



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Σχολή Μηχανολόγων Μηχανικών
Τομέας Ρευστών
Εργαστήριο Θερμικών Στροβιλομηχανών

Διδακτορική Διατριβή

**ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΠΡΟΒΛΗΜΑΤΩΝ ΑΕΡΟΔΥΝΑΜΙΚΗΣ-
ΑΕΡΟΕΛΑΣΤΙΚΟΤΗΤΑΣ ΣΕ ΕΠΕΞΕΡΓΑΣΤΕΣ ΚΑΡΤΩΝ ΓΡΑΦΙΚΩΝ**

Ξενοφών Σ. Τρομπούκη

Αθήνα 2012

Επιβλέπων : **Κυριάκος Χ. Γιαννάκογλου**, Καθηγητής ΕΜΠ

Περίληψη

Σκοπός της παρούσας διδακτορικής διατριβής είναι η δημιουργία μεθόδων και λογισμικού για την επίλυση 2Δ/3Δ προβλημάτων αεροδυναμικής/αεροελαστικότητας. Βασικός στόχος είναι η σημαντική μείωση του υπολογιστικού κόστους και του χρόνου αναμονής του μηχανικού, κατά τη φάση ανάλυσης ή σχεδιασμού αεροδυναμικών μορφών, ιδίως σε προβλήματα μεγάλης κλίμακας, με τη βέλτιστη χρήση των επεξεργαστών καρτών γραφικών (GPUs).

Τα τελευταία 10 χρόνια οι GPUs, έχουν εξελιχθεί σε πολυπύρηνες παράλληλες μονάδες επεξεργασίας κοινής μνήμης, με υπολογιστική ισχύ υπερδεκαπλάσια οποιασδήποτε σύγχρονης κεντρικής μονάδας επεξεργασίας (CPU). Οι GPUs που χρησιμοποιήθηκαν στην παρούσα διατριβή στηρίζονται στην αρχιτεκτονική CUDA που έχει αναπτυχθεί από την NVIDIA. Συγκεκριμένα, χρησιμοποιήθηκαν τα μοντέλα καρτών γραφικών GeForce GTX 280, 285 και Tesla M2050. Οι Tesla M2050 (αποτελούν το πλέον σύγχρονο τύπο κάρτας γραφικών της παραπάνω εταιρίας τη στιγμή που γράφεται η διατριβή αυτή) που χρησιμοποιήθηκαν ανήκουν σε συστοιχία από κάρτες γραφικών της ΜΠΥΡ&Β/ΕΘΣ που αποτελείται από 4 διασυνδεδεμένους υπολογιστικούς κόμβους (blade servers). Κάθε υπολογιστικός κόμβος περιέχει 3 Tesla M2050. Συνολικά, η συστοιχία αποτελείται από 12 Tesla M2050, οι οποίες χρησιμοποιήθηκαν για την παράλληλη επίλυση μεγάλης κλίμακας αεροδυναμικών/αεροελαστικών προβλημάτων.

Στη διδακτορική διατριβή αναπτύχθηκε GPU-κώδικας υψηλής παράλληλης απόδοσης που επιλύει τις 2Δ χρονικά μόνιμες ή μη-μόνιμες εξισώσεις Navier-Stokes συμπιεστού ρευστού και τις 3Δ χρονικά μόνιμες ή μη-μόνιμες εξισώσεις Euler, σε δυναμικά μεταβαλλόμενα μη-δομημένα/υβριδικά πλέγματα. Η μοντελοποίηση της τύρβης, όπου χρειάστηκε, έγινε με το μοντέλο μίας εξίσωσης των Spalart-Allmaras. Η ολοκλήρωση των εξισώσεων της ροής γίνεται σύμφωνα με την τεχνική των πεπερασμένων όγκων με κεντροκομβική αποθήκευση των μεταβλητών της ροής. Ο κώδικας που αναπτύχθηκε αναπαράγει τα αριθμητικά αποτελέσματα υπάρχοντος CPU-επιλύτη των εξισώσεων Navier-Stokes, επαρκώς πιστοποιημένου στο πλαίσιο παλαιότερων διατριβών στη ΜΠΥΡ&Β/ΕΘΣ. Συνεπώς, τουλάχιστον όσον αφορά την αεροδυναμική ανάλυση, έμφαση δόθηκε αποκλειστικά στη μείωση του υπολογιστικού χρόνου.

