ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΚΡΗΤΗΣ - ΠΑΡΑΡΤΗΜΑ ΧΑΝΙΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΦΥΣΙΚΩΝ ΠΟΡΩΝ & ΠΕΡΙΒΑΛΛΟΝΤΟΣ
ΕΡΕΥΝΗΤΙΚΟ ΚΕΝΤΡΟ ΦΥΣΙΚΗΣ ΠΛΑΣΜΑΤΟΣ & LASERS - CPPL

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

«Βελτιστοποίηση Συστήματος Συστοιχίας Υπολογιστών σε Hardware και Software, σε Περιβάλλον LINUX, για την Παράλληλη Επεξεργασία Προβλημάτων Προσομοίωσης Υψηλών Απαιτήσεων»

του Μπαρούτσου Ανδρέα

Επιβλέπων καθηγητής:
Δρ. Βασίλειος Δημητρίου, Επίκουρος Καθηγητής

Χανια 2016

Περίληψη

Στην παρούσα πτυχιακή εργασία αναλύεται η δομή και λειτουργία των συστοιχιών ηλεκτρονικών υπολογιστών, των cluster, για παράλληλη επεξεργασία προσομοιώσεων υψηλών υπολογιστικών απαιτήσεων. Παρουσιάζεται το hardware και το software που τα επιλέγονται να εγκατασταθούν και να δώσουν ζωή σε 2 cluster. Αναλύεται η εγκατάσταση και οι ρυθμίσεις των παραμέτρων του λειτουργικού συστήματος και του απαραίτητου software προσομοίωσης και παράλληλης επεξεργασίας - MPI. Το λειτουργικό σύστημα Linux και το OpenMPI ακολουθεί η εγκατάσταση των ANSYS και Ls-Dyna καθώς και των EPOCH αλλά και Pluto και Χοο-pic, λογισμικών Πεπερασμένων Στοιχείων, PIC αλλά και μαγνητο-υδροδυναμικής - MHD, αντίστοιχα.

Παρουσιάζεται η Μέθοδος των Πεπερασμένων Στοιχείων - FEM και ακολουθούν αντιπροσωπευτικές προσομοιώσεις στα 2 cluster που συντέθηκαν. Η απόδοση της παράλληλης επίλυσης αναλύεται στα ANSYS και Ls-Dyna για προβλήματα FEM και αντίστοιχα στο Pluto για MHD. Η καταγραφή και ανάλυση των προσομοιώσεων αυτών οδηγεί σε σημαντικά συμπεράσματα τα οποία με τη σειρά τους οδηγούν στην βελτιστοποίηση των παραμέτρων software και hardware των 2 cluster άρα και στη συνολική απόδοση της παράλληλης επεξεργασίας.

Abstruct

This thesis is focused on the study of the structure, the operation and administration of computer clusters dedicated to parallel processing of high computational demanding simulations. The hardware and software chosen to be installed, is presented and analyzed for each of the 2 clusters that are built. Each installation and the parameters of the operating systems and of all of the necessary simulation and parallel processing – MPI software are described.

After the LINUX operating system and OpenMPI software stabilization, follows the installation of ANSYS, Ls-Dyna, and also EPOCH, PLUTO and Xoo-pic, programs of Finite Element Method - FEM, PIC and magneto-hydrodynamic – MHD simulations. Parallel processing of FEM is described and representative simulations in the 2 clusters are following. The recording and analysis of the performance of the parallel processing in ANSYS, Ls-Dyna for FEM and in PLUTO in MHD problems, leads to important conclusions for the efficiency of the 2 cluster systems. These conclusions are used for the optimization of the parallel simulation process.

Ευχαριστίες

Η παρούσα πτυχιακή εργασία είναι η επισφράγιση των σπουδών μου στο ΤΕΙ Κρήτης και συγκεκριμένα στο τμήμα Μηχανικών Φυσικών Πόρων και Περιβάλλοντος στα Χανιά. Θα ήθελα να ευχαριστήσω την σύζυγό μου, Χριστίνα, για την τεράστια υπομονή που επέδειξε τόσο χρόνια, για τον χρόνο που ξόδεψε ώστε να με βοηθάει να διορθώνω τα άπειρα ορθογραφικά και συντακτικά λάθη μου και για τις ατελείωτες ώρες που έλειπα από το σπίτι ή ήμουν απορροφημένος στον ηλεκτρονικό υπολογιστή, για να εκπονήσω αυτήν την πτυχιακή εργασία. Επίσης, θέλω να ευχαριστήσω τον επιβλέποντα καθηγητή μου Δρ. Βασίλειο Δημητρίου, για την στήριξη, την καθοδήγηση που μου έδωσε και για τον κόπο που κατέβαλε, καθ' όλη την διάρκεια της εκπόνησης. Δεν θα πρέπει να ξεχάσω να ευχαριστήσω τον Προϊστάμενό μου Υποπλοίαρχο (Ε) Α. Κουνδουράκη για την αμέριστη συμπαράσταση και διευκόλυνση ως προς τις σπουδές μου και την σύνταξη της παρούσας πτυχιακής εργασίας, στο προσωπικό του εργαστηρίου CPPL, τον Αλέξανδρο Σκουλάκη και τον Καθηγητή Δρ. Μιχάλη Ταταράκη, για την υπομονή και την βοήθεια σε όποια προβλήματα προέκυψαν.

Πίνακας περιεχομένων

ΚΕΦΑΛΑΙΟ 1. Cluster και παράλληλη επεξεργασία

1.1. Εισαγωγή

Η παρούσα πτυχιακή εργασία, έχει ως σκοπό την ανάλυση και βελτιστοποίηση της παράλληλης επεξεργασίας υπολογιστικών προβλημάτων προσομοίωσης μέσω συστήματος συστοιχίας ηλεκτρονικών υπολογιστών (computer cluster).

Η δημιουργία υπολογιστικού συστήματος, το οποίο θα έχει την δυνατότητα να εκτελεί παράλληλη επεξεργασία δεδομένων, απαιτεί λεπτομερή σχεδιασμό, ρεαλιστική ανάλυση των αναγκών, προσεκτική επιλογή του υλικού (hardware), του λογισμικού (software) και του λειτουργικού συστήματος (operating system - OS) που θα χρησιμοποιήσουμε. Την ίδια στιγμή, πρέπει να αποφασιστεί ποιες παραχωρήσεις θα γίνουν, για να είναι το υπολογιστικό μας σύστημα περισσότερο ή λιγότερο ασφαλές, ευέλικτο και φιλικό στο χρήστη. Όλα αυτά, αναμφίβολα, συνθέτουν ένα πρόβλημα, που στην παρούσα περίπτωση, άλλοτε ήταν ιδιαίτερα απλή η αντιμετώπισή του και άλλοτε ήταν αναγκαία η λήψη αποφάσεων που επηρέασαν την λειτουργικότητά του.

Η χρήση computer cluster, ή απλούστερα cluster, είναι πλέον κοινή πρακτική για εκπαιδευτικά ιδρύματα, κρατικές υπηρεσίες ανάπτυξης τεχνολογίας και μεγάλες εταιρίες που στοχεύουν στην έρευνα και την αξιοποίηση της τεχνολογίας. Σκοπός τους είναι η αποδοτικότερη, οικονομικά και χρονικά, εφαρμογή της επιστημονικής έρευνας στον πραγματικό κόσμο. Αξιοποιώντας την συνδυασμένη επεξεργαστική ισχύ ενός cluster, καθιστούμε αυτόματα την ανάπτυξη νέων συστημάτων και νέων τεχνολογιών μία διαδικασία που κρύβει λιγότερες εκπλήξεις, εκμεταλλευόμαστε τα υλικά που έχουμε στο έπακρο, ενώ μειώνουμε τους πόρους, πρώτες ύλες και ενέργεια, που θα σπαταλούσαμε αν χρησιμοποιούσαμε τις παραδοσιακές μεθόδους προσομοίωσης μέσω πραγματικών μοντέλων. Πλέον, τα δυνατότερα και αποδοτικότερα υπολογιστικά συστήματα είναι κατασκευασμένα με την μέθοδο της συστοιχίας ηλεκτρονικών υπολογιστών. Χαρακτηριστικό παράδειγμα, στην Εικόνα 1, ο *Tianhe-2* [1] στο National Supercomputing Center στην Guangzhou στην Κίνα, που αποτελείται

από 16.000 υπολογιστές (computer nodes) και είναι το ταχύτερο υπολογιστικό σύστημα στον πλανήτη, με ταχύτητα 33,86 PFLOPS .

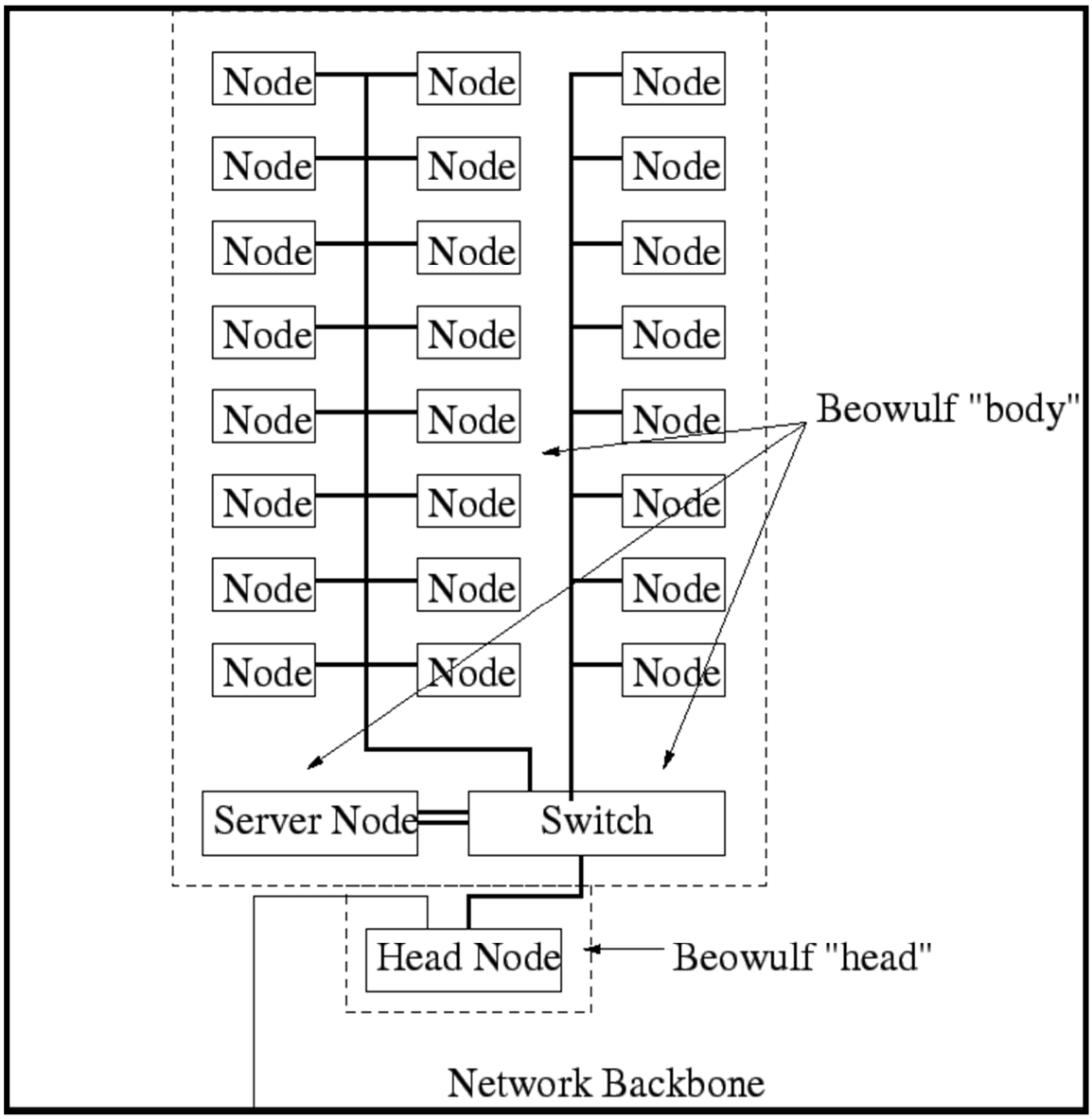
Εικόνα 1. Ο ταχύτερος υπολογιστής στον κόσμο Tianhe-2.

1.2. Συστοιχίες ηλ/κών υπολογιστών – cluster

Cluster, είναι ο τύπος του υπολογιστικού συστήματος, το οποίο θα χρησιμοποιήσουμε για την πραγματοποίηση της παράλληλης επεξεργασία δεδομένων.

Στα Αγγλικά, cluster σημαίνει σύνολο ομοειδών αντικειμένων που βρίσκονται κοντά τοποθετημένα μεταξύ τους. Στην πληροφορική, cluster είναι μία ομάδα μη στενά συζευγμένων (αυτόνομων) ηλεκτρονικών υπολογιστών, οι οποίοι όμως εργάζονται όλοι μαζί σαν σύνολο και μία οντότητα. Η έννοια αυτής της μορφής υπολογιστικού συστήματος όμως, φαίνεται καλύτερα από αυτό που μπορούμε να αποκομίσουμε. Απλά, με ένα σύνολο ηλεκτρονικών υπολογιστών χαμηλών ή μέσων δυνατοτήτων, μπορούμε, χρησιμοποιώντας τεχνικές παράλληλης επεξεργασίας δεδομένων, να παράγουμε έργο και να έχουμε επεξεργαστική ισχύ που θα παρήγαγε ένας ηλεκτρονικός υπολογιστής πολλαπλάσιας αξίας. Τον υπολογιστή αυτό, φυσικά, θα μας ήταν πολύ πιο δύσκολο να διαχειριστούμε και να συντηρήσουμε λόγω της μοναδικότητάς του, γιατί συνήθως κατασκευάζεται κατά παραγγελία. Επιπροσθέτως, απαιτεί λειτουργικό σύστημα και software, το οποίο και αυτό έχει δημιουργηθεί ειδικά για αυτό το μηχάνημα, ώστε να καλύπτει τις ξεχωριστές απαιτήσεις του. Είναι προφανής η οικονομία στο κόστος του hardware, αλλά και σε χρόνο για την εκπαίδευση προσωπικού για τη διαχείριση και τη συντήρηση του. Όλο αυτό το

εγχείρημα, της δημιουργίας cluster και της χρήσης του, μας θυμίζει ένα ρητό που το περιγράφει με μεγάλη ακρίβεια : «Ισχύς εν τη ενώσει». Στην Εικόνα 2 παρουσιάζεται τυπική σύνδεση κόμβων σε αντιπροσωπευτικό Beowulf cluster [2].



Εικόνα 2. Σύνδεση κόμβων σε αντιπροσωπευτικό Beowulf cluster.

Ένα cluster αποτελείται από **κόμβους** ή **node**, όπως είναι η αγγλική ορολογία. Κόμβος είναι ο κάθε ξεχωριστός ηλεκτρονικός υπολογιστής που συνιστά το cluster. Όμοια κάθε κόμβος αποτελείται από 1 ή περισσότερους **πυρήνες** ή **cores**, όπου πυρήνας είναι η κάθε ανεξάρτητη μονάδα επεξεργασίας, που μπορεί να τρέχει εντολές αυτόνομα.

Έτσι, η ανάγκη για αύξηση της επεξεργαστικής ισχύος, συγκρατώντας το κόστος όσο το δυνατόν χαμηλότερα και παράλληλα διατηρώντας την αξιοπιστία

του συστήματος, ενώ την ίδια στιγμή το υλικό να είναι εμπορικού τύπου COTS (commercial off-the-self), οδήγησε στην εύρεση μεθόδων αξιοποίησης των υπαρχόντων υπολογιστικών συστημάτων. Η μέθοδος με την οποία γίνεται η υλοποίηση του clustering, είναι σχετικά απλή. Ένας αριθμός προσωπικών υπολογιστών διασυνδεδεμένοι μέσω γρήγορου τοπικού δικτύου, με όμοιο λειτουργικό σύστημα και μία εφαρμογή η οποία αναλαμβάνει την ενορχήστρωση όλων των nodes, ώστε να λειτουργούν σαν μία οντότητα, όπου node ονομάζουμε την κάθε ξεχωριστή οντότητα - υπολογιστικό σύστημα που συμμετέχει στο cluster. Αν και πλέον έχουν αναπτυχθεί διάφορες τεχνικές και κατηγορίες clustering που διαφέρουν μεταξύ τους, η βασική ιδέα δεν έχει αλλάξει. Το τελικό συμπέρασμα, είναι ότι η διαχείριση είναι κεντρική και το αποτέλεσμα που θέλουμε προέρχεται από την συνδυασμένη επεξεργαστική ισχύ των nodes που το αποτελούν. Ένα παράδειγμα Cluster από COTS [3] ηλεκτρονικούς υπολογιστές παρουσιάζεται στην Εικόνα 3.

Εικόνα 3. παράδειγμα Cluster από COTS.

1.3. Η εξέλιξη των cluster

Ιστορικά, σύμφωνα με τον Gregory Pfister, στο βιβλίο του «In search of Clusters» [4], τα Cluster εφευρέθηκαν από τους χρήστες των υπολογιστών στην δεκαετία του 1950 -1960, γιατί δεν μπορούσαν να ικανοποιηθούν οι απαιτήσεις τους με τη χρήση ενός μόνο υπολογιστή. Η πρώτη επίσημη αναφορά σε παράλληλη επεξεργασία έγινε από τον Gene Amdahl της IBM το 1967. Αργότερα, με την ανάπτυξη των δικτύων υπολογιστών, η εμφάνιση του ARPANET έκανε πραγματικότητα το πρώτο, μη επίσημο, cluster που διασυνέδεε 4 υπολογιστικά κέντρα. Αυτό αποτέλεσε το παράδειγμα για τον τρόπο διασύνδεσης και επικοινωνίας των μετέπειτα computer cluster. Το πρώτο εμπορικά διαθέσιμο cluster, ήταν το ARCnet της Datapoint το 1977, που εξελίχθηκε σε εμπορική αποτυχία. Αντίθετα το VAXcluster της DEC που εμφανίστηκε το 1984 με λειτουργικό σύστημα VAX/VMS, εξακολουθεί να υπάρχει σαν VMScluster από την IBM μέσω του εγχειρήματος openVMS. Τα δύο προηγούμενα συστήματα έρχονταν με κοινόχρηστο σύστημα αρχείων και περιφερειακά. Όμως, δεν θα είχαμε φτάσει σε τέτοιο επίπεδο εξέλιξης, αν δεν είχε εφευρεθεί το κατάλληλο software, που αποτέλεσε το κλειδί για την αλματώδη ανάπτυξη των cluster, αλλά και την κατακρήμνιση της σχέσης του λόγου κόστους-ισχύος, που έφερε η δυνατότητα χρήσης απλών προσωπικών υπολογιστών στην συστοιχία ενός cluster. Αυτό ήταν το PVM (Parallel Virtual Mashine) το 1989, που βασιζόμενο στο πρωτόκολλο TCP/IP, επέτρεψε την άμεση δημιουργία υπερυπολογιστών, αρκεί να ήταν συνδεδεμένα σε ένα δίκτυο TCP/IP. Η εξέλιξη λογισμικού τέτοιου τύπου, μας επέτρεψε να εμφανιστεί το 1994 ένας νέος τύπος cluster, το Beowulf cluster. Το Beowulf cluster (βλ. Εικόνες 2. και 3), πρωτοπαρουσιάστηκε από τους Thomas Sterling και Donald Becker στη NASA[3]. Αποτελεί ένα σφιχτά συνδεδεμένο σύνολο ηλεκτρονικών υπολογιστών, με κοινή διαχείριση μέρους του συστήματος αρχείων, στο οποίο έχει εγκατασταθεί software PVM και MPI (Message Passing Interfase)[5]. Είναι η πλέον σκληροπυρηνική μορφή cluster, όπου το harware και το software είναι πανομοιότυπα σε όλα τα nodes, ενώ τα τμήματα του λειτουργικού συστήματος και του συστήματος αρχείων μοιράζονται εξίσου σε όλο το cluster. Έτσι, μπορεί να συμπεριφέρεται σαν μία οντότητα, σαν ένα υπολογιστικό σύστημα και όχι σαν ένα σύνολο υπολογιστών που συνεργάζονται στενά. Άρα, είναι

προφανές γιατί είναι η πιο συνηθισμένη κατηγορία που χρησιμοποιείται για παράλληλη επεξεργασία προσομοιώσεων υψηλών απαιτήσεων.

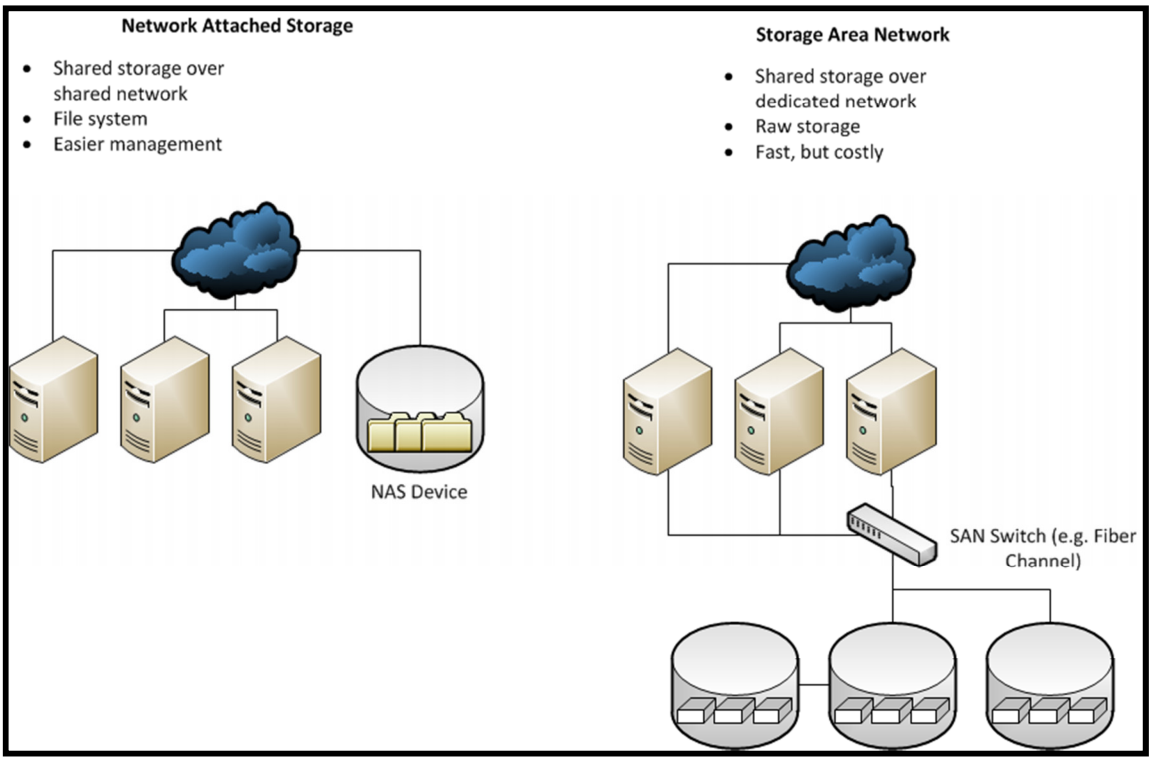
1.4. Το hardware των cluster

Στα σύγχρονα cluster, η επιλογή του hardware γίνεται βάση της οικονομικής δυνατότητας του εκάστοτε ενδιαφερόμενου. Μπορούμε να χρησιμοποιήσουμε παλιούς ηλεκτρονικούς υπολογιστές χαμηλών δυνατοτήτων, είτε υπέρ-υψηλής απόδοσης HPC (High Performance Computer) nodes, με επικοινωνία μέσω δικτύου οπτικών ινών και κοινόχρηστο σύστημα αρχείων, το οποίο υλοποιείται μέσω ειδικών συνδέσεων μεγάλου εύρους ζώνης. Πλέον, είναι εξαιρετικά σπάνιο φαινόμενο, το να μην είναι x86-64 η αρχιτεκτονική των nodes που το αποτελούν, δηλαδή να μην υποστηρίζουν σετ εντολών 64bit. Αυτό, οφείλεται στο γεγονός ότι η αρχιτεκτονική αυτή μπορεί να υποστηρίξει μέχρι 64TB φυσικής μνήμης (RAM) σε λειτουργικό σύστημα LINUX, ενώ διπλασιάζεται το μήκος των καταχωρητών. Το αποτέλεσμα είναι η αύξηση της ταχύτητας επεξεργασίας, κυρίως σε εφαρμογές οι οποίες απαιτούν υψηλή επεξεργαστική ισχύ, όπως είναι και οι εφαρμογές προσομοιώσεων. Με βάση την σημαντικότητα της συνεισφοράς τους σε ένα cluster, περιγράφονται τα βασικότερα δομικά του στοιχεία: i. Συστήματα αποθήκευσης, ii. Δικτυακή διασύνδεση, iii. Επεξεργαστές και iv. Μνήμη RAM.

1.4.1. Συστήματα αποθήκευσης και σκληροί δίσκοι

Εκτός όμως από την αρχιτεκτονική, υπάρχουν και άλλα σημεία που δίνεται έμφαση στο υλικό. Αν εξαιρέσουμε τα σύστημα που είναι κατασκευασμένα για εκπαιδευτικούς λόγους ή από λάτρεις των ηλεκτρονικών υπολογιστών, θα δούμε ότι έχει κυριαρχήσει η χρήση συστημάτων NAS (Network-Attached Storage) και SAN (Storage Area Network). Και τα δύο μας παρέχουν την δυνατότητα να έχουμε κοινόχρηστο σύστημα αρχείων στα node, αλλά με διαφορετική προσέγγιση, ενώ παράλληλα μας δίνουν την δυνατότητα να χρησιμοποιήσουμε τεχνικές διασφάλισης των δεδομένων μας όπως είναι το RAID (Retundant Array of Independent Disks). Στην Εικόνα 4, μπορούμε να δούμε σχηματικά το βασικό τρόπο με τον οποίο χειρίζονται τον αποθηκευτικό χώρο ένα σύστημα NAS και ένα σύστημα SAN [6]. Τα συστήματα NAS είναι πιο ευέλικτα, όντας σχεδιασμένα να

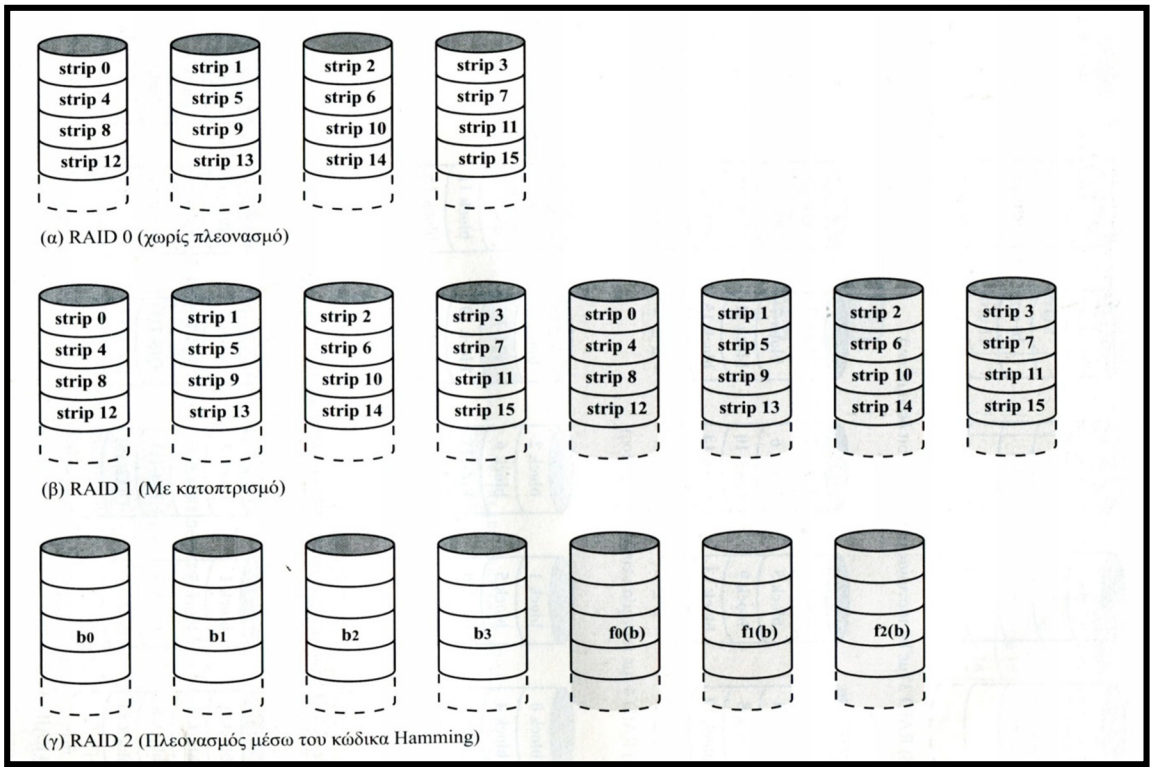
εξυπηρετούν ετερογενείς πελάτες. Αυτό γίνεται επειδή μπορεί να γίνει χρήση διαφόρων πρωτοκόλλων διαμοίρασης αρχείων, χωρίς να περιορίζεται μόνο σε ένα κάθε φορά, γιατί όλοι οι εξυπηρετούμενοι το αναγνωρίζουν σαν εξυπηρετητή αρχείων (File Server). Μπορεί σε οικονομικά συστήματα να μην είναι εξειδικευμένη συσκευή, αλλά ένας από τους υπολογιστές στον οποίο έχουν εγκατασταθεί οι κατάλληλες υπηρεσίες και πρωτόκολλα, για παράδειγμα NFS (Network File System) για συστήματα LINUX. Από την άλλη πλευρά, τα συστήματα SAN είναι εξειδικευμένα υψηλής απόδοσης, με την ανάλογη επίπτωση στο κόστος τους. Δεν λειτουργούν με τη λογική του File Server, αλλά παρουσιάζουν τον αποθηκευτικό χώρο που προσφέρουν σαν ένα σκληρό δίσκο, στον οποίο έχει ευθύνη ο εξυπηρετούμενος για την επιλογή του συστήματος αρχείων (File System)[7]. Το γεγονός αυτό, απαιτεί χρήση ειδικών πρωτοκόλλων, ώστε να μην καταστρέφονται τα δεδομένα κάθε φορά που αποκτά πρόσβαση κάποιος εξυπηρετούμενος. Αυτά είναι στην συντριπτική πλειοψηφία των συστημάτων τα Fibre-Channel και iSCSI. Το Fibre-Channel, είναι ένα ολοκληρωμένο σύστημα το οποίο περιλαμβάνει δίκτυο οπτικών ινών, ειδικά σχεδιασμένο για αποκλειστική χρήση από τη συσκευή που παρέχει το SAN και τους εξυπηρετούμενους από αυτό. Στο Fibre-Channel, χρησιμοποιείται το Fibre-Channel Protocol, το οποίο μεταφέρει εντολές SCSI στο δίκτυό του. Μπορεί να διασφαλίσει την ταχύτατη μεταφορά δεδομένων, την μέγιστη εκμετάλλευση των πόρων που διατίθενται και την απρόσκοπτη συνεχή μεταφορά δεδομένων, ακόμα και αν χαθεί πάνω από το μισό μέρος του συστήματος, λόγω βλάβης ή αστοχίας υλικού. Εναλλακτικά, μπορεί να λειτουργήσει και σε δίκτυο συστραμμένων καλωδίων. Εφευρέθηκε το 1988, ώστε να αντικαταστήσει το δύσχρηστο σύστημα HIPPI και πήρε έγκριση από την διεθνή ένωση ANSI το 1994. Είχε ως ανταγωνιστή το SSA (Serial Storage Architecture) της IBM, του οποίου όμως τελικά επικράτησε.



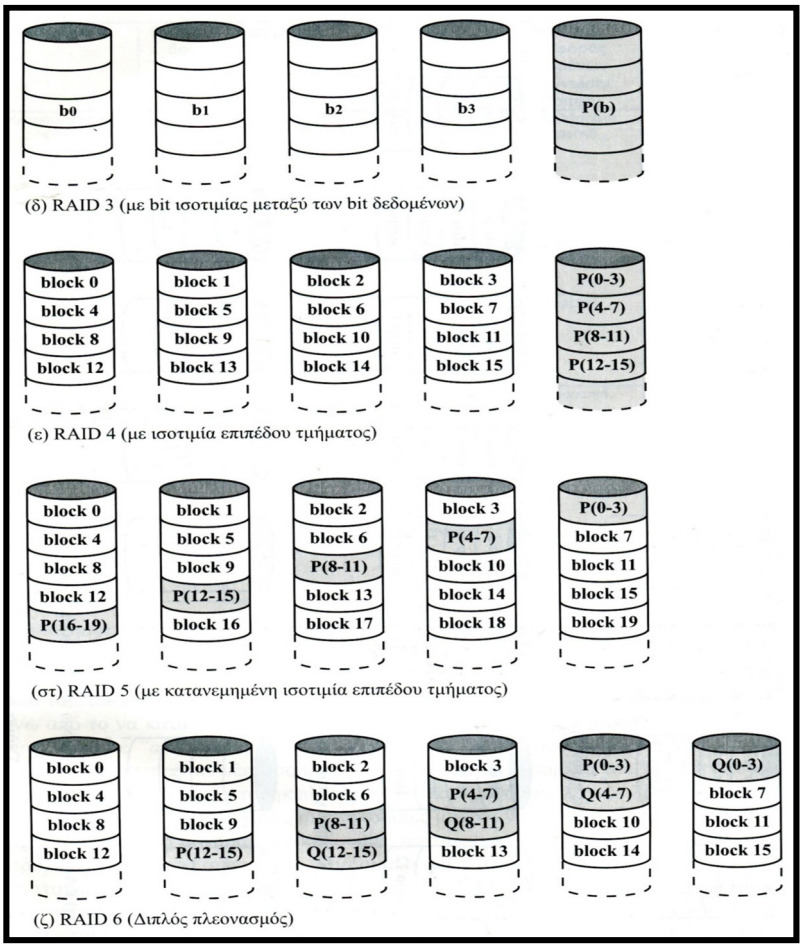
Εικόνα 4. Παραδείγματα NAS και SAN.

Το Fibre-Channel, είναι πλέον το πλέον συνηθισμένο υψηλής ταχύτητας δίκτυο για SAN, προσφέροντας ταχύτητες μεταγωγής δεδομένων μέχρι 6400Mbps. Αντίθετα, το iSCSI (internet Small Computer System Interface), δεν απαιτεί εξειδικευμένο δίκτυο για την λειτουργία του, αλλά μεταφέρει εντολές SCSI σε οποιοδήποτε δίκτυο TCP/IP. Έτσι, μπορεί να λειτουργήσει σε συστήματα SAN που οι υπολογιστές και οι συστοιχίες δίσκων βρίσκονται σε απομακρυσμένα σημεία, χωρίς τη δυνατότητα απευθείας διασύνδεσης.

Στις Εικόνες 5 και 6 παρουσιάζονται τα επίπεδα του RAID σύμφωνα με το βιβλίο Οργάνωση & Αρχιτεκτονική Υπολογιστών του William Stallings [8].

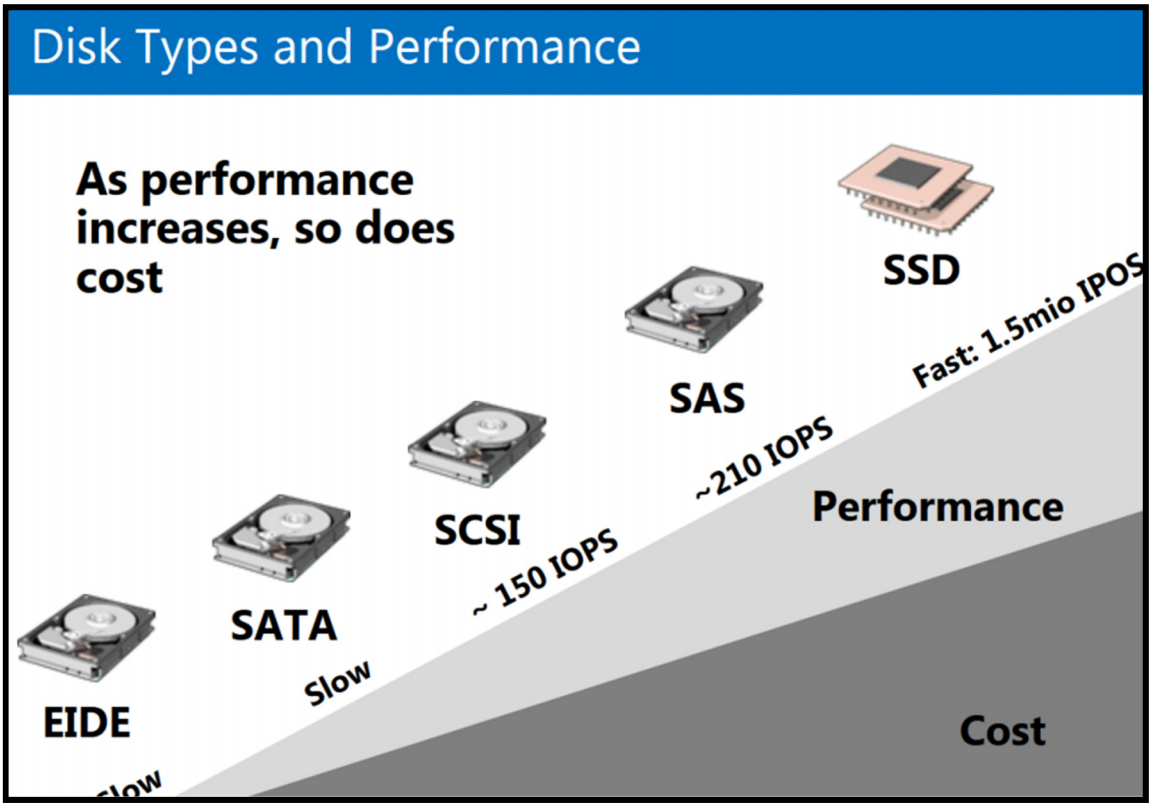


Εικόνα 5. Τα επίπεδα του RAID - a.



Εικόνα 6. Τα επίπεδα του RAID - b.

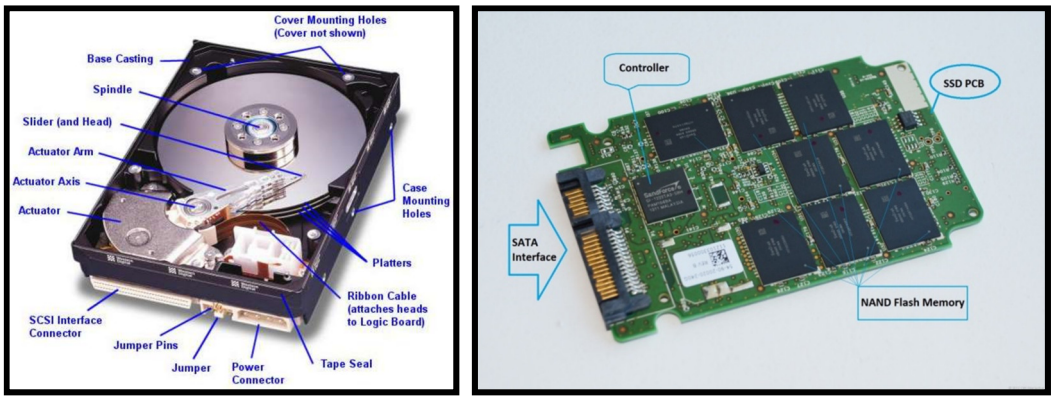
Ένα άλλο σημείο που χαρακτηρίζει ένα cluster, είναι η τεχνολογία των σκληρών δίσκων που χρησιμοποιεί στα nodes και στις συστοιχίες σκληρών δίσκων του. Είναι δεδομένο ότι η τεχνολογία IDE είναι πλέον παρωχημένη και δεν μπορεί πλέον να θεωρείται σαν επιλογή, ακόμα και για τα παλαιότερα συστήματα. Αντίθετα, έχουμε μία ποικιλία ανάλογα με τις κατά περίπτωση ανάγκες. Δίσκοι 3,5” SATA των 7200 rpm και 10000 rpm ακόμα και δίσκοι SAS (Serial Attached SCSI) των 15000 rpm, οι οποίοι προσφέρουν μεγάλες επιδόσεις και αποτελούν ουσιαστικά τη σειριακή εξέλιξη της τεχνολογίας SCSI. Ταυτόχρονα, συχνή είναι η εμφάνιση σκληρών δίσκων 2,5”, είτε 5400 rpm είτε 7200 rpm, σε nodes που ανήκουν σε ολοκληρωμένα συστήματα cluster, λόγω του μικρού μεγέθους τους, αλλά και της αυξημένης αντοχής τους, αφού σχεδιάστηκαν να αντέχουν τις υψηλές θερμοκρασίες και τις καταπονήσεις που δέχεται ένας φορητός υπολογιστής.



Εικόνα 7. Σχέση κόστους τεχνολογίας σκληρών δίσκων με την απόδοση.

Επίσης, μία τεχνολογία που υπάρχει αρκετά χρόνια αλλά πρόσφατα έγινε προσιτή στο κοινό, είναι οι σκληροί δίσκοι SSD (Solid State Drive). Για πάρα πολλά χρόνια, η χρήση τους ήταν μονόδρομος σε αεροναυτικά, διαστημικά και

ειδικά βιομηχανικά προγράμματα, λόγω της απαράμιλλης αντοχής τους σε μηχανικές καταπονήσεις. Χωρίς αυτούς δεν θα μπορούσαμε να βλέπουμε σύγχρονα πολεμικά αεροπλάνα να πετούν, ούτε να έχουμε αποστολές στον πλανήτη Άρη από ειδικά Robot. Τα χαρακτηριστικά τους αυτά, τα οφείλουν στο γεγονός ότι δεν περιέχουν κινητά μέρη, δηλαδή δεν αποτελούνται από περιστρεφόμενες πλάκες επιστρωμένες από σιδηρομαγνητικό υλικό, ούτε έχουν κινητές κεφαλές, αλλά μνήμες τύπου NAND. Με απλά λόγια, χωρίς να είμαστε απόλυτα σωστοί, μπορούμε να απλουστεύσουμε την περιγραφή μας, λέγοντας ότι λειτουργούν όπως τα φορητά USB stick που έχουμε οι περισσότεροι από εμάς. Όμως, οι δίσκοι SSD έδωσαν στο ευρύ κοινό κάτι που ήταν ένα άπιαστο όνειρο: εξαιρετικά υψηλές ταχύτητες μεταγωγής δεδομένων. Παρόλο που είναι μικρότερη η χωρητικότητά τους και πολύ μεγαλύτερη η σχέση κόστους ανά GB, η αγορά τους και η χρήση τους διευρύνεται συνεχώς. Η πληθώρα επιλογών στα μέσα μαζικής αποθήκευσης, μας δίνει και την δυνατότητα να σχεδιάσουμε καλύτερα το σύστημα που ικανοποιεί τις ανάγκες μας, αλλά και να διατηρήσουμε το κόστος στα πλαίσια των δυνατοτήτων μας. Αν μας ενδιαφέρει το κοινόχρηστο σύστημα αρχείων να έχει πολύ μεγάλη χωρητικότητα, μπορούμε, απλά και οικονομικά, να χρησιμοποιήσουμε δίσκους SATA 7200 rpm, που δίνουν την καλύτερη σχέση κόστους ανά GB. Αντίθετα, αν θέλουμε ταχύτητα, τότε επιλέγουμε δίσκους SAS, αυξάνοντας το κόστος ή μειώνοντας την συνολική χωρητικότητα. Αν η ταχύτητα είναι μονόδρομος, τότε η επιλογή δεν είναι άλλη από τους δίσκους SSD. Στα nodes η πλέον συνηθισμένη επιλογή είναι οι SSD και λιγότερο οι SAS, λόγω του αυξημένου μεγέθους τους, εκτός αν ο βασικότερος παράγοντας είναι το κόστος οπότε η επιλογή θα είναι SATA. Στην Εικόνα 7 [9] παρουσιάζεται η σχέση τεχνολογίας σκληρών δίσκων σε σχέση με το κόστος τους και στην Εικόνα 8 τα χαρακτηριστικά των μαγνητικών και SSD δίσκων αποθήκευσης [10], [11].

