31268 Web Systems

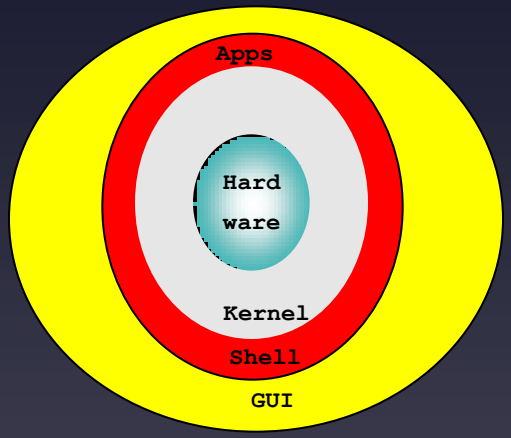
Week 2: Operating Systems

Part 3: File Storage

31268 Web Systems

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

Week 2

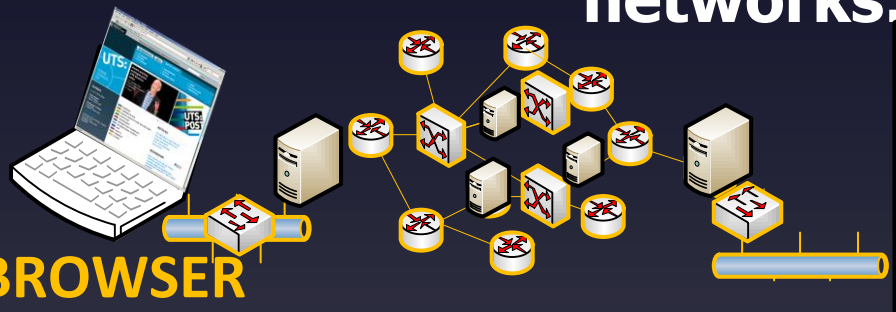


Operating Systems
File Storage

The web…

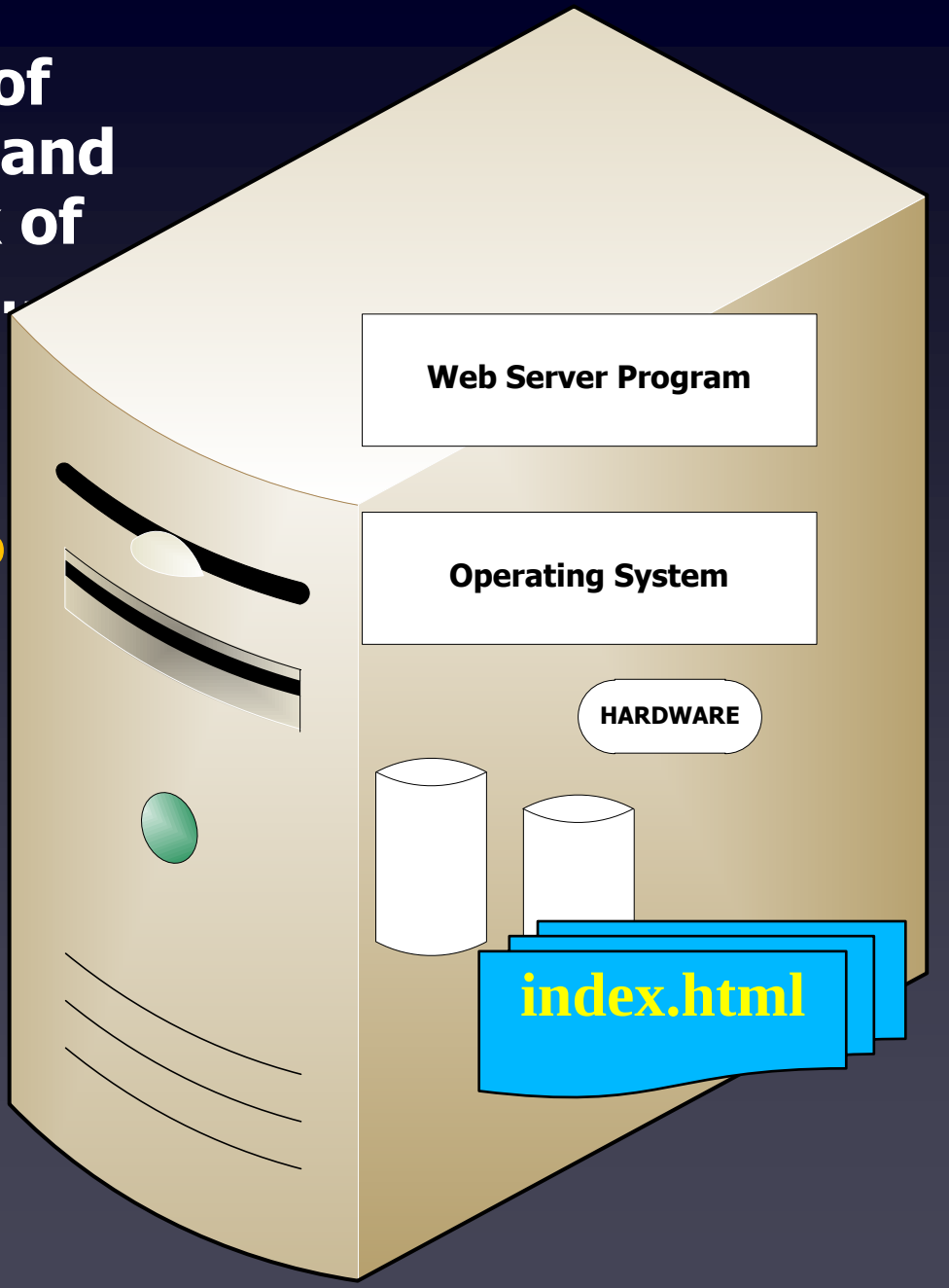
- Is It …

A bunch of computers and a network of networks.



BROWSER

… and a whacking **big** computer running the **web site program** on an **operating system** running on **HARDWARE**

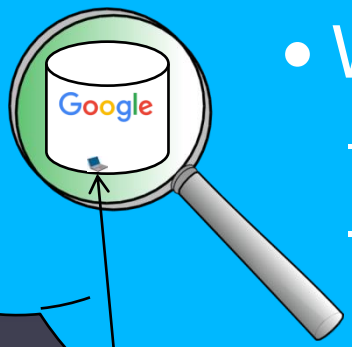


The Web

WWW ++

- How big is the internet??
  - 1.003 Billion web sites? ++
    - http://news.netcraft.com/archives/category/web-server-survey/

→ But each website consists of many html pages and images??



- What about Google?
  - Est 2014 indexed 200Tb, 2016??
  - But this is estimated 0.004% of the internet

- What about Laptop?

Chris' Laptop: 400,311 files/directories = 178Gb

The Web

- So many files!
- Question: Does the web, or google, or UTS, or your laptop store all the files in 1 single directory?

```
C:\> dir   web:\
934,856,356,384,959,437,893,947,373,248,094,
837,833,417,456,885,789,347,567,890  file(s)
1 dir(s)
∞ bytes free
```

File systems and file manipulation

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

How is data stored on computers of the internet?

Google: A LOT of drives connected to LOTS of servers in 20+ data centres around the world..