Βασικό μειονέκτημα των σύγχρονων GPUs σε σχέση με τις σημερινές CPUs είναι η περιορισμένη ενδιάμεση (cache) μνήμη τους. Οι cache μνήμες είναι μνήμες ταχείας προσπέλασης και χρησιμοποιούνται για την προσωρινή αποθήκευση δεδομένων. Όσο μεγαλύτερο είναι το μέγεθος της συνολικής cache μνήμης που έχει μία μονάδα επεξεργασίας τόσο ταχύτερη γίνεται η εκτέλεση ενός κώδικα σ' αυτή. Επομένως, η παράλληλη απόδοση οποιουδήποτε GPU-κώδικα είναι άρρηκτα συνδεδεμένη με θέματα διαχείρισης της μνήμης της GPU. Σημειώνεται ότι η επιλογή της κεντροκομβικής διατύπωσης σε μη-δομημένα πλέγματα αποτελεί τη δυσκολότερη περίπτωση σε θέματα διαχείρισης της μνήμης της GPU, συγκριτικά με την εναλλακτική χρήση δομημένων πλεγμάτων ή της κεντροκυβελικής διατύπωσης σε μη-δομημένα πλέγματα. Αυτό οφείλεται στη μη-δομημένη αρίθμηση των κόμβων του πλέγματος και στο μεταβλητό πλήθος γειτόνων ανά κόμβο.

Ο πρώτος GPU-επιλύτης που προγραμματίστηκε προέκυψε απλά ξαναγράφοντας τον υπάρχοντα CPU-επιλύτη στη γλώσσα του περιβάλλοντος προγραμματισμού της CUDA (C++ με την προσθήκη ορισμένων ειδικών συναρτήσεων για τη διαχείριση της GPU). Παρόλο που ο εν λόγω GPU-επιλύτης μείωσε το χρόνο επίλυσης προβλημάτων αεροδυναμικής (σε σχέση με τον αντίστοιχο CPU-κώδικα), η παράλληλη απόδοσή του ήταν αρκετά χαμηλή. Για την αύξησή της έγιναν οι ακόλουθες ενέργειες:

- Προγραμματίστηκαν και αξιολογήθηκαν, ως προς τον χρόνο εκτέλεσης, διαφορετικές τεχνικές υπολογισμού των αριθμητικών διανυσμάτων της ροής.
- Έγινε σωστή/προσεκτική διαχείριση των διαθέσιμων μνημών της GPU (texture, constant, shared και global/device). Για παράδειγμα, ο προγραμματισμός στηρίχθηκε, κατά πολύ, στο ότι δεδομένα στα οποία έχουν πρόσβαση διεργασίες που εκτελούνται ταυτόχρονα πρέπει να καταλαμβάνουν κοντινές (ή ακόμα και διαδοχικές) θέσεις μνήμης.
- Το λογισμικό επίλυσης αναδομήθηκε πλήρως, με τρόπο ώστε οι σειριακές διεργασίες να εκτελούνται στη CPU ενώ, ταυτόχρονα, άλλες ανεξάρτητες διεργασίες να εκτελούνται παράλληλα στη GPU.
- Το πλέγμα χωρίστηκε σε υποχωρία και οι κόμβοι του κάθε υποχωρίου επαναριθμήθηκαν και έπειτα ταξινομήθηκαν με βάση το πλήθος των γειτονικών σε αυτούς κόμβους.
- Χρησιμοποιήθηκε αριθμητική μικτής ακρίβειας. Σύμφωνα με αυτή, τα μητρώα του αριστερού μέλους των διακριτοποιημένων εξισώσεων της ροής αποθηκεύονται σε μεταβλητές απλής ακρίβειας, μειώνοντας τον αριθμό προσβάσεων στην κεντρική μνήμη της GPU. Αυτό συμβάλλει στην αύξηση της παράλληλης απόδοσης του GPU-κώδικα χωρίς να αλλοιώνεται η ακρίβεια των αποτελεσμάτων, η οποία καθορίζεται αποκλειστικά από τον τρόπο υπολογισμού των υπολοίπων των εξισώσεων της ροής. Τα τελευταία υπολογίζονται με την επιβαλλόμενη ακρίβεια δεύτερης τάξης.

