Will you recover your latest data in case of crash or data corruption?

Pascal Ruckebusch

M81

pruckebusch@m81.eu

Agenda

- Usual situation for Database protection
- What kind of Backups
- Use FlashCopy / Snapshot
- Use Journals

- 3 case studies
  - One standalone system without HA
  - Two partitions, with HA (HW or SW)
  - Combine FlashCopy, Backups and journals

Database protection

Data is the heart of every company

- IT department is in charge to protect it.
- Data must:
  - Be protected from any outage
  - Remain accessible for users, locally and remotely
  - Be available 24/7
  - Stay uncorrupted

Database protection

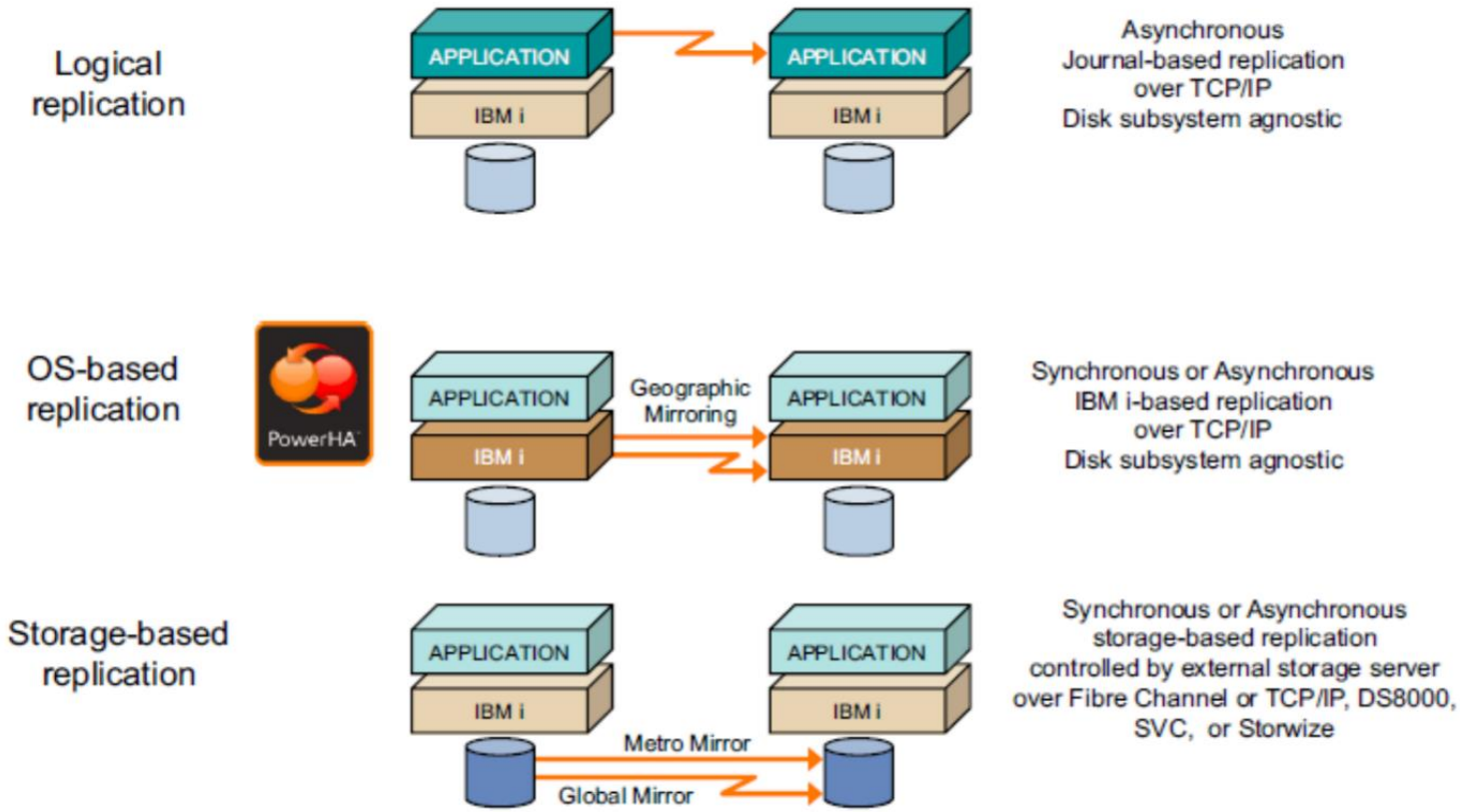
- Hardware protections
  - Electrical redundancy
  - Disk protection (RAID5, RAID6, Mirror, …)
  - Redundant and protected Network access
  - Telecommunications

Database protection

- Backups
  - To physical tapes (LTO)
  - To virtual tapes (VTL)
    - Local
    - With remote replication
  - To network
  - Standard commands, or with BRMS or another product
    - Every product finally uses standard commands SAVSYS, SAVLIB, …
    - Except some network backup products
- Must be done as often as possible
- But time consuming

Database protection

- Remote copy



Database protection

- Journaling
  - Log every change made to the DataBase
  - “AFTER only” image
    - Standard option
    - Allows to know what has changed, and to reapply changes
  - “BEFORE and AFTER” image
    - Needs more disk space
    - Allows also rollback in most cases (but not all)
  - Often configured in modern applications
  - Very easy to setup, without to change the programs
  - Not very easy to use

Database protection

- Security to access the Database is also very important
  - Prevent unauthorized user to connect
  - More and more access protocols necessary
    - 5250
    - FTP
    - http/https
    - ODBC/JDBC
    - …
  - Managed by Firewall outside the partition
  - Several products allow to monitor security and access inside the partition

So, What is the problem ?

- IBM i is a VERY **stable** system
- It works “by itself”, with close to no maintenance
- Data Base is fully integrated to the system
- Reputation to have “No viruses”
  - At least not known at this time
- HA solutions (HW or SW) prevent from system outage

Result:

A very well protected database

What else ?

So, What is the problem ?

Data deletion
Loss of integrity
Corruption
Bug …

So, What is the problem ?

- A new version of a program (Interactive or Batch)
- Tested and validated on a test partition with a reduced set of data

- On production partition,
  - A bug is detected, but only after a few hours
  - A lot of data in one or more tables have been corrupted

➔ Use the last backup, and come back 24h ago

So, What is the problem ?

- An environment with only ONE partition
- Production, development and tests are done on the same partition
- A programmer wants to test a new program, and forgets to change the Library List
  - Test program runs on Production table
  - Potentially many records corrupted

➔ Use the last backup, and come back 24h ago

So, What is the problem ?

- A user is about to be fired, and is angry about this
- He finds a way to create another user profile with enough rights
- On evening, CLRPFM for 2 or 3 important tables
  - Every internet order of the day are lost
  - Operation done by an unknown user, hard to find who made this operation

➔ Use the last backup, and come back 24h ago

So, What is the problem ?

Data deletion, loss of integrity, corruption, bug…

- Generally heavy impacts
  - A lot of working hours under heavy stress attempting to recover an 'acceptable' situation.
  - Data lost and/or consistency lost
  - Need to restart from the last backup
    - Up to 24 hours of data lost

- **Loss of time, money, image … and credibility for IT**

So, What is the problem ?

Hight Availability doesn’t help in these cases

- Made to protect against a system outage
- The corrupted data are instantly transferred to the second partition
- Same damage on both partition

What are the solutions

What are the solutions

- Backups
  - Traditional
  - Save while Active
- FlashCopy / Snapshot
- Journals
  - Roll back
  - Reapply over restored tables

Backups

Standard Backups

- Run backups at least once per day
- Standard backups are very stable
- Would be interesting to run several backups per day, but not possible
  - Save time is generally very long
  - Tables are locked during backup

- In case of corruption
  - Up to 24 hours of data lost

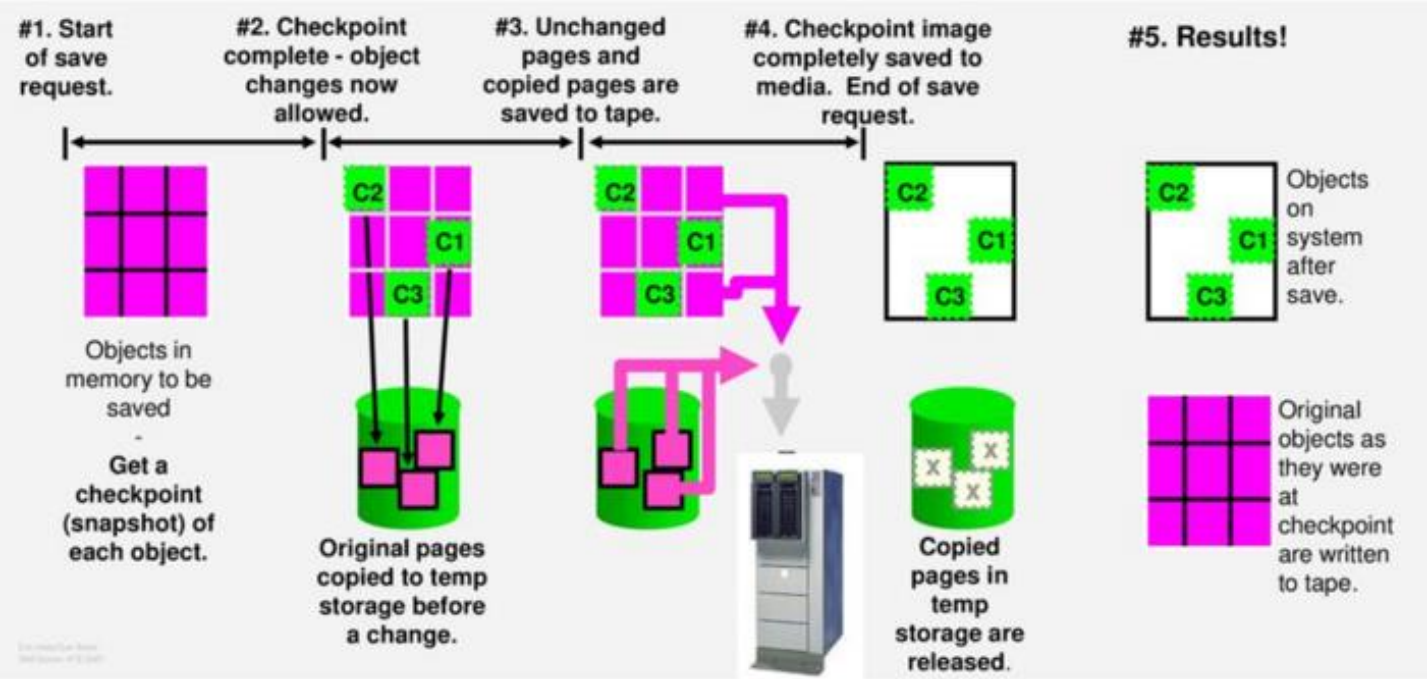
Backups with Save While Active

- Standard IBM I Backup commands with “Save While Active” parameter

```
Save active  . . . . . SAVACT        > *SYNCLIB
Save active wait time:               SAVACTWAIT
  Object locks . . . . . .           > 120
  Pending record changes . . . .     > *LOCKWAIT
  Other pending changes  . . . .     > *LOCKWAIT
Save active message queue  . . . SAVACTMSGQ   > QSYSOPR
  Library  . . . . . .                          *LIBL
```