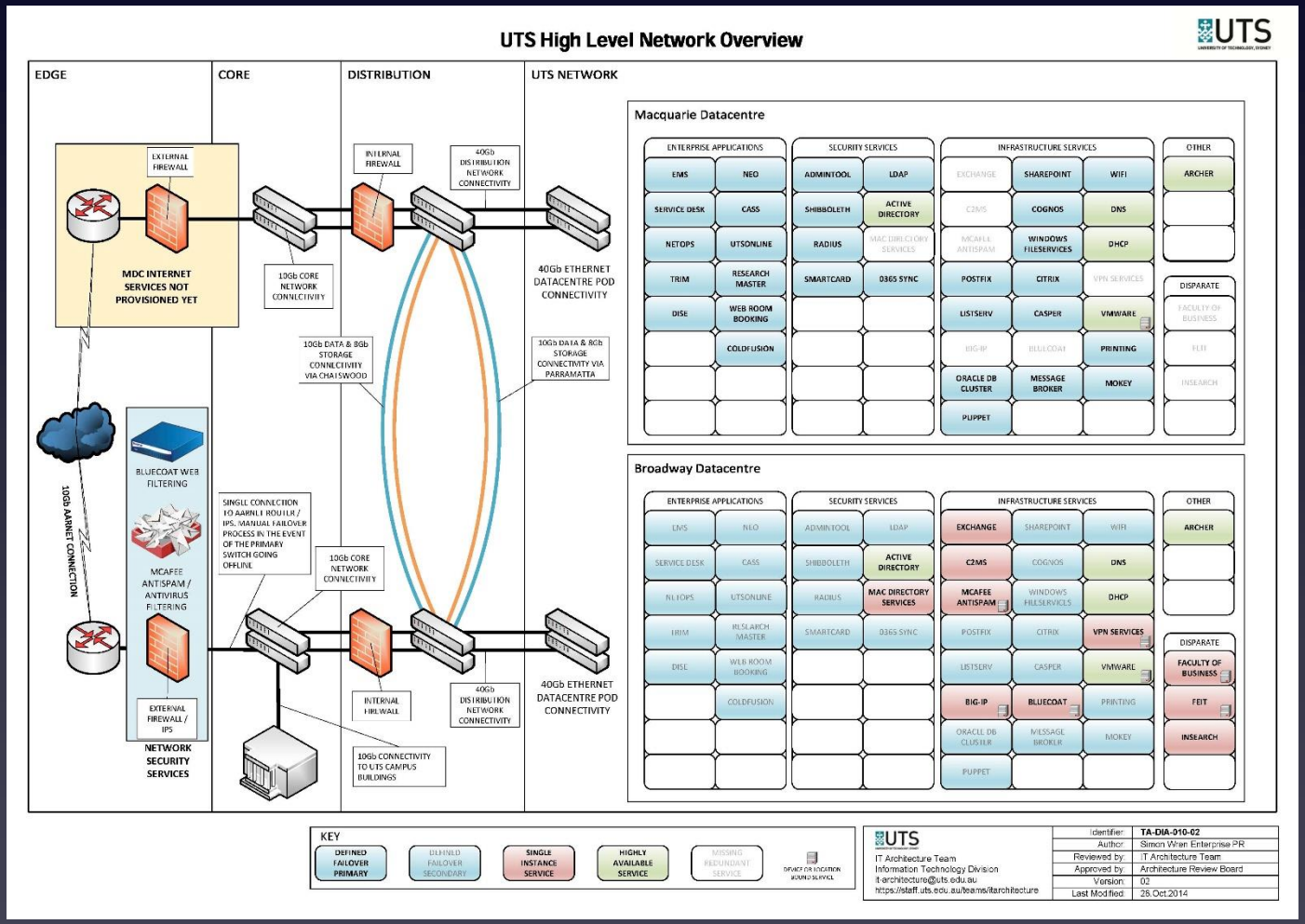
→Called a distributed file system

UTS: we have 2 data centres – in Building 1 and in Macquarie Park

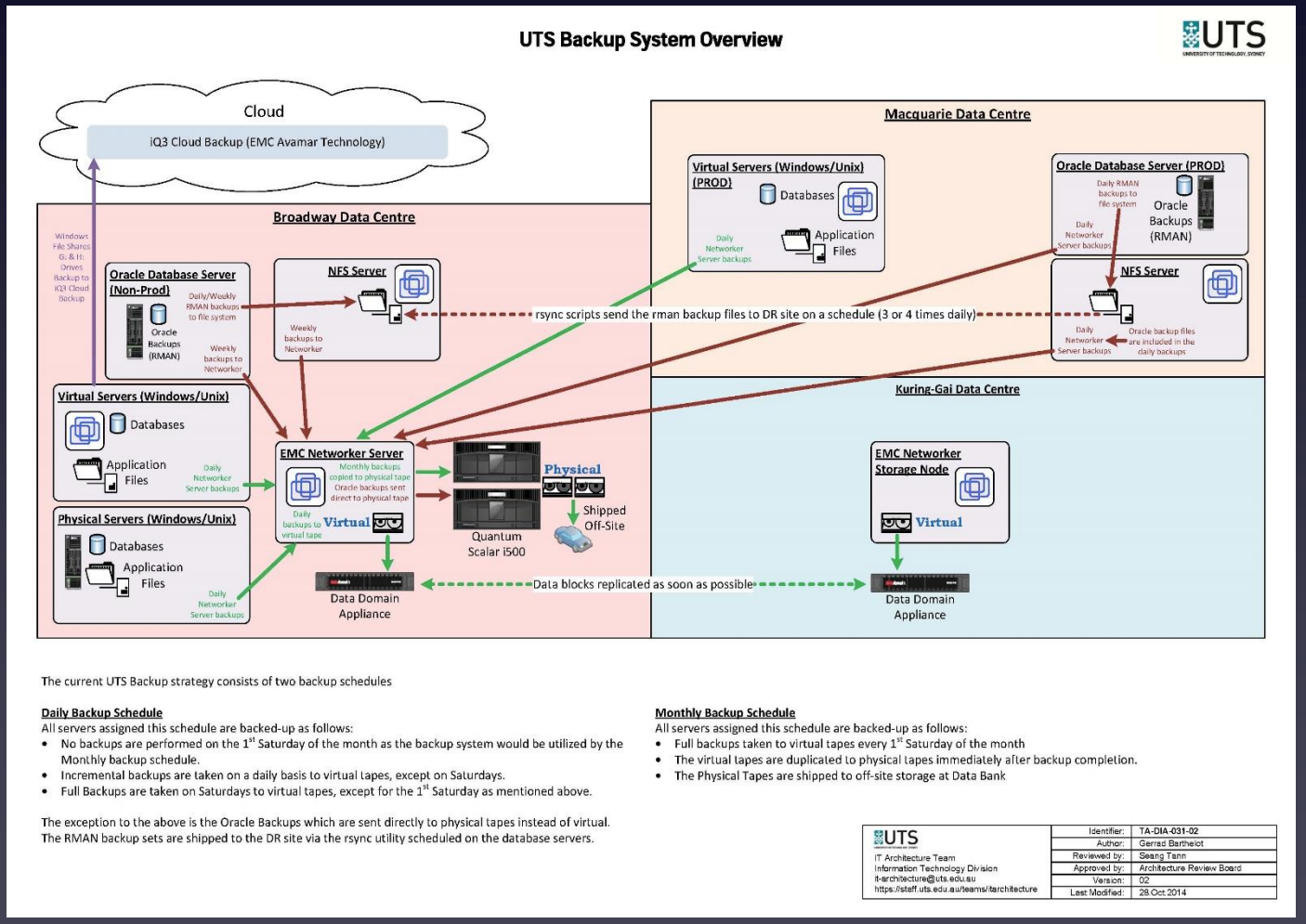
→ Central file system called a SAN

→ plus drives on various servers

UTS Data Centre



UTS Data Centre



The current UTS Backup strategy consists of two backup schedules

Daily Backup Schedule

All servers assigned this schedule are backed-up as follows:

- No backups are performed on the 1st Saturday of the month as the backup system would be utilized by the Monthly backup schedule.
- Incremental backups are taken on a daily basis to virtual tapes, except on Saturdays.
- Full Backups are taken on Saturdays to virtual tapes, except for the 1st Saturday as mentioned above.

The exception to the above is the Oracle Backups which are sent directly to physical tapes instead of virtual. The RMAN backup sets are shipped to the DR site via the rsync utility scheduled on the database servers.

