CHAPTER 12

Mass-Storage Systems

(60 Questions)

01: What do ***Magnetic Disks*** provide? How to work?

02: What is ***Transfer Rate***?

03: What is ***Positioning Time***?

04: What is ***Head Crash***?

05: What does ***Host Controller use*** in computer? What are Busses including?

06: Draw "***Moving-Head Disk Mechanism***"?

07: What is the ***Magnetic Disk Performance***?

08: What are the ***attributes of Solid-State Disks***?

09: What are the ***attributes of Magnetic Tape***?

10: What is ***Sector 0***?

11: How many ***ways for*** computers can **access to disk storage**?

12: What is the difference between ***Host-Attached Storage*** and ***Network-attached Storage***?

13: What is ***Fiber Channel (FC)***?

14: What can ***Storage Array do***?

15: What does ***Storage Array have***?

16: What is ***Storage Area Network (SAN)?***

17: What are the ***features of Storage Area Network (SAN)***?

18: What is ***Network Attached Storage (NAS)***?

19: What are ***NFC and CIFS***?

20: What is the difference between ***SAN*** and ***NAS***?

21: What is ***iSCSI***?

22: What is ***Disk Bandwidth***?

23: How many ***sources for disk I/O request***?

24: What does ***I/O request include***?

25: What does ***SSTF select***?

26: What is ***SSTF Scheduling***?

27: What is the ***function of SCAN Scheduling***?

28: Sometimes ***SACN algorithm called*** by?

29: What will it happen if requests are ***uniformly dense in SCAN Scheduling***?

30: What does ***C-SCAN provide***?

31: What are ***the functions of C-SCAN***?

32: What are ***LOOK and C-LOOK Scheduling***? What is the reason for using it?

33: How we choose the best one of many ***disk-scheduling algorithms***?

34: What is ***Low-level formatting***?

35: What are the two ***steps to use a disk to hold files***?

36: How are ***Disk I/O*** and ***File System I/O*** done?

37: What is ***Raw disk access*** working for?

38: What is the function of ***Boot Block***?

39: What is ***Bootstrap Loader***?

40: What is ***Sector Spring use for***?

41: What is ***Swap-Space***?

42: What ***Swap-Space could be***?

43: What is ***Swap-Space Management***?

44: What if a ***system runs out of Swap-Space***?

45: What is ***RAID***?

46: What are ***the two factors*** that ***mean time to failure is the loss of data***?

47: Why ***frequently RAID combined with NVRAM***?

48: How many ***levels*** that ***RAID is arranged into***?

49: What does Disk Striping use?

50: What do RAID Schemes improve?

51: What are the differences between ***Mirroring***, ***Striped mirrors*** and ***Block Interleaved parity***?

52: What is the difference between Mirroring and Striping?

Mention their advantages and disadvantages?

53: What is a ***Snapshot***?

54: What is ***Replication***?

55: What is ***Hot Spare disk***?

56: Can RAID alone prevent or detect data corruption or other errors?

57: What is ***Checksums***?

58: Mention the two needed things to implement Stable Storage?

59: A disk write results in one of three outcomes. What are they?

60: The system must maintain two physical blocks for each logical block. An output operation is executed as follows:

End of Questions (Chapter 12).

Chapter 12

Mass-Storage Structure

1: **الأقراص الممغنطة** توفر جزء كبير (bulk) من التخزين الثانوي للكمبيوترات الحديثة. فلأقراص يمكن ان تكون قابلة للإزالة

وأما المحركات ترفق مع الكمبيوتر عبر I/O bus. فعند الاستخدام يدور محرك motor بسرعة عالية وهي تحتوي على:

- محركات تدور من 60 إلى 250 مرة في كل ثانية.
- معدل النقل (Transfer Rate) وهو معدل تدفق البيانات بين القرص وجهاز الكمبيوتر.
- وضعية الوقت (Positioning Time) أو وقت الوصول العشوائي: وهو وقت حركة ذراع القرص إلى الأسطوانة المرغوبة **seek time** ووقت القطاع المرغوب فيها (Desired Sector) للتدير تحت رأس القرص **rotational latency.**
- التحام الرأس (Head Crash) وينتج من اتصال رأس القرص مع سطح القرص.

2: **معدل النقل (Transfer Rate)** وهو معدل تدفق البيانات بين القرص وجهاز الكمبيوتر.

