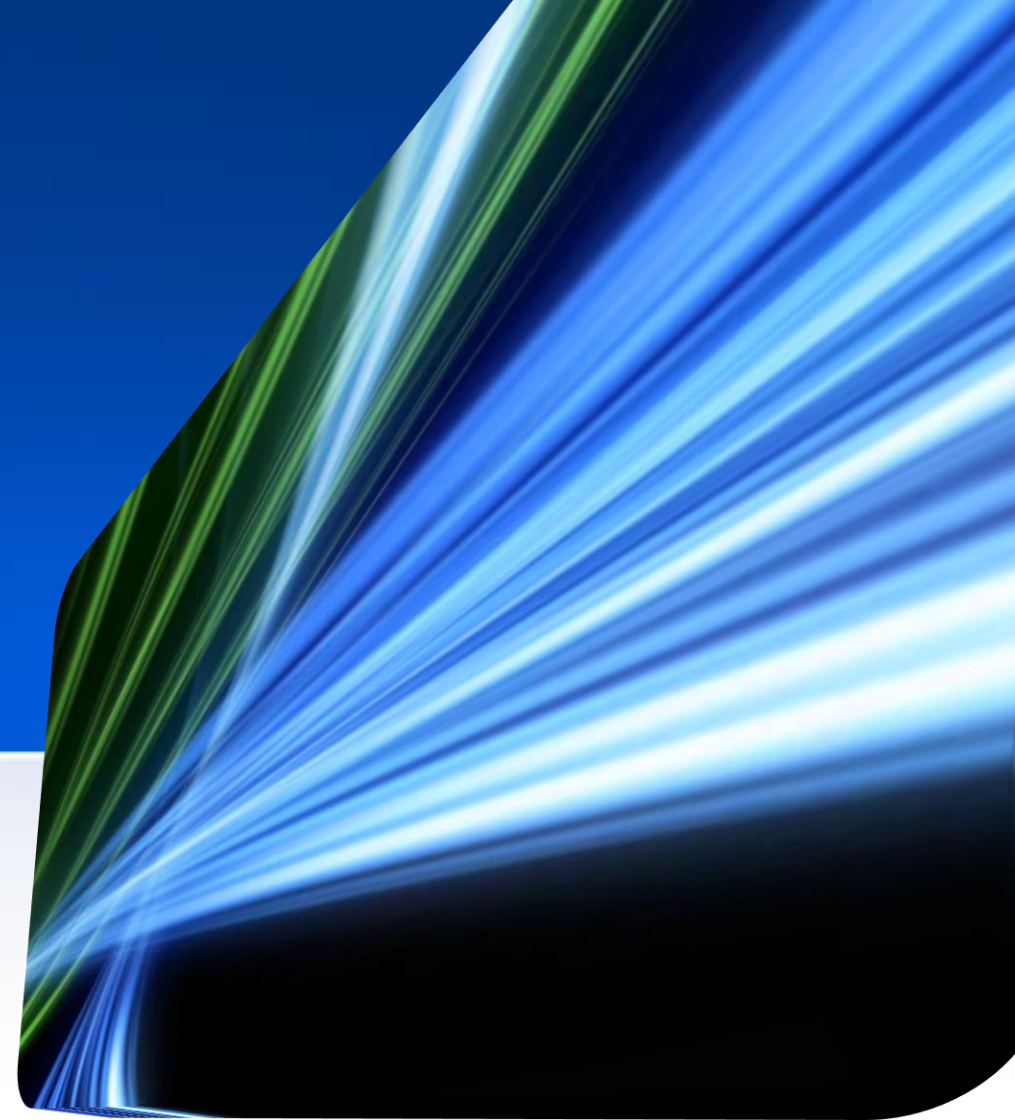


Λειτουργικά Συστήματα
6ο εξάμηνο ΣΗΜΜΥ
Ακ. έτος 2020-2021

**Συστήματα μαζικής
αποθήκευσης**

Παν. Τσανάκας



Συστήματα μαζικής αποθήκευσης

- ◆ Δρομολόγηση HDD και NVM
- ◆ Εντοπισμός & διόρθωση σφαλμάτων
- ◆ Διαχείριση συσκευών αποθήκευσης