Οι ενέργειες που αναφέρθηκαν επέβαλαν, εκ των πραγμάτων, την πλήρη αναδόμηση του αρχικού GPU-κώδικα. Ο τελικός GPU-κώδικας έχει περίπου 20 φορές υψηλότερη παράλληλη απόδοση ως προς τον αρχικό. Για τον υπολογισμό της παράλληλης απόδοσης του τελικού GPU-κώδικα υπολογίστηκαν ροές γύρω από μεμονωμένες αεροτομές χρησιμοποιώντας πλέγματα διαφορετικού πλήθους κόμβων και διαφορετικές GPUs (GTX 280, 285 και Tesla M2050). Φάνηκε ότι η επιτάχυνση στην πρόλεξη της ροής που δίνει η χρήση μίας GPU αντί ενός πυρήνα μίας σημερινής CPU εξαρτάται από το μοντέλο της ροής (ατριβής, τυρβώδης), τη διάσταση του πλέγματος και, φυσικά, τον τύπο της κάρτας γραφικών. Συγκεκριμένα, οι εφαρμογές μεγάλης κλίμακας βοηθούνται περισσότερο από τη χρήση GPUs, καθώς οι προλέξεις στα μεγαλύτερα πλέγματα εμφανίζουν μεγαλύτερη παράλληλη επιτάχυνση. Για το λόγο αυτό οι εφαρμογές που παρουσιάζονται στη συνέχεια έχουν διαφορετικές παράλληλες επιταχύνσεις. Η μεγαλύτερη επιτάχυνση που καταγράφηκε στην επίλυση 2D τυρβώδους συμπιεστής ροής είναι 60x, 90x και 110x εκτελώντας τον GPU-κώδικα σε μία Tesla M2050 και χρησιμοποιώντας αριθμητική διπλής, μικτής και απλής ακρίβειας, αντίστοιχα. Οι ίδιοι υπολογισμοί αν πραγματοποιούνταν σε μία GTX 285 θα έδιναν επιτάχυνση μικρότερη κατά περίπου 30%. Η μέγιστη επιτάχυνση που καταγράφηκε στην επίλυση των 3D εξισώσεων Euler είναι 40x, 55x και 90x χρησιμοποιώντας μία Tesla M2050 και αριθμητική διπλής, μικτής και απλής ακρίβειας αντίστοιχα. Οι αντίστοιχες

επιταχύνσεις χρησιμοποιώντας μία GTX 285 είναι μικρότερες κατά περίπου 25%. Υπενθυμίζεται ότι η χρήση αριθμητικής μικτής ακρίβειας αναπαράγει τα αποτελέσματα της διπλής ακρίβειας συνιστώσας του GPU-κώδικα. Στον υπολογισμό των επιταχύνσεων που παρουσιάστηκαν προηγουμένως, ο CPU-κώδικας χρησιμοποιεί αριθμητική διπλής ακρίβειας και εκτελείται σε έναν πυρήνα μίας Intel Xeon CPU στα 2.00 GHz, με 4096 MByte cache μνήμη.

Επικουρικά, και για λόγους πληρότητας, στη διατριβή πραγματοποιήθηκε και η ανάπτυξη GPU-επιλύτων των εξισώσεων ατρίβους συμπίεστης ροής σε δομημένα πλέγματα ή κάνοντας χρήση της κεντροκυβελικής διατύπωσης της τεχνικής των πεπερασμένων όγκων σε μη-δομημένα πλέγματα. Επιπλέον, η διατριβή συνεισέφερε στην αποδοτική μεταφορά κωδικών επίλυσης των 2Δ εξισώσεων ασυμπίεστης ροής και του συζυγούς προβλήματος από τη CPU στη GPU.

