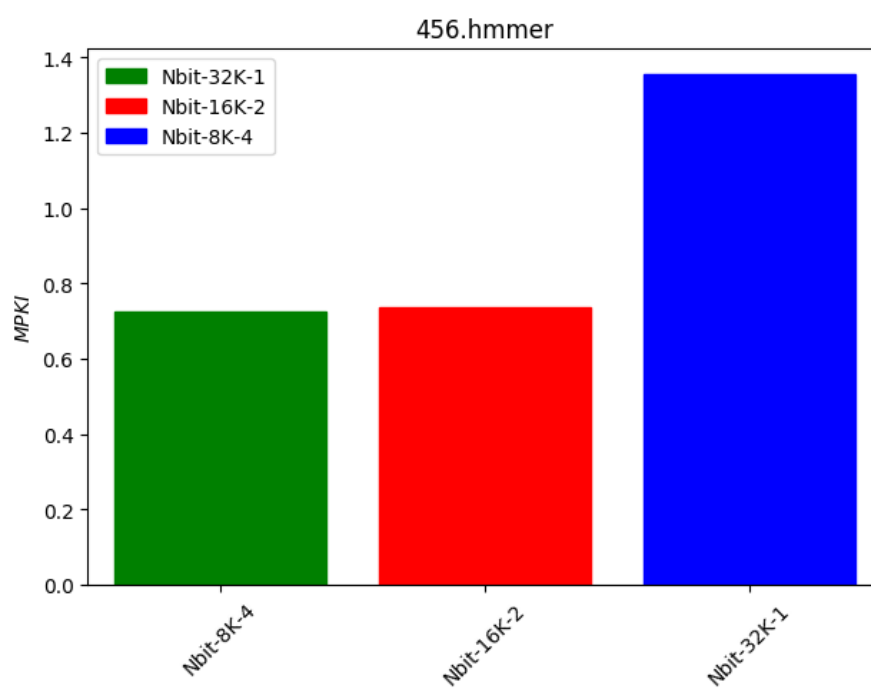
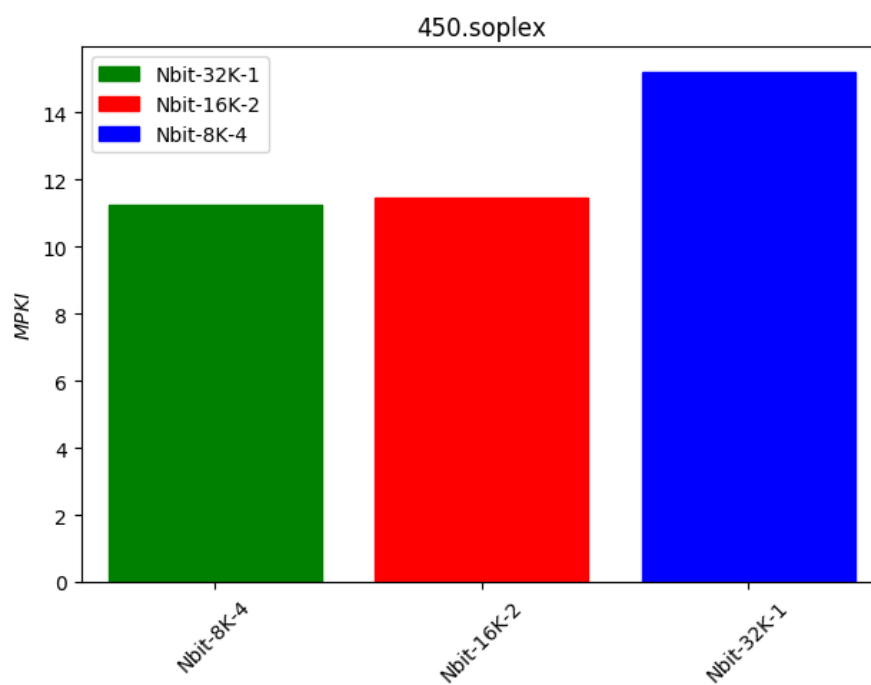


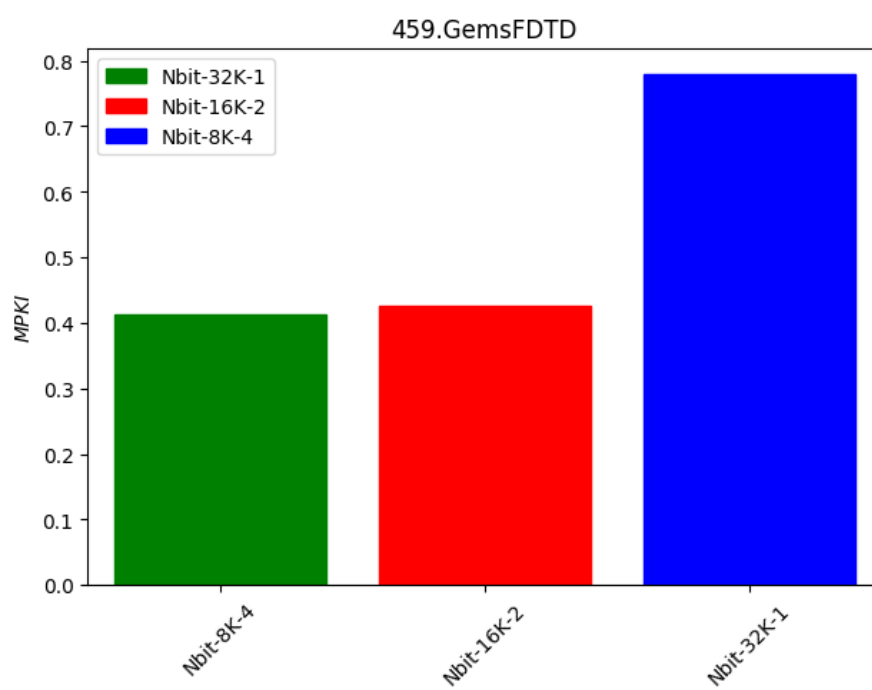
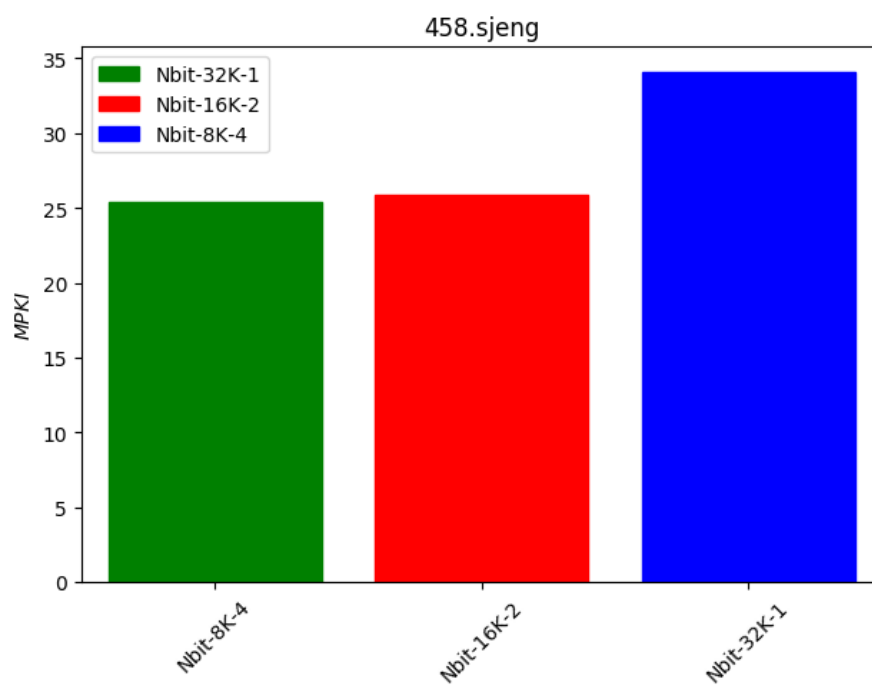
Παρακάτω παραθέτουμε ξεχωριστά για κάθε benchmark το αντίστοιχο διάγραμμα για τους  $n$ -bit predictors για **Hardware** = 32K bits,  $n = \{1,2,4\}$  με μεγαλύτερη λεπτομέρεια, για καλύτερη επισκόπηση και εξαγωγή συμπεσμάτων

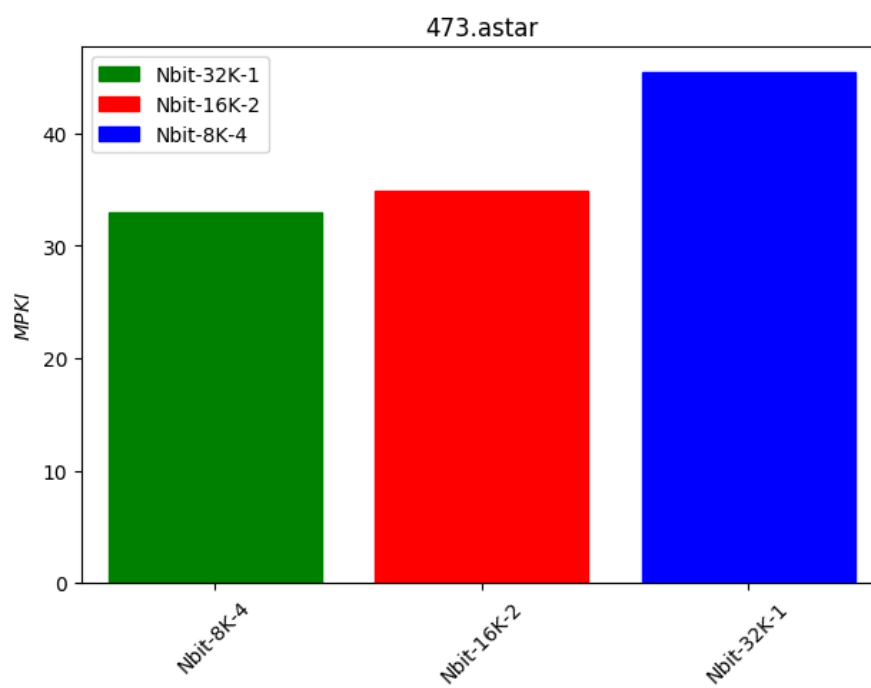
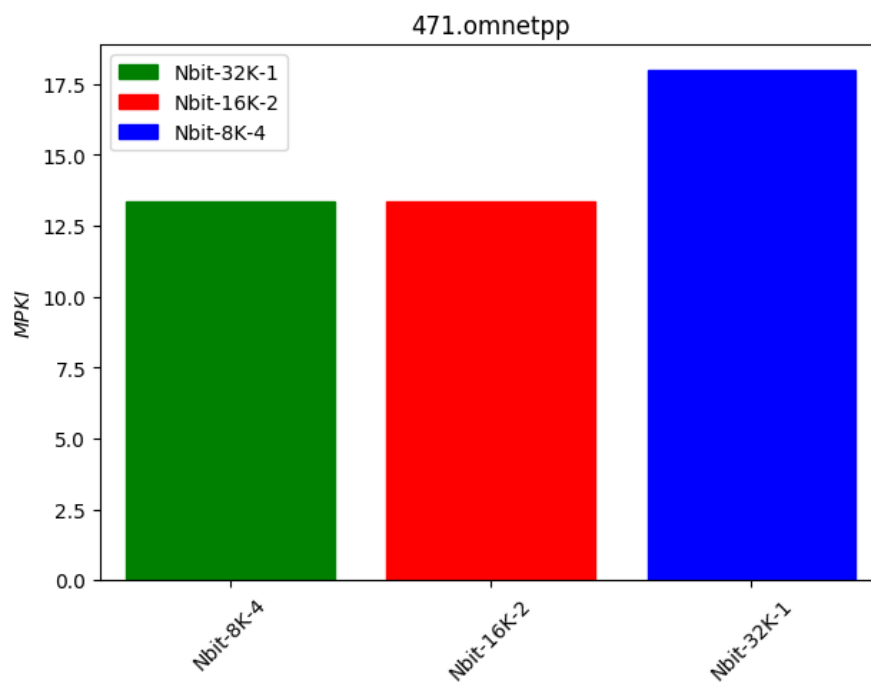


