

Data Recovery Solutions - A Paradigm Shift

by Gordon Hahlweg - EMC Corporation

Thirty years ago, before I was born (well, almost) tape capture/restore was considered the ‘Be all to end all’ of TPF data recovery solutions. Fifteen years ago, when TPF was introduced to me as the ‘opportunity to see the world’, ACPDB was emerging as the logical or partial database recovery solution. Just a few years ago, CRUISE was released as the replacement for TPFDF capture/restore. Tape capture and CRUISE were acceptable solutions for their time, mostly because there were no alternatives.

Tape capture, together with many logging tapes, creates a complete point in time copy (to the nearest second) of a production TPF database on tape. CRUISE creates a copy of a partial or logical database on tape. Although CRUISE guarantees the integrity of individual sub-files, CRUISE does not create a point in time copy of the entire logical database. I leave it to the reader to decide on the completeness of either CRUISE or tape capture as a solution for data recovery in a TPF environment.

While you ponder the completeness of your current data recovery solutions, let me tell you about the not so new alternative; not so new because the technology has been available for nearly a decade, and the related TPF control software for about 2 or 3 years depending on whether we’re talking about remote copies or local copies. So why, if the TPF software has been around for so long, are you only now beginning to hear about it? The answer to that question is quite simply that any paradigm shift in TPF takes about 3 years to gain acceptance as a concept and between 5 and 10 years to be implemented within the TPF production environment. Anyone heard of the Coupling Facility? Great concept, great relief for those performance bottlenecks designed into MPLF, the connect path and the connect device. Plan on waiting another 3-8 years.

Before I turn this into a commentary on MPLF, let’s get back to TPF data recovery and the software solution alternatives to tape capture; EMC TimeFinder for TPF, and EMC SRDF Controls for TPF. TimeFinder and SRDF can be best described as a combination of host software and Symmetrix microcode, which enable the creation of local (TimeFinder) or remote (SRDF) point in time copies of one or more MDBF subsystems.

TimeFinder and SRDF TPF data structures are created by initialization routines using the module file status table as a guide to the TPF configuration, and a low level API to determine Symmetrix configuration information. Once the data structures have been initialized any EMC DR software functional entry will operate on all online and offline modules in all online EMC control units. Optionally, an EMC DR software functional entry can operate on one or a range of modules in any online EMC control unit. All low level API calls which control TimeFinder and SRDF microcode features result in CCWs issued on the traditional channel I/O interface. Hence, all functional entries operate only with online EMC control units.

The following procedure creates a point in time copy and then restores only the changed tracks since that copy was created. The configuration has been kept simple for ease of illustration. The availability of some product features mentioned in the procedure is dependent on microcode release and Symmetrix model. Product feature availability and other details are documented in the EMC TPF software product guides.

1. Illustrate the initialized configuration with a control unit summary display. Each control unit entry refers to a single Sub System ID (SSID). The control units of type “Remote” refer to remote subsystems found across SRDF links. These “Remote” Symmetrix are excluded from the following procedure.

ZUTIM DIS K CU

```
E1TQ0001I CU Keypoint Display
Serial #      Type  Model Ucod OSDA  MOD  SSN  GKD  BCV1  Cnt  STD1  Cnt
000184505047  Local  8430  5567 3300 0100 BSS No  0048  72  0000  36
000184500315  Remote 8430  5567 3300 0100 BSS No  0048  36  0000  36
100184505047  Local  8430  5567 33A2 0102 BSS No  00D8  72  0090  36
100184500315  Remote 8430  5567 33A2 0102 BSS No  00B4  36  006C  36
End of Display
```

- Establish BCV Group 2. This command assigns the BCV as the next available mirror of a standard Symmetrix device and copies the entire contents of the standard device to the BCV. The Establish operation is initiated for each online and offline module for every local online logical control unit defined to TimeFinder for TPF. The Establish also opens a Symmetrix Differential Data Facility (SDDF) session for the TPF module and its point in time copy. The SDDF session records track changes between the TPF module and the point in time copy. Once the SDDF session exists for this pair of volumes subsequent synchronization operations need only be incremental using the Re-establish and Incremental Restore.