3: **وضعية الوقت (Positioning Time)** أو وقت الوصول العشوائي: وهو وقت حركة ذراع القرص إلى الأسطوانة المرغوبة

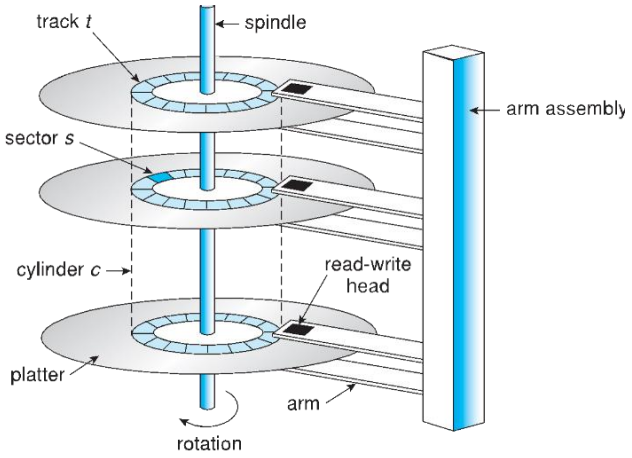
Seek time ووقت القطاع المرغوب فيها (Desired Sector) للتدير تحت رأس القرص **rotational latency.**

4: **التحام الرأس (Head Crash)** وينتج من اتصال رأس القرص مع سطح القرص.

5: **Host Controller** في الكمبيوتر يستخدم bus للتحدث إلى disk controller المبنية في المحرك أو في مصفوفة التخزين.

Busses لديها أنواع مختلفة فيه تشمل: EIDE, ATA, SATA, USB, Fiber Channel, SCSI, SAS, Firewire

6: آلية تحرك الرأس بالقرص (Moving-Head Disk Mechanism):



7: **أداء القرص الممغنط (Magnetic Disk Performance):**

a) معدل النقل – 6 GB/sec
b) فاعلية معدل النقل – 1 GB/sec
c) Seek Time من 3ms إلى 12ms.
d) معدل مقياس Seek Time او الحساب بناءا على 1/3 من tracks
e) Latency بناءا على سرعة spindle.
f) معدل latency = ½ latency.

8: **الأقراص الصلبة (Solid-State Disks) من سماتها:**

1) تعتبر ذاكرة Nonvolatile وتستخدم مثل المحرك الصلب.
2) فعالة أكثر من HDDs.
3) تكلفتها تزداد لكل MB.
4) ربما يكون عمره أقصر.
5) سعته أقل.
6) أسرع بكثير.
7) Busses سرعته بطيئة جدا، وتتصل مباشرة إلى PCI.
8) أجزاءه غير متحركة، فهو إذا ليس seek time أو rotational latency.

9: **الشريط الممغنط (Magnetic Tape) من سماته:**

1) كان في السابق متوسط التخزين الثانوي.
2) مستمر نسبيا ويحمل كميات كبيرة من البيانات.
3) وقت الوصول بطيء.
4) الوصول العشوائي ابطأ 1000 مرة من القرص.
5) يستخدم بشكل أساسي لأجل backup.
6) يحتفظ في spool and wound.

10: **القطاع 0 (Sector 0):** هو القطاع الأول من المسار الأول (first track) في الأسطوانة الابعد (outermost cylinder).

11: **الكمبيوترات تصل إلى قرص التخزين بطريقتين:**

الأولى: عبر منافذ الادخال والإخراج أو host-attached storage، وهو شائع في الأنظمة الصغيرة.

الثانية: عبر remote host في ملفات النظام الموزعة، وهي تعود إلى network-attached storage.

12: **الفرق بين host and network-attached storage:**

- Host-attached storage: هو تخزين موصول من خلال منافذ الادخال والإخراج المحلية.
- Network-attached storage: هو جهاز لنظام تخزين أغراض خاصة التي يتم الوصول إليها عن بعد عبر شبكة البيانات.

13: **قناة الألياف (fiber channel):** هي بنية تسلسلية عالية السرعة التي يمكن أن تعمل عبر الالياف البصرية أو عبر الكابلات

النحاسية ذو الأربع الموصلات (a four-conductor).

14: **مصفوفة التخزين (Storage Array)** بإمكانه إرفاق الاقراص (attach disks) أو تنظيم الاقراص (array of disks).

15: **مصفوفة التخزين (Storage Array)** لها تحكمات تقوم بتوفير ميزات الـ hosts المرفقة:

- منافذ لاتصال hosts to array.
- الذاكرة، برامج التحكم.
- عدد قليل من الآلاف من الاقراص.
- RAID, hot spares, hot swap
- التخزين المشترك وهو أكثر كفاءة.
- ميزات موجودة في بعض أنظمة الملفات (Snaphots, clones, thin provisioning, replication, deduplication)

16: **شبكة منطقة التخزين (SAN: Storage Area Network):** هو عبارة عن مصفوفة تخزين أو أكثر. وتكون متصلة مع

مفتاح (switch) او أكثر من قنوات الألياف (Fiber Channel).