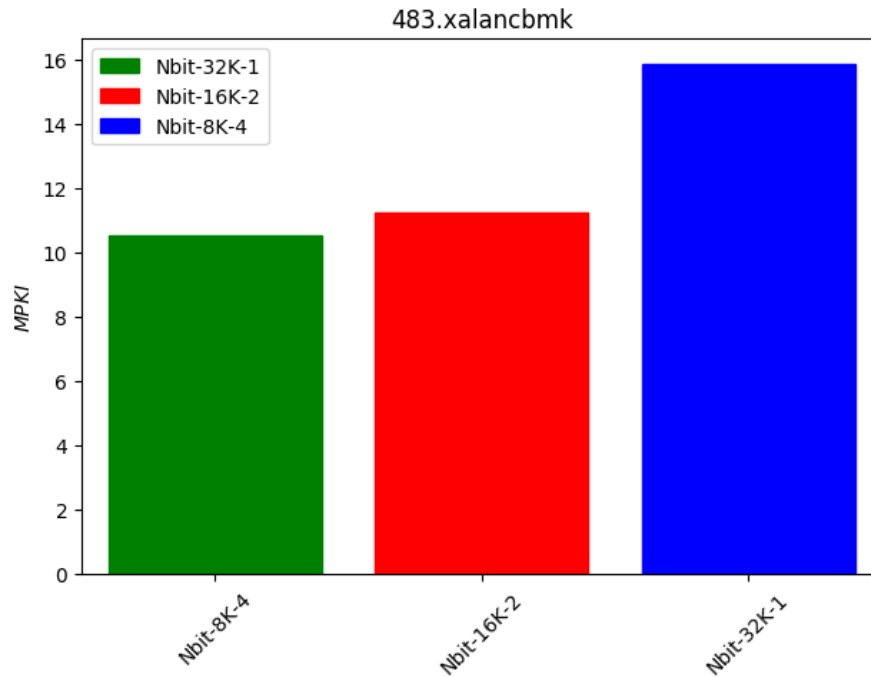












• **Συμπεράσματα:**

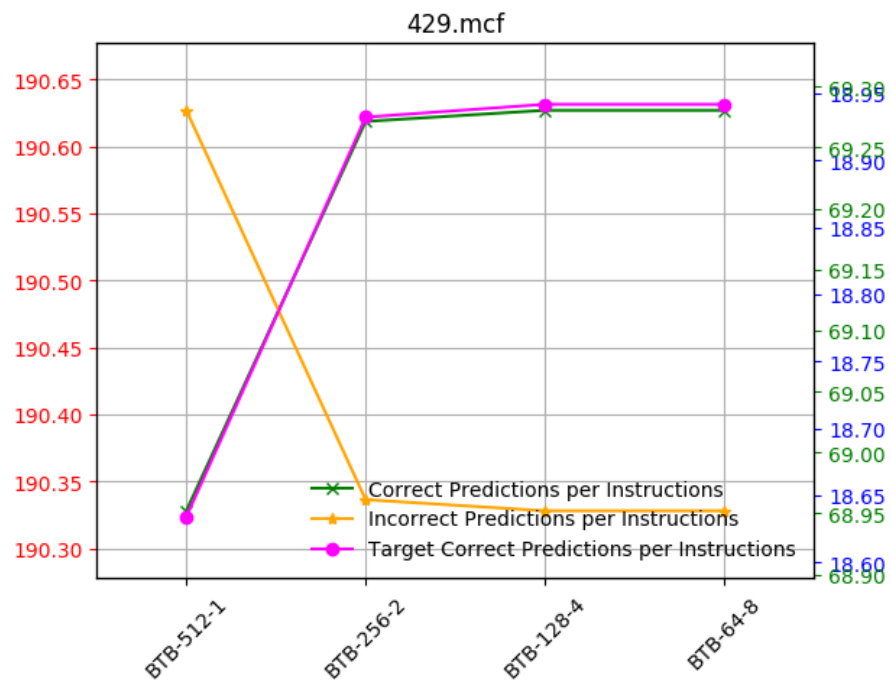
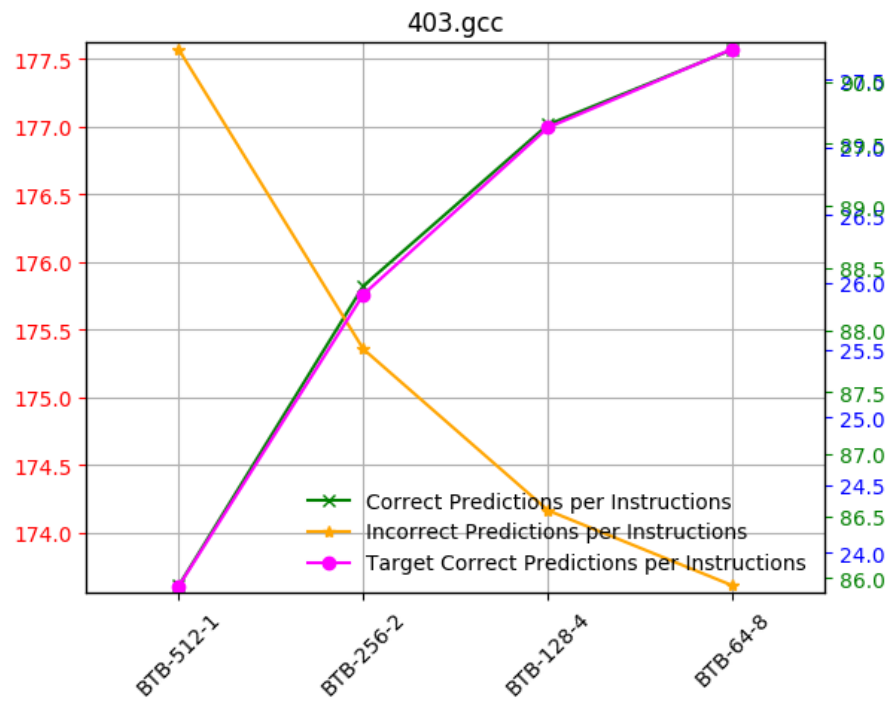
1. Σημειώνεται πως η αύξηση του bits συνδέεται άρρηκτα και πλήρως με την αύξηση του απαιτούμενου hardware που απαιτεί η εφαρμογή, καθώς ο αριθμός των BHT entries διατηρείται σταθερός.
2. Όπως και προηγουμένως ο **predictor για  $n = 1$  bit** έχει την χειρότερη απόδοση σε όλα τα benchmarks. Επιπλέον, η απόδοση για  $n = 2$  &  $n = 4$  bits κυμαίνεται στα ίδια επίπεδα, αλλά είναι σχετικά καλύτερη από αυτή για  $n = 1$  bit.
3. Βάσει της παραπάνω διερεύνησης και αποτελεσμάτων, επιλέγουμε τον **2-predictor** καθώς κατά μέσο όρο παρουσιάζει σχετικά καλύτερη απόδοση από όλους τους υπόλοιπους! Αξίζει να ανφερθεί πάντως ότι κατά μια γενική ομολογία, στην πραγματικότητα αυτό που χρησιμοποιείται είναι κυρίως ένα κράμα διαφορετικών τύπων predictors, όπως πχ υβριδικό από 2-bit counter και global predictor, το οποίο **βελτιώνει την απόδοση, αλλά οδηγεί σε αυξημένες καθυστερήσεις (tradeoff)**

### 3.3 Μελέτη του BTB

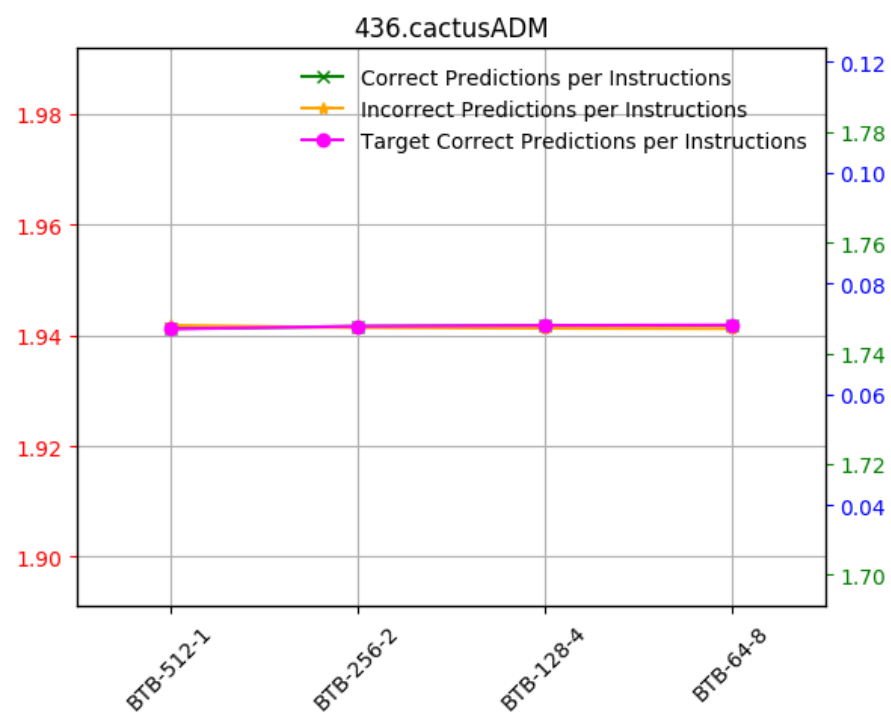
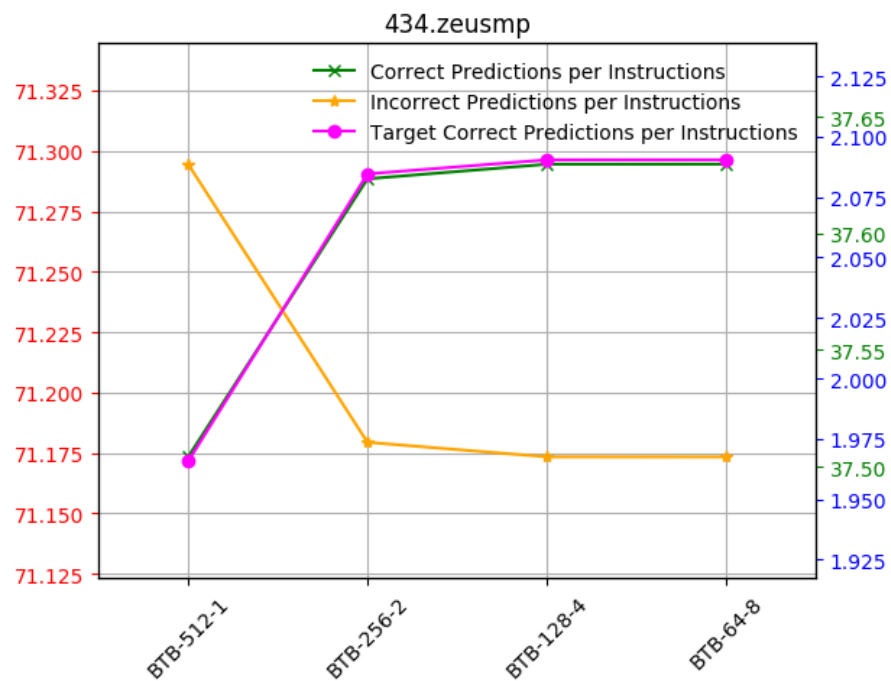
- Στο **Τρίτο Μέρος** της πειραματικής αξιολόγησης για την παρούσα άσκηση, εξετάζουμε την **ακρίβεια πρόβλεψης** κάνοντας χρήση ενός **Branch Target Buffer (BTB)**. Συγκεκριμένα, εκτελούμε προσομοιώσεις για τις παρακάτω περιπτώσεις:

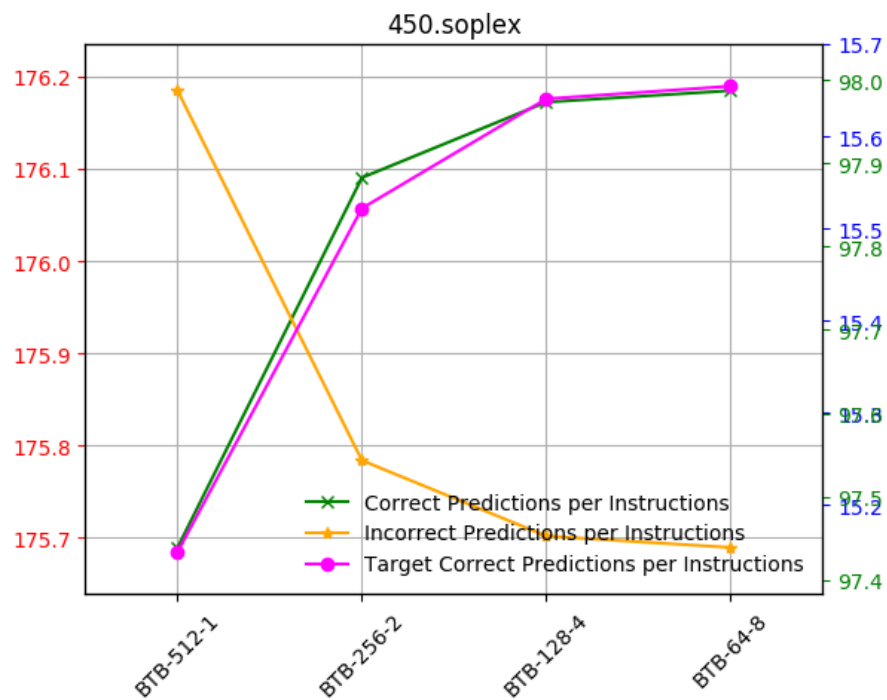
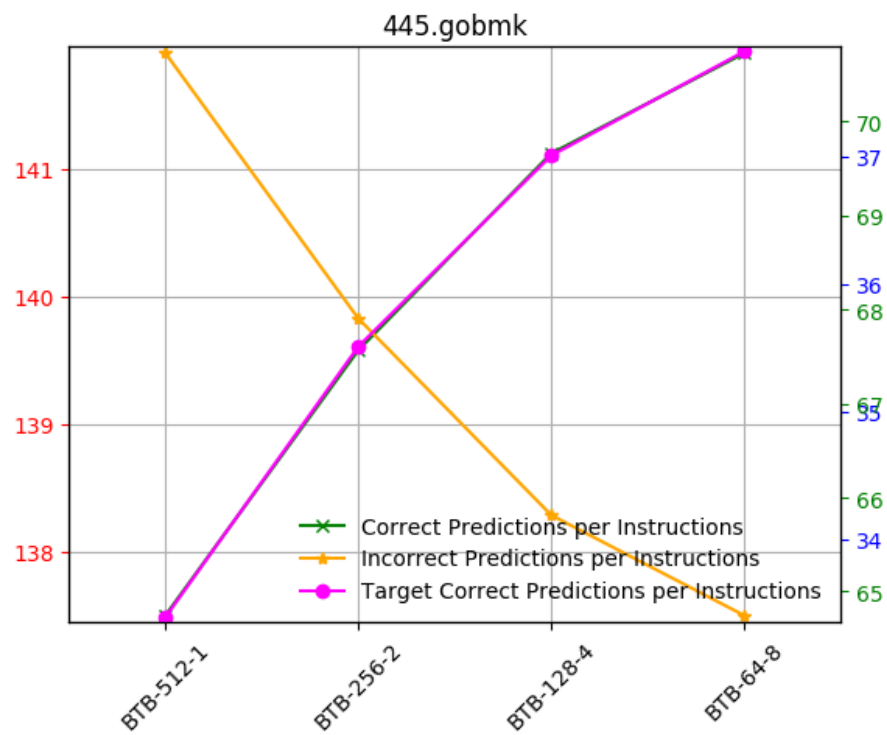
BTB entries	BTB associativity
512	1
256	2
128	4
64	8

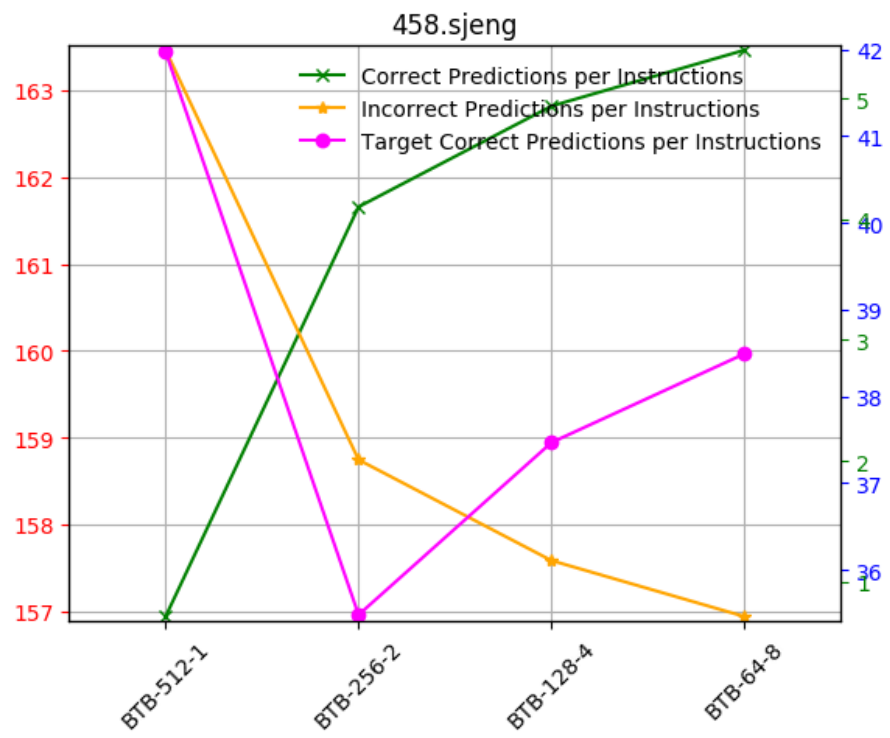
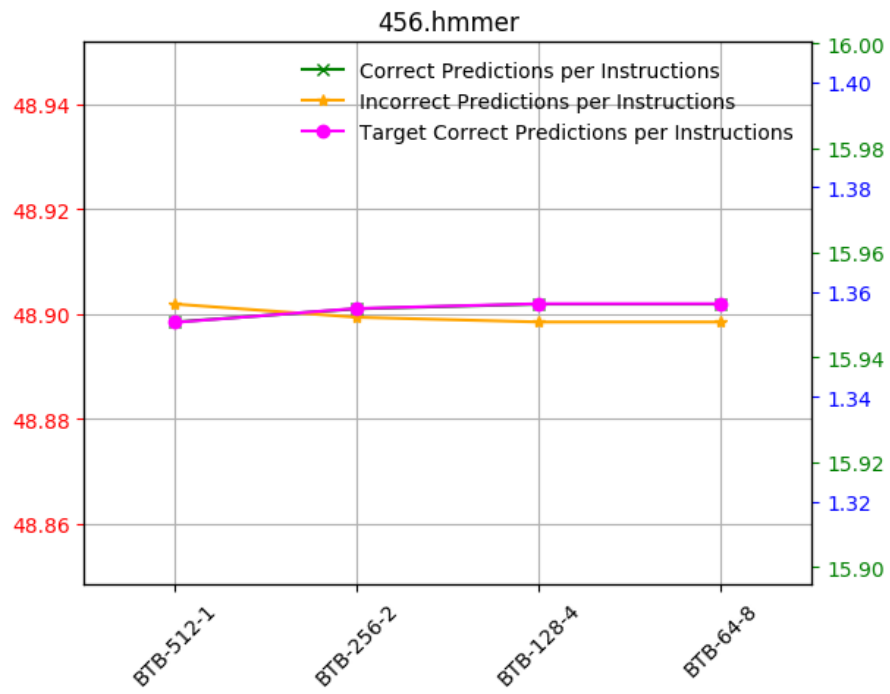
Παρακάτω παραθέτουμε ξεχωριστά για κάθε benchmark το αντίστοιχο διάγραμμα για **[BTB entries, BTB associativity] = [512-1, 256-2, 128-4, 64-8]** με μεγαλύτερη λεπτομέρεια, για καλύτερη επισκόπηση και εξαγωγή συμπερασμάτων

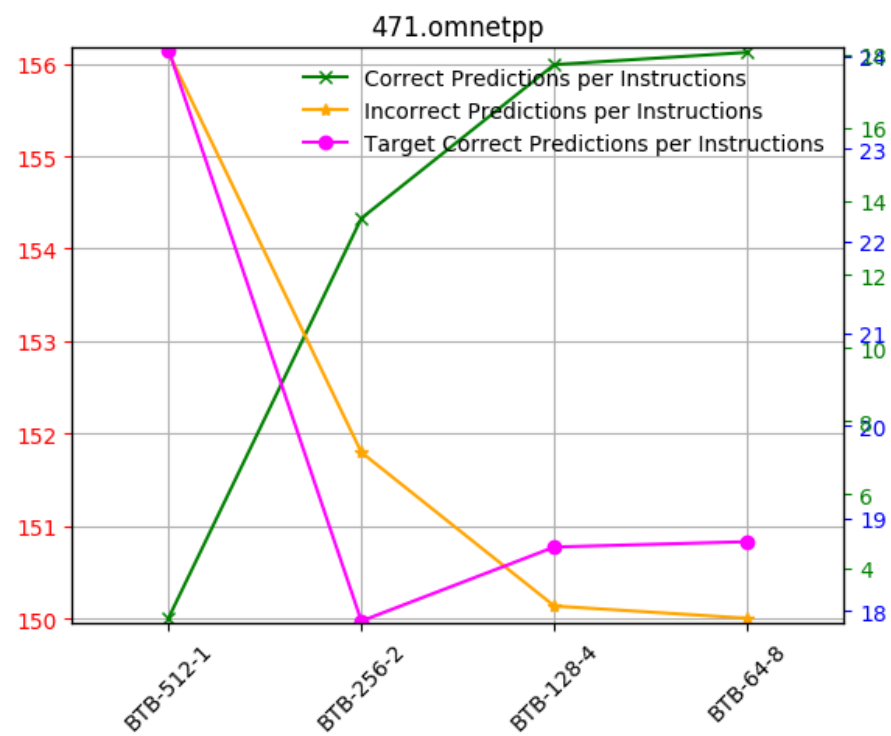
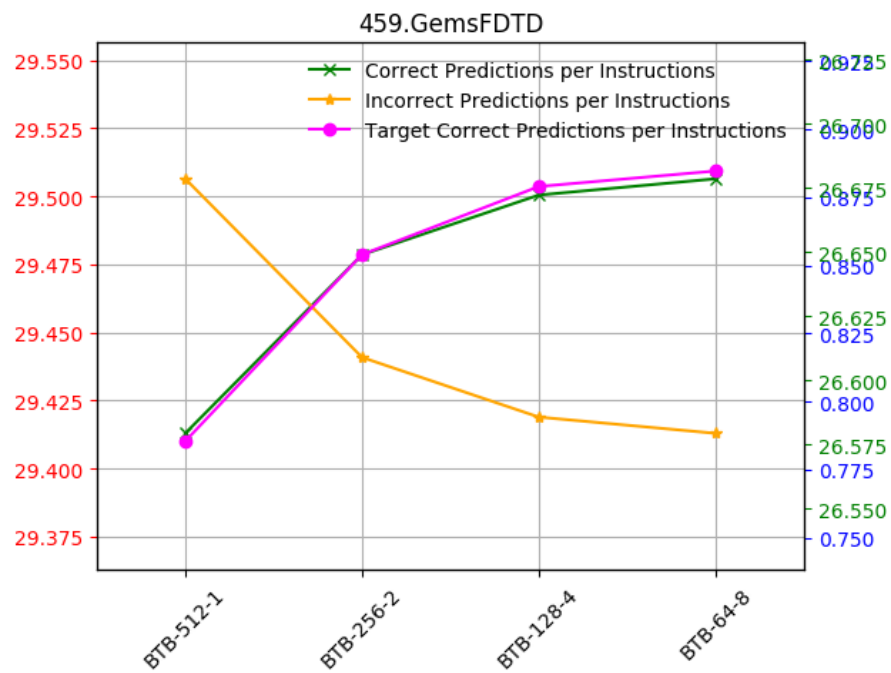


