

Κεφάλαιο 1

Εισαγωγή

1.1 Οι Μεγάλες Σύγχρονες Επιστημονικές Προκλήσεις

Πέραν της παρατήρησης, της θεωρίας και του πειράματος, η σύγχρονη επιστήμη χαρακτηρίζεται και από την αριθμητική προσομοίωση, η οποία αποτελεί το σημαντικότερο εργαλείο της. Όπου η παρατήρηση και το πείραμα κρίνονται ως ανέφικτα ή ανεπαρκή προς επιβεβαίωση ή απόρριψη μιας θεωρητικής υπόθεσης η οποία αφορά ένα πολύπλοκο επιστημονικό πρόβλημα, η καταφυγή σε μεθόδους αριθμητικής προσομοίωσης αποτελεί σήμερα κοινή επιστημονική πρακτική.

Τα σύγχρονα πολύπλοκα επιστημονικά προβλήματα συνήθως ονομάζονται "μεγάλες προκλήσεις" (grand challenges). Η κατηγοριοποίησή τους, κατά τον Levin, είναι η ακόλουθη:

1. Κβαντοχημεία, στατιστική μηχανική και σχετικιστική φυσική.
2. Κοσμολογία και αστροφυσική.
3. Υπολογιστική δυναμική των ρευστών και τυρβώδης ροή.
4. Σχεδιασμός υλικών και υπεραγωγιμότητα.

5. Βιολογία, φαρμακολογία, γενετική μηχανική, αναδίπλωση πρωτεΐνης, δραστηριότητα ενζύμου και προτυποποίηση κυττάρου.
6. Ιατρική και προτυποποίηση ανθρωπίνων οργάνων και οστών.
7. Καιρός της υδρογείου και περιβαλλοντική προτυποποίηση.

Η αριθμητική προσομοίωση των μεγάλων προκλήσεων απαιτεί τεράστια υπολογιστική ισχύ, την οποία οι ακολουθιακοί υπολογιστές δεν διαθέτουν. Προς συνειδητοποίηση του γεγονότος αυτού, ας αναφερθούν τα ακόλουθα παραδείγματα:

1. Ερευνητές στο Oak Ridge National Laboratory των Η.Π.Α. χρησιμοποίησαν νόμους της κβαντομηχανικής για να αναπτύξουν ένα πρόγραμμα υπολογιστή, το οποίο έχει τη δυνατότητα πρόβλεψης των δομικών, ταλαντωτικών και ηλεκτρονικών ιδιοτήτων νέων υλικών. Η ολοκλήρωση ενός μόνον σχετικού υπολογιστικού πειράματος απαιτεί 150 tf (teraflops - τρισεκατομμύρια πράξεις κινητής υποδιαστολής).

2. Ένας γήινος δορυφόρος μπορεί να δημιουργήσει απαίτηση οργάνωσης, αποθήκευσης και ανάλυσης όγκου ψηφιακών δεδομένων ενός TByte (1 Terabyte – 1 τρισεκατομμύριο bytes) ανά ημέρα.

3. Η παραγωγή ειδικών εντυπώσεων στις κινηματογραφικές ταινίες απαιτούν χειρισμό 1,8 GBytes (gigabytes – δισεκατομμύρια bytes) οπτικών δεδομένων ανά δευτερόλεπτο.

Εκτιμάται ότι οι μεγάλες προκλήσεις, προκειμένου να αντιμετωπιστούν, απαιτούν υπολογιστές με:

- ταχύτητα: 1 TFlop/s (Teraflop/s – 10^{12} floating points operations per second - 1 τρισεκατομμύριο πράξεις κινητής υποδιαστολής ανά δευτερόλεπτο),
- κύρια μνήμη: 1 TByte,
- ταχύτητα εισόδου-εξόδου: 1 TByte/sec

κατ' ελάχιστο. Υπολογιστές τέτοιων προδιαγραφών έχουν ήδη σχεδιαστεί και κατασκευαστεί με βάση την έννοια της παραλληλίας, και ως εκ τούτου καλούνται παράλληλοι.