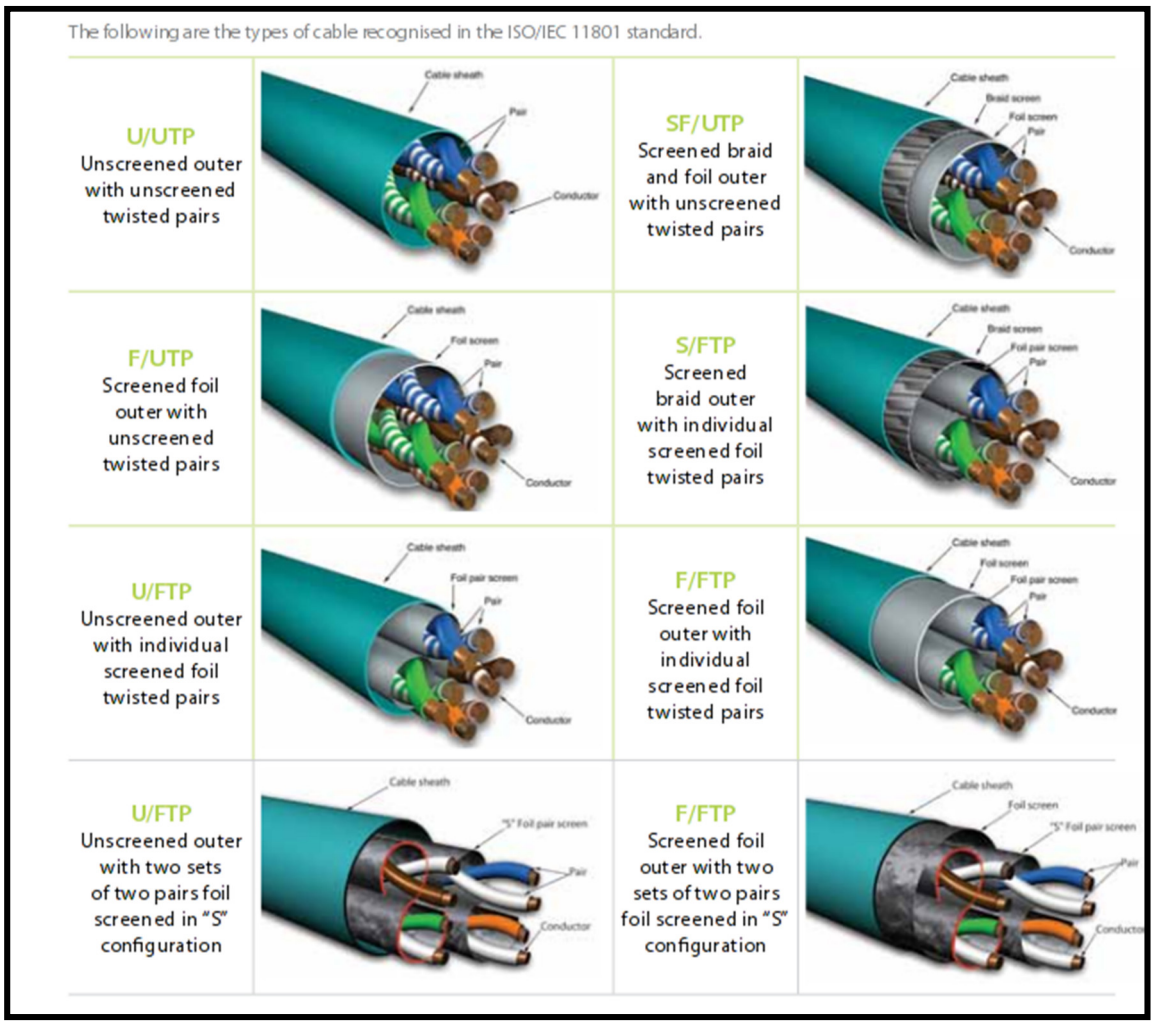


Εικόνα 8. Ανατομία μαγνητικού (περιστροφικού) και SSD σκληρού δίσκου.

1.4.2. Διασύνδεση – δικτύωση υπολογιστών

Ο ρόλος του δικτύου που θα χτιστεί μεταξύ των nodes είναι πολύ σημαντικός. Η σημερινή τεχνολογία, προσφέρει επιλογές ανάλογα με την ταχύτητα, την απόσταση μεταξύ τους, την ασφάλεια και την πιστότητα στα δεδομένα που ανταλλάσσονται, αλλά και τον πόσο ηλεκτρονικό θόρυβο έχουμε στο χώρο που έχουμε εγκαταστήσει τον εξοπλισμό μας. Οι δυνατές επιλογές που μπορούμε να έχουμε στο επίπεδο του τοπικού δικτύου, είναι η χρήση συστραμμένων καλωδίων, η χρήση οπτικών ινών, ή ο συνδυασμός αυτών. Στο επίπεδο των δικτύων ευρείας περιοχής, έχουμε τις οπτικές ίνες, προπληρωμένες γραμμές, δορυφορικά κυκλώματα και άλλες διάφορες επιλογές ανάλογα με τις ανάγκες, τα διαθέσιμα μέσα και το εύρος ζώνης που χρειάζεται. Εξετάζοντας τις ανάγκες μας για την διασύνδεση ενός cluster, μπορούμε να δούμε ότι έχουμε την οικονομική επιλογή των συστραμμένων καλωδίων. Τα ενεργά στοιχεία του δικτύου μας (switch, router) είναι οικονομικότερα, ενώ η κατασκευή των αντίστοιχων καλωδίων (οριζόντια και κάθετη καλωδίωση) είναι αρκετά εύκολη και μπορεί να γίνει απλά χωρίς ιδιαίτερες γνώσεις. Το εύρος ζώνης που είναι διαθέσιμο, είναι 10Mbps, 100Mbps, 1Gbps, αν και πλέον τα 10Mbps θεωρούνται παρελθόν, ενώ για οποιαδήποτε σύστημα απαιτήσεων το εύρος ζώνης 1Gbps είναι μονόδρομος. Δυστυχώς όμως, οι παραπάνω ταχύτητες στα δίκτυα που χρησιμοποιούν τον χαλκό ως μέσω αποστολής και λήψης δεδομένων δεν είναι διασφαλισμένες. Η απόκτηση και η χρήση ενεργών στοιχείων, αλλά και η καλωδίωση που είναι απαραίτητη ώστε να υποστηρίζει αυτές τις ταχύτητες, είναι επιρρεπείς στις παρεμβολές EMI (Electromagnetic Interference). Επίσης, το μήκος των καλωδίων δεν μπορεί να

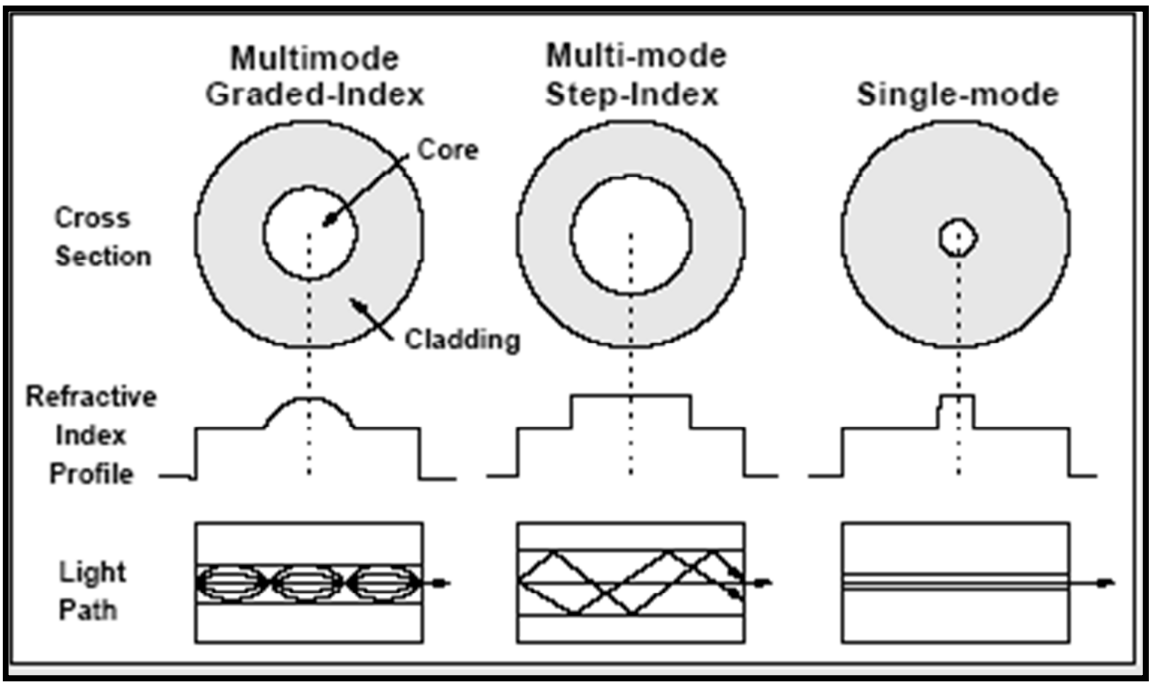
ξεπερνά τα 100 μέτρα. Οι παρεμβολές EMI, μπορούν να προέρχονται από τα καλώδια τροφοδοσίας ή κάποια ηλεκτρική συσκευή που υπάρχει στον χώρο, ενώ ακριβώς για τον ίδιο λόγο, η τεχνολογία αυτή είναι μη ασφαλής στην υποκλοπή δεδομένων. Για την αντιμετώπισή των παρεμβολών, κατασκευάστηκαν καλώδια χαλκού τύπου STP (Screened Twisted Pair), FTP(Foiled Twisted Pair), καθώς και ο συνδυασμός τους S/FTP, αλλά για την πλήρη λειτουργικότητά τους, απαιτείται ξεχωριστή γείωση για αυτά στο κτήριο, καθώς και τακτική γείωσή τους στους καλωδιαδρόμους, που σχεδόν σε καμία εγκατάσταση δεν υφίστανται. Τα είδη των συστραμμένων καλωδίων [12] παρουσιάζονται στην Εικόνα 9.



Εικόνα 9. Είδη καλωδίων συστρεμμένων ζευγών.

Πιο αποδοτική επιλογή, είναι φυσικά η χρήση οπτικών ινών. Είναι ανεπηρέαστες από οποιαδήποτε εξωτερική ηλεκτρομαγνητική παρεμβολή και παρέχουν εξαιρετικά υψηλή αντοχή ενάντια στην υποκλοπή των δεδομένων που διακινούν. Πρωτοπαρουσιάστηκαν τη δεκαετία του 1840 στο Παρίσι, από τους Daniel

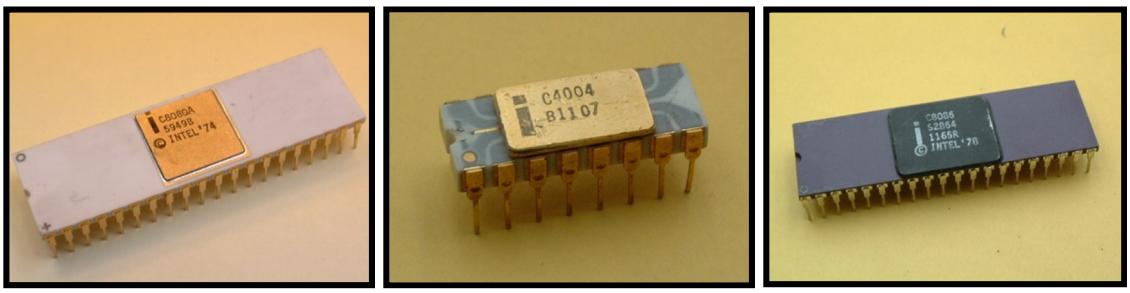
Colladon και Jacques Babinet. Σε επικοινωνιακή εφαρμογή πρώτη φορά χρησιμοποιήθηκαν στο φωτόφωνο, που εφευρέθηκε από τους Alexander Graham Bell και Summer Tainter το 1880 στο εργαστήριο Volta της Ουάσινγκτον. Όπως φαίνεται στην Εικόνα 10, διακρίνονται σε πολύτροπες (multi-mode), οι οποίες είναι αυτές που χρησιμοποιούνται κυρίως σε τοπικά δίκτυα και μπορούν να αναλυθούν μέσων των εξισώσεων την οπτικής, και σε μονότροπες (single-mode), οι οποίες έχουν πολύ μικρότερη διατομή, περιγράφονται μέσω των εξισώσεων των ηλεκτρομαγνητικών κυμάτων και προορίζονται κυρίως για μητροπολιτικά και ευρείας περιοχής δίκτυα [13]. Οι πολύτροπες οπτικές ίνες, ανάλογα με το πρότυπο εκπομπής, μπορούν να φτάσουν σε ταχύτητες από 100Mbps μέχρι και 100Gbps. Σε αυτό, βασικό ρόλο παίζει το μήκος κύματος του διερχόμενου φωτός και το είδος αυτού, δηλαδή αν προέρχεται από LED ή από LASER. Η πιο συνηθισμένη χρήση τους, γίνεται στα 1Gbps και στα 10Gbps με χρήση πομπών φωτός LED. Φυσικά, το κόστος απόκτησης και κατασκευής τοπικού δικτύου οπτικών ινών είναι μεγάλο, ενώ απαιτείται υψηλή εξειδίκευση και εργαλεία μεγάλης ακρίβειας για να τερματιστούν τα καλώδια οπτικών ινών. Γενικότερα, ένα ανεπηρέαστο και γρήγορο τοπικό δίκτυο στο cluster μας, θα επιτρέπει την απροβλημάτιστη και γρηγορότερη εκτέλεση των εφαρμογών παράλληλης επεξεργασίας που του έχουμε εγκαταστήσει.



Εικόνα 10. Διαφοροποίηση Multimode – Single-mode οπτικών ινών.

1.4.3. Κεντρική μονάδα επεξεργασίας - CPU

Τελευταία αλλά και πιο σημαντικά, έρχονται η καρδιά και η μνήμη κάθε υπολογιστικού συστήματος, ο κεντρικός επεξεργαστής ή CPU (Central Prossesing Unit) και η μνήμη RAM (Random Access Memory). Σε ένα cluster, όλα τα προηγούμενα μέρη του hardware που εξετάσαμε, υπάρχουν ώστε να εξυπηρετούν τους διαθέσιμους κεντρικούς επεξεργαστές όπου, με την βοήθεια της μνήμης, να μπορέσουν εκτελέσουν τους κατάλληλους υπολογισμούς και τελικά να μας παρουσιάσουν το ανάλογο αποτέλεσμα. Ιστορικά, η σημερινή μορφή και δομή των CPU, άρχισε να υφίσταται με την εφεύρεση των ολοκληρωμένων κυκλωμάτων, η οποία υλοποιήθηκε με τις εφευρέσεις του Federico Faggin και την παρουσίαση του πρώτου εμπορικά διαθέσιμου μικροεπεξεργαστή 4004 από την Intel. Ο βασικός σχεδιασμός τους αποτελείται από την αριθμητική και λογική μονάδα ALU (Arithmetical and Logic Unit), την μονάδα ελέγχου Control Unit, τους καταχωρητές Registers, τη μονάδα χρονισμού και τους διαύλους δεδομένων Data Bus, με βασικό αυτόν που τον διασυνδέει με την μνήμη RAM [14]. Αν και έχουν περάσει πολλά στάδια στην εξέλιξη των κεντρικών επεξεργαστών, αυτή την στιγμή η πλειοψηφία των υπό χρήση στα συστήματα επιδόσεων και στους προσωπικούς υπολογιστές είναι απόγονοι του 8086 της Intel. Φυσικά, υπήρξαν στην διαδρομή διάφορες τεχνολογίες, όπως ο SPARK της SUN, η σειρά POWER της IBM, τα οποία πλέον βρίσκονται σπάνια σε παλιάς τεχνολογίας συστήματα, οι Vector Prossesor που έχουν εξαιρετική απόδοση στην επεξεργασία γραφικών και τους βρίσκουμε σε διάσημες παιχνιδομηχανές (PlayStation) και σε εξομοιωτές γραφικών στρατιωτικών κυρίως εφαρμογών, και οι επεξεργαστές Tesla της NVIDIA που μέσω της τεχνολογίας CUDA, παραχωρείται επεξεργαστική ισχύς του επεξεργαστή γραφικών GPU (Graphics Prossesing Unit) στο υπόλοιπο σύστημα.



Εικόνα 11. CPU’ s πρώτης γενιάς.

Αν και δεν υπάρχει περιορισμός για ποιο είδος CPU θα χρησιμοποιήσουμε, συνήθως για τη δημιουργία cluster προτιμούμε να χρησιμοποιούμε αυτές που έχουν σχεδιαστεί ειδικά για χρήση σε server. Είναι σχεδιασμένες για να παρέχουν μεγαλύτερη πιστότητα, μεγαλύτερο χρόνο ζωής και κύκλο χρήσης, ενώ παρέχουν διόρθωση σφαλμάτων με χρήση κατάλληλης RAM.

Εικόνα 12. CPU τελευταίας γενιάς.

1.4.4. Η μονάδες μνήμης - RAM

Παράλληλα με τις CPU, εξελίχθηκαν οι μνήμες RAM. Πλέον οι συχνότητα λειτουργίας τους ξεπερνά το 1Ghz για τη μνήμη τεχνολογίας DDR3 (Double Data Rate), ενώ οι προηγούμενες τεχνολογίες πλέον δεν παράγονται. Σημαντικό ρόλο παίζει η ποσότητα της μνήμης. Τα πολύπλοκα μαθηματικά μοντέλα πεπερασμένων στοιχείων τα οποία περιγράφουν ένα σύστημα που θέλουμε να προσομοιώσουμε, καθώς και τα ενδιάμεσα αποτελέσματα, χρειάζονται τεράστια ποσότητα διαθέσιμης μνήμης RAM. Αν δεν μπορεί το υπολογιστικό σύστημά μας να παρέχει τη μνήμη RAM, τότε το λειτουργικό σύστημα στρέφεται στην χρήση προσωρινής μνήμης στον σκληρό δίσκο, την Swap, η οποία όμως είναι πολύ πιο αργή από την RAM. Αυτός είναι ένας από τους λόγους που τα συστήματα 64bit κατέκτησαν την αγορά των cluster, υποστηρίζοντας ποσότητα μνήμης RAM πάνω από 3Gb αντίθετα από τα συστήματα 32bit. Άρα, είναι φανερό ότι η ταχύτητα της CPU, καθώς και η ταχύτητα και η ποσότητα της μνήμης RAM, είναι καθοριστικοί παράγοντες για το χρόνο που απαιτείται από ένα υπολογιστικό πρόβλημα.

Εικόνα 13. Μνήμη RAM.

1.5. Το software των cluster

Πέρα από το hardware, ένα cluster για να λειτουργήσει χρειάζεται το λειτουργικό σύστημα και software. Το software αυτό, έχει σκοπό να εκμεταλλευτεί όσο το δυνατόν καλύτερα το υλικό, ώστε να αποκομίσουμε το βέλτιστο δυνατό αποτέλεσμα, όσο το δυνατόν πιο γρήγορα.

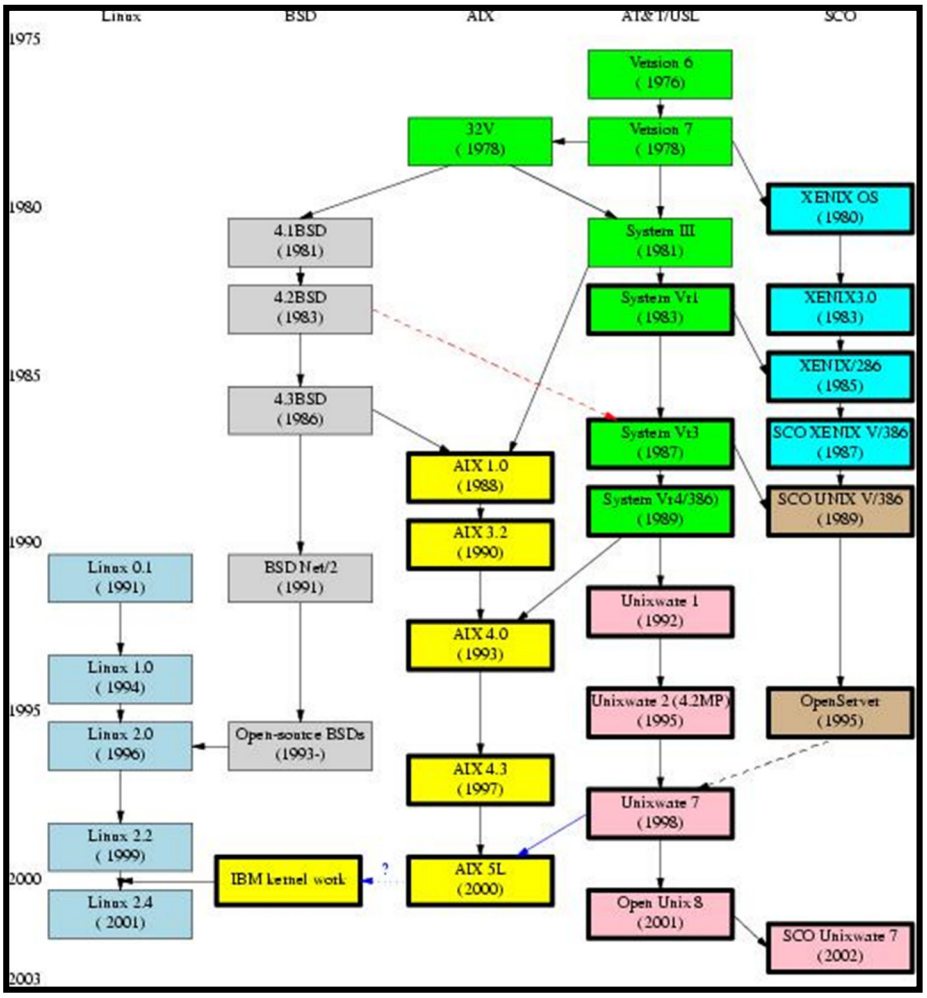
1.5.1. Το λειτουργικό σύστημα

Ο θεμέλιος λίθος του λογισμικού, είναι το λειτουργικό σύστημα. Λειτουργικό Σύστημα κατά τον Andrew S. Tanenbaum στο βιβλίο του «Σύγχρονα Λειτουργικά Συστήματα» [15] είναι: «ένα στρώμα λογισμικού, του οποίου η αποστολή είναι η διαχείριση όλων των συσκευών και η παροχή προγραμμάτων στον χρήστη, για απλούστερη διασύνδεση με το hardware». Δηλαδή, είναι ένα σύνολο προγραμμάτων, που λειτουργεί ως ενδιάμεσος ανάμεσα στην μηχανή και το χρήστη, ενώ προσπαθεί να αξιοποιήσει όσο καλύτερα γίνεται το διαθέσιμο hardware. Η ιστορία των λειτουργικών συστημάτων, ξεκινάει από την στιγμή που υλοποιήθηκε ο πρώτος ηλεκτρονικός υπολογιστής. Εξελίχθηκαν παράλληλα, ώστε να μπορούν να πληρούν τις εκάστοτε απαιτήσεις και εκμεταλλεύονται το υπάρχον διαθέσιμο hardware.

1.5.2. Λειτουργικά συστήματα UNIX – LINUX

Σταθμοί στην ιστορία τους ήταν η ανάπτυξη του UNICS, που μετονομάστηκε από τον Ken Thompson σε UNIX, στις αρχές της δεκαετίας του 1970, το MS-DOS το 1981 από την Microsoft, τον μετέπειτα κολοσσό στο χώρο του λογισμικού, το MAC OS από τη εταιρία Apple το 1984, που ήταν το πρώτο λειτουργικό σύστημα με γραφικό περιβάλλον και τέλος το LINUX που παρουσιάστηκε από τον Linus Torvalds στις 5 Οκτωβρίου 1991 και αποτελεί ένα τύπου UNIX λειτουργικό σύστημα. Η πορεία εξέλιξης του Unix σε Linux παρουσιάζεται στην Εικόνα 14 [16]. Το LINUX είναι το πρώτο πραγματικά ελεύθερο λειτουργικό σύστημα, το οποίο οποιοσδήποτε μπορεί να χρησιμοποιήσει και να το αναπτύξει βάση των αδειών της GNU General Public License. Η εμφάνιση του προκάλεσε επανάσταση στην τεχνολογία των λειτουργικών συστημάτων. Ομάδες εθελοντών, ερασιτεχνών προγραμματιστών, αλλά και μεγάλες εταιρίες έχουν συνεισφέρει στην ραγδαία εξάπλωσή του, κυρίως σε servers και λιγότερο στους προσωπικούς υπολογιστές. Αυτό συμβαίνει γιατί εκμεταλλεύεται καλύτερα τους πόρους, είναι παραμετροποιήσιμο σε πολύ μεγάλο βαθμό, χωρίς να περιέχει σημεία που να είναι μαύρα κουτιά όπως είναι π.χ. στα Windows, ενώ ο μεγάλος αριθμός επαγγελματιών ή μη προγραμματιστών που το αναπτύσσουν, το βελτιώνουν διαρκώς ώστε να εισάγει καινοτομίες ταχύτερα από άλλα λειτουργικά συστήματα.

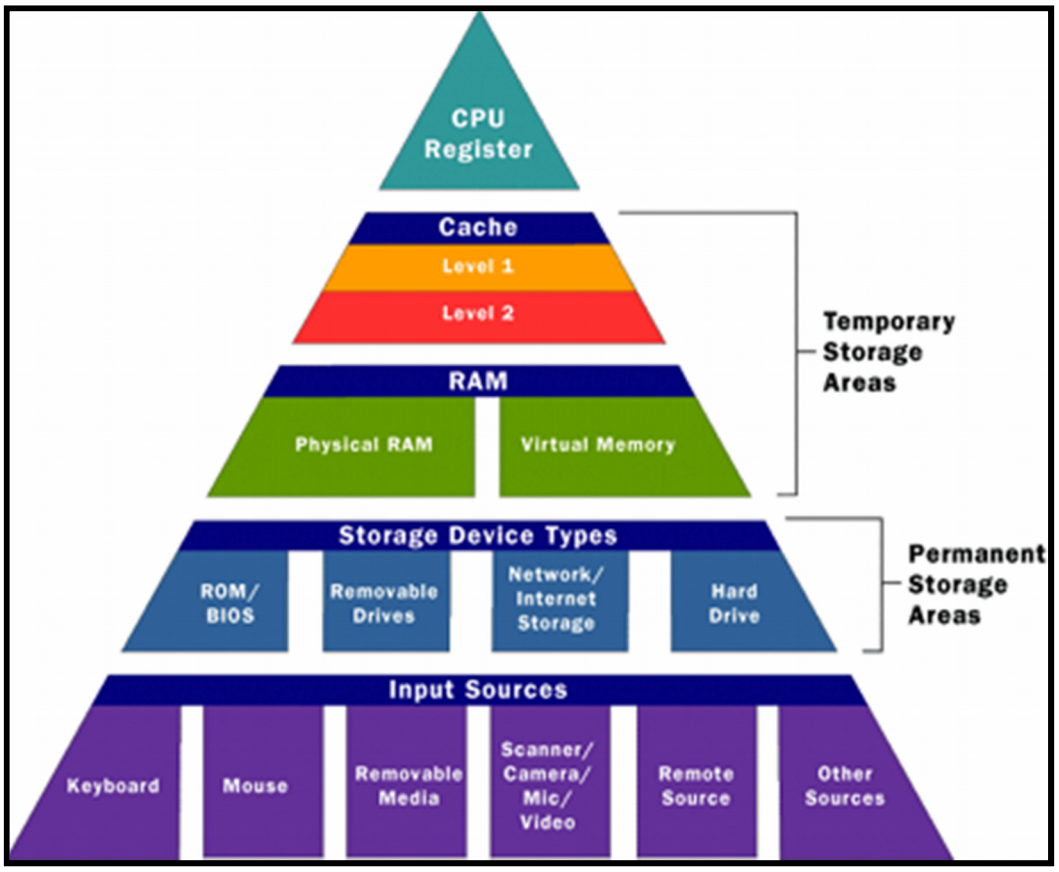
Παράλληλα, προσφέρει δυνατότητα λειτουργίας χωρίς γραφικό περιβάλλον, είτε με γραφικό περιβάλλον επιλογής μας και την ταυτόχρονη χρήση από πολλούς διαφορετικούς χρήστες. Το LINUX είναι πλέον η πρώτη επιλογή για χρήση σε υπερυπολογιστές και σε cluster, λόγω της ευελιξίας του να μπορεί να καταναλώνει ελάχιστους πόρους όταν υπάρχει ανάγκη για τέτοια διαμόρφωση.



Εικόνα 14. Η πορεία των λειτουργικών συστημάτων UNIX – LINUX.

1.5.3. Δομή και λειτουργία λειτουργικού συστήματος

Βασικό συστατικό στοιχείο των λειτουργικών συστημάτων είναι ο πυρήνας ή kernel. Το kernel είναι το ενδιάμεσο στρώμα ανάμεσα στο υλικό και τον χρήστη ή τις εφαρμογές που εκτελεί ο χρήστης. Διαχειρίζεται τις συσκευές, τη μνήμη RAM και την CPU με τέτοιο τρόπο, ώστε να γίνεται η καλύτερη δυνατή χρήση τους. Γύρω από το kernel, υπάρχουν τα κελύφη ή shells, τα οποία ανάλογα με τα προνόμια της εφαρμογής ή του χρήστη, του επιτρέπουν να πλησιάσει στις δυνατότητες που του παρέχει ο πυρήνας. Αυτό φυσικά γίνεται για λόγους ασφαλείας, ώστε μία λανθασμένη επιλογή να μην οδηγήσει το σύστημα σε κατάρρευση.



Εικόνα 15. Ιεραρχία λειτουργικού συστήματος.

Παράλληλα, είναι υπεύθυνο να ελέγχει και να δίνει προτεραιότητα στις διεργασίες, να μοιράζει τον χρόνο που θα έχει η καθεμία σε κάθε κύκλο στην CPU και ανάλογα με τις δυνατότητες του hardware να ελέγχει την πολυεπεξεργασία, δηλαδή την ταυτόχρονη εκτέλεση πολλαπλών εφαρμογών. Στο επίπεδο της μνήμης, παραχωρεί τον χώρο που χρειάζεται κάθε εφαρμογή και αποφασίζει τι θα στείλει στην εικονική μνήμη ή swap στον σκληρό δίσκο και πότε. Ένα σύγχρονο λειτουργικό σύστημα, εκτελεί ταυτόχρονα εκατοντάδες ελέγχους σε συσκευές software και περιφερειακά, που θα ήταν πολύ κουραστικό απλά να τα αναφέρουμε. Η ιεραρχία λειτουργίας του λειτουργικού συστήματος απεικονίζεται στην Εικόνα 15 [17].

1.5.4. Software παράλληλης επεξεργασίας

Πέρα από το λειτουργικό σύστημα, ένα cluster με προσανατολισμό την παράλληλη επεξεργασία προβλημάτων, απαιτεί την χρήση επιπλέον λογισμικού. Αρχικά, θα χρειαστούμε ένα πρόγραμμα που θα κάνει cluster να λειτουργεί σαν

μία οντότητα. Όπως προαναφέραμε να ένα από τα πρώτα προγράμματα που έκαναν αυτή την δουλειά είναι το PVM (Parallel Virtual Mashine). Σε λειτουργικά συστήματα, όπως το LINUX, έχουμε έτοιμα εργαλεία για την λειτουργία και συντήρηση cluster, που κάνουν ευκολότερη και αποδοτικότερη την εργασία του προσωπικού που το ελέγχει. Έπειτα, θα χρειαστούμε μία εφαρμογή που θα εκτελεί την παράλληλη επεξεργασία και θα διαμοιράζει την προς εκτέλεση εφαρμογή στις CPU του κάθε node. Επειδή είναι σπάνιο οι CPU να αποτελούνται από ένα πυρήνα ή core, θα αναφερόμαστε στα core αφού κάθε ένα είναι και μπορεί να λειτουργήσει σαν ανεξάρτητη CPU. Τα προγράμματα που είναι σχεδιασμένα να επιτελούν αυτή την λειτουργία, είναι τα προγράμματα Διασύνδεσης Μεταβίβασης Μηνυμάτων MPI (Message Passing Interface). Το MPI ουσιαστικά είναι ένα πρωτόκολλο επικοινωνίας για την εκτέλεση παράλληλης επεξεργασίας. Είναι σχεδιασμένο για να συνεργάζεται με εφαρμογές γραμμένες σε γλώσσα προγραμματισμού C και Fortran. Αναλαμβάνει να συγχρονίσει τα node και τα core τους, να από-δομήσει την προς εκτέλεση εφαρμογή, να την διανείμει στα διαθέσιμα core και όταν είναι έτοιμα τα αποτελέσματα, να τα συλλέξει και να τα ανασυνθέσει, ώστε η εφαρμογή που εκτελέστηκε να παρουσιάσει τα αποτελέσματα. Υπάρχουν διάφορες διαθέσιμες εφαρμογές MPI, άλλες είναι ελεύθερες προς χρήση από το καθένα και άλλες είναι εμπορικά διαθέσιμες, αποτελώντας πνευματική ιδιοκτησία μεγάλων εταιριών του χώρου της πληροφορικής. Μερικές από αυτές είναι, το OpenMPI, το MPICH, το OpenMP, που είναι ελεύθερα διαθέσιμα και το Intel MPI και το IBM Platform MPI, που είναι μόνο εμπορικά διαθέσιμα. Εφόσον το MPI εγκαταστάθηκε και λειτουργεί, το σύστημά μπορεί να εκτελέσει παράλληλη επεξεργασία, οπότε και ακολουθεί το software που θα κάνει την επίλυση των προβλημάτων προσομοίωσης.

ΚΕΦΑΛΑΙΟ 2. Cluster hardware και Software

2.1. Διαθέσιμο hardware για την κατασκευή συστοιχίας ηλεκτρονικών υπολογιστών

Στόχος της παρούσας εργασίας είναι η δημιουργία και βελτιστοποίηση συστήματος συστοιχίας υπολογιστών για την παράλληλη επεξεργασία προβλημάτων προσομοίωσης υψηλών απαιτήσεων σε περιβάλλον LINUX. Η πρώτη παράμετρος που εξετάζουμε είναι το διαθέσιμο hardware. Οι διαθέσιμες συνθέσεις είναι δύο. Η πρώτη, Συστοιχία -Α, αποτελείται από τέσσερις βιομηχανικού τύπου ηλεκτρονικούς υπολογιστές, οι οποίοι βρίσκονται τοποθετημένοι εντός ικριώματος και ένα επιτραπέζιο ηλεκτρονικό υπολογιστή. Η δεύτερη Συστοιχία -Β, αποτελείται από ένα ολοκληρωμένο σύστημα της Hewlett-Packard, το HP BladeSystem c7000, στο οποίο έχει προστεθεί ένα σύστημα SAN.

2.1.1. Συστοιχία -Α

Η πρώτη Συστοιχία, αποτελείται από 5 συνολικά ηλεκτρονικούς υπολογιστές. Όπως φαίνεται και στην Εικόνα 16, στην ουσία, πρόκειται για όμοια μηχανήματα στα οποία οι διαφορές επικεντρώνονται: α) στο κουτί, β) Στην κάρτα γραφικών και γ) στην ύπαρξη δεύτερης κάρτας δικτύου.

Εικόνα 16. Η Συστοιχία –Α.

Οι 5 υπολογιστές αποτελούνται από μητρική κάρτα ASUS P7H55 (Εικόνα 17), η οποία διαθέτει chipset το H55 της Intel, δέχεται επεξεργαστές i7, i5 και i3 socket LGA1156, με πυρήνες Lynnfield και Clarkdale. Κεντρικό επεξεργαστή έχουν τον Intel i7 870, ο οποίος είναι χρονισμένος στα 2,93Ghz, διαθέτει 4 πυρήνες, 2 νήματα σε κάθε πυρήνα και 8MB μνήμη cache. Φυσικά, είναι αρχιτεκτονικής 64bit και μπορεί να συνεργαστεί με μέχρι 16GB μνήμη RAM. Η Μνήμη RAM αποτελείται από 2 DIMM των 2 GB DDR3, άρα έχουμε συνολικά 4 GB, και είναι χρονισμένη στα 1333 Mhz.

Εικόνα 17. Μητρική της Συστοιχίας –Α.

Σκληρό δίσκο διαθέτουν τον Caviar Blue WD5000AAKS της Western Digital, χωρητικότητας 500GB, με πρωτόκολλο επικοινωνίας SATA 2, μεγέθους 3,5” και ταχύτητας περιστροφής 7200rpm. Μία από τις διαφορές είναι ότι διαθέτουν διαφορετικό κουτί. Οι βιομηχανικού τύπου, έχουν κουτί με προδιαγραφές για τοποθέτηση εντός ικριώματος και ύψος 2U, ενώ ο επιτραπέζιος, έχει ένα κοινό

κουτί επιτραπέζιου ηλεκτρονικού υπολογιστή (Εικόνα 16). Η επόμενη διαφορά, είναι η εγκατεστημένη κάρτα γραφικών. Οι βιομηχανικού τύπου, διαθέτουν την Mobility Radeon HD 4500 της ATI, χαμηλού προφίλ με μνήμη γραφικών 512MB, ενώ ο επιτραπέζιος, έχει την 8800GT της NVIDIA με 1GB μνήμη γραφικών. Τέλος, οι 4 ηλεκτρονικοί υπολογιστές βιομηχανικού τύπου, διαθέτουν επιπλέον κάρτα δικτύου, η οποία βρίσκεται τοποθετημένη σε ένα από τα PCI slots τους.

2.1.2. Συστοιχία -B

Στη δεύτερη Συστοιχία που παρουσιάζεται στην Εικόνα 18, βρίσκουμε το HP BladeSystem c7000 [18]. Αυτό, είναι ένα ολοκληρωμένο σύστημα server της Hewlett-Packard.

Εικόνα 18. Η Συστοιχία –B

Η Συστοιχία –Β αποτελείται από:

a) **16 nodes ProLiant BL260c G5**
b) **2 BladeSystem c7000 Onboard Administrator**
c) **2 HP 1/10Gb VC-Enet Module**
d) **2 HP 4Gb VC-FC Module**
e) **10 Active Cool 200 Fan**
f) **6 HP BladeSystem c-Class P/S**
g) **1 Insight Display**
h) **2 HP StorageWorks 4/8 SAN Switch της Hewlett-Packard**
i) **2 HP StorageWorks MSA 2000, στα οποία είναι τοποθετημένα σε θήκες εύκολης αφαίρεσης 14 σκληροί δίσκοι SAS Cheetah 15K.5 της Seagate.**

Το HP BladeSystem c7000, βρίσκεται τοποθετημένο εντός ικριώματος της Hewlett-Packard, στο οποίο είναι τοποθετημένα τα **h)** και **i)** υποσυστήματα.

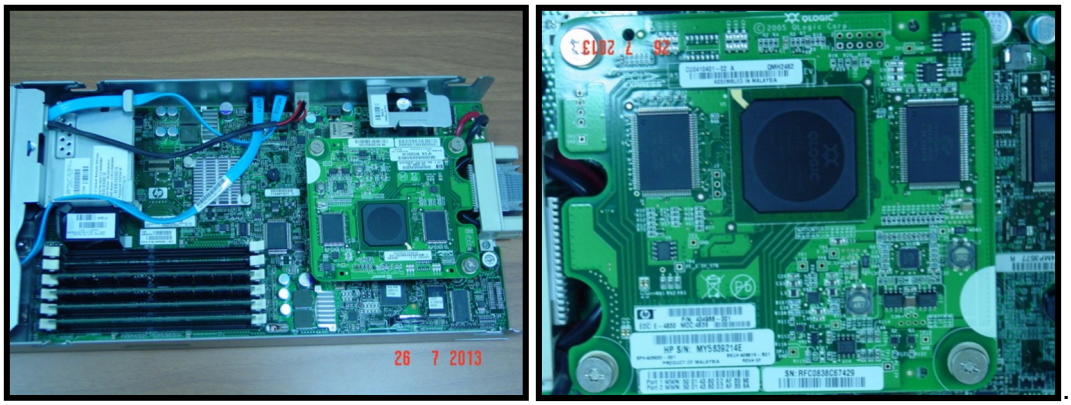
a) ProLiant BL260c G5

Τα 16 ProLiant BL260c G5, είναι τα nodes που παρέχει το HP BladeSystem c7000 και παρουσιάζονται στην Εικόνα 19. Είναι τοποθετημένα στο εμπρός τμήμα αυτού σε 2 σειρές των οκτώ, ενώ το κουτί τους είναι ταυτόχρονα και αποσπώμενη θήκη, όπως παρουσιάζεται στην Εικόνα 20. Πρόκειται για ηλεκτρονικούς υπολογιστές με δυνατότητες server. Έχουν BIOS αυξημένων δυνατοτήτων, με δυνατότητα απομακρυσμένου χειρισμού μέσω εικονικού KVM (Keyboard - Video – Mouse), RAID και παραμετροποίηση εγκατεστημένων περιφερειακών. Σε αυτά βρίσκουμε εγκατεστημένους 2 κεντρικούς επεξεργαστές (CPU) Xeon E5405, (Εικόνα 22) που είναι χρονισμένοι στα 2 Ghz. Ο καθένας από αυτούς διαθέτει 4 πυρήνες και 12MB μνήμη cache, άρα κάθε node αποτελείται από 8 πυρήνες. Φυσικά, είναι αρχιτεκτονικής 64bit και μπορεί να συνεργαστεί με μνήμη RAM τεχνολογίας αυτόματης διόρθωσης σφαλμάτων ECC. Η Μνήμη RAM είναι τεχνολογίας DDR2 ECC, αποτελείται από 2 DIMM των 4 GB και 2 DIMM των 512MB, άρα έχουμε συνολικά 9 GB, ενώ είναι χρονισμένη στα 667 Mhz. Διαθέτουν 2 σκληρούς δίσκους MHZ2120BS G1 της Fujitsu, όπως φαίνεται στην Εικόνα 22, χωρητικότητας 120GB έκαστος, με πρωτόκολλο επικοινωνίας SATA , μεγέθους 2,5” και ταχύτητας περιστροφής 5400rpm. Η κάρτα γραφικών είναι η ES1000 της ATI, η οποία είναι Onboard και παρέχει βασικές υπηρεσίες απεικόνισης γραφικών. Υπάρχουν 2 κάρτες δικτύου, οι οποίες συνδέονται απ' ευθείας μέσω του κεντρικού βύσματος της μητρικής των ProLiant BL260c G5 στο εσωτερικό δίκτυο του HP

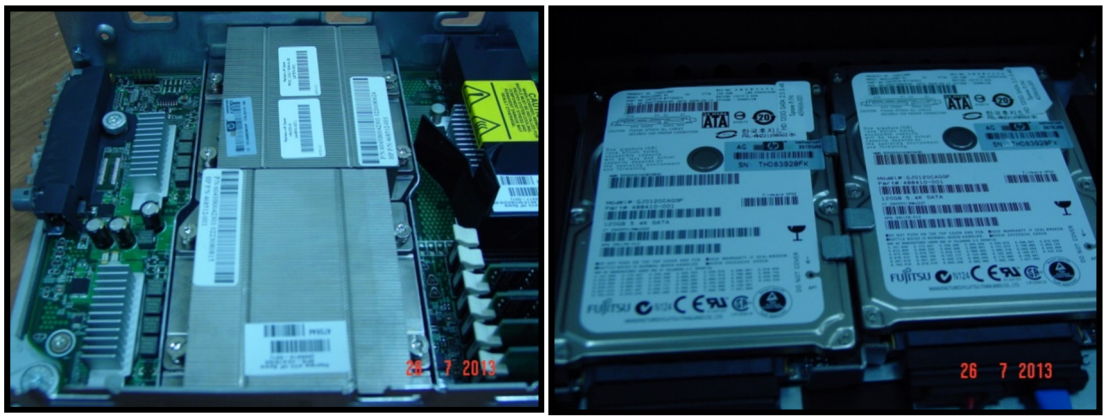
BladeSystem c7000 και από εκεί στα HP 1/10Gb VC-Enet Module. Τέλος, στην ειδική θύρα επέκτασης της μητρικής κάρτας mezzanine slot, είναι εγκατεστημένη η κάρτα QHM2462 της QLOGIC (Εικόνα 21), η οποία παρέχει 2 θύρες fibre-channel FC σε κάθε node. Αυτές, συνδέονται απευθείας, μέσω του κεντρικού βύσματος της μητρικής των ProLiant BL260c G5, στο εσωτερικό δίκτυο του HP BladeSystem c7000 και από εκεί στα HP 4Gb VC-FC Module. Η γενική άποψη της μητρικής κάρτας του ProLiant BL260c G5, φαίνεται στην Εικόνα 21.