Στόχοι

- ◆ Περιγραφή της φυσικής δομής και της σύνδεσης με τη χρήση των συσκευών
- ◆ Εξήγηση χαρακτηριστικών επίδοσης συσκευών αποθήκευσης
- ◆ Αξιολόγηση αλγορίθμων δρομολόγησης E/E

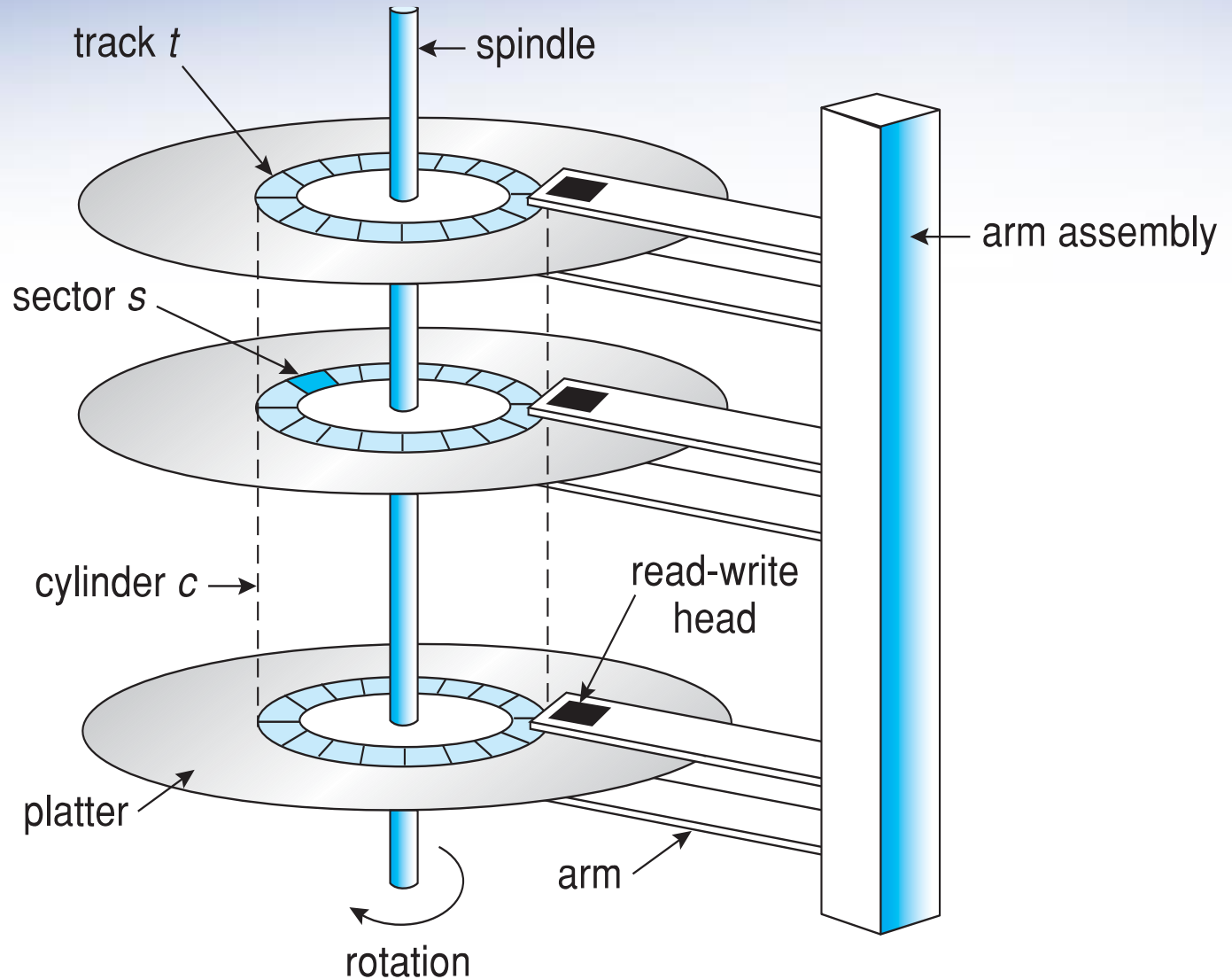
Μηχανισμός σκληρού δίσκου

Επιφάνειες (platters)