- Allows to save the tables and keep users working after phase 1 (checkpoint)
- But still need to stop the users for a short break at the beginning of the process
- Objects that are locked by users during the phase 1 of the SWA process may not be saved

Backups with Save While Active



- No activity on Database during phase 1
- The data on tape is an image of the database at the end of phase 1.

Backups with Save While Active

```
/* End users jobs  */
CALL        PGM(END_USERS)

CLRMSGQ     MSGQ(SAVE_ACT)
MONMSG      MSGID(CPF0000)


STRSAVSYNC SYNCID(TOGETHER) NUMSYNC(2)

SBMJOB      CMD(SAVLIB LIB(*ALLUSR) DEV(TAP01) +
              SAVACT(*SYNCLIB) SAVACTMSGQ(SAVE_ACT) +
              SYNCID(TOGETHER)) JOB(SAVLIB) JOBQ(QBATCH1)

SBMJOB      CMD(SAV DEV('/qsys.lib/tap02.devd') +
              OBJ(('/*')) SAVACT(*SYNC) +
              SAVACTMSGQ('/qsys.lib/qgpl.lib/SAVE_ACT.msgq') +
              SYNCID(TOGETHER)) JOB(SAVIFS) JOBQ(QBATCH2)
```

Backups with Save While Active

```
/* We read the first message     */
 RCVMSG      MSGQ(SAVE_ACT) MSGTYPE(*LAST) WAIT(60) +
               RMV(*NO) KEYVAR(&MSGKEY) MSGID(&MSGID)

/* Initialize the counter of checkpoint messages received */
 CHGVAR      VAR(&COUNT) VALUE(0)

/* Initialize the time limit   */
 CHGVAR      VAR(&MAXDELAY) VALUE(600)

 RTVSYSVAL   SYSVAL(QHOUR) RTNVAR(&QHOUR)
 RTVSYSVAL   SYSVAL(QMINUTE) RTNVAR(&QMINUTE)

 CHGVAR      VAR(&QHOURN) VALUE(&QHOUR)
 CHGVAR      VAR(&QMINUTEN) VALUE(&QMINUTE)
 CHGVAR      VAR(&START) VALUE((&QHOURN * 3600) + (&QMINUTEN * 60))
```

Backups with Save While Active

```
WAIT_CHECK: /* We read next message until checkpoint  */
            RCVMSG     MSGQ(SAVE_ACT) MSGTYPE(*NEXT) +
                         MSGKEY(&MSGKEY) WAIT(60) RMV(*NO) +
                         KEYVAR(&NEWMSGKEY) MSGID(&MSGID)

            CHGVAR     VAR(&MSGKEY) VALUE(&NEWMSGKEY)

            IF         COND(&MSGID *EQ 'CPI3712') THEN(DO)
              CHGVAR     VAR(&COUNT) VALUE(&COUNT +1)
            ENDDO

           /* We check that it's not too long        */
           /*  otherwise, we start the users         */
            RTVSYSVAL  SYSVAL(QHOUR) RTNVAR(&QHOUR)
            RTVSYSVAL  SYSVAL(QMINUTE) RTNVAR(&QMINUTE)

            CHGVAR     VAR(&QHOURN) VALUE(&QHOUR)
            CHGVAR     VAR(&QMINUTEN) VALUE(&QMINUTE)
            CHGVAR     VAR(&END) VALUE((&QHOURN * 3600) + (&QMINUTEN * 60))
```

Backups with Save While Active

```
            IF          COND((&END - &START) > &MAXDELAY) THEN(DO)
              /* It was too long. We start the users    */
              GOTO        CMDLBL(STR_USERS)
            ENDDO

           /* We wait for 2 checkpoint messages    */
            IF          COND(&COUNT < 2) THEN(DO)
              GOTO        CMDLBL(WAIT_CHECK)
            ENDDO


           /* Checkpoint has been reached             */
           /* Users can work again                    */
STR_USERS:  CALL        PGM(STR_USERS)
```

Incremental Backups

- Incremental backup is the same as Standard backup, except that SOME objects are not saved.

- Backups only changed objects since last full backup
- In a big history file, if only one record has changed, the entire file is backed-up.

➔ Definitely not a good option to protect against data corruption

Backups

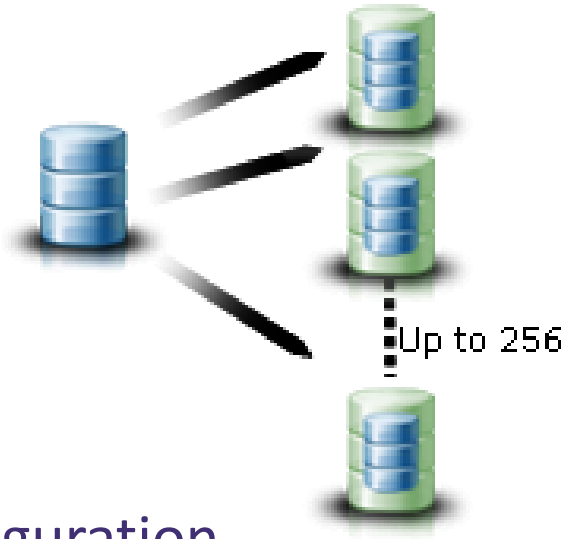
- Conclusion
  - Possible to reduce downtime for backup by using “Save While Active”
  - But not really possible to run several backups per day
  - Still a risk of objects locked
  - Additional risk if there is no tape available

- In case of corruption
  - Up to 24 hours of data lost

FlashCopy / Snapshot

FlashCopy / Snapshot

- External storage provides Snapshot functions (called FlashCopy in IBM world)
- Creates an INSTANT copy of the source disks
- Possible to create one image of production disks every hour
  - No performance issue
  - No impact on users
  - Can be automated with some configuration and ssh instructions sent to the storage system
  - Needs about 20% of additional disk space for 24 hours (data are not fully copied using Thin provisioning)



FlashCopy / Snapshot

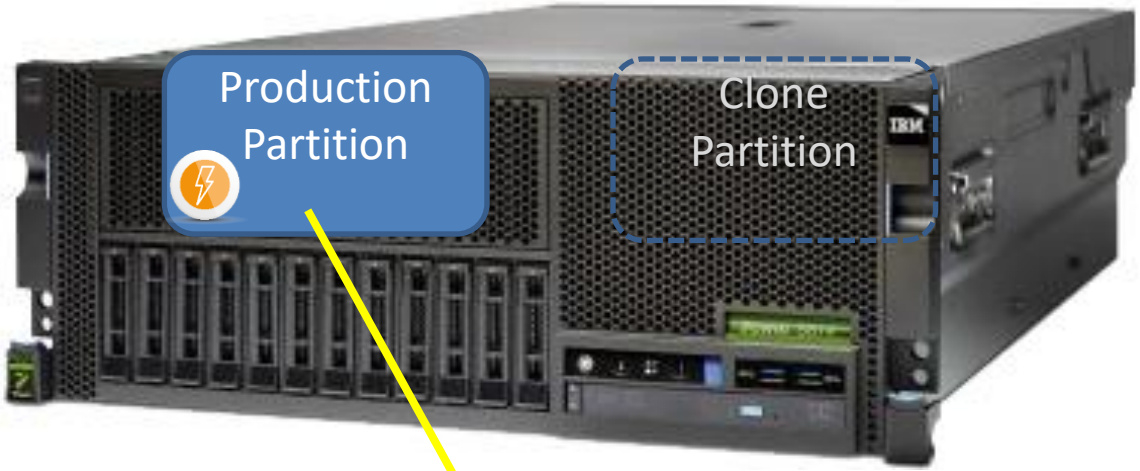
- In case of corruption of the Database
  - Determine the set of disks with correct data
  - Show this set of disk to a new partition
  - Start this partition in manual mode (will be an abnormal IPL)
  - Save the table(s) that have been corrupted, and restore them to the production partition

FlashCopy / Snapshot for backups

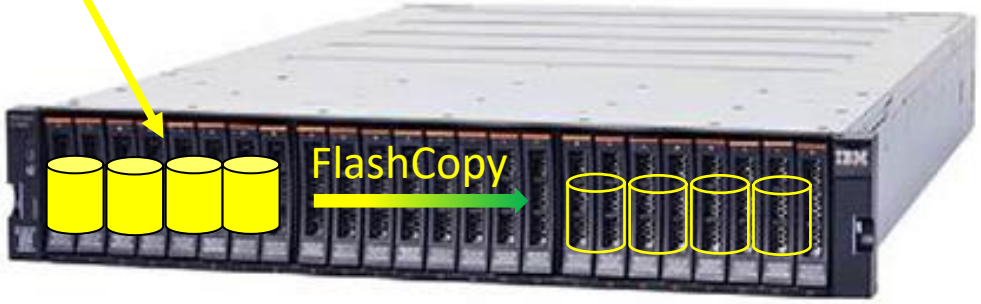
- Take advantage of the FlashCopy/Snapshot to run the backup on a Clone partition
  - Stop users for 1 or 2 minutes, or even don't stop them
  - Possible to run a SAVE21 every day

- Associate both technics
  - Start 1 snapshot for backup
  - And start 23 snapshots for database recovery

How does it work ?