Εικόνα 19. Συστοιχία των 16 ProLiant BL260c G5.

Εικόνα 20. Το ProLiant BL260c G5 όταν αφαιρεθεί από την συστοιχία.



Εικόνα 21. Μητρική κάρτα και κάρτα QHM2462.



Εικόνα 22. Κεντρικοί Επεξεργαστές και σκληροί δίσκοι.

b) BladeSystem c7000 Onboard Administrator

Αυτές οι συσκευές, είναι υπεύθυνες για τον έλεγχο και τη διαχείριση όλων των συσκευών που περιλαμβάνονται στο HP BladeSystem c7000 (Εικόνα 23). Βρίσκονται στο πίσω τμήμα του και είναι εύκολα αποσπώμενες από αυτό. Μέσω ιστοσελίδας, μπορούμε δικτυακά να ελέγξουμε την καλή λειτουργία, τις παραμέτρους και τον τρόπο πρόσβασης σε καθετί που αποτελεί το HP BladeSystem c7000. Η μία λειτουργεί σαν active και η άλλη stand-by, με διαφορετικές διευθύνσεις IP, ώστε να διασφαλίζεται η απρόσκοπτη λειτουργία του υπολογιστικού συστήματος σε περίπτωση βλάβης ή αστοχίας υλικού. Είναι σε άμεση επικοινωνία με το Insight Display, το οποίο είναι ο τρόπος να επικοινωνεί αν δεν έχει επιτευχθεί επικοινωνία μέσω ιστοσελίδας.

c) HP 1/10Gb VC-Enet Module

Τα HP 1/10Gb VC-Enet Module, είναι η συσκευές που αναλαμβάνουν να γεφυρώσουν το εσωτερικό δίκτυο του HP BladeSystem c7000 με τον έξω κόσμο.

Βρίσκονται στο πίσω τμήμα του και είναι εύκολα αποσπώμενες από αυτό, όπως βλέπουμε στην Εικόνα 23. Καθεμία μπορεί να προγραμματιστεί ανεξάρτητα, ανάλογα με τις ανάγκες μας, αλλά χωρίς την σωστή χρήση τους δεν μπορεί να λειτουργήσουν σωστά οι κάρτες δικτύου των ProLiant BL260c G5. Αποτελούνται από 8 θύρες ethernet RG45 και 2 θύρες 10Gbps για μεταξύ τους δικτύωση. Η ρύθμισή τους γίνεται μέσω ξεχωριστής ιστοσελίδας (δεν μπορεί να γίνει απευθείας από τον Onboard Administrator), η οποία ονομάζεται Virtual Connect Manager.

d) HP 4Gb VC-FC Module

Τα HP 4Gb VC-FC Module, είναι οι συσκευές που αναλαμβάνουν να συνδέσουν τις fibre-channel QHM2462 κάρτες μέσω του εσωτερικού δικτύου του HP BladeSystem c7000 με τα fibre-channel switch και τελικά με τις συσκευές SAN HP StorageWorks MSA 2000. Όπως και οι HP 1/10Gb VC-Enet Module, βρίσκονται στο πίσω τμήμα του (Εικόνα 23) και αποσπώνται με ευκολία από αυτό. Καθεμία μπορεί να προγραμματιστεί ανεξάρτητα, ανάλογα με τις ανάγκες μας μέσω του Virtual Connect Manager. Αποτελούνται από τέσσερις υποδοχές για SFP (Small Form Pluggable Transceiver) οπτικών ινών, όπου βρίσκονται τέσσερα SFP HP 4G SW, για οπτικές ίνες με διπλό LC βύσμα.

Εικόνα 23. HP BladeSystem c7000,πάνω προς τα κάτω: 5 Active Cool 200 Fan, 2 HP 1/10Gb VC-Enet Module, 2 HP 4Gb VC-FC Module, 2 BladeSystem c7000 Onboard Administrator.

e) Active Cool 200 Fan

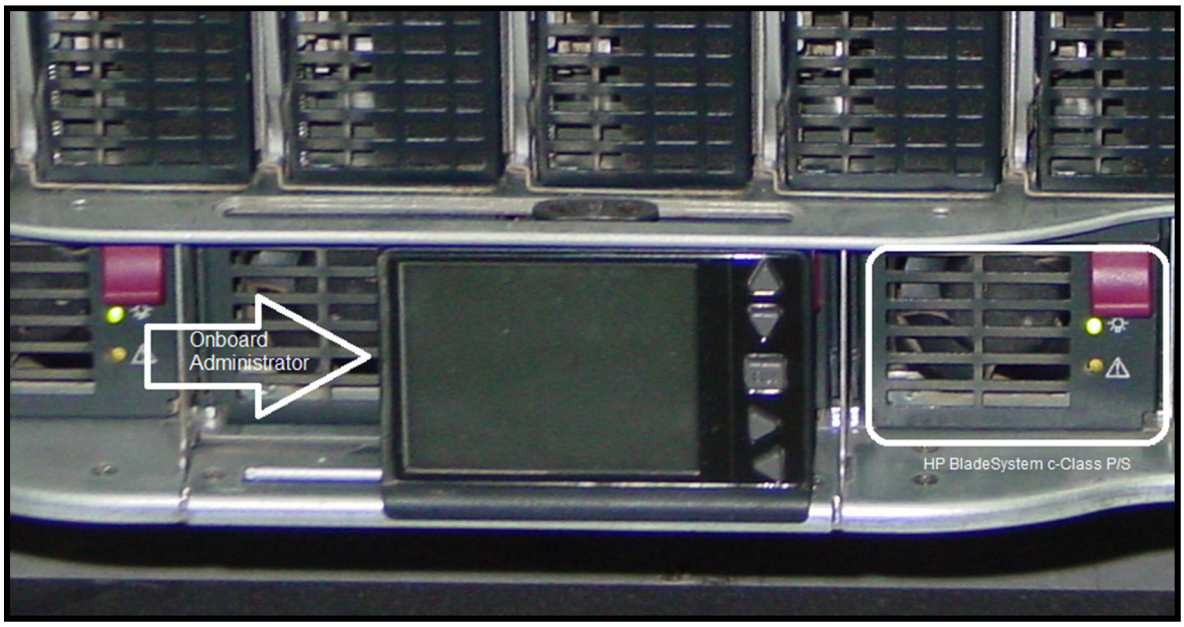
Αυτοί είναι οι ανεμιστήρες που αντλούν τον θερμό αέρα από το εσωτερικό του HP BladeSystem c7000, στέλνοντάς τον στο πίσω τμήμα του ικριώματος. Η ταχύτητα περιστροφής τους ελέγχεται από το σύστημα ψύξης και ενέργειας του HP BladeSystem c7000. Είναι τοποθετημένοι στο πίσω τμήμα του, σε δύο σειρές των πέντε, η μία στο πάνω μέρος και η άλλη στο κάτω μέρος. Μπορούμε να τις διακρίνουμε πολύ εύκολα στην Εικόνα 23.

f) HP BladeSystem c-Class P/S

Το σύστημά μας, αναμφίβολα, χρειάζεται τροφοδοσία για να λειτουργήσει. Αυτό, το παρέχουν τα HP BladeSystem c-Class P/S, όπως παρουσιάζονται στην Εικόνα 24. Το καθένα μπορεί να αποδώσει μέχρι 2250W στο σύστημα, ενώ μέσω διάφορων τεχνικών του Onboard Administrator μπορεί να συνεργαστούν και τα έξι τροφοδοτικά, ώστε να υπάρχει δυνατότητα συνεχούς λειτουργίας σε περίπτωση βλάβης ή αστοχίας κάποιου.

g) Insight Display

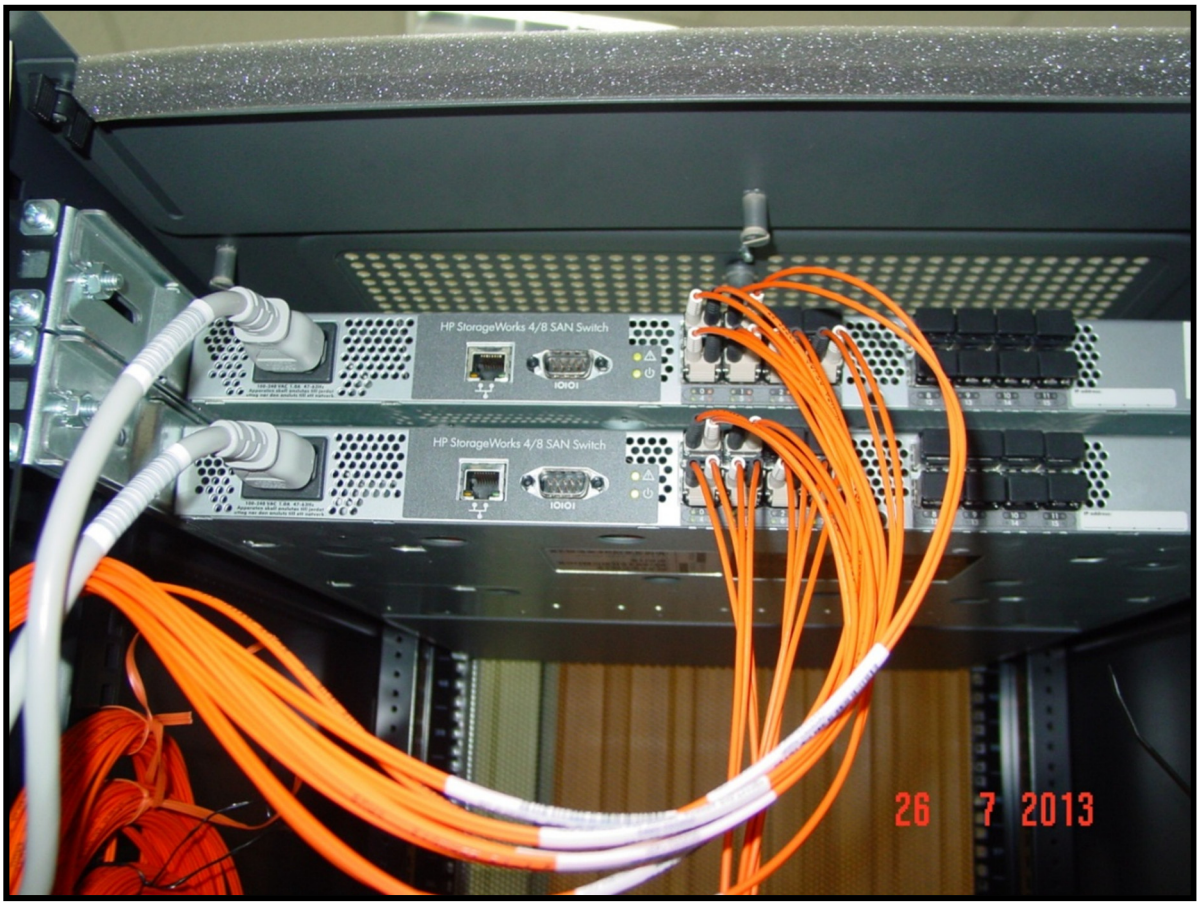
Το Insight Display (Εικόνα 24), είναι ένα interface ανάμεσα στον χρήστη – συντηρητή του HP BladeSystem c7000. Μπορεί να διεξάγει βασικό έλεγχο για σφάλματα, να πραγματοποιήσει την αρχική παραμετροποίηση του συστήματος, αλλά και μία εναλλακτική μέθοδο για επισκόπηση και επαναφορά σε λειτουργία του HP BladeSystem c7000, σε περίπτωση απώλειας τις πρόσβασης μέσω του δικτύου.



Εικόνα 24. Insight Display και HP BladeSystem c-Class P/S.

h) HP StorageWorks 4/8 SAN

Τα HP StorageWorks 4/8 SAN, είναι εξειδικευμένα fibre-channel switch (Εικόνα 25). Είναι υπεύθυνα για να συνδέουν τις HP 1/10Gb VC-Enet Module με τα HP StorageWorks MSA 2000, ενώ δεν έχουν την δυνατότητα να συνδεθούν σε συμβατικά δίκτυα υπολογιστών. Η ρύθμισή τους μέσω java-applet ή telnet δεν είναι ιδιαιτέρα εύκολη και απαιτεί αρκετά μεγάλο βαθμό εξειδίκευσης, λόγω της απαίτησης για λεπτομερή αποτύπωση του τρόπου δικτύωσης των συνδεδεμένων συσκευών σε αυτό. Το καθένα από αυτά αποτελείται από δύο οκτάδες υποδοχών για SFP οπτικών ινών, όπου βρίσκονται δεκαέξι SFP HP 4G SW (Εικόνα 27), για οπτικές ίνες με διπλό LC βύσμα. Από αυτές, μόνο οι οκτώ είναι ενεργές σε κάθε switch, λόγω της άδειας χρήσης που έχει χρησιμοποιηθεί για την ενεργοποίησή τους. Το κάθε HP StorageWorks 4/8 SAN λειτουργεί ανεξάρτητα από το άλλο, ώστε σε περίπτωση βλάβης ή αστοχίας υλικού να μην διακοπεί η διασύνδεση με τα HP StorageWorks MSA 2000.



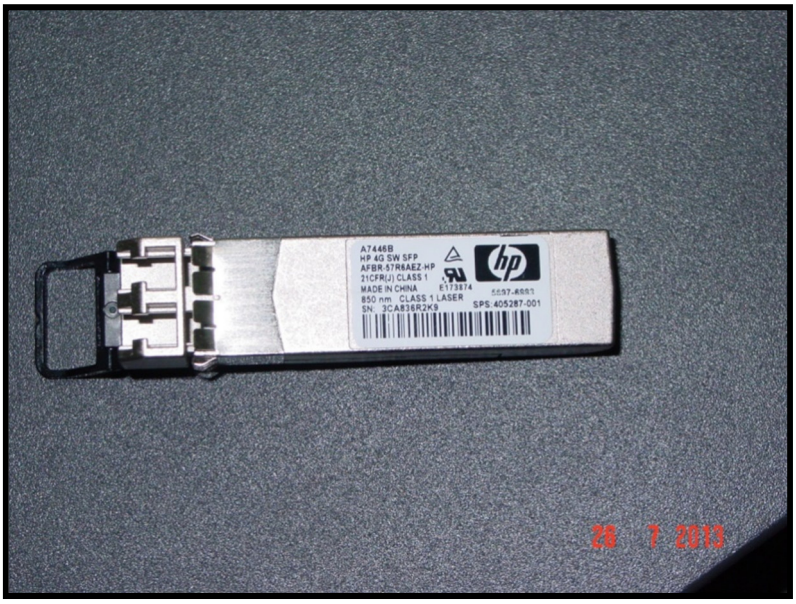
Εικόνα 25. HP StorageWorks 4/8 SAN.

i) HP StorageWorks MSA 2000

Η καρδιά του συστήματος SAN που συνεργάζεται με το HP BladeSystem c7000, χτυπάει στο HP StorageWorks MSA 2000, όπως βλέπουμε στην Εικόνα 26. Είναι η συσκευή στην οποία βρίσκονται οι σκληροί δίσκοι του SAN, όπου γίνεται η διαχείριση και ο έλεγχος καλής λειτουργίας σε επίπεδο hardware του συστήματος. Αποτελείται από 2 μονάδες, οι οποίες συνδέονται με δίκτυο 10Gbps, όπου η μία ορίζεται master και η άλλη slave. Η καθεμία μπορεί να δεχτεί μέχρι δώδεκα σκληρούς δίσκους 3,5”. Υπάρχουν συνολικά δεκατέσσερις σκληροί δίσκοι SAS Cheetah 15K.5 των 300GB έκαστος, όπου οι 2 είναι στο slave HP StorageWorks MSA 2000 και οι υπόλοιποι δώδεκα βρίσκονται στο master HP StorageWorks MSA 2000. Το καθένα έχει 2 τροφοδοτικά με δυνατότητα εναλλαγής τους, χωρίς όμως να απενεργοποιηθεί η συσκευή. Η επικοινωνία με το υπόλοιπο σύστημα SAN, γίνεται μέσω 2 υποδοχών SFP στο καθένα (συνολικά τέσσερις) στις οποίες βρίσκονται SFP HP 4G SW, (Εικόνα 27) για οπτικές ίνες με διπλό LC βύσμα. Παράλληλα, υπάρχει μία θύρα ethernet και μία σειριακή RS-232 θύρα στο κάθε

HP StorageWorks MSA 2000, με σκοπό την παραμετροποίηση και την επισκόπηση της καλής λειτουργίας των συσκευών, μέσω εφαρμογής ιστοσελίδας ή telnet αντιστοίχως.

Εικόνα 26. HP StorageWorks MSA 2000.



Εικόνα 27. SFP HP 4G SW.

2.2. Software για την συστοιχία ηλεκτρονικών υπολογιστών

Στο επίπεδο του λογισμικού, σε αντίθεση με το hardware, έχουμε μία ευελιξία επιλογών. Οι πρώτες διαθέσιμες δυνατότητες αυτών βρίσκονται στο

επίπεδο του λειτουργικού συστήματος. Η αρχική επιλογή, ήταν για χρήση λειτουργικού συστήματος LINUX, αφού οι άδειες χρήσης και η απόκτηση άδειας για χρήση Windows σε cluster για παράλληλη επεξεργασία είναι απαγορευτικού κόστους, ενώ η χρήση τρίτου λειτουργικού συστήματος δεν θα επέτρεπε την πλήρη αξιοποίηση του λογισμικού προσομοίωσης.

Οι επιλογές σε λειτουργικό σύστημα LINUX, αν και οι διαθέσιμες διανομές είναι εκατοντάδες, περιορίζονται. Οι κατασκευαστές εξειδικευμένου λογισμικού στο 95% των περιπτώσεων, παρέχουν πλήρη υποστήριξη στις εμπορικές διανομές της Novell-SUSE, που έχουν την ονομασία SLES (Suse Linux Enterprise Server) και στις διανομές της Red Hat που έχουν εμπορική ονομασία Red Hat Enterprise Linux Server ή RHEL. Οι πιο πρόσφατες εκδόσεις τους, είναι η SLES 11 SP2 και η Red Hat Enterprise Linux Server 6. Και τα δύο έχουν απομακρυνθεί ελάχιστα από τα ορθόδοξη φιλοσοφία των διανομών LINUX, χρησιμοποιώντας σαν διαχειριστή πακέτων τη μέθοδο RPM (Red Hat Package Manager). Με αυτό τον τρόπο, αυτοματοποιείται η εγκατάσταση και η διαχείριση του λογισμικού και του λειτουργικού συστήματος.

Παράλληλα με τις προαναφερθείσες διανομές, οι οποίες παρέχουν εγγυημένη υποστήριξη, υπάρχουν και άλλες που προέρχονται από αυτές, χρησιμοποιώντας την άδεια διανομής του λειτουργικού συστήματος LINUX. Υπάρχουν δηλαδή διανομές, οι οποίες έχουν εξάγει τις εμπορικές, πατενταρισμένες και ότι αποτελεί πνευματική ιδιοκτησία από την SLES και την Red Hat Enterprise Linux Server ή χρησιμοποιούν κάτι από αυτά μετά από ειδική άδεια, προσφέροντας περίπου παρόμοια λειτουργικότητα και συμβατότητα με τις εμπορικά διαθέσιμες. Έτσι, μπορούν με μηδενικό κόστος να προσφέρουν αρκετά, ενώ παράλληλα, λόγω της ελεύθερης ανάπτυξής τους, μπορούν να γίνουν πεδίο δοκιμών νέων τεχνολογιών λογισμικού και εκμετάλλευσης αυτού, γεγονός που στα εμπορικά διαθέσιμα δεν μπορεί να γίνει, λόγω της έμφασης που δίνεται στην σταθερότητα και την ασφάλεια του λειτουργικού συστήματος. Αυτές οι διανομές, είναι η από την πλευρά της Red Hat οι :

a) **CentOS, που αποτελεί τον πιο δημοφιλή κλώνο του Red Hat**
b) **ClearOS**
c) **Oracle Linux**
d) **Scientific Linux, μία διανομή που δημιουργήθηκε για να καλύψει τις ανάγκες του CERN και του εργαστηρίου Fermi.**

e) **Yellow Dog Linux**
f) **Fedora Core**

ενώ, η μόνη διανομή που προέρχεται από την SLES, είναι το OpenSUSE.

2.3. Software προσομοίωσης πεπερασμένων στοιχείων και βοηθητικό software.

Πριν από την τελική επιλογή λειτουργικού συστήματος, πρέπει να συνυπολογίσουμε τις απαιτήσεις του λογισμικού το οποίο θα χρησιμοποιήσουμε για την εκτέλεση προσομοιώσεων υψηλών απαιτήσεων FEM. Το ANSYS είναι το πρώτο software του είδους, παρουσιάστηκε στα μέσα της δεκαετίας του 1970 και είναι το δημοφιλέστερο. Μπορεί να διαχειριστεί μέχρι ένα εκατομμύριο βαθμούς ελευθερίας σε κάθε εξεταζόμενο στοιχείο, ενώ με την συνδρομή του LS-DYNA καλύπτει το σύνολο των απαιτήσεων των ερευνητικών κέντρων για ταυτόχρονη ανάλυση διαφόρων φυσικών παραμέτρων. Ταυτόχρονα επιλέγεται συμβατό software MPI με εξέχουσα δωρεάν διανομή το OpenMPI. Επόμενο βήμα είναι η επιλογή του κατάλληλου λειτουργικού συστήματος με βάση τις παραπάνω προϋποθέσεις. Σύμφωνα με την ANSYS, υπάρχει συμβατότητα και με το SLES 11 SP2 και με το Red Hat Enterprise Linux Server 6. Λόγω της υπάρχουσας εμπειρίες με τις διανομές της SUSE, προτιμήθηκε αυτή η διανομή LINUX.

Κύριο χαρακτηριστικό του SUSE, είναι κυρίως η εφαρμογή YAST που διαθέτει. Αυτή η εφαρμογή, παρέχει κεντρική διαχείριση όλων των παραμέτρων του λειτουργικού συστήματος, κάνοντας χρήση γραφικού περιβάλλοντος. Έτσι, γίνεται εύκολο ακόμα και για κάποιον με μικρή εμπειρία να επέμβει και να προβεί στις απαραίτητες ενέργειες. Επίσης, ένα άλλο χαρακτηριστικό, είναι η ύπαρξη πρόσθετων πακέτων add-on, μέσω των οποίων, μπορούμε να διαμορφώσουμε λεπτομερώς τα στοιχεία που επιθυμούμε να δώσουμε στον εξυπηρετητή που ετοιμάζουμε. Έτσι, η λειτουργία High Availability [19], η οποία είναι απαραίτητη για συστήματα που συνεργάζονται με δίσκους fibre-channel, εγκαθίσταται με όλα τα διαθέσιμα πακέτα, με ελάχιστο κόπο και σε μεγάλο βαθμό προ-ρυθμισμένη. Οι ελεύθερες διανομές OpenSUSE, διατηρούν το YAST, το οποίο είναι από διακριτά χαρακτηριστικά σε σχέση με άλλες διανομές LINUX. Παράλληλα, έχουν τις

τελευταίες εκδόσεις σε κάθε διαθέσιμο software, γεγονός που μας δίνει την δυνατότητα να εκμεταλλευτούμε νέες τεχνολογίες. Το κόστος όμως σε όλα αυτά, είναι η μειωμένη σταθερότητα του συστήματος, αφού μπορεί να προκύψουν προβλήματα που η κοινότητα δεν τα έχει επιλύσει ακόμα. Επίσης, δεν υφίστανται πακέτα add-on και όλα πρέπει να γίνουν χειροκίνητα από μηδενική βάση.

ΚΕΦΑΛΑΙΟ 3. Λειτουργικό σύστημα και λογισμικά

3.1. Εγκατάσταση λειτουργικού συστήματος και software

Στο εργαστήριο έχουμε στη διάθεσή μας δύο συστήματα, τις συνθέσεις Α και Β που περιγράφηκαν στο προηγούμενο Κεφάλαιο. Σε αυτά θα εγκατασταθεί λειτουργικό σύστημα και το software, για τις προσομοιώσεις αριθμητικών προβλημάτων προσομοίωσης υψηλών απαιτήσεων. Οι δύο βασικές προϋποθέσεις που πρέπει να καλυφθούν είναι:

a) η σταθερότητα του υπολογιστικού συστήματος και την πλήρη λειτουργικότητα του.
b) η εκμετάλλευση νέων τεχνολογιών και ευκολιών στην αλληλεπίδραση, την παραμετροποίηση του υπολογιστικού συστήματος, αλλά και την προστασία του συστήματος από δικτυακές απειλές

Τα παραπάνω, σε συνδυασμό με το hardware, όπως αναλυτικά περιγράφηκε στο προηγούμενο κεφάλαιο, μας οδηγούν στα εξής:

- στη **Συστοιχία -Α**, δηλαδή τους 4 βιομηχανικού τύπου ηλεκτρονικούς υπολογιστές και τον επιτραπέζιο. Θέλοντας να εκμεταλλευτούμε τις σύγχρονες δυνατότητες hardware, μιας και πρόκειται για σύγχρονους υπολογιστές, με πολύ ισχυρές CPU, αλλά λίγους πυρήνες συνολικά, επιλέγουμε την εγκατάσταση λειτουργικού συστήματος ανοικτού κώδικα (open source) και συγκεκριμένα την διανομή Opensuse. Έτσι, θα μπορέσουμε να εκμεταλλευτούμε στον καλύτερο δυνατό βαθμό τις νέες τεχνολογικές εξελίξεις στο λειτουργικό σύστημα και στο software, την αυξημένη ασφάλεια, αλλά και το αναβαθμισμένο ποιοτικά γραφικό περιβάλλον. Παράλληλα, οι νέες αναβαθμίσεις σε σχέση με πιο παλιές εκδόσεις λειτουργικού συστήματος, θα αυξήσουν την προστασία του συστήματος.

- στην **Συστοιχία -B**, επιλέγουμε την εγκατάσταση πιο σταθερού, αλλά λιγότερο ευέλικτου λειτουργικού συστήματος, με σκοπό την σταθερότητα και την πλήρη εκμετάλλευση του λογισμικού προσομοίωσης, χωρίς να διακινδυνεύουμε την εμφάνιση ασυμβατότητας και αστάθειας του λογισμικού. Αυτό συμβαίνει γιατί η Συστοιχία Β, αν και όχι σύγχρονη, (είναι 3 χρόνια παλαιότερη από την Συστοιχία Α), μπορεί να μας αποδώσει πολύ μεγαλύτερη επεξεργαστική ισχύ, με τους 128 πυρήνες που διαθέτει καθώς και με τον κοινόχρηστο SAN διασυνδεδεμένο με δίκτυο fibre-channel. Επομένως, οδηγούμαστε στην χρήση εμπορικά διαθέσιμης διανομής LINUX της SLES 11 SP2, την οποία και υποστηρίζει επίσημα η ANSYS.

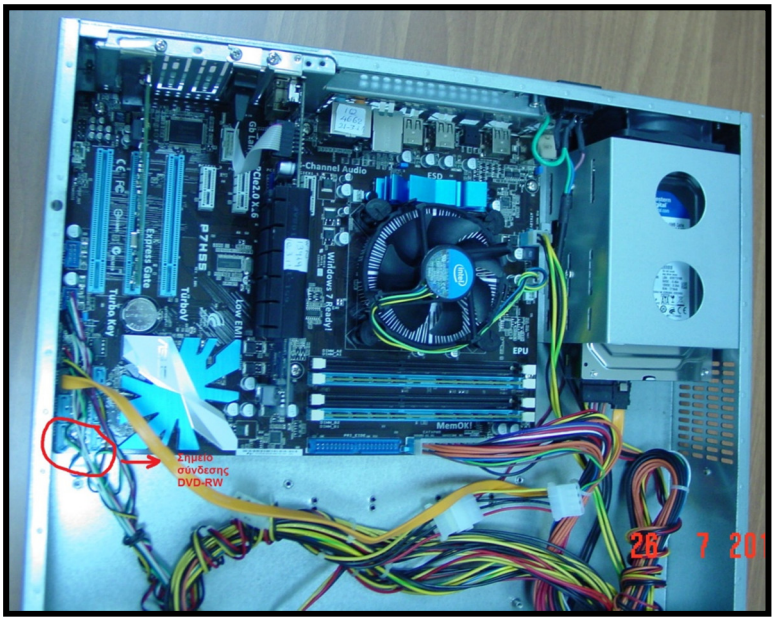
3.2. Εγκατάσταση λειτουργικού συστήματος στη Συστοιχία -A

Αρχικά παρατηρούμε πως στη Συστοιχία Α (Εικόνα 16) οι υπολογιστές, εκτός του επιτραπέζιου, δεν έχουν οπτικούς δίσκους, με αποτέλεσμα να δυσχεραίνεται η προσπάθεια εγκατάστασης λειτουργικού συστήματος. Η λύση δόθηκε με την χρήση αντάπτορα USB to IDE – SATA και ενός DVD-RW όπως βλέπουμε στις εικόνες 28 και 29

Εικόνα 28. αντάπτορα USB σε IDE – SATA.

Παρά τη χρήση του παραπάνω αντάπτορα στο δεύτερο ηλεκτρονικό υπολογιστή βιομηχανικού τύπου, η παραπάνω διάταξη δεν λειτουργούσε σωστά, γιατί το BIOS δεν αναγνώριζε τον οπτικό δίσκο κατά την εκκίνηση. Έτσι, αποφασίστηκε να συνδεθεί απευθείας στην μητρική κάρτα (Εικόνα 30) μετακινώντας το επάνω κάλυμμα και το κουτί του υπολογιστή έξω από το ικρίωμα, ενώ για τροφοδοσία επιλέχθηκε να χρησιμοποιηθεί αυτή από τον αντάπτορα USB σε IDE – SATA, για να αυξηθεί η ευελιξία τοποθέτησης του DVD-RW.

Εικόνα 29. DVD-RW.



Εικόνα 30. Μητρική κάρτα.

Το λειτουργικό σύστημα εγκαθίσταται με όμοιο τρόπο και στους πέντε ηλεκτρονικούς υπολογιστές του Cluster της Συστοιχίας Α. Έτσι θα έχουμε ένα ενιαίο υπολογιστικό σύστημα που για παράλληλη επεξεργασία. Όπως και προαναφέρθηκε, η διανομή που επιλέχθηκε είναι η OpenSuse 12.2, η οποία είναι ώριμη και σταθερή. Το DVD του λειτουργικού συστήματος εισάγεται στο DVD-RW και επιλέγεται να εκκινήσει ο ηλεκτρονικός υπολογιστής από τον οδηγό οπτικού μέσου. Στην Εικόνα 31 έχουμε τις επιλογές που μας παρέχει ο διαχειριστής εγκατάστασης.

Εικόνα 31. Διαχειριστής εγκατάστασης - Φόρτωση πυρήνα εγκατάστασης και οδηγών συσκευών

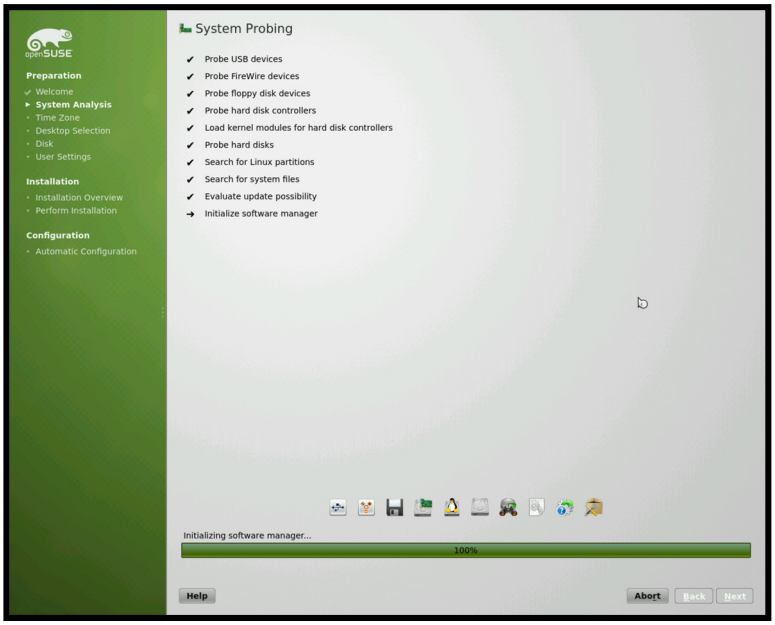
Επιλέγεται Installation και μετά επιλέγεται η γλώσσα εγκατάστασής, όπου στην παρούσα περίπτωση είναι τα Αγγλικά και δεν χρειάζεται να τροποποιηθεί. Πατώντας F3 μπορεί να επιλεχθεί η ανάλυση της οθόνης εγκατάστασης σε διάφορες υποστηριζόμενες αναλύσεις ή σε κατάσταση χωρίς γραφικά,. Η ανάλυση είναι 1280 X 1024. Ακολούθως φορτώνεται ο πυρήνας του λειτουργικού συστήματος, ο οποίος με τη σειρά του θα βρει και θα φορτώσει τις απαιτούμενες εφαρμογές και οδηγούς συσκευών, ώστε να ξεκινήσει η εγκατάσταση του λειτουργικού συστήματος. Στην Εικόνα 31 παρατηρούμε την εκκίνηση του διαχειριστή εγκατάστασης και των οδηγών συσκευών.

Όλα αυτά παίρνουν περίπου 5 λεπτά να εκτελεστούν ίσως λίγο λιγότερο, αν πρόκειται για σύγχρονο σύστημα με γρήγορη μνήμη, αλλά ίσως και παραπάνω αν υπάρχει συσκευή πού απαιτεί οδηγό με ιδιαίτερες απαιτήσεις, όπως είναι, για παράδειγμα, οι ελεγκτές SCSI. Αυτό το μίνι λειτουργικό, ο διαχειριστής εγκατάστασης που θα μας καθοδηγήσει στα επόμενα βήματα, βρίσκεται στην μνήμη του συστήματος (RAM). Αν για οποιοδήποτε λόγο γίνει αντιληπτό ότι δεν λειτουργεί σωστά ή ότι τερματίζεται απροειδοποίητα, τότε πρέπει να επιλεχθεί να μην κάνουμε χρήση του γραφικού συστήματος, αλλά εγκατάσταση μέσω κονσόλας.

Το πρώτο βήμα είναι η επιλογή της γλώσσας εγκατάστασης και της γλώσσας του πληκτρολογίου που χρησιμοποιείται. Στα δύο αυτά drop-box επιλέγεται η Αγγλική ΗΠΑ γλώσσα, για να αποφευχθεί κάθε πιθανότητα τυχόν ασυμβατότητας με το software που θα εγκατασταθεί, αλλά και για να μπορεί ο οποιοδήποτε χρήστης να εκμεταλλευτεί το υπολογιστικό σύστημα.

Στη συνέχεια, γίνεται ανίχνευση σε όλες της συσκευές οι οποίες έχουν την δυνατότητα αποθήκευσης δεδομένων (USB – Firewire – Floppy Disk – Σκληροί

Δίσκοι) καθώς και στους ελεγκτές σκληρών δίσκων. Αφού ολοκληρωθεί αυτή η ενέργεια, γίνεται ανίχνευση διαμερισμάτων με συστήματα αρχείων LINUX στους δίσκους και ακολουθεί εκκίνηση και σύνδεση του διαχειριστή λογισμικού με το repository δεδομένων του οπτικού δίσκου εγκατάστασης του OpenSuse 12.2. Η παραπάνω διαδικασία γίνεται αυτόματα (Εικόνα 32).

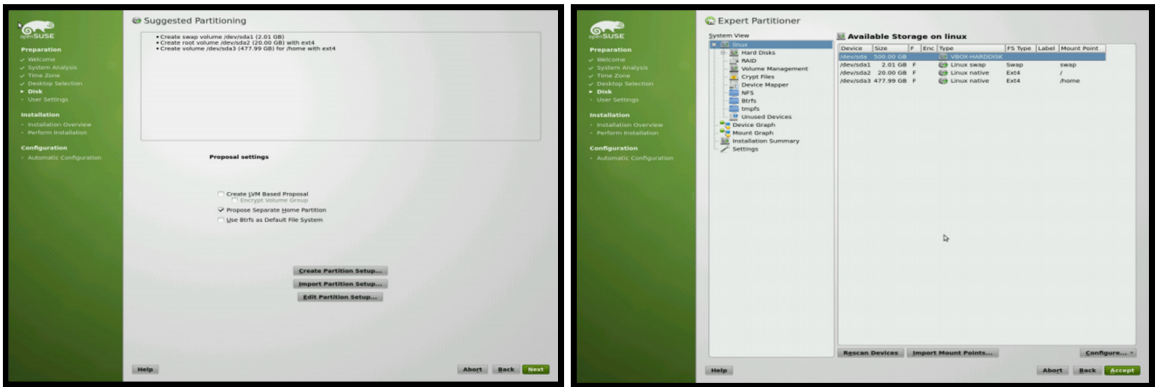


Εικόνα 32. Διαχειριστής εγκατάστασης – ανίχνευση συστήματος.

Επόμενο βήμα είναι η ρύθμιση της ώρας και της ημερομηνίας του συστήματός μας. Η επιλογή αυτή προκαλεί την εμφάνιση μιας οθόνη στην οποία μπορεί να ρυθμιστεί ο χρόνος του συστήματος χειροκίνητα ή μέσω εξυπηρετητή NTP. Ο όσο το δυνατόν ακριβέστερος συγχρονισμός των ηλεκτρονικών υπολογιστών μας είναι μια απαραίτητη συνθήκη ώστε το cluster να λειτουργεί με προβλεπόμενο τρόπο και να μην εμφανίζει παράξενες καθυστερήσεις ή και αλλοιωμένα αποτελέσματα. Η ταυτόχρονη χρήση πολλών κεντρικών μονάδων επεξεργασίας για την εκτέλεση παράλληλης επεξεργασίας δεδομένων και προσομοίωσης, απαιτεί την όσο το δυνατόν πιο μικρή απόκλιση από τον χρόνο αναφοράς. Παράλληλα δεν απαιτείται επεξεργαστική ισχύς για την ανάλυση και συναρμολόγηση επεξεργασμένων δεδομένων τα οποία έχουν χρονική απόκλιση, αλλά και χρήση προσωρινής ή μη μνήμης, φαινόμενο που θα έχει αντίκτυπο στον συνολικό χρόνο που θα απαιτηθεί για την εκτέλεση της προσομοίωσης. Αυτό γίνεται εμφανές και από την καθυστέρηση (lag) κατά την εκκίνηση των προγραμμάτων προσομοίωσης πέραν του συνηθισμένου, λόγω της μικρότερης

ταχύτητας επικοινωνίας των απομακρυσμένων μονάδων κεντρικής επεξεργασίας σε σχέση με την ταχύτητα εντός του κάθε ηλεκτρονικού υπολογιστή. Για το λόγο αυτό κάνουμε τη βέλτιστη επιλογή για το χρονισμό του συστήματός μας με εξυπηρετητή NTP Synchronize with NTP Server". Στο πεδίο με όνομα "NTP Server Address" συμπληρώνουμε την διεύθυνση του εξηπηρετητη NTP, που στην περίπτωσή μας είναι ntp.teicrete.gr ή gr.rool.ntp.gr και πατάμε "Synchronize now". Έπειτα, αφού έχουμε επιλεγμένες τις επιλογές ""Run NTP as deamon" και "Save NTP Configuration" μπορούμε να προχωρήσουμε στο επόμενο βήμα πατώντας "Accept". Προεπιλεγμένο γραφικό περιβάλλον είναι το KDE που μαζί με το GNOME είναι τα πιο λειτουργικά και δοκιμασμένα περιβάλλοντα στις διανομές LINUX.

Ακολουθεί η επιλογή κατανομής του χώρου του σκληρού δίσκου – Partitioning. Φτάνοντας σε αυτό το σημείο, παρατηρούμε την πρόταση του λογισμικού εγκατάστασης του λειτουργικού συστήματος για την κατάτμηση του σκληρού δίσκου. Αν αυτό που προτάθηκε δεν είναι σύμφωνο με τον σχεδιασμό που έχει αποφασιστεί, πατώντας το κουμπί Edit Partition Setup ανοίγει η οθόνη όπου υπάρχει η δυνατότητα αλλαγής του προτεινόμενου διαμερισμού του σκληρού δίσκου, Εικόνα 33.



Εικόνα 33. Διαχειριστής εγκατάστασης – επιλογές συστήματος αρχείων.

Έχει προ-αποφασιστεί ότι οι κατατμήσεις στους οποίους θα χωριστεί ο σκληρός δίσκος είναι : 9GB SWAP , 100 GB στον ριζικό τόμο (/), και 391 GB περίπου για τον τόμο των χρηστών (/home), για ένα σκληρό δίσκο 500GB. Έτσι, θα υπάρχει αρκετός αποθηκευτικός χώρος για προγράμματα, υπεραρκετή μνήμη SWAP για να μην καταρρεύσει το σύστημα από έλλειψη μνήμης αν καταναλωθεί όλη η μνήμη RAM και ο χώρος για τον χρήστη για τις ρυθμίσεις και τα αρχεία του.

Για να πραγματοποιηθούν τα παραπάνω, επιλέγουμε το Hard Disks στην περιοχή System View. Επόμενη επιλογή είναι ο σκληρός δίσκος με την οποία εμφανίζεται η προτεινόμενη διαμέριση του σκληρού δίσκου. Η επόμενη επιλογή είναι να διαγράψουμε καθένα από τους τόμους αυτούς. Αφού διαγραφούν όλοι οι τόμοι, θα δημιουργήσουμε νέους, στο μέγεθος που έχει προαποφασιστεί πατώντας Add. Σε αυτό το σημείο το σύστημα ρωτάει τον χρήστη αν θέλει να δημιουργήσει πρωτεύον ή εκτεταμένο διαμέρισμα. Εφ' όσον δεν υπάρχει ανάγκη για αυξομείωση μεγέθους, επιλέγεται το μέγεθος του πρωτεύοντος διαμερίσματος. Η διαφορά μεταξύ τους είναι ότι εκτεταμένο διαμέρισμα μπορεί να περιέχει πολλούς τόμους σε αυτό, ενώ επιτρέπεται και η αυξομείωση του μεγέθους τους, ενώ στο πρωτεύον έχουμε ένα μόνο τόμο και συνολικά μέχρι τέσσερα πρωτεύοντα διαμερίσματα σε ένα δίσκο. Προχωρώντας υπάρχει η επιλογή του μεγέθους του διαμερίσματος που μπορεί να γίνει με δύο τρόπους. Ο πρώτος είναι με την απευθείας εγγραφή του μεγέθους σε MB ή σε GB. Ο δεύτερος είναι ο τομέας εκκίνησης και ο τομέας τερματισμού του διαμερίσματος. Και οι δύο είναι ισοδύναμοι, με την προτίμηση να πηγαίνει στον πρώτο, λόγω της μεγαλύτερης ευκολίας ορισμού του μεγέθους. Το επόμενο βήμα οδηγεί στον τύπο διαμόρφωσης του τόμου και το πού ο τόμος θα συνδεθεί στο σύστημα - mounting.

Ο πρώτος τόμος που δημιουργείται, είναι αυτός που θα χρησιμοποιηθεί από το λειτουργικό σύστημα σαν μνήμη SWAP. Η μνήμη SWAP είναι ο χώρος στον σκληρό δίσκο, ο οποίος προορίζεται να χρησιμοποιηθεί σαν επέκταση την φυσικής μνήμης του υπολογιστικού συστήματος με λειτουργικό LINUX. Ο τρόπος χρήσης της είναι απλός: όταν εξαντληθεί η φυσική μνήμη ή μνήμη RAM, τότε γίνεται χρήση του τόμου SWAP αυξάνοντας κατά πολύ την μνήμη, ώστε να μην καταρρεύσει το υπολογιστικό σύστημα. Αυτή η λειτουργία έχει όμως μεγάλο κόστος στην ταχύτητα, αφού η διαφορά ταχύτητας μεταξύ, μνήμης RAM και σκληρού δίσκου είναι πολύ μεγάλη.