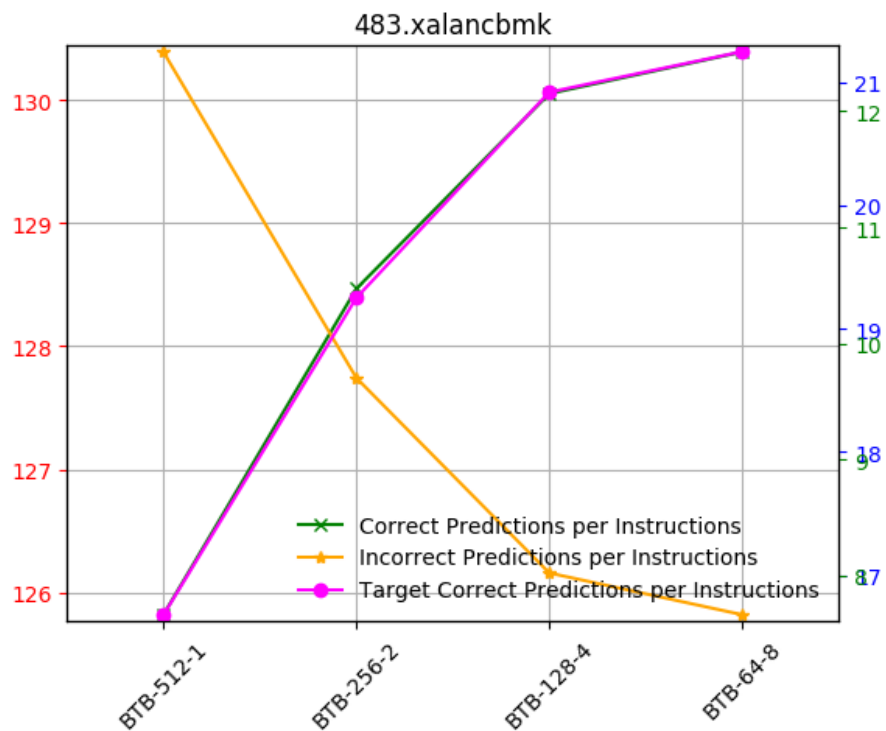
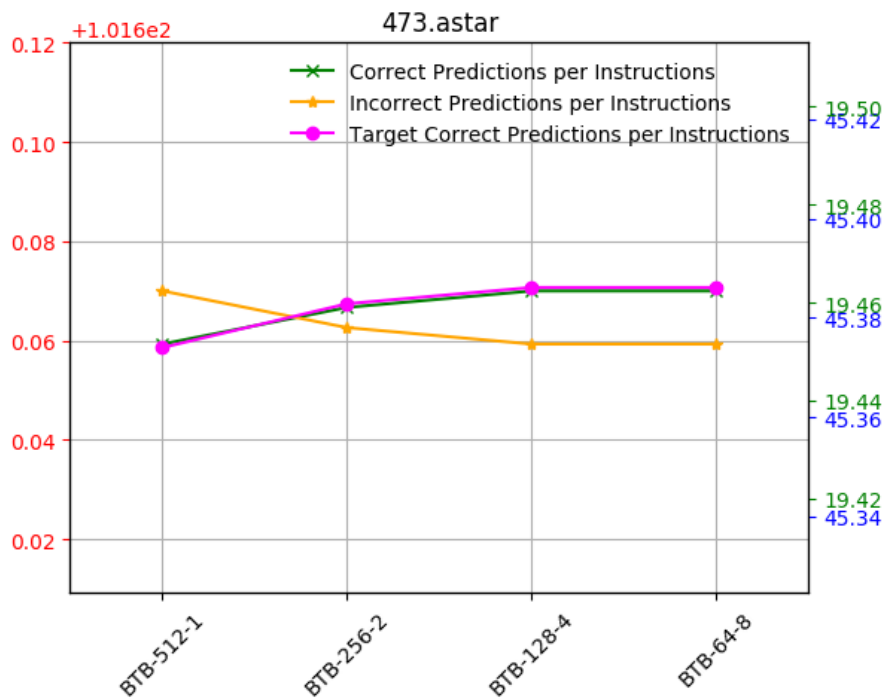












• Συμπεράσματα:

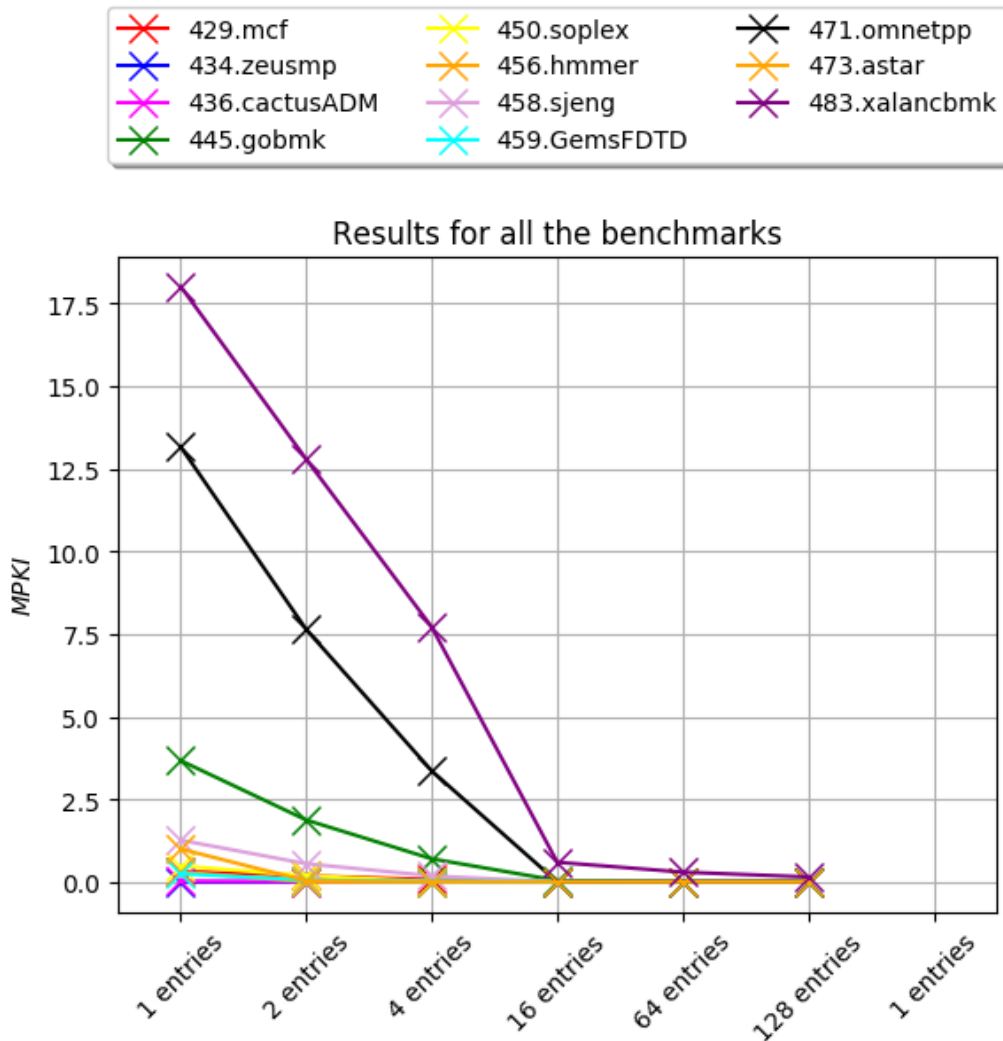
1. Αρχικά να σημειωθεί πως τα BTB entries έχουν **μεγαλύτερο υπολογιστικό κόστος** από BHT entries , αλλά κάνουν redirect fetches σε πολύ νωρίτερο στάδιο στο pipeline και μπορούν να **επιταχύνουν τα indirect branches**. Δεν ξεχνάμε όμως ότι τα BHT entries διατηρούν περισσότερα entries και έχουν μεγαλύτερη ακρίβεια.
2. Παρ'όλα αυτά η συγκεκριμένη μέθοδος θεωρείται αρκετά **αποτελεσματική**. Προφανώς βέβαια υπάρχουν σφάλματα λανθασμένου στόχου διακλάδωσης, τα οποία οφείλονται στην μη αποθήκευση τους στον πίνακα παρά την σωστή απόφαση για την πορεία της διακλάδωσης.
3. Οι predictors φτάνουν παρόμοια επίπεδα απόδοσης κατά μέσο όρο για οποιονδήποτε απο τους 4 παραπάνω συνδυασμούς entries και associativity.
4. Τέλος, αξίζει να σημειωθεί ο **τρόπος αντικατάστασης** για υψηλό associativity: Η κάθε εντολή θα πρέπει να αντικατασταθεί αυτόματα από κάποια άλλη στο BTB πίνακα με **ίδια διεύθυνση**. Αντιθέτως, απαιτούνται τόσες εντολές branch (με ίδια θέση στον πίνακα BTB) = associativity

### 3.4 Μελέτη του RAS

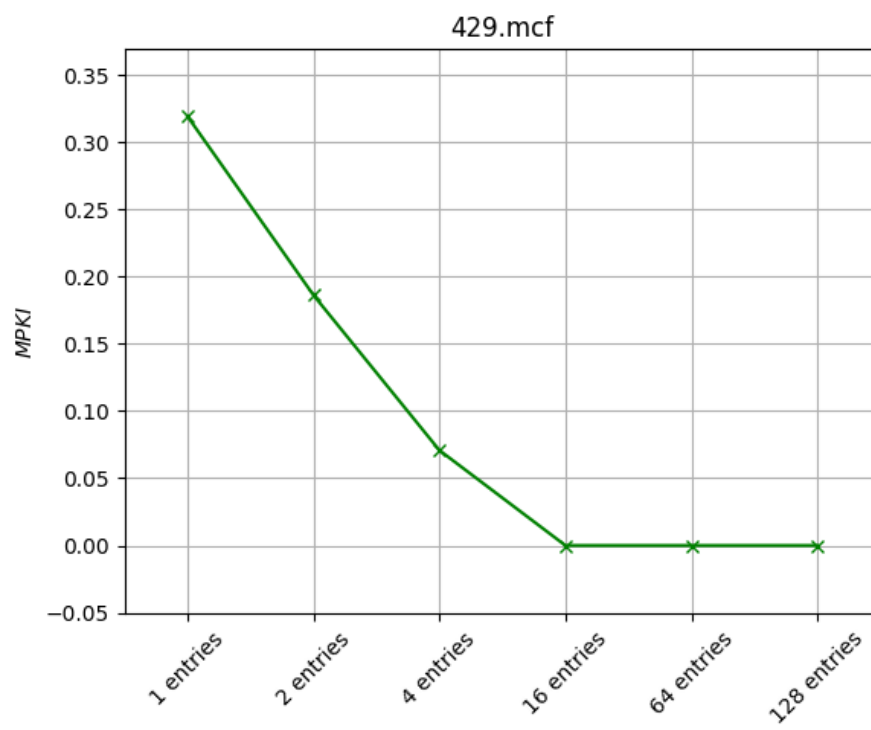
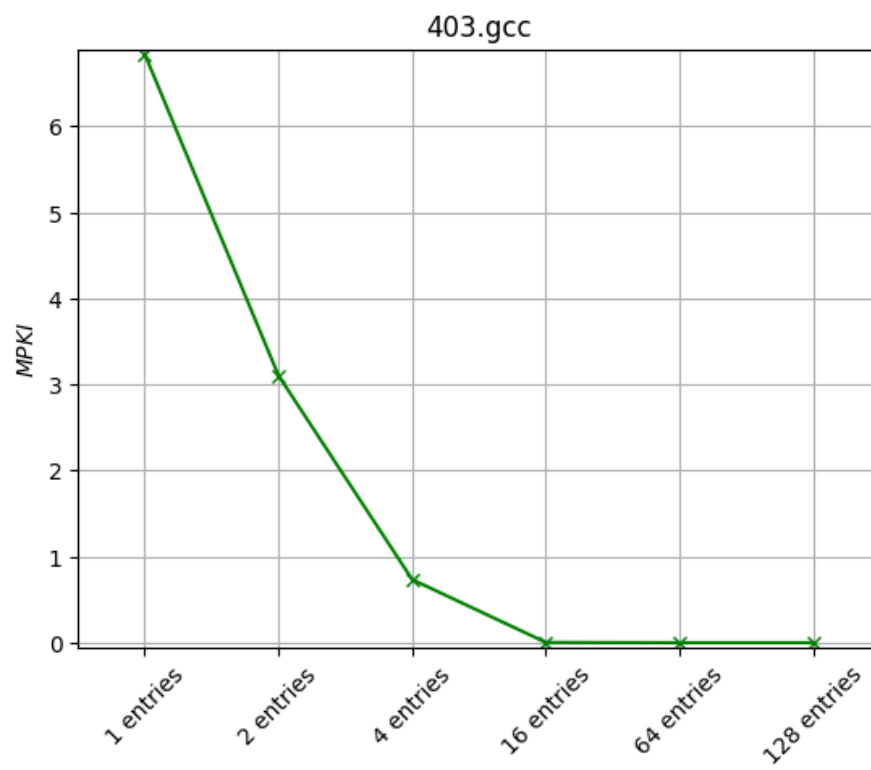
- Στο **Τέταρτο Μέρος** της πειραματικής αξιολόγησης για την παρούσα άσκηση, εξετάζουμε το ποσοστό αστοχίας που προκύπτει κάνοντας χρήση διαφορετικού αριθμού εγγράφων RAS.
- Για την υλοποίηση της RAS, η οποία είναι πολύ απλή και αποτελεί μέθοδο για την **πρόβλεψη εντολών return**, χρησιμοποιούμε στο **"ras.h"** (με μικρές προσαρμογές) μία στοίβα (FILO) η οποία λειτουργεί ως εξής: Σε κάθε κλήση προσθέτουμε διευθύνσεις και σε κάθε *return* αφαιρούμε.
- Συγκεκριμένα, εκτελούμε προσομοιώσεις για τις παρακάτω περιπτώσεις:

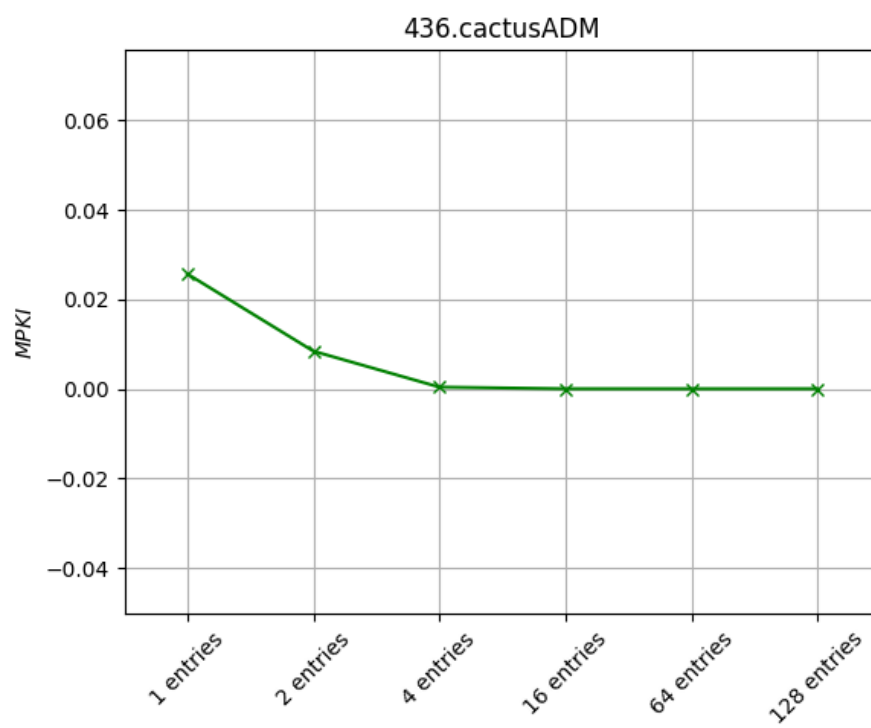
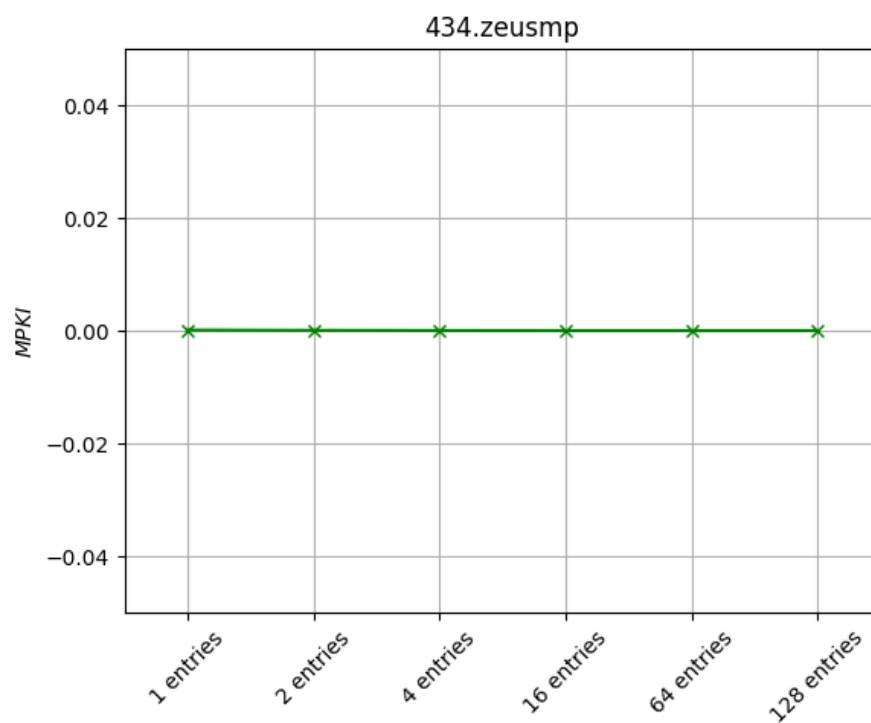
Εγγραφές στην RAS
1
2
4
8
16
64
128

Αποτελέσματα των *n*-bits predictors για  $\#RAS = \{1, 2, 4, 8, 16, 64, 128\}$  για 12 benchmarks ΣΥΝΟ-ΛΙΚΑ

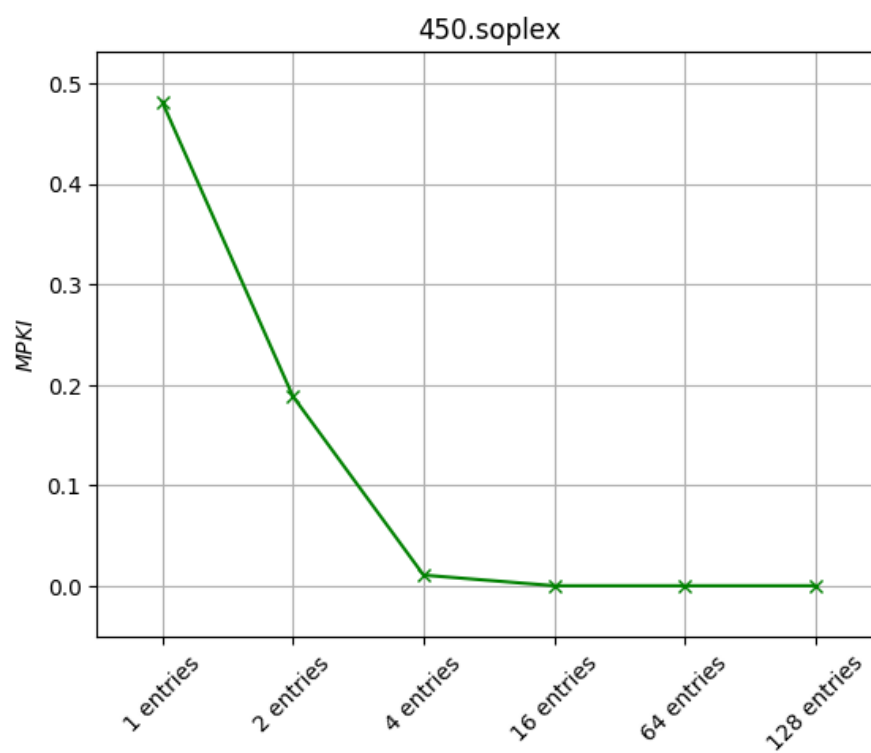
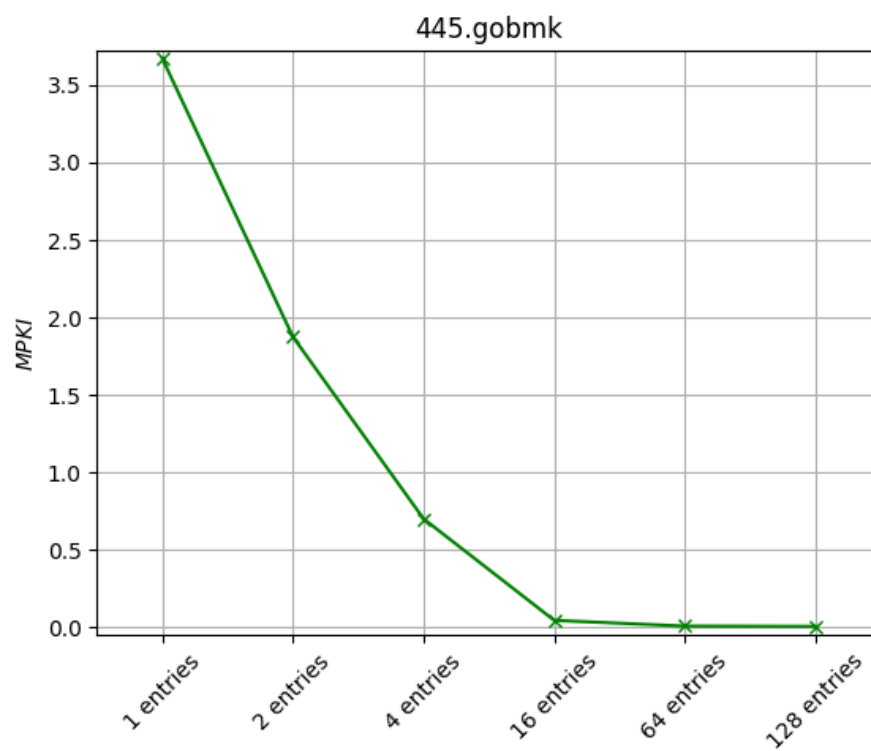


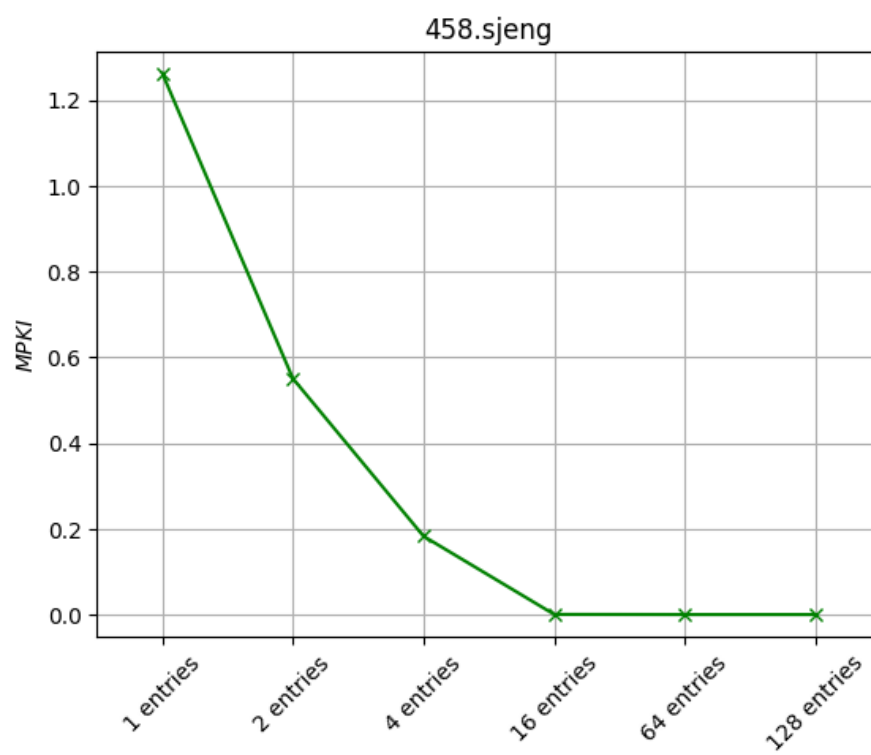
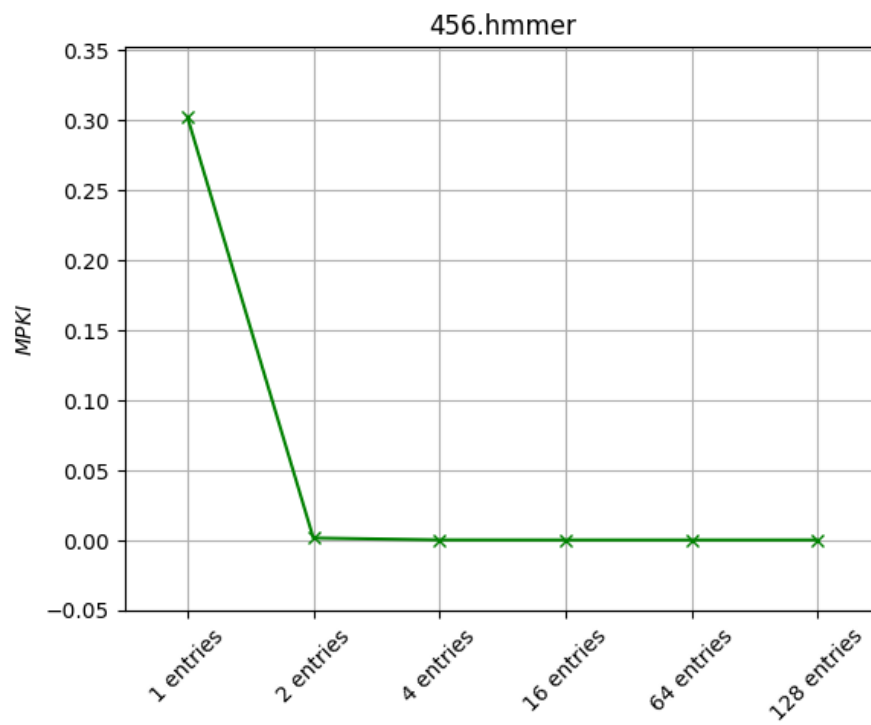
Παρακάτω παραθέτουμε ξεχωριστά για κάθε benchmark το αντίστοιχο διάγραμμα για  $\#RAS = \{1, 2, 4, 8, 16, 64, 128\}$  με μεγαλύτερη λεπτομέρεια, για καλύτερη επισκόπηση και εξαγωγή συμπεσμάτων