IBM Power Systems



External Stockage

IBM

DELL EMC

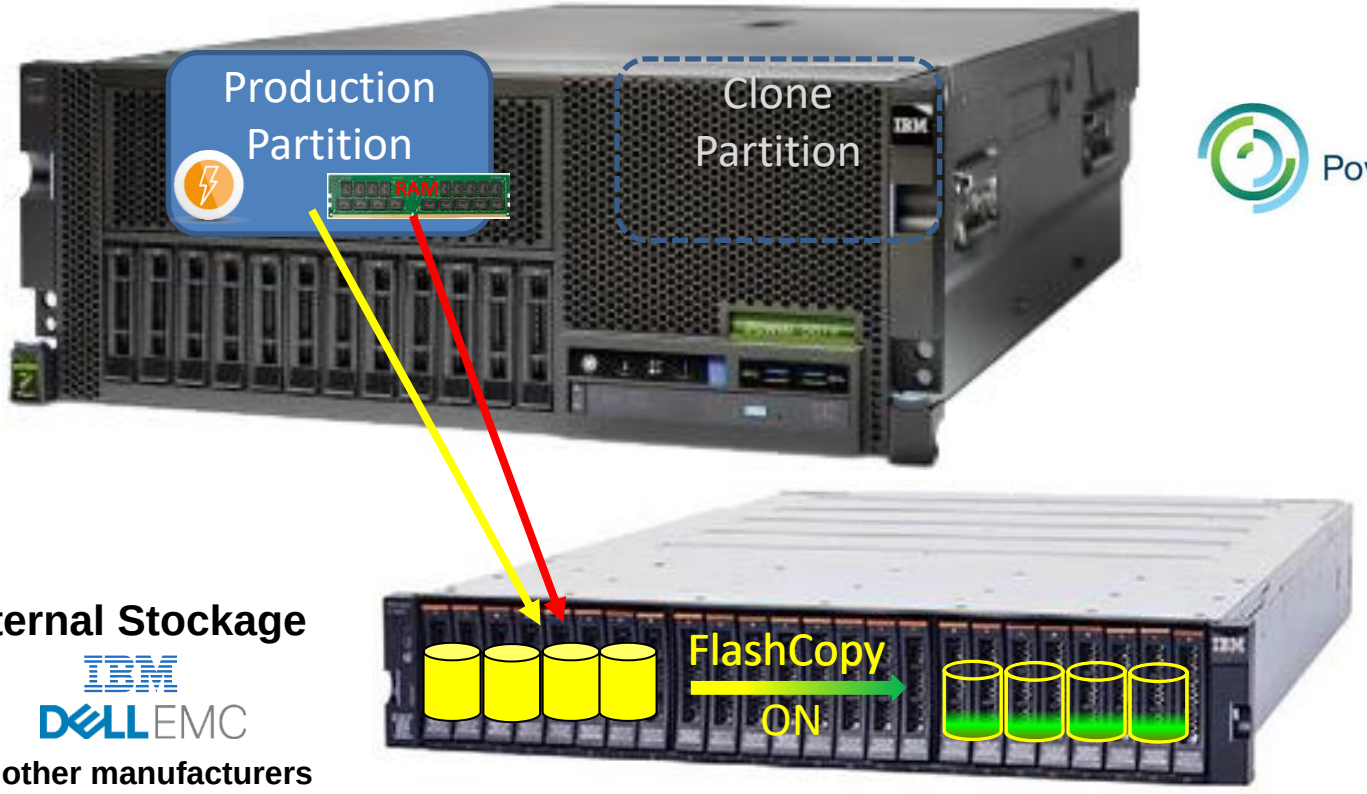
And other manufacturers

How does it work ?

```
                    Partition cloning (FLCLONE)

Indiquez vos choix, puis appuyez sur ENTREE.

Environment  . . . . . .          B_DAY, B_WEEK, SAVE21, ...
```



IBM Power Systems

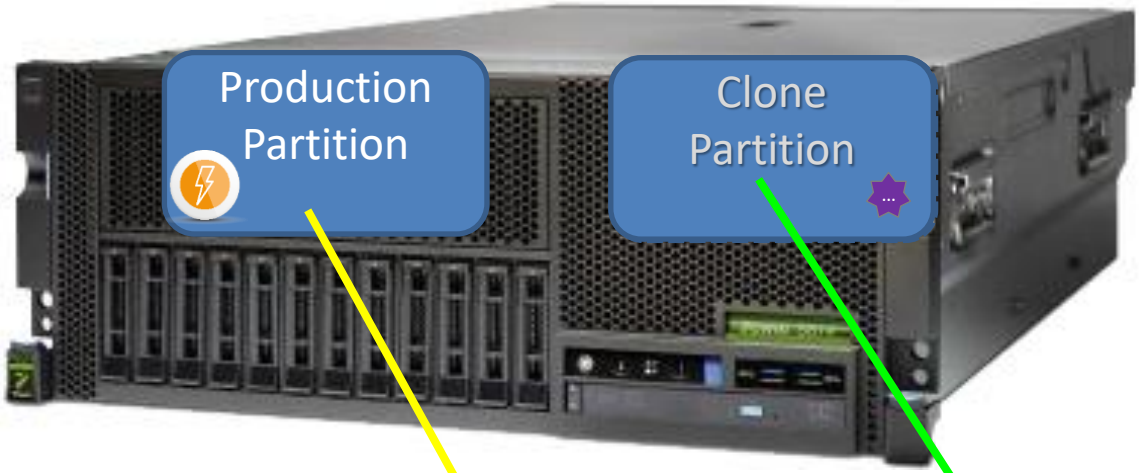
External Stockage

IBM
DELL EMC

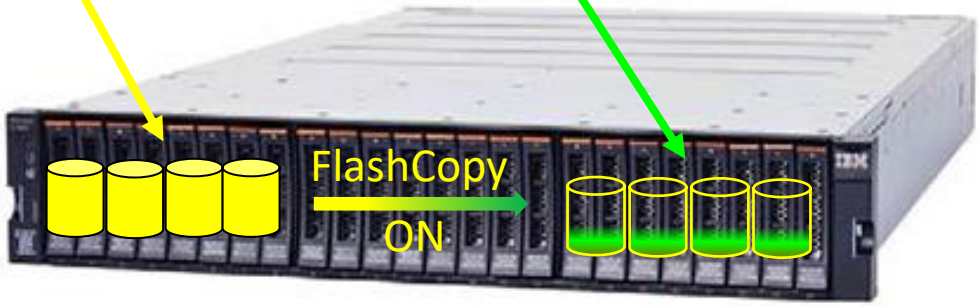
And other manufacturers

How does it work ?