Από την πλευρά της αεροδυναμικής ανάλυσης, ο GPU-επιλύτης χρησιμοποιήθηκε για τη μελέτη (α) της αλληλεπίδρασης του οριακού στρώματος με κύμα κρούσης (buffet) στην αεροτομή OAT15A και (β) της επίδρασης σύνθετης δέσμης (synthetic jet) στη μείωση της ζώνης αποκόλλησης κατάντι εμποδίου. Επιλύθηκε, επιπλέον, η ροή μη-συνεκτικού ρευστού γύρω από αεροσκάφος και σε πετύγωση υπερηχητικού συμπίεστη. Η επιτάχυνση στην πρόλεξη της ροής από τη χρήση μίας GPU σε αυτές τις εφαρμογές είναι μεταξύ 40x και 45x. Στις εφαρμογές αυτές χρησιμοποιήθηκαν οι τότε διαθέσιμες GeForce GTX 285. Αξιοσημείωτο είναι το γεγονός ότι οι συγκεκριμένες κάρτες γραφικών έχουν τοποθετηθεί σε παλιούς-ξεπερασμένους υπολογιστές, αναβαθμίζοντας τους.

Έχοντας αναπτύξει έναν αξιόπιστο πολύ γρήγορο επιλύτη (τη μικτής ακρίβειας εκδοχή του GPU-κώδικα), επόμενο βήμα ήταν η χρήση αυτού ως λογισμικού αξιολόγησης σε εξελκτικό αλγόριθμο (EA, λογισμικό EASY), όπου απαιτείται μεγάλο πλήθος αξιολογήσεων. Έτσι, πραγματοποιήθηκε η βελτιστοποίηση των παραμέτρων ενεργού ελέγχου της ροής (μέσω συνεχούς αναρρόφησης) σε αγωγή με διαμόρφωση. Ο EA συντόνιζε περισσότερες από μία GPU και ανέθετε την αξιολόγηση της εκάστοτε υποψήφιας λύσης σε μία κάρτα. Επομένως, το κέρδος από τη χρήση GPUs αντί CPUs υπερτίθεται σε εκείνο λόγω της ταυτόχρονης αξιολόγησης υποψήφιων λύσεων σε κάθε γενιά του EA.

Στοχεύοντας στην εκμετάλλευση της ταχύτητας εκτέλεσης της λιγότερο ακριβής και ακριβούς εκδοχής απλής ακρίβειας του GPU-κώδικα, προτάθηκε διεπίπεδο σχήμα βελτιστοποίησης καθοδηγούμενο από το λογισμικό EASY, όπου ο απλής ακρίβειας GPU-επιλύτης χρησιμοποιείται για μία γρήγορη πρώτη ανίχνευση του χώρου των υποψήφιων λύσεων στο χαμηλό επίπεδο. Ο αποδεκτής ακρίβειας και, ως εκ τούτου, πιο αργός (αλλά πολύ πιο γρήγορος σε σχέση με το αντίστοιχο CPU-λογισμικό) μικτής ακρίβειας GPU-κώδικας χρησιμοποιείται για την αξιολόγηση των υποψήφιων λύσεων του υψηλού επιπέδου, από όπου προκύπτει το μέτωπο Pareto των μη-κυριαρχούμενων λύσεων (στην περίπτωση πολυκριτηριακής βελτιστοποίησης) ή η βέλτιστη λύση (στην περίπτωση μονοκριτηριακής βελτιστοποίησης). Το χαμηλό και το υψηλό επίπεδο ανταλλάσσουν μικρό πλήθος από τις καλύτερες ανά επίπεδο υποψήφιες λύσεις, οι οποίες επαναξιολογούνται από το λογισμικό αξιολόγησης του επιπέδου. Οι αξιολογήσεις και των δύο επιπέδων γίνονται ταυτόχρονα σε όλες τις διαθέσιμες GPUs. Ο ρυθμός επικοινωνίας των EA ανά επίπεδο καθορίζεται από το χρήστη. Ο διεπίπεδος EA χρησιμοποιήθηκε για το σχεδιασμό μεμονωμένης αεροτομής και αεροτομής συμπίεστη με στόχους τη μεγιστοποίηση του συντελεστή άνωσης/ελαχιστοποίηση του συντελεστή οπισθέλκουσας και την ελαχιστοποίηση του συντελεστή απωλειών, αντίστοιχα. Έτσι, το κέρδος από τη χρήση του διεπίπεδου αλγορίθμου, σε σχέση με τη χρήση ενός κλασικού EA (ενός επιπέδου), υπερτίθεται σε εκείνο από τη χρήση των GPUs. Με τις εφαρμογές αυτές αναδεικνύεται ένας καλός τρόπος χρήσης του απλής ακρίβειας GPU-επιλύτη, αν και ο τελευταίος, από μόνος του, δεν είναι σήμερα αποδεκτός για χρήση σε προβλήματα υπολογιστικής ρευστοδυναμικής. Σημειώνεται ότι, για τα πλέγματα που