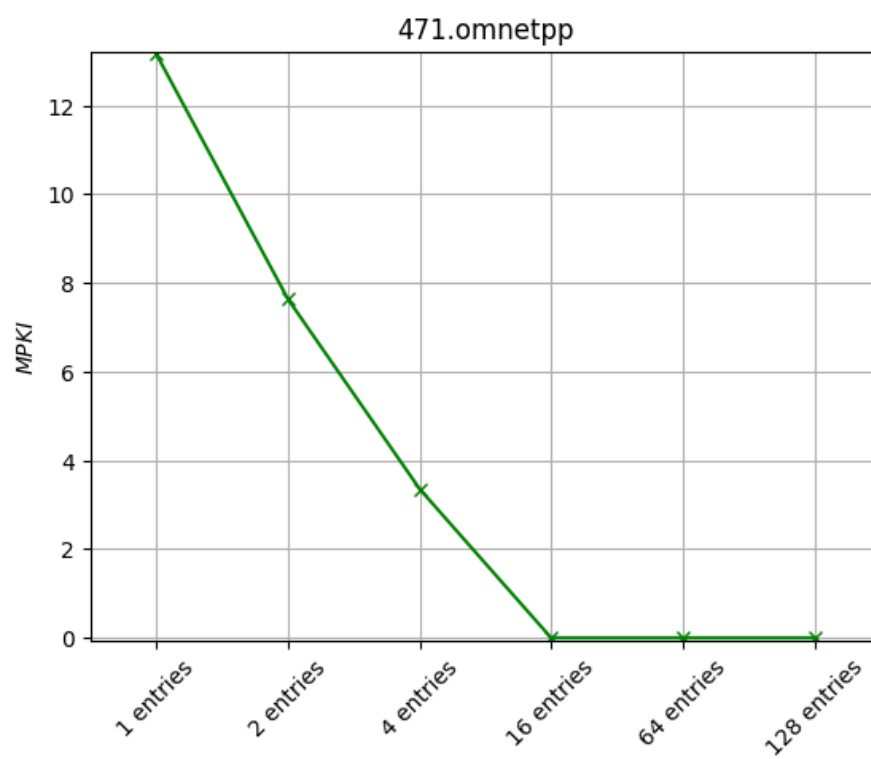
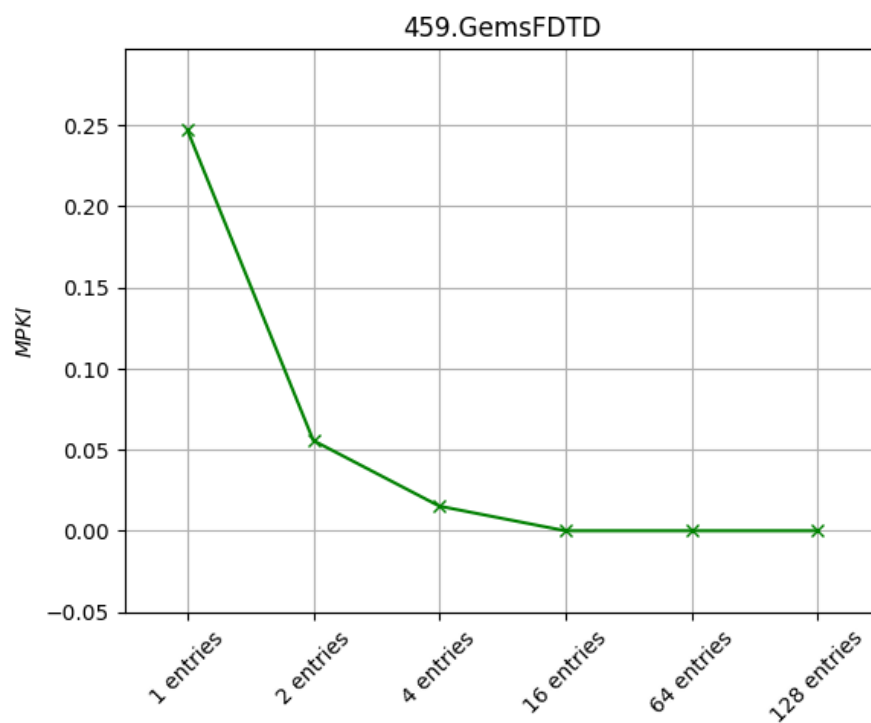


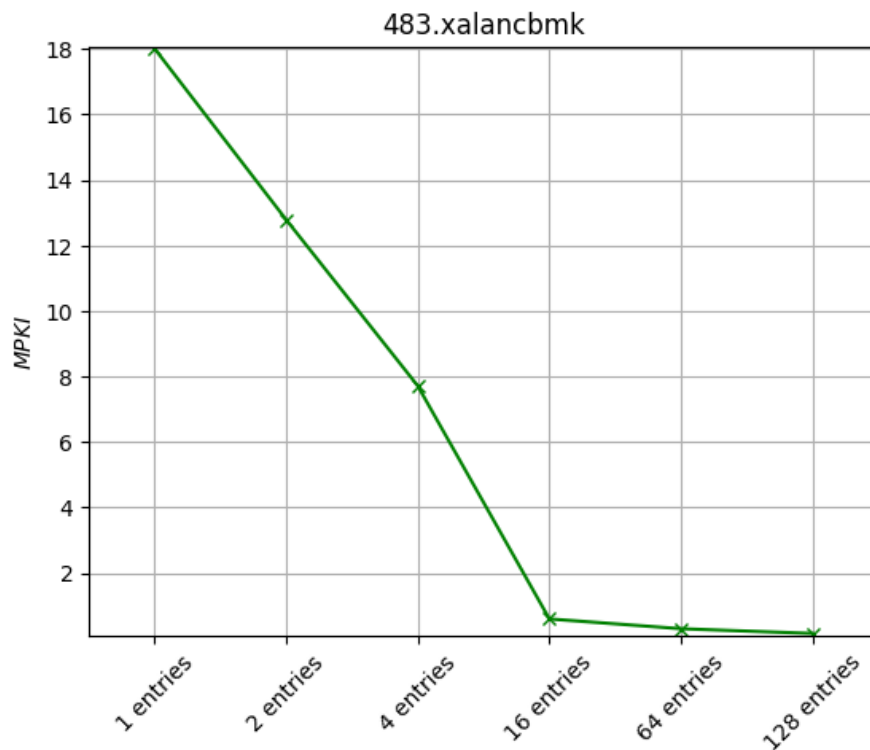
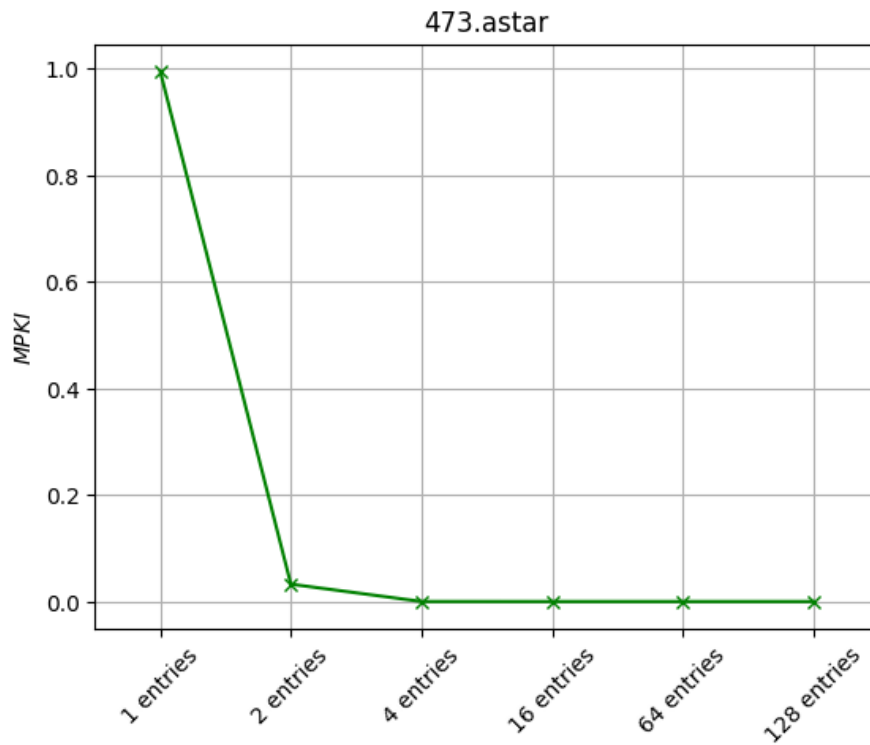












• Συμπεράσματα:

1. Κατά μια γενική ομολογία και κάνοντας μια πολύ απλή σύγκριση με τις γραφικές MPKI που αφορούν την (3.2) Ενότητα - Μελέτη n-bit predictors, παρατηρούμε χαμηλότερα επίπεδα κατά μέσο όρο, με τα χειρότερα benchmarks να αγγίζουν 18% σε αντίθεση με προηγούμενως που κάποια ξεπερνούσαν και το 35%.
2. Όταν  $\#RAS = 1$ , έχουμε παρ' όλα αυτά πολύ μεγάλο MPKI σχεδόν για όλα τα benchmarks, ενώ όταν αυξήσουμε τον αριθμό αυτόν σε 2, παρατηρείται **μεγάλη μείωση του δείκτη**, ενώ για περισσότερα από 2 έγγραφα σημειώνεται αρκετά μικρότερων κλιμακίων βελτίωση.
3. Τονίζουμε ότι οι παραπάνω προσομοιώσεις ήταν καταλυτικές για να κατανοήσουμε ότι η αύξηση

του  $\# RAS$  δεν μπορεί να οδηγήσει σε μείωση της απόδοσης σε καμία των περιπτώσεων, άρα παρατηρούμε μόνο διατήρηση ή και βελτίωση της απόδοσης