TCP/IP
Scheduled jobs
Tape Drives
resources
Users subsystems
…



IBM Power Systems



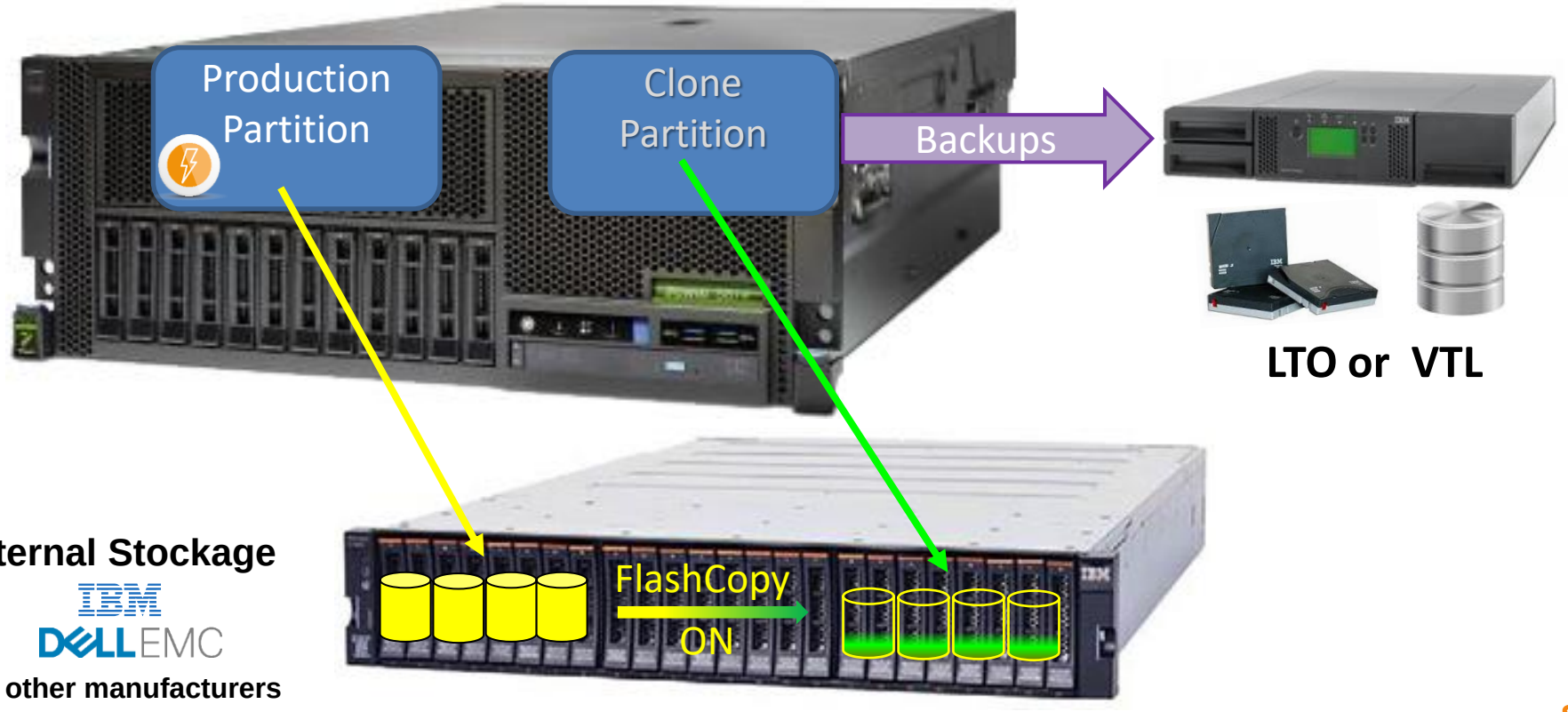
External Stockage

IBM
DELL EMC

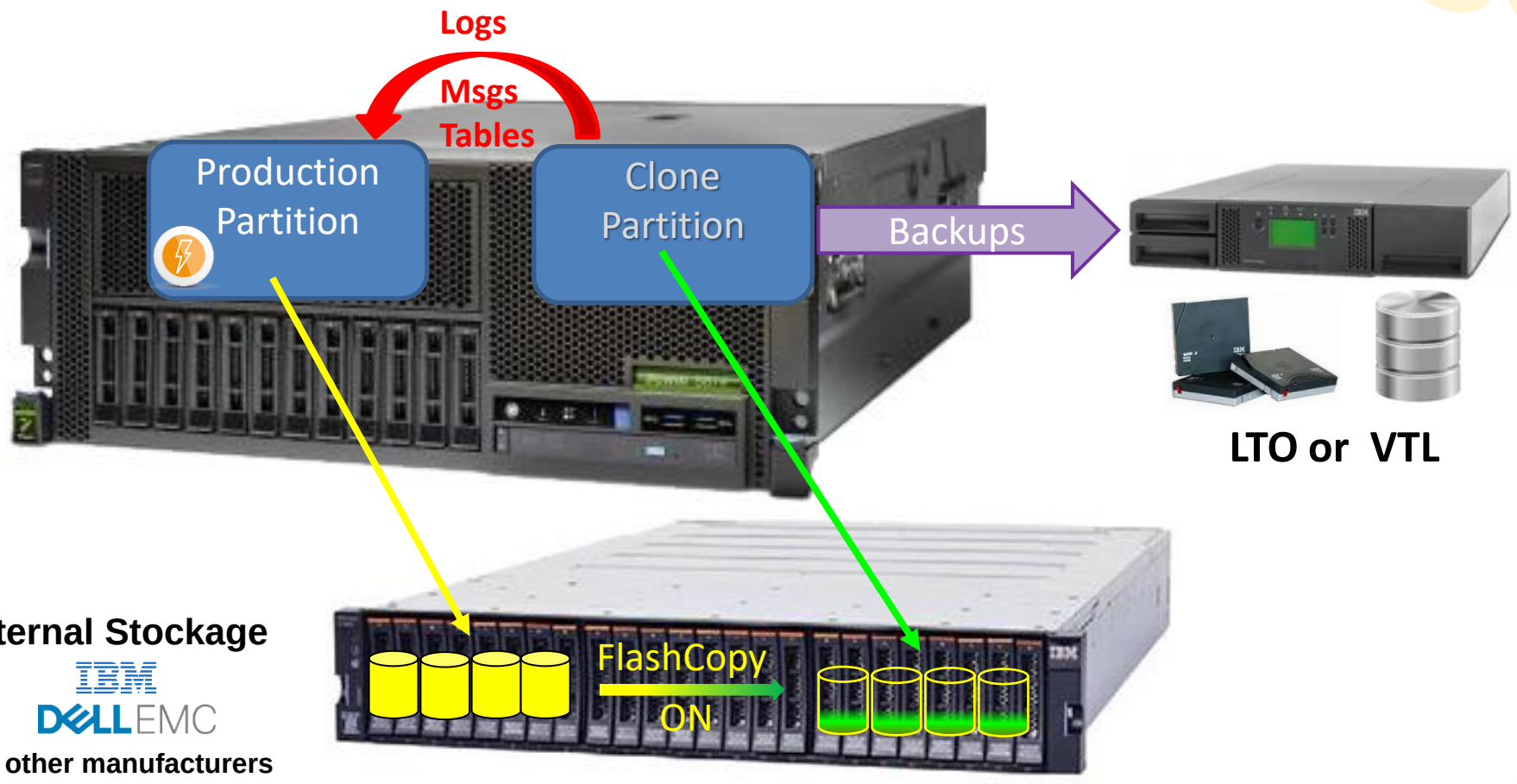
And other manufacturers

How does it work ?

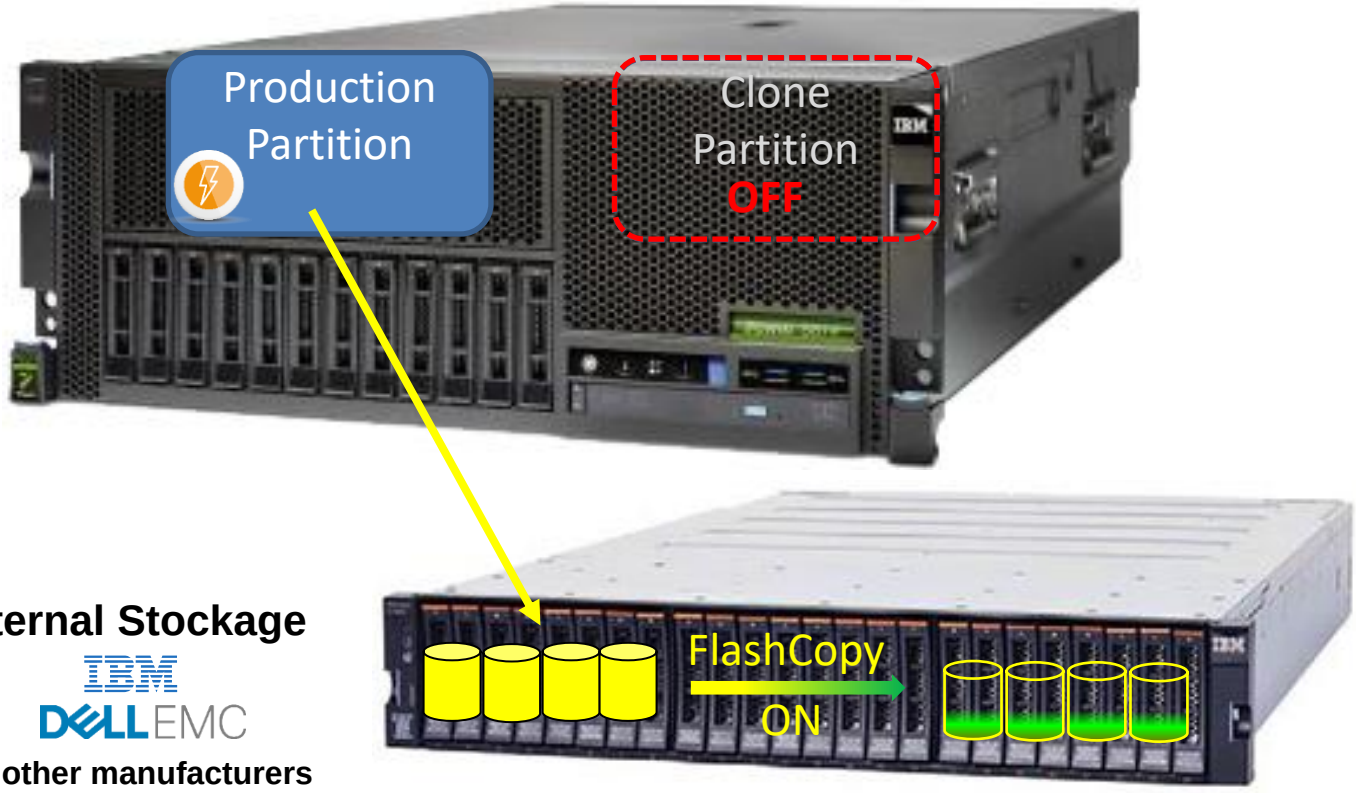
With BRMS
Without BRMS
SAVE21
…



How does it work ?



How does it work ?