The Re-establish command reassigns the BCV as the next available mirror of the standard device to which it was assigned before it was Split. Any tracks written on the BCV and/or standard device while the BCV was Split from the standard device are copied from the standard device to the BCV. The Incremental Restore command reassigns the BCV as next available mirror of the standard device to which it was assigned before the Split. Any tracks written on the BCV and/or standard device while the BCV was Split from the standard device are copied from the BCV to the standard device.

When TimeFinder for TPF starts issuing the Establish operation to one or more standard devices in a logical control unit it issues the UTIM0000I message. When TimeFinder for TPF has completed issuing the Establish operation to one or more standard devices in a logical control unit it issues the UTIM0001I message. Once TimeFinder for TPF has completed issuing the TimeFinder operation to all online logical control units, the TimeFinder Monitor is initiated. The TimeFinder monitor reviews and reports a summary status of all modules to which the operation was initiated. The TimeFinder Monitor will reinitiate itself at a user-defined interval, which defaults to 3 minutes, until it determines that the intended final Establish status is achieved.

ZUTIM EST G 2 ALL

```

UTIM0024I BCV Group keypoint validation completed
UTIM1000I TimeFinder started issuing Establish for CU 000184505047
UTIM1000I TimeFinder started issuing Establish for CU 100184505047
UTIM1001I TimeFinder completed issuing Establish for CU 000184505047
UTIM1001I TimeFinder completed issuing Establish for CU 000184505047
UTIM1031I Local TimeFinder Status Display
BCV Group: 2 Base Operation: Establish
  Start Time : 21.52.45 Date : 05/28/01

```

CU Serial #	Opr SDA	Complete	In Progress	Not Started	Opr RC Summary
000184505047	3300	0	36	0	0000
300184505047	33A2	0	36	0	0000

```

End of Display

```

- Split BCV Group 2 leaving the point in time copy ready on the channel. The Split command splits the BCV mirror from the standard Symmetrix device and makes it available (with the data from the standard device with which it was paired) to hosts through its separate device address. The Split operation is initiated for each online and offline module for every local online logical control unit defined to TimeFinder for TPF. When TimeFinder for TPF starts issuing the Split operation to one or more standard devices in a logical control unit it issues the UTIM0000I message. When TimeFinder for TPF has completed issuing the Split operation to one or more standard devices in a logical control unit it issues the UTIM0001I message. It is important to note that the Split is accomplished with only one I/O per logical subsystem. The multi-instant Split occurs at a single point in time for all modules in a single logical control unit. Completion of the point in time copy is a background operation. The URDY option can be used when TPF has no channel access (as defined by an IOCDS or CTK0) to the point in time copy, and duplicate VOLSERS are not a concern. At the time the modules are marked as Split, the point in time copy is accessible to any other host having channel access to it. Should the host with access to the point in time copy attempt to access a track not yet physically copied, the track will immediately be copied making it available to the requesting host.

The multi-instant Split mechanism results in a negligible net Split time in a logical control unit and seconds system wide, dependent on the number of logical control units being operated on in the TPF complex. The net Split time is measured from the first UTIM0000I message to the last UTIM0001I message. Should this granularity not be acceptable to the user, traditional exception recording can be started just prior to issuing the ZUTIM SPLIT and stopped following TimeFinder Monitor termination. This gross Split time is commonly in the order of 4 minutes dependent on the number of logical control units being

operated on in the TPF complex. The gross Split time consists primarily of the time to validate the TimeFinder for TPF data structures. TimeFinder for TPF also includes a Split user exit, which can be used to code a VFA flush routine. The user exit is entered just prior to the processing of the TimeFinder operation at the logical control unit level.

Once TimeFinder has completed issuing the multi-instant Split operation to all online logical control units, the TimeFinder Monitor is initiated. The TimeFinder monitor reviews and reports a summary status of all modules to which the operation was initiated. The TimeFinder Monitor will terminate immediately after the first initiation if it determines that the background Split has completed or is in progress for all relevant modules.