17: **من مميزات شبكة منطقة التخزين (SAN):**

1) الاتصال مع مفتاح (switch) أو أكثر من قنوات الالياف.
2) Hosts أيضا يرفق (attach) مع المفاتيح (switches).
3) وحدة التخزين المتوفرة عبر LUN Masking من مصفوفات معينة إلى خوادم معينة.
4) السهولة في إضافة وإزالة وحدة التخزين، وإمكانية إضافة host وتخصيصه في وحدة التخزين.

18: **وحدة التخزين المتصلة بالشبكة (NAS: Network-Attached Storage):** هي وحدة التخزين المتوفرة عبر الشبكة بدلا من الاتصال المحلي مثل bus. وهي مرفقة إلى انظمة الملفات عن بعد.

19: **NFS & CIFS** هما بروتوكولات الاكثر شيوعا وتنفذ عبر نداءات الاجراءات عن بعد (RPCs: Remote Procedure Calls) بين host ووحدة التخزين TCP عادة أو UDP على شبكة IP.

20: **الفرق بين شبكة منطقة التخزين (SAN) ووحدة التخزين المتصلة بالشبكة (NAS):**

- **شبكة منطقة التخزين (SAN):** يتعرف عليه الجهاز العميل على أنه كتلة (مثل القرص الصلب المحلي)، وتستطيع التحكم به من خلال تقسيم الكتلة إلى مجموعة أجزاء ()، وتركيب file system عليه، وإعادة فرمته الكتلة والاجزاء مثل أي قرص صلب محلي.
- وحدة التخزين المتصلة بالشبكة (NAS): يقدم النظام السيرفر file system القابل للاستخدام مباشرة عبر الشبكة المحلية أو الخارجية، ويتم التعرف عليه من قبل النظام العميل على انه file system وليس كتلة أو قرص صلب.

21: **iSCSI:** هو بروتوكول يستخدم شبكة IP ليحمل بروتوكول SCSI. وتكون مرفقة إلى الاجهزة blocks عن بعد.

22: **Disk Bandwidth:** هو مجموع ارقام البايتات (bytes) المنقولة، والمقسمة بمجموع الوقت بين الطلب الاول من الخدمة

وحتى الانتهاء من اخر نقل.

23: هناك العديد من **مصادر (disk I/O request)** مثل: OS / عمليات النظام / عمليات المستخدمين.

24: **I/O request** يشمل طريقة الادخال أو الاخراج، وعنوان الذاكرة وعدد القطاعات للنقل. ونظام التشغيل يقوم بصيانة صف

الطلبات لكل قرص او جهاز. Idle disk بإمكانه العمل مباشرة على I/O request واذا كان القرص مشغول هذا يعني:

Work must queue.

25: **Shortest Seek Time First (SSTF)** يقوم باختيار الطلب الاقصر استغراقا للوقت من رأس الموقع الحالي.

26: **جدولة الـ SSTF:** هو شكل من أشكال جدولة SJF؛ قد يسبب التعطيل (starvation) لبعض الطلبات.

27: **جدولة الـ SCAN:** ذراع القرص يبدأ من نهاية القرص ويتحرك صوب النهاية الاخرى، والخدمات تقوم بالطلب حتى تصل

إلى النهاية الأخرى من القرص، عندما تنعكس حركة الرأس فإن الخدمة تستمر.

28: **SCAN Algorithm** يطلق عليه بعض الاحيان elevator algorithm.

29: لو كانت الطلبات في خوارزمية SCAN كثيفة وغير موحدة فإن الطلبات الموجودة في نهايات القرص يكون انتظارها أطول.

30: **خوارزمية C-SCAN** تقوم بتوفير وقت الانتظار بشكل أكثر انتظاما (uniform wait time) من الـ SCAN.

31: **وظيفة C-SCAN:** أن الرأس يتحرك من نهاية القرص إلى النهاية الاخرى، وخدمة الطلبات تقوم بالطلب كما هو. وعندما يصل

إلى النهاية الاخرى، ومع ذلك، فإنه يعود بشكل مباشر إلى بداية القرص من خير خدمة اي طلب عند طريق العودة.

32: **LOOK** نسخة من SCAN، **C-LOOK** نسخة من C-SCAN.