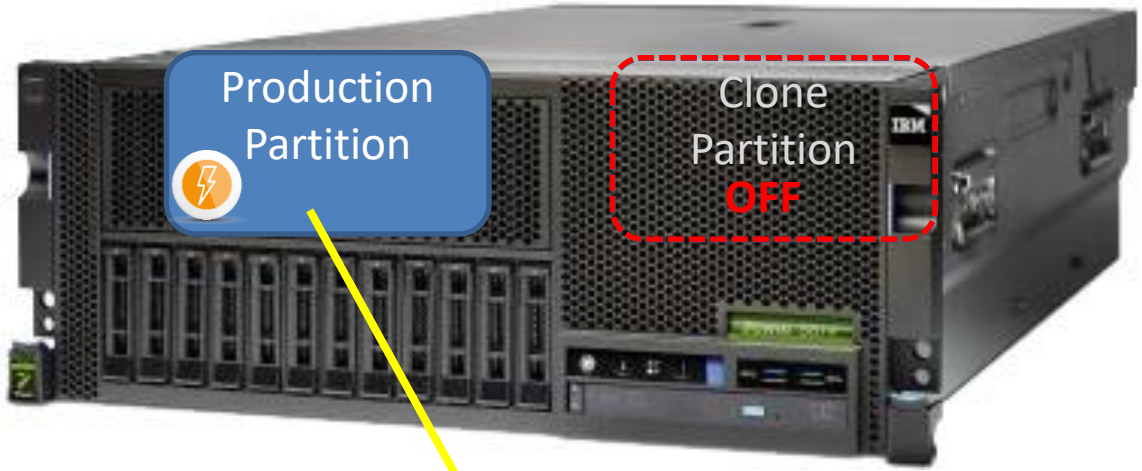
External Stockage

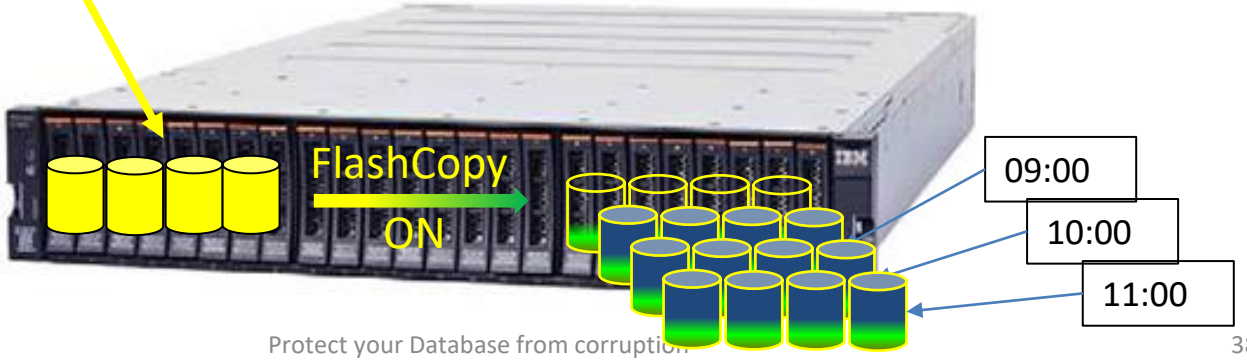
IBM

DELL EMC

And other manufacturers

How does it work ?





External Stockage

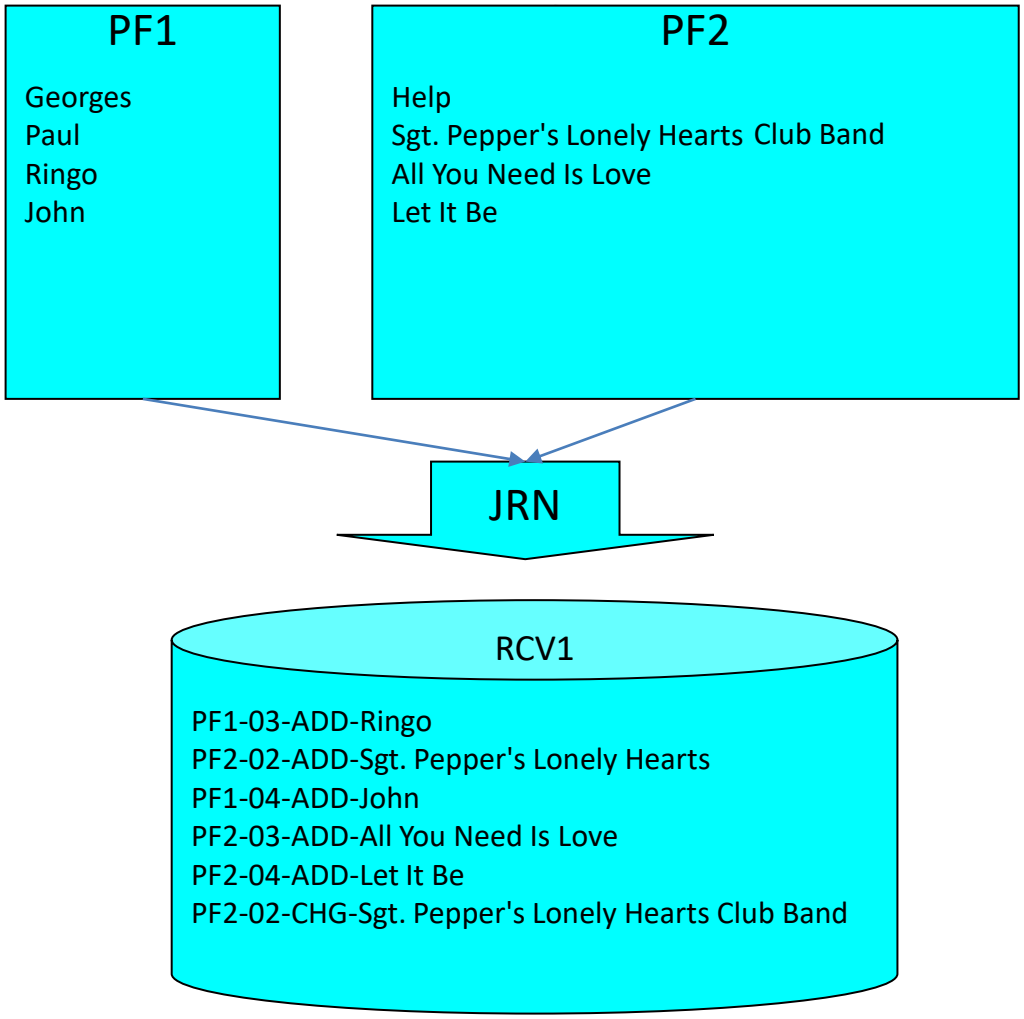
IBM

DELL EMC

And other manufacturers

Journals

What is a Journal ?



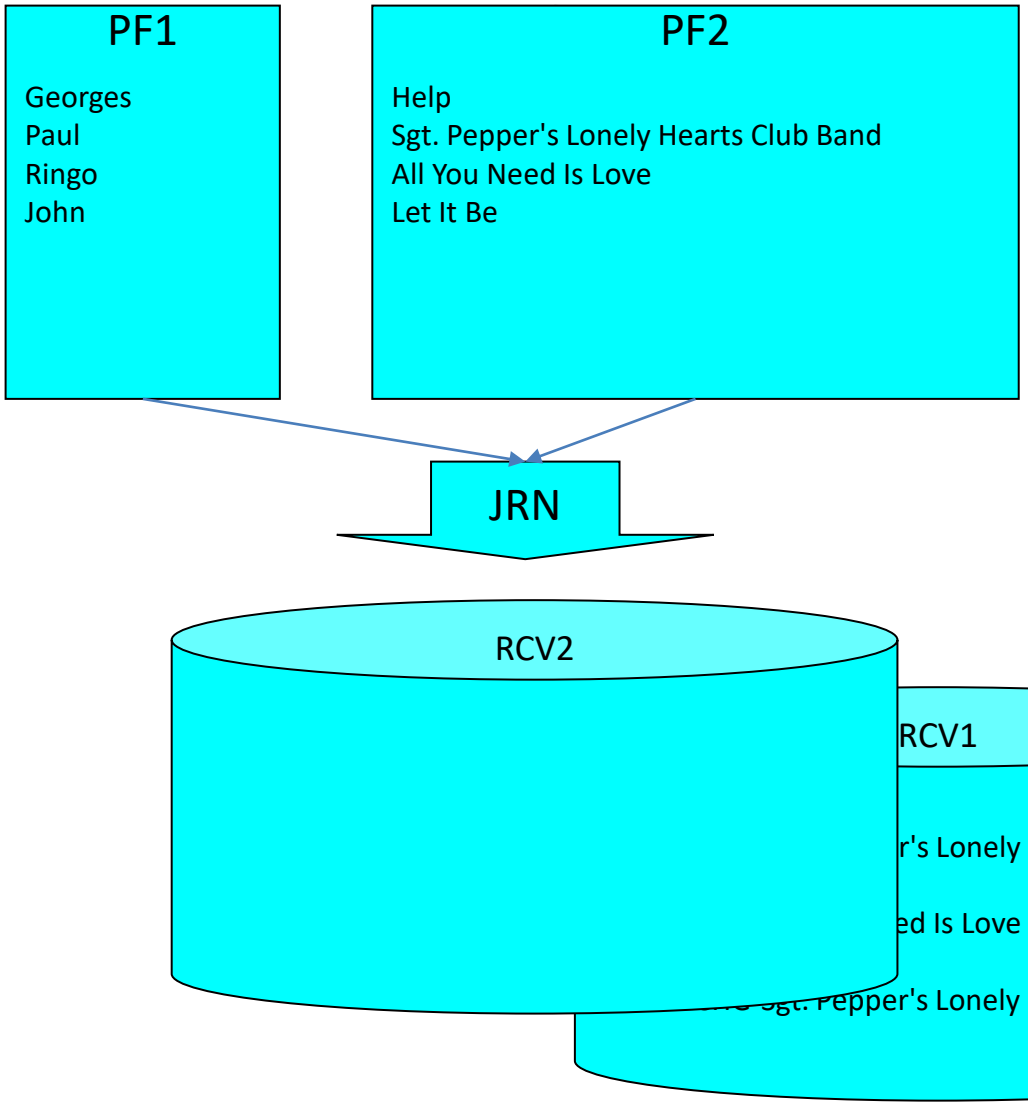
CRTJRNRCV JRNRCV(RCV1)

CRTJRN JRN(JRN) JRNRCV(RCV1)

STRJRNPF FILE(PF1 PF2) JRN(JRN)

CHGJRN JRN(JRN) JRNRCV(*GEN)

What is a Journal ?




```
CRTJRNRCV JRNRCV(RCV1)

CRTJRN JRN(JRN) JRNRCV(RCV1)

STRJRNPF FILE(PF1 PF2) JRN(JRN)

CHGJRN JRN(JRN) JRNRCV(*GEN)
```

What is a Journal ?

- Fully integrated with the system, and DB2 for I
- Start journaling with a command depending of the type of object

STRJRNPF	Start Journal Physical File	Files and Tables
STRJRNOBJ	Start Journal Object	*DTAARA and *DTAQ
STRJRNAP	Start Journal Access Path	Access path
STRJRNLIB	Start Journal Library	All objects *FILE,*DTAARA and *DTAQ of a library
STRJRN	Start Journal	IFS

- No need to change the programs
  - Unless use of Commitment control

What is a Journal ?

```
  Start Journal Physical File (STRJRNPF)
Record images  . . . . . .  *AFTER
```

- Image AFTER only
  - Only the new value is recorded in the journal
  - Allows to
    - Have a log of every change in the database

      DSPJRN
    - Re-apply changes after restoring a previous version of the database

      APYJRNCHG

What is a Journal ?

