FAILURE CLASSIFICATION, LOG-BASED RECOVERY & SHADOWING

DATABASE MANAGEMENT SYSTEM

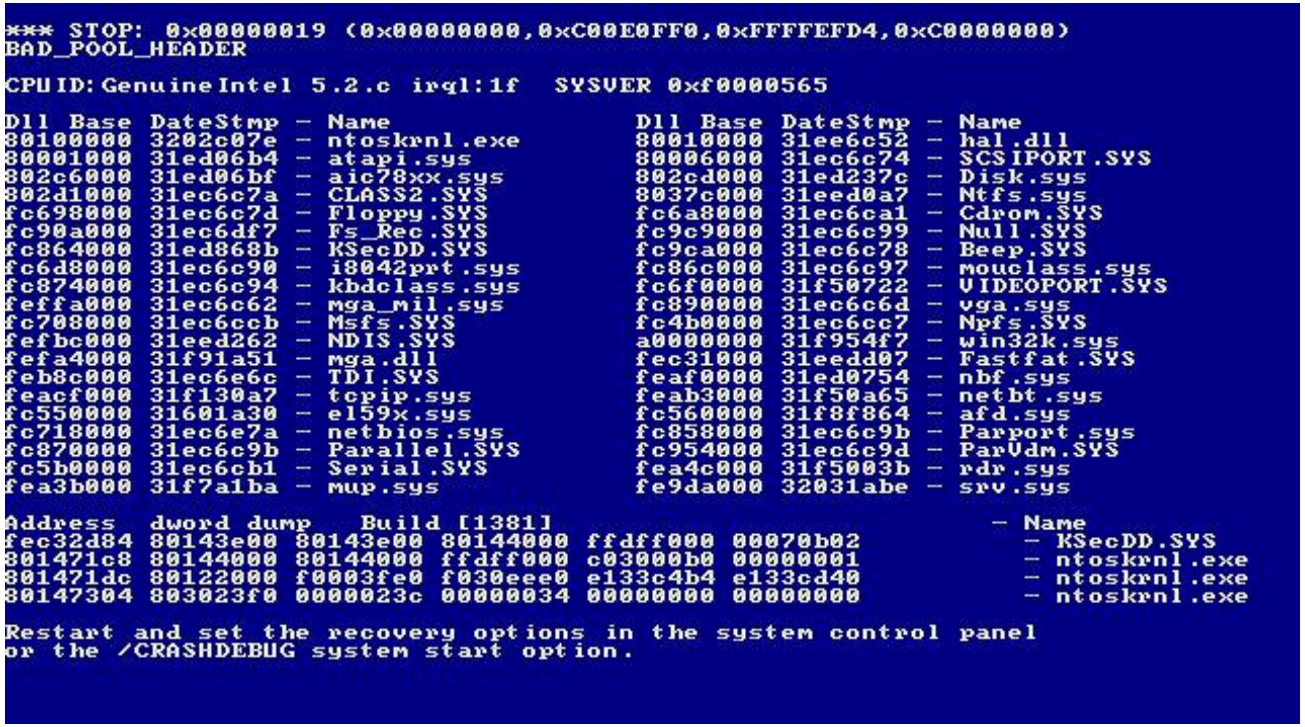
Sujan Tamrakar

Failure Classification

- Computer is subject to failure from variety of causes like
  - Disk crash
  - Power outage
  - Software error / logical issue
  - Physical damage / hardware defects
  - Improper usage
  - Environmental factor
- Results occurring from failures:
  - Loss of information
  - Corruption of information
  - Physical damage
  - Defame of Image/Dignity
  - Alteration in program logic
  - Fatalities

Failure Classification

- Types of failures (after generalization)
  1. Transaction failure
  2. System crash
  3. Disk failure



```
*** STOP: 0x00000019 (0x00000000,0xC00E0FF0,0xFFFFEFD4,0xC0000000)
BAD_POOL_HEADER

CPUID:GenuineIntel 5.2.c irql:1f  SYSVER 0xf0000565

Dll Base DateStmp - Name               Dll Base DateStmp - Name
80100000 3202c07e - ntoskrnl.exe       80010000 31ee6c52 - hal.dll
80001000 31ed06b4 - atapi.sys          80006000 31ec6c74 - SCSIPORT.SYS
802c6000 31ed06bf - aic78xx.sys        802cd000 31ed237c - Disk.sys
802d1000 31ec6c7a - CLASS2.SYS         8037c000 31eed0a7 - Ntfs.sys
fc698000 31ec6c7d - Floppy.SYS         fc6a8000 31ec6ca1 - Cdrom.SYS
fc90a000 31ec6df7 - Fs_Rec.SYS         fc9c9000 31ec6c99 - Null.SYS
fc864000 31ed868b - KSecDD.SYS         fc9ca000 31ec6c78 - Beep.SYS
fc6d8000 31ec6c90 - i8042prt.sys       fc86c000 31ec6c97 - mouclass.sys
fc874000 31ec6c94 - kbdclass.sys       fc6f0000 31f50722 - VIDEOPORT.SYS
feffa000 31ec6c62 - mga_mil.sys        fc890000 31ec6c6d - vga.sys
fc708000 31ec6ccb - Msfs.SYS           fc4b0000 31ec6cc7 - Npfs.SYS
fefbc000 31eed262 - NDIS.SYS           a0000000 31f954f7 - win32k.sys
fefa4000 31f91a51 - mga.dll            fec31000 31eedd07 - Fastfat.SYS
feb8c000 31ec6e6c - TDI.SYS            feaf0000 31ed0754 - nbf.sys
feacf000 31f130a7 - tcpip.sys          feab3000 31f50a65 - netbt.sys
fc550000 31601a30 - el59x.sys          fc560000 31f8f864 - afd.sys
fc718000 31ec6e7a - netbios.sys        fc858000 31ec6c9b - Parport.sys
fc870000 31ec6c9b - Parallel.SYS       fc954000 31ec6c9d - ParVdm.SYS
fc5b0000 31ec6cb1 - Serial.SYS         fea4c000 31f5003b - rdr.sys
fea3b000 31f7a1ba - mup.sys            fe9da000 32031abe - srv.sys

Address  dword dump   Build [1381]                           - Name
fec32d84 80143e00 80143e00 80144000 ffdff000 00070b02          - KSecDD.SYS
801471c8 80144000 80144000 ffdff000 c03000b0 00000001          - ntoskrnl.exe
801471dc 80122000 f0003fe0 f030eee0 e133c4b4 e133cd40          - ntoskrnl.exe
80147304 803023f0 0000023c 00000034 00000000 00000000          - ntoskrnl.exe

Restart and set the recovery options in the system control panel
or the /CRASHDEBUG system start option.
```