1.2 Εξέλιξη της Παράλληλης Επεξεργασίας Δεδομένων

Η έννοια της παραλληλίας, την οποία υπό μορφή αξιώματος ενσωμάτωσε ο Ευκλείδης στην ομώνυμη γεωμετρία του, είναι πρωτογενής. Σε οποιαδήποτε κλίμακα μεταξύ της κοσμολογίας και της κβαντομηχανικής, ο κόσμος συμπεριφέρε-

ται ως οργανωμένος μέσα σε τρεις διαστάσεις χώρου και μια διάσταση χρόνου. Τα γεγονότα, με τα οποία συνυπάρχουν αντικείμενα και δραστηριότητες, συμβαίνουν στο χώρο και στο χρόνο. Είναι δυνατόν σε γεγονότα να συμβαίνουν στον ίδιο χώρο, το ένα μετά το άλλο μέσα στο χρόνο, δηλαδή κατ' ακολουθία, και εξίσου δυνατόν είναι σε γεγονότα να συμβαίνουν σε διαφορετικούς χώρους στον ίδιο χρόνο, δηλαδή παράλληλα. Κατά φυσική συνέπεια, οι υπολογιστικές μηχανές με δυνατότητες παράλληλης επεξεργασίας δεδομένων χαρακτηρίζονται ως παράλληλοι υπολογιστές.

Δύο πρώιμοι παράλληλοι υπολογιστές σχεδιάστηκαν από τον Daniel Slotnick στο Πανεπιστήμιο του Illinois: ο Solomon, ο οποίος κατασκευάστηκε από την Westinghouse Electric Company κατά τις αρχές της δεκαετίας του 1960, και ο ILLIAC IV, ο οποίος συναρμολογήθηκε από την Burroughs Corporation κατά τις αρχές της δεκαετίας του 1970.

Μέσα στη δεκαετία του 1970, δύο καλά τεκμηριωμένοι παράλληλοι υπολογιστές, ο C.mmp και ο Cm*, κατασκευάστηκαν στο Πανεπιστήμιο Carnegie-Mellon. Στις αρχές της δεκαετίας του 1980, ερευνητές στο Caltech κατασκεύασαν τον Cosmic Cube, τον πρόγονο των πολυ-υπολογιστών (multicomputers) οι οποίοι κατασκευάστηκαν από τις εταιρείες Ametek, Intel και nCUBE. Αλλά εμπορικοί παράλληλοι υπολογιστές, κατασκευασμένοι με μικροεπεξεργαστές, εμφανίστηκαν περί τα μέσα της δεκαετίας του 1980.

Η απόδοση των μικροεπεξεργαστών αυξήθηκε ταχύτερα από την απόδοση άλλων ειδών επεξεργαστών. Μέχρι τα μέσα της δεκαετίας του 1970, οι βασικές αρχιτεκτονικές πρόοδοι, όπως είναι η bit-παράλληλη μνήμη, η bit-παράλληλη αριθμητική, η κρυφή μνήμη (cache memory), οι δίαυλοι, η διαφυλλωμένη μνήμη (interleaved memory), η σωλήνωση εντολών, οι πολλαπλές λειτουργικές μονάδες, οι σωληνωμένες λειτουργικές μονάδες και η σωλήνωση δεδομένων, είχαν ήδη συμπεριληφθεί στους σχεδιασμούς υπερυπολογιστών. Έκτοτε, αύξηση της απόδοσης επεξεργαστή σήμαινε μείωση χρόνου του κύκλου εντολής. Αυτό έγινε ιδιαίτερα δύσκολο, εφόσον η ταχύτητα των ηλεκτρονικών κυκλωμάτων περιορίζεται από την ταχύτητα του φωτός. Κατ' αντιδιαστολή, οι πρώιμοι μικροϋπολογιστές, οι οποίοι δεν είχαν ενσωματώσει όλες τις αρχιτεκτονικές προόδους στο σχεδιασμό τους, είχαν μεγάλα περιθώρια εξέλιξης, η οποία ανελευκτικά οδήγησε σε εντυπωσιακές αποδόσεις αυτήν την κλάση των υπολογιστών.

Η σύγκλιση στη σχετική απόδοση μεταξύ μικροϋπολογιστών και παραδοσιακών υπερυπολογιστών κατέληξε στην ανάπτυξη εμπορικά βιώσιμων παράλληλων υπολογιστών, οι οποίοι αποτελούνται από δεκάδες, εκατοντάδες ή και χιλιάδες μικροεπεξεργαστές και λέγονται μαζικά παράλληλοι. Σε πλήρη αποδοτικότητα, μαζικά παράλληλοι υπολογιστές, όπως ο Paragon XP/S της Intel και ο CM-5

της Thinking Machines, υπερακοντίζουν την ταχύτητα των παραδοσιακών υπερυπολογιστών με έναν επεξεργαστή, όπως του Cray Y/MP και του SX-3 της NEC.