Τροχιές (tracks)

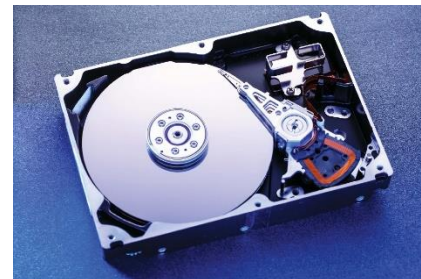
Κύλινδροι (cylinders)

Τομείς (sectors)



Γενική δομή

- ◆ **HDDs** περιστρεφόμενες μαγνητικές επιφάνειες, κάτω από κινούμενες κεφαλές A/E
 - 60 μέχρι 250 περιστροφές ανα sec
 - **Ρυθμός μεταφοράς, π.χ. 6 Gbps**
 - **Χρόνος τυχαίας προσπέλασης (random-access time),**
 - Χρόνος αναζήτησης (π.χ. 9 ms) + Χρόνος περιστροφής (1/2 του χρόνου πλήρους περιστροφής, π.χ. 2 ms)



Μη πτητικές μνήμες NVM

- Πιο αξιόπιστες από HDDs
- Πιο ακριβές ανά MB
- Μάλλον μικρότερο χρόνο ζωής
- Μικρότερο χρήσιμο χώρο
- Αλλά, είναι **πολύ** πιο γρήγορες
- Όχι μηχανικά μέρη, μηδενικός χρόνος αναζήτησης

Μη πτητικές μνήμες - NVM

- ◆ Ανάγνωση και εγγραφή σε “σελίδες”, αλλά δεν μπορεί να γίνει άμεση επανεγγραφή
- Προηγείται η διαγραφή (σε μεγάλο χώρο), και μετά ακολουθεί η εγγραφή σε μία «σελίδα»
- Δεν μπορούν να γίνουν απεριόριστες «διαγραφές» – $\sim 100,000$
- Εξισορρόπηση φθοράς - Wear leveling