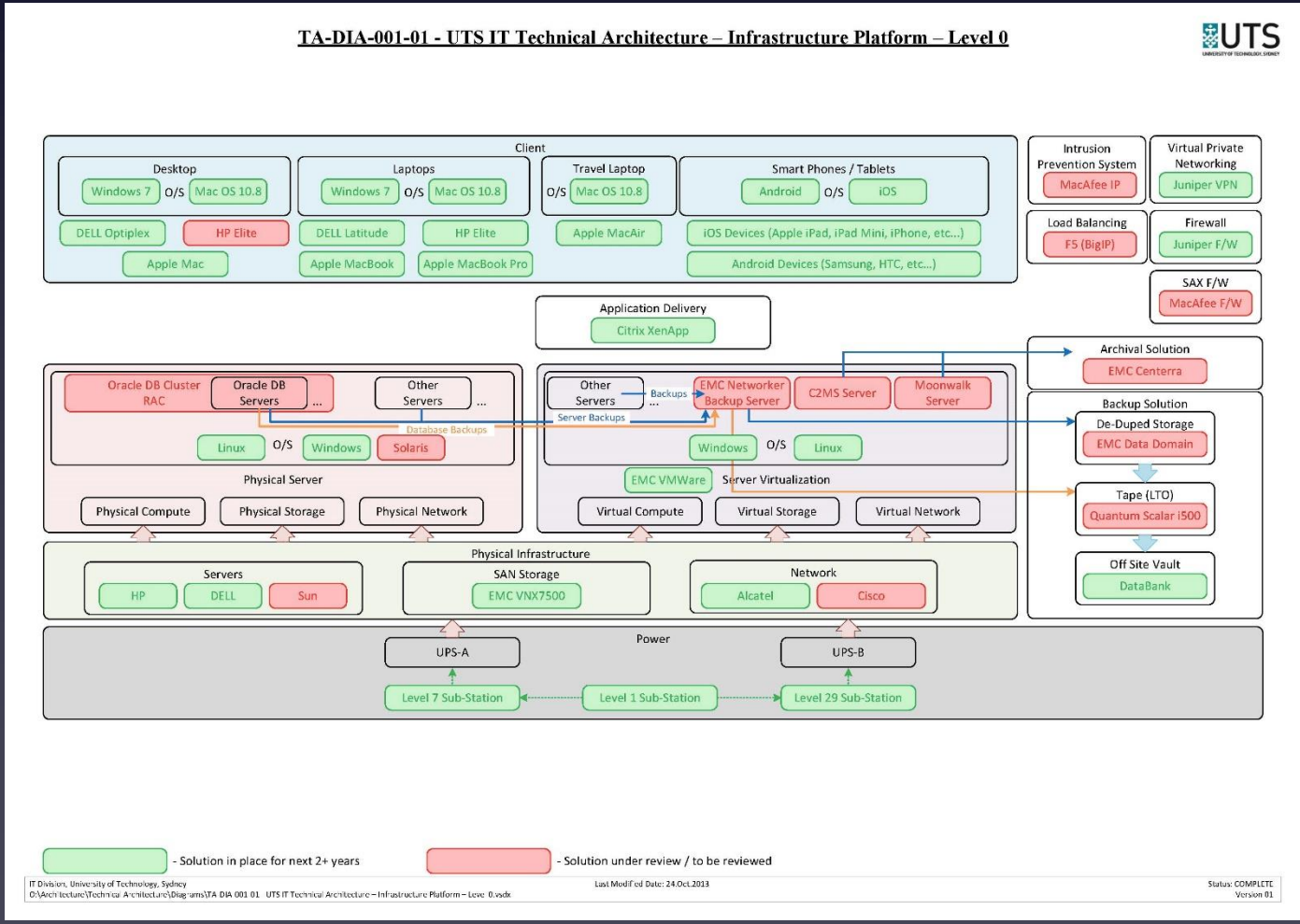
Monthly Backup Schedule

All servers assigned this schedule are backed-up as follows:

- Full backups taken to virtual tapes every 1st Saturday of the month
- The virtual tapes are duplicated to physical tapes immediately after backup completion.
- The Physical Tapes are shipped to off-site storage at Data Bank

UTS		
IT Architecture Team Information Technology Division it-architecture@uts.edu.au https://staff.uts.edu.au/teams/itarchitecture	Identifier:	TA-DIA-031-02
	Author:	Gerrad Barthelot
	Reviewed by:	Seang Tann
	Approved by:	Architecture Review Board
	Version:	02
	Last Modified:	28.Oct.2014

UTS Data Centre



Drives, Disks etc

- Typical personal computer has 1 physical storage device

→ typically a Hard disk eg: Seagate, WD, Samsung, Maxtor e.g 1Tb $80

→ Solid State Drives (SSD) are becoming very popular: e.g 120Gb $80

File systems and file manipulation

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

Let's look at how hard disk (& SSD!) is managed and organized by an operating system:

- Disk physical structure
- Disk logical structures
- File allocation methods

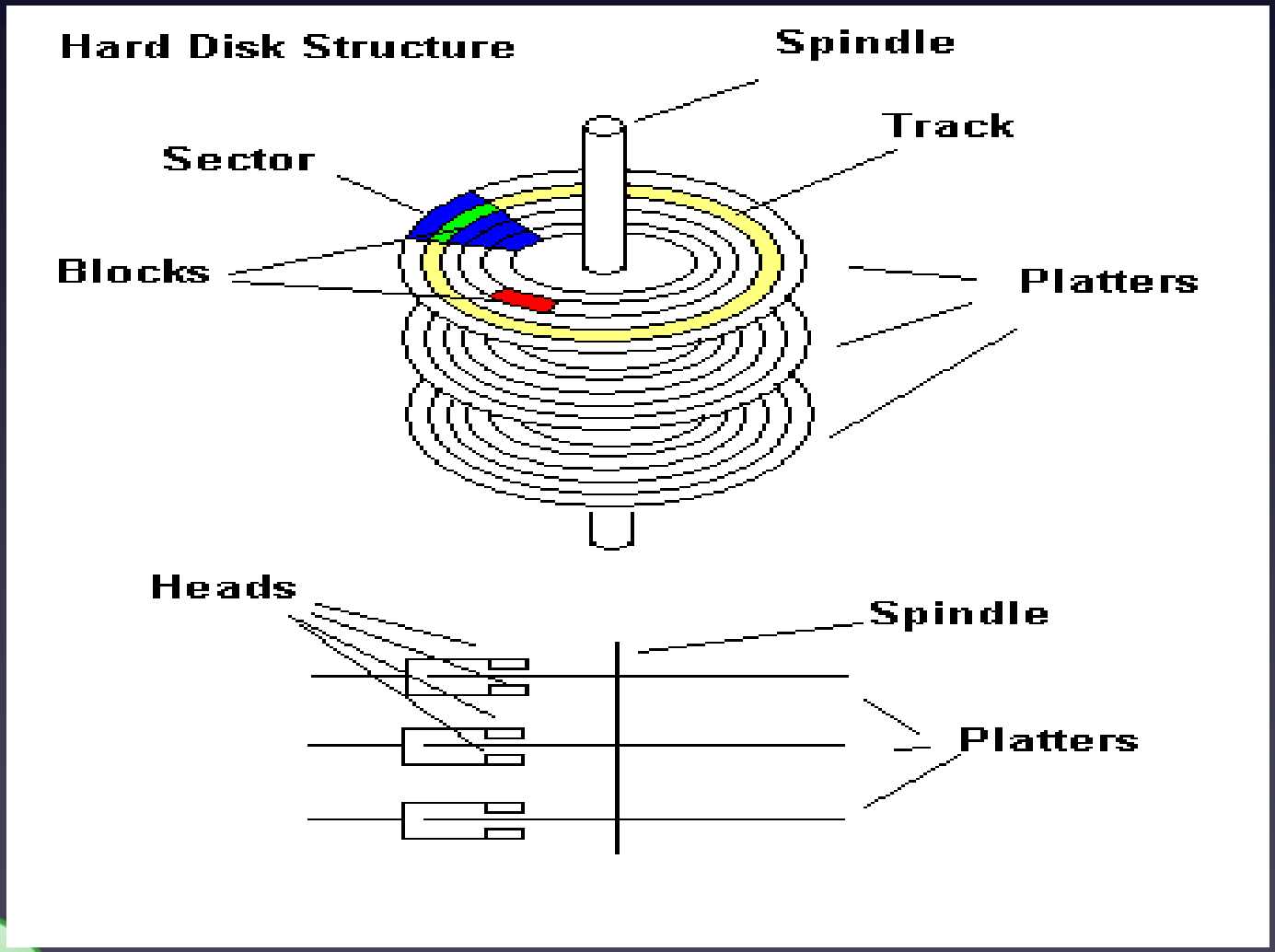
Disk Physical Structure:

A physical hard disk is organized into:

- **Tracks**: Concentric rings on the platter.
- **Heads**: reads data from a platter
- **Cylinders**: Collection of all tracks on platters
  - which are horizontally in the same position.
- **Sectors**: part of a track for data

Low Level Hard-Disk Data Storage

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology



Disk Structure

- A disk is a stack of magnetic platters.
  - This stack is divided into cylinders
  - each cylinder contains circular tracks
    - which are in turn divided into sectors.