Τον Δεκέμβριο του 1996 πραγματοποιήθηκε η ιστορική διάσπαση του υπολογιστικού φράγματος του 1 TFlop/s από τον υπερυπολογιστή Intel ASCI Teraflops. Αυτός ο μαζικά παράλληλος υπολογιστής είναι εγκατεστημένος στο Sandia National Laboratory των Η.Π.Α., στο New Mexico. Διαθέτει 9200 επεξεργαστές Pentium Pro 200MHz, 573 GBytes μνήμη συστήματος, και 2,25 TBytes αποθήκευση δίσκου. Καταναλώνει 850 Kilowatts κατά μέγιστο, ζυγίζει 44 τόνους, απαιτεί 300 τόνους ψυκτικού εξοπλισμού, περιλαμβάνει 86 μικρούς θαλάμους, οι οποίοι καταλαμβάνουν έκταση 1728 τετραγωνικών ποδών, και η διασύνδεση μεταξύ αυτών των θαλάμων και των κόμβων του συστήματος επιτυγχάνεται με καλώδια μήκους δύο μιλίων. Συμπληρώνει 40 δισεκατομμύρια υπολογισμούς σε ένα πεντηκοστό του δευτερολέπτου, και μια χρήση του είναι η επιβεβαίωση της ασφάλειας, της αξιοπιστίας και της αποτελεσματικότητας της πυρηνικής αποθήκης των Η.Π.Α. μέσω αριθμητικής προσομοίωσης αντί πυρηνικών δοκιμών. Το υπολογιστικό αυτό σύστημα χρησιμοποιήθηκε ήδη για την εκτίμηση του αποτελέσματος της πρόσκρουσης ενός κομήτη πλάτους ενός χιλιομέτρου στον Ατλαντικό Ωκεανό. Την 11η Ιουνίου 1997, η Intel Corporation ανήγγειλε ότι, με τη χρήση του βιομηχανικού προτύπου μεθόδου μέτρησης Linpack, ο υπερυπολογιστής της αυτός επέτυχε απόδοση 1,34 TFlop/s.

Την 28η Οκτωβρίου 1998, η IBM παρέδωσε στην κυβέρνηση των Η.Π.Α. έναν υπερυπολογιστή της, ο οποίος έχει δυνατότητα εκτέλεσης 3,9 TFlop/s, είναι δηλαδή 15000 φορές ταχύτερος από τον κοινό προσωπικό υπολογιστή. Η ανάπτυξη του υπολογιστή αυτού, ο οποίος ονομάστηκε "Blue Pacific", έγινε από την IBM και το Εθνικό Εργαστήριο Lawrence Livermore του Υπουργείου Ενέργειας των Η.Π.Α. στην Καλιφόρνια. Ο Blue Pacific έχει 2,6 Tbytes μνήμη, 80000 φορές τη μνήμη ενός κοινού προσωπικού υπολογιστή, και θα μπορούσε να αποθηκεύσει όλα τα βιβλία της Βιβλιοθήκης του Κογκρέσου. Ένας άνθρωπος με μια αριθμομηχανή θα χρειαζόταν 63000 χρόνια για την πραγματοποίηση τόσων υπολογισμών, όσων ο Blue Pacific θα πραγματοποιούσε σε ένα δευτερόλεπτο.

Την 28η Ιουνίου 2000, η IBM ανακοίνωσε την ολοκλήρωση της κατασκευής του υπερυπολογιστή της Accelerated Strategic Computing Initiative White, ή ASCI White, μια μαζικά παράλληλη μηχανή, η οποία έχει 8192 επεξεργαστές και μέγιστη απόδοση 12,3 TFlop/s.

Τον Νοέμβριο του 2004 ο ταχύτερος υπολογιστής στον κόσμο ήταν ο BlueGene/L της IBM με μέγιστη απόδοση 70,7 TFlop/s και 32.768 επεξεργαστές. Ο υπολογιστής BlueGene/P, δεύτερη γενιά της οικογένειας υπολογιστών Blue Gene της IBM, ο οποίος κατασκευάστηκε το 2007, περιέχει 212.992 επεξεργα-

στές και επιτυγχάνει μέγιστη απόδοση 478,2 TFlop/s. Το 2008 η IBM κατασκεύασε τον Roadrunner ο οποίος χρησιμοποιεί τους αντίστοιχους με τον BlueGene/L τυπικούς επεξεργαστές και επιπλέον οκταπύρηνους επεξεργαστές Cell της IBM. Ο Roadrunner έσπασε το φράγμα του PFlop (1015 Flop/s) και πέτυχε μέγιστη απόδοση 1,02 PFlop/s με 122.400 επεξεργαστές. Εγκαταστάθηκε το 2008 στο εθνικό εργαστήριο του Λος Άλαμος στο Νέο Μεξικό των Η.Π.Α. με σκοπό την προσομοίωση πυρηνικών δοκιμών. Ο Roadrunner χρησιμοποιήθηκε μέχρι τον Απρίλιο του 2013, οπότε ανακοινώθηκε από το εργαστήριο του Λος Άλαμος η απόσυρσή του.