Γνωρίζοντας τα παραπάνω, μπορούμε να προχωρήσουμε στην συμπλήρωση των χαρακτηριστικών του τόμου που θα δημιουργηθεί. Στο πρώτο πεδίο, Formating Options, έχουμε τις επιλογές να διαμορφώσουμε τον τόμο αυτό και σε ποιο σύστημα αρχείων και αν θα κρυπτογραφηθεί. Επιλέγεται η διαμόρφωση του τόμου ως SWAP. Το δεύτερο πεδίο Mounting Options, παρέχει τις επιλογές σύνδεσης του τόμου. Οι παρεχόμενες επιλογές είναι: η σύνδεση του

τόμου και σε ποιο σημείο του δέντρου του συστήματος αρχείων, και η μη σύνδεση του τόμου. Η επιλογή που θα γίνει είναι σύνδεση του τόμου που δημιουργείται σαν SWAP.

Παρομοίως για τα άλλα 2 διαμερίσματα που θα δημιουργήσουμε, θα γίνουν οι ίδιες ενέργειες. Μπορούμε σε γραφικό περιβάλλον να δούμε την διαμόρφωση του ριζικού τόμου σε σύστημα αρχείων ext4 καθώς και το σημείο σύνδεσης στο σύστημα αρχείων που είναι η ρίζα / καθώς και τον τόμο του συνδέεται στους χρήστες /home που και σε αυτόν έχει επιλεγεί να διαμορφωθεί με ext4 σύστημα αρχείων. Το ext4 είναι μια πολύ καλή επιλογή συστήματος αρχείων, γιατί είναι ο άμεσος απόγονος των συστημάτων αρχείων ext2 και ext3, με πολλά χρόνια εξέλιξης από την κοινότητα του ελεύθερου λογισμικού. Παρουσιάζουν ελάχιστα σφάλματα -bugs και το πλεονέκτημα άριστης επιλογή για συστήματα που μπορεί να τερματίσουν βίαια, π.χ.: διακοπή ρεύματος ή κατάρρευση του λειτουργικού συστήματος από έλλειψη πόρων. Το τελευταίο είναι ιδιαίτερα σημαντικό, καθώς επιδιώκεται το 100% της απόδοσης της συστοιχίας υπολογιστικών συστημάτων ενώ είναι αδύνατη η αποφυγή κατάρρευσης του λειτουργικού συστήματος, ειδικά σε αρχικά στάδια. Σε τέτοια περίπτωση το σύστημα κατά την εκκίνησή του εκτελεί την εφαρμογή fsck, η οποία διορθώνει αυτόματα τα σφάλματα και στέλνει τα διορθωμένα αρχεία στον φάκελο lost and found. Μόλις ολοκληρωθεί η δημιουργία των τόμων, βλέπουμε την επισκόπηση για το πως θα διαμερισθεί ο σκληρός δίσκος του συστήματος. Αν βεβαιωθούμε ότι όλα είναι σωστά, σύμφωνα με τον αρχικό σχεδιασμό, αποδεχόμαστε τα παραπάνω πιέζοντας Accept.

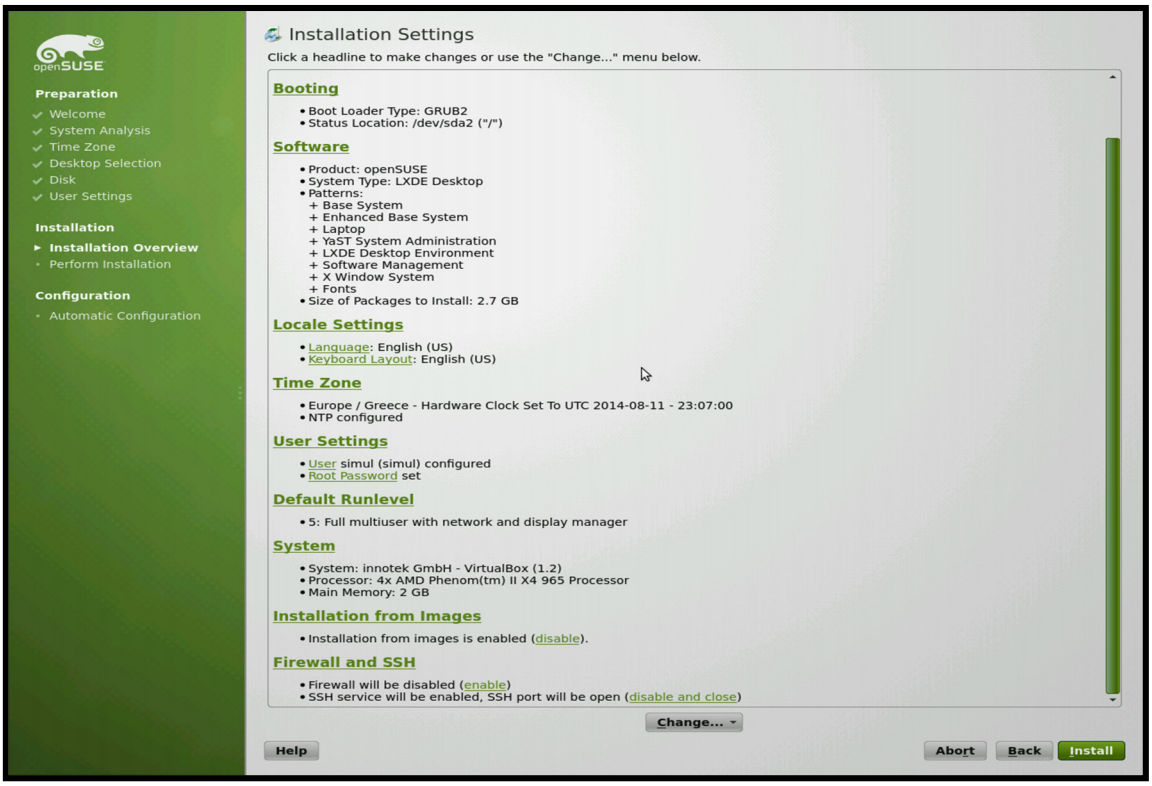
Στη συνέχεια καθορίζεται η διαχείριση και οι χρήστες του συστήματος Εξ' ορισμού ο διαχειριστής σε κάθε λειτουργικό σύστημα LINUX ή UNIX έχει το όνομα root και το μόνο που μας επιτρέπεται είναι να ορίσουμε το κωδικό του. Συνεπώς, αρχικά πρέπει να ορίσουμε τους ή τον χρήστη του συστήματος. Αυτός είναι ο simul και κατόπιν ορίζουμε το κωδικό του χρήστη αυτού. Παράλληλα, υπάρχει η επιλογή ο κωδικός αυτός να είναι ο κωδικός του χρήστη root, που το από-επιλέγουμε, ενώ επιλέγουμε ο χρήστης simul να κάνει αυτόματη είσοδο στο λειτουργικό σύστημα. Αυτό μας βοηθάει κατά την απομακρυσμένη πρόσβαση, γιατί έτσι θα υπάρχει προ-ενεργοποιημένο το software που απαιτείται για να γίνει αυτή. Ακολούθως εμφανίζεται η οθόνη της περίληψης εγκατάστασης λειτουργικού συστήματος.

Στην Εικόνα 34, μπορούμε να δούμε ότι μας παρέχεται η δυνατότητα να επέμβουμε σε αρκετές παραμέτρους του λειτουργικού συστήματος. Η πρώτη ενέργεια που θα πρέπει να εκτελεστεί είναι να ελέγξουμε ότι η υπηρεσία και η πόρτα της υπηρεσίας SSH είναι ενεργοποιημένη και ανοικτή αντίστοιχα. Αυτό είναι αναγκαίο για να μπορεί να υπάρχει απομακρυσμένος έλεγχος του συστήματος από την πρώτη στιγμή χωρίς να απαιτείται επί τόπου παρέμβαση. Πρέπει να αποφευχθεί να απενεργοποιηθεί το firewall, ώστε να μην είναι μπλοκαρισμένες οι θύρες οι οποίες απαιτούνται για την λειτουργία του λογισμικού που θα εγκατασταθεί. Φυσικά εάν γίνει απαραίτητη η χρήση firewall, τότε η ενεργοποίησή του μέσα από το λειτουργικό σύστημα δεν είναι καθόλου δύσκολη.

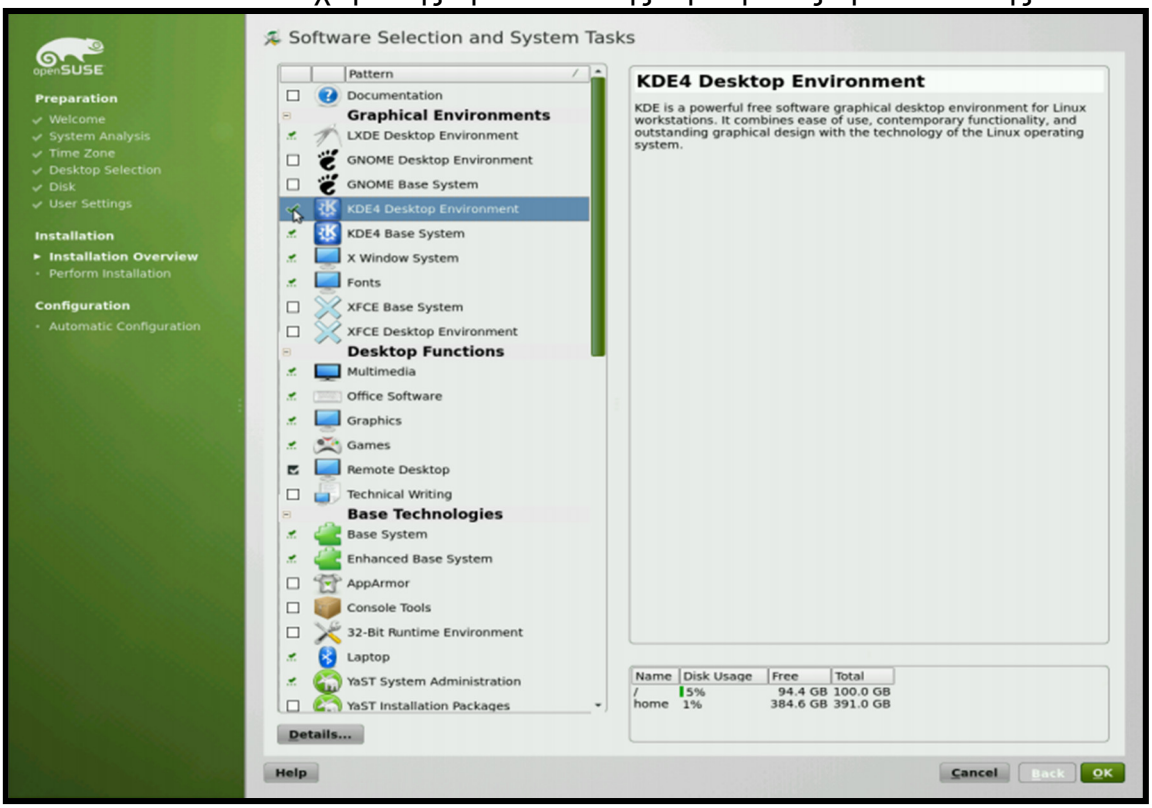
Άλλο ένα σημείο στο οποίο θα έπρεπε να εστιάσουμε την προσοχή είναι οι ρυθμίσεις του διαχειριστή εκκίνησης ή boot-loader. Κατά προτίμηση θα πρέπει να εγκατασταθεί ο διαχειριστής εκκίνησης GRUB2, ο οποίος είναι αυτός με την μεγαλύτερη υποστήριξη από την κοινότητα του ελεύθερου λογισμικού και είναι εύκολα παραμετροποιήσιμος. Επίσης το σημείο εγκατάστασης του είναι συνήθως η ρίζα του συστήματος αρχείων /, αφού έτσι υπάρχει πολύ μικρή πιθανότητα αποτυχίας. Εναλλακτικά μπορεί να εγκατασταθεί στο MBR - Main Boot Record του σκληρού δίσκου, αλλά με μεγάλη πιθανότητα να προκληθεί κατάρρευση αυτού και απώλεια όλων των δεδομένων που έχουν αποθηκευτεί. Παρατηρώντας την περίληψη εγκατάστασης λειτουργικού συστήματος, βλέπουμε την επιλογή Installation from Images. Αν η επιλογή αυτή είναι ενεργοποιημένη τότε θα ελαττωθεί ο χρόνος εγκατάστασης του λειτουργικού συστήματος, αλλά πέραν αυτού δεν έχει καμία άλλη επίπτωση στις επιδόσεις του υπό εγκατάσταση συστήματος.

Αν όλες οι παράμετροι του υπό εγκατάσταση λειτουργικού συστήματος είναι σωστές, τότε το μόνο που χρειάζεται να γίνει είναι η εγκατάσταση του επιθυμητού software. Πατώντας Change και κατόπιν Software οδηγούμαστε στην οθόνη που βλέπουμε στην Εικόνα 35. Σε αυτή βλέπουμε την επισκόπηση των πακέτων software που απαρτίζει το λειτουργικό σύστημα, τα περιφερειακά του προγράμματα και τα προγράμματα που έχουν σκοπό την εξυπηρέτηση των αναγκών του χρήστη. Αυτά είναι διαθέσιμα άμεσα στο ψηφιακό μέσο εγκατάστασης του λειτουργικού συστήματος. Για τη διευκόλυνση μας αλλά και για μεγαλύτερη ευελιξία επιλέγουμε να εγκατασταθεί επιπλέον το γραφικό περιβάλλον

KDE. Πατώντας την επιλογή Details οδηγούμαστε στην οθόνη χειροκίνητης αναζήτησης πακέτων λογισμικού που παρέχει το SUSE.

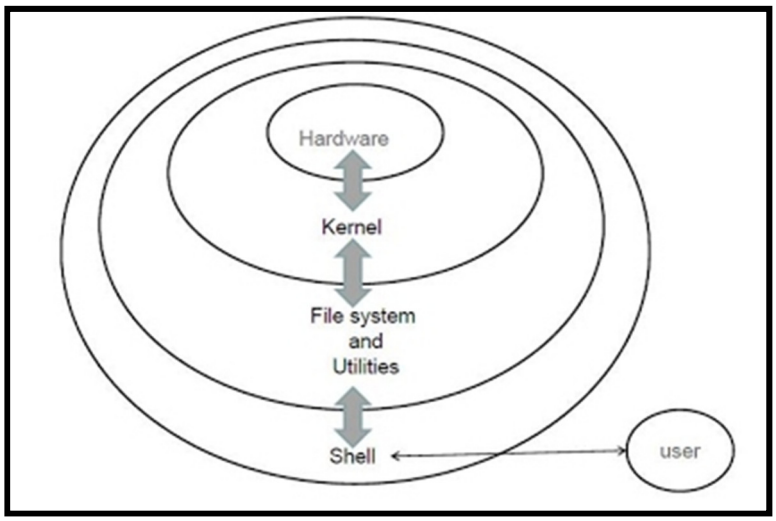


Εικόνα 34. Διαχειριστής εγκατάστασης – ρυθμίσεις εγκατάστασης .



Εικόνα 35. Διαχειριστής εγκατάστασης – Επιλογή πακέτων συστήματος.

Το πρώτο και ένα από τα βασικότερα εργαλεία είναι ο πηγαίος κώδικας του πυρήνα του λειτουργικού συστήματος - kernel-source. Το Kernel source είναι απαραίτητο για την άμεση επικοινωνία λογισμικού με τον πυρήνα του λειτουργικού συστήματος. Για παράδειγμα οι οδηγοί συσκευών -drivers, πρέπει να μεταφράσουν τον κώδικά τους αλλά και τον κώδικα του λειτουργικού συστήματος με τρόπο τέτοιο ώστε να υπάρχει απόλυτη επικοινωνία και διασύνδεση. Ένα σύγχρονο λειτουργικό σύστημα αποτελείται από πολλά επίπεδα και ενδιάμεσα σε αυτά τα επίπεδα έχουμε τα κελύφη, τα οποία είναι τα σημεία στα οποία επικοινωνούν οι εφαρμογές με το λειτουργικό [20]. Αυτή η δομή παρουσιάζεται στην Εικόνα 36.

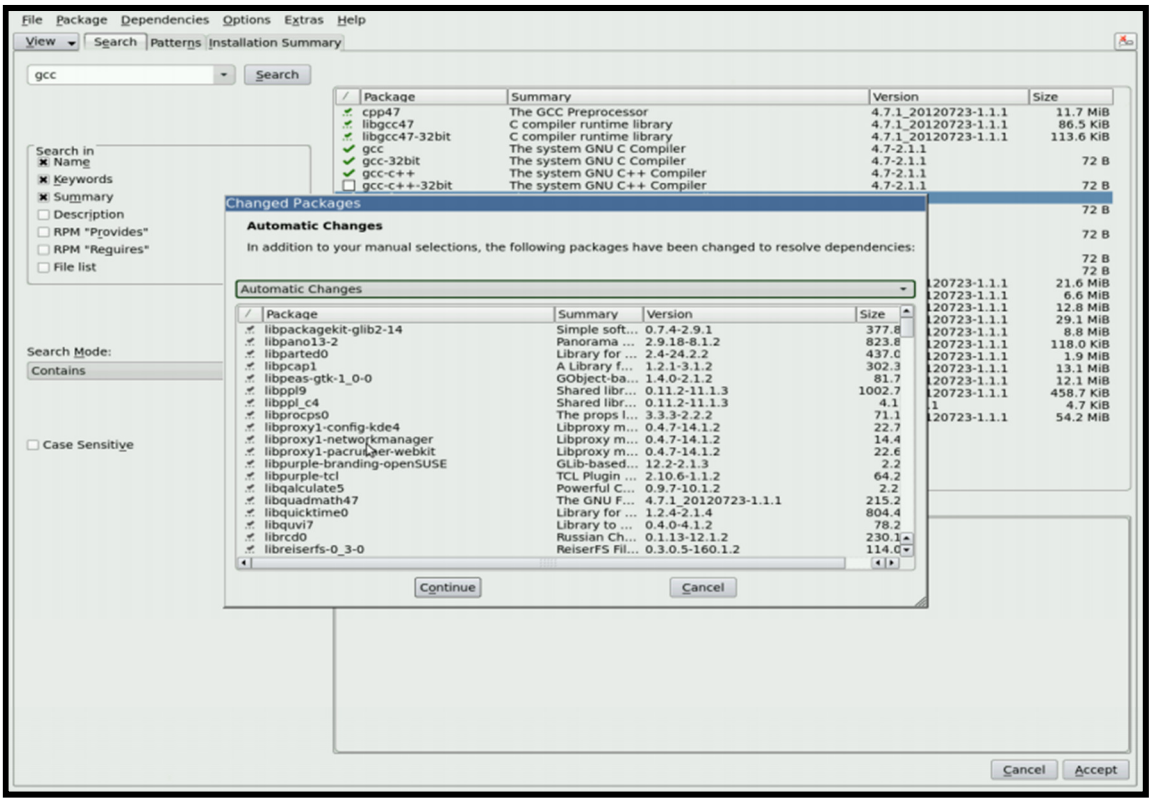


Εικόνα 36. Κελύφη LINUX

Το σχήμα της Εικόνα 36, δείχνει ότι ο πυρήνας του λειτουργικού είναι το άμεσα υπεύθυνο τμήμα για την επικοινωνία του υλικού με το software όπου βρίσκονται τα κελύφη, τα οποία είναι υπεύθυνα για την επικοινωνία του κάθε στρώματος. Αυτά είναι διάφανα στον χρήστη ή στον προγραμματιστή, εκτός από αυτά που είναι υπεύθυνα για την αλληλεπίδραση με τον χρήστη.

Σημαντικά πακέτα λογισμικού που πρέπει να εγκαταστήσουμε ώστε να είναι δυνατό το μετέπειτα compillation λογισμικού (μετάφραση) από το λειτουργικό σύστημα, είναι το πακέτο make, που έχει σκοπό να συνδέει τον μεταφραστή με το λειτουργικό σύστημα. Το ρόλο του compiler θα παίξει το πακέτο gcc με συνδυασμό με τις βιβλιοθήκες – γλώσσες προγραμματισμού που θα επιλέξουμε. Οι βιβλιοθήκες αυτές είναι οι C++ και η fortran, που κυρίως στην δεύτερη είναι γραμμένα τα περισσότερα προγράμματα υπολογιστικών μαθηματικών και

προσομοίωσης, ενώ η βιβλιοθήκη της γλώσσας C είναι εξ' ορισμού μέρος του λειτουργικού συστήματος.



Εικόνα 37. Διαχειριστής εγκατάστασης – επίλυση εξαρτήσεων επιπλέων πακέτων συστήματος

στην Εικόνα 37. Δεν πρέπει να ξεχνάμε ότι τα λειτουργικά συστήματα τύπου UNIX δεν είναι μονολιθικά όπως είναι τα WINDOWS αλλά modular. Κάθε πακέτο που εγκαθίσταται έχει την ανάγκη από συγκεκριμένα κομμάτια για να προσαρμοστεί στο παζλ του λειτουργικού συστήματος, τα οποία κατά την εγκατάστασή του απαιτεί την ύπαρξή τους. Πιέζοντας Install περνάμε στην τελική πια διαδικασία της εγκατάστασης του λειτουργικού συστήματος.Η διαδικασία είναι πλήρως αυτοματοποιημένη. Αρχικά γίνεται προετοιμασία του σκληρού δίσκου, δημιουργία των διαμερισμάτων και των τόμων μέσα σε αυτά. Ο δίσκος κατατμείται αρχικά στα διαμερίσματα, μετά διαμορφώνονται οι τόμοι με το σύστημα αρχείων που έχουμε επιλέξει σε κάθε τόμο και τέλος συνδέεται ο καθένας από αυτού στο σημείο που έχει επιλεγεί σε προηγούμενα βήματα. Το τελευταίο βήμα, είναι η αυτόματη ρύθμιση του συστήματος, η σταθεροποίηση, δηλαδή η εγκατάσταση των

συσκευών, του δικτύου και τελευταίο ο διαχειριστής εκκίνησης -boot loader. Τέλος στην Εικόνα 38 παρουσιάζεται η επιφάνεια εργασίας του OpenSUSE.

Πλέον η Συστοιχία –Α αποτελείται από τα παρακάτω Node:
a) simulation1, b) simulation2, c) simulation3, d) simulation4, e) simulation5

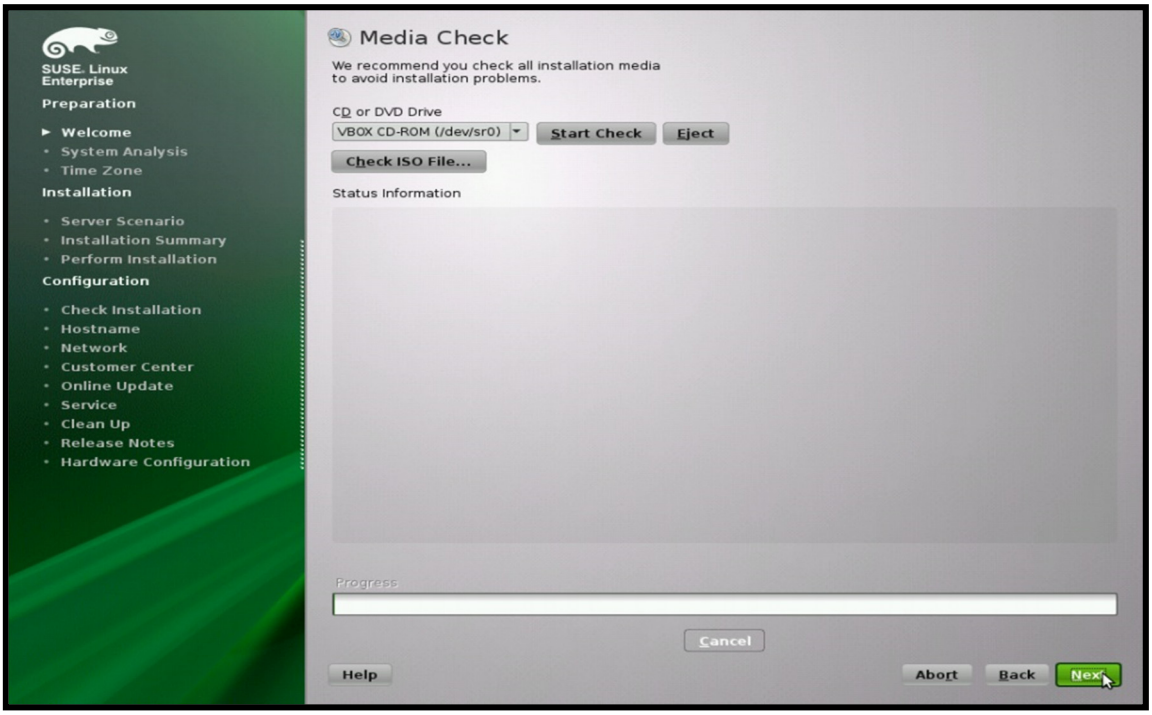
Εικόνα 38. Επιφάνεια εργασίας OpenSUSE.

3.3. Εγκατάσταση λειτουργικού συστήματος Συστοιχία -Β

Στη δεύτερη Συστοιχία, παρατηρούμε την απουσία οπτικού οδηγού. Αυτό σημαίνει ότι για μία ακόμη φορά θα κάνουμε χρήση του αντάπτορα USB to IDE – SATA και ενός DVD-RW όπως τον είδαμε στην Εικόνα 28 στην αρχή του κεφαλαίου. Σε αυτή τη Συστοιχία, σκοπός μας είναι η σταθερότητα και η μέγιστη συμβατότητα του λειτουργικού συστήματος, του software προσομοίωσης και του hardware, οπότε θα επιλέξουμε για λειτουργικό σύστημα το SLES 11 SP2.

Η διαδικασία είναι παρόμοια με την αντίστοιχη της Συστοιχίας Α. Θα παρατεθούν με εικόνες και θα σχολιαστούν μόνο τα βήματα στα οποία υπάρχουν διαφοροποιήσεις.

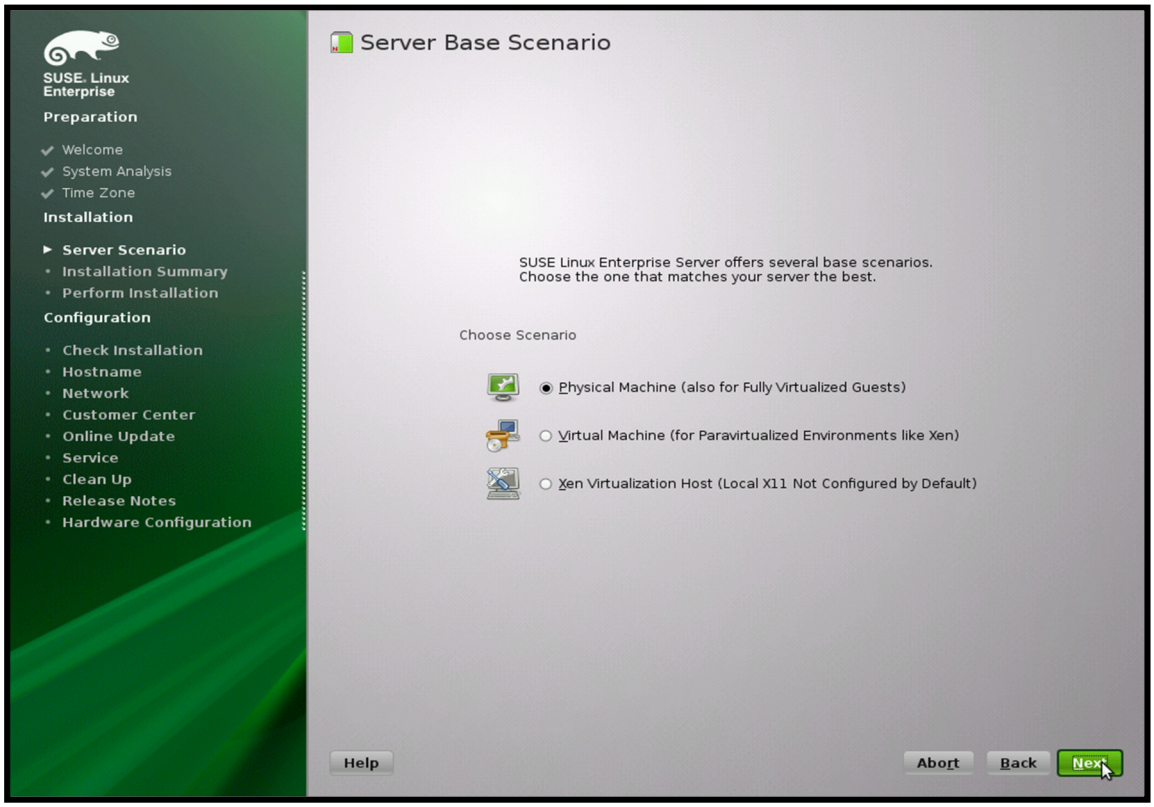
Αμέσως μετά την αποδοχή της συμφωνίας εγκατάστασης και την επιλογή της γλώσσα εγκατάστασης, εμφανίζεται η οθόνη ελέγχου ακεραιότητας του μέσου εγκατάστασης Εικόνα 39. Ο λόγος που γίνεται αυτό είναι ότι η συγκεκριμένη διανομή είναι εμπορική και απευθύνεται σε επαγγελματίες, οι οποίοι δεν θα είναι ικανοποιημένοι όταν μετά από 1 περίπου ώρα απασχόλησης διαπιστώσουν ότι το dvd εγκατάστασης έχει κάποιο σφάλμα και ότι θα πρέπει να τερματιστεί η εγκατάσταση. Παράλληλα, παρέχεται η δυνατότητα να ελεγχθούν και εικόνες CD ή DVD τύπου ISO ώστε να καλυφθούν όλες οι δυνατές περιπτώσεις μεθόδων εγκατάστασης.



Εικόνα 39. Διαχειριστής εγκατάστασης – έλεγχος μέσου εγκατάστασης.

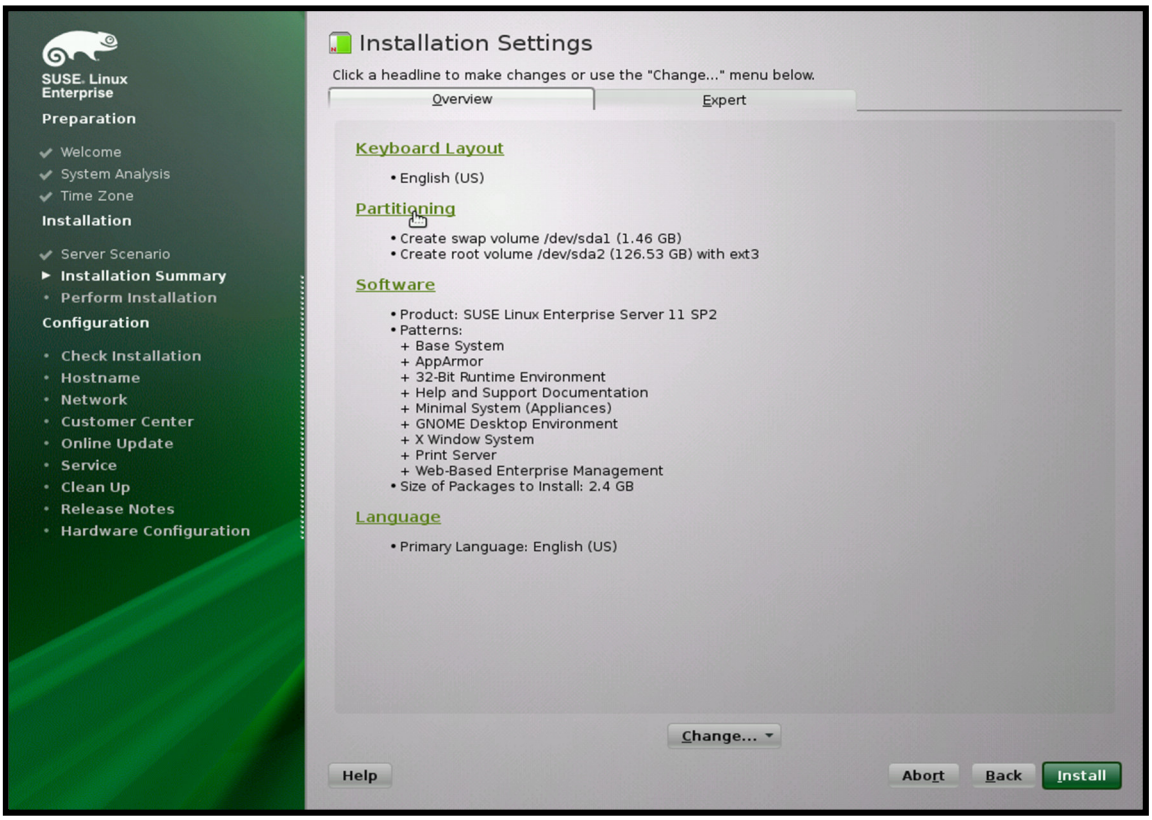
Ένα σύγχρονος εξυπηρετητής μπορεί να εγκατασταθεί σαν φυσικός ή σαν εικονικός (virtual). Αυτή η οθόνη Εικόνα 40, εμφανίζεται αμέσως μετά την ρύθμιση του χρόνου και του NTP server. Η επιλογή αυτή είναι σημαντική γιατί με την σύγχρονη υπολογιστική ισχύ μπορούμε να λειτουργούμε πολλαπλούς εξυπηρετητές στο ίδιο hardware. Στην περίπτωσή μας θα επιλεχθεί να

εγκατασταθεί σαν φυσικός server, αφού εμείς χρειαζόμαστε όλη την υπολογιστική ισχύ για την παράλληλη επεξεργασία. Στη συνέχεια πατώντας next περνάμε στην περίληψη εγκατάστασης Εικόνα 41.



Εικόνα 40. Διαχειριστής εγκατάστασης – σενάριο εξυπηρετητή.

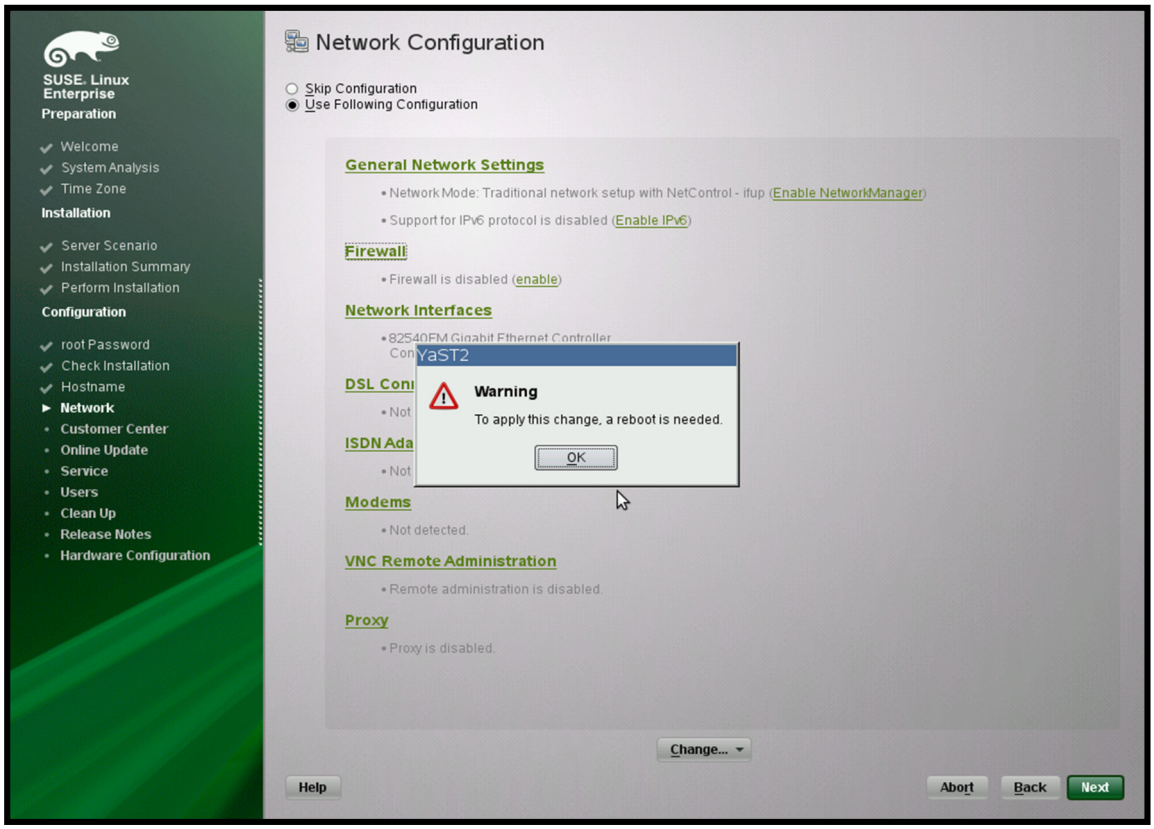
Άλλη μία διαφοροποίηση είναι ότι η κατάτμηση και διαμόρφωση συστήματος αρχείων γίνεται στις ρυθμίσεις εγκατάστασης. Όπως βλέπουμε στην οθόνη της περίληψη εγκατάστασης, Εικόνα 41, πατώντας την επιλογή partitioning οδηγούμαστε στην οθόνη προετοιμασία σκληρών δίσκων. Άλλη μια διαφοροποίηση είναι ότι θα δημιουργήσουμε μόνο swap και ριζικό / διαμέρισμα, με μέγεθος 18Gb και 110GB αντίστοιχα. Αυτό συμβαίνει γιατί θα κάνουμε χρήση της συστοιχίας των σκληρών δίσκων του fibre-chanell σαν κοινόχρηστο αποθηκευτικό χώρο, /home, μετά την ολοκλήρωση της εγκατάστασης του λειτουργικού συστήματος. Η υπόλοιπη διαδικασία της περίληψης εγκατάστασης είναι όμοια με της Συστοιχίας -A.



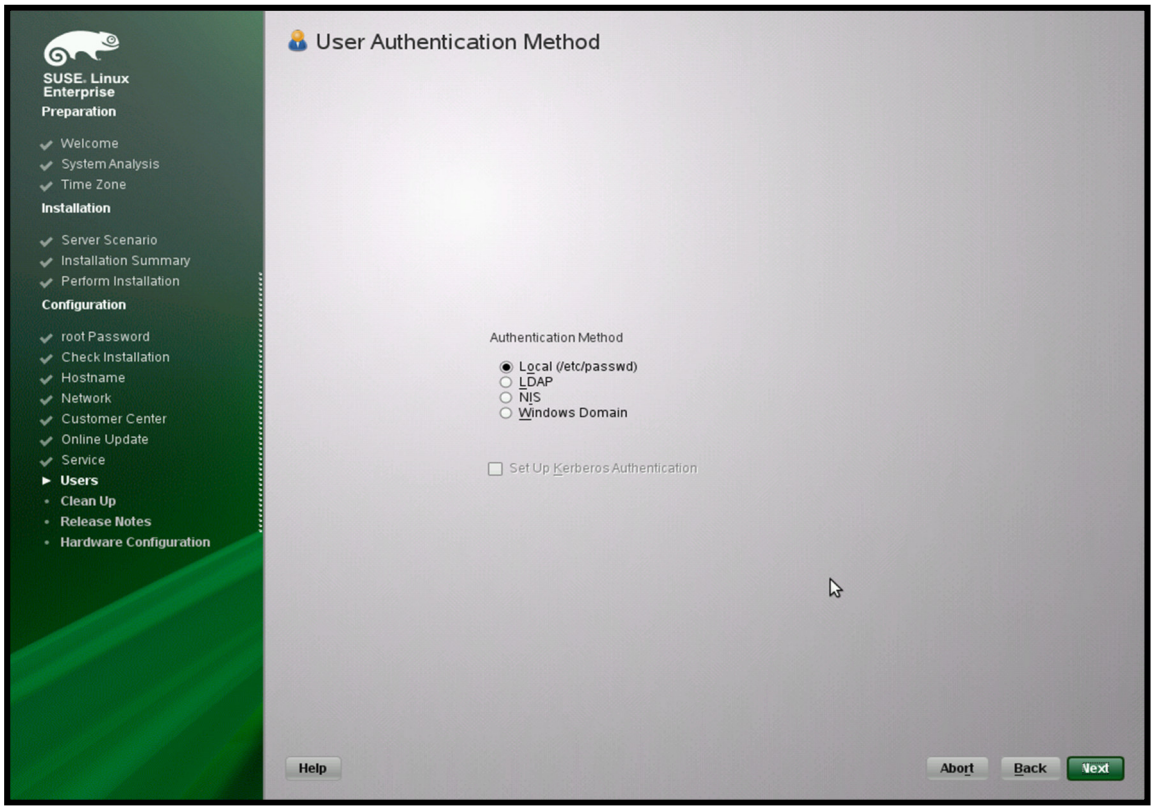
Εικόνα 41. Διαχειριστής εγκατάστασης – ρυθμίσεις εγκατάστασης.

Με την ολοκλήρωση την εγκατάστασης των πακέτων του λειτουργικού συστήματος, το σύστημα κάνει επανεκκίνηση. Εδώ ξεκινάει μία ακόμα διαφοροποίηση: η παραμετροποίηση του συστήματος γίνεται μετά την εγκατάσταση και είναι πάρα πολύ αναλυτική. Αρχικά, ζητάει τον κωδικό του υπερχρήστη (root).

Έπειτα, στην επόμενη οθόνη εισάγεται το όνομα του υπολογιστή καθώς και το όνομα του domain στο οποίο ανήκει. Ακολούθως, εισερχόμαστε στην ρύθμιση δικτύου, Εικόνα 42. Εκεί απενεργοποιείται το IPv6, ενεργοποιείται το SSH και απενεργοποιείται το firewall. Αφού πιεστεί το OK στην προτροπή για απενεργοποίηση του IPv6, μετά από επανεκκίνηση εμφανίζεται η οθόνη δοκιμής στο διαδίκτυο, την οποία θα παρακάμψουμε. Η επόμενη οθόνη, έχει την ρύθμιση των υπηρεσιών δικτύου, η οποία επίσης παρακάμπτεται.



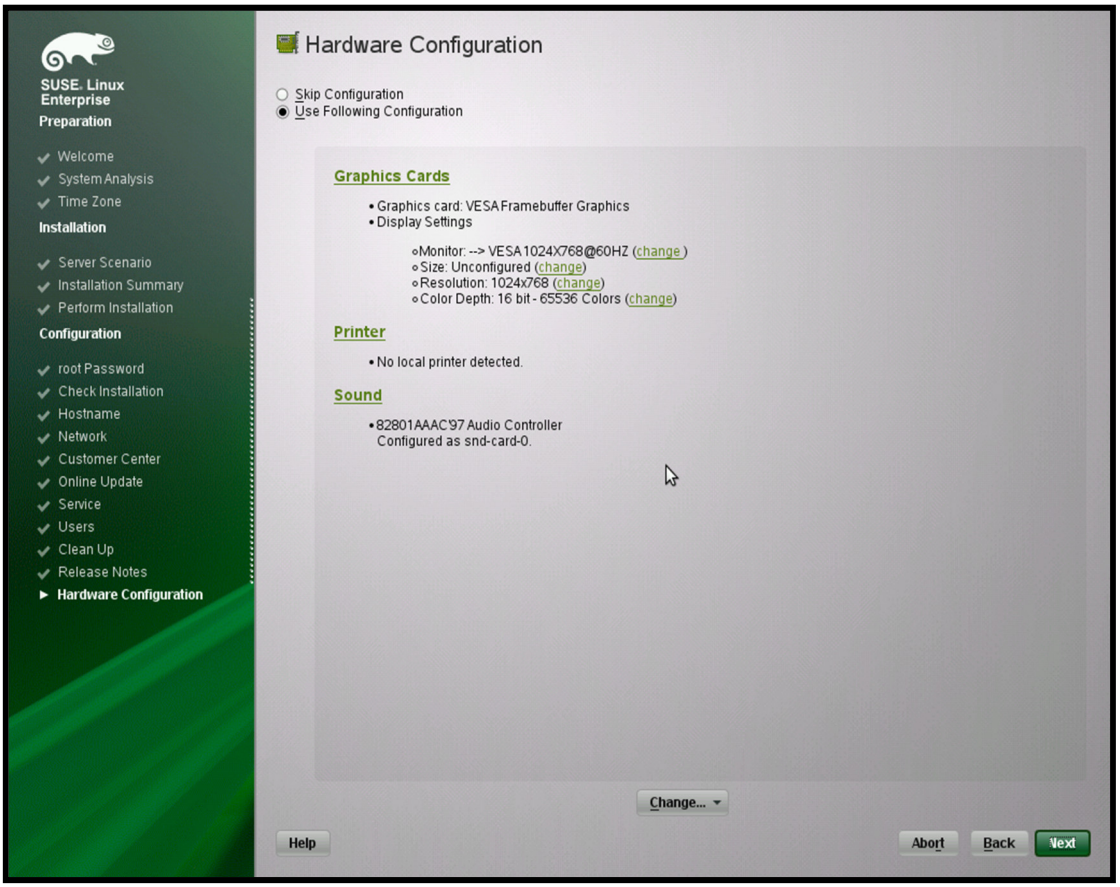
Εικόνα 42. Ρυθμίσεις λειτουργικού συστήματος – ρύθμιση δικτύου.