Disk Structure

- A disk is a stack of magnetic platters.
  - This stack is divided into cylinders
  - each cylinder contains circular tracks
    - which are in turn divided into sectors.
- Read/write operations are provided by the disk heads
  - These move concurrently along the fixed disk arm.
- The disk itself rotates with constant angular velocity to provide access to every sector.

Disk Structure

- SSD have no moving parts, but emulate a rotating hard disk!
- New technologies to make better use of SSD eg PCIe SSD, Non-Volatile Memory express (NVMe)

Disk Formatting

- Formatting is the operation which creates the **physical** disk structure.
- Formatting is organizing and marking the surface of a disk into tracks, sectors , and cylinders.
- http://www.tldp.org/LDP/sag/html/g3130.html

Disk Formatting

- Formatting is the operation which creates the **physical** disk structure.
- Formatting is organizing and marking the surface of a disk into tracks, sectors , and cylinders.
- It is also sometimes (incorrectly) a term used to signify the action of writing a filesystem to a disk (especially in the MS Windows/MS DOS world).

File systems and file manipulation

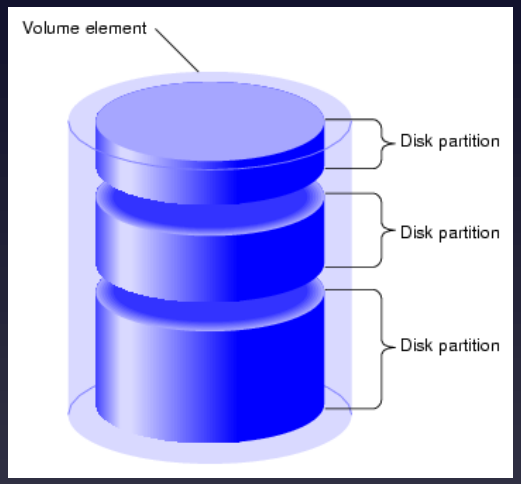
UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

Let's look at the following parts of a file system:

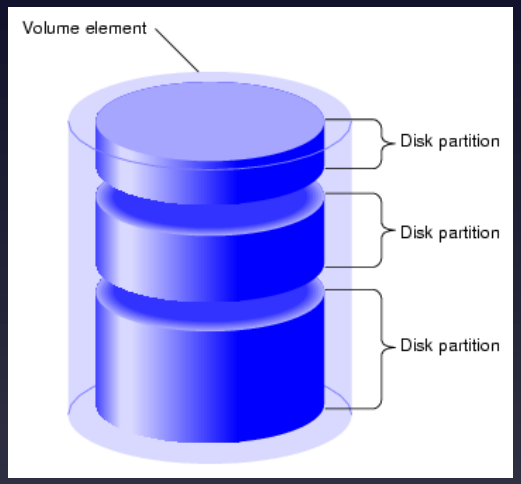
- Disk physical structure
- **Disk logical structures**
- File allocation methods

Disk Logical Structure

- **Partitions**: Disks can be subdivided into partitions
  – each is an independent storage device.



Disk Logical Structure

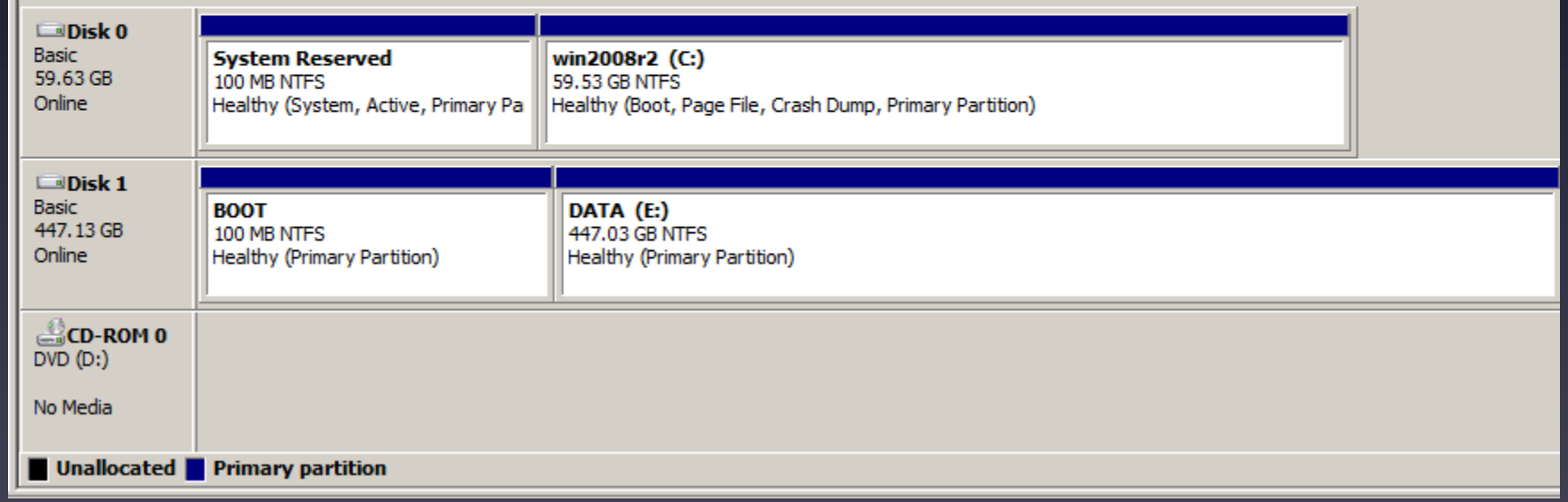


- **Partitions**: Disks can be subdivided into partitions
  – each is an independent storage device.

- **Blocks**: The operating system views all the disk space as an array of fixed size logical blocks.
  – A logical block is the smallest unit of data to transfer.

Example: Windows Partitions

Volume	Layout	Type	File System	Status	Capacity	Free Space	% Free	Fault Tolerance	Overhead
BOOT	Simple	Basic	NTFS	Healthy (Primary Partition)	100 MB	84 MB	84 %	No	0%
DATA (E:)	Simple	Basic	NTFS	Healthy (Primary Partition)	447.03 GB	320.42 GB	72 %	No	0%
System Reserved	Simple	Basic	NTFS	Healthy (System, Active, Primary Partition)	100 MB	71 MB	71 %	No	0%
win2008r2 (C:)	Simple	Basic	NTFS	Healthy (Boot, Page File, Crash Dump, Primary Partition)	59.53 GB	21.41 GB	36 %	No	0%



File systems and file manipulation

Let's look at the following parts of a file system:

- Disk physical structure
- Disk logical structures
- **File allocation methods**

File Allocation

- Say I have a file called ...

phonebook.txt

1	Solomon A	155 Kent St Millers Point 2000 (02) 9241 5300
2	Solomon A	70 B High St North Sydney 2060 (02) 9957 3280
3	Solomon A	10 Thyra Rd Palm Beach 2108 (02) 9974 2145
4	Solomon A	23 Cadigal Ave Pyrmont 2009 (02) 9518 8979
5	Solomon A	112 Chuter Ave Ramsgate Beach 2217 (02) 9529 5676
6	Solomon A	15 Faraday Ave Rose Bay 2029 (02) 9371 9293
7	Solomon A & G	8 Comeroy Crs Frenchs Forest 2086 (02) 9451 1051
8	Solomon A J	30 Raymond St Eastwood 2122 (02) 9876 2476
9	Solomon A J	155 Kent St Sydney 2000 (02) 9241 5888
10	Solomon A L	51 Darley Rd Randwick 2031 (02) 9399 5161
11	Solomon A M	Fleet St Carlton 2218 (02) 9593 0620