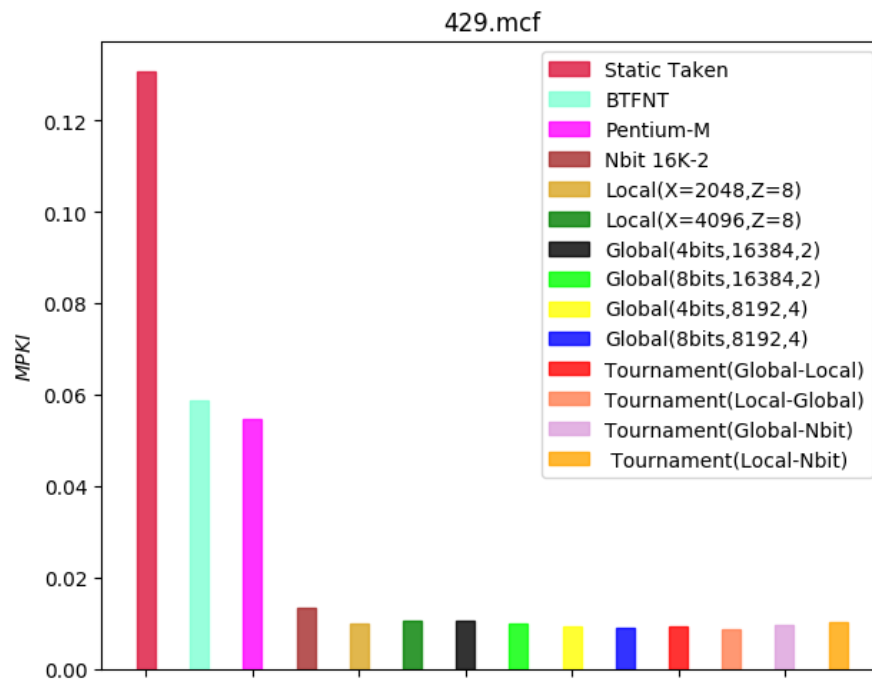
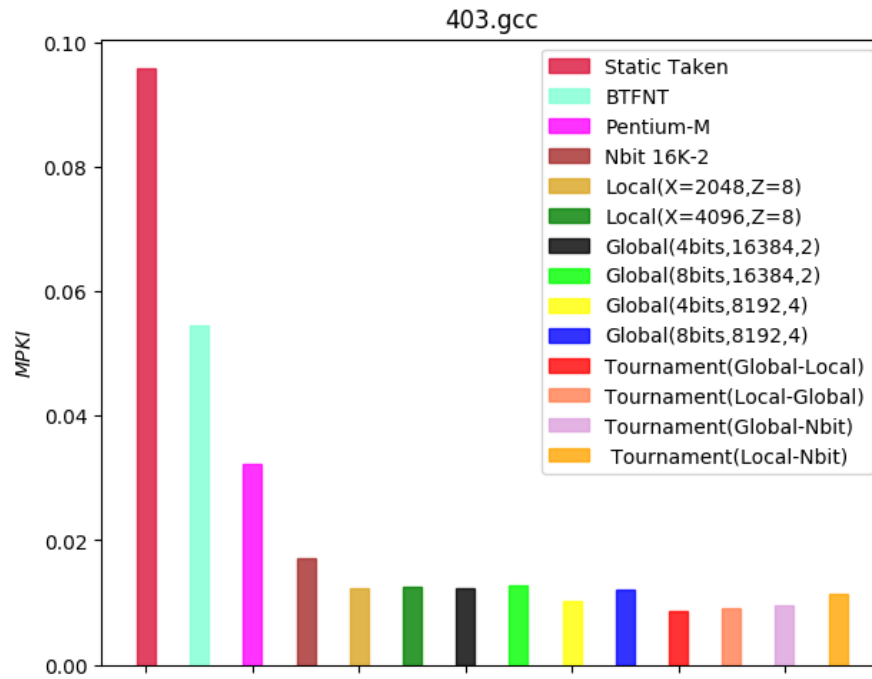
4. Με αντικειμενικό κριτήριο το κόστος του υλικού, επιλέγουμε υλοποίηση που κάνει χρήση  $\# RAS = 4$  ή  $8$ . (εφόσον δεν παρατηρείται και μεγάλη βελτωση με χρήση περισσότερων εγγράφων)

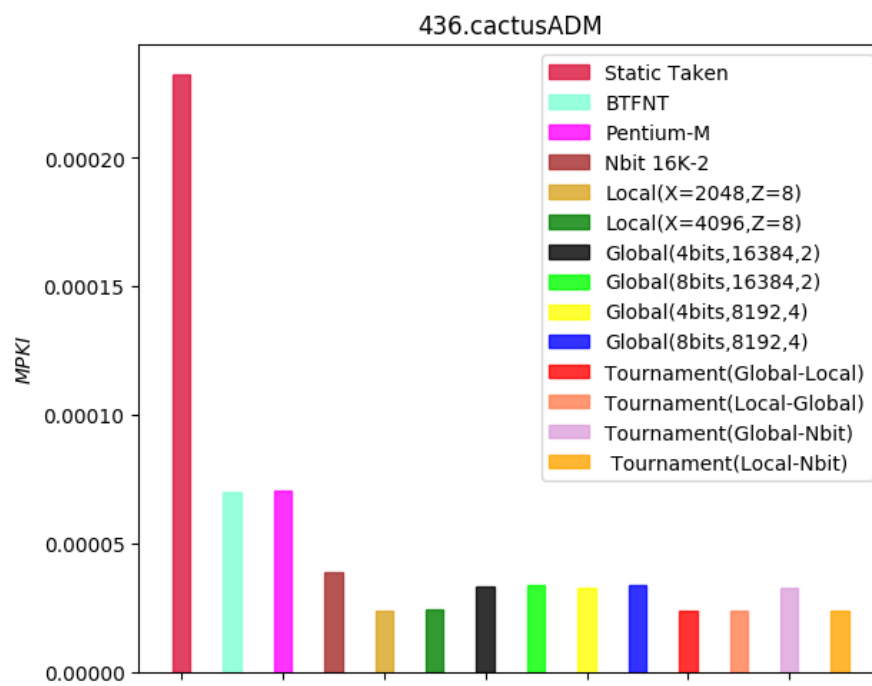
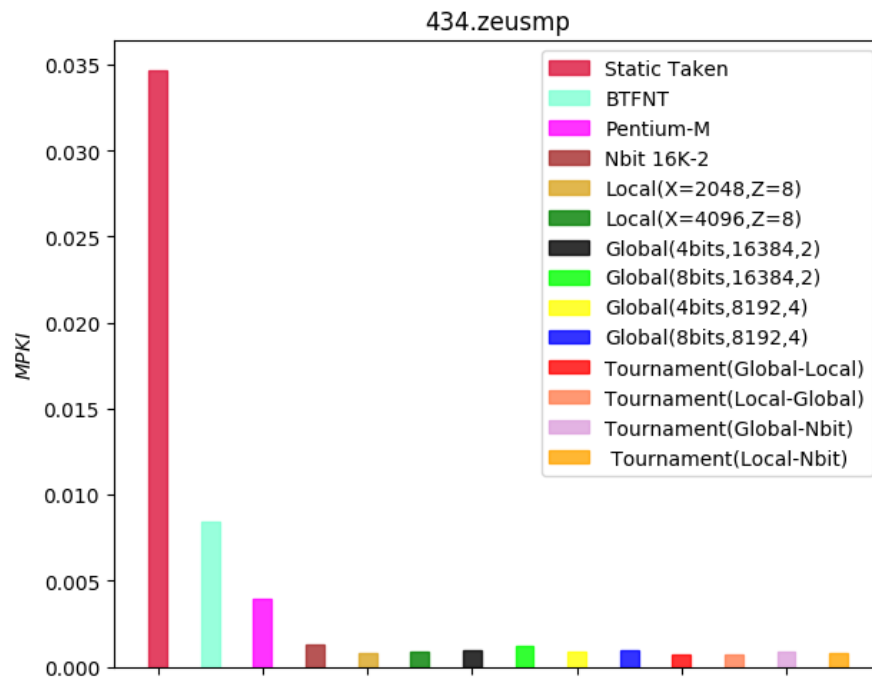
### 3.5 Σύγκριση διαφορετικών predictors

- Στο Πέμπτο Μέρος και τελευταίο της πειραματικής αξιολόγησης για την παρούσα άσκηση, εξετάζουμε την απόδοση διαφορετικών predictors. Συγκεκριμένα

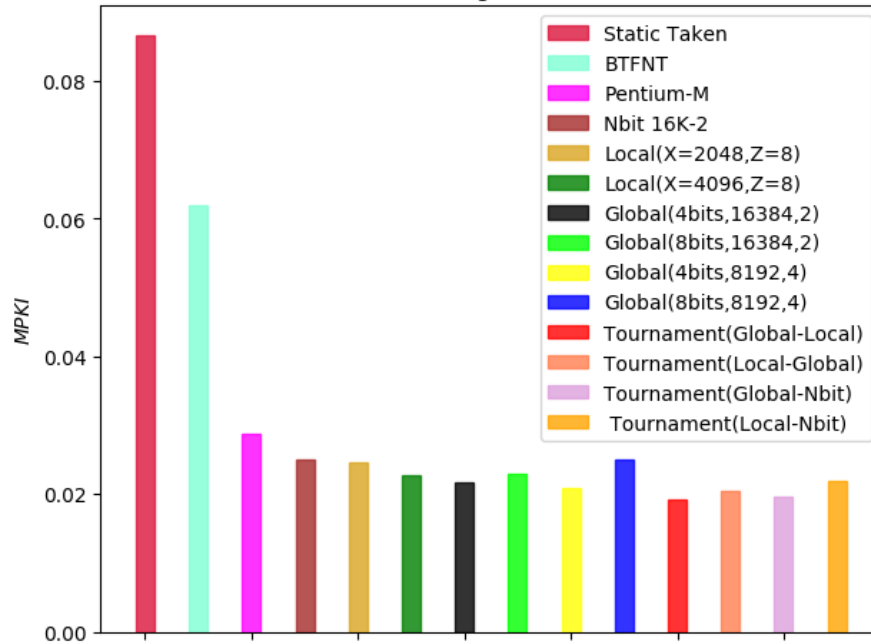
<b>Predictors to compare</b>
<b>Static Taken</b>
<b>Pentium-M</b>
<b>Static BTFNT</b>
<b>Best N-bit predictor from part 4.2 → 16K-2</b>
<b>Local-History 2-level</b> [ PHT entries = 8192, PHT n-bit counter length = 2, BHT entries = 2048, BHT entry length = 8 ]
<b>Local-History 2-level:</b> [ PHT entries = 8192, PHT n-bit counter length = 2, BHT entries = 4096, BHT entry length = 4 ]
<b>Global-History 2-level:</b> [ PHT entries = 16384, PHT n-bit counter length = 2, BHR length = 4 ]
<b>Global-History 2-level:</b> [ PHT entries = 16384, PHT n-bit counter length = 2, BHR length = 8 ]
<b>Global-History 2-level:</b> [ PHT entries = 16384, PHT n-bit counter length = 4, BHR length = 4 ]
<b>Global-History 2-level:</b> [ PHT entries = 16384, PHT n-bit counter length = 4, BHR length = 8 ]
<b>Tournament Hybrid predictor Global(optimal)-Local(optimal)</b>
<b>Tournament Hybrid predictor Local(optimal)-Global(optimal)</b>
<b>Tournament Hybrid predictor Local(optimal)-N-bit(optimal)</b>
<b>Tournament Hybrid predictor GLocal(optimal)-N-bit(optimal)</b>