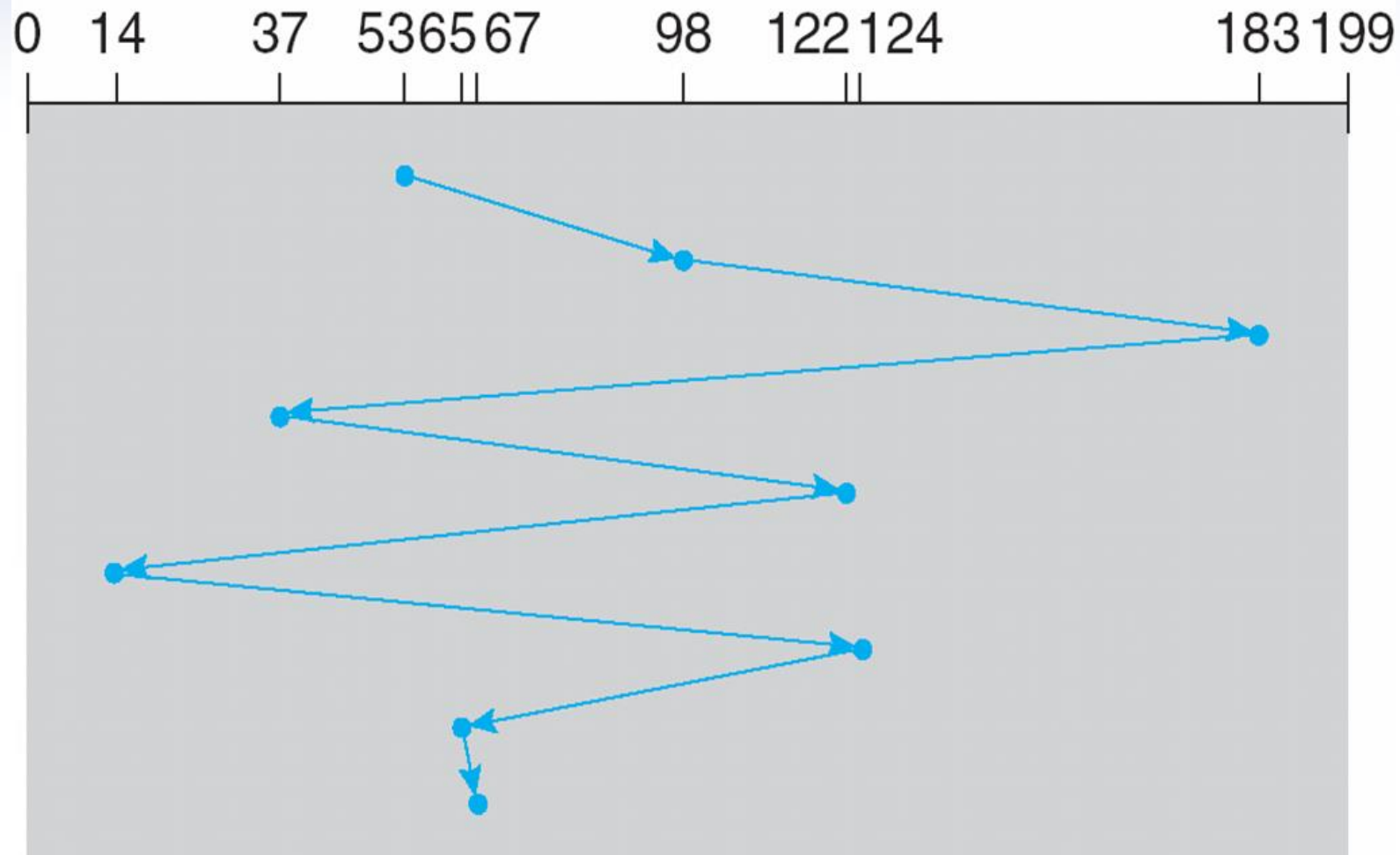
Απεικόνιση διευθύνσεων

- ◆ Δίσκος – Μεγάλος μονοδιάστατος πίνακας από **λογικά πλαίσια** (logical blocks)
- ◆ **Πλαίσιο** – η ελάχιστη μονάδα μεταφέρσιμης πληροφορίας
- ◆ Απεικόνιση **λογικών πλαισίων** σε **φυσικά πλαίσια** (οδηγός συσκευής)
- ◆ Τομέας 0 – ο πρώτος τομέας της πρώτης τροχιάς του πιο εξωτερικού κυλίνδρου.
- ◆ Συνεχόμενη απεικόνιση των υπόλοιπων τομέων των άλλων κυλίνδρων Κ.Ο.Κ.
 - Εξαιρούνται οι κατεστραμμένοι τομείς
 - Όχι σταθερό πλήθος τομέων ανά τροχιά