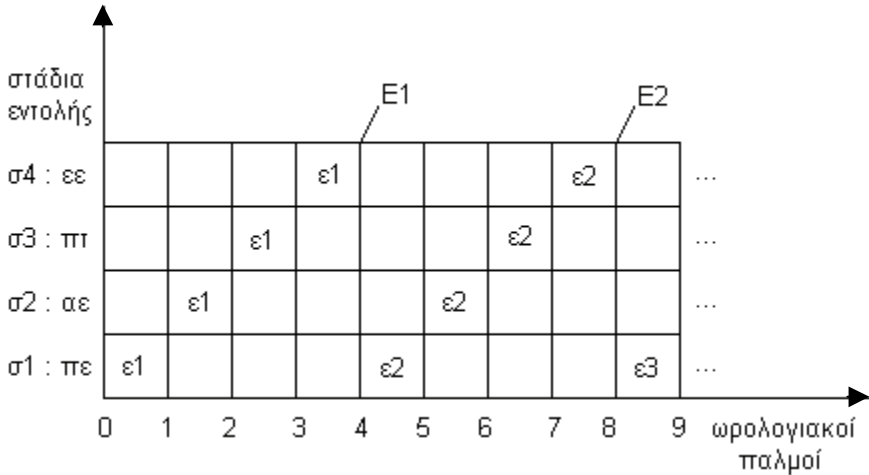
Στις 8 Φεβρουαρίου 2011, η IBM ανακοίνωσε την κατασκευή του υπερυπολογιστή BlueGene/Q, τρίτη γενιά της οικογένειας υπολογιστών Blue Gene. Το 2012, η μέγιστη απόδοση του BlueGene/Q ήταν στα 16,320 PFlop/s. Σύμφωνα με την top500 λίστα (www.top500.org), η οποία δημοσιεύεται τον Ιούνιο και το Νοέμβριο κάθε έτους και περιλαμβάνει ταξινομημένους με βάση την απόδοσή τους, τους 500 υψηλότερης απόδοσης εμπορικά διαθέσιμους υπερυπολογιστές στον κόσμο, τον Ιούνιο του 2013 ο υπερυπολογιστής με την υψηλότερη απόδοση στον κόσμο ήταν ο Tianhe-2, ο οποίος αναπτύχθηκε από το Εθνικό Πανεπιστήμιο Τεχνολογίας της Άμυνας της Κίνας. Η μέγιστη απόδοσή του είναι 33,86 PFlop/s.

1.3 Οι Έννοιες της Σωλήνωσης, του Παραλληλισμού Δεδομένων και του Παραλληλισμού Ελέγχου

Η παράλληλη επεξεργασία αποβλέπει στον ταυτόχρονο χειρισμό δεδομένων, τα οποία ανήκουν σε μία ή περισσότερες διεργασίες οι οποίες λύνουν το ίδιο πρόβλημα, και επομένως στην επίτευξη της *συνδρομικότητας (concurrency)* σε έναν υπολογισμό, ήτοι στην αύξηση του αριθμού των λειτουργιών που εκτελούνται σε κάθε χρονική στιγμή. Η σωλήνωση όπως και ο παραλληλισμός δεδομένων επιτυγχάνουν την αύξηση της συνδρομικότητας σε έναν υπολογισμό.

Σωληνωμένος υπολογισμός (pipelined computation) ή σωλήνωση. Ένας σωληνωμένος υπολογισμός διαίρειται σε έναν αριθμό βημάτων τα οποία καλούνται *στάδια (stages)*. Κάθε στάδιο αποτελεί ένα συγκεκριμένο τμήμα του υπολογισμού. Η έξοδος ενός σταδίου αποτελεί την είσοδο του επόμενου σταδίου. Αν θεωρήσουμε ότι κάθε στάδιο απαιτεί τον ίδιο χρόνο για να ολοκληρωθεί, τότε είναι εύκολο να παρατηρήσουμε ότι η βελτίωση του συνολικού χρόνου του υπολογισμού είναι ίση με έναν πολλαπλασιαστικό παράγοντα που ισούται με τον αριθμό των σταδίων της σωλήνωσης.