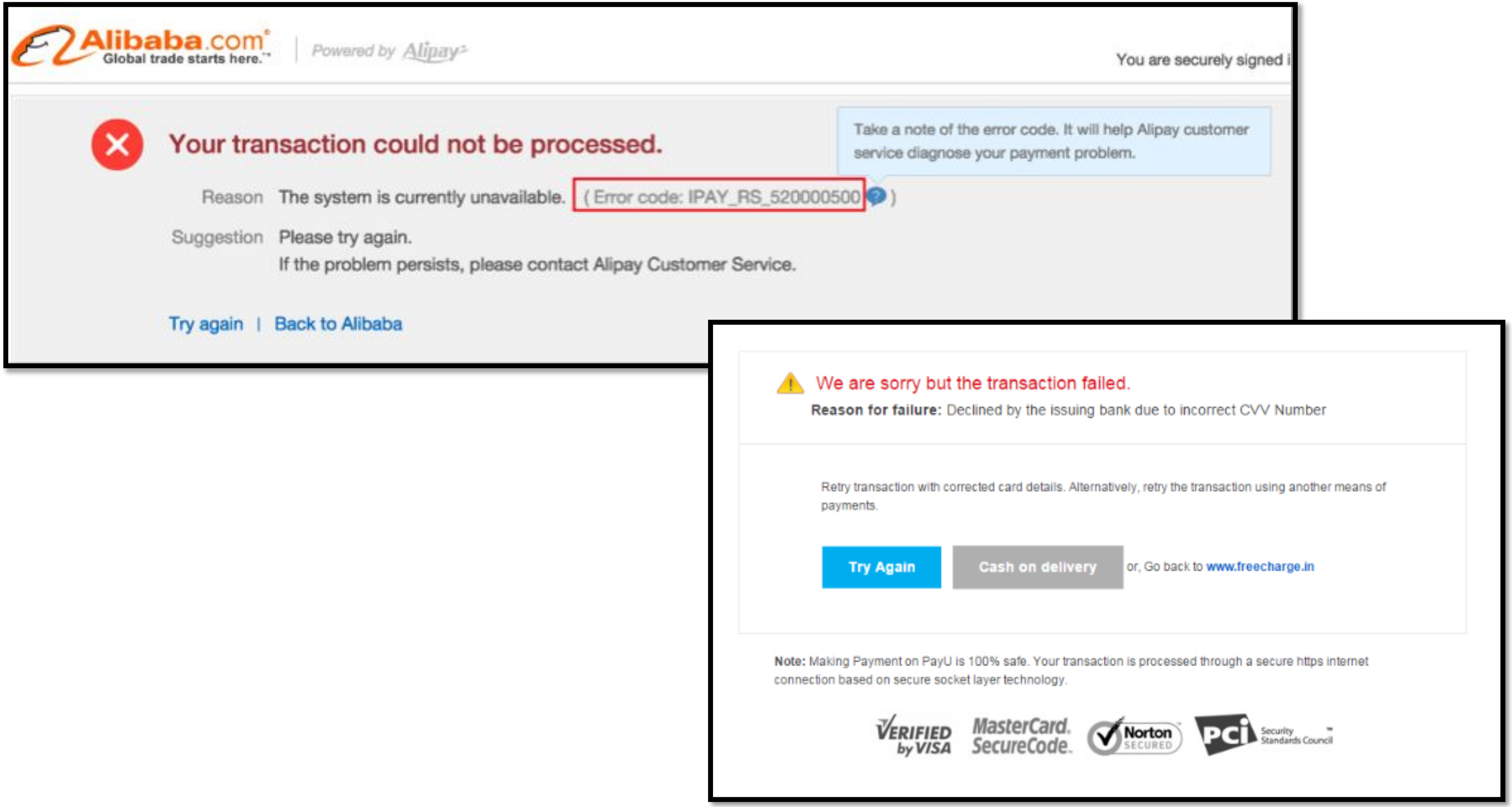
Failure Classification

❖ **Transaction failure**

- A transaction has to abort when it fails to execute or when it reaches a point from where it can't go any further.
- This is called transaction failure where only a few transactions or processes are hurt.

Reasons for a transaction failure could be −

- **Logical errors** − Where a transaction cannot complete because it has some code error or any internal error condition like bad input, data not found, integer overflow, resource limit exceeded, divide by zero, logical programming error, etc.
- **Syntax errors** − Where the database system itself terminates an active transaction because the DBMS is not able to execute it, or it has to stop because of some system condition. For example, in case of deadlock or resource unavailability, the system aborts an active transaction.