File Allocation

- **Block**: Space is allocated to a file as one or more blocks

phonebook.txt

Block

1	Solomon A	155 Kent St Millers Point 2000 (02) 9241 5300
2	Solomon A	70 B High St North Sydney 2060 (02) 9957 3280
3	Solomon A	10 Thyra Rd Palm Beach 2108 (02) 9974 2145
4	Solomon A	23 Cadigal Ave Pyrmont 2009 (02) 9518 8979
5	Solomon A	112 Chuter Ave Ramsgate Beach 2217 (02) 9529 5676
6	Solomon A	15 Faraday Ave Rose Bay 2029 (02) 9371 9293
7	Solomon A & G	8 Comeroy Crs Frenchs Forest 2086 (02) 9451 1051
8	Solomon A J	30 Raymond St Eastwood 2122 (02) 9876 2476
9	Solomon A J	155 Kent St Sydney 2000 (02) 9241 5888
10	Solomon A L	51 Darley Rd Randwick 2031 (02) 9399 5161
11	Solomon A M	Fleet St Carlton 2218 (02) 9593 0620

File Allocation

- **Block**: Space is allocated to a file as one or more blocks

phonebook.txt

Block

1	Solomon A	155 Kent St Millers Point 2000 (02) 9241 5300
2	Solomon A	70 B High St North Sydney 2060 (02) 9957 3280
3	Solomon A	10 Thyra Rd Palm Beach 2108 (02) 9974 2145
4	Solomon A	23 Cadigal Ave Pyrmont 2009 (02) 9518 8979
5	Solomon A	112 Chuter Ave Ramsgate Beach 2217 (02) 9529 5676
6	Solomon A	15 Faraday Ave Rose Bay 2029 (02) 9371 9293
7	Solomon A & G	8 Comeroy Crs Frenchs Forest 2086 (02) 9451 1051
8	Solomon A J	30 Raymond St Eastwood 2122 (02) 9876 2476
9	Solomon A J	155 Kent St Sydney 2000 (02) 9241 5888
10	Solomon A L	51 Darley Rd Randwick 2031 (02) 9399 5161
11	Solomon A M	Fleet St Carlton 2218 (02) 9593 0620

- **Directory**: is a table of information that the OS uses to locate blocks associated with files on a disk.

File Allocation

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

- **Block**: Space is allocated to a file as one or more blocks

phonebook.txt

Block

1	Solomon A	155 Kent St Millers Point 2000 (02) 9241 5300
2	Solomon A	70 B High St North Sydney 2060 (02) 9957 3280
3	Solomon A	10 Thyra Rd Palm Beach 2108 (02) 9974 2145
4	Solomon A	23 Cadigal Ave Pyrmont 2009 (02) 9518 8979
5	Solomon A	112 Chuter Ave Ramsgate Beach 2217 (02) 9529 5676
6	Solomon A	15 Faraday Ave Rose Bay 2029 (02) 9371 9293
7	Solomon A & G	8 Comeroy Crs Frenchs Forest 2086 (02) 9451 1051
8	Solomon A J	30 Raymond St Eastwood 2122 (02) 9876 2476
9	Solomon A J	155 Kent St Sydney 2000 (02) 9241 5888
10	Solomon A L	51 Darley Rd Randwick 2031 (02) 9399 5161
11	Solomon A M	Fleet St Carlton 2218 (02) 9593 0620

- **Directory**: is a table of information that the OS uses to locate blocks associated with files on a disk.

/home/cw/

phonebook.tx

students.txt

staff.txt

Problem: Time to access file

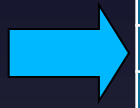
- Here's my phonebook

1	Solomon A	155 Kent St Millers Point 2000 (02) 9241 5300
2	Solomon A	70 B High St North Sydney 2060 (02) 9957 3280
3	Solomon A	10 Thyra Rd Palm Beach 2108 (02) 9974 2145
4	Solomon A	23 Cadigal Ave Pyrmont 2009 (02) 9518 8979
5	Solomon A	112 Chuter Ave Ramsgate Beach 2217 (02) 9529 5676
6	Solomon A	15 Faraday Ave Rose Bay 2029 (02) 9371 9293
7	Solomon A & G	8 Comeroy Crs Frenchs Forest 2086 (02) 9451 1051
8	Solomon A J	30 Raymond St Eastwood 2122 (02) 9876 2476
9	Solomon A J	155 Kent St Sydney 2000 (02) 9241 5888
10	Solomon A L	51 Darley Rd Randwick 2031 (02) 9399 5161
11	Solomon A M	Fleet St Carlton 2218 (02) 9593 0620

- Question:
  - How long (# disk accesses) does it take to find the block containing the 3rd Solomon?

Problem: Time to access file

- Here's my phonebook



1	Solomon A	155 Kent St Millers Point 2000 (02) 9241 5300
2	Solomon A	70 B High St North Sydney 2060 (02) 9957 3280
3	Solomon A	10 Thyra Rd Palm Beach 2108 (02) 9974 2145
4	Solomon A	23 Cadigal Ave Pyrmont 2009 (02) 9518 8979
5	Solomon A	112 Chuter Ave Ramsgate Beach 2217 (02) 9529 5676
6	Solomon A	15 Faraday Ave Rose Bay 2029 (02) 9371 9293
7	Solomon A & G	8 Comeroy Crs Frenchs Forest 2086 (02) 9451 1051
8	Solomon A J	30 Raymond St Eastwood 2122 (02) 9876 2476
9	Solomon A J	155 Kent St Sydney 2000 (02) 9241 5888
10	Solomon A L	51 Darley Rd Randwick 2031 (02) 9399 5161
11	Solomon A M	Fleet St Carlton 2218 (02) 9593 0620

- Question:
  - How long (# disk accesses) does it take to find the block containing the 3rd Solomon?

File Allocation Methods

Answer:

It all depends on the way the file is allocated.

File Allocation Methods

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

Answer:

It all depends on the way the file is allocated.

There are three common types of file allocation:

- **Contiguous** Allocation
- **Chained** or **Linked** Allocation
- **Indexed** Allocation (e.g inode)

File Storage - Contiguous

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

- A single contiguous set of blocks is allocated to a file at the time of file creation.

03 So	lo mo	n A	m 10	Th yr	a rd	Pa lm	Be ac	21 08	(02)9	97 4-	21 45
1	2	3	4	5	6	7	8	9	10	11	12	13

- To access information in block B, this information resides at block number starting block + B

File Storage - Contiguous

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

- A single contiguous set of blocks is allocated to a file at the time of file creation.

03 So	lo mo	n A	m 10	Th yr	a rd	Pa lm	Be ac	21 08	(02)9	97 4-	21 45
1	2	3	4	5	6	7	8	9	10	11	12	13

- To access information in block B, this information resides at block number starting block + B
- Supports random access: you know exactly where every block is after the starting block.
- Fragmentation of unused space (external fragmentation) will occur, needs compaction.
- Often used in magnetic tapes rather than disks

Contiguous Animation

Chained Allocation

- File is written as a collection of non-contiguous blocks
- File is implemented as a linked list of blocks