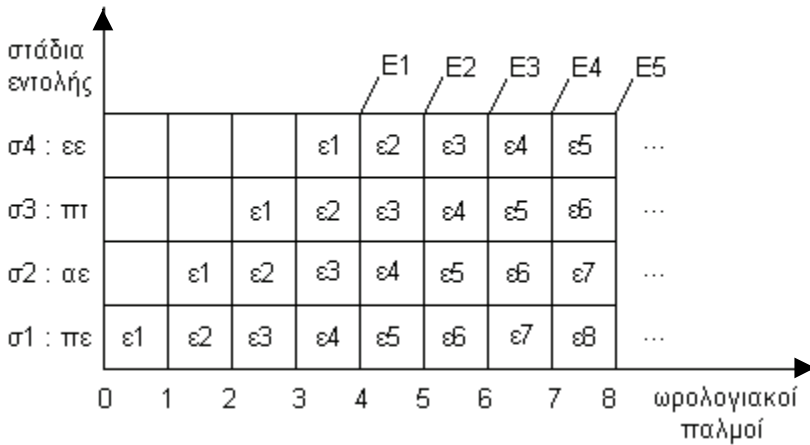
Προκειμένου να γίνει καλύτερα κατανοητή η έννοια της σωλήνωσης, ας θεωρήσουμε το Σχήμα 1.1, το οποίο απεικονίζει τα συνήθη τέσσερα διαδοχικά στάδια $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ διέλευσης μιας εντολής ϵ μέσα από έναν επεξεργαστή:



Σχήμα 1.1

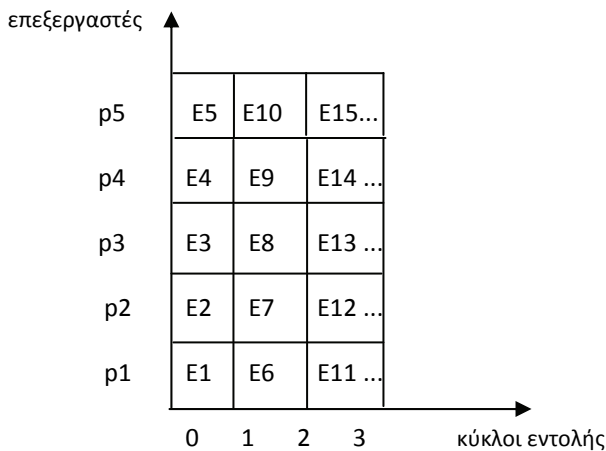
Στο σ_1 πραγματοποιείται η προσκόμιση της εντολής ϵ ($\pi\epsilon$), στο σ_2 η αποκωδικοποίηση της ϵ ($\alpha\epsilon$), στο σ_3 η προσκόμιση του τελεστέου ($\pi\iota$), και στο σ_4 η εκτέλεση της ϵ ($\epsilon\epsilon$). Υποθέτοντας ότι τα χρονικά διαστήματα $[i, i+1]$, $i=0,1,2,\dots$ είναι ίσα, στον παλμό 4 έχουμε την έξοδο $E1$ της πρώτης εντολής ϵ_1 από τον επεξεργαστή, στον παλμό 8 έχουμε την έξοδο $E2$ της δεύτερης εντολής ϵ_2 , κ.ο.κ.

Αν η διέλευση των εντολών ϵ_i , $i=1,2,\dots$ πραγματοποιείται από τον επεξεργαστή κατά το Σχήμα 1.2, ο επεξεργαστής χαρακτηρίζεται ως σωληνωμένος (*pipelined*), ενώ όλη η σχετική διεργασία ως σωλήνωση (*pipelining*). Ο χρόνος ολοκλήρωσης όλων των σταδίων μιας εντολής καλείται κύκλος εντολής (*instruction cycle*), ενώ ο χρόνος ολοκλήρωσης κάθε σταδίου της, κύκλος σωλήνα (*pipeline cycle*). Παρατηρούμε ότι στον παλμό 4 έχουμε την έξοδο $E1$ της πρώτης εντολής ϵ_1 από τον επεξεργαστή, στον παλμό 5 έχουμε την έξοδο $E2$ της δεύτερης εντολής ϵ_2 , στον παλμό 6 έχουμε την έξοδο $E3$ της τρίτης εντολής ϵ_3 , κ.ο.κ. Γενικεύοντας, μετά τον παλμό 3 έχουμε σε κάθε παλμό την έξοδο μιας εντολής, ήτοι μετά τον πρώτο κύκλο μηχανής έχουμε την έξοδο τεσσάρων εντολών σε κάθε κύκλο μηχανής.



Σχήμα 1.2

Παραλληλισμός δεδομένων (data parallelism). Είναι η χρήση πολλών λειτουργικών μονάδων προκειμένου να εφαρμοσθεί η ίδια λειτουργία ταυτόχρονα στα στοιχεία ενός συνόλου δεδομένων. Εάν υποθέσουμε ότι δεν υπάρχει επιβάρυνση (overhead) στην εκτέλεση των λειτουργιών που να συσχετίζεται με την αύξηση των λειτουργικών μονάδων, τότε εάν ο αριθμός των λειτουργικών μονάδων αυξηθεί κατά έναν πολλαπλασιαστικό παράγοντα c , ο χρόνος ολοκλήρωσης των λειτουργιών αυξάνεται επίσης κατά έναν πολλαπλασιαστικό παράγοντα c .