χρησιμοποιήθηκαν στις παραπάνω εφαρμογές, κάθε αξιολόγηση με το λογισμικό απλής ακρίβειας χρειάζεται περίπου 25% λιγότερο χρόνο σε σχέση με αντίστοιχη αξιολόγηση με το λογισμικό μικτής ακρίβειας.

Για την επίλυση αεροελαστικών προβλημάτων χρησιμοποιήθηκαν δυναμικά πλέγματα που ακολουθούν την κίνηση/παραμόρφωση της επιφάνειας του εξεταζόμενου αεροδυναμικού σώματος προσαρμοζόμενα στην εκάστοτε μορφή αυτής. Η χρήση της μεθόδου στρεπτικών ελατηρίων εξασφαλίζει τη διατήρηση της ποιότητας του πλέγματος μετά από την προσαρμογή αυτού. Αρχικά, πιστοποιήθηκε ο GPU-επιλύτης (όταν χρησιμοποιούνται δυναμικά πλέγματα) σε εφαρμογή εξωτερικής αεροδυναμικής όπου αεροτομή εκτελεί εξαναγκασμένη περιστροφική ταλάντωση γύρω από το 1/4 της χορδής της. Η σύγκριση με πειραματικά αποτελέσματα είναι ικανοποιητική. Στη συνέχεια, ο GPU-επιλύτης χρησιμοποιήθηκε για τη μελέτη της αεροελαστικής συμπεριφοράς αεροτομής, η κίνηση της οποίας περιορίζεται από ένα στρεπτικό και ένα γραμμικό ελατήριο. Υπολογίστηκε το όριο της περιοχής ευσταθούς λειτουργίας της αεροτομής (flutter boundary). Η σύγκριση με υπολογιστικά αποτελέσματα άλλων ερευνητικών ομάδων είναι πολύ καλή. Επιπλέον, πραγματοποιήθηκε η μελέτη της αεροελαστικής συμπεριφοράς αεροτομής με τρεις βαθμούς ελευθερίας. Η ακμή εκφυγής της αεροτομής θεωρήθηκε ως πτερύγιο ελέγχου, η κίνηση του οποίου περιορίζεται από ένα επιπλέον στρεπτικό ελατήριο. Επιπλέον, επιτεύχθηκε αύξηση της περιοχής ευσταθούς λειτουργίας με τον έλεγχο της γωνίας της ακμής εκφυγής (ως προς το κύριο σώμα της αεροτομής), μέσω ανάδρασης των μεταβλητών κατάστασης της αεροτομής. Επισημαίνεται ότι η εύρεση του ορίου της περιοχής ευσταθούς λειτουργίας μίας αεροτομής απαιτεί την επίλυση μεγάλου πλήθους μη-μόνιμων ροών, κάτι που θα ήταν εξαιρετικά δαπανηρό αν χρησιμοποιούνταν μόνο CPUs. Η επιτάχυνση που δίνει η χρήση των Tesla M2050 είναι 50x.