Alibaba.com
Global trade starts here.™
Powered by Alipay
You are securely signed i

Your transaction could not be processed.

Take a note of the error code. It will help Alipay customer service diagnose your payment problem.

Reason The system is currently unavailable. (Error code: IPAY_RS_520000500 ?)

Suggestion Please try again.
If the problem persists, please contact Alipay Customer Service.

Try again | Back to Alibaba

We are sorry but the transaction failed.

Reason for failure: Declined by the issuing bank due to incorrect CVV Number

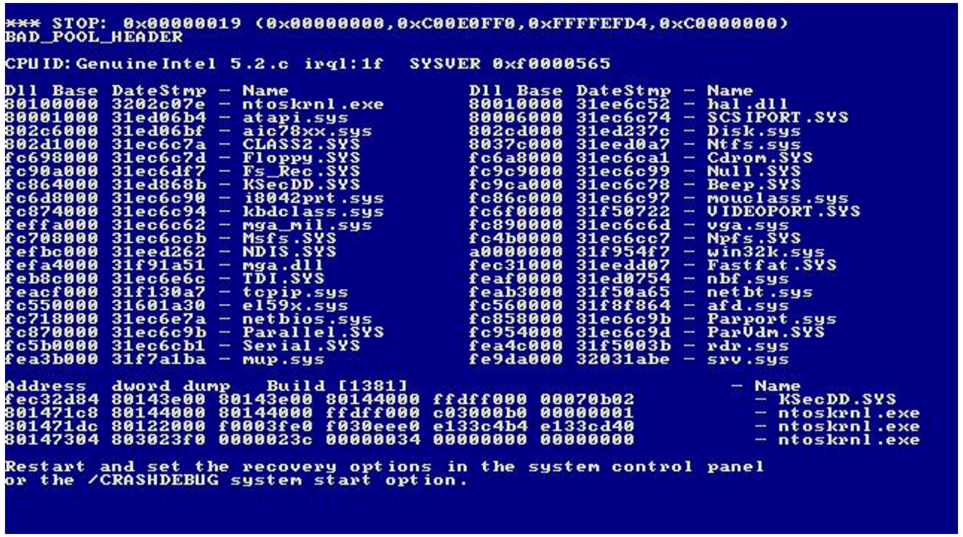
Retry transaction with corrected card details. Alternatively, retry the transaction using another means of payments.

Note: Making Payment on PayU is 100% safe. Your transaction is processed through a secure https internet connection based on secure socket layer technology.

Failure Classification

❖ System Crash

- These are problems – external to the system – that may cause the system to stop abruptly and cause the system to crash.
- For example, interruptions in power supply may cause the failure of underlying hardware or software failure.
- Examples may include operating system errors, hardware architecture.



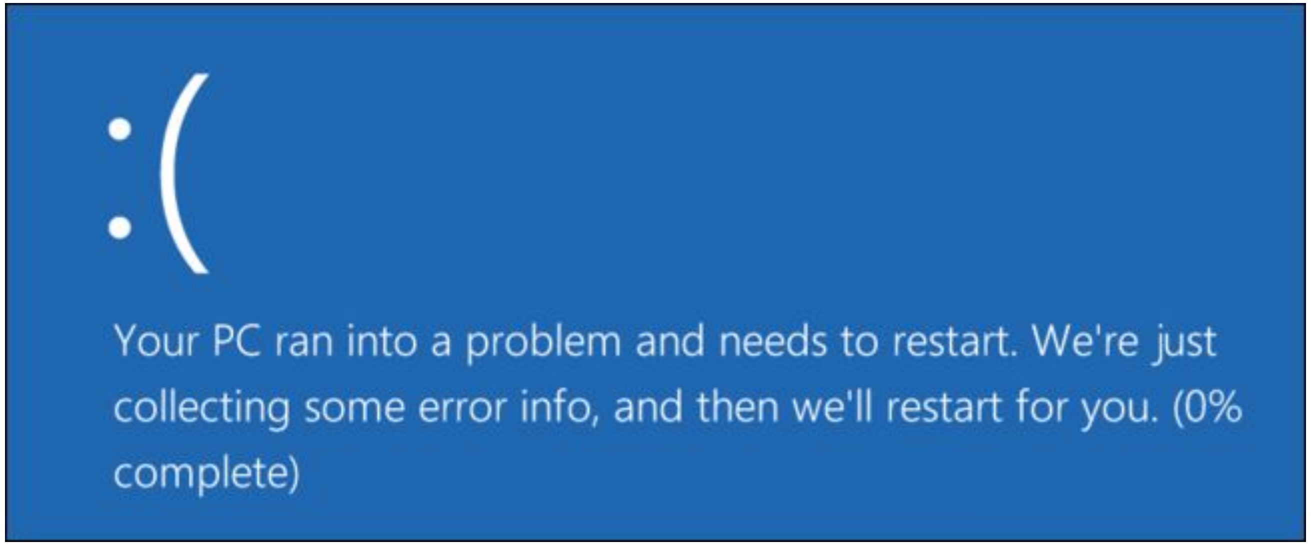
```
*** STOP: 0x00000019 (0x00000000,0xC00E0FF0,0xFFFFEFD4,0xC0000000)
BAD_POOL_HEADER

CPUID:GenuineIntel 5.2.c irql:1f   SYSVER 0xf0000565

Dll Base DateStmp - Name               Dll Base DateStmp - Name
80100000 3202c07e - ntoskrnl.exe       80010000 31ee6c52 - hal.dll
80001000 31ed06b4 - atapi.sys          80006000 31ec6c74 - SCSIPORT.SYS
802c6000 31ed06bf - aic78xx.sys        802cd000 31ed237c - Disk.sys
802d1000 31ec6c7a - CLASS2.SYS         8037c000 31eed0a7 - Ntfs.sys
fc698000 31ec6c7d - Floppy.SYS         fc6a8000 31ec6ca1 - Cdrom.SYS
fc90a000 31ec6df7 - Fs_Rec.SYS         fc9c9000 31ec6c99 - Null.SYS
fc864000 31ed868b - KSecDD.SYS         fc9ca000 31ec6c78 - Beep.SYS
fc6d8000 31ec6c90 - i8042prt.sys       fc86c000 31ec6c97 - mouclass.sys
fc874000 31ec6c94 - kbdclass.sys       fc6f0000 31f50722 - VIDEOPORT.SYS
feffa000 31ec6c62 - mga_mil.sys        fc890000 31ec6c6d - vga.sys
fc708000 31ec6ccb - Msfs.SYS           fc4b0000 31ec6cc7 - Npfs.SYS
fefbc000 31eed262 - NDIS.SYS           a0000000 31f954f7 - win32k.sys
fefa4000 31f91a51 - mga.dll            fec31000 31eedd07 - Fastfat.SYS
feb8c000 31ec6e6c - TDI.SYS            feaf0000 31ed0754 - nbf.sys
feacf000 31f130a7 - tcpip.sys          feab3000 31f50a65 - netbt.sys
fc550000 31601a30 - el59x.sys          fc560000 31f8f864 - afd.sys
fc718000 31ec6e7a - netbios.sys        fc858000 31ec6c9b - Parport.sys
fc870000 31ec6c9b - Parallel.SYS       fc954000 31ec6c9d - ParVdm.SYS
fc5b0000 31ec6cb1 - Serial.SYS         fea4c000 31f5003b - rdr.sys
fea3b000 31f7a1ba - mup.sys            fe9da000 32031abe - srv.sys

Address  dword dump   Build [1381]                            - Name
fec32d84 80143e00 80143e00 80144000 ffdff000 00070b02           - KSecDD.SYS
801471c8 80144000 80144000 ffdff000 c03000b0 00000001           - ntoskrnl.exe
801471dc 80122000 f0003fe0 f030eee0 e133c4b4 e133cd40           - ntoskrnl.exe
80147304 803023f0 0000023c 00000034 00000000 00000000           - ntoskrnl.exe

Restart and set the recovery options in the system control panel
or the /CRASHDEBUG system start option.
```