Σχήμα 1.3

Προκειμένου να αντιδιαστείλουμε τις έννοιες της σωλήνωσης και του παραλληλισμού δεδομένων, υποθέτουμε τώρα ότι διαθέτουμε μία μηχανή με π.χ. πέντε (μη σωληνωμένους) επεξεργαστές, τους p_1, p_2, p_3, p_4, p_5 , οι οποίοι εργάζονται παράλληλα προς επεξεργασία των $\epsilon_i, i=1,2,\dots$. Επίσης, υποθέτουμε ότι οι εντολές $\epsilon_i, i=1,2,\dots$ μπορούν να εκτελεσθούν παράλληλα (στην πραγματικότητα, αποτελούν την ίδια εντολή, η οποία εργάζεται σε διαφορετικά δεδομένα). Τότε, ανά κύκλο εντολής, ήτοι ανά τέσσερις ωρολογιακούς παλμούς, θα έχουμε την έξοδο πέντε εντολών, μια από κάθε επεξεργαστή, κατά το Σχήμα 1.3.

Επιτάχυνση (Speedup). Ως επιτάχυνση μιας μηχανής με ενσωματωμένη σωλήνωση ή παραλληλία, ως θεωρηθεί επί του παρόντος το κλάσμα με αριθμητή τον απαιτούμενο χρόνο εκτέλεσης ενός υπολογισμού από τον πλέον αποτελεσματικό ακολουθιακό αλγόριθμο, και παρονομαστή τον απαιτούμενο χρόνο εκτέλεσης του ίδιου υπολογισμού από την μηχανή με ενσωματωμένη σωλήνωση ή παραλληλία. Παρατηρούμε ότι, όσο μεγαλύτερη είναι η επιτάχυνση ενός παράλληλου ή σωληνωμένου υπολογισμού, τόσο μικρότερος είναι ο χρόνος εκτέλεσής του.

Με βάση την προηγούμενη έννοια της επιτάχυνσης, η υπεροχή του παραλληλισμού δεδομένων σε σχέση με τη σωλήνωση είναι προφανής. Ωστόσο αυτός δεν είναι άνευ προβλημάτων, όσον αφορά τα αντίστοιχα υπολογιστικά συστήματα. Τα κυριότερα των προβλημάτων αυτών σχετίζονται με:

- τον τρόπο επικοινωνίας και συγχρονισμού των μονάδων επεξεργασίας
- τον έλεγχο ταυτόχρονης πρόσβασης στην κοινή μνήμη ή τις συσκευές εισόδου-εξόδου
- τον τρόπο ανάθεσης των προγραμμάτων στις μονάδες επεξεργασίας
- τον τρόπο περιγραφής παράλληλων προγραμμάτων
- τη διαδικασία ανάπτυξης και συντήρησης παράλληλου λογισμικού

Πιθανοί συνδυασμοί σωλήνωσης και παραλληλισμού δεδομένων, προς επίτευξη μεγαλύτερης απόδοσης, παρουσιάζουν εξαιρετική πολυπλοκότητα, ιδιαίτερα καθώς διευρύνεται το πλήθος των επεξεργαστών.

Παραλληλισμός ελέγχου (control parallelism). Ενώ παραλληλισμός δεδομένων θεωρείται η εφαρμογή της ίδιας λειτουργίας ταυτόχρονα στα στοιχεία ενός συνόλου δεδομένων, ο παραλληλισμός ελέγχου επιτυγχάνεται με την εφαρμογή διαφορετικών λειτουργιών σε διαφορετικά σύνολα δεδομένων ταυτόχρονα. Η ροή των δεδομένων μεταξύ των διεργασιών ενός αλγορίθμου παραλληλισμού ελέγχου μπορεί να είναι ιδιαιτέρως πολύπλοκη. Εάν ο κατευθυνόμενος γράφος της ροής δεδομένων (data-flow digraph) ενός τέτοιου υπολογισμού είναι τόσο απλός ώστε να αποτελεί ένα κατευθυνόμενο μονοπάτι, τότε έχουμε ένα σωλη-

νωμένο υπολογισμό. Επομένως, η σωλήνωση μπορεί να θεωρηθεί ως ειδική περίπτωση του παραλληλισμού ελέγχου.

Αξίζει να σημειωθεί ότι οι περισσότεροι παράλληλοι αλγόριθμοι, οι οποίοι επιλύουν σύνθετα προβλήματα, αξιοποιούν και τα δύο είδη παραλληλισμού, ήτοι τον παραλληλισμό δεδομένων και τον παραλληλισμό ελέγχου.