Για την επίλυση μεγάλης κλίμακας προβλημάτων αεροδυναμικής/αεροελαστικότητας, μία GPU δεν επαρκούσε εξαιτίας της περιορισμένης έκτασης της κεντρικής της μνήμης (1 GByte ή 3 GByte, αντίστοιχα, για τις GeForce GTX 285 και Tesla M2050). Για το λόγο αυτό, ο GPU-κώδικας παραλληλοποιήθηκε σε πολλές GPUs του ίδιου, ή διαφορετικών υπολογιστικών κόμβων. Η διαχείριση των GPUs του ίδιου υπολογιστικού κόμβου ανατίθεται στο ίδιο CPU-thread. Για την επικοινωνία μεταξύ δύο ή περισσότερων GPUs διαφορετικών υπολογιστικών κόμβων χρησιμοποιήθηκε το πρωτόκολλο παράλληλης επικοινωνίας MPI. Ο παράλληλος σε πολλές κάρτες γραφικών GPU-επιλύτης χρησιμοποιήθηκε για την επιτυχή υλοποίηση των υπολογισμών ερευνητικού προγράμματος χρηματοδοτούμενου από την Ευρωπαϊκή Ένωση, σχετικού με το σχεδιασμό ενός νέου αεροσκάφους τύπου Blended-Wing-Body (BWB). Τα αεροσκάφη αυτού του τύπου χαρακτηρίζονται από πολλά πλεονεκτήματα από αεροδυναμικής πλευράς, όπως μεγαλύτερος λόγος άνωσης προς οπισθέλκουσα, μικρότερη κατανάλωση καυσίμου ανά επιβάτη και μίλι και μικρότερη εκπομπή θορύβου σε σχέση με συμβατικά αεροσκάφη ίδιων προδιαγραφών. Χρησιμοποιήθηκε ο GPU-επιλύτης της διατριβής για τη μελέτη της αεροδυναμικής συμπεριφοράς του αεροσκάφους όταν προεπιλεγμένη επιφάνεια ελέγχου ή ολόκληρη η κατασκευή εξαναγκάζεται σε περιοδική κίνηση/παραμόρφωση. Αν και η χρήση GPUs δεν ήταν επιβεβλημένη, επιλέχθηκε να χρησιμοποιηθεί ο GPU-κώδικας για τη μείωση του χρόνου επίλυσης των ζητούμενων αεροδυναμικών/αεροελαστικών προβλημάτων. Συγκεκριμένα, τα προβλήματα ροής που μελετήθηκαν επιλύονται 6 φορές πιο γρήγορα σε 3 Tesla M2050 του ίδιου υπολογιστικού κόμβου αντί σε 8 οκταπύρηνες CPUs (σύνολο 64 CPU-πυρήνες).

Συνοψίζοντας, στις εφαρμογές που αναφέρθηκαν, η επιτάχυνση από τη χρήση των GPUs κυμαίνεται από 40x έως 110x, ανάλογα με το χρησιμοποιούμενο μοντέλο ροής, τη διάσταση του πλέγματος, την κάρτα γραφικών και την ακρίβεια της αριθμητικής που χρησιμοποιείται (διπλής, μικτής ή απλής). Η επιτάχυνση αυτή είναι ιδιαίτερα σημαντική αφενός στην αεροδυναμική ή αεροελαστική ανάλυση σωμάτων αφετέρου δε (και κυρίως) στη βελτιστοποίηση-σχεδιασμό όπου ο απαιτούμενος αριθμός κλήσεων του λογισμικού

αξιολόγησης αυξάνει σημαντικά. Με τον τρόπο αυτό η χρήση GPUs, αντί CPUs, διευρύνει το εύρος των βιομηχανικών εφαρμογών μεγάλης κλίμακας που μπορούν να γίνουν σε αποδεκτό χρόνο και μειώνει σημαντικά το χρόνο αναμονής του μηχανικού. Από τη χρήση των GPUs ευνοούνται περισσότερο προβλήματα μεγαλύτερης κλίμακας, καθώς καταγράφουν μεγαλύτερες επιταχύνσεις. Κλείνοντας, αξ σημειωθεί ότι η γνώση που αποκτήθηκε μπορεί να χρησιμοποιηθεί στην επίλυση προβλημάτων άλλων επιστημονικών περιοχών που διέπονται από μ.δ.ε. διατυπωμένες σε μη-δομημένα πλέγματα.