Δρομολόγηση δίσκου - FCFS

queue = 98, 183, 37, 122, 14, 124, 65, 67

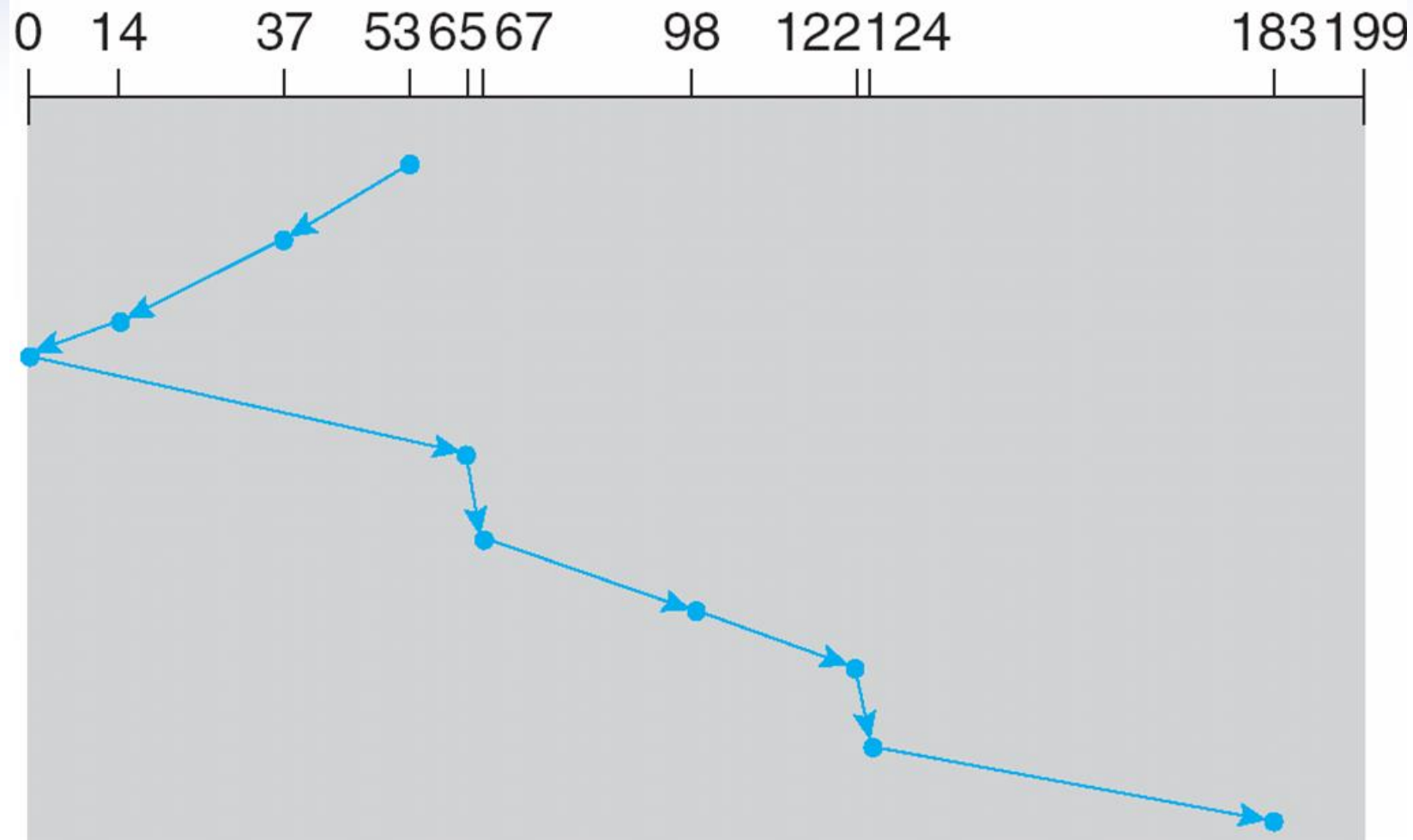
head starts at 53



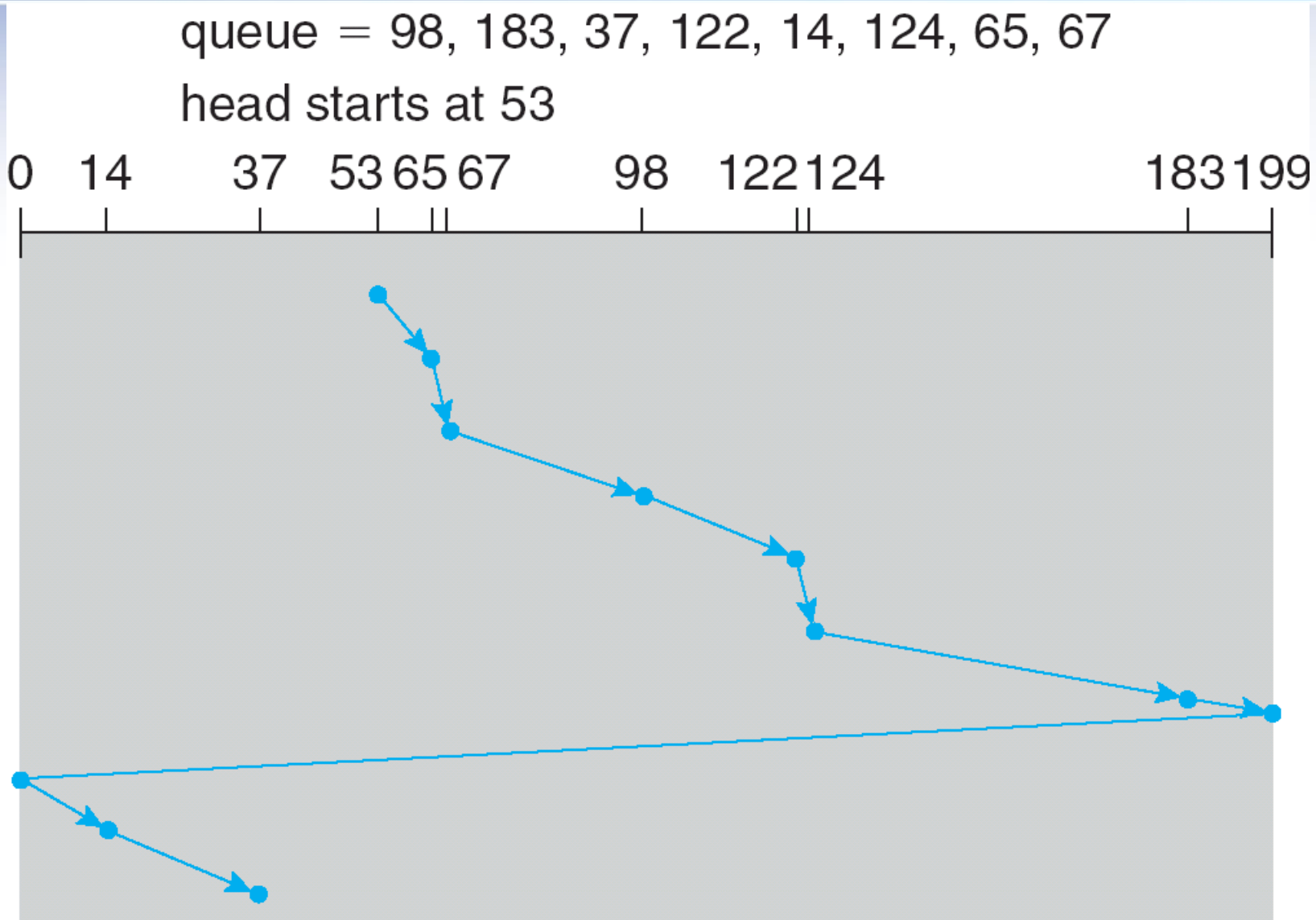
Δρομολόγηση δίσκου - SCAN

queue = 98, 183, 37, 122, 14, 124, 65, 67

head starts at 53



Δρομολόγηση δίσκου C-SCAN



Δρομολόγηση NVM

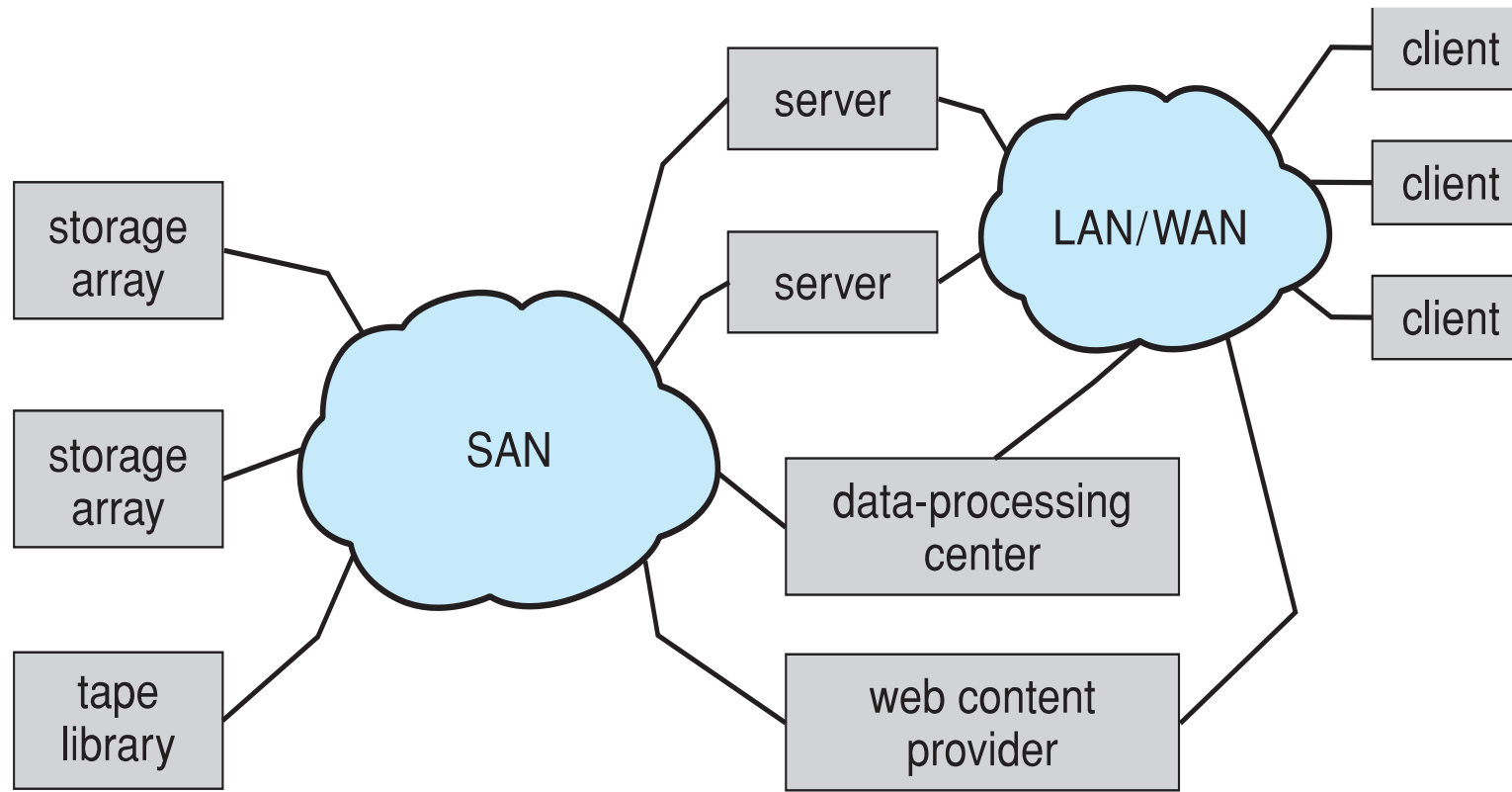
- ◆ Χωρίς μηχανικά μέρη, έχει μικρά περιθώρια για βελτιώσεις.
- Συγχώνευση γειτονικών αιτημάτων
- Wear leveling

Αποθήκευση στο νέφος

- ◆ Αποθήκευση σε Κέντρα Δεδομένων, μέσω Διαδικτύου