- Image BEFORE and AFTER
  - Old and new values are recorded in the journal
  - Allows to
    - Roll back the changes made in the database

      RMVJRNCHG
    - In some conditions. For example: with CLRPFM, old value is not recorded
    - Use Commitment control in the programs

How to use Journals

- In case of corruption on the Database
  - Restore the files to recover (or the entire database)
  - Analyze the journal and find the exact position where you want to return (DSPJRN)
    - This is the hardest part of the process
  -
```
APYJRNCHG  JRN(JRN)
           FILE((MYLIB/*ALL))
           RCVRNG(RCV1 RCV6)
           FROMENTLRG(*LASTSAVE)
           TOENTLRG(125885)
           OBJERROPT(*END)
```
- Must be done in the original library
- Lot of stress to do it on Production partition

How to use Journals

- Keep journal receivers long enough
  - Delete them after more than 24 h
  - Detach and save them frequently
  - Save to tape, or save to SAVF and export them
- Don’t forget the other servers
  - IBM I is often connected to other servers (EDI, …)
  - Database of all servers are synchronized
  - Look for the same process in other databases

Case 1

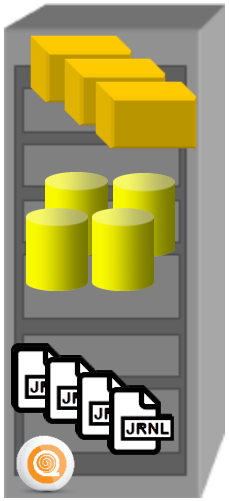
One standalone system
Without HA

Case 1: Standalone system without HA

- Description of the environment
  - One system, with only one partition
  - No HA solution because
    - It's too expensive to maintain 2 systems
    - Contract to get a new server if a crash occurs
    - 24 h without IT is acceptable (for this application)
- In case of crash
  - Changes in database since last backup are lost

Case 1: Standalone system without HA

Server
or IBM i partition
Production



1/ Regular system backups

2/ Daily backups

3/Every N minutes, detach Journal receivers and export them

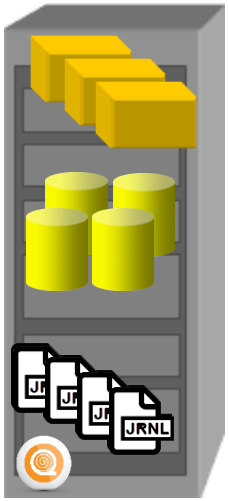
4/ Storage on Server FTP or NFS

Possible replication

Remote site

Case 1: Standalone system without HA

Server
or IBM i partition
Production



The cartridges
(physical or virtual)
are stored securely
off-site

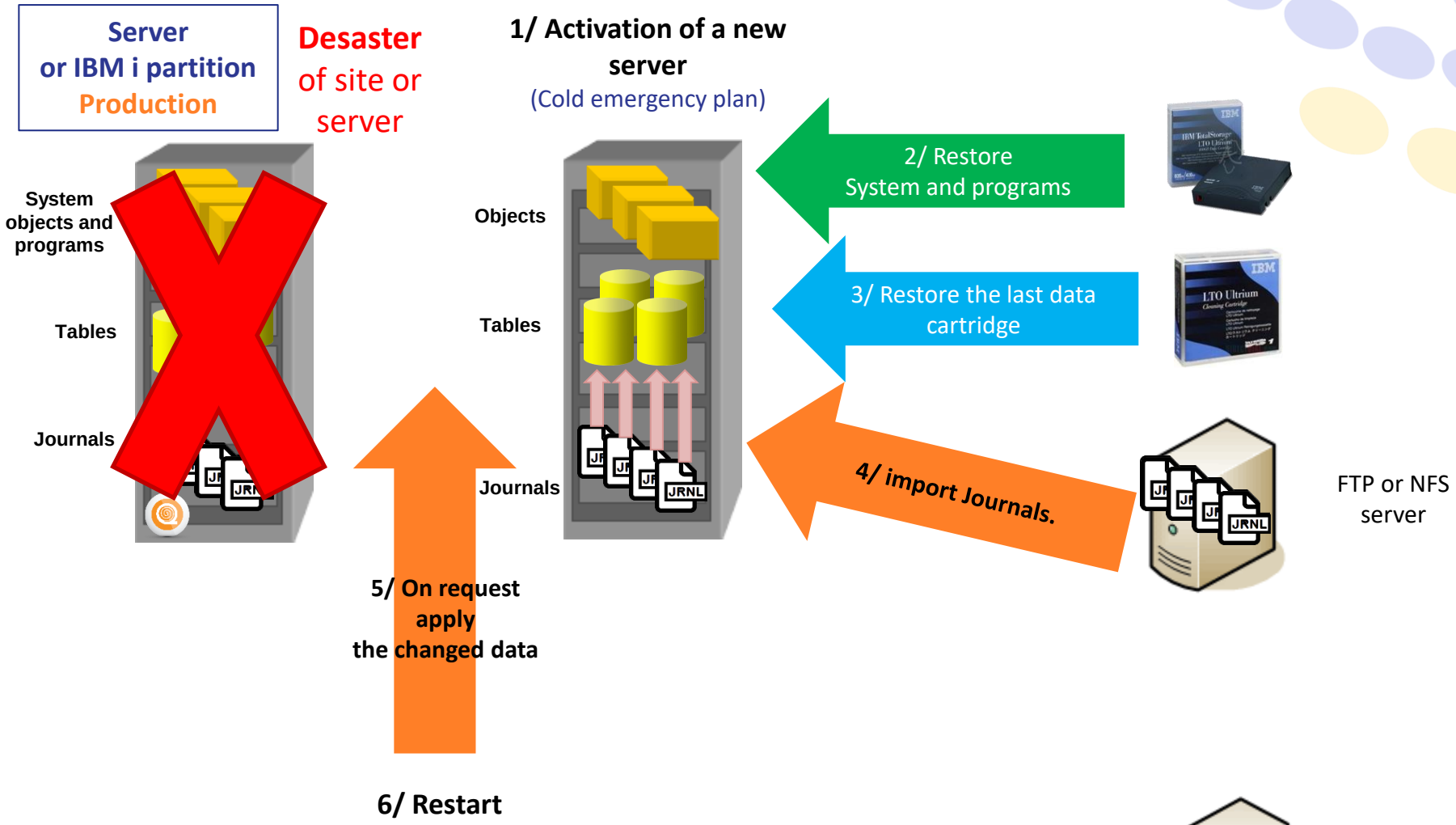
Journals are stored securely
offsite

Server
FTP or NFS

Possible replication

Remote site

Case 1: Standalone system without HA



Only a few minutes of data loss!

Case 2

Two partitions,
With HA (HW or SW)

Case 2: 2 partitions with HA

- Description of the environment
  - One production partition replicated to a second partition
  - Software or Hardware replication
  - Data are synchronously replicated on second partition
- In case of system crash
  - Activity is transferred from main partition to HA partition
  - Only a few minutes lost
- In case of Data corruption
  - Restore the tables from last backup
  - Up to 24 h of changes lost
  - Users are stopped while recovering the data

Case 2: 2 partitions with HA

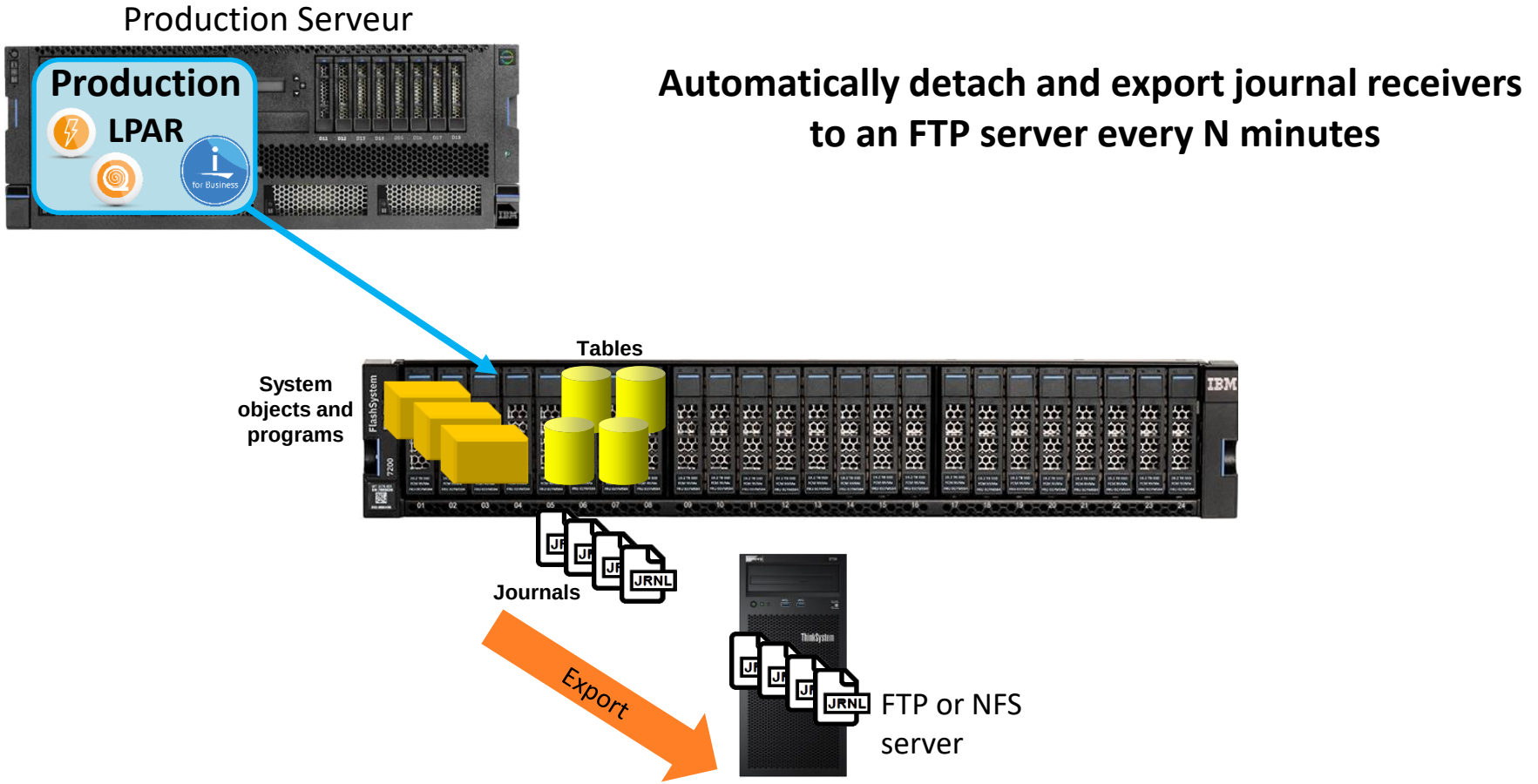
- In case of Data corruption
  - Restore the corrupted files or libraries
  - Apply the changes
- If the journals receivers have been exported to an external server, it also prevent from an “angry person” to clear the files and delete the receivers