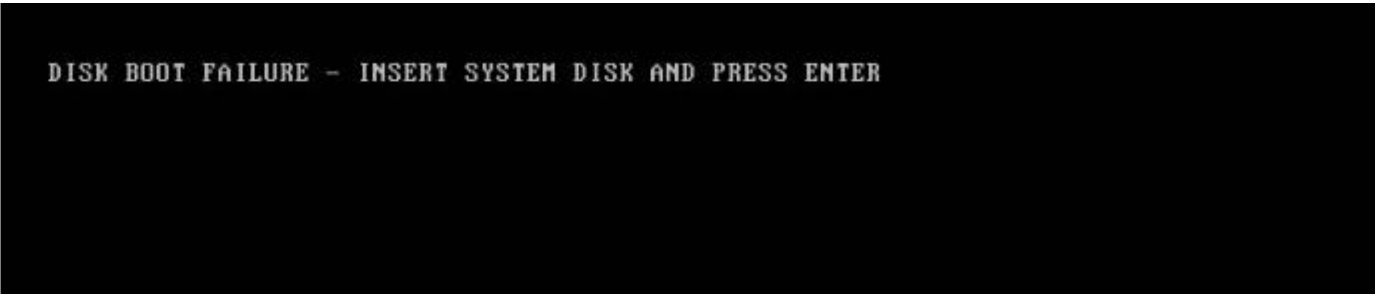
:(

Your PC ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you. (0% complete)

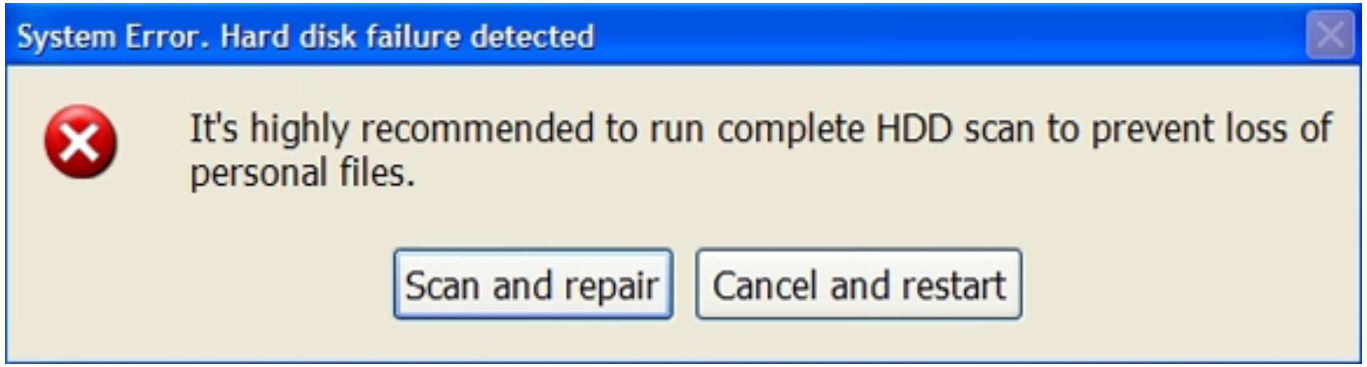
Failure Classification

❖ Disk Failure

- In early days of technology evolution, it was a common problem where hard-disk drives or storage drives used to fail frequently.
- Disk failures include formation of bad sectors, unreachability to the disk, disk head crash or any other failure, which destroys all or a part of disk storage.