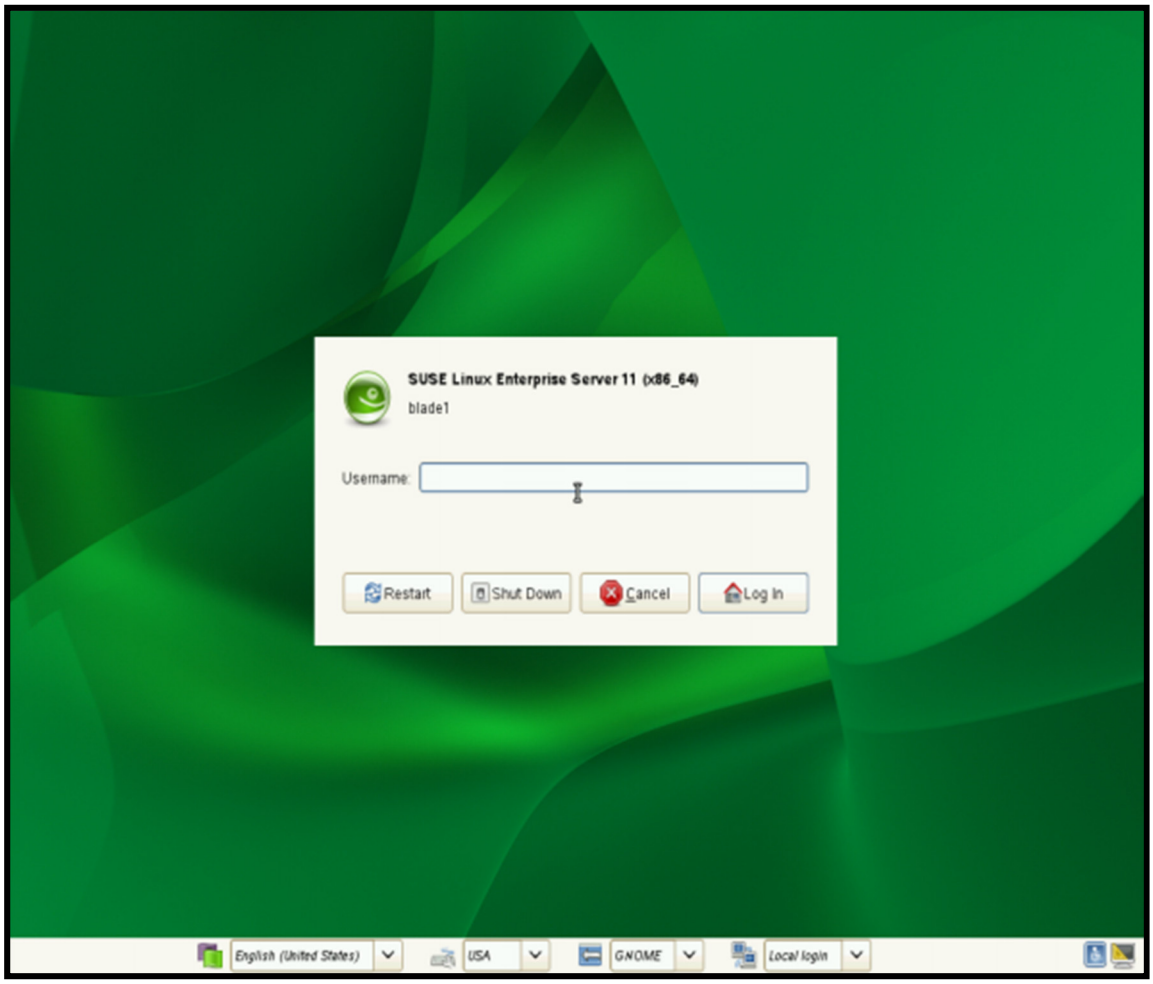
Εικόνα 43. Ρυθμίσεις λειτουργικού συστήματος – μέθοδος αυθεντικοποίησης.

Το επόμενο βήμα, Εικόνα 43, μας ζητάει την μέθοδο ορισμού των χρηστών. Θα επιλεγεί η τοπική μέθοδος. Έπειτα εισάγεται το όνομα χρήστη και τον κωδικό χρήστη. Αν ο κωδικός δεν πληρεί τις προϋποθέσεις, τότε εμφανίζεται ένα μήνυμα που το αναφέρει, όπου αποδεχόμαστε ή όχι την απλότητα του κωδικού πρόσβασης. Μόλις τελειώσουμε την εισαγωγή του χρήστη, περνάμε στην αποθήκευση και καταγραφή των πληροφοριών στην ρύθμιση του συστήματος, και αμέσως μετά, την εμφάνιση των σημειώσεων έκδοσης του λειτουργικού συστήματος.

Μετά περνάμε στην οθόνη ρύθμισης hardware, Εικόνα 44, περιλαμβάνει τη ρύθμιση της κάρτας γραφικών, της κάρτας ήχου και την εγκατάστασης και ρύθμιση των συνδεδεμένων εκτυπωτών. Το σύστημα είναι αυτοματοποιημένο και μόνο όταν δεν έχει γίνει η κατάλληλη επιλογή απαιτείται η επέμβαση του χρήστη. Τέλος, στην επόμενη οθόνη, αναφέρεται το πέρας της εγκατάστασης του λειτουργικού συστήματος και μετά εμφανίζεται η οθόνη σύνδεσης χρήστη, Εικόνα 45, που μας φανερώνει ότι όλα έγιναν σωστά και ότι το λειτουργικό σύστημα είναι έτοιμο προς χρήση.



Εικόνα 44. Ρυθμίσεις λειτουργικού συστήματος – ρύθμιση hardware.



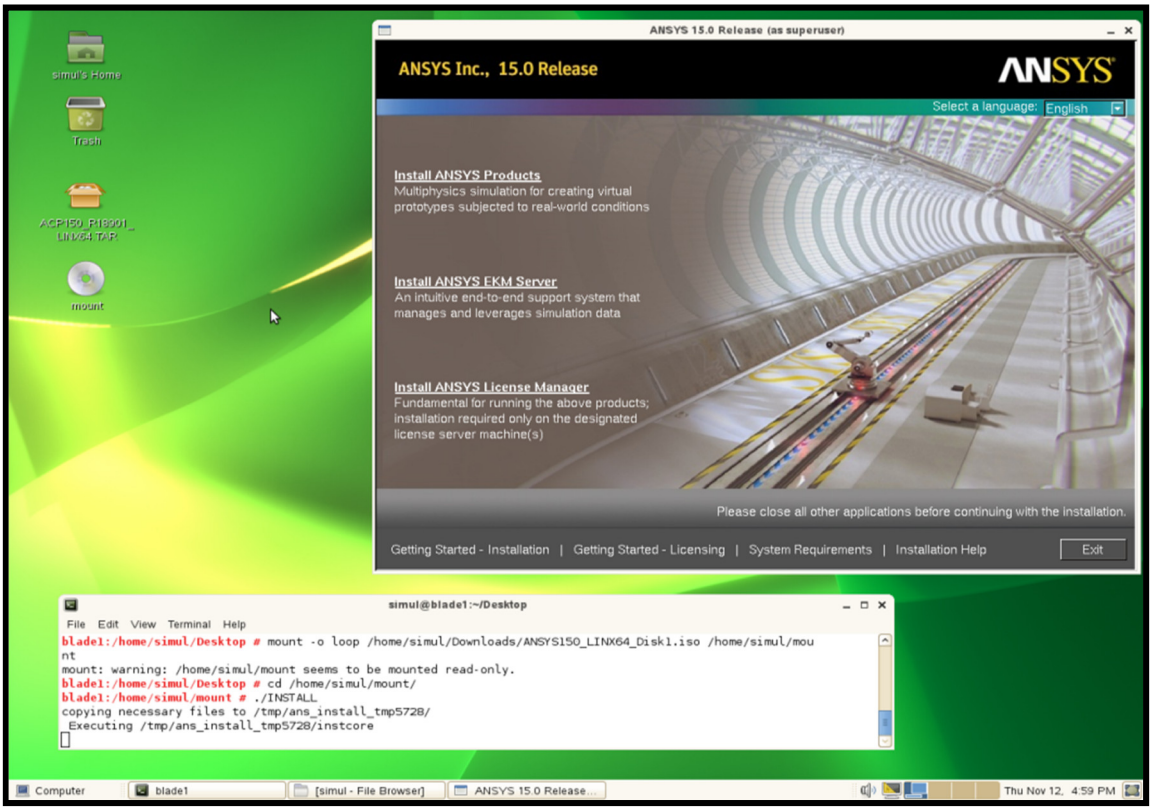
Εικόνα 45. Οθόνη σύνδεσης χρήστη.

Έτσι η Συστοιχία –Β αποτελείται από16 node: τα blade1, blade2, …, blade16

3.3 Εγκατάσταση Software Προσομοιώσεων Πεπερασμένων Στοιχείων

Στις δύο συνθέσεις που έχουμε στη διάθεσή μας, θα εγκατασταθούν 2 λογισμικά Προσομοιώσεων Πεπερασμένων Στοιχείων, το ANSYS και έπειτα το LS-DYNA. Το ANSYS αποτελείται, στην έκδοση για Linux, από 3 εικόνες δίσκων[21]. Μπορούμε είτε να τις μετατρέψουμε σε DVD είτε να εργαστούμε με αυτές, αφού στο Linux η χρήση εικόνων γίνετε εύκολα και χωρίς χρήση τρίτου λογισμικού. Έτσι, έχοντας κατεβάσει τις 3 εικόνες DVD από την ιστοσελίδα της ANSYS, προχωράμε στην εγκατάσταση.

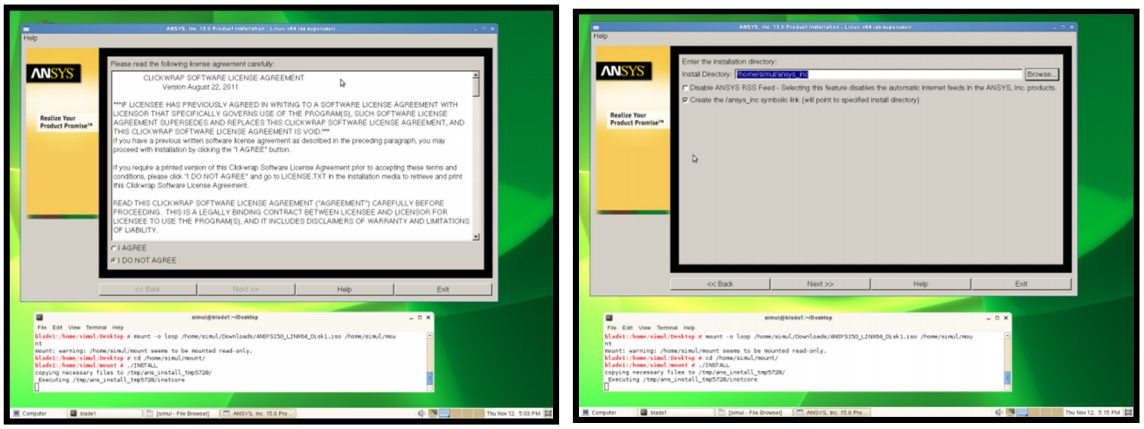
Οι εικόνες αποθηκεύονται στον φάκελο /home/simul/Downloads. Παράλληλα, δημιουργούμε ένα φάκελο με όνομα mount στο /home/simul, όπου θα γίνει mount το .ISO αρχείο. Αυτή είναι μια ενέργεια στην οποία συσκευές ή εικόνες συσκευών γίνονται ορατές προς χρήση από το λειτουργικό σύστημα.



Εικόνα 46. Γραφικό περιβάλλον εγκατάστασης ANSYS.

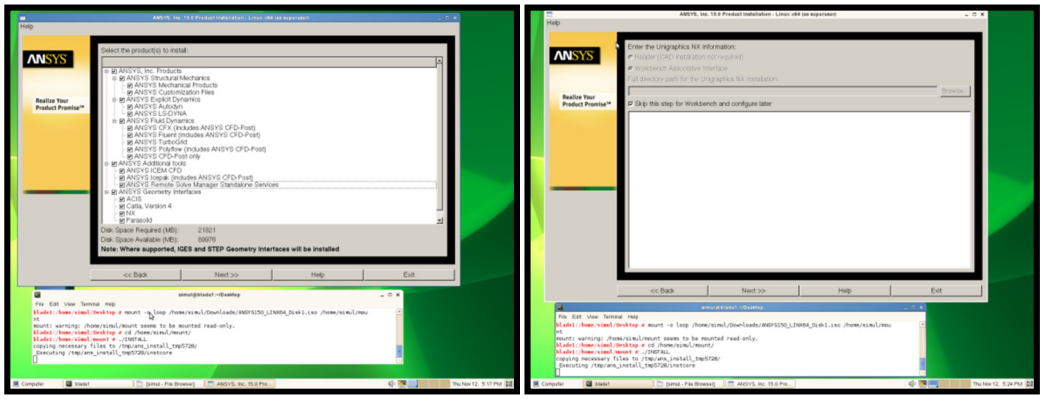
Αρχικά πληκτρολογούμε έχοντας δικαιώματα υπερχρήστη # mount -o loop /home/simul/Downloads/ANSYS150_LINX64_Disk1.iso /home/simul/mount. Αυτή

η εντολή λέει στο κέλυφος να κάνει mount την πρώτη Εικόνα, το DVD 1, στον φάκελο mount. Μας επιστρέφεται ένα μήνυμα ότι ο φάκελος είναι μόνο για ανάγνωση. Μπαίνουμε στο φάκελο mount και τρέχουμε το εκτελέσιμο αρχείο της εγκατάστασης με την εντολή # ./INSTALL. Τότε, εμφανίζεται το γραφικό σύστημα εγκατάστασης του ANSYS. Στην Εικόνα 46 μπορούμε να δούμε τις εντολές στο κέλυφος του λειτουργικού συστήματος, καθώς και την πρώτη οθόνη του γραφικού συστήματος εγκατάστασης.



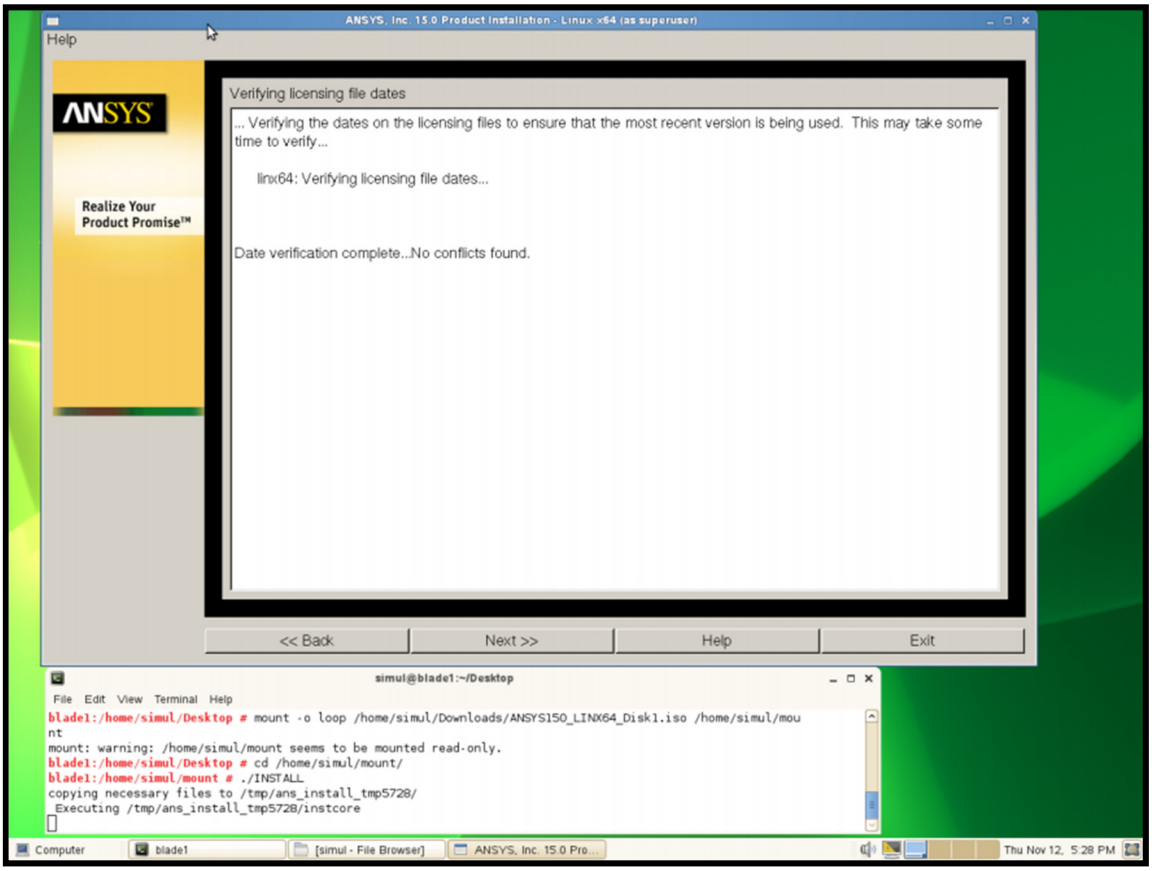
Εικόνα 47. Άδεια χρήσης – Φάκελος εγκατάστασης ASNYS.

Έχοντας αφήσει την γλώσσα εγκατάστασης στα Αγγλικά. πατάμε το κουμπί Install ANSYS Products. Τότε, εμφανίζεται μια νέα οθόνη, Εικόνα 47, στην οποία μπορούμε να δούμε την συμφωνία χρήσης, την οποία πρέπει να αποδεχτούμε για να προχωρήσει η διαδικασία. Πιέζουμε I AGREE και κατόπιν Next, για να περάσουμε στην επόμενη οθόνη.



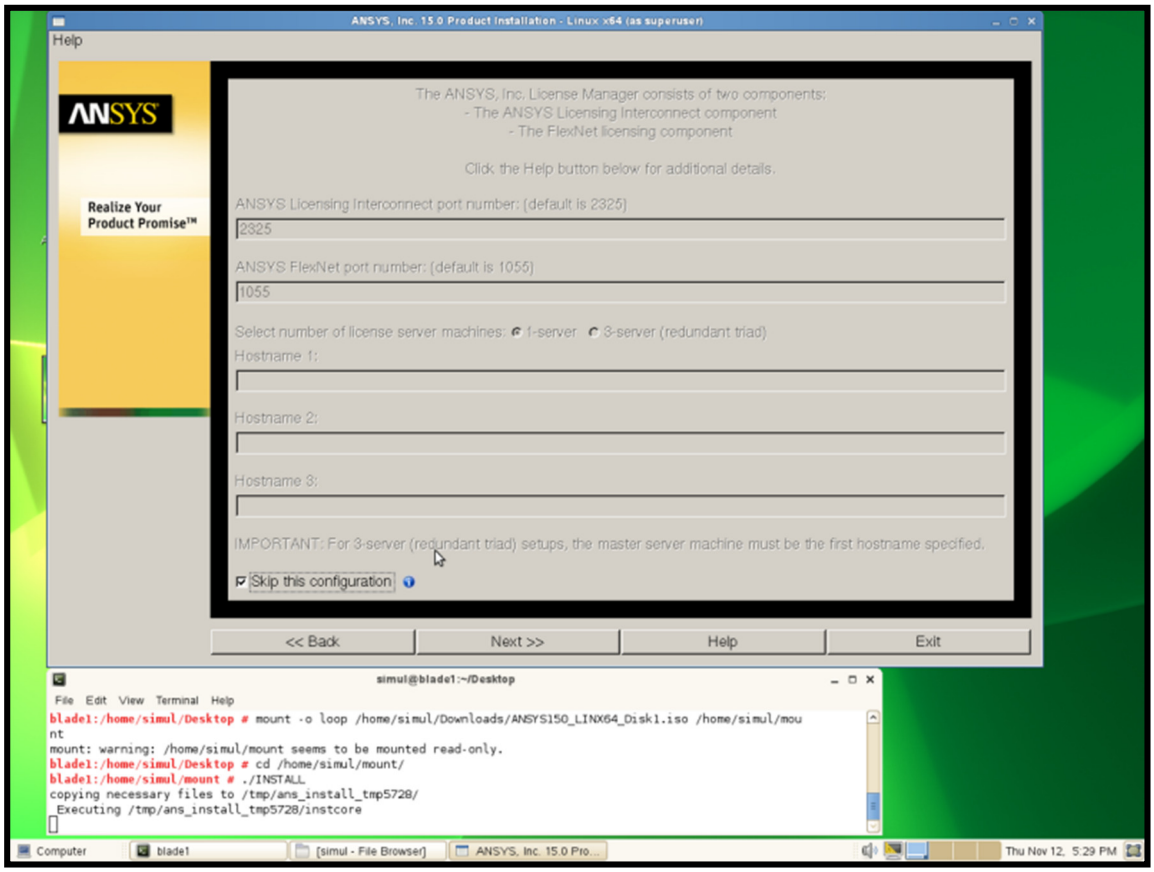
Εικόνα 48. Πακέτα εγκατάστασης ANSYS – Ρύθμιση Unigraphics NX.

Σε αυτή την οθόνη, Εικόνα 47, ορίζουμε το φάκελο στον οποίο θα το εγκαταστήσουμε. Έχει σαν προεπιλογή τον /usr/ansys_inc, ο οποίος στην περίπτωσή μας δεν μας εξυπηρετεί, λόγω της παράλληλης επεξεργασίας, γι' αυτό και θα επιλέξουμε τον φάκελο /home/simul/ansys_inc. Επίσης, προσέχουμε να είναι επιλεγμένη η εντολή για δημιουργία συντόμευσης στον φάκελο /ansys_inc και πιέζουμε Next. Επιλέγουμε όλα τα πακέτα εγκατάστασης, όπως φαίνεται στην Εικόνα 48, και πιέζοντας Next περνάμε στην επόμενη οθόνη, όπου εμφανίζεται η ρύθμιση του Unigraphics NX και παραλείπουμε πιέζοντας skip και μετά Next, Εικόνα 48.

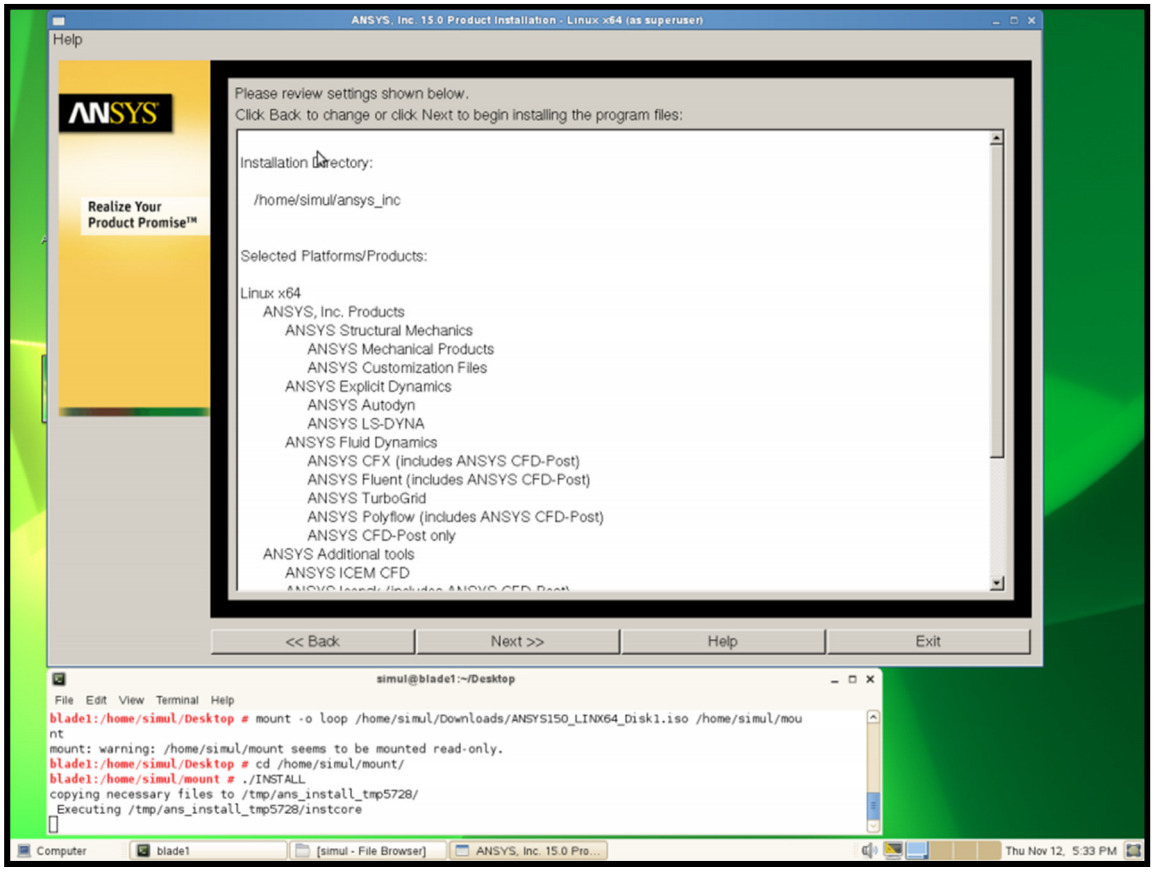


Εικόνα 49. Έλεγχος ημερομηνιών αδειών.

Ακολουθεί έλεγχος των ημερομηνιών στα αρχεία αδειών χρήσης, Εικόνα 49, και προχωράμε, πιέζοντας Next. Στη συνέχεια, εμφανίζεται η οθόνη εγκατάστασης του διαχειριστή αδειών, Εικόνα 50, που επίσης παραλείπουμε, γιατί είναι προτιμότερο να εγκατασταθεί μετά, εφ' όσον απαιτείται η χρήση του και δεν λειτουργήσει η χειροκίνητη εγκατάσταση της άδειας χρήσης. Έτσι πιέζουμε και skip και μετά Next.

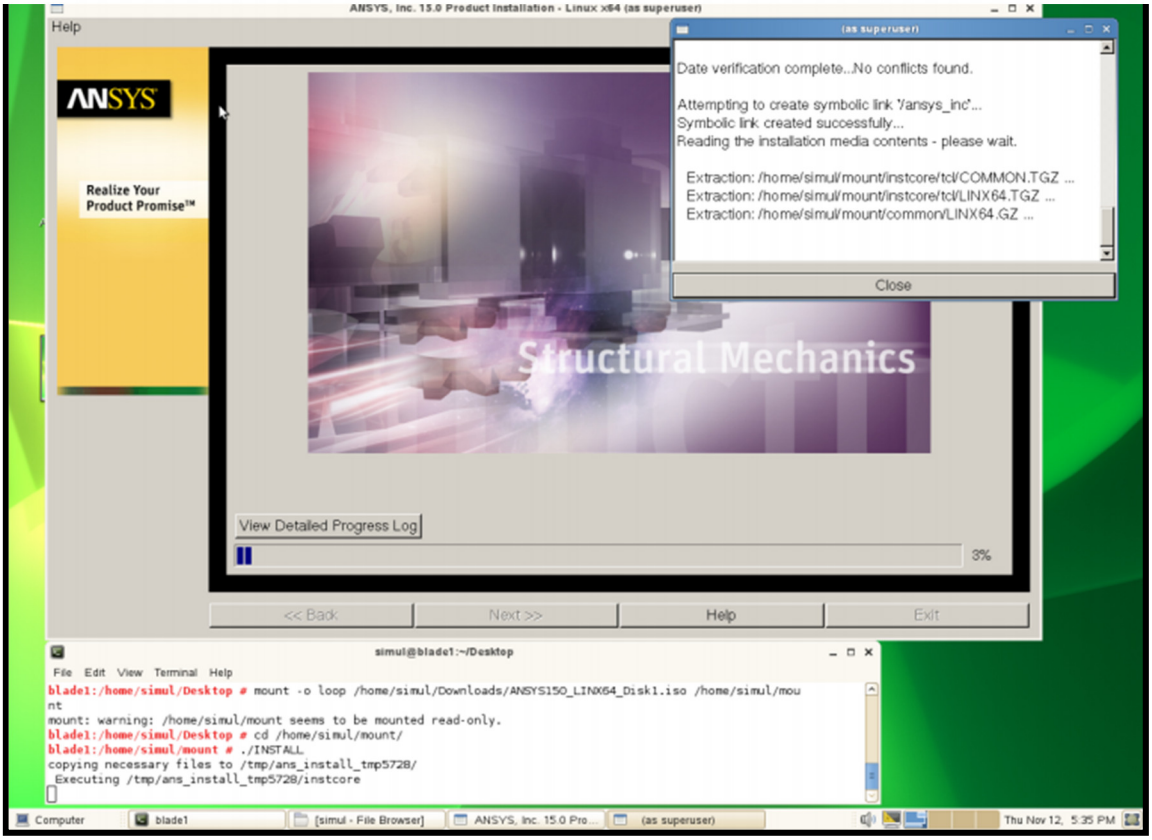


Εικόνα 50. Διαχειριστής αδειών ANSYS.



Εικόνα 51. Ανασκόπηση εγκατάστασης.

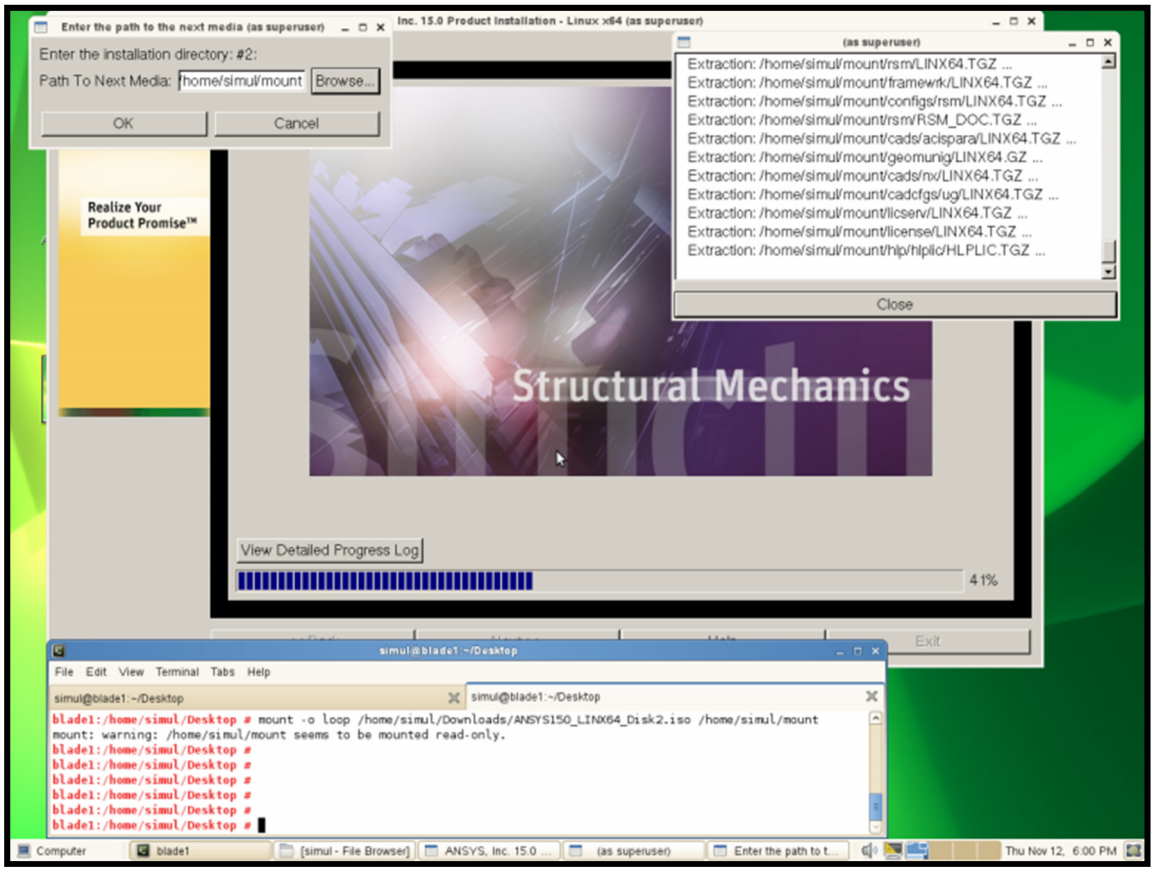
Περνώντας στην επόμενη οθόνη, Εικόνα 51, μας γίνεται ανασκόπηση της εγκατάστασης, πιέζουμε πάλι Next και προχωράμε στην εγκατάσταση.



Εικόνα 52. Επισκόπηση εγκατάστασης ANSYS.

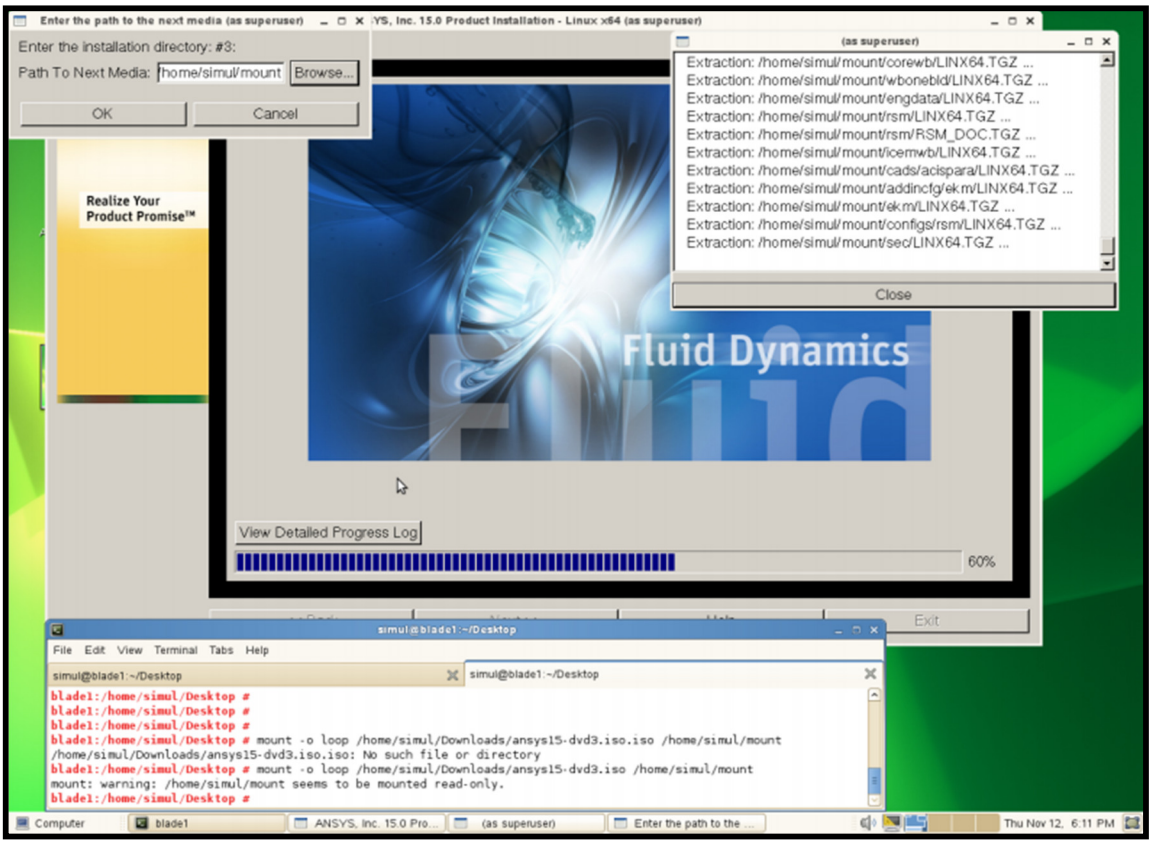
Στην οθόνη που εμφανίζεται, Εικόνα 52, βλέπουμε τα πακέτα που εγκαθίστανται και την εξέλιξη της εγκατάστασης. Όταν εμφανιστεί το μήνυμα για εισαγωγή του δεύτερου dvd, ανοίγουμε μία νέα κονσόλα, μπαίνουμε με δικαιώματα διαχειριστή και πληκτρολογούμε:

#mount -o loop /home/simul/Downloads/ANSYS150_LINX64_Disk2.iso/home/simul/mount.

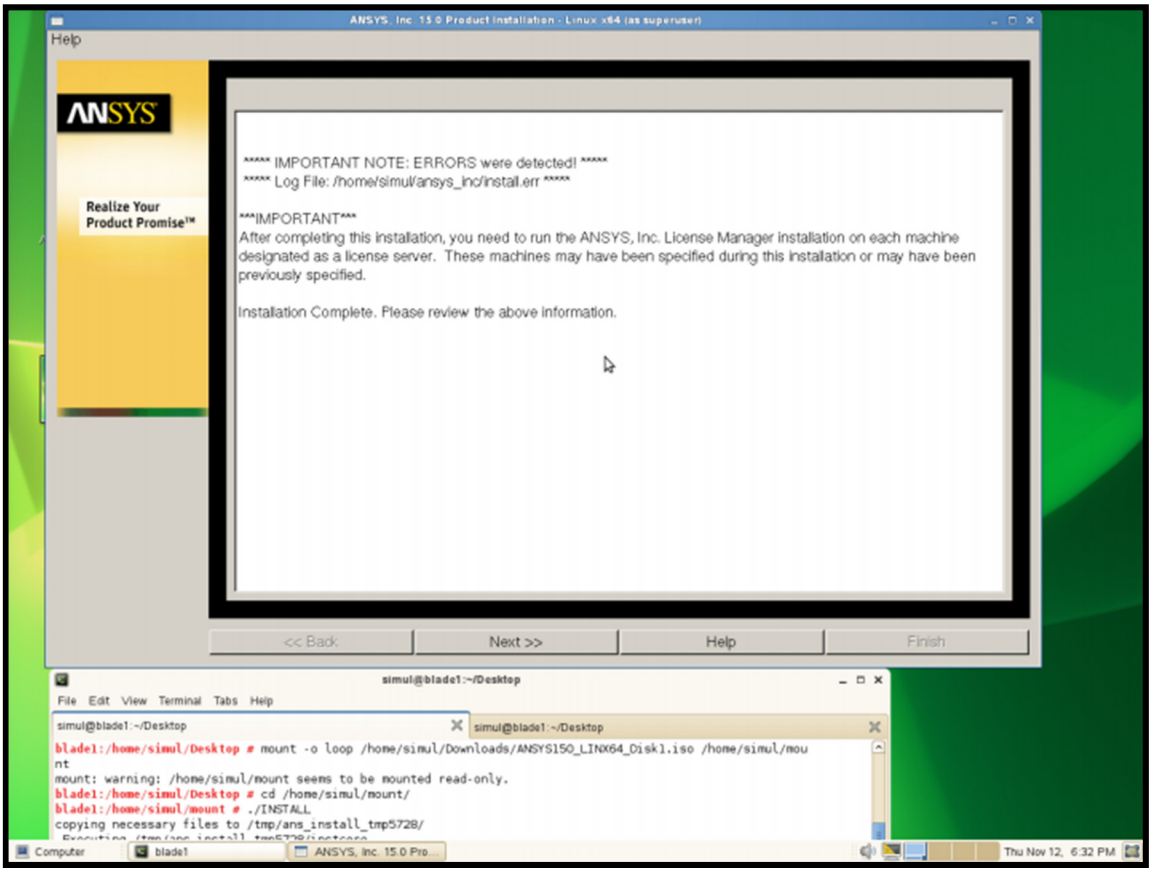


Εικόνα 53. Αίτημα εισαγωγής δεύτερου DVD.

Δίνουμε OK στο παράθυρο που ζητάει την διαδρομή του επόμενου μέσου, όπως φαίνεται στην Εικόνα 53. Η εγκατάσταση συνεχίζεται κανονικά, μέχρι που θα ζητηθεί το τρίτο DVD, Εικόνα 54, όπου και ενεργούμε παρομοίως και η εγκατάσταση ολοκληρώνεται. Εμφανίζεται το αρχείο καταγραφής, Εικόνα 55, όπου μας ενημερώνει για σφάλματα που τυχόν προέκυψαν και πατάμε πάλι Next.

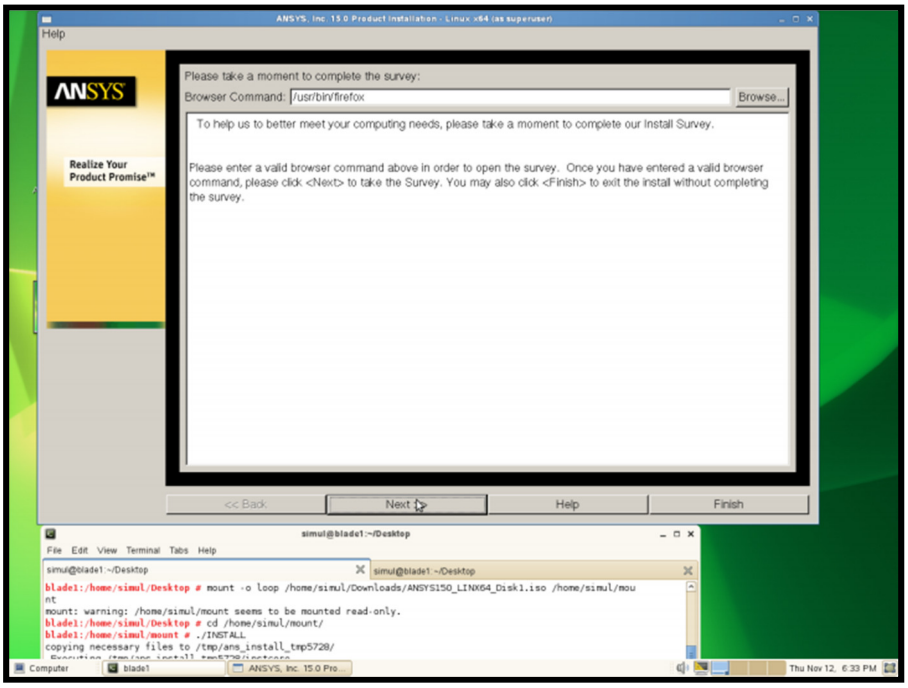


Εικόνα 54. Αίτημα εισαγωγής τρίτου DVD.



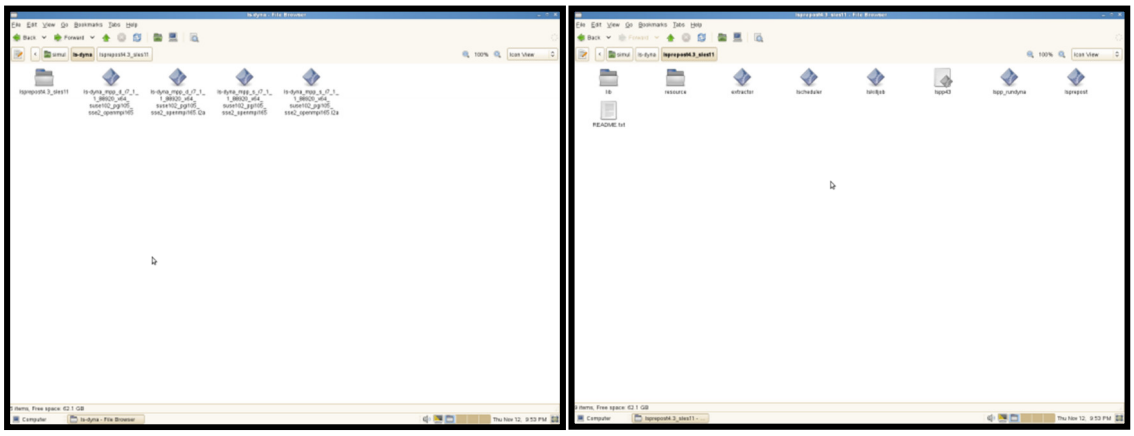
Εικόνα 55. Αρχείο καταγραφής σφαλμάτων.