ZUTIM SPL G 2 ALL URDY

```

UTIM0024I BCV Group keypoint validation completed
UTIM1000I TimeFinder started issuing Split for CU 000184505047
UTIM1000I TimeFinder started issuing Split for CU 100184505047
UTIM1001I TimeFinder completed issuing Split for CU 000184505047
UTIM1001I TimeFinder completed issuing Split for CU 100184505047
UTIM1031I Local TimeFinder Status Display
BCV Group: 2 Base Operation: Split
  Start Time : 00.28.11 Date : 05/29/01

```

CU Serial #	Opr SDA	Complete	In Progress	Not Started	Opr RC Summary
000184505047	3300	3	33	0	0000
300184505047	33A2	6	30	0	0000

```

End of Display

```

Not to throw too much at you at once, but this is a good place to quickly illustrate a simple extension of this DR solution to include maintenance of a remote Point In Time Copy. The point in time copy, which has just been created, can be configured as the source for a remote or target point in time copy. Most likely, only that part of the point in time copy representing a simplex TPF system would be configured as the source for a target point in time copy. This target point in time copy might be a half test system. Without going into too much detail, the synchronization of source to target can be set up so as to enable an automatic synchronization of only the changed tracks source to the target at the time of the Split. Assuming the local point in time copy is configured as SRDF source volumes and a SRDF target volumes exist on the other side of one or more SRDF links, we can add one parameter to the previous Split command to achieve our goal of automatic synchronization of changed tracks only.

ZUTIM SPL G 2 ALL URDY DIFS

To disable automatic propagation of changed source tracks to the target, add one more parameter.

ZUTIM SPL G 2 ALL URDY DIFS DPR2

- Restore only the changed tracks since the last Split, and maintain the integrity of the point in time copy. The PROT parameter on the ZUTIM INCRESTORE or RESTORE command will leave the point in time copy modules write disabled, and ensure the integrity of the point in time copy. In order to maintain the integrity of the TimeFinder TPF data structures, the record ID of the TimeFinder data structure must be defined as VFA delay file, and use processor locking. The TPF complex must be folded down to one processor and this utility processor must be in 1052 or restart.

The Incremental Restore operation is initiated for each online and offline module for every local online logical subsystem defined to TimeFinder. When TimeFinder for TPF starts issuing the Incremental Restore operation to one or more standard devices in a logical control unit it issues the UTIM0000I message. When TimeFinder for TPF has completed issuing the Incremental Restore operation to one or more standard devices in a logical control unit it issues the UTIM0001I message. Once the TimeFinder monitor indicates that all modules are engaged in the incremental restore process, the TPF system has access to all data on the point in time copy from which the system is being restored. Should TPF attempt to access a track, which has not yet been restored, that track is immediately retrieved from the point in time copy module and made accessible to TPF.

IPL the utility processor to Restart or 1052 state.

ZRTDM MODIFY RECID-A386,VFAF-DELAY,LOCKF-PROC

ZUTIM INCRESTORE G 2 ALL PROT

```

UTIM0024I BCV Group keypoint validation completed
UTIM1000I TimeFinder started issuing Inc. Restore for CU 000184505047
UTIM1000I TimeFinder started issuing Inc. Restore for CU 100184505047
UTIM1001I TimeFinder completed issuing Inc. Restore for CU 000184505047
UTIM1001I TimeFinder completed issuing Inc. Restore for CU 100184505047
UTIM1031I Local TimeFinder Status Display
BCV Group: 2 Base Operation: Inc. Restore
  Start Time : 22.28.11 Date : 05/29/01

```

CU Serial #	SDA	Opr Complete	In Progress	Not Started	Opr RC Summary
000184505047	3300	3	33	0	0000
300184505047	33A2	3	33	0	0000

End of Display

5. IPL and Restart the TimeFinder Monitor for the Incremental Restore. The IPL will refresh the core copy of the restored TPF keypoints. Since the TimeFinder data structures have been filed over the restored data structures, the VFA delay file and processor locking attributes can be removed from the TimeFinder record ID RIAT definition. The ZUTIM RESTART command will recognize the fact that all modules are already in the required state. TimeFinder will not issue any operation for any online or offline module already in the required state.