- Use HA partition to recover the data instead of production partition
  - Less stress

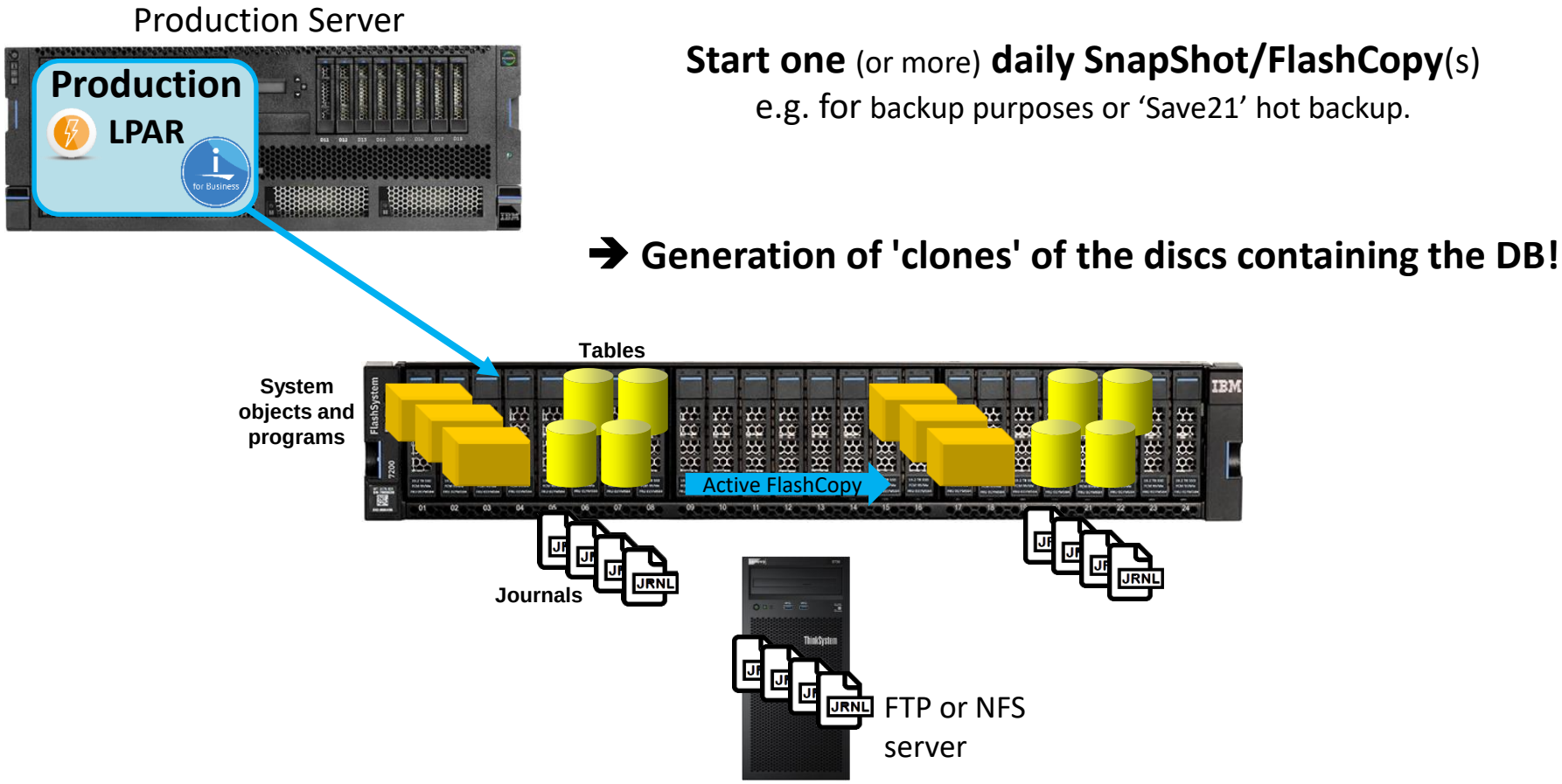
Case 3

Combine FlashCopy, Backups and journals

Case 3: Combine FlashCopy, Backup and journals



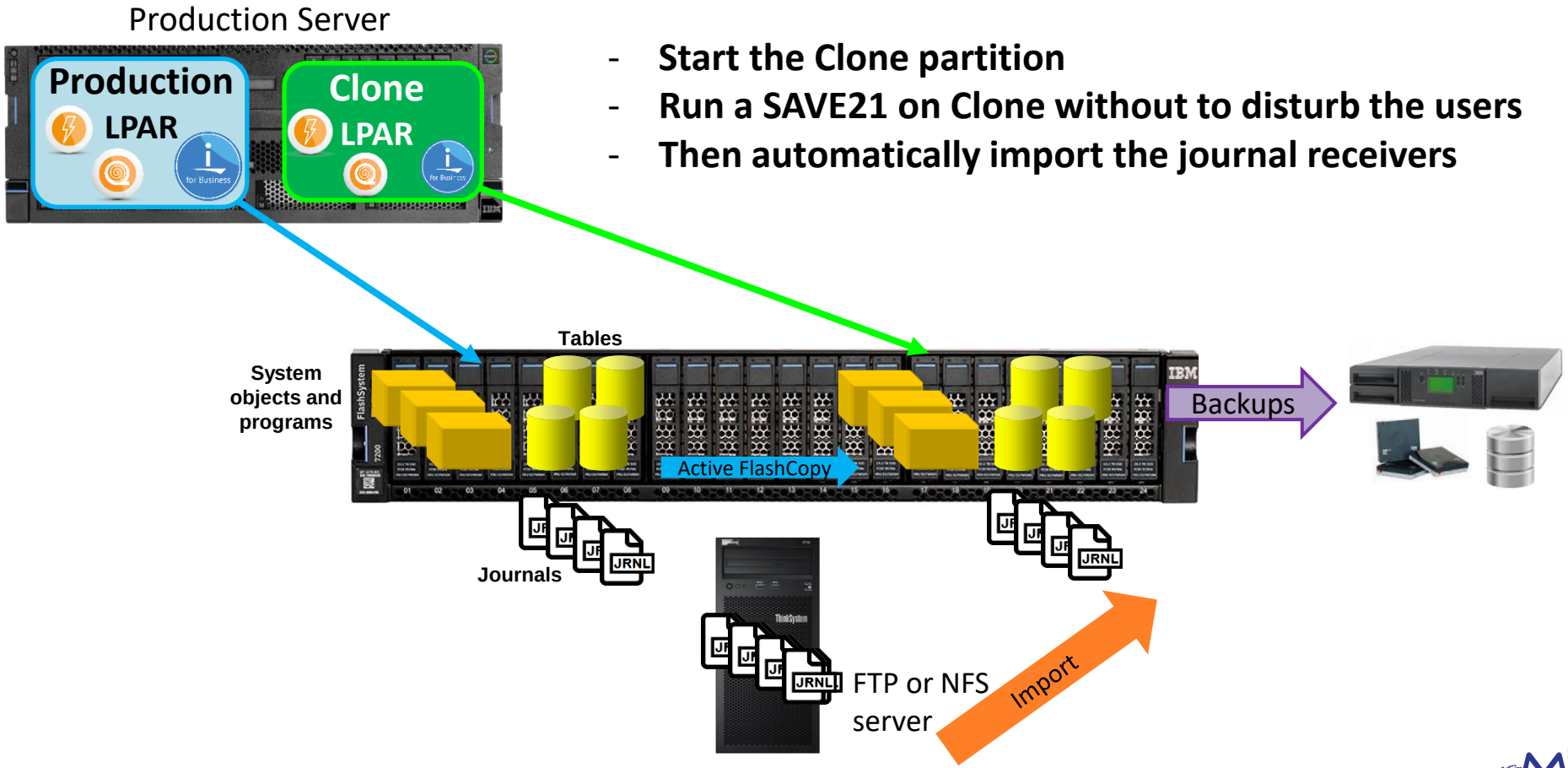
Case 3: Combine FlashCopy, Backup and journals