- الذراع فقط يذهب بقدر الامكان إلى الطلب الاخير في كل اتجاه وبعد ذلك يعكس الاتجاه مباشرة من غير الذهاب إلى المكان الاول لجميع الطرق إلى نهاية القرص.

33: **اختيار خوارزمية جدولة القرص:**

- SSTF هو شائع وله جاذبية طبيعية (a natural Appeal).
- SCAN and C-SCAN يؤدي بشكل أفضل للأنظمة التي تقع عليها الحمولة الثقيلة في القرص ونسبة التعطيل فيها أقل.
- الاداء يعتمد على عدد الطلبات وانواعها.
- طلبات خدمة القرص يمكن أن يتأثر (influenced) بواسطة طرقة تخصيص الملف.
- خوارزمية جدولة القرص ينبغي ان تكون مكتوبة كأنماط مفصلة من نظام التشغيل، والسماح لها بتغيير الموقع مع خوارزمية مختلفة عند الضرورة.
- إما SSTF أو LOOK هو الخيار المقنع للخوارزمية الافتراضية (default algorithm).

34: **Low-level formatting or physical formatting:** يقوم بتقسيم القرص إلى قطاعات والتي تجعل من المتحكم بالقرص

القراءة والكتابة. وكل قطاع يستطيع ان يتمسك بمعلومات الـ header بالإضافة إلى البيانات وكود تصحيح الاخطاء (ECC: Error Correction Code).

35: قبل استخدام القرص للتمسك بالملفات، نظام التشغيل لا يزال يحتاج إلى تسجيل هياكل بياناته الخاصة على القرص ويأتي ذلك

بخطوتين:

- **الخطوة الاولى:** Partition حيث يقسم القرص إلى مجموعة واحدة أو أكثر من الاسطوانات، وجميعها تكون معالجة كقرص منطقي.
- **الخطوة الثانية:** logical formatting وهو انشاء نظام ملف.

36: **لزيادة الكفاءة معظم انظمة الملفات** يجمع blocks إلى مجموعات (cluster):

- قرص I/O ينجز في الـ blocks.
- ملف I/O ينجز في الـ clusters.

37: **Raw Disk Access:** هي للتطبيقات التي تريد ان تدير ال block الخاص بها. وتبعد نظام التشغيل بعيد عنها مثل قواعد البيانات.

38: **Boot Black:** يقوم بتهيئة (initializes) النظام، وتكون مخزنة في الـ ROM.

39: **Bootstrap Loader** هو برنامج مخزن في boot blocks of boot partition.

40: **Sector Spring:** يستخدم للتعامل مع الـ blocks السيء.

41: **مساحة المبادلة (Swap-space):** هي ذاكرة افتراضية تستخدم مساحة القرص امتدادا للذاكرة الرئيسية. وهو أقل شيوعا الان بسبب زيادة سعة الذاكرة.

42: **مساحة المبادلة (Swap-Space)** بإمكانها ان تكون منتقاة (carved out) من نظم الملفات الطبيعية. أو تكون قسم لقرص منفصل وهو الأكثر شيوعا.

43: **إدارة مساحة المبادلة (Swap-Space Management):**

- BSD4.3 يقوم بتخصيص مساحة مبادلة عندما تبدأ العملية بالتمسك بـ text segment أو data segment.
- النواة (kernel) تستخدم swap maps لتتبع استخدام مساحة المبادلة.
- Solaris 2 تقوم بتخصيص مساحة المبادلة فقط عندما تجبر صفحة كثيرة الأخطاء بالخروج من الذاكرة الفعلية. وليس عندما يتم انشاء صفحة الذاكرة الافتراضية أولا.

44: بعض الأنظمة ستسمح بتعدد مساحات المبادلة.

45: **RAID** هو تصفيف مكرر من الأقراص الغير مكلفة (redundant array of inexpensive disks) وهو تقنية لكيفية التعامل مع الأقراص الصلبة للوصول إلى زيادة سعة التخزين والأداء العالي وتوفر المعلومات والتحكم بكيفية تخزين المعلومات و Hot Swap أي انه يمكن تغيير أي قرص صلب معطوب في أي وقت دون إطفاء السيرفر.

46: **أثناء حدوث الفشل (main time to failure) في mirrored volume** وذلك بفقدان البيانات فإنه يعتمد على عاملين:

a) أثناء حدوث الفشل في الأقراص الشخصية (mean time to failure of the individual disks)
b) أثناء الإصلاح (mean time to repair) وهنا يأخذ معدل الوقت الذي يتم فيه تبديل القرص الذي باء بالفشل وتخزين البيانات عليها.