Chained Allocation

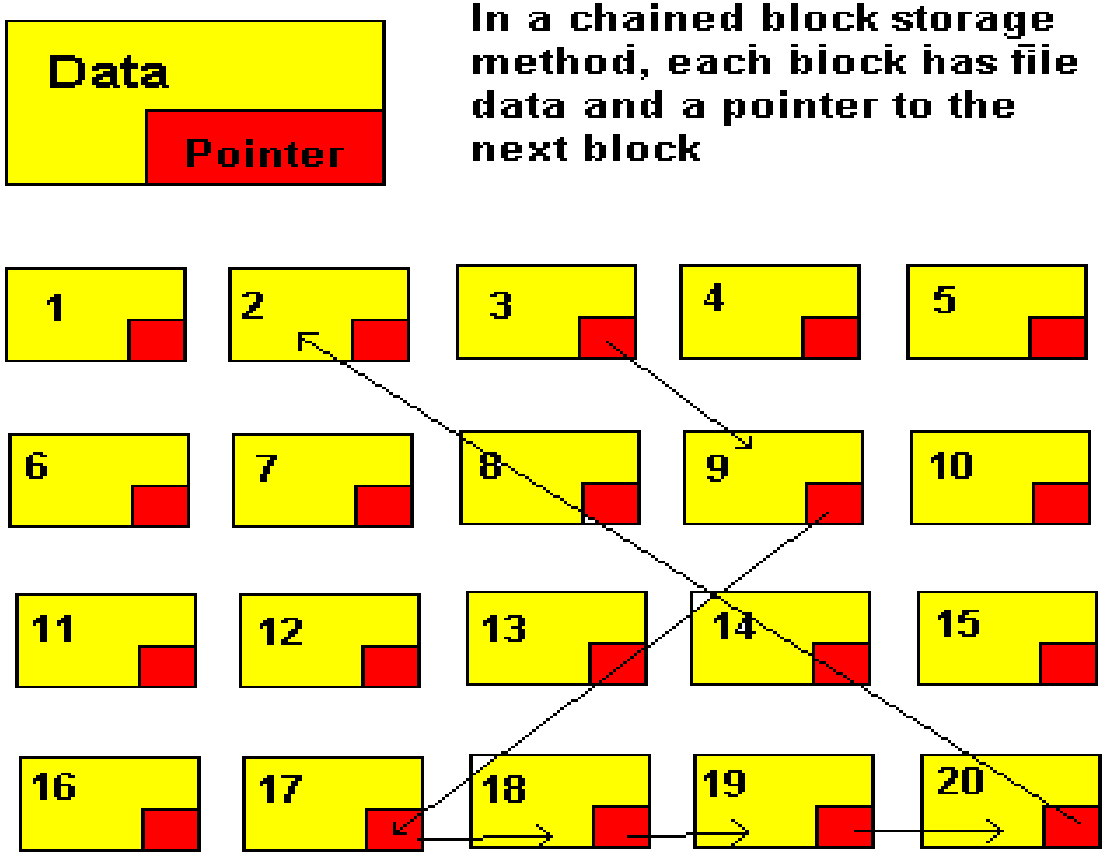
- File is written as a collection of non-contiguous blocks
- File is implemented as a linked list of blocks
- Each block contains a (pointer to) the address of next block.
  - Last block contains invalid (negative) number (End-Of-File marker)

Chained Allocation

- File is written as a collection of non-contiguous blocks
- File is implemented as a linked list of blocks
- Each block contains a (pointer to) the address of next block.
  - Last block contains invalid (negative) number (End-Of-File marker)
- Directory entry contains the head (starting) block number and length of the file
- Chained is good for sequential access, bad for random access

File Storage – Chained Allocation

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology



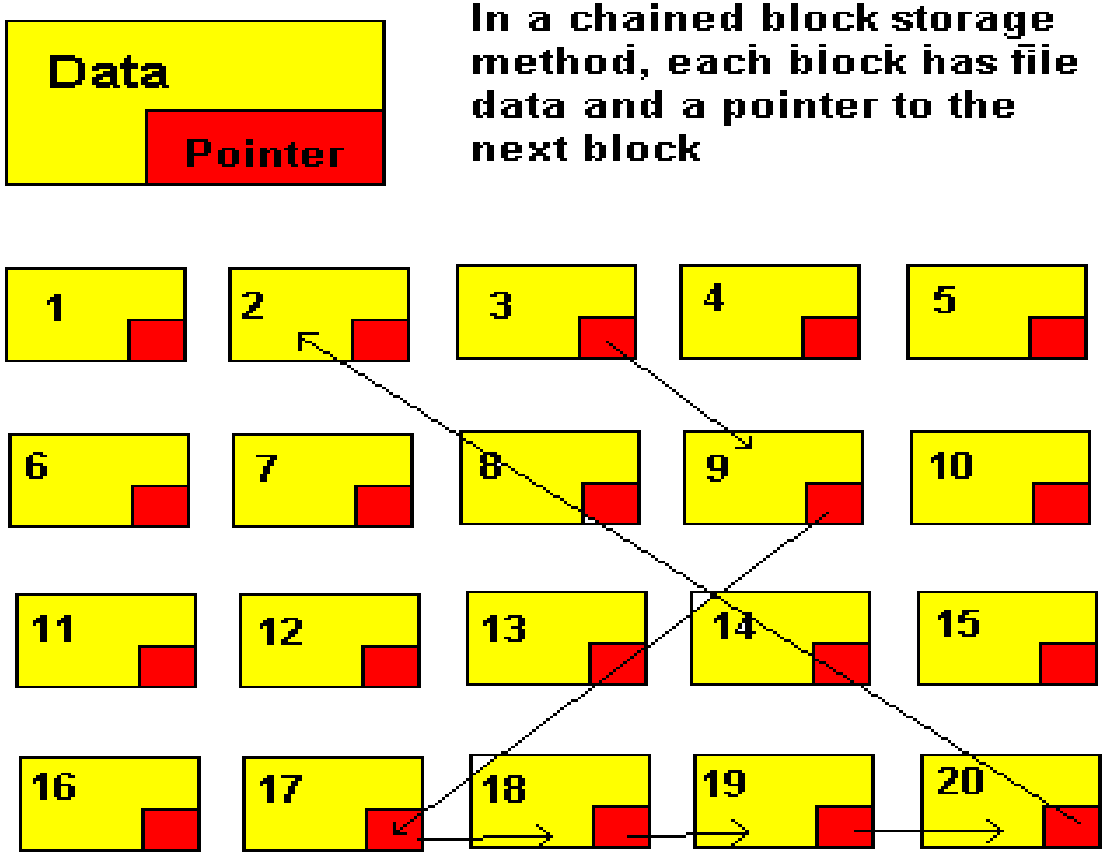
A simple chained block storage method overcomes the disadvantages of a contiguous file storage system. However, a chained block system has its own disadvantages.

To get to data in block N the operating system has to read N blocks. In large files this makes access to the end blocks very slow.