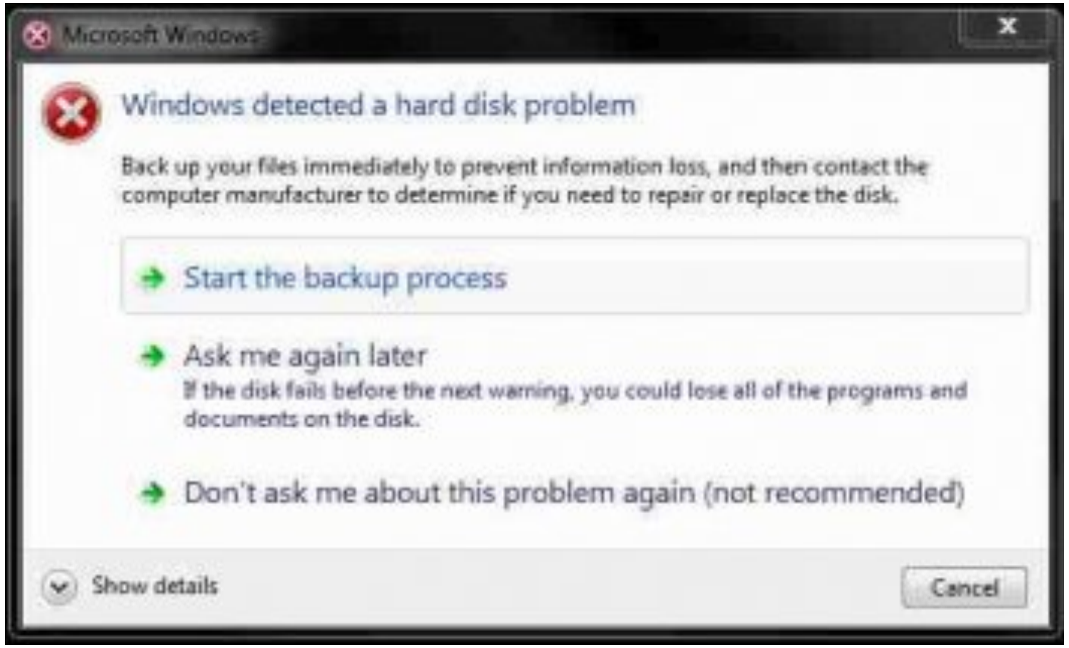


```
DISK BOOT FAILURE - INSERT SYSTEM DISK AND PRESS ENTER
```



System Error. Hard disk failure detected

It's highly recommended to run complete HDD scan to prevent loss of personal files.



Microsoft Windows

Windows detected a hard disk problem

Back up your files immediately to prevent information loss, and then contact the computer manufacturer to determine if you need to repair or replace the disk.

→ Start the backup process

→ Ask me again later

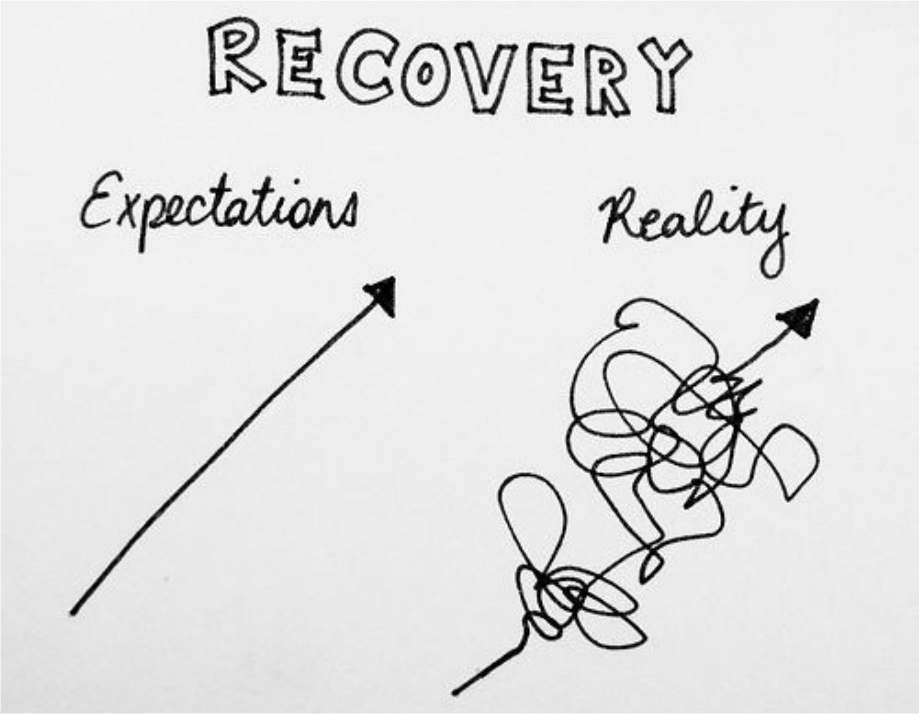
If the disk fails before the next warning, you could lose all of the programs and documents on the disk.

→ Don't ask me about this problem again (not recommended)

Failure Classification

Failures like system crash, transaction error, exceptions detected by transactions are more common than disk failure and physical problems (catastrophes).