Κλιμάκωση (scalability). Η κλιμάκωση ως έννοια, εκφράζει τη σταδιακή επέκταση. Ένας αλγόριθμος είναι *κλιμακώσιμος (scalable)*, αν το επίπεδο παραλληλισμού του αυξάνεται τουλάχιστον γραμμικά με το μέγεθος του προβλήματος, το οποίο σχετίζεται μ' αυτόν. Μια αρχιτεκτονική είναι κλιμακώσιμη, αν εξακολουθεί να συνεπάγεται την ίδια απόδοση ανά επεξεργαστή, καθώς αυξάνεται ο αριθμός των επεξεργαστών, μολονότι χρησιμοποιήθηκε σε μεγαλύτερου μεγέθους πρόβλημα. Η αλγοριθμική και η αρχιτεκτονική κλιμάκωση είναι σημαντικές, διότι επιτρέπουν στο χρήστη την επίλυση μεγαλύτερου μεγέθους προβλημάτων στον ίδιο χρόνο, δια της χρήσης ενός παράλληλου υπολογιστή με περισσότερους επεξεργαστές.

Οι αλγόριθμοι παραλληλισμού δεδομένων είναι πιο κλιμακώσιμοι από τους αλγόριθμους παραλληλισμού ελέγχου, διότι το επίπεδο του παραλληλισμού ελέγχου είναι συνήθως μια σταθερά, ανεξάρτητη του μεγέθους του προβλήματος, ενώ το επίπεδο παραλληλισμού δεδομένων είναι μια αύξουσα συνάρτηση του μεγέθους του προβλήματος.

Ως θετική ή αρνητική κλιμάκωση χαρακτηρίζεται και η δυνατότητα αύξησης ή μείωσης αντίστοιχα των συστατικών ενός υπολογιστικού συστήματος, ανάλογα με τις ανάγκες. Τέτοια κλιμακωτά συστήματα λειτουργούν ως σύνολα *υπηρετών (servers)*, οι οποίοι εξυπηρετούν ειδικούς σκοπούς εντός της συνάθροισης "υλικό - λογισμικό - ανθρώπινο δυναμικό", ενώ ένας οποιοσδήποτε υπηρέτης μπορεί να είναι ο ίδιος ένας παράλληλος υπολογιστής. Στα συστήματα αυτά διακρίνονται:

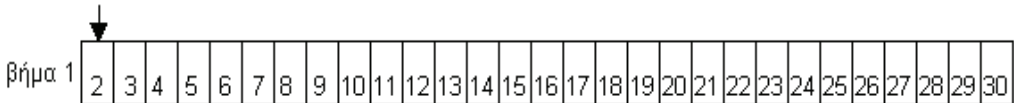
- ο υπηρέτης υπολογισμών (computation server)
- ο υπηρέτης δεδομένων/αποθήκευσης (data/storage server)
- ο υπηρέτης πρόσβασης/επικοινωνίας (access/communication server)
- ο υπηρέτης ορατικότητας (visualization server: εξυπηρετεί οπτικές απεικονίσεις)

Κλιμάκωση μπορεί να υφίσταται και στις *παράλληλες ιδεατές μηχανές (parallel virtual machines)*, οι οποίες είναι σύνολα ετεροκλήτων υπολογιστικών συστημάτων, στα οποία επιτυγχάνεται παράλληλη λειτουργία μέσω ειδικού λογισμικού.

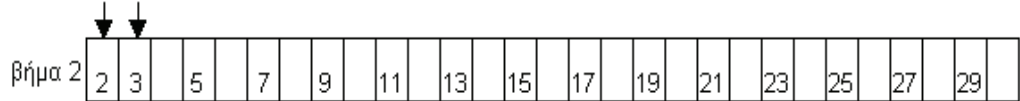
1.4 Το Κόσκινο του Ερατοσθένη

Πρόκειται για τον κλασικό αλγόριθμο εύρεσης πρώτων αριθμών, ο οποίος οφείλεται στον Αλεξανδρινό μαθηματικό, φιλόσοφο, αστρονόμο, φιλόλογο και ποιητή Ερατοσθένη (275-195 π.Χ.), και ο οποίος προσφέρεται για τη διερεύνηση μεθόδων παραλληλισμού.