File Storage – Chained Allocation

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

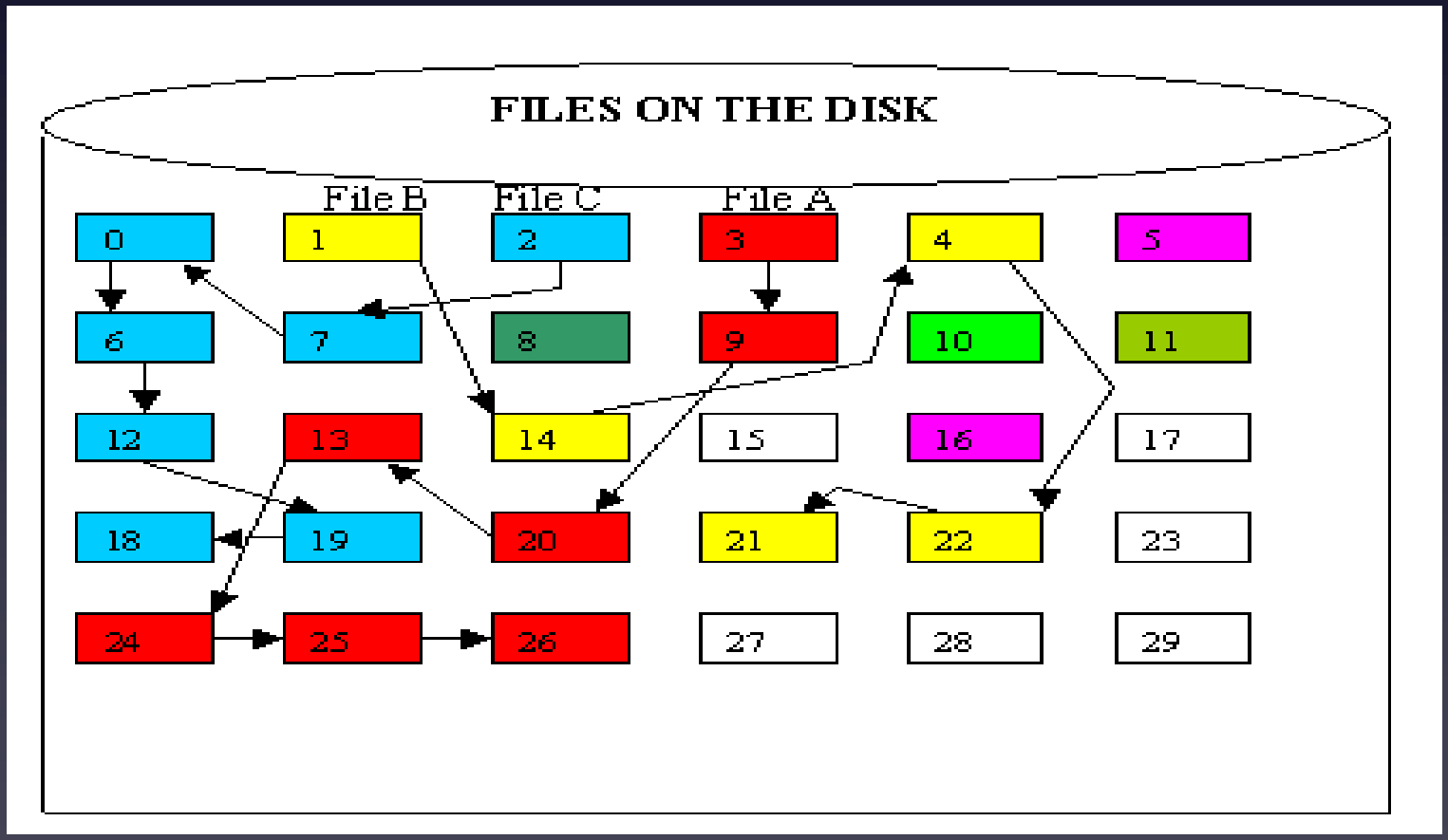
Sequence is: 3, 9, 17, 18, 19, 20, 2



A simple chained block storage method overcomes the disadvantages of a contiguous file storage system. However, a chained block system has its own disadvantages.

To get to data in block N the operating system has to read N blocks. In large files this makes access to the end blocks very slow.

Example of chained allocation

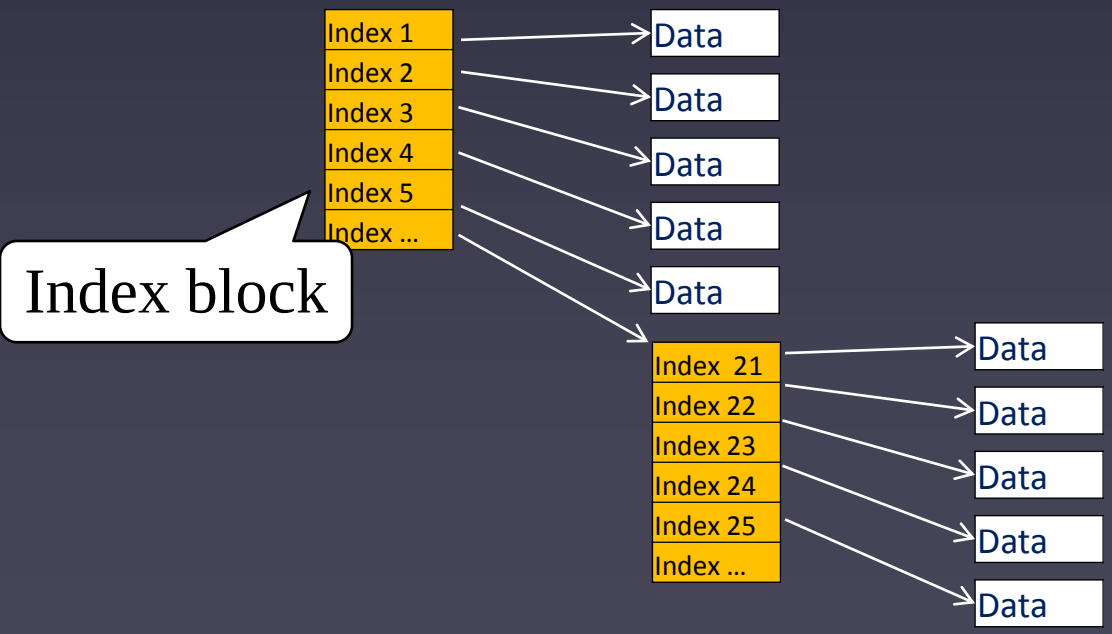


Linked animation

Indexed Allocation

- ➔ "Tree" based allocation system
- A special "index" data block will contain a list of data blocks# for the file

- If the file is too big, the "index" data block will point to other "index" data blocks

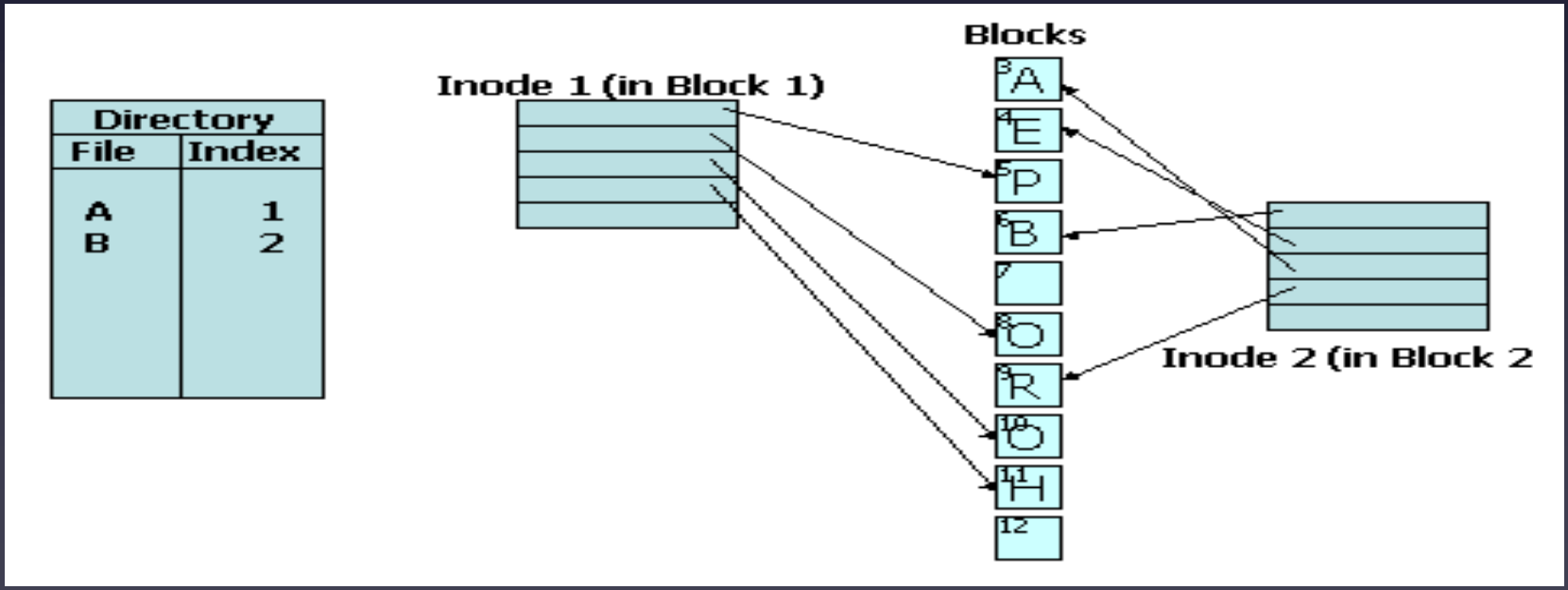


Indexed Animation

File storage: Indexed – UNIX

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

- On UNIX, files are stored on blocks called an **inode**
- Each file/directory is referenced by an inode



- Very efficient use of space and fast to read blocks

iNodes and Directories in Unix

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Information Technology

- The inode system is used in unix.
- All inodes are numbered.
- Special blocks/file on disk called a directory
- Directories contain the names of files and the inode number for the file.
- Notice that the inode structure is a **tree**

What does an inode store?