For safety, recovery disks should be maintained at regular interval of time to restore the data in case of failure.



There are two types of techniques, which can help a DBMS in recovering as well as maintaining the atomicity of a transaction −

- Maintaining the logs of each transaction, and writing them onto some stable storage before actually modifying the database.
- Maintaining shadow paging, where the changes are done on a volatile memory, and later, the actual database is updated.

Log Based Recovery

- Log is a sequence of records, which maintains the records of actions performed by a transaction.
- It is important that the logs are written prior to the actual modification and stored on a stable storage media, which is failsafe.
- Log-based recovery works by−
  - Keeping the log file on a stable storage media.
- Log-based recovery works when−
  - a transaction enters the system and starts execution, it writes a log about it.
  - a transaction modifies an item X, it write logs.
  - a transaction finishes, it logs.

Log Based Recovery

Types of log records:

a) **Update log records:**

Describes a single db write, it has following fields;

i. Transcription identifier: unique identifier of transaction that performed the write operation.
ii. Data item identifier: unique identifier of the data item written. Typically, a location on disk of data item.
iii. Old value: value of data item prior to the write
iv. New value: value that the data item will have after the write.

b) **Special log records:**

Records significant events during transaction which includes start of transaction commit or abort transaction.

Log Based Recovery

- Before the write is done, it is essential that the log record for write be created
- After log writing is successful, we can output the modification to the db
- We have the ability to undo a modification that has already been output to db.
- We undo it by the help of log records which has old value & old state info.
- It is essential to have log records in a stable storage in order to be useful for recovery purpose.

Log Based Recovery

- The database can be modified using two approaches –

1. **Deferred database modification** – All logs are written on to the stable storage and the database is updated when a transaction commits.

2. **Immediate database modification** – Each log follows an actual database modification. That is, the database is modified immediately after every operation.

Log Based Recovery

1. **Deferred database modification**

- It ensures transaction atomicity by recording all db modification in the log but deferring the execution of all write operations of a transaction until the transaction commits
- When a transaction commits, the information on the log associated with the transaction is used in executing the deferred writes.
- If the transaction is aborted then info on log is ignored.
- Using the log, system can handle any failure & can be used for data recovery and return to previous state.
- Executing the redo operation for any transaction any number of times should result the same result as executed once.

Log Based Recovery

1. **Deferred database modification**

Let T_0 be transaction that transfers Rs. 50 from account A to B.

T_0 : read(A);
A:= A-50;
write(A);
read(B);
B:= B+50;
write(B);
	Balance: A-> 1000; B-> 2000
	<T_0 start>
	<T_0, A, 950>
	<T_0, B, 2050>
	<T_0, commit>
	Balance: A-> 950; B-> 2050
		Value of A is changed in db only after record <T_0, A, 950> has been placed in the log & final commit action is performed.

Log Based Recovery

2. Immediate database modification

- Allows db modifications to be output to db while the transaction is still in active state.
- Data modifications written by active transactions are called uncommitted modifications.
- Applies all updates directly to the db.
- If crash occurs, system uses info in log in restoring the state of system to old state.

Log Based Recovery

2. Immediate database modification

<T_0 start>	Before starting of execution, system writes to log.
<T_0 , A, 1000, 950>	During execution, write(X) operation is written to logs.
<T_0 , B, 2000, 2050>	
<T_0 commit>	When all updates are performed.

- Actual update to db is done only after log record is written.
- Using log, system can handle any failure that doesn’t result in loss of info.
- After a failure has occurred, the recovery scheme consults the log to determine which transactions need to be redone & which need to be undone.

Log Based Recovery

2. **Immediate database modification**

- ✓ Undo(T_1):
  - ✓ Restores values updated by Transaction T1 to old values.
  - ✓ Transaction needs to be undone if log contains record <T1 start> but does not contain <T1 commit>

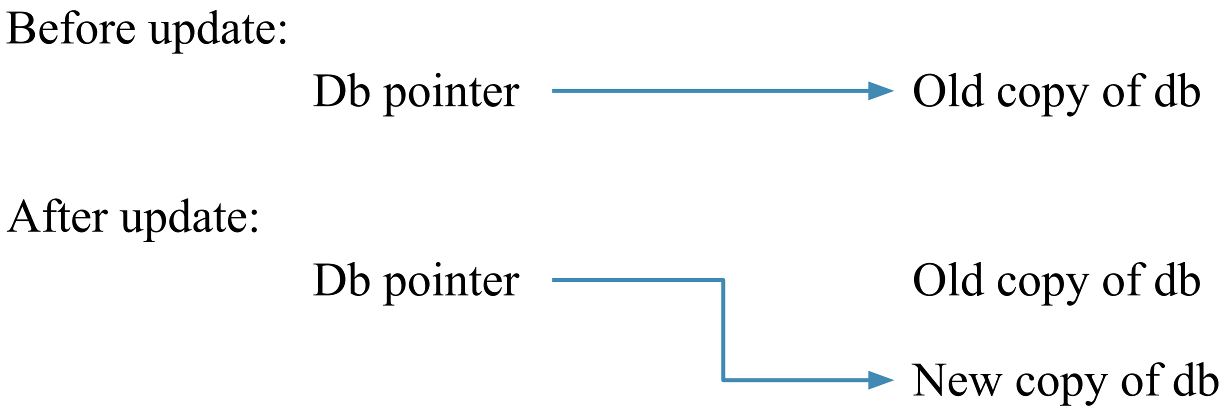
- ✓ Redo(T_1):
  - ✓ Sets values updated by Transaction T1 to new values.
  - ✓ Transaction needs to be redone if log contains both record <T1 start> and record <T1 commit>

Shadowing

- a.k.a Shadow paging
- Scheme based on making copies of db, called shadow copies
- Assumes only one transaction is active at a time
- Assumes db as simple a file on a disk
- A pointer called db pointer is maintained on disk which points to current copy of db
- Mechanism
  - Transaction that wants to update db will create a complete copy of db
  - All updates are done on new db copy, not touching the old original copy (shadow copy)
  - If transaction needs to be aborted, then simply new copy is deleted & db pointer points to old original copy which is not affected at all.