Τέλος στην επόμενη οθόνη, Εικόνα 56, πατάμε Finish και ολοκληρώνουμε την εγκατάσταση.



Εικόνα 56. Τέλος εγκατάστασης.

Η εγκατάσταση του LS-DYNA είναι απλούστερη [22]. Αρκεί να αντιγραφούν τα εκτελέσιμα αρχεία του σε ένα φάκελο, όπως και του LS-PrePost, το οποίο χρειάζεται για την προετοιμασία του μοντέλου που θα προσομοιωθεί. Παράλληλα, θα απαιτηθεί software MPI, όπου επιλέγεται το OpenMPI.. Δημιουργούμε ένα φάκελο με το όνομα ls-dyna, και απλά αντιγράφουμε τα εκτελέσιμα του LS-DYNA και του LS-PrePost εκεί, όπως φαίνεται στην Εικόνα 57. Η εκτέλεση αυτών γίνεται απλά ανοίγοντας μια κονσόλα.



Εικόνα 57. Εκτελέσιμα αρχεία LS-DYNA.

Εκτός από το ANSYS και το LS-DYNA, στις Συνθέσεις Α και Β εγκαταστήσαμε 3 ακόμη προγράμματα προσομοίωσης μοντέλων πεπερασμένων στοιχείων. Τα προγράμματα αυτά είναι τα:

- Xoopic [23] μαζί με το XGrafix
- PLUTO [24]
- EPOCH [25]

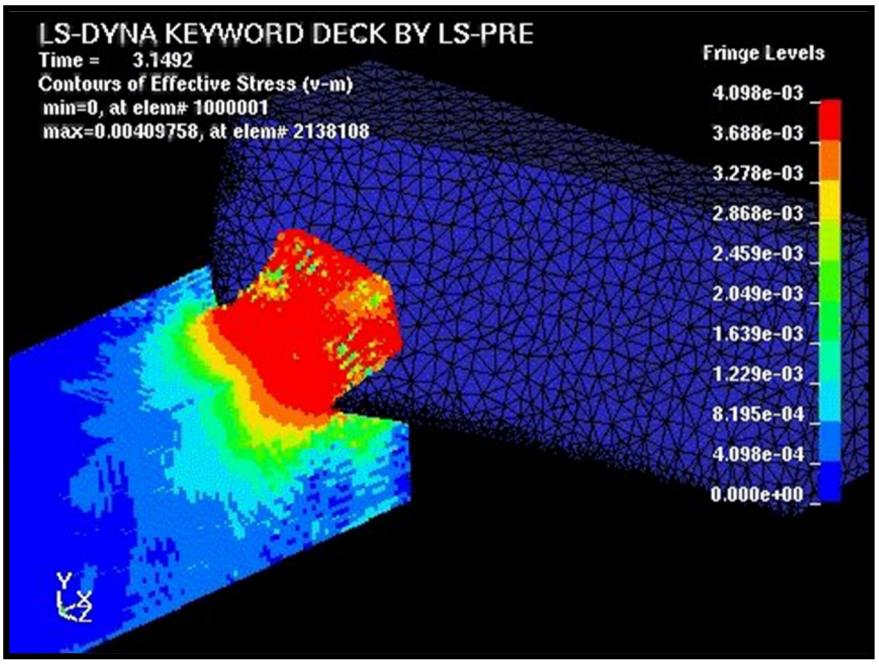
Αυτά τα 3 προγράμματα διεύρυναν το πεδίο των δυνατοτήτων προσομοιώσεων των συστημάτων. Είναι εξειδικευμένα για προσομοιώσεις σε προβλήματα μαγνητουδροδυναμικής αστροφυσικής -MHD, ο PLUTO και σε προβλήματα σωματιδιακής φυσικής -PIC, οι Xoopic και EPOCH. Τα 3 αυτά προγράμματα εγκαταστάθηκαν στην Συστοιχία-Α στο /simulation1/ και στην Συστοιχία-Β στο /home/simul/ με όνομα φακέλου το όνομα του προγράμματος.

ΚΕΦΑΛΑΙΟ 4. Προσομοιώσεις και αξιολόγηση

4.1. Η Μέθοδος Πεπερασμένων Στοιχείων & προσομοιώσεις στα cluster

Η έλευση των Η/Υ, επέτρεψε την ανάπτυξη και την υλοποίηση της θεωρίας της Μεθόδου των Πεπερασμένων Στοιχείων του Ritz, ο οποίος ανέπτυξε τις αρχές της, και του Galerkin, που εμβάθυνε την θεωρεία της Μεθόδου Πεπερασμένων Στοιχείων. Το όνομα «Πεπερασμένα Στοιχεία», χρησιμοποιήθηκε για πρώτη φορά από τον Clough καθηγητή του Berkeley το 1960, ενώ σημαντικό ρόλο στη θεμελίωση των Πεπερασμένων στοιχείων έπαιξε το βιβλίο του I. Αργύρη «Ενεργειακά Θεωρήματα και η Μέθοδος των Μητρώων» [26].

Ήταν εξαιρετικά δύσκολο να αναπτυχθεί αυτή η μέθοδος νωρίτερα, γιατί απαιτεί τη χρήση ισχυρών υπολογιστικών μηχανών. Αυτό είχε σαν αποτέλεσμα την ύπαρξη προβλημάτων τα οποία θεωρούταν άλυτα από την επιστημονική κοινότητα. Η ανάπτυξη των ηλεκτρονικών υπολογιστών και η ανάγκη για γρήγορη και ακριβή επίλυση προβλημάτων στην αεροναυπηγική, την δεκαετία του 1950, έδωσαν το έναυσμα στην ταχύτατη ανάπτυξη της Μεθόδου των πεπερασμένων Στοιχείων στα ερευνητικά κέντρα.



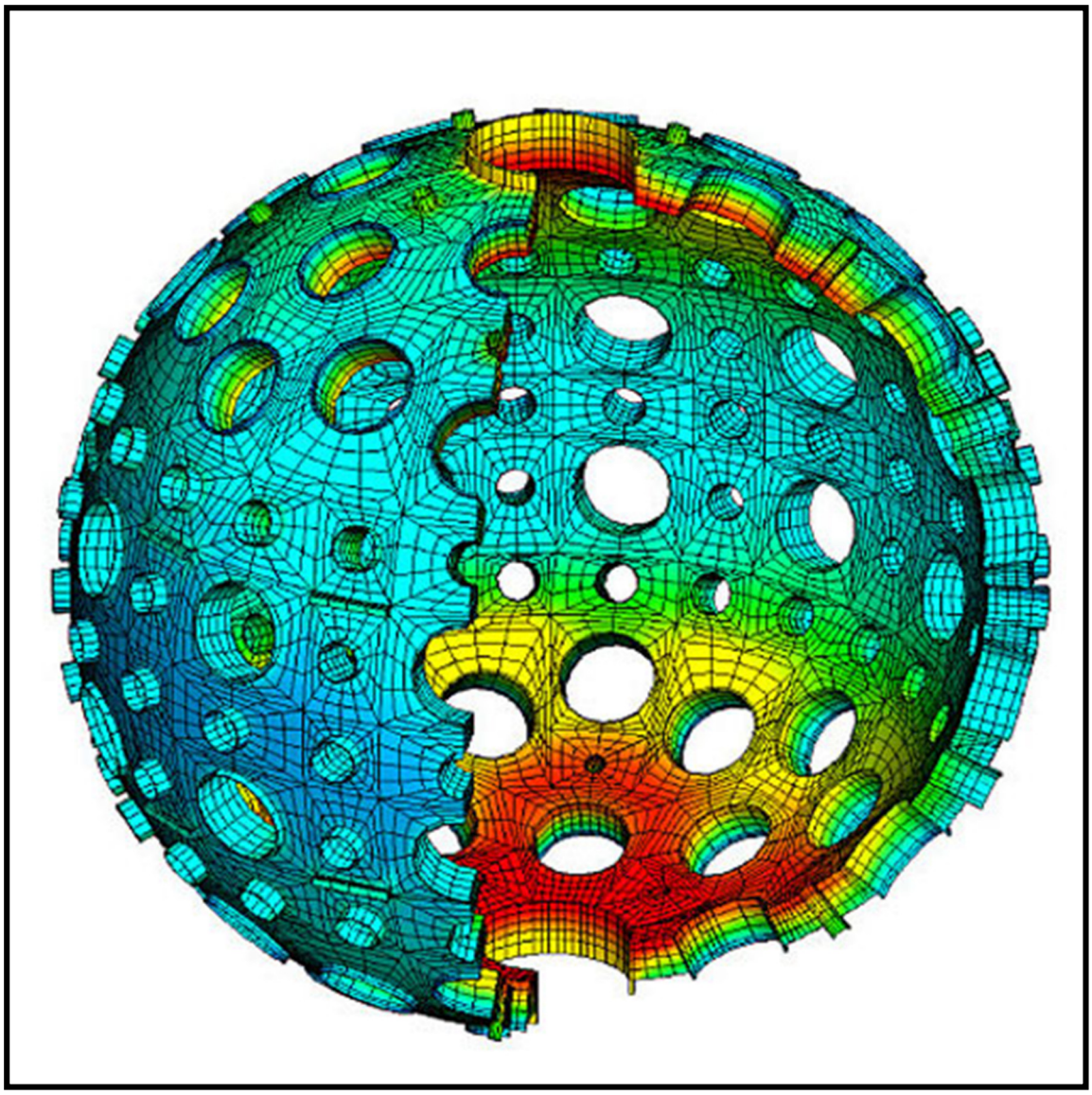
Εικόνα 58. Προσομοίωση σε LS-DYNA.

Όπως αναφέρεται στην ιστοσελίδα της Wikipedia: «Μέθοδος Πεπερασμένων Στοιχείων είναι η αριθμητική μέθοδος για την εύρεση προσεγγιστικών λύσεων σε προβλήματα οριακών τιμών μερικών διαφορικών εξισώσεων» [27]. Είναι δηλαδή μία μέθοδος η οποία μπορεί να δώσει αποτελέσματα τα οποία τείνουν στα πραγματικά, με βαθμό ακρίβειας ανάλογο με το βαθμό προσέγγισης στο πρόβλημα που είναι προς επίλυση. Το μειονέκτημα της είναι ότι όσο αυξάνεται ο βαθμός προσέγγισης, πολλαπλασιάζεται η απαιτούμενη επεξεργαστική ισχύς του υπολογιστικού συστήματος που θα εκτελέσει την επίλυση. Αυτό το πρόβλημα στις μέρες μας τείνει να εξαλειφτεί, λόγο της ραγδαίας αύξησης της υπολογιστικής ισχύος των ηλεκτρονικών υπολογιστών και την εφαρμογή του νόμου του Moore. Έτσι, μπορούμε από ένα απλό στατικό πρόβλημα, να περάσουμε σε Multiphysics προβλήματα, με πολλά φυσικά φαινόμενα να αναλύονται ταυτόχρονα, ενώ διατηρείται υψηλός βαθμός προσέγγισης. Η πιο διαδεδομένη μέθοδος πλέον για επίλυση πεπερασμένων στοιχείων είναι η χρήση computer clusters, όπως περιγράφηκε στα προηγούμενα κεφάλαια.

Η προσέγγιση επίλυσης μπορεί να περιγραφεί απλοϊκά, μεταφράζοντας το τόξο της περιφέρειας ενός κύκλου σε ένα σύνολο ευθειών. Όσο πιο πολλές είναι οι ευθείες αυτές, τόσο πιο πολύ το σχήμα που δημιουργείται πλησιάζει σε όψη το τόξο που θέλουμε να περιγράψουμε. Αυτού του τύπου η προσέγγιση μπορεί να γίνει σε μία, σε δύο ή σε τρεις διαστάσεις. Με αυτόν τον τρόπο, δημιουργείται το πλέγμα του μοντέλου που θέλουμε να παραστήσουμε με πεπερασμένα στοιχεία. Το μοντέλο αποτελείται από ένα πλήθος πολυγώνων (συνήθως τρίγωνα), τα οποία λέγονται στοιχεία, ενώ τα σημεία που ενώνονται οι κορυφές των πολυγώνων λέγονται κόμβοι. Οι κόμβοι ως τοπικά οριακά σημεία, είναι τα σημεία όπου οι οριακές συνθήκες του ενός στοιχείου περνάνε στα επόμενα, οπότε το πλήθος τους και οι αρχικές παράμετροι του δικτυώματος παίζουν σημαντικό ρόλο στην ποιότητα του αποτελέσματος και στην ακρίβεια των αποτελεσμάτων.

Ένα παράδειγμα για την μορφή και τη φύση του πλέγματος, είναι ο τρόπος που απεικονίζει τα τρισδιάστατα γραφικά ένας ηλεκτρονικός υπολογιστής. Αυτό εμφανίζεται σαν ένα σύνολο πολυγώνων (τρίγωνα), τα οποία στα μάτια του χρήστη δείχνουν σαν ένα ενιαίο σύνολο με καμπύλες και σχήμα. Ένα τέτοιο

παράδειγμα παρουσιάζεται στην Εικόνα 59 [28], ενώ ένα παράδειγμα επιλυμένου προβλήματος πζρουσιάζεται στην εικόνα 58 [29].



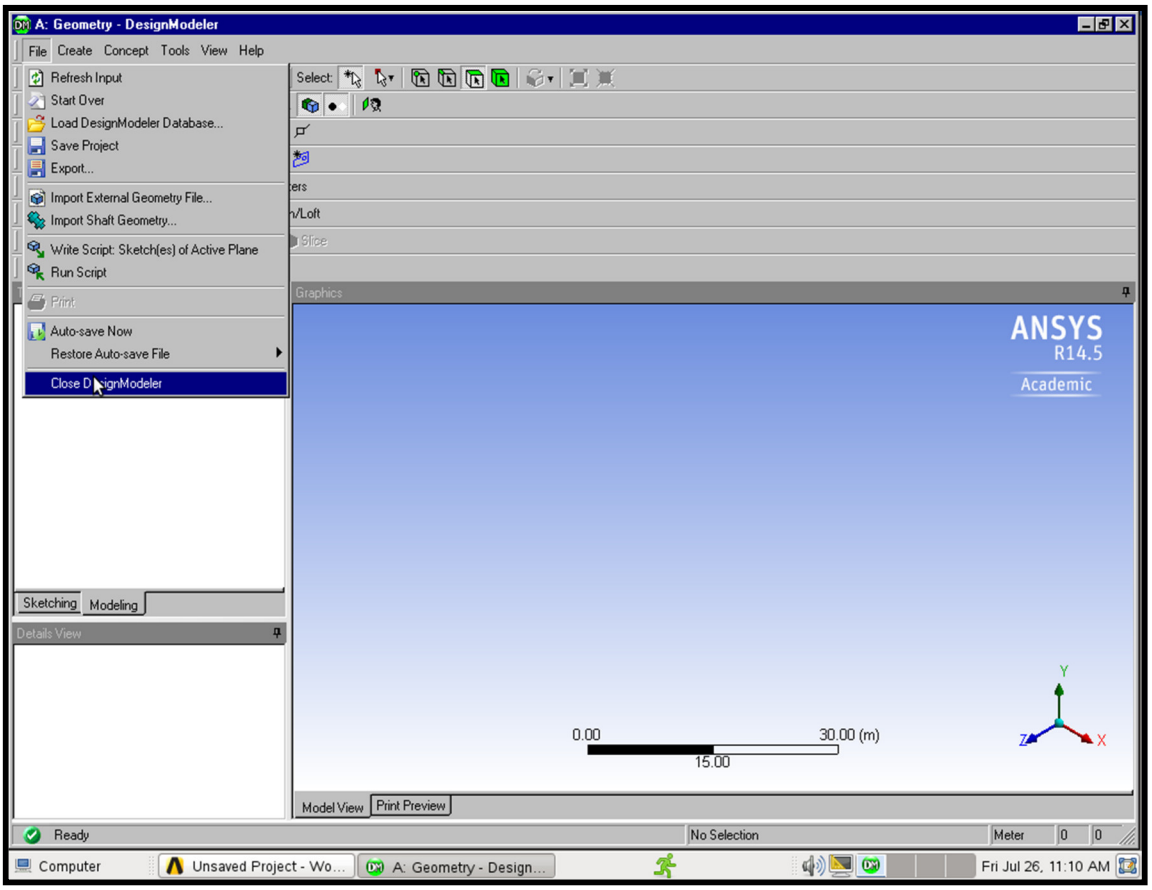
Εικόνα 59. Το μοντέλο του αντιδραστήρα σύντηξης NIS.

Για την εφαρμογή της μεθόδου υπάρχουν 4 διακριτά βήματα. Αυτά μπορούμε να τα εκτελέσουμε είτε από την ίδια την εφαρμογή επίλυσης, είτε από εξωτερικές εφαρμογές που συνεργάζονται με τον επιλύτη των πεπερασμένων στοιχείων:

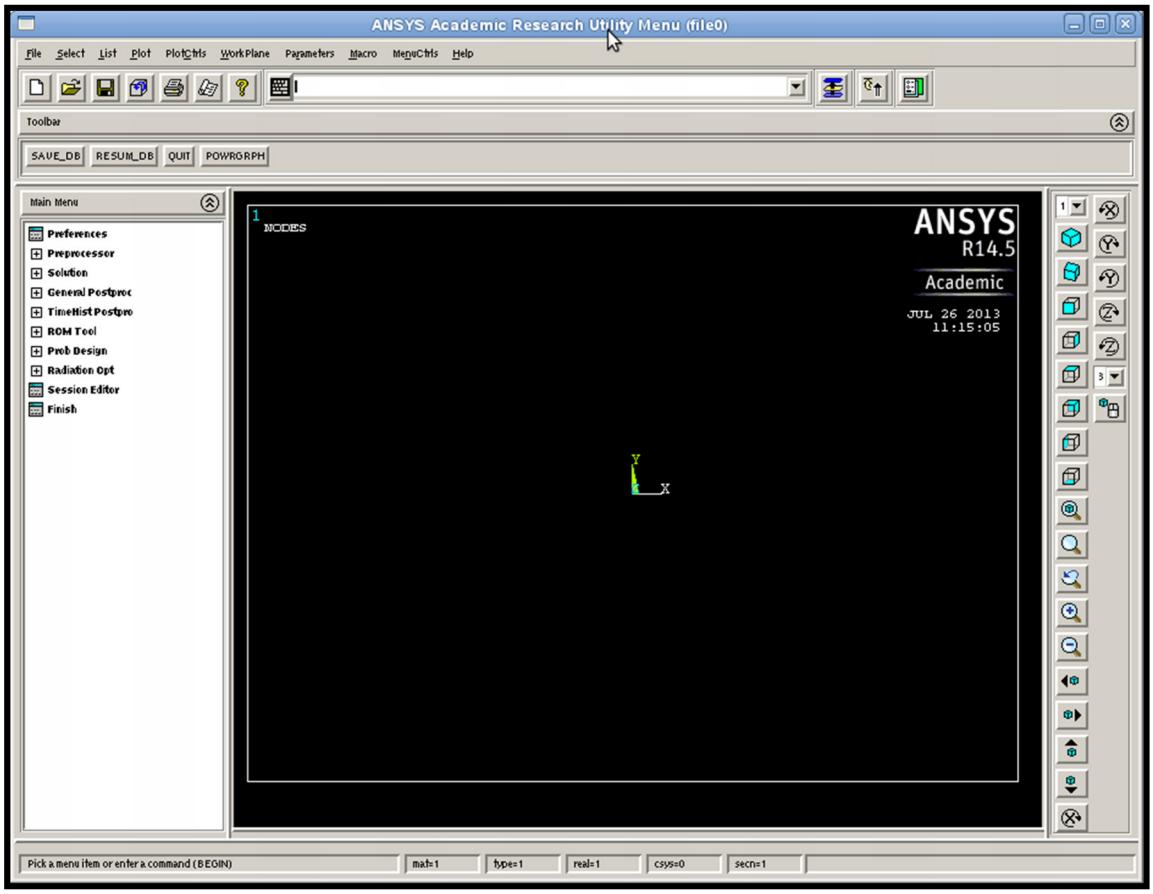
1. **Η κατασκευή της γεωμετρίας**
2. **Προ-επεξεργασία.** Ορισμός πλέγματος και καθορισμός παραμέτρων.
3. **Η επίλυση του προβλήματος**, μοντέλου.
4. **Η αξιολόγηση και η επεξεργασία των αποτελεσμάτων**.

Η γεωμετρία του μοντέλου, μπορεί να κατασκευαστεί με τη χρήση προγράμματος CAD, το οποίο, είτε μπορεί να είναι εξωτερικό (τρίτου

κατασκευαστή), είτε να παρέχεται από τον ίδιο το δημιουργό του επιλύτη. Με αυτή την ενέργεια γίνεται η δημιουργία του μοντέλου, που είναι τρισδιάστατο, δισδιάστατο ή μονοδιάστατο. Απαιτείται η χρήση αυτών, ώστε το μοντέλο να είναι ακριβές, να απεικονίζεται με κάθε λεπτομέρεια, αλλά και να μπορεί με την ίδια ευκολία να προστεθούν ή να μεταβληθούν τμήματα σε αυτό. Στις Εικόνες 60 και 61, βλέπουμε το software κατασκευής γεωμετρίας του ASNYS Workbench και APDL αντίστοιχα.



Εικόνα 60. Προ-επεξεργασία μοντέλου ANSYS Workbench.

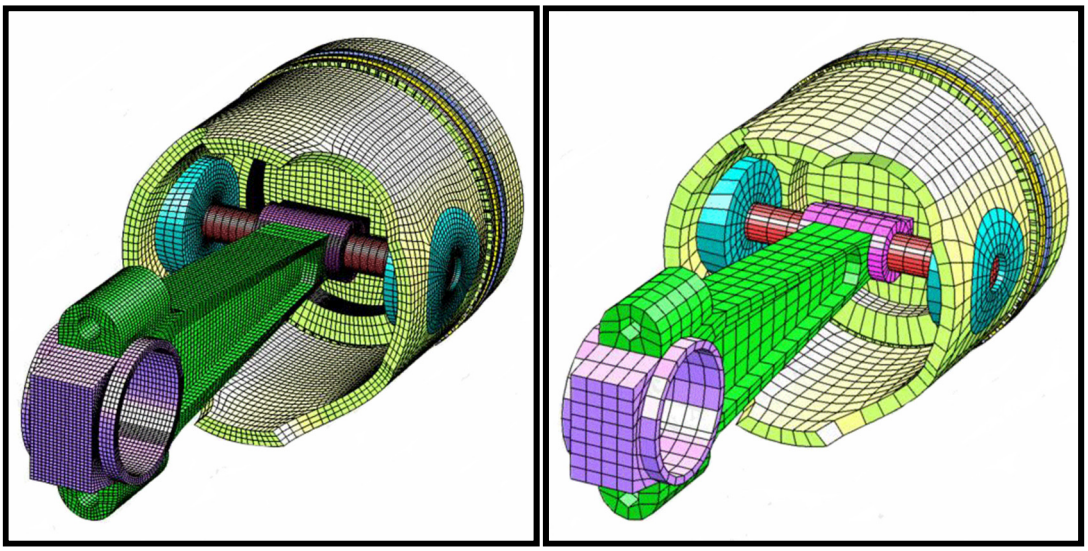


Εικόνα 61. Προ-επεξεργασία μοντέλου APDL.

Σε αυτό το σημείο, πρέπει να οριστούν τα βασικά φαινόμενα που διέπουν το πρόβλημα, τα οριακά σημεία, οι ιδιότητες των υλικών και η μορφή των στοιχείων του πλέγματος. Για αυτή την ενέργεια γίνεται χρήση προγραμμάτων, τα οποία λέγονται προεπεξεργαστές. Αυτοί είναι πολλές φορές ενσωματωμένοι στους επιλύτες, ή είναι εξωτερικά προγράμματα που παρέχονται από τους κατασκευαστές των επιλυτών. Καθορίζονται:

- Τα υλικά
- Το σχήμα
- Οι φυσικοί νόμοι
- Ο αριθμός των πεπερασμένων στοιχείων
- Η πυκνότητα του πλέγματος
- Η διάρκεια του χρόνου επίλυσης
- Οι αρχικές συνθήκες

Μπορούμε να δούμε την διαφορά στην πυκνότητα πλέγματος για το ίδιο μοντέλο, όπως παρουσιάζεται στην Εικόνα 62 [30].



Εικόνα 62. Πυκνό και αραιό πλέγμα.

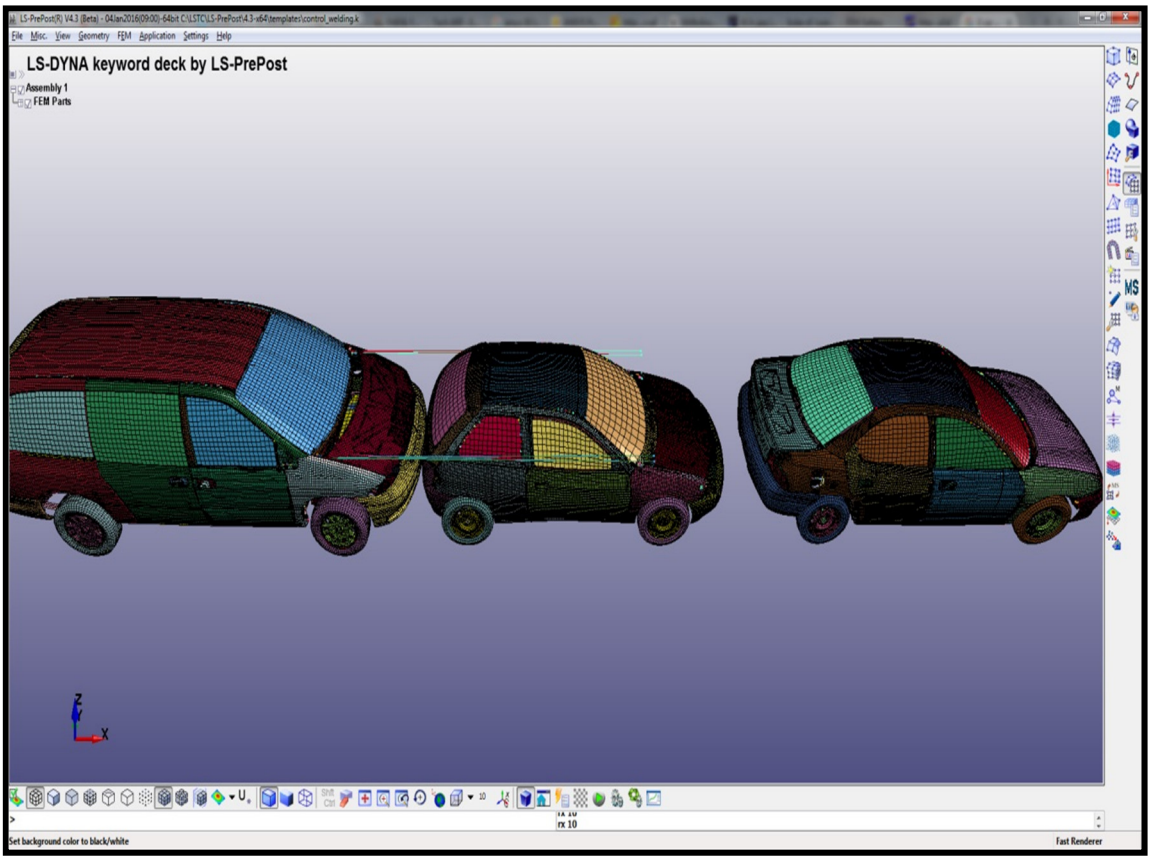
Αφού έχουμε προετοιμάσει το μοντέλο, εισάγοντας σε αυτό όλα τα απαραίτητα δεδομένα, έρχεται το επόμενο βήμα της χρήσης του επιλύτη. Σε αυτό το σημείο σημαντικό ρόλο παίζει το υπολογιστικό σύστημα που είναι διαθέσιμο. Η επεξεργαστική ισχύς, η μνήμη, η ταχύτητα του δικτύου υπολογιστών και η ταχύτητα του συστήματος αποθήκευσης (σκληροί δίσκοι), είναι τα βασικά χαρακτηριστικά που πρέπει να ληφθούν υπόψη, γιατί αυτά είναι που θα καθορίσουν τον χρόνο που θα απαιτηθεί για να πάρουμε το αποτέλεσμα. Είναι συνηθισμένο η επίλυση να διακόπτεται βίαια, με διακοπή λειτουργίας του υπολογιστικού συστήματος, αν η ετοιμασία στην προεπεξεργασία δεν είναι σωστά ορισμένη και απαιτούνται περισσότεροι πόροι από αυτούς που μπορεί να δώσει το σύστημα. Το αποτέλεσμα της επίλυσης αποθηκεύεται σε αρχεία, ώστε αυτά να μπορούν να αναλυθούν από τον χρήστη στο επόμενο βήμα, αυτό της αξιολόγησης των αποτελεσμάτων.

4.2. Προσομοιώσεις στα cluster σε προβλήματα FEM, Ls-Dyna και ANSYS APDL και η απόδοση της παράλληλης επεξεργασίας

Αφου εγκαταστάθηκαν και ελέχθηκαν hardware και Software μελετούμε την επίδοσή τους στην επίλυση προβλημάτων με παράλληλη επεξεργασία, στοχεύοντας στη βελτιστοποίηση σε hardware και software, ώστε να έχουμε ταχύτερα και πιο ακριβή αποτελέσματα. Το ίδιο μοντέλο, με διαφορετικές

παραμέτρους μοντελοποίησης και επίλυσης υποβάλλεται σε διαδοχικές προσομοιώσεις.

Το μοντέλο πεπερασμένων στοιχείων που επιλέχθηκε για τις δοκιμές είναι το 3cars_shell2_150ms [31]. Αφορά τη συγκρουση 3 αυτοκινήτων όπως παρουσιάζεται στην Εικόνα 63 και παρέχεται από την LS-DYNA ελεύθερα, ώστε να μπορεί να γίνει δοκιμή του υπολογιστικού συστήματος και έλεγχος για την ορθή λειτουργία του επιλύτη. Αυτό χρησιμοποιήθηκε για να πάρουμε μετρήσεις, ως προς το χρόνο επίλυσης του προβλήματος του μοντέλου κάτω από διαφορετικές συνθήκες και παράλληλα να ελέγχουμε τα υπολογιστικά συστήματα που έχουν δημιουργηθεί για υψηλής απόδοσης παράλληλη επεξεργασία, αν λειτουργούν σωστά κάτω από οριακές συνθήκες. Με αυτόν τον τρόπο, θα μπορέσουμε να βγάλουμε ένα γενικό συμπέρασμα, ως προς την ποιότητα του καθενός cluster και να βρούμε τις δυνατότητες βελτιστοποίησης αυτών. Η τεχνική που χρησιμοποιήθηκε, είναι η MPP (Multiple Parallel Processing). Αυτό σημαίνει ότι πρόκειται για τεχνική η οποία είναι σχεδιασμένη, για παράλληλη επεξεργασία με χρήση πολλών ανεξάρτητων ηλεκτρονικών υπολογιστών. Η μέθοδος δηλαδή που είναι η πλέον κατάλληλη για τα τις συνθέσεις -A και -B, τα cluster που χτίσαμε. Το LS-DYNA για την λειτουργία του απαιτεί να οριστούν δύο παράμετροι μνήμης, το **mem1** και το **mem2**. Το **mem1** είναι η ποσότητα της μνήμης που απαιτεί ο κύριος πυρήνας που κάνει αποδόμηση του προβλήματος και **mem2** είναι η ποσότητα της μνήμης που χρειάζεται ο κάθε πυρήνας για την επίλυση του προβλήματος[22].Οι ποσότητες αυτές εκφράζονται σε words, όπου 8bytes=1word σε υπολογιστή 64bit. Οι χρόνοι επίλυσης καταγράφονται σε αρχείο κειμένου, που μετά το πέρας των προσομοιώσεων, εισήχθησαν στο πρόγραμμα Origin [32], στο οποίο έγινε η ανάλυση των επιδόσεων και η δημιουργία των γραφικών παραστάσεων που ακολουθούν. Για την άμεση και εύκολη αντιστοιχία και σύγκριση των χρόνων προσομοίωσης, οι ίδιοι/συγκρίσιμοι αριθμοί πυρήνων όπου επιλέχθηκαν, και χρησιμοποιήθηκαν σε κάθε επίλυση, υπογραμμίζονται.



Εικόνα 63. Το μοντέλο 3 car crash.

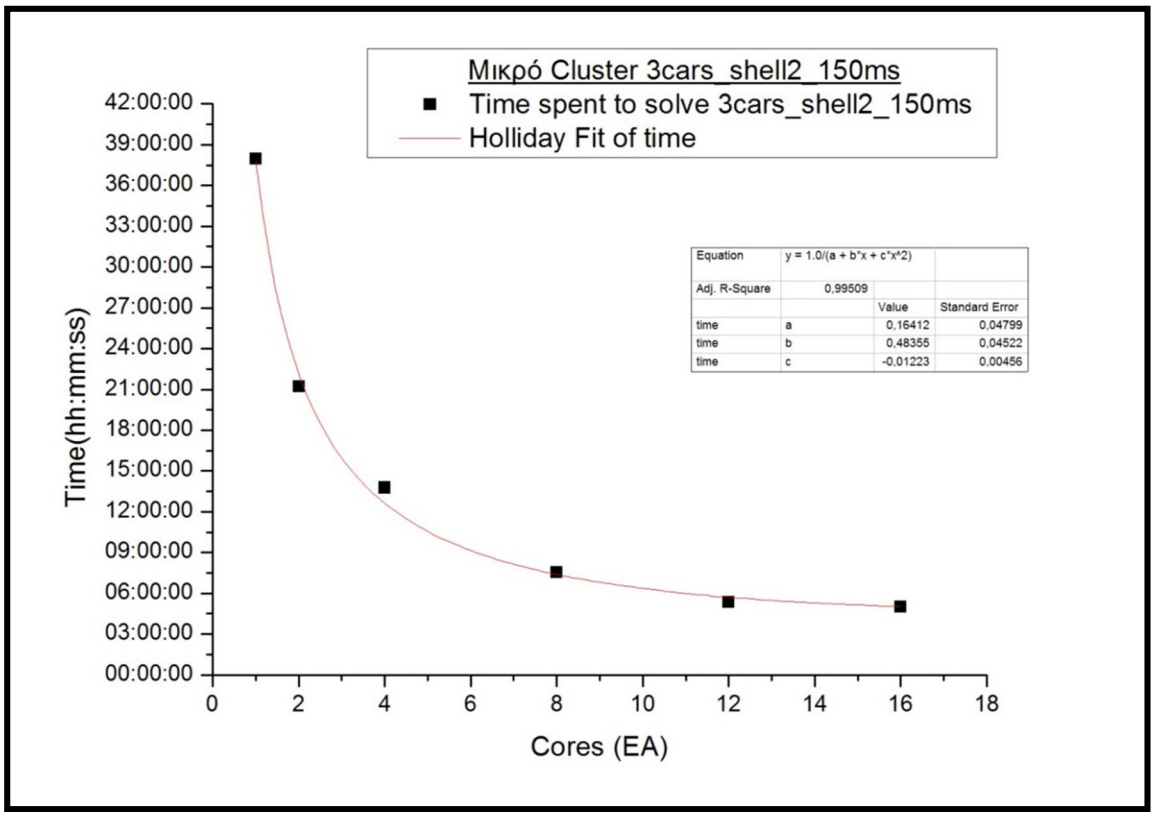
4.2.1 3-car crash στην Συστοιχία –A Ls-Dyna.

Όπως περιγράφηκε, η Συστοιχία -A αποτελείται από συνολικά 5 ηλεκτρονικούς υπολογιστές. Σε αυτή πέρα από τις πρώτες δοκιμές, που έγιναν καθαρά για λόγους ελέγχου της σωστή λειτουργίας του λογισμικού και του εξυπηρετητή αδειών χρήσης (license manager), κρατήθηκαν χρόνοι για κάθε προσομοίωση που έληξε επιτυχημένα.

Αρχικά βλέπουμε, Εικόνα 65, την καμπύλη της ταχύτητας περάτωσης της επίλυσης ως προς τον αριθμό των διαθέσιμων πυρήνων στον επιλύτη. Όλες οι μετρήσεις που παρουσιάζονται στο παρακάτω διάγραμμα, έχουν γίνει με ορισμό μνήμης : **mem1=200M και mem2=60M.**

Αποτελέσματα χρόνων επίλυσης σε σχέσης με τους χρησιμοποιούμενους πυρήνες:

Πυρήνες	Χρόνος επίλυσης (HH:mm:ss)
1	37:58:15
2	21:13:32
4	13:45:44
8	07:32:04
12	05:21:28
16	04:59:43



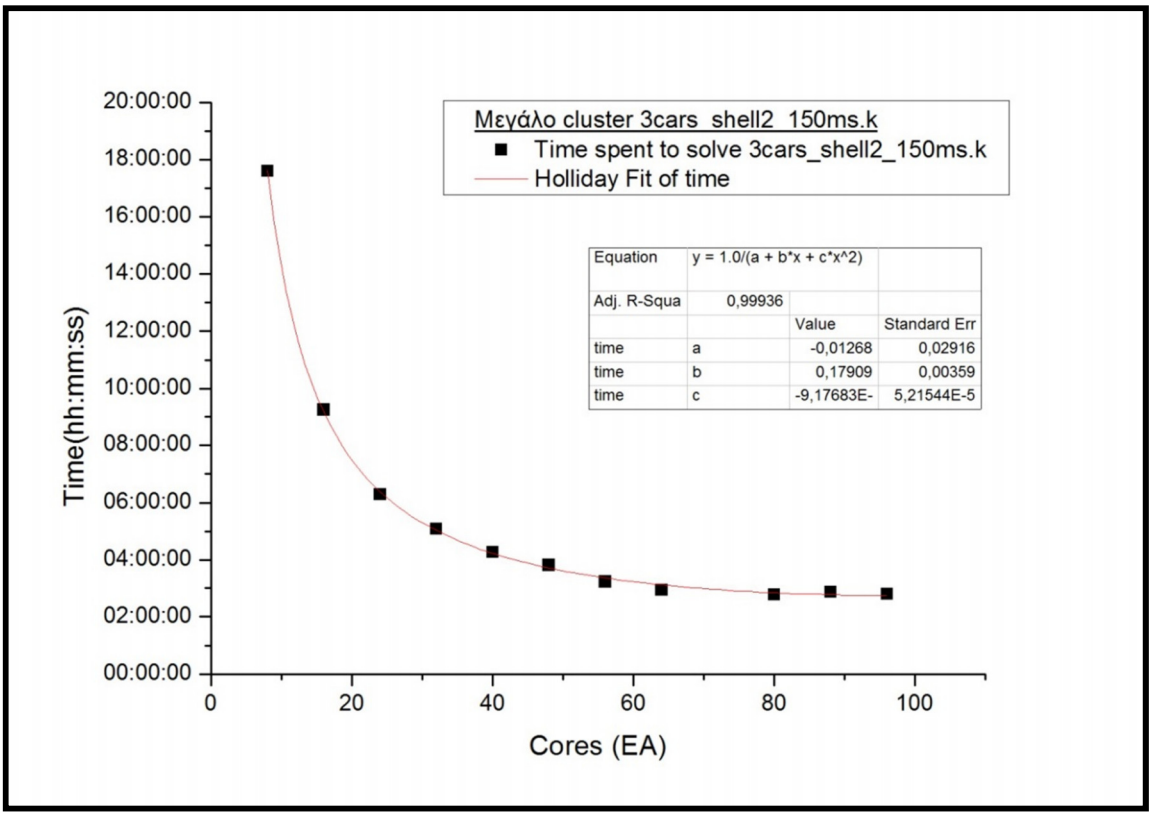
Εικόνα 65. Συστοιχία –A, 3 car crash - LS-DYNA.

Σε αυτό το διάγραμμα, μπορεί να παρατηρηθεί ότι η αύξηση του αριθμού των πυρήνων που παρέχονται στον επιλύτη, δεν έχει γραμμική επίδραση, αλλά αντίστροφη πολυονυμική της μορφής $y = \frac{1}{a+bx+cx^2}$. Η προσαρμογή είναι εξαιρετικά καλή, όπως μας δείχνει ο συντελεστής προσαρμογής R2, ο οποίος είναι 0,99509, ενώ το σφάλμα είναι πολύ μικρό. Έτσι γίνεται φανερό ότι, για επιτευχθεί μεγάλη αύξηση της ταχύτητας χρειαζόμαστε πραγματικά πολύ μεγάλα υπολογιστικά συστήματα. Δεν έγινε πλήρης χρήση των δυνατοτήτων της Συστοιχίας Α, γιατί οδηγούσε σε κατάρρευση του συστήματος λόγω έλλειψης μνήμης και υπολογιστικής ισχύος σε συστήματα που διαχειρίζονται το cluster.

4.2.2 3-car crash στην Συστοιχία –B Ls-Dyna.

Όπως περιγράφηκε, η Συστοιχία -A αποτελείται από συνολικά 16 ηλεκτρονικούς υπολογιστές. Σε αυτή πέρα από τις πρώτες δοκιμές, που έγιναν καθαρά για λόγους ελέγχου της σωστή λειτουργίας του λογισμικού και του εξυπηρετητή αδειών χρήσης (license manager), κρατήθηκαν χρόνοι για κάθε προσομοίωση που έληξε επιτυχημένα. Η επίλυσή έγινε με δύο διαφορετικούς τρόπους. Πρώτα με το LS-DYNA και μετά με το ANSYS APDL χρησιμοποιώντας το module LS-DYNA Multiphysics.

Στην Εικόνα 66 μπορούμε να παρατηρήσουμε την καμπύλη ταχύτητας περάτωσης της επίλυσης, ως προς τον αριθμό των διαθέσιμων πυρήνων στον επιλύτη. Όλες οι μετρήσεις που παρουσιάζονται στο παρακάτω διάγραμμα έχουν γίνει με ορισμό μνήμης**: mem1=200M και mem2=60M.** Ακολουθώντας την ίδια φιλοσοφία, στην Εικόνα 66 απεικονίζονται γραφικά οι επιδόσεις σε χρόνους σε σχέση με τον αριθμό πυρήνων που δεσμεύτηκαν.



Εικόνα 66. Συστοιχία –B, 3 car crash με LS-DYNA.

Πυρήνες	Χρόνος επίλυσης (HH:mm:ss)
8	17:36:59
16	09:15:09
24	06:17:20
32	05:04:38
40	04:16:05
48	03:48:48
56	03:13:51
64	02:56:13
80	02:46:36
88	02:52:22
96	02:48:57

Σε αυτό το διάγραμμα παρατηρούμε παρόμοιο αποτέλεσμα με το αντίστοιχο της συστοιχίας -A. Μπορεί να παρατηρηθεί και πάλι ότι η αύξηση του αριθμού των πυρήνων που παρέχονται στον επιλύτη, δεν έχει γραμμική επίδραση, αλλά αντίστροφη πολυονυμική της μορφής $\mathbf{y} = \frac{\mathbf{1}}{\mathbf{a+bx+cx^2}}$. Η προσαρμογή είναι εξαιρετικά καλή, όπως μας δείχνει ο συντελεστής προσαρμογής R2, ο οποίος είναι 0,99936, ενώ το σφάλμα είναι πολύ μικρό.

Επίσης και στην Συστοιχία -B, δεν έγινε πλήρης χρήση των δυνατοτήτων της, γιατί οδηγούσε σε κατάρρευση του συστήματος λόγω έλλειψης μνήμης και υπολογιστικής ισχύος σε συστήματα διαχείρισης του cluster.