- File Metadata: Size, Owner id, Group id, Permissions, Timestamps
  ➔ *NOTE: inode does NOT store name of file*!!!
- Pointers to the blocks that store the files data

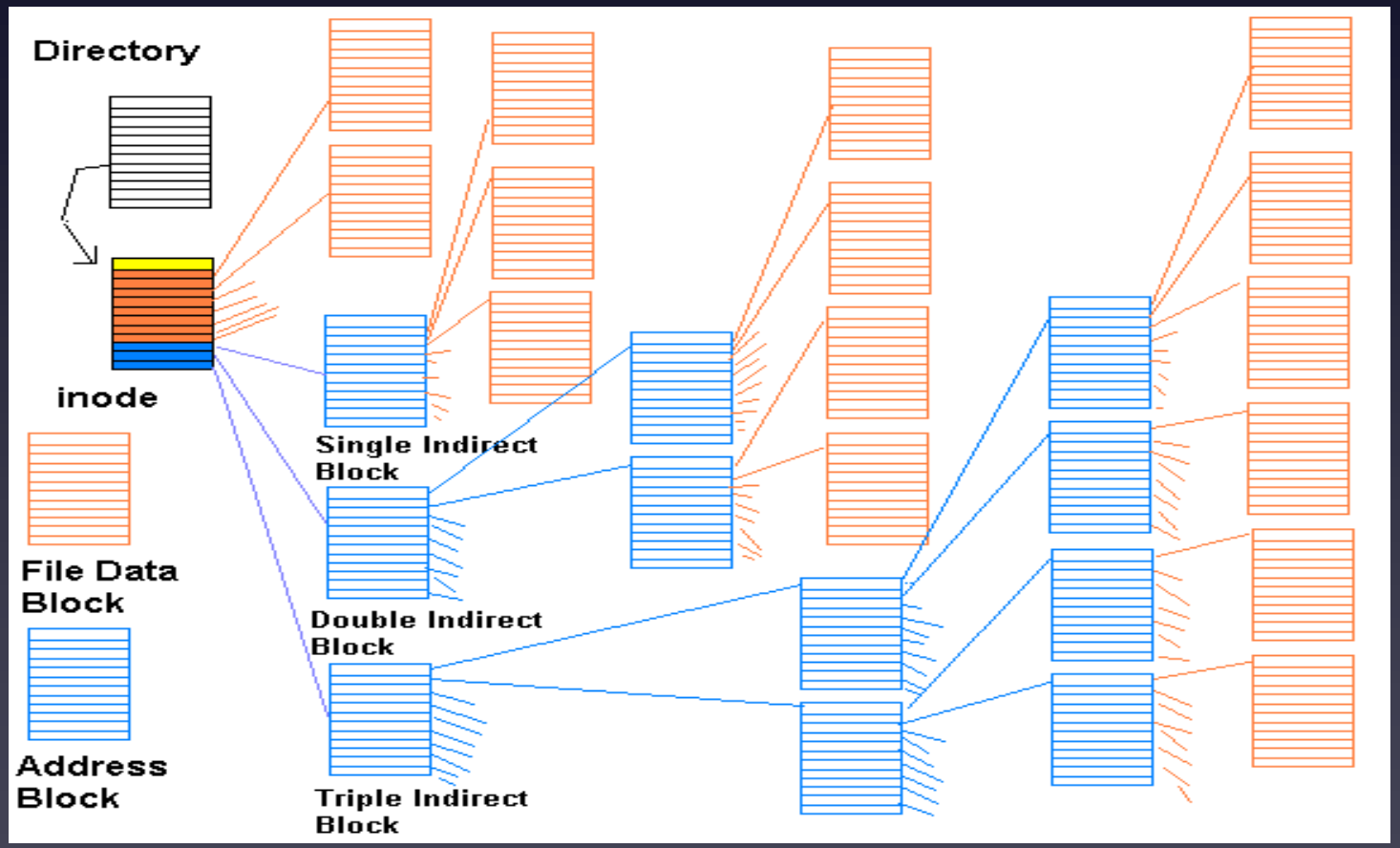
What does an inode store?

- File Data
- Pointers to the blocks that store the files data
- (optionally)
  - "Single indirect block"
    → pointer to a disk block → contains an index of pointers to data blocks
  - "double indirect block"
    → Pointer to more "single indirect blocks"
  - "triple indirect block"
    → pointer to more "double indirect blocks"

File Storage iNodes

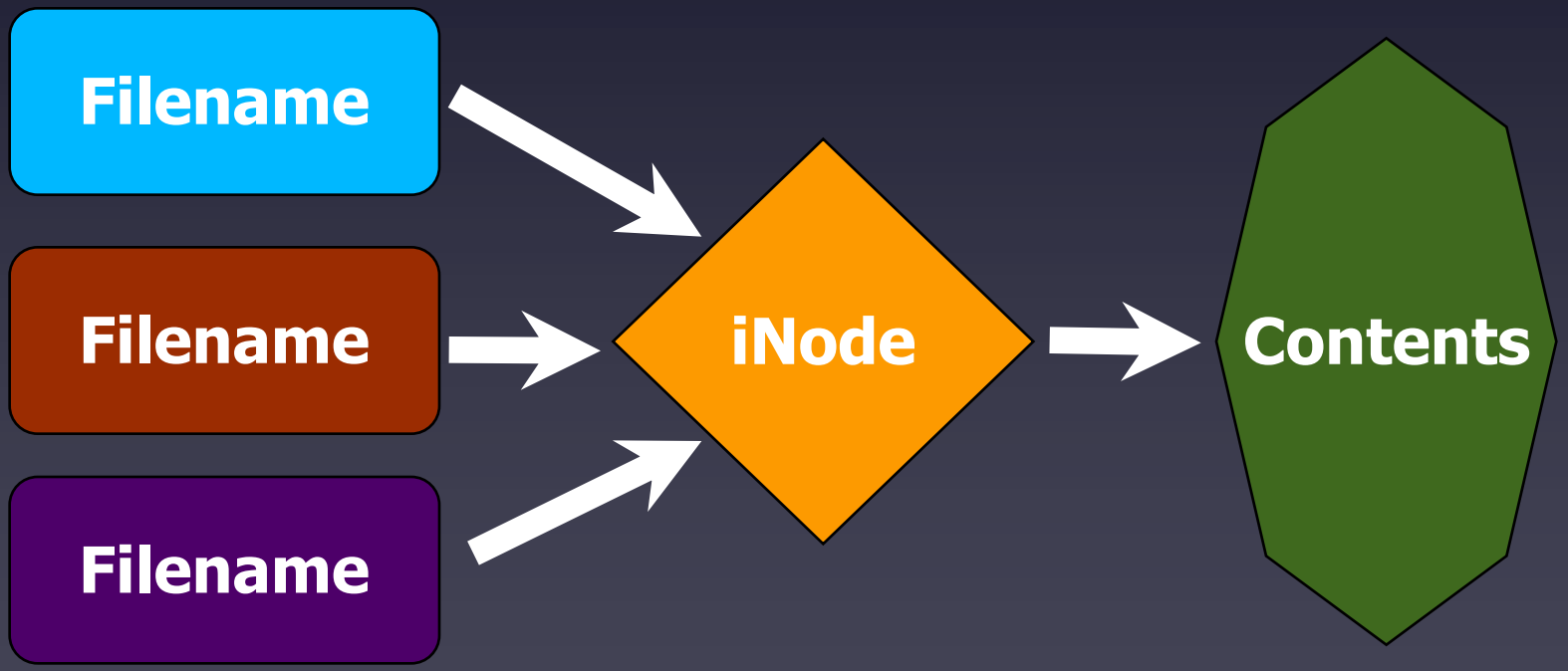
File Data: Size, Owner id, Group id, Permissions, Timestamps
Pointer to File Data Block 1
Pointer to File Data Block 2
Pointer to File Data Block 3
Pointer to File Data Block 4
Pointer to File Data Block 5
Pointer to File Data Block 6
Pointer to File Data Block 7
Pointer to File Data Block 8
Pointer to File Data Block 9
Pointer to File Data Block …. & so on
Pointer to Single indirect block
Pointer to Double indirect block
Pointer to Triple indirect block

iNodes


Directory
inode
File Data Block
Address Block
Single Indirect Block
Double Indirect Block
Triple Indirect Block

Why are iNodes so cool?

Their structure allows for some interesting behavior with respect to moving and deleting files.



Conclusion

- Which file allocation type depends on your needs?
- Contiguous is great for "direct" storage & tiny file systems
- Chained/Linked is good for archival (eg backup)
- Indexed is the only reasonable option for large systems