Αποτελέσματα MPKI για όλους τους παραπάνω predictors
--

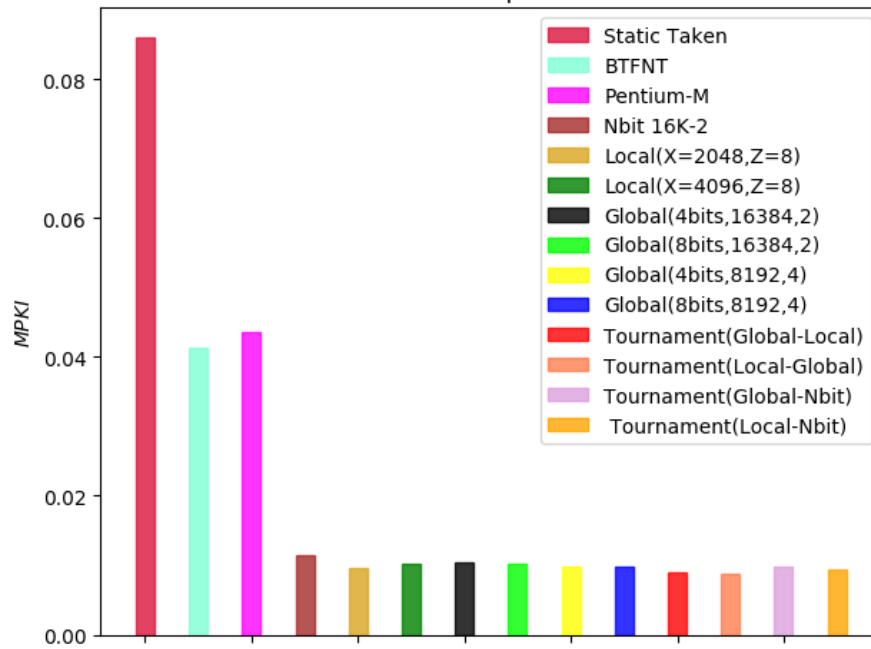




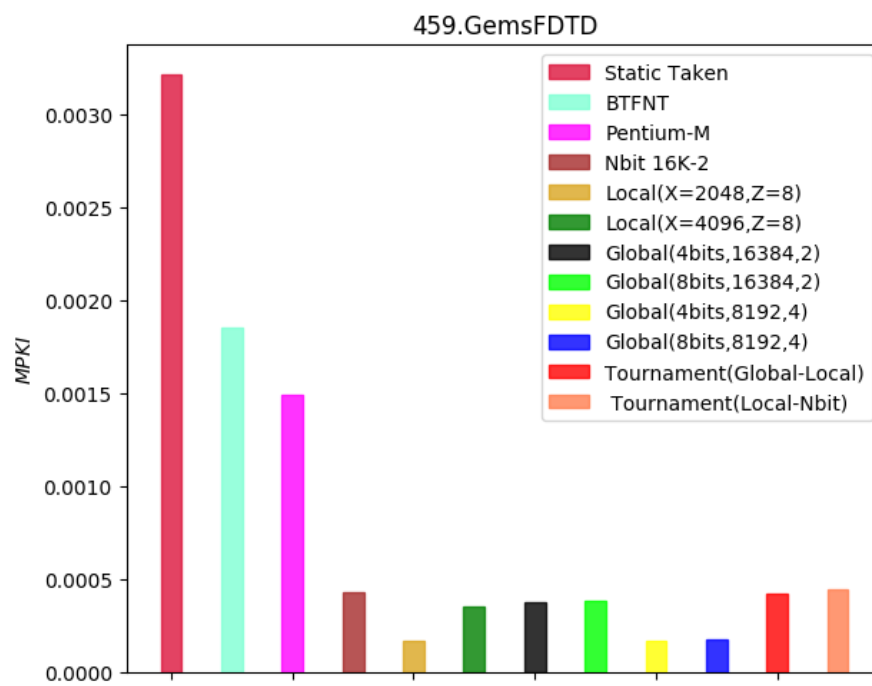
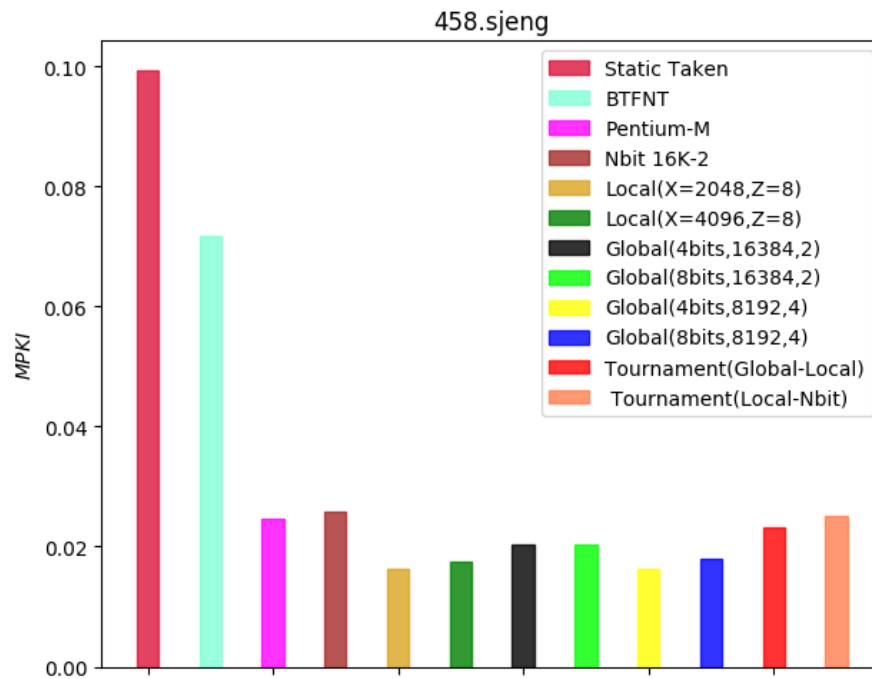
445.gobmk

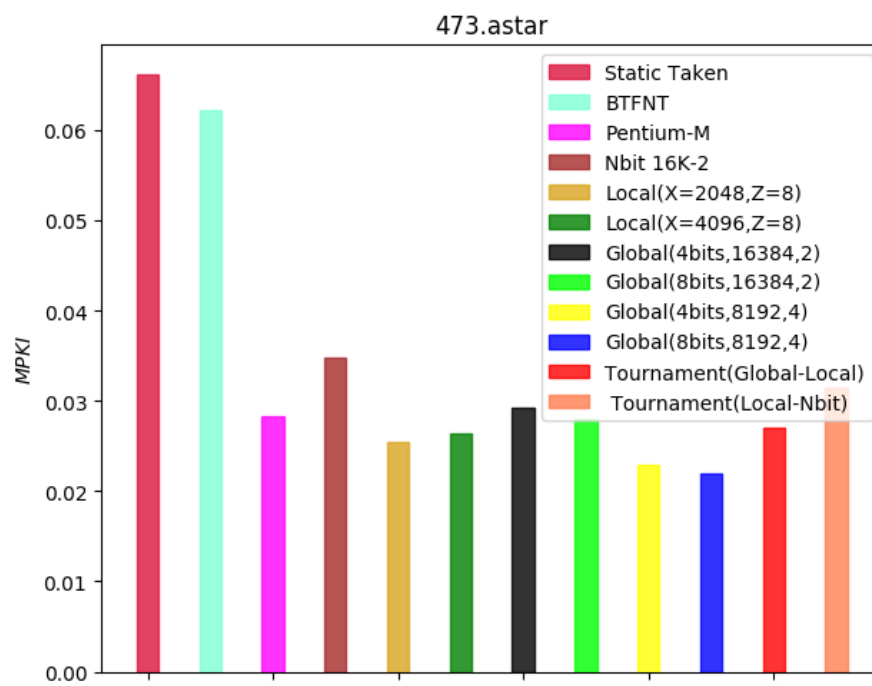
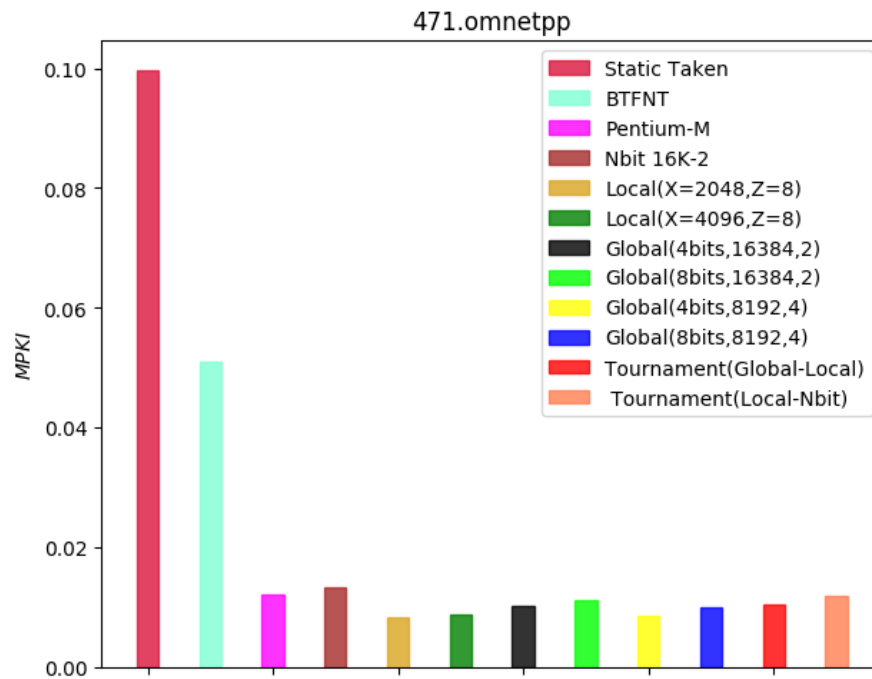


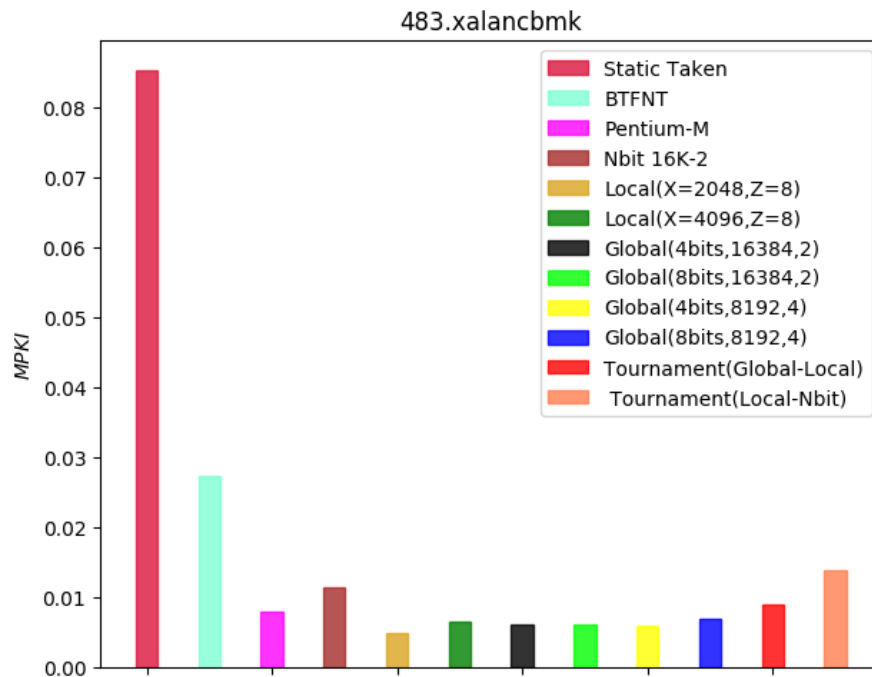
450.soplex











• **Συμπεράσματα:**

1. Ο **Static Taken predictor** έχει την χειρότερη δυνατή απόδοση για όλα τα benchmarks , και αυτό γιατί αλγοριθμικά θεωρεί ότι τα branch δεν θα εκτελεστούν ενώ έχουμε μεγάλου βρόγχους.
2. Ο **BTFNT predictor** παρουσιάζει κατά γενική ομολογία καλύτερη απόδοση για όλα τα benchmarks συγκριτικά με **Static Taken predictor**, αλλά τα επίπεδα του δείκτη MKPI παραμένουν ακόμα αρκετά υψηλά
3. Εν συνεχεία οι **N-bit**, **Pentium-M** και όλοι οι **Global History** και **Local History** οδηγούν σε πολύ χαμηλότερες τιμές του δείκτη MKPI συγκριτικά με τους 2 προαναφερθέντες, αλλά δεν παρουσιάζεται σαφές pattern για εξαγωγή άμεσου αποτελέσματος για την απόδοσή τους, οπότε η ταξινόμηση τους θα γίνεται για την απόδοσή του κατά μέσο όρο για όλα τα benchmarks.
4. Οι **tournament predictors**, παρουσιάζουν επίσης καλή απόδοση , είναι υβριδικοί και χρησιμοποιούν μια συνδυασμένη έκδοση **Global**, **Local** και **N-bit** predictors.

*Ταξινόμηση των predictors με αντικειμενικό κριτήριο την κατά μέσο όρο απόδοσή τους πάνω στα 12 benchmarks*

1. **Global History (8 bits)** predictor
2. **Local History (8 bits)** predictor
3. **Local History (4 bits)** predictor
4. **Global History (4 bits)** predictor
5. **Tournament (Local - Global)** predictor
6. **Tournament (Global - Local)** predictor
7. **Tournament (Global - Nbit)** predictor
8. **Tournament (Local - Nbit)** predictor
9. **N-bit 16K-2** predictor
10. **Pentium-M** predictor
11. **Static BTFNT** predictor
12. **Static Taken** predictor