Note that ZUTIM RESTART can be used to restart the issuing of a TimeFinder operation after an unplanned outage, or to review the state summary of the TimeFinder complex following normal termination of the TimeFinder Monitor as a result of the Symmetrix achieving the requested state for all modules. It is recommended that the user always enter the general status display, ZUTIM DIS Status to determine which operation will be restarted.

If required, the exception recording tape(s) (plural only if running loosely coupled and no XCP merge utility is available), can be restored following the Restart command. Otherwise, the TPF system has access to the point in time copy. The network can be restarted and TPF can be cycled to Norm State. If the TPF system is loosely coupled, all other TPF processors can be IPL'd and cycled to Norm State. The TimeFinder Monitor will reinitiate itself at a user-defined interval, which defaults to 3 minutes, until it determines that the intended final Incremental Restore status is achieved. Once the TimeFinder Monitor determines that the TPF production modules have been fully synchronized to the point in time copy it will issue a completed message and terminate.

IPL the utility processor to 1052 state.

ZRTDM MODIFY RECID-A386,VFAF-NO,LOCKF-DASD

ZUTIM RESTART

```

UTIM0024I BCV Group keypoint validation completed
UTIM1000I TimeFinder started issuing Inc. Restore for CU 000184505047
UTIM1000I TimeFinder started issuing Inc. Restore for CU 100184505047
UTIM1001I TimeFinder completed issuing Inc. Restore for CU 000184505047
UTIM1001I TimeFinder completed issuing Inc. Restore for CU 100184505047
UTIM1031I Local TimeFinder Status Display
BCV Group: 2 Base Operation: Inc. Restore
  Start Time : 22.28.11 Date : 05/29/01

```

CU Serial #	SDA	Opr Complete	In Progress	Not Started	Opr RC Summary
000184505047	3300	9	27	0	0000
300184505047	33A2	11	25	0	0000

End of Display

ZCYCL NORM ALL

6. Split BCV Group 2 leaving the point in time copy ready on the channel. The Split operation is initiated for each online and offline module for every local online logical control unit defined to TimeFinder for TPF. When TimeFinder for TPF starts issuing the Split operation to one or more standard devices in a logical control unit it issues the UTIM0000I message. When TimeFinder for TPF has completed issuing the Split operation to one or more standard devices in a logical control unit it issues the UTIM0001I message. Review step 3 for a more comprehensive explanation of the multi-instant Split operation. Note that since the PROT parameter was used on the INCRESTORE command, the point in time copy is the same as it was following the Split in step 3.

ZUTIM SPL G 2 ALL URDY

```

UTIM0024I BCV Group keypoint validation completed
UTIM1000I TimeFinder started issuing Split for CU 000184505047
UTIM1000I TimeFinder started issuing Split for CU 100184505047
UTIM1001I TimeFinder completed issuing Split for CU 000184505047
UTIM1001I TimeFinder completed issuing Split for CU 100184505047
UTIM1031I Local TimeFinder Status Display
BCV Group: 2 Base Operation: Split
  Start Time : 23.11.34 Date : 05/29/01
      Opr   _____ Operation Status_____ Opr RC
CU Serial #  SDA  Complete In Progress  Not Started  Summary
000184505047 3300   21           15           0           0000
300184505047 33A2   19           17           0           0000
End of Display

```

The steps outlined above are general in nature and are not likely exactly the way it would be done at your particular site, however they should serve to illustrate why there is no need to wait out the next 5-8 years to make this work for your company. EMC TimeFinder for TPF and EMC SRDF Controls for TPF are not only an alternative to tape capture/restore, rather as I will illustrate in the next installment of Data Recovery Solutions, EMC Data Recovery Software is the comprehensive DR Solution for TPF.