Shadowing

- OS makes sure all pages of new copy of db have been written to the disk.
- After the pages are written to disk, db system updates the db-pointer to new copy of db; the new copy then becomes the current copy of db and then old copy of db is deleted.
- Either all updates of the transaction are reflected, or none of the effects are reflected.



Before update:

Db pointer → Old copy of db

After update:

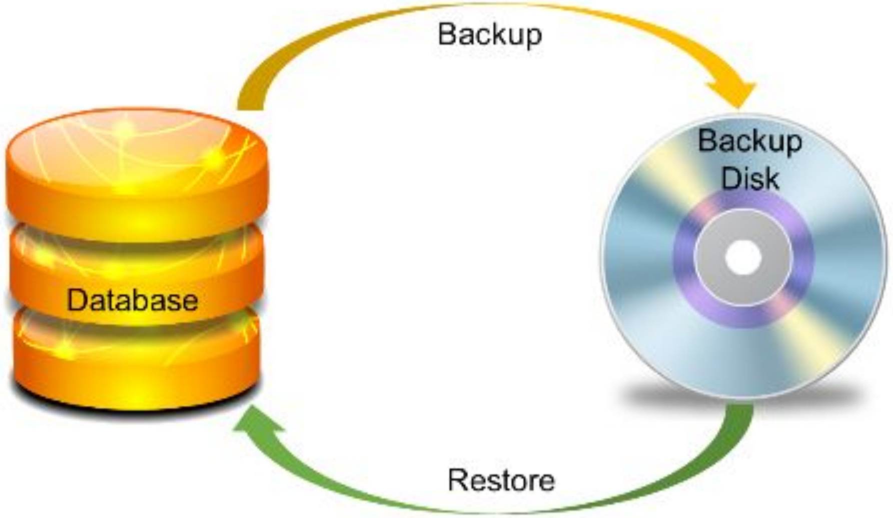
Db pointer

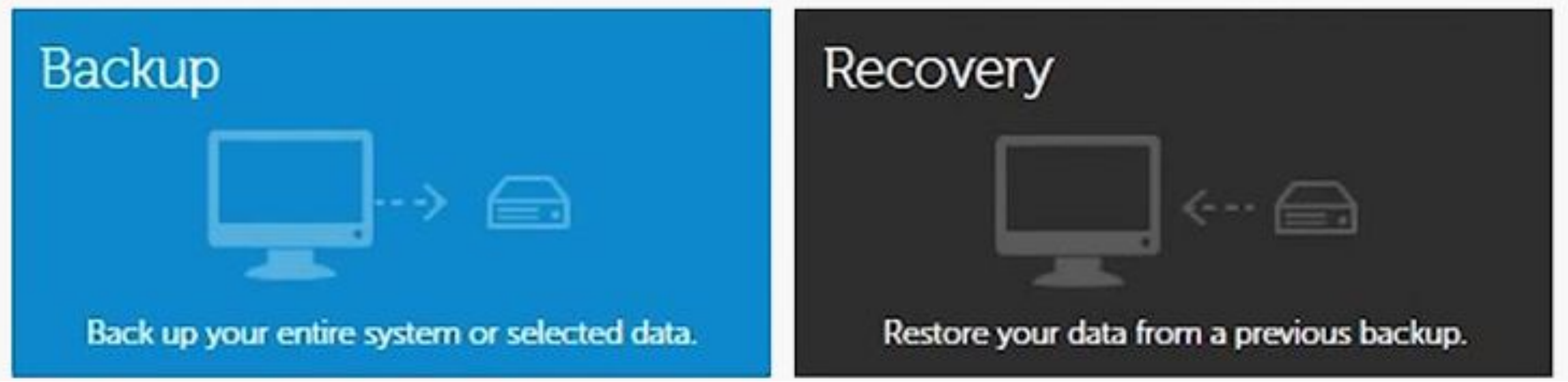
Old copy of db

New copy of db

Fig. Shadow copy technique for Atomicity & Durability.

Backup / Recovery





Backup

Back up your entire system or selected data.