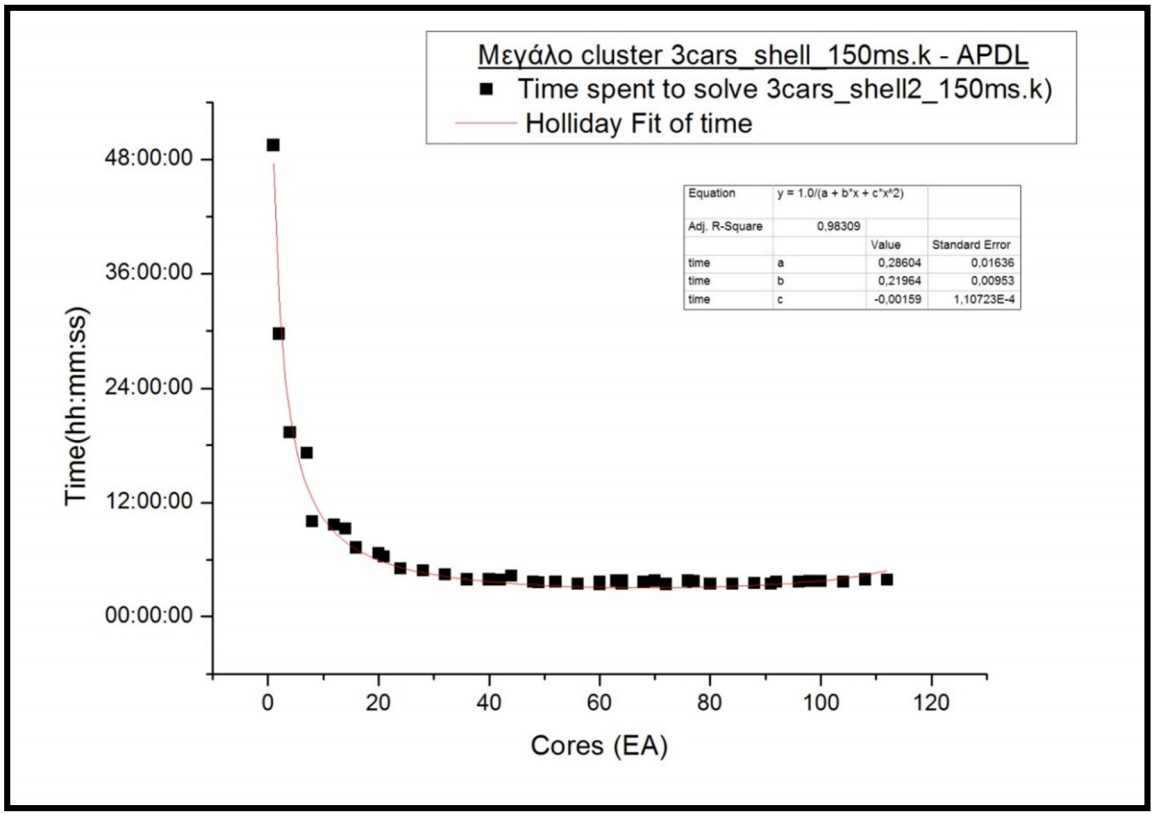
4.2.3 3-car crash στην Συστοιχία -B ANSYS APDL.

Το ίδιο μοντέλο επιλύεται και πάλι στη Συστοιχία -B με τη χρήση όμως του ANSYS APDL. Σε αυτήν την περίπτωση, παρουσιάζονται διαφοροποιήσεις σε σχέση με τις 2 προηγούμενες. Συγκεκριμένα, μετά τους 48 πυρήνες παρουσιάζεται μια σχετική σταθεροποίηση στον χρόνο επίλυσης. Η προσομοίωση έγινε με ρύθμιση μνήμης 10.000.000 (words) και το node με όνομα blade15 είχε 24Gb RAM.

Στην Εικόνα 67, μπορούμε να παρατηρήσουμε την καμπύλη ταχύτητας περάτωσης της επίλυσης, ως προς τον αριθμό των διαθέσιμων πυρήνων στον επιλύτη.

Αποτελέσματα χρόνων επίλυσης σε σχέσης με τους χρησιμοποιούμενους πυρήνες:

Πυρήνες	Χρόνος επίλυσης (HH:mm:ss)
1	49:30:53
2	29:41:44
4	19:20:51
7	17:09:43
8	10:01:38
12	09:40:13
14	09:15:58
16	07:16:08
20	06:39:28
21	06:19:17
24	05:03:47
28	04:49:51
32	04:25:05
36	03:55:21
40	03:54:35
42	03:51:38
44	04:18:25
48	03:41:07
49	03:35:56
52	03:39:51
56	03:26:34
60	03:24:20 με 9 Nodes
60	03:38:22 με 10 nodes
63	03:46:40
64	03:27:14 με 10 nodes
64	03:45:04 με 11 nodes
68	03:40:57
70	03:45:24
72	03:24:14
76	03:45:47
77	03:42:05
80	03:28:29
84	03:26:21
88	03:32:39
91	03:27:30
92	03:40:26
96	03:40:01
98	03:44:31
100	03:44:51
104	03:39:37
108	03:54:53
112	03:50:44



Εικόνα 67. Συστοιχία –B, 3 car crash με APDL.

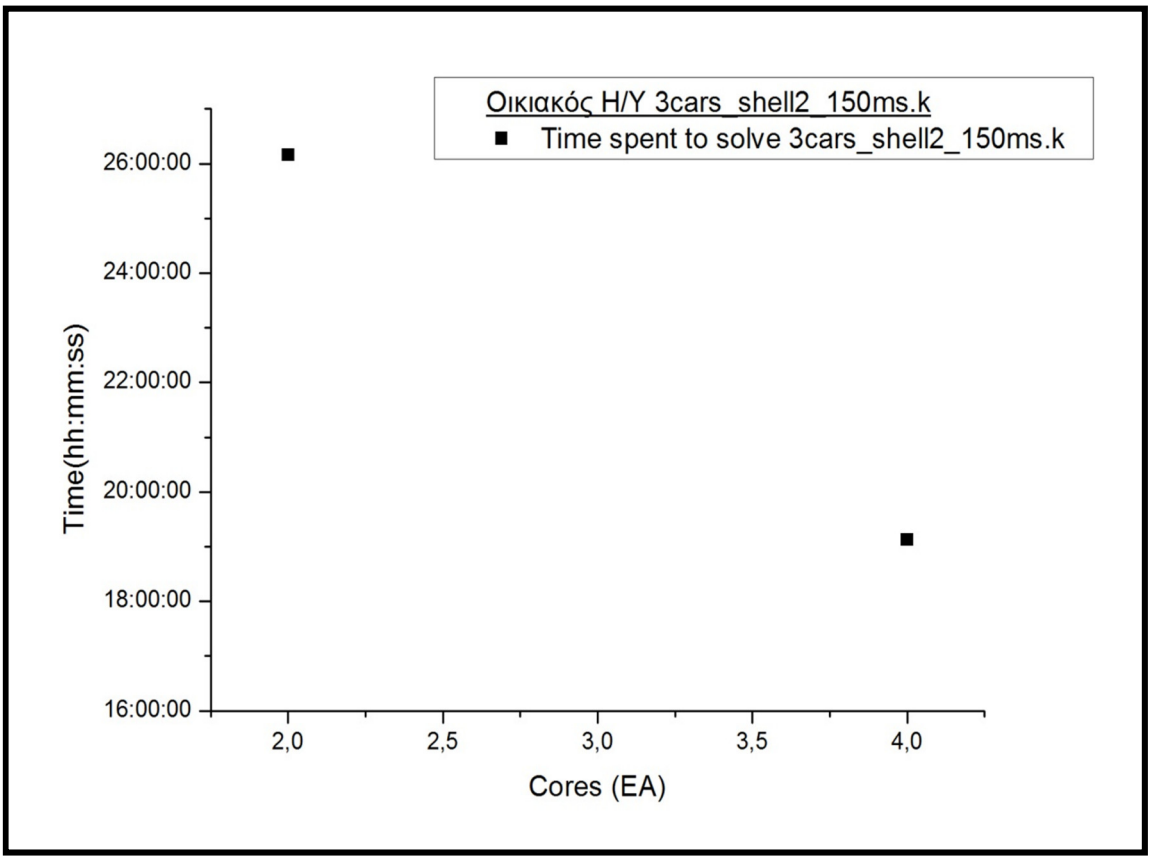
Σε αυτό το διάγραμμα μπορεί να παρατηρηθεί ότι η αύξηση του αριθμού των πυρήνων που παρέχονται στον επιλύτη, δεν έχει γραμμική επίδραση, αλλά αντίστροφη πολυονυμική της μορφής $\mathbf{y} = \frac{\mathbf{1}}{\mathbf{a+bx+cx^2}}$ και πάλι. Η προσαρμογή είναι εξαιρετικά καλή, όπως μας δείχνει ο συντελεστής προσαρμογής R2, ο οποίος είναι 0,98309 , ενώ το σφάλμα είναι πολύ μικρό. Επιπρόσθετα, στην περίπτωση αυτή έγινε δυνατή η χρήση όλων των διαθέσιμων πυρήνων του συστήματος (14 nodes X 8 cores το καθένα), ενώ 2 nodes δεν χρησιμοποιήθηκαν. Το ένα (blade14) εκτελούσε το χρέος του domain controller του cluster και το άλλο (blade13) γιατί δεν είχε επαρκή διαθέσιμη μνήμη, αφού την χρησιμοποιήσαμε για να δημιουργήσουμε ένα node με 24Gb RAM.

Άλλο ένα αξιοσημείωτο φαινόμενο, είναι ότι όσο πιο πολλά node κάναμε χρήση με το ίδιο αριθμό πυρήνων, τόσο ο χρόνος περάτωσης αυξανόταν. Έτσι, για 60 πυρήνες με 9 nodes είχαμε χρόνο 3h 24min 20sec, ενώ με 10 node 3h 38min 22sec. Επίσης, παρόμοια αποτελέσματα στο χρόνο περάτωσης υπήρξαν για 64 πυρήνες, για 10 node 3h 27min 14sec και για 11 node 3h 43min 24sec.

4.2.4 3-car crash σε PC ANSYS APDL.

Για λόγους σύγκρισης και ανάλυσης της ταχύτητας επίλυσης του προβλήματος, σε σχέση με αυτή που θα χρειαζόταν αν χρησιμοποιούσαμε ένα μέσο προσωπικό ηλεκτρονικό υπολογιστή, επιλύσαμε ξανά το ίδιο πρόβλημα. Το PC που ανέλαβε να υλοποιήσει την προσομοίωση αποτελείται από 8Gb RAM DDR2 και ένα τετραπύρηνο κεντρικό επεξεργαστή AMD Phenom II 965. Σε αυτήν την περίπτωση δεν ήταν δυνατό να πάρουμε πάνω από δύο χρόνους επίλυσης, λόγω του μικρού αριθμού πυρήνων και του μεγάλου χρόνου για την περάτωση της επίλυσης. Γι' αυτό άλλωστε δεν έχει γίνει προσαρμογή καμπύλης, όπως στις προηγούμενες περιπτώσεις. Τους χρόνους αυτούς μπορούμε να τους δούμε στην Εικόνα 68:

Πυρήνες	Χρόνος επίλυσης (HH:mm:ss)
2	**26:09:24**
4	**19:07:44**



Εικόνα 68. Προσωπικός υπολογιστής, 3 car crash με APDL.

Τα αποτελέσματα είναι εντός των αναμενόμενων ορίων. Ειδικότερα είναι απόλυτα συγκρίσιμα με αυτά της Συστοιχίας -B, ενώ υστερούν όπως αναμενόταν της Συστοιχίας -A, λόγω της ανώτερης υπολογιστικής ισχύος που διαθέτει αυτή ανά χρησιμοποιούμενο πυρήνα.

4.3. Ανάλυση απόδοσης παράλληλης επεξεργασίας σε προβλήματα FEM

Το πρώτο που μπορούμε να παρατηρήσουμε, είναι η μεγάλη απόκλιση που παρατηρείται στη Συστοιχία -A και την Συστοιχία -B ως προς τους χρόνους επίλυσης. Στους 8 πυρήνες υπάρχει χρονική διαφορά 10 ωρών περίπου (Συστοιχία -A: 7h 32min 4sec – Συστοιχία -B: 17h 36min 59sec) και στους 16 πυρήνες υπάρχει χρονική διαφορά περίπου 4 ωρών και 15 λεπτών (Συστοιχία -A: 4h 59min 43sec – Συστοιχία -B: 9h 15min 9sec), με επίλυση με το LS-DYNA. Το άλλο που παρατηρούμε, είναι ότι η αύξηση των πυρήνων δεν σημαίνει ανάλογη αύξηση στην ταχύτητα επίλυσης. Ένα άλλο χαρακτηριστικό είναι, ότι κάθε πυρήνας απαιτούσε την ίδια ποσότητα μνήμης RAM για την επίλυση, ανεξάρτητα από τον συνολικό αριθμό των πυρήνων που έγινε χρήση για την επίλυση του προβλήματος. Η σημασία αυτής της παρατήρησης, έγκειται στο γεγονός ότι, πρέπει να υπάρχει έλεγχος, τουλάχιστον στα πρώτα βήματα της επίλυσης, ώστε να μην υπάρξει κατάρρευση του συστήματος από εξάντληση της RAM. Προχωρώντας και ταυτόχρονα παρατηρώντας τους χρόνους επίλυσης με APDL, στην Συστοιχία -B, είναι φανερή η σταθεροποίηση αυτών χρησιμοποιώντας 48 πυρήνες ή περισσότερους. Παράλληλα, στην ίδια περίπτωση μπορεί να παρατηρηθεί ότι από τους 92 πυρήνες και πάνω, υπάρχει μία ανοδική τάση στον χρόνο επίλυσης. Κάτι παρόμοιο μπορούμε να δούμε στην περίπτωση επίλυσης με LS-DYNA στην Συστοιχία -B. Αν και η καμπύλη χρόνου περάτωσης είναι πολύ ομαλή, Εικόνα 66, έχουμε σταθεροποίηση των χρόνων περάτωσης επίλυσης στους 80 πυρήνες. Πρέπει να δούμε επίσης, ότι οι μετρήσεις που πήραμε από τον προσωπικό ηλεκτρονικό υπολογιστή είναι σχεδόν ταυτόσημες με αυτές της Συστοιχίας -B και υστερούν της Συστοιχίας -A, γεγονός που θα μας βοηθήσει στην σύγκριση των δυνατοτήτων των συστημάτων μας. Τέλος υπάρχει απόκλιση στον χρόνο περάτωσης μεταξύ APDL και LS-DYNA, με το LS-DYNA να πλεονεκτεί στους πολλούς πυρήνες και το APDL στους λίγους.

Η απόκλιση των χρόνων περάτωσης επίλυσης μεταξύ της Συστοιχίας -Α και της Συστοιχίας -Β είναι εμφανής. Στους 8 πυρήνες υπάρχει χρονική διαφορά 10 ωρών περίπου (Συστοιχία -Α: 7h 32min 4sec – Συστοιχία -Β: 17h 36min 59sec) και στους 16 πυρήνες υπάρχει χρονική διαφορά περίπου 4 ωρών και 15 λεπτών (Συστοιχία -Α: 4h 59min 43sec – Συστοιχία -Β: 9h 15min 9sec), με επίλυση με το LS-DYNA. Αυτό μας φανερώνει ότι 1 node της Συστοιχίας -Α, είναι ισχυρότερο από 2 node της Συστοιχίας -Β. Δηλαδή, 4 πυρήνες (Intel i7 870) και μνήμη 4Gb DDR3 είναι πιο αποδοτικά από τουλάχιστον 16 πυρήνες (Xeon E5405) και 18Gb ECC DDR2 μνήμης. Επίσης τα node της Συστοιχίας -Α είναι COTS (Components Of The Self), δηλαδή αποτελούνται από κομμάτια που μπορούμε να τα προμηθευτούμε από οποιοδήποτε κατάστημα πώλησης ηλεκτρονικών υπολογιστών. Αντίθετα, η Συστοιχία -Β αποτελείται από αποκλειστικά κομμάτια για εξυπηρετητές. Αυτό μας κάνει φανερό ότι, η τεχνολογική εξέλιξη και η κατοχή σύγχρονων υπολογιστικών μηχανών έχει άμεση συνέπεια στην ταχύτητα με την οποία θα εξαχθούν τα αναμενόμενα αποτελέσματα.

Αναλύοντας τη δομή των δύο αυτών συνθέσεων, οι διαφορές τους που θα μπορούσαν να επηρεάσουν την ταχύτητα επίλυσης είναι :

α. Η τεχνολογία και η ταχύτητα της κεντρικής μονάδας επεξεργασίας.
β. Η τεχνολογία και η ταχύτητα της μνήμης.
γ. Η ταχύτητα του σκληρού δίσκου κάθε μονάδας.

Στα παραπάνω, τα node της Συστοιχίας -Α υπερτερούν στην τεχνολογία σε όλα τα πεδία, εκτός από την ποσότητα της μνήμης RAM και στην δυνατότητα της αυτόματης διόρθωσης λαθών που έχει η Συστοιχία -Β. Για την δεύτερη περίπτωση, πρέπει να πούμε ότι αναφερόμαστε στον σκληρό δίσκο που τρέχει το λειτουργικό σύστημα και η μνήμη SWAP της Συστοιχίας -Β και όχι στον δίσκο SAN που λειτουργεί με τεχνολογία fibrechannel. Επίσης, είναι γνωστό ότι η μνήμη DDR2 έχει μικρή απόκλιση στην ποσοστό που επηρεάζει την ταχύτητα ενός υπολογιστικού συστήματος σε σχέση με την μνήμη DDR3. Αυτό συμβαίνει γιατί, ενώ φαινομενικά αυξάνεται ο χρονισμός της μνήμης RAM, αυξάνεται ταυτόχρονα και το Latency (καθυστέρηση). Επομένως, μπορούμε με σχετική ασφάλεια να οδηγηθούμε ότι ο λόγος της απόκλισης αυτής, είναι η μεγάλη διαφορά που

υπάρχει στην υπολογιστική ισχύ μεταξύ των κεντρικών επεξεργαστών των nodes των δύο cluster. Αυτό βεβαίως δεν αποκλείει ότ,ι ο πιο αργός σκληρός δίσκος και η πιο αργή RAM δεν έχουν το ποσοστό της ευθύνης τους στους αργούς χρόνους επίλυσης.

Από τα παραπάνω 3 διαγράμματα, είναι εύκολο να παρατηρηθεί, ότι το να βάλουμε ένα μεγάλο αριθμό πυρήνων, δεν σημαίνει πως θα υπάρχει κατ' ανάγκην αντίστοιχη βελτίωση στο χρόνο επίλυσης της προσομοίωσης πεπερασμένων μαθηματικών, που έχουμε εισάγει στο υπολογιστικό μας σύστημα. Τα διαγράμματα της Εικόνας 65 και της Εικόνας 66 και το αρχικό τμήμα του διαγράμματος της Εικόνας 67, αποτυπώνουν αυτό το συλλογισμό. Εδώ γίνεται άμεσα αντιληπτή σε όλους η ανάγκη για cluster, τα οποία αποτελούνται από χιλιάδες κεντρικούς επεξεργαστές και τεράστια ποσότητα μνήμης RAM, για να προσομοιωθούν αστροφυσικά και παγκόσμια φαινόμενα, όπως το παγκόσμιο κλίμα. Αυτό γιατί, για να καταφέρουμε να διπλασιάσουμε την αποτελεσματικότητα μιας υπολογιστικής μηχανής θα πρέπει να προσφέρουμε πολλαπλάσια επεξεργαστική παράλληλη ισχύ. Έτσι, μπορεί να βελτιωθεί η εξέλιξη της τεχνολογίας και της συνολικής ωφέλιμης επεξεργαστικής ισχύος, με την αύξηση της επεξεργαστικής ισχύος ανά πυρήνα ή βρίσκοντας μεθόδους οι οποίες θα μας επιτρέψουν να συνδέουμε οικονομικά χιλιάδες node, ξεπερνώντας στον περιορισμό της ταχύτητας του δικτύου που θα διασυνδέονται αυτοί και των συστημάτων αποθήκευσης δεδομένων.

Η χρήση της μνήμης RAM, αν και δεν είναι ορατή στα παρατιθέμενα στοιχεία, παρουσίαζε μία ιδιομορφία. Για την επίλυση του κάθε μοντέλου, απαιτούταν η ίδια ποσότητα μνήμης ανά πυρήνα. Έτσι για κάθε πυρήνα στο πρόβλημά μας χρειαζόμασταν περίπου 1,2Gb μνήμης ,RAM ενώ για τον πυρήνα 0 (τον πρώτο ο οποίος εκτελεί και καθήκοντα ελεγκτή της διαδικασίας) περίπου 1,5Gb. Αυτό έφτανε τα συστήματά μας στα όριά τους. Έτσι γινόταν διαλλακτικά χρήση της μνήμης SWAP, γεγονός που οδηγούσε σε ραγδαία ελάττωση της ταχύτητας επεξεργασίας ή σε ορισμένες περιπτώσεις στη Συστοιχία -B, στην πλήρη κατάρρευση του cluster. Ο τρόπος που γινόταν η διαχείριση μνήμης RAM, αν και προβλέψιμος ως προς τη συνολική ποσότητα που απαιτούταν, ανάγκαζε τουλάχιστον για τα πρώτα στάδια συνεχή παρακολούθηση των πόρων του συστήματος. Αυτό ενισχυόταν από την επιλογή που υπάρχει για το LS-DYNA, το

AUTO_MEMORY. Η συγκεκριμένη επιλογή, επιτρέπει στο LS-DYNA να δεσμεύει αυτόματα όση μνήμη RAM απαιτεί, οποιαδήποτε στιγμή και σε οποιοδήποτε σημείο την επίλυσης, το οποίο έπρεπε να ενεργοποιηθεί για τον λόγο που θα εξηγήσουμε στην παράγραφο της βελτιστοποίησης.

Η γραφική παράσταση της Εικόνας 67 δείχνει ότι μετά τους 48 πυρήνες οι χρόνοι περάτωσης επίλυσης σταθεροποιούνται. Αυτό σε ιδανικές συνθήκες δεν θα έπρεπε να συμβαίνει. Θα έπρεπε έστω και σε πάρα πολύ μικρή κλίμακα να έχουμε μείωση των χρόνων αυτών, και αντίθετα με την προηγούμενη παραδοχή, βλέπουμε ότι μετά τους 92 πυρήνες υπάρχει η τάση για αύξηση του χρόνου περάτωσης επίλυσης, γεγονός που είναι εντελώς εκτός κάθε θεωρητικής προσέγγισης για την λειτουργία ενός cluster για παράλληλη επεξεργασία μοντέλων πεπερασμένων στοιχείων. Αυτή η μη φυσιολογική συμπεριφορά, θα πρέπει να έχει κάποια αίτια. Τα αίτια αυτά όμως, θα πρέπει να επικεντρώνονται σε σημεία που επενεργούν ταυτόχρονα σε όλα τα προς χρήση node και σε όλους τους προς χρήση πυρήνες. Τα μόνα που πληρούν αυτές τις προϋποθέσεις, είναι το δίκτυο μεταξύ των υπολογιστών και ο κοινόχρηστος δίσκος SAN. Προσπαθώντας να εντοπίσουμε το πρόβλημα μπορούμε να θεωρήσουμε ότι το ποσοστό που αναλογεί στον κοινόχρηστο δίσκο SAN είναι μικρό. Αυτό συμβαίνει γιατί ο τρόπος λειτουργίας, το μέγεθος του εύρους του δικτύου οπτικών ινών fibrechannel και η χρήση δύο συνδέσεων σε κάθε node προς τον κοινόχρηστο δίσκο, έχουν σχεδιαστεί ώστε να έχουν ελάχιστη καθυστέρηση και μέγιστη ταχύτητα διαμεταγωγής δεδομένων από δεκάδες χρήστες ταυτόχρονα. Παράλληλα, η κάλυψη του κοινόχρηστου δίσκου από δεδομένα την στιγμή των πειραμάτων, ήταν περίπου στο 30%, πολύ μικρό ποσοστό για να έχει σαν επίπτωση κάποια σημαντική πτώση στην ταχύτητα εγγραφής και ανάγνωσης πράγμα που οδηγεί στο δίκτυο των υπολογιστών.

Η ονομαστική ταχύτητα του δικτύου Ethernet είναι 1Gbps. Αυτό σημαίνει ότι, μπορεί πρακτικά να δουλεύει στο 50% της ονομαστικής του ταχύτητας, σε full duplex mode. Την ίδια στιγμή, η κίνηση των δεδομένων αυξάνει με εκθετικό ρυθμό για κάθε node που χρησιμοποιείται. Αυτό συμβαίνει γιατί έχουμε μαζική μετακίνηση δεδομένων προς τον ελεγκτή -controller της παράλληλης επεξεργασίας, ταυτόχρονα με τα σήματα ελέγχου του cluster και τον όγκο τον δεδομένων που θα εγγραφούν στον σκληρό δίσκο. Ιδίως για το τελευταίο, η συγκέντρωση δεδομένων

και η αποστολή τους για εγγραφή στον δίσκο NAS, απαιτεί πολλά πακέτα και τεράστιο όγκο δεδομένων που μπορεί να φτάσει τα δεκάδες MB ανά δευτερόλεπτο. Έτσι, βλέπουμε ότι αγγίζουμε τα όρια του hardware που έχουμε διαθέσιμο.

Μερικώς όμοια συμπεριφορά με τους χρόνους περατώσεως επίλυσης με APDL στη Συστοιχία -B, παρατηρείται στην περίπτωση του LS-DYNA, στην Συστοιχία -B. Αυτό που είναι ορατό στην καμπύλη της Εικόνας 66, είναι ότι μετά τους 80 πυρήνες έχουμε σταθεροποίηση των χρόνων περατώσεως επίλυσης. Δηλαδή, σημαίνει πως, παρ' όλη την αύξηση της συνολικής επεξεργαστικής ισχύος, για λόγους προφανώς παρόμοιους με αυτούς που είχαμε στην περίπτωση των 48 και 92 πυρήνων με APDL, δεν έχουμε μείωση των χρόνων αυτών. Η διαπίστωση αυτή, μας βοηθάει να επαληθεύσουμε την παρατήρηση στην προηγούμενη παράγραφο, ότι δεν είναι τυχαίο φαινόμενο. Παράλληλα όμως, μπορούμε να δούμε ότι υπάρχει διαφορετική προσέγγιση στην χρήση και εκμετάλλευση των πόρων, μεταξύ LS-DYNA και της εφαρμογής του ANSYS, APDL, η οποία έχει σαν επιλογή τους επιλύτες -solvers του LS-DYNA.

Η χρήση του προσωπικού ηλεκτρονικού υπολογιστή έγινε για λόγους επαλήθευσης των αποτελεσμάτων. Σαν τεχνολογία, το επίπεδό του βρίσκεται σε αυτό της Συστοιχίας -B. Ανήκουν στην ίδια γενιά hardware, οπότε θα έπρεπε οι επιδόσεις σαν συστήματα να είναι παρόμοιες, αφού έχουν την ίδια υπολογιστική ισχύ ανά πυρήνα, παρ' όλο που το πλήθος των μετρήσεων είναι μικρό. Πάραυτα, η τάση που βλέπουμε είναι αρκετή για να εξαχθούν αποτελέσματα. Η σύγκλιση των χρόνων είναι ορατή και η μικρή απόκλιση οφείλεται σε επιμέρους ιδιαιτέρα χαρακτηριστικά, όπως είναι η ταχύτητα του σκληρού δίσκου. Όμως, το αποτέλεσμα που περιμένουμε το βλέπουμε, δηλαδή ότι οι χρόνοι είναι παρόμοιοι. Άρα, μπορούμε να εξάγουμε συμπεράσματα για το πόσο επηρρεάζουν άλλα χαρακτηριστικά ενός υπολογιστικού συστήματος, όπως η ταχύτητα του σκληρού δίσκου.

Αποκλίσεις παρουσιάζονται μεταξύ του APDL και του LS-DYNA, για την επίλυση του ίδιου μοντέλου πεπερασμένων στοιχείων, ενώ θεωρητικά για το ίδιο πρόβλημα θα έπρεπε να δαπανάται περίπου ο ίδιος χρόνος και στις δύο εφαρμογές. Αυτό που παρατηρούμε όμως είναι διαφορετικό και για να γίνει πιο ορατό, παρατίθενται οι χρόνοι περάτωσης επίλυσης για τους κοινούς αριθμούς πυρήνων, για APDL και LS-DYNA.

Πυρήνες	APDL (HH:mm:ss)	LS-DYNA (HH:mm:ss)
8	10:01:38	17:36:59
16	07:16:08	09:15:09
24	05:03:47	06:17:20
32	04:25:05	05:04:38
40	03:54:35	04:16:05
48	03:41:07	03:48:48
56	03:26:34	03:13:51
64	03:27:14	02:56:13
80	03:28:29	02:46:36
88	03:32:39	02:52:22
96	03:40:01	02:48:57

Το παράδοξο σε αυτή την παρατήρηση είναι, ότι ενώ μέχρι τους 56 πυρήνες έχουμε προβάδισμα (μικρότερους χρόνους περατώσεως επίλυσης) από το APDL, από τους 56 πυρήνες και μετά βλέπουμε αυτό να αντιστρέφεται προς όφελος του LS-DYNA. Χρησιμοποιώντας παρόμοιο συλλογισμό με τις προηγούμενες παρατηρήσεις, προσπαθούμε να εστιάσουμε στις διαφοροποιήσεις μεταξύ του APDL και του LS-DYNA. Στο επίπεδο του hardware, δεν υπάρχει καμία διαφοροποίηση, άρα το αποκλείουμε. Οπότε θα πρέπει να κατευθυνθούμε στις διαφοροποιήσεις στο software. Σε επίπεδο του λογισμικού, υπάρχουν μερικά σημεία που μπορούμε να στηριχθούμε για να ερμηνεύσουμε αυτήν την διαφορά. Αρχικά, είναι το ίδιο το software προσομοίωσης πεπερασμένων στοιχείων. Το APDL λειτουργεί έχοντας γραφικό περιβάλλον, ενώ το LS-DYNA λειτουργεί αποκλειστικά σε κονσόλα. Είναι προφανές ότι υπάρχει διαφορετική προσέγγιση στον τρόπο που γίνεται η διαχείριση των πόρων και στο πλήθος των δεδομένων που εγγράφονται στον σκληρό δίσκο ή και που μετακινούνται μεταξύ των πυρήνων. Το δεύτερο σημείο, στο επίπεδο του λογισμικού, είναι το πρόγραμμα MPI για την παράλληλη επεξεργασία που χρησιμοποιείται. Στο APDL γίνεται χρήση του PLATFORM MPI της IBM, ενώ στο LS-DYNA χρησιμοποιήσαμε το OpenMPI. Οι κατασκευαστές του PLATFORM MPI, ισχυρίζονται στην ιστοσελίδα τους (http://www-03.ibm.com/systems/platformcomputing/products/mpi/), ότι είναι το πιο αποδοτικό MPI. Ίσως αυτή να είναι και η αδυναμία του. Αφού εκμεταλλεύεται στο έπακρο τους πόρους του συστήματος, όταν εμπλέκονται πολλά node, τότε το πλήθος των ενεργειών που εκτελεί θα είναι τόσο μεγάλο, που θα έχει το αντίθετο αποτέλεσμα. Εκείνη την στιγμή κάποιος πόρος θα παρουσιάσει

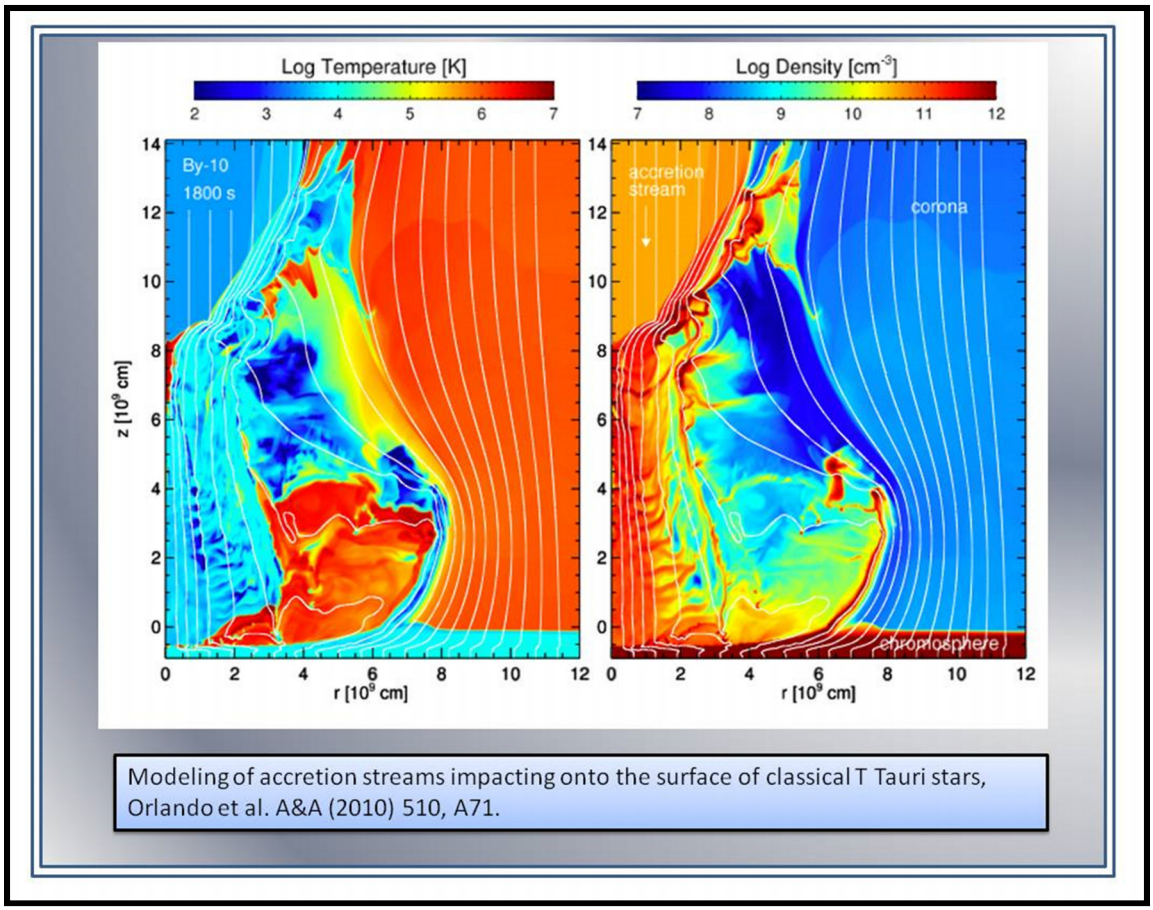
κορεσμό και θα έχουμε μείωση αντί για αύξηση της ταχύτητας. Αντίθετα, το ελεύθερο OpenMPI σε συνδυασμό με το LS-DYNA σε συνθήκες υψηλών απαιτήσεων και πάρα πολλών πυρήνων, έχει καλύτερη απόδοση και σε κατάσταση κορεσμού των πόρων -bottleneck κερδίζουμε περίπου μία ώρα ή περίπου 30% στο χρόνο περάτωσης επίλυσης, για το ίδιο μοντέλο πεπερασμένων στοιχείων.

4.4. Προσομοιώσεις στα cluster σε προβλήματα μαγνητο-υδροδυναμικής MHD στον PLUTO και η απόδοση της παράλληλης επεξεργασίας

Θέλοντας να μελετήσουμε την παράλληλη επεξεργασία MHD μοντέλων, έγινε χρήση του λογισμικού MHD προσομοιώσεων PLUTO. Η μεθοδολογία εργασίας ήταν παρόμοια με την αυτή που έγινε με το μοντέλο FEM 3 car crash. Η επίλυση έγινε και στις 2 Συνθέσεις που διαθέτουμε, ενώ οι χρόνοι κρατήθηκαν σε αρχείο που μετά το πέρας των απαιτούμενων περιπτώσεων εισήχθησαν στο περιβάλλον Origin, όπου και έγινε η δημιουργία των απαιτούμενων γραφικών παραστάσεων και η ανάλυσή τους. Το PLUTO εξάγει τα αποτελέσματα σε αρχεία τύπου *.vtk. Αυτά, συνήθως, έχουν μέγεθος συνολικά μερικών δεκάδων Gb. Το επόμενο βήμα είναι να εισαχθούν σε περιβάλλον ανάλυσης, όπως είναι το VISIT του Lawrence Livermore National Laboratory. Από εκεί μπορούμε να πάρουμε την οπτικοποίηση των αποτελεσμάτων, σαν αυτό που βλέπουμε στην Εικόνα 69 [33].

Αξιοσημείωτο για το PLUTO, είναι ότι πριν την προσομοίωση, το πρόβλημα μπορεί να γίνει compile για σειριακή ή παράλληλη επεξεργασία, βελτιστοποιώντας την συμπεριφορά του προβλήματος και του λογισμικού στις συνθήκες επίλυσης. Ταυτόχρονα, παρόλο που για να εκκινήσει η επίλυση γίνεται κλήση του OpenMPI, το PLUTO διατηρεί την διαχείριση των πυρήνων και την μέθοδο που θα διαμοιραστεί το πρόβλημα σε αυτούς. Εκτός από την παράλληλη επεξεργασία, μπορεί πριν γίνει compile να ρυθμιστούν αρκετές συνθήκες, μεταβλητές και παράμετροι. Σε όλα αυτά πρέπει να προστεθεί ένα σύνηθες πρόβλημα που παρουσιάζεται σε αυτό το software. Είναι πολύ εύκολο, κατά την διάρκεια της προσομοίωσης, οι διαστάσεις του πλέγματος επίλυσης να πάρουν αρνητικές τιμές. Μόλις γίνει αυτό, πραγματοποιείται βίαιος τερματισμός του προγράμματος. Όμως δεν βγαίνει κάποιο μήνυμα λάθους από το PLUTO, αλλά εμφανίζεται μήνυμα

σφάλματος από το λειτουργικό σύστημα για χρήση απαγορευμένου τμήματος της μνήμης.



Εικόνα 69. Επίλυση προβλήματος PLUTO.

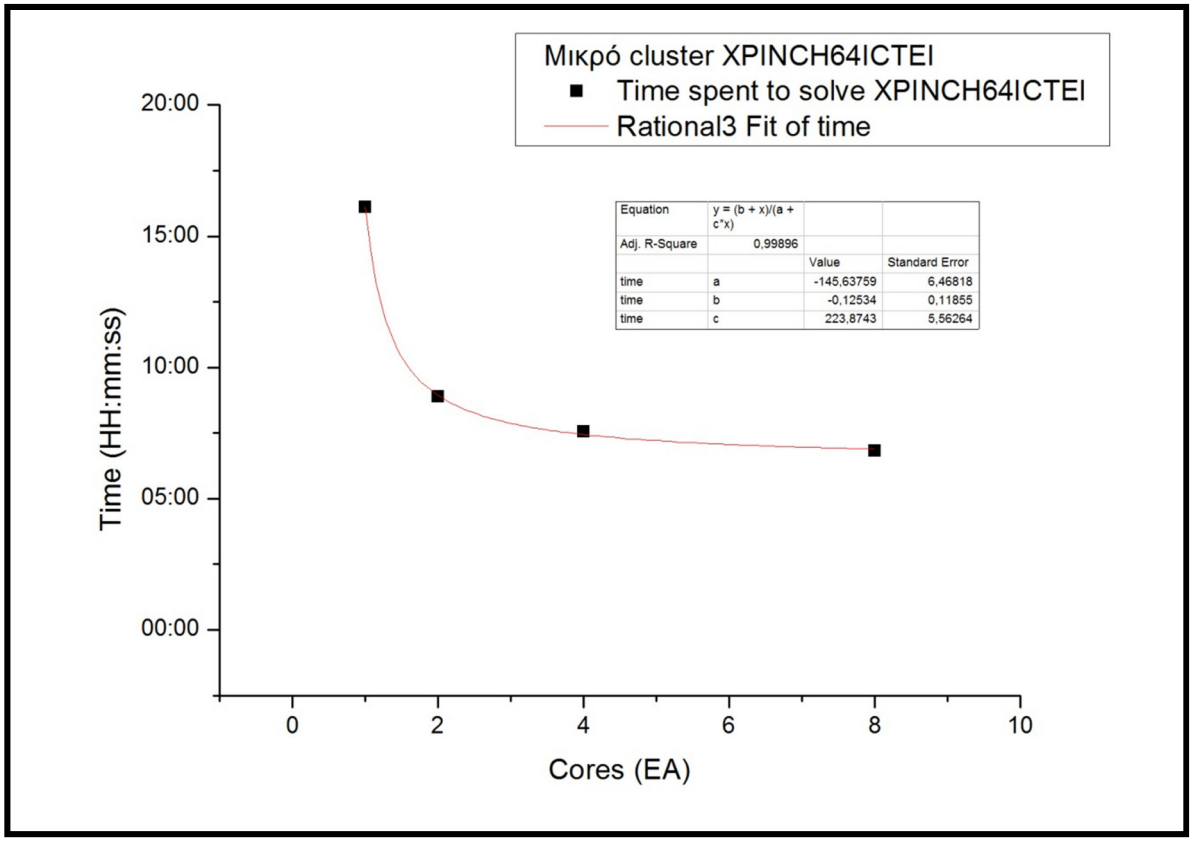
Το μοντέλο που επιλέχθηκε για τις δοκιμές είναι το xpinch64 και αφορά τμήμα προσομοιώσεων διδακτορικής διατριβής σε συμπεριφορά πλάσματος παραγόμενου σε πειράματα Pinch, του κ. Γ. Κουνδουράκη. Αφού συγκεντρώσουμε τα απαραίτητα δεδομένα, τα οποία στην περίπτωσή μας είναι οι χρόνοι περατώσεως και η μέγιστη RAM ανά πυρήνα, τα εισάγουμε στο Origin. Εκεί κάνουμε επεξεργασία των δεδομένων, κατασκευάζουμε τα γραφήματα και τις καμπύλες προσαρμογής.

4.4.1 PLUTO MHD στη Συστοιχία -A

Στην περίπτωση της Συστοιχίας -A, τα δεδομένα αυτά είναι λίγα. Ο λόγος είναι η τεράστια ποσότητα δεδομένων που αποθηκεύονται κατά την διάρκεια της προσομοίωσης και επίλυσης του προβλήματος. Το αποτέλεσμα ήταν, πως όταν γινόταν χρήση τρίτου node, δηλαδή χρησιμοποιούσαμε πάνω από 8 πυρήνες, το cluster κατέρρεε, γιατί ο κοινόχρηστος δίσκος NFS, δεν είχε αρκετή ταχύτητα να διαχειριστεί τα δεδομένα και παράλληλα το δίκτυο έφτανε στα όρια του κορεσμού. Σε όλες τις προσομοιώσεις, το μοντέλο έγινε compile για παράλληλη επεξεργασία με OpenMPI.