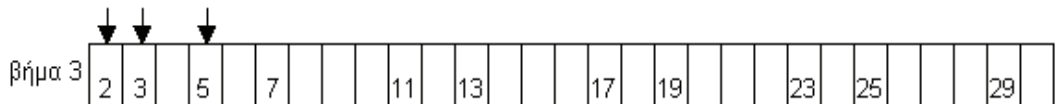
Το Κόσκινο του Ερατοσθένη επιλύει το πρόβλημα της εύρεσης των πρώτων αριθμών, οι οποίοι είναι μικρότεροι ενός θετικού και ακεραίου αριθμού N . Το Κόσκινο αρχίζει με μια γραμμική παράθεση των φυσικών αριθμών $2, 3, \dots, N$ και μετά εξαλείφει διαδοχικούς πρώτους και πολλαπλάσιά τους, ενώ τελειώνει αφού εξαλειφθούν πολλαπλάσια του μεγαλύτερου πρώτου, ο οποίος είναι μικρότερος ή ίσος της τετραγωνικής ρίζας του N . Το ακόλουθο παράδειγμα, για $N=30$, απεικονίζει τη λειτουργία του Κοσκίνου:



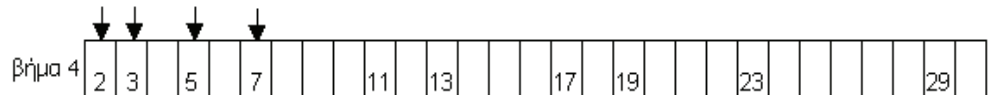
Πρώτος είναι ο επόμενος μη επισημασμένος φυσικός αριθμός: 2 (ο οποίος κατόπιν επισημαίνεται).



Εξάλειψη όλων των πολλαπλασίων του 2, αρχίζοντας με τον 2^2 . Πρώτος είναι ο επόμενος μη επισημασμένος φυσικός αριθμός: 3 (ο οποίος κατόπιν επισημαίνεται).



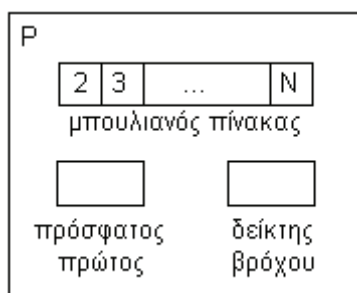
Εξάλειψη όλων των πολλαπλασίων του 3, αρχίζοντας με τον 3^2 . Πρώτος είναι ο επόμενος μη επισημασμένος φυσικός αριθμός: 5 (ο οποίος κατόπιν επισημαίνεται).



Εξάλειψη όλων των πολλαπλασίων του 5, αρχίζοντας με τον 5^2 . Πρώτος είναι ο επόμενος μη επισημασμένος φυσικός αριθμός: 7 (ο οποίος κατόπιν επισημαίνεται). Εφόσον $7^2 > 30$, ο αλγόριθμος τελειώνει. Οι επισημασμένοι και όλοι οι φυσικοί αριθμοί, οι οποίοι απόμειναν, είναι οι μικρότεροι του 30 διαδοχικοί πρώτοι αριθμοί.

Η εξάλειψη όλων των πολλαπλασίων, αρχίζοντας κάθε φορά με το τετράγωνο του προηγούμενου πρώτου, κάνει το Κόσκινο δυσκολόχρηστο, όταν το N παίρνει μεγάλες τιμές, π.χ. εκατοντάδων ψηφίων η καθεμία. Πάντως, πιο εύχρηστοι αλγόριθμοι παραγοντοποίησης κάνουν χρήση τεχνικών Κοσκίνου κατά διαφορετικούς τρόπους. Ας θεωρήσουμε το Σχήμα 1.4, το οποίο απεικονίζει την ακολουθιακή εκτέλεση του Κοσκίνου από έναν επεξεργαστή p . Τρεις ουσιαστικές δομές δεδομένων γίνονται αντικείμενο διαχείρισης:

- Ένας μονοδιάστατος μπουλιανός πίνακας, του οποίου τα στοιχεία αντιστοιχούν στους φυσικούς αριθμούς, οι οποίοι κοσκινίζονται.
- Ένας ακέραιος, ο οποίος αντιστοιχεί στον πλέον προσφάτως ευρεθέντα πρώτο.
- Ένας ακέραιος, ο οποίος χρησιμοποιείται ως δείκτης βρόχου, αυξανόμενος καθώς πολλαπλάσια του τρεχούμενου πρώτου, ως σύνθετοι αριθμοί, εξαλείφονται.



Σχήμα 1.4

Στη συνέχεια θα μελετήσουμε την εκτέλεση του Κοσκίνου πρώτα σε ένα πρότυπο παραλληλισμού ελέγχου και στη συνέχεια σε ένα πρότυπο παραλληλισμού δεδομένων. Προχωρούμε τώρα σ' ένα πρότυπο παραλληλισμού ελέγχου, κατά το Σχήμα 1.5.