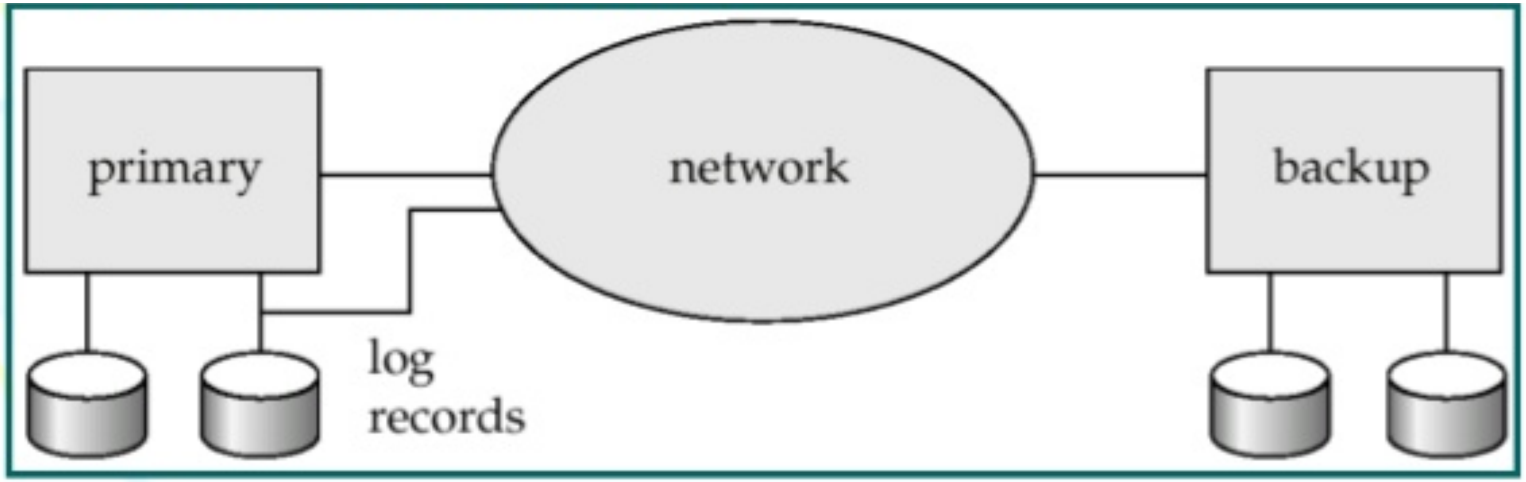
Recovery

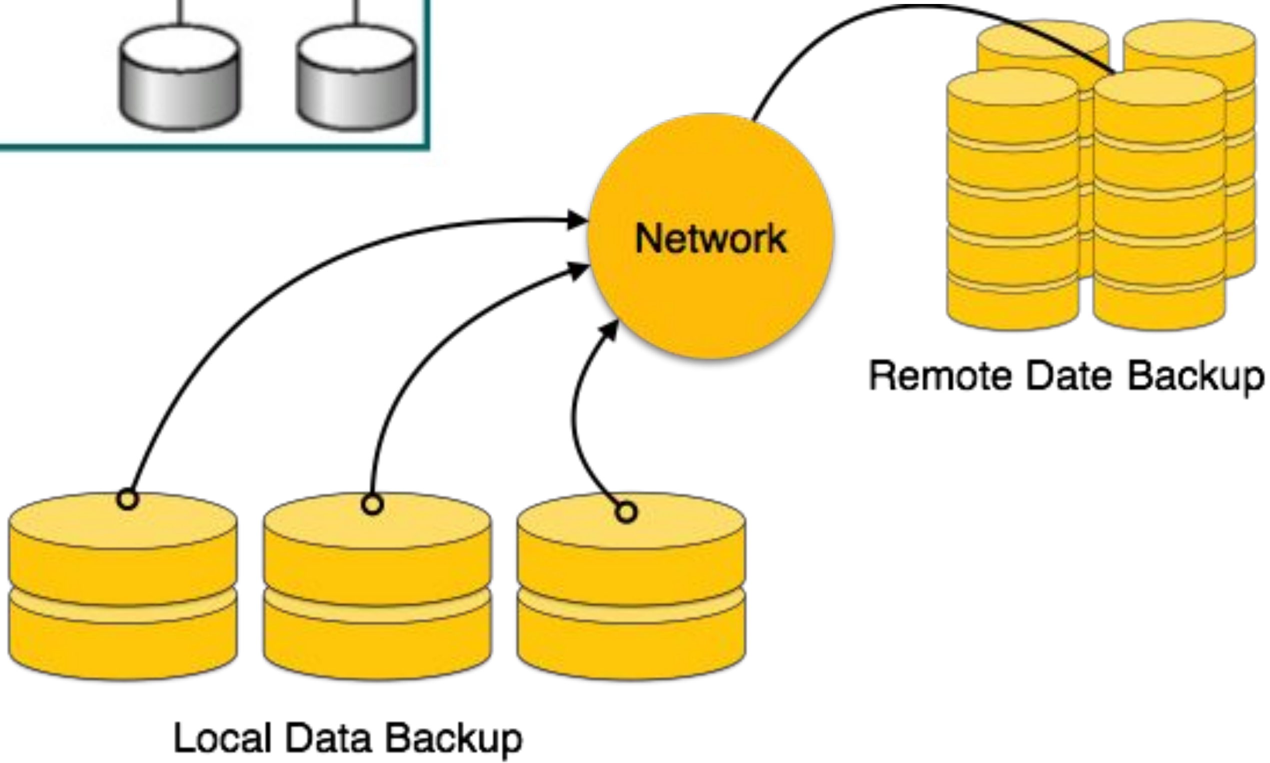
Restore your data from a previous backup.

Backup / Recovery

- A backup utility is used to create a backup copy of source db, usually by dumping the entire db onto destination tape or optical storage or other media.
- Backup copy can be used to restore the data in case of failure.
- Incremental backups are often used where only changes since the last backups are recorded; it saves space and is better than full backup.
- Backup should be done periodically either manually or automatically.
- A new log is started after each backup.
- Recovery will restore the db to some consistent state making system appear of that instant.

Remote backup system





Remote backup system

- Data from a primary site a.k.a. local site are replicated at secondary site.
- Secondary site is known as the Remote Backup System (RBS).
- RBS provides high availability by allowing transaction processing to continue even if the primary site is destroyed.
- RBS is needed in order to be safe from various security issues including environmental disaster & loss of data.
- RBS must be kept synchronized with primary site (since updates are performed at primary site)
- Whenever Primary site fails, Remote Backup System takes over the processing either automatically or manually.

Thank you.