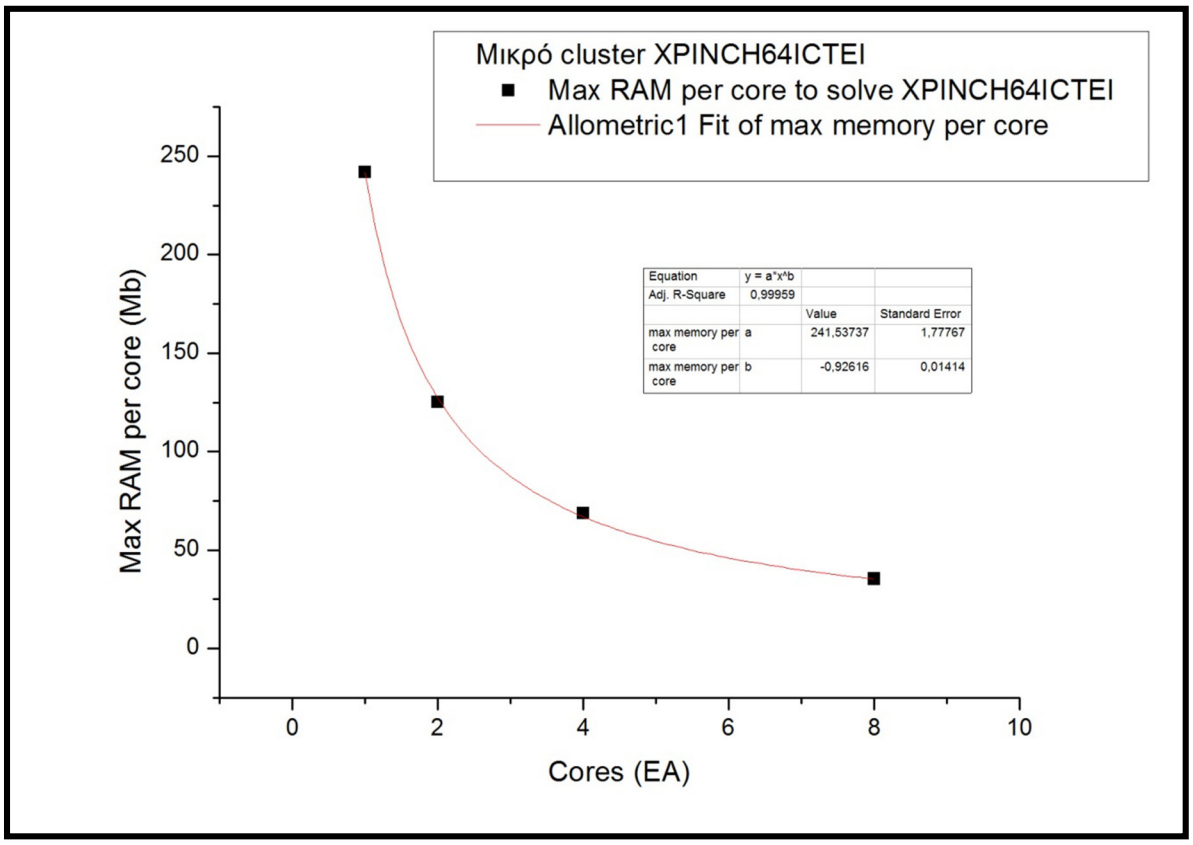
Αποτελέσματα χρόνων επίλυσης σε σχέσης με τους χρησιμοποιούμενους πυρήνες:

Πυρήνες	Χρόνος επίλυσης (mm:ss)	Μέγιστη RAM ανά πυρήνα (MB)
1	16:15	242,06
2	8:54	125,19
4	7:33	68,53
8	6:49	35,44



Εικόνα 70. Συστοιχία -A, xpinch64, time - cores.

Στο διάγραμμα της Εικόνας 70 μπορεί να παρατηρηθεί ότι η αύξηση του αριθμού των πυρήνων που παρέχονται στον επιλύτη, δεν έχει γραμμική επίδραση, αλλά της μορφής $\mathbf{y = \frac{b+x}{a+cx}}$. Η προσαρμογή είναι εξαιρετικά καλή, όπως μας δείχνει ο συντελεστής προσαρμογής R2, ο οποίος είναι 0,99896, ενώ το σφάλμα είναι σχετικά μικρό. Υπάρχει διαφοροποίηση, με την σχέση αριθμού πυρήνων – χρόνου περατώσεως επίλυσης, σε σχέση με όλες τις προηγούμενες αντίστοιχες σχέσεις στις μετρήσεις με το 3 car crash. Όμως, ο πάρα πολύ μικρός αριθμός δειγμάτων δεν μας επιτρέπει να έχουμε ασφαλή συμπεράσματα.



Εικόνα 71. Συστοιχία -A, xpinch64, RAM per core - cores.

Στο διάγραμμα της Εικόνας 71, μπορεί να παρατηρηθεί ότι η αύξηση του αριθμού των πυρήνων που παρέχονται στον επιλύτη, δεν έχει γραμμική επίδραση στην μείωση του μέγιστου μεγέθους μνήμης RAM, αλλά της μορφής $\mathbf{y = ax^b}$. Η προσαρμογή είναι εξαιρετικά καλή, όπως μας δείχνει ο συντελεστής προσαρμογής R2, ο οποίος είναι 0,99959, ενώ το σφάλμα είναι μικρό. Υπάρχει διαφοροποίηση, με την σχέση αριθμού πυρήνων – μέγιστου μεγέθους μνήμης RAM, σε σχέση με

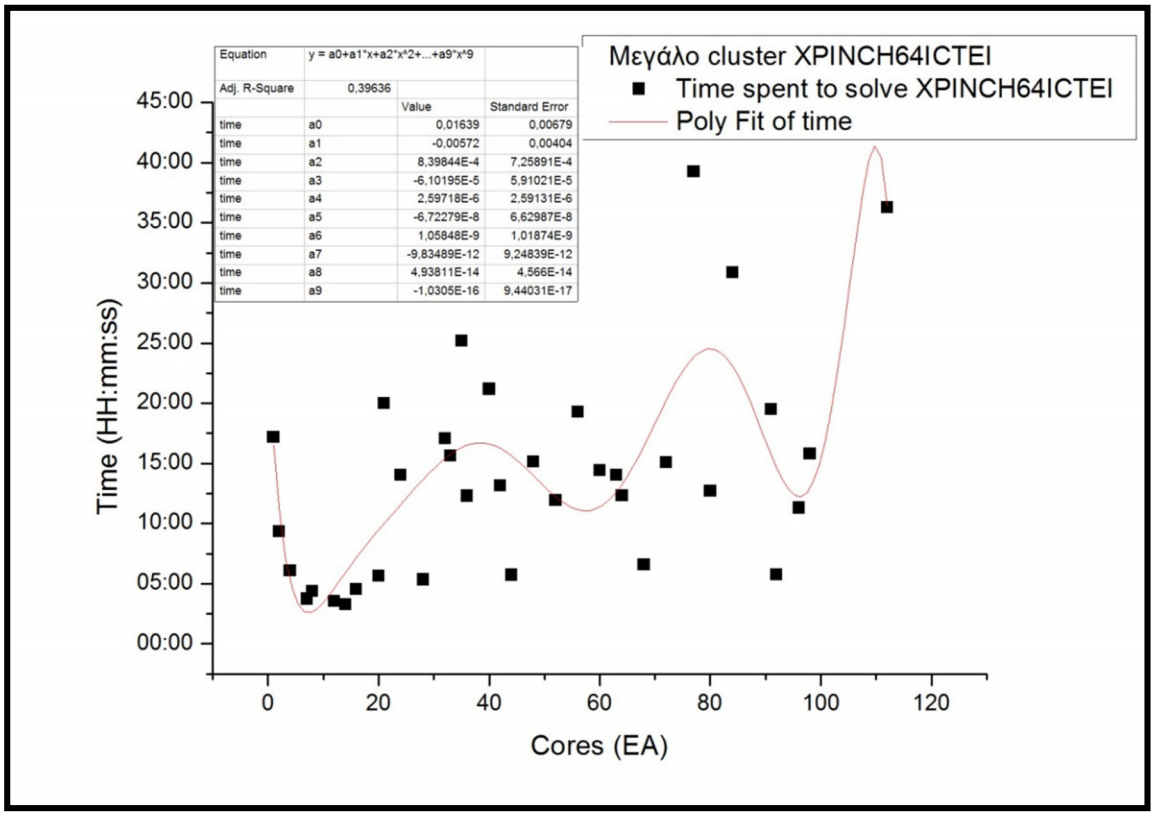
όλες τις προηγούμενες αντίστοιχες σχέσεις στις μετρήσεις με το 3 car crash. Όπως και στο προηγούμενο γράφημα, ο πάρα πολύ μικρός αριθμός δειγμάτων δεν μας επιτρέπει να έχουμε ασφαλή συμπεράσματα. Παρατηρούμε όμως διαφορετική συμπεριφοράς σε σχέση με το ANSYS APDL και το LS-DYNA. Μπορούμε κάνουμε διάφορες υποθέσεις, όπως ότι οφείλεται στη μειωμένη απαίτηση σε πόρους, εκτός από την ταχύτητα εγγραφής δεδομένων στον σκληρό δίσκο NFS και από την ταχύτητα του δικτύου. Η ταχύτητα μεταγωγής δεδομένων στο δίκτυο αφορά μόνο δεδομένα για εγγραφή στον δίσκο NFS, και όχι για ανταλλαγή δεδομένων μεταξύ των πυρήνων. Όμως, όλα αυτά παραμένουν εικασίες μέχρι να εξεταστεί η περίπτωση της Συστοιχίας -B, η οποία έχει τον απαραίτητο αριθμό δειγμάτων για να εξαχθούν ασφαλή συμπεράσματα.

4.4.2 PLUTO MHD στη Συστοιχία -B

Έχοντας εργαστεί με παρόμοιο τρόπο με την Συστοιχία -A, θα περιμέναμε τα αποτελέσματα και οι καμπύλες πυρήνων – χρόνων περατώσεως επίλυσης και πυρήνων – μέγιστης ποσότητας RAM ανά πυρήνα, να κυμαίνονται σε παραπλήσια επίπεδα. Αυτό όμως δεν συμβαίνει, οπότε και θα ανατρέξουμε σε κάθε στοιχείο του hardware, για να ερμηνευτούν τουλάχιστον οι μη αναμενόμενες, στην καμπύλη πυρήνων – χρόνων περατώσεως επίλυσης, μετρήσεις που έχουν ληφθεί. Υπήρχαν περιπτώσεις που δεν μπορούσε το σύστημα να μας δώσει χρόνο και δεν λήφθηκαν υπ’ όψιν, ενώ σε δύο περιπτώσεις οι χρόνοι που πήραμε ήταν υπερβολικά μεγάλοι και απορρίφθηκαν. Σε όλες τις προσομοιώσεις, το μοντέλο έγινε compile για παράλληλη επεξεργασία με OpenMPI.

Αποτελέσματα χρόνων επίλυσης σε σχέσης με τους χρησιμοποιούμενους πυρήνες:

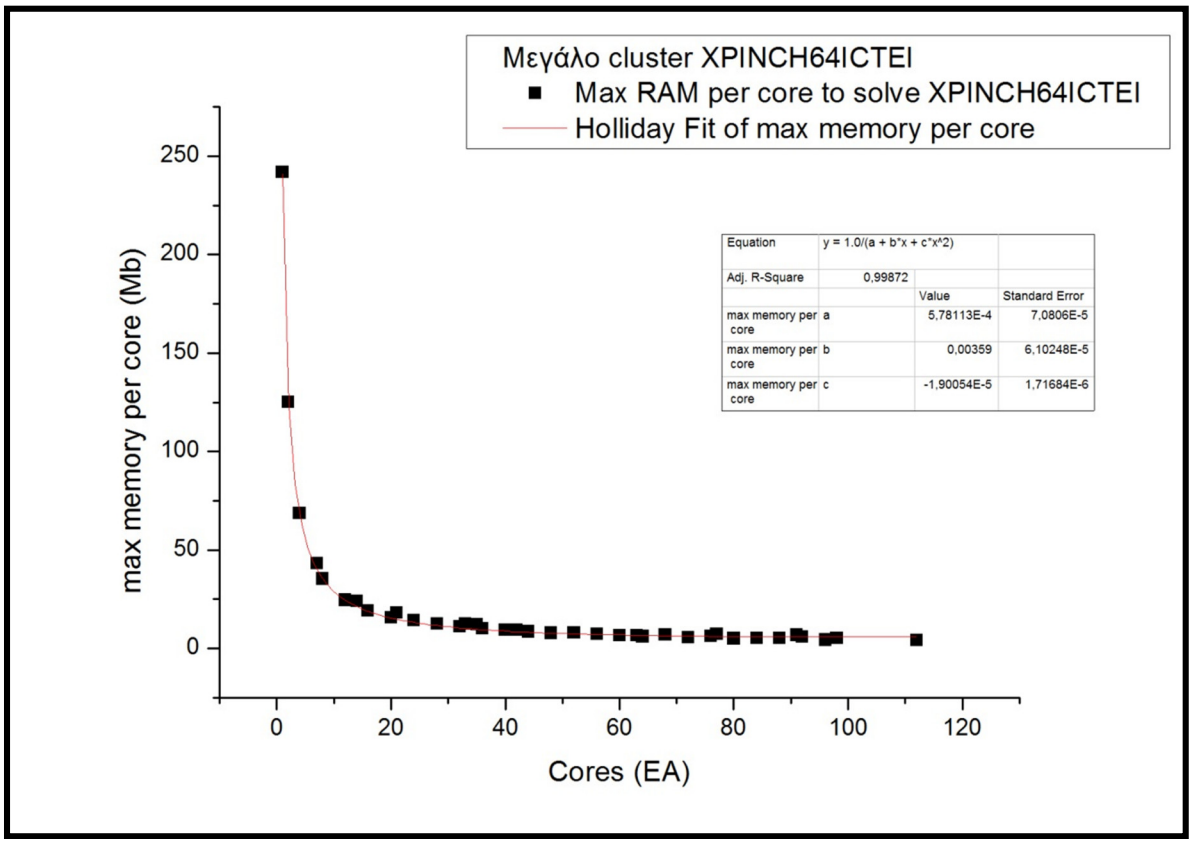
Πυρήνες	Χρόνος επίλυσης (HH:mm:ss)	Μέγιστη RAM ανά πυρήνα (MB)	
1	00:17:38	242,06	
2	00:09:21	125,19	
4	00:06:04	68,53	
7	00:03:44	43,38	
8	00:04:23	35,44	
12	00:03:34	24.48	
14	00:03:16	23.91	
16	00:04:33	18.99	
20	00:05:38	15.7	
21	00:20:00	18.06	
24	00:14:04	14.33	
28	00:05:21	12.41	
32	00:17:04	11.23	
33	00:15:38	12.48	(blade5)
33	00:15:43	12.48	(blade16)
35	00:25:12	12.22	
36	00:12:18	9.9	
40	00:21:11	9.31	
42	00:13:08	9.40	
44	00:05:44	8.58	
48	00:15:08	7.77	
49	Υπερβολικά πολύς χρόνος	-----	
52	00:11:57	8.03	
56	00:33:00	7.39	
60	00:14:25	6.45	
63	00:14:04	6.51	
64	00:12:19	6.04	
68	00:06:35	6.93	
70	Υπερβολικά πολύς χρόνος	-----	
72	00:15:07	5.57	
76	03:45:06	6.38	
77	00:39:17	7.17	
80	00:12:43	5.02	
84	00:30:54	5.13	
88	01:40:24	5.15	
91	00:19:31	6.73	
92	00:05:47	5.83	
96	00:11:19	4.34	
98	00:15:49	5.38	
112	00:36:19	4	



Εικόνα 72. Συστοιχία -B, XPINCH64ICTEI, time - cores.

Στο παραπάνω διάγραμμα, Εικόνα 72, μπορεί να παρατηρηθεί πολύ μεγάλη διασπορά στα αποτελέσματα της σχέσης αριθμού πυρήνων – χρόνου περατώσεως επίλυσης. Αυτό δεν μας επιτρέπει να γίνει η προσαρμογή των δεδομένων με αποτελεσματικότητα. Η καλύτερη καμπύλη προσαρμογής, που τείνει να δώσει μια υποτυπώδη αναπαράσταση των δεδομένων έχει την μορφή του πολυωνύμου: $y = a + a_1x + a_2x^2 + \quad + a_9x^9$. Παρά τη χρήση ενάτου βαθμού ο συντελεστής προσαρμογής R2 είναι 0.39636, που φανερώνει ότι έχουμε ασθενή συσχέτιση των δεδομένων. Άρα, το φαινόμενο μπορεί να περιγραφεί γραφικά με μεγάλο βαθμό πιθανής απόκλισης. Οι μετρήσεις για τους 76 και 88 πυρήνες παραλείφθηκαν, γιατί ήταν εξαιρετικά μεγάλες και στατιστικά μπορούν να θεωρηθούν ως μη ικανοποιητικές. Επιπλέον, οι μετρήσεις για τους πυρήνες 49 και 70 δεν ολοκληρώθηκαν ποτέ, αν και έγιναν αρκετές επαναλήψεις για επίλυση με αυτό τον αριθμό πυρήνων. Παρατηρώντας τον τρόπο διασποράς δεδομένων στο διάγραμμα, μπορεί κάποιος να παρατηρήσει ότι υπάρχουν 2 ομαδοποιήσεις αυτών. Η πρώτη, είναι αυτή που δημιουργούν οι μετρήσεις από τους πυρήνες 1 εώς 20, 28, 44, 68 και 92. Η δεύτερη ομάδα, αποτελείται από όλες τις υπόλοιπες μετρήσεις ανά αριθμό πυρήνων.

Στο διάγραμμα της Εικόνας 73, παρατηρούμε τα αποτελέσματα και την καμπύλη της σχέσης αριθμού πυρήνων – Μέγιστης ποσότητας μνήμης RAM ανά πυρήνα. Η καμπύλη είναι αντίστροφη πολυονυμική της μορφής $\mathbf{y} = \frac{\mathbf{1}}{\mathbf{a+bx+cx^2}}$. Η προσαρμογή είναι εξαιρετικά καλή, όπως μας δείχνει ο συντελεστής προσαρμογής R2, ο οποίος είναι 0,99872 , ενώ το σφάλμα είναι πάρα πολύ μικρό. Η καμπύλη μπορεί σχεδόν απόλυτα να μας περιγράψει πως η αύξηση των πυρήνων σχετίζεται με την μείωση της μέγιστης ποσότητας μνήμης RAM ανά πυρήνα. Οι όποιες διαφοροποιήσεις, ή μικρό αυξήσεις, παρατηρούνται γύρω από τους πυρήνες που υπάρχει απότομη μεταβολή του χρόνου περατώσεως επίλυσης.



Εικόνα 73. Συστοιχία -B, XPINCH64ICTEI, RAM per core - cores.

4.5. Ανάλυση απόδοσης παράλληλης επεξεργασίας σε προβλήματα MHD

Με βάση την ανάλυση της Παραγράφου 4.3 βλέπουμε πως έχουμε μείωση της απαιτούμενης RAM ανά πυρήνα, η οποία ακολουθεί συγκεκριμένο πρότυπο. Συγκρίνοντας τις τιμές στη Συστοιχία -Α και στην Συστοιχία -Β βλέπουμε ότι είναι όμοιες. Συμπερασματικά, μπορούμε να ορίσουμε, ότι για την μνήμη RAM ισχύει η προσέγγιση που έχει γίνει στην Συστοιχία -Β, δηλαδή $\mathbf{y = \frac{1}{a+bx+cx^2}}$. Το αντίθετο συμβαίνει στην ανάλυση ως προς το χρόνο περατώσεως επίλυσης. Στα γραφήματα του αριθμού των πυρήνων – χρόνου περατώσεως επίλυσης παρατηρούμε, ότι μέχρι τους 20 πυρήνες, έχουμε ένα σαφές μοτίβο, και στις δύο συνθέσεις. Μετά όμως από αυτό το όριο υπάρχει πολύ μεγάλη διασπορά αποτελεσμάτων, εκτός από αυτά των πυρήνων 28, 44, 68 και 92. Μεταξύ τους φαίνεται ότι υπάρχει και κάποιου είδους μαθηματική σχέση με βήμα 14+N*10, όπου N είναι ο αριθμός του βήματος. Επιπροσθέτως, υπάρχουν περιπτώσεις, που για συγκεκριμένο αριθμό πυρήνων δεν μπορεί να επιλυθεί το μοντέλο, 49 και 70 πυρήνες, και περιπτώσεις που η επίλυση για να επιτευχθεί απαιτήθηκε υπερβολικά πολύς χρόνος, π.χ. 78 και 88 πυρήνες.

Το μοντέλο με το οποίο μελετούμε την μέγιστη ποσότητα μνήμης RAM ανά πυρήνα, είναι πολύ ισχυρό και ακριβές. Παρ’ όλα αυτά, εμφανίζονται κάποιες μικρές ανακολουθίες κοντά στον αριθμό των πυρήνων που έχουμε τις μεγαλύτερες μεταβολές στον χρόνο περατώσεως επίλυσης. Παράλληλα, βλέπουμε ένα μοτίβο στους πυρήνες 1 έως 20, 28, 44,68, και 92, όπως έχει αναφερθεί παραπάνω. Έχοντας υπ’ όψιν τα φαινόμενα που εμφανίζονται ως προς το hardware σε όλη την ανάλυση των παρατηρήσεων για το μοντέλο 3 car crash, καθώς και το μέγεθος των πληροφοριών που αποθηκεύονται στον δίσκο που απαιτεί η επίλυση με το PLUTO, μπορούμε να εξάγουμε κάποια συμπεράσματα. Το PLUTO διαχειρίζεται αυτόνομα την μέθοδο, με την οποία το πρόβλημα διανέμεται στους πυρήνες σε επιμέρους υπο-προβλήματα. Αυτό συμπεριλαμβάνει, τον αριθμό των επιπέδων κατάτμησης και ανασύνθεσης του μοντέλου, καθώς και την ισορροπία μεταξύ αυτών, μέσω του όγκου της απαιτούμενης επεξεργαστικής ισχύος για κάθε υπό-πρόβλημα. Οπότε, όσο πιο μικρό τμήμα έχει να εκτελέσει ο κάθε πυρήνας, τόσο

μεγαλύτερη η πιθανότητα να απαιτεί πολύ λιγότερη ή πολύ περισσότερη επεξεργαστική ισχύ από τον μέσο όρο, άρα και αντίστοιχα αλλαγμένο χρόνο για να περατώσει το έργο του. Αυτό θα προκαλέσει αντίστοιχη καθυστέρηση, όταν θα υπάρχουν συνθήκες απόκλισης από τον μέσο όρο χρόνου επίλυσης. Άρα, όταν ξεπεράστηκε το όριο των 20 πυρήνων και πλέον, ο κάθε πυρήνας εκτελεί πιο μικρού φάσματος εργασία, ως προς το πρόβλημα. Μόνο όταν ισχύει το μοτίβο στους πυρήνες 1 έως 20, 28, 44,68, και 92 υπάρχει ισόποσος επιμερισμός του προβλήματος σε υπό-προβλήματα, δημιουργώντας έτσι δυσκολία στην επιλογή του βέλτιστου αριθμού πυρήνων. Αντίθετα, στην περίπτωση της μέγιστης ποσότητα μνήμης RAM ανά πυρήνα, έχουμε την δυνατότητα να εκτελέσουμε μοντέλα με τεράστιες απαιτήσεις σε μνήμη RAM. Τόσο μεγάλα που δεν μπορούμε να τα επιλύσουμε σε κανένα κοινό προσωπικό ηλεκτρονικό υπολογιστή.

4.6. Βελτίωση απόδοσης παράλληλης επεξεργασίας

Οι μέθοδοι βελτίωσης της απόδοσης των 2 cluster αφορούν βελτιστοποίηση στο hardware και στο software. Οι αναλύσεις που θα γίνουν θα αφορούν και τα δύο προβλήματα, 3 car crash και xpinch64, στα οποία βασίστηκε η μελέτη της συμπεριφοράς των cluster που δημιουργήσαμε, για παράλληλη επεξεργασία. Ο παράλληλος προγραμματισμός[34], η απόδοση του και η μελέτη του είναι ένα από τα αντικείμενα του νόμου Gustafson και του νόμου Amdahl[35], οι οποίοι, πολύ απλοποιημένα, μας δίνουν, ο καθένας, από μία σχέση για την αύξηση στην καθυστέρηση, λόγο της αύξησης του αριθμού των πυρήνων. Άρα η αύξηση στην πολυπλοκότητα του υλικού, έχει επίδραση στην απόδοση των συστοιχιών -A και –B.

4.6.1 Βελτίωση του hardware

Η βελτιστοποίηση στο hardware είναι το σημείο που μπορούμε να εφαρμόσουμε τις περισσότερες προσπάθειες για την πραγματοποίηση του στόχου μας, ο οποίος είναι η αύξηση της απόδοσης των δύο Συνθέσεών. Δυστυχώς, λόγω έλλειψης οικονομικών πόρων, ήταν αδύνατον να εφαρμοστούν τα περισσότερα από αυτά που θα χρειαζόταν, για να δούμε αποτελέσματα. Όσες ενέργειες

μπόρεσαν να υλοποιηθούν, έγιναν με κόστος την απώλεια της λειτουργικότητας κάποιου από τα node, είτε ολοκληρωτικά είτε μερικώς.

Βασικό σημείο αναφοράς είναι η μνήμη. Αυτή έχει δύο σκέλη, τη μνήμη αποθήκευσης, δηλαδή τους σκληρούς μας δίσκους και τη μνήμη τυχαίας προσπέλασης, δηλαδή τη μνήμη RAM. Έχοντας υπ' όψιν τις παρατηρήσεις από τα δύο μοντέλα που προσομοιώθηκαν, βρήκαμε την αιτία για μερικά από τα προβλήματά μας. Η ποσότητα της μνήμης RAM, η ταχύτητα των σκληρών δίσκων και λιγότερο η ταχύτητα της μνήμης RAM, είχαν σημαντικό ρόλο στις επιδόσεις ταχύτητας επίλυσης των μοντέλων.

Στην περίπτωση της Συστοιχίας -Α υπάρχει μικρή ποσότητα μνήμης RAM, 4 Gb DDR3, με αποτέλεσμα να γίνεται χρήση της μνήμης SWAP, στους σκληρούς δίσκους. Όμως, οι σκληροί δίσκοι ήταν αρκετά γρήγοροι και οι αναγκαίες λειτουργίες ελάχιστες. Το μόνο μέτρο που πάρθηκε, ήταν η εξαίρεση του ηλεκτρονικού υπολογιστή simulation1, που διαμοίραζε τις άδειες χρήσης, από την διαδικασία επίλυσης, ώστε να μην καταρρέει αυτή η λειτουργία και κάνει αυτόματο σταμάτημα επίλυσης το APDL και το LS-DYNA. Αντίθετα, στην Συστοιχία -Β, αν και η μνήμη RAM είναι περισσότερη (9Gb DDR2 ECC), όταν οποιοδήποτε node βρισκόταν στα όρια της εξάντλησης των πόρων και γινόταν εκτεταμένη χρήση της μνήμης SWAP, κατέρρεαν οι υπηρεσίες διαχείρησης του cluster που έτρεχαν σε αυτό. Αυτό οφειλόταν κατά μεγάλο ποσοστό στην χαμηλή ταχύτητα των σκληρών δίσκων των node, της Συστοιχίας -Β. Αρχικά έχαναν στην διασύνδεση με τον σκληρό δίσκο SAN. Το επόμενο στάδιο ήταν η αδυναμία επικοινωνίας με τον domain controller του cluster. Αυτόματα τότε κατέρρεε η παράλληλη επεξεργασία. Τέλος, λόγω της τεράστιου αριθμού πακέτων δεδομένων προς τον domain controller, για επανασύνδεση κατέρρεε όλο το cluster. Η λύση δόθηκε με την μεταφορά της μνήμης RAM από 2 node (blade13 και blade14) στο blade15, έχοντας σαν αποτέλεσμα την παύση χρήσης του blade13 λόγω έλλειψης μνήμης και τη χρήση του blade14 μόνο σαν domain controller me 1Gb μνήμης RAM. Έτσι όμως αποκτήσαμε ένα node στο οποίο μπορούμε να δοκιμάσουμε τα μοντέλα, χωρίς να κινδυνεύουμε να έχουμε βίαιο τερματισμό ή πάγωμα του συστήματος. Με αυτόν τον τρόπο μειώθηκαν κατά πολύ οι βίαιοι τερματισμοί των λογισμικών ή το κόλλημα του λειτουργικού συστήματος και στις δύο Συνθέσεις. Όμως, δεν πρέπει να ξεχνάμε ότι η άφθονη RAM είναι αναγκαία συνθήκη, ώστε να μην γίνεται χρήση

της μνήμης SWAP στους σκληρούς δίσκους, που οδηγεί σε ραγδαία μείωση της ταχύτητας.

Στο θέμα των σκληρών δίσκων τα πράγματα είναι απλά καθώς είναι το πιο αργό στοιχείο ενός σύγχρονου ηλεκτρονικού υπολογιστή. Με τις διαθέσιμες χωρητικότητες, αν υπάρχει η οικονομική δυνατότητα, η χρήση των δίσκων SSD (Solid State Drive), είναι μονόδρομος. Διαφορετικά, επιλέγουμε την λύση των περιστροφικών σκληρών δίσκων, πολύ υψηλής ταχύτητας περιστροφής ή SAS (Serial SCSI). Με αυτόν τον τρόπο, τα δεδομένα θα αποθηκεύονταν με πολύ μεγάλους ρυθμούς, φαινόμενο που μας προβλημάτισε στην περίπτωση του PLUTO. Αντίθετα, όταν θα έκανε το σύστημα χρήση της μνήμης SWAP, η πτώση της ταχύτητας θα ήταν αισθητά μικρότερη. Δυστυχώς, δεν ήταν δυνατό να γίνει οποιαδήποτε προσπάθεια βελτιστοποίησης σε αυτό το αντικείμενο, λόγω κόστους.

Ένα ακόμη σημείο που μπορεί να γίνει βελτίωση, είναι το δίκτυο Ethernet. Ο κορεσμός του αποτελούσε σημαντικό πρόβλημα. Στη Συστοιχία -B τα πράγματα είναι απλά. Το δίκτυο είναι ενσωματωμένο στο σασί που τοποθετούνται τα node, επομένως δεν υπήρχε περιθώριο παρέμβασης. Αντίθετα, στην Συστοιχία -A έγινε προσπάθεια αντικατάστασης των καλωδίων FTP, με νέα. Όμως, το αποτέλεσμα ήταν πενιχρό γιατί δεν υπήρχε η δυνατότητα για καλώδια FTP CAT6, στα οποία πραγματικά το δίκτυο θα απέδιδε τα μέγιστα. Η μόνη παρέμβαση στο δίκτυο Ethernet που επέδωσε ορατά αποτελέσματα, ήταν η αντικατάσταση της κάρτας δικτύου του node simulation, κατά τη διάρκεια την δημιουργίας της Συστοιχίας -A, που έλυσε πολλά προβλήματα δυσλειτουργίας.

Επιγραμματικά μόνο, θα πρέπει να αναφερθεί, ότι το βασικό στοιχείο που εκφράζει την ταχύτητα ενός υπολογιστικού συστήματος είναι ο κεντρικός επεξεργαστής (CPU). Όπως είναι πασιφανές στην σύγκριση μεταξύ Συστοιχίας -A και Συστοιχίας -B, η γενιά και η ταχύτητα του κεντρικού υπολογιστή έχει τεράστια επίπτωση στην απόδοση. Φυσικά, η αναβάθμιση της CPU, δεν είναι οικονομικά συμφέρουσα, αλλά η χρήση παρωχημένων συστημάτων θα έχει σημαντική επίπτωση στην απόδοση, με παράδειγμα τον προσωπικό ηλεκτρονικό υπολογιστή και την σχέση του με την Συστοιχία -B.

4.6.2 Βελτίωση του software

Στο software, οι παρεμβάσεις που μπορούν να οδηγήσουν σε βελτιστοποίηση είναι περιορισμένες γιατί δεσμευόμαστε από τον τρόπο λειτουργίας του λειτουργικού συστήματος και του λογισμικού προσομοίωσης. Μία κίνηση που έγινε, είναι η ρύθμιση μίας μεταβλητής συστήματος, για το LS-DYNA, η οποία ονομάζεται AUTO_MEMORY. Αυτή η μεταβλητή, έδωσε τη δυνατότητα στο LS-DYNA, να δεσμεύει αυτόματα όση μνήμη RAM απαιτεί για την επίλυση του κάθε μοντέλου. Αυτό, μας επέτρεψε να αποφύγουμε τον βίαιο τερματισμό του παραπάνω λογισμικού, λόγω ελλιπούς δεσμευμένης μνήμης RAM, αλλά είχε και σαν παρενέργεια ότι το LS-DYNA μπορεί να δεσμεύσει όλη τη μνήμη RAM και SWAP προκαλώντας την κατάρρευση ολόκληρου του συστήματος. Σπάνιο μεν, αλλά μας ανάγκασε να τοποθετηθεί παραπάνω μνήμη στο node blade15, θέτοντας εκτος λειτουργίας τα nodes blade13 και blade14.

Ένα άλλο σημείο πάνω στο οποίο έγινε προσπάθεια παρέμβασης, ήταν η χρήση του καταλληλότερου προγράμματος MPI. Το PLATFORM MPI εκμεταλλεύεται στο έπακρο τους πόρους και είναι ταχύτερο σε λίγους πυρήνες, αντίθετα το OpenMPI λειτουργεί καλύτερα σε περισσότερους πυρήνες, μειώνοντας το φαινόμενο bottle-neck. Δυστυχώς, η προσπάθεια απέτυχε, γιατί κανένα από τα 3 λογισμικά δεν επέτρεπε την χρήση άλλου λογισμικού MPI, από αυτό για το οποίο είχε σχεδιαστεί.

ΚΕΦΑΛΑΙΟ 5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα αποτελέσματα και οι παρατηρήσεις που συλλέξαμε, μπορούν να μας διευκολύνουν, να καταλήξουμε για τον τρόπο που θα δημιουργήσουμε, ένα σύστημα συστοιχίας ηλεκτρονικών υπολογιστών, για παράλληλη επεξεργασία, σε περιβάλλον LINUX. Δίνουν σαφείς πληροφορίες για το πώς θα εγκαταστήσουμε τα προγράμματα επεξεργασίας μοντέλων πεπερασμένων στοιχείων και μέσω του πειραματισμού, ποια βήματα ακολουθήσαμε ή θα έπρεπε να ακολουθήσουμε ώστε να σταθεροποιηθεί και να βελτιστοποιηθεί η λειτουργία του. Ταυτόχρονα εξάχθηκαν συμπεράσματα σύμφωνα με τα οποία, θα μπορούμε να έχουμε εκτίμηση για την συμπεριφορά και απόδοση των συστοιχιών ηλεκτρονικών υπολογιστών που δημιουργήσαμε, ανάλογα με το πρόβλημα.

Η ανάλυση των αποτελεσμάτων έδειξε την εμφάνιση του φαινομένου bottle-neck και στις δύο συνθέσεις, οδηγώντας στο συμπέρασμα πως η απόδοση, ιδίως αν θέλουμε να έχουμε μεγάλης κλίμακας παραλληλισμό, βασίζεται, εκτός από την ποσότητα, αλλά και στην ποιότητα του hardware που έχουμε. Αυτό το συμπέρασμα, ενισχύεται από τα ευρήματα της διαφοράς απόδοσης μεταξύ Συστοιχίας -Α και Συστοιχίας -Β, σε συνάρτηση με τις μετρήσεις, που κάναμε σε προσωπικό ηλεκτρονικό υπολογιστή. Άρα, η τεχνολογική εξέλιξη του διαθέσιμου υλικού είναι, ίσως, το πιο σημαντικό εργαλείο που μπορούμε να έχουμε στα χέρια μας. Ακόμα και εξειδικευμένοι ηλεκτρονικοί υπολογιστές, όπως είναι τα node της Συστοιχίας -Β, είναι λιγότερο αποδοτικά, από προσωπικούς ηλεκτρονικούς υπολογιστές με τελευταίας γενιάς hardware. Για να είμαστε πιο συγκεκριμένοι, η ποσότητα και η τεχνολογία της μνήμης RAM, η ταχύτητα των σκληρών δίσκων, η ταχύτητα του δικτύου και κυρίως η ταχύτητα και η τεχνολογία της κεντρικής μονάδας επεξεργασίας είναι τα πιο καίρια σημεία επιρροής στην επίδοση. Απαραίτητο στοιχείο πάντα είναι η βαθιά γνώση του διαθέσιμου hardware και η γνώση αρχιτεκτονικής υπολογιστών, η γνώση των διαθέσιμων τεχνολογιών και των δυνατοτήτων του υλικού για αναβάθμιση. Μόνο με αυτόν τον τρόπο, μπορεί να γίνει η κατάλληλη επιλογή λειτουργικού συστήματος και του αναγκαίου λογισμικού που θα επιλεχθούν. Η γνώση αυτή απαιτεί συνεχή ενημέρωση και ανανέωση πληροφοριών αλλά και συνεχή και χρόνια επαφή με το αντικείμενο.

Για την επιτυχημένη επιλογή λειτουργικού συστήματος πρέπει να υπάρχει σαφής εκτίμηση του hardware. Αυτή η λογική εφαρμόστηκε και στην παρούσα εργασία. Προχωρήσαμε μετά την ανάλυση του hardware στην ενδεδειγμένη επιλογή λειτουργικών συστημάτων, και ακολούθησε η εγκατάστασή τους δημιουργώντας από ανεξάρτητους ηλεκτρονικούς υπολογιστές ένα Beowulf cluster ηλεκτρονικών υπολογιστών, μία συστοιχία με σκοπό την παράλληλη επεξεργασία. Έγινε χρήση επίκαιρου και λιγότερο σταθερού λειτουργικού συστήματος (OpenSUSE), στη Συστοιχία -A, ενώ στην Συστοιχία -B σταθερότερου αλλά περιορισμένης ευελιξίας λειτουργικού (SLES 11 SP2), με την δυνατότητα όμως καλύτερης χρήση του hardware.

Για τη καλύτερη λειτουργία των συστημάτων μας, έγιναν αλλαγές και δοκιμές, οι οποίες μας ανάγκασαν να εμβαθύνουμε ακόμα περισσότερο, στην παράλληλη σχέση της μηχανής με το software που υπάρχει σε αυτή και το ζητούμενο αυτής της σχέσης, που είναι οι αποδοτικές προσομοιώσεις αριθμητικών προβλημάτων. Για την ανάλυση και την όσο το δυνατόν γενίκευση των αποτελεσμάτων, έγινε χρήση 3 ειδικών λογισμικών προσομοίωσης, ANSYS-APDL, LS-DYNA και PLUTO.

Για την επίτευξη της σταθερότητας του λογισμικού, στην Συστοιχία -A, απομονώθηκε το node simulation1, από την παράλληλη επεξεργασία, και του ανατέθηκε ο ρόλος του εξυπηρετητή αδειών χρήσης και κοινόχρηστοι σκληρού δίσκου NFS. Αυτό μείωσε κατά πολύ τον βίαιο τερματισμό των προγραμμάτων επίλυσης πεπερασμένων στοιχείων, αφού διαφορετικά, η εξάντληση του εύρους του δικτύου υπολογιστών, δεν επέτρεπε την έγκαιρη αποστολή των αδειών χρήσης στα node simulation2-5. Στην Συστοιχία -B έγινε κάτι πιο πολύπλοκο. Επειδή η έλλειψη πόρων οδηγούσε το cluster σε κατάρρευση, η οποία απαιτούσε αρκετές ώρες μέχρι να επανέλθει στην αρχική κατάσταση, έγινε μεταφορά μνήμης από 2 node (blade13, blade14) στο blade15, το οποίο απέκτησε 24Gb μνήμη RAM. Παράλληλα, στο node blade14 εγκαταστάθηκε 1GB μνήμης RAM και συνέχισε να λειτουργεί σαν domain controller των υπηρεσιών cluster της Συστοιχίας -B. Τα αποτελέσματα ήταν πολύ ικανοποιητικά, επιτρέποντας την βέλτιστη λειτουργία και την εκτέλεση πλήθους επιτυχών προσομοιώσεων.

Τέλος, είδαμε την διαφοροποίηση στην απόδοση μεταξύ δύο διαφορετικών λογισμικών MPI. Το PLATFORM MPI είναι πιο αποδοτικό και εκμεταλλεύεται στο

έπακρο τους διαθέσιμους πόρους, ενώ το OpenMPI παρουσιάζει καλύτερη συμπεριφορά και αποδοτικότητα, όταν έχουμε φτάσει ή ξεπεράσει, τα όρια των πόρων. Έτσι για μικρό αριθμό πυρήνων το PLATFORM MPI έχει καλύτερη απόδοση, ενώ όταν έχουμε πολλούς πυρήνες για παράλληλη επεξεργασία προτιμότερο είναι το OpenMPI.

Βιβλιογραφία

[1] Tianhe-2, http://www.scmp.com/lifestyle/technology/article/1263602/chinas-tianhe-2-supercomputer-unseats-us-titan-worlds-fastest, last accessed 5-2016.
[2] http://pixshark.com/beowulf-cluster-sterling.htm, last accessed 5-2016
[3] Beowulf cluster, http://en.wikipedia.org/wiki/Beowulf_cluster, last accessed 5-2016
[4] In search of clusters: the coming battle in lowly parallel computing, Gregory F. Pfister, Univ. of Carlifornia, Berkeley and IBM, Prentice-Hall, Inc. Upper Saddle River,1995
[5] How to Build a Beowulf: A Guide to the Implementation and Application of PC Clusters, William Stallings, Τζιόλα, 2011
[6] Παράδειγμα SAN – NAS, http://www.turbotekcomputer.com/resources/small-business-it-blog/bid/58074/Difference-Between-NAS-and-SAN-3-Considerations, last accessed 5-2016
[7] NAS, https://en.wikipedia.org/wiki/Network-attached_storage, last accessed 5-2016
[8] Οργάνωση & Αρχιτεκτονική Υπολογιστών, William Stallings, Τζιόλα, 2003
[9] Κόστος - απόδοση σκληρών δίσκων, http://slideplayer.com/slide/1517683/ , last accessed 5-2016
[10] Σκληρός δίσκος HDD, http://www.cheadledatarecovery.co.uk/how-do-hard-disk-drives-work/, last accessed 5-2016
[11] Σκληρός δίσκος SSD, http://i.i.cbsi.com/cnwk.1d/i/tim/2013/03/01/SSD.jpg, last accessed 5-2016
[12] http://electronics.stackexchange.com/questions/120737/purpose-of-outer-sheath-in-foiled-twisted-pair-network-cable, last accessed 5-2016
[13] http://www.infocellar.com/networks/fiber-optics/fiber.htm, last accessed 5-2016
[14] Αρχιτεκτονική των Υπολογιστών, Μια Δομημένη Προσέγγιση, Andrew S. Tanenbaum, «Κλειδάριθμος» 2000
[15] Σύγχρονα Λειτουργικά Συστήματα του Andrew S. Tanenbaum, «Κλειδάριθμος» 2004
[16] Unix και Linux http://torstefanshome.googlecode.com/svn/trunk/projects/unixPres/img,last accessed 5-2016
[17] Ιεραρχία λειτουργικού συστήματος http://www.mindpride.net/root/Extras/how-stuff-works/how_computer_memory_works.htm, last accessed 5-2016
[18] HP BladeSystem c7000 http://www8.hp.com/gr/el/products/enclosures/product-detail.html?oid=1844065 last accessed 5-2016
[19] High Availability https://www.suse.com/documentation/sle_ha/book_sleha/?page=/documentation/sle_ha/book_sleha/data/book_sleha.html last accessed 5-2016

[20] Κελύφη Linux (http://www.embeddedsystemonline.com/operating-system/linux/introduction-to-linux-unix/1-history-and-background) last accessed 5-2016
[21] ANSYS documents, http://www.ansys.com/staticassets/ANSYS/staticassets/support/platform-support-14.5-detailed-summary.pdf , last accessed 5-2016
[22] LS-DYNA documents, http://www.dynasupport.com/manuals, last accessed 5-2016
[23] xoopic, http://ptsg.egr.msu.edu/pub/codes/xoopic/oopic_manual/xoopic.html, last accessed 5-2016
[24] PLUTO, http://plutocode.ph.unito.it/ , last accessed 5-2016
[25] EPOCH, http://www.ccpp.ac.uk/epoch/epoch_user.pdf , last accessed 5-2016
[26] Πεπερασμένα στοιχεία Δρ Πασχάλης Κ. Γκότσης Εκδόσεις ΖΗΤΗ 2005
[27] Πεπερασμένα Στοιχεία, https://en.wikipedia.org/wiki/Finite_element_method last accessed 5-2016
[28] Ο αντιδραστήρας NIS, http://www.truegrid.com/femgallery.html pdf , last accessed 5-2016
[29] Αποτέσλεσμα ανάλυσης με LS-DYNA, http://www.ls-dyna.com/pages/images/613c.jpg , last accessed 5-2016
[30] Διαφορά στην πυκνότητα του δικτυώματος, http://www.truegrid.com/femgallery.html , last accessed 5-2016
[31] 3 car crash, http://www.lstc.com/ , last accessed 5-2016
[32] Origin, http://www.originlab.com/ , last accessed 5-2016
[33] Επίλυση προβλήματος PLUTO, http://plutocode.ph.unito.it/images/Diapositiva23.JPG , last accessed 5-2016
[34] Παράλληλος προγραμματισμός, https://en.wikipedia.org/wiki/Parallel_computing , last accessed 6-2016
[35] Νόμοι Gustafson & Amdahl, http://www.tem.uoc.gr/~vagelis/Courses/Scientific_Computing/Ch3_Theory.pdf , last accessed 6-2016