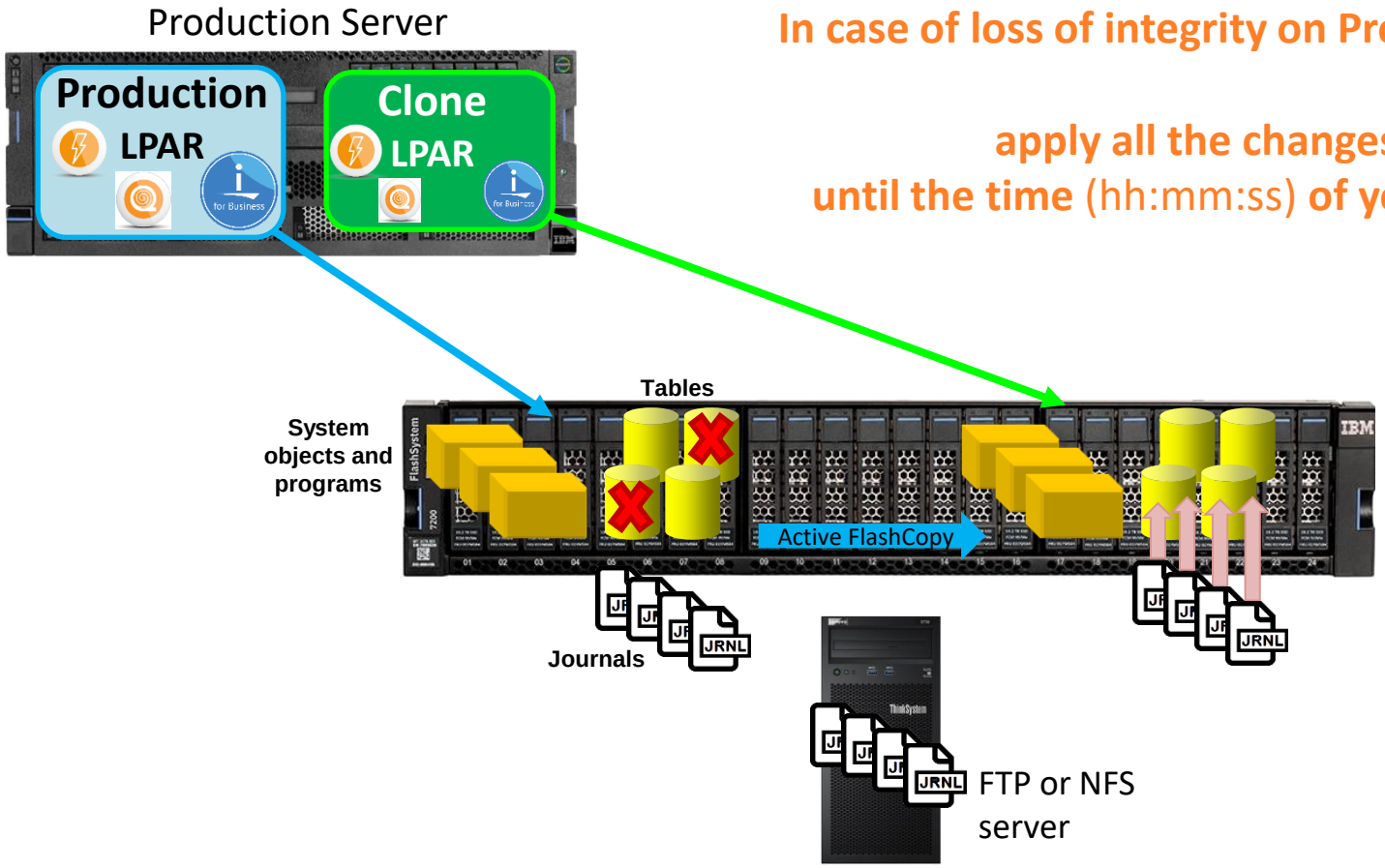
Start one (or more) **daily SnapShot/FlashCopy**(s)
e.g. for backup purposes or 'Save21' hot backup.

➔ **Generation of 'clones' of the discs containing the DB!**

Case 3: Combine FlashCopy, Backup and journals



Case 3: Combine FlashCopy, Backup and journals

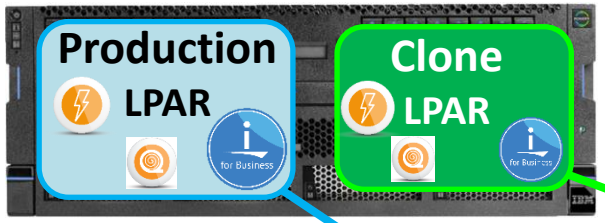


In case of loss of integrity on Prod database

apply all the changes
until the time (hh:mm:ss) **of your choice**

Case 3: Combine FlashCopy, Backup and journals

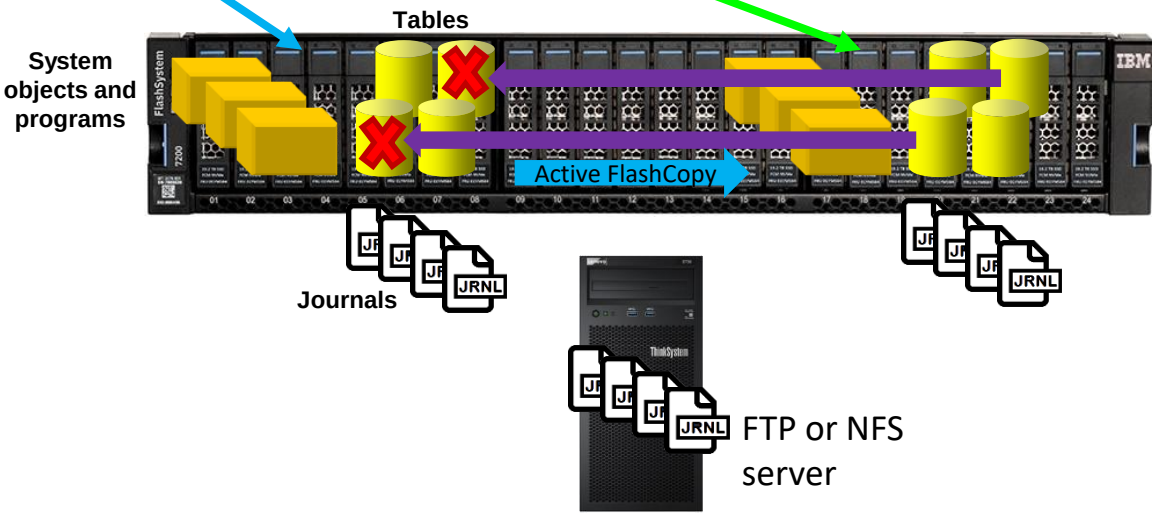
Production Server



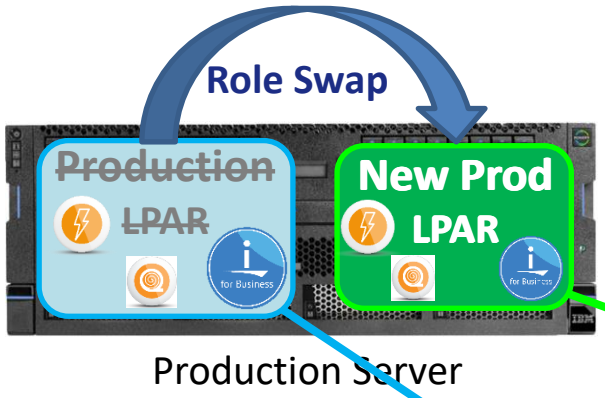
Then, in order to restart on an uncorrupted Database

Case 1 : only a few tables are corrupted

➔ **Backup, Transfer and Restore these tables to Prod partition**



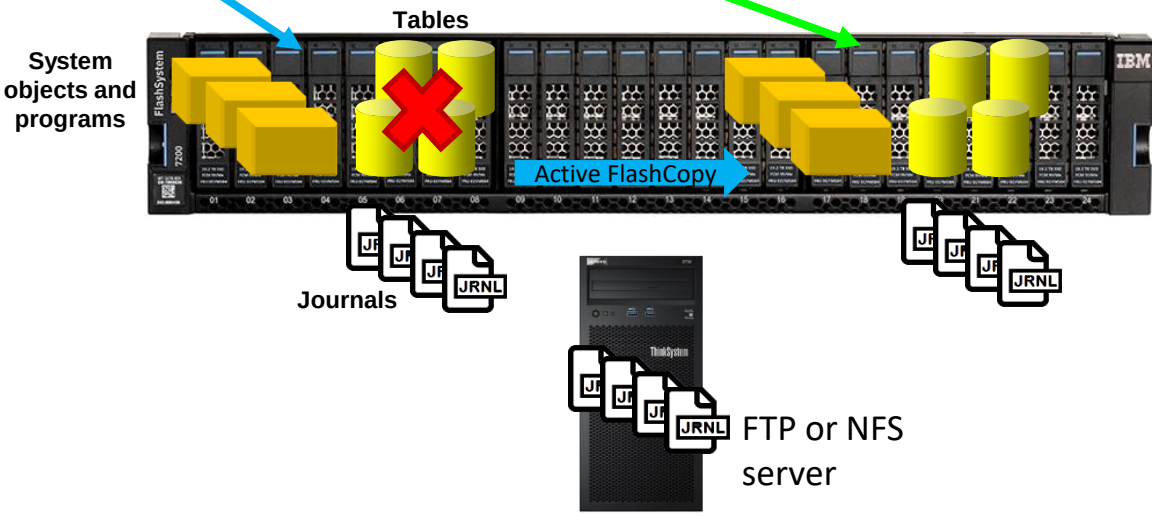
Case 3: Combine FlashCopy, Backup and journals



Then in order to restart on an uncorrupted Database

Case 2 : all the Database is corrupted

➔ **reverse the roles:**
the Clone partition becomes PROD



Contacts

Website
www.m81.eu

Trade and partner relations (French and English)
Bruno Tavern
+33 620 27 22 17
btaverne@m81.eu

Technical (French and English)
Pascal Ruckebusch
+33 689 05 76 55
pruckebusch@m81.eu

Commercial and technical DACH (German and English)
Karl Heinz Prisching
+43 676 633 06 05
khprisching@m81.eu

"**M81**" (or Bode galaxy) a very beautiful **spiral galaxy...**
located **12 million light years from Earth!**

End

See you soon we hope!

M81 : IBM i independant software vendor (ISV)

3 products : Flash for i, Recover for i, Control for i

Visit our booth : www.iug-ipower.com/sponsor-m81/