- ◆ Η πρόσβαση στηρίζεται κυρίως σε API
 - Dropbox, Amazon S3, Microsoft OneDrive, Apple iCloud, Google Drive κλπ
 - Πληρέστερη διαχείριση αστοχιών και καθυστερήσεων

Storage Area Networks

- ◆ Πολύ μεγάλοι αποθηκευτικοί χώροι
- ◆ Πολλαπλοί Διακομιστές – Πολλαπλές μονάδες αποθήκευσης – Κέντρα Δεδομένων



Αποθήκευση αντικειμένων

Object storage

- ◆ Για αποθήκευση δεδομένων μεγάλης κλίμακας
- ◆ Ξεκινάμε από τις συσκευές και τοποθετούμε σε αυτές αντικείμενα
 - Object είναι απλώς ένας περιέκτης δεδομένων
 - Δεν υπάρχει τρόπος να πλοηγηθείς στα δεδομένα για να βρείς αντικείμενα

Τυπική ακολουθία

- Δημιουργία αντικειμένου – Απόκτηση ObjectID
- Προσπέλαση αντικειμένου μέσω ObjectID
- Διαγραφή αντικειμένου με το ObjectID

Αποθήκευση αντικειμένων

Object storage

- ◆ Συστήματα όπως το **Hadoop file system (HDFS)** και το **Ceph** ορίζουν **πού** αποθηκεύονται τα αντικείμενα και χειρίζονται μηχανισμούς προστασίας
 - Συχνά αποθηκεύοντας N αντίγραφα, με κατανομή σε N συστήματα, στην αποθηκευτική συστάδα (storage cluster)
 - **Οριζοντίως κλιμακώσιμα (scalable)**
 - **Προσπέλαση με τη διεύθυνση, χωρίς δομή**