47: **RAID يرتبط باستمرار مع NVRAM** لتحسين أداء الكتابة.

48: RAID ينظم في 6 مستويات مختلفة.

49: **Disk Striping** يستخدم مجموعة من الأقراص كوحدة تخزين واحدة.

50: **RAID Schemes** يقوم بتحسين الأداء وأيضا بتحسين صلابة نظام التخزين بواسطة تخزين البيانات المتكررة.

51: ***Mirroring:*** أو ما يسمى shadowing (RAID 1) يبقي نسخة مشابهة في كل قرص.

Striped mirrors (RAID 1+0) او Mirrored Stripes (RAID 0+1): يقوم بتوفير أداء عالي وصلابة عالية.

Block interleaved parity (RAID 4,5,6): يستخدم التكرير (redundancy) أقل بكثير.

52: **الفرق بين Mirroring and striping:**

- **Mirroring:** يقوم بتخزين نسخة من نفس البيانات على أقراص مختلفة. ولو قرص صلب واحد تعطل فسوف تجد نسخة أخرى محفوظة في القرص الاخر.
- **Striping:** يقوم بتخزين المعلومات أو البيانات على كل قرص. أي يقوم بتقسيم البيانات ومن ثم تخزينها في أقراص صلبة مختلفة. ولو تعطل أي قرص فسوف تخسر جميع البيانات. وتستخدم لتطوير السرعة لأن كل قرص سيقوم بمعالجة البيانات الموجودة لديه.

ومن أبرز صفاتهم وعيوبهم:

Striping	Mirroring
Striping: is storing the data or information into each hard drive to read and write the data faster.	Mirroring: is storing the same copies of data or information for different hard drives. Which increases data security from loss or damage.
Advantages: • Increasing the speed of read and write the data. • Improving the performance and efficiency. • Easy to implement. • There is no overhead because the storage of al storage is used.	Advantages: • Has a backup; for any damaged or failed for any disks, you will get it from the other disk. • Increasing the data from any loss or damages. •
Disadvantages: • Any damaged or failed in any disks, you will lose all data. • Not recommended to use for mission critical systems.	Disadvantages: • In theory, slower for processing the data. • It is expensive. • When the power outages before writing is complete on both disks. They will be incompatible. It will write only one copy in block then to the other.

53: **Snapshot:** هو عرض نظام الملفات قبل مجموعة من التغيرات التي قد تحدث.

54: **النسخ المتماثل (Replication):** هو نسخ متماثل تلقائي من الكتابات بين المواقع المنفصلة:

- للتكرار واستعادة الكوارث.
- يمكن ان تكون متلازمة وغير متلازمة.

55: **Hot Spare Disk:** هو غير مستخدمة ومستخدمة تلقائيا بواسطة منتج الـ RAID لو فشل القرص في تبديل القرص المعطوب وأعاد بناء مجموعة RAID لو أمكن.

56: **RAID بمفرده** لا يستطيع أن يمنع أو يكتشف البيانات المعطوبة أو الأخطاء الأخرى. فقط مع الأقراص التي باءت بالفشل.

57: **Checksums:** تحفظ مع المؤشر في object، للكشف إذا كان الـ object هو الصحيح أو ما إذا كان قد تغير.

58: **تنفيذ Stable-Storage يكون على طريقتين:**

a) **تكرار المعلومات** على أكثر من nonvolatile storage media مع وسائط الفشل (failure modes) المستقلة.
b) **تحديث المعلومات** في controlled manner للتأكيد بأننا نستطيع استعادة بيانات الـ stable بعد حدوث أي فشل خلال
نقل البيانات أو عند عملية الاستعادة.

59: **كتابة القرص لديه 3 نتائج (outcomes):**

1) **الانتهاء بنجاح (Successful completion):** عند كتابة البيانات بشكل صحيح على القرص.
2) **الفشل الجزئي (Partial Failure):** الفشل قد يحصل من منتصف عملية النقل لذا بعض القطاعات قد كتب عليها البيانات
الجديدة، والقطاعات التي كتب عليها قد تكون معطوبة.
3) **الفشل التام (Total failure):** الفشل الذي قد يحصل قبل عملية حدوث الكتابة على القرص، لذا قيم البيانات السابقة على
القرص تبقى على حالها.

60: **النظام يقوم بصيانة 2 physical blocks لكل block منطقي ويعمل الاتي:**

a) الكتابة إلى 1^{st} physical.
b) عند نجاحه، يتم الكتابة إلى 2^{nd} physical.
c) يظهر فقط العناصر المكتملة بعد اكتمال الكتابة الثانية بنجاح.

END of chapter 12