POWER NX842 Compression for Db2

Ultrafast, CPU-Free Backup and Log Archive Compression

Version 1.0, September 29, 2017

This article describes what NX842 compression is about and how you can use it to minimize space and time when creating backup images and archive logs for Db2®. The impressive results prove that a well-matched combination of hardware and software is the right approach to achieve the best results.

Joern Klauke - Db2 Advanced Support Europe & MEA, IBM Germany
jklauke@de.ibm.com
Holger Hellmuth – Db2 Development, IBM Germany R&D
hel@de.ibm.com

Table of Contents

Table of Figures

Management Summary

NX842 hardware-accelerated compression for backup and log file archiving in Db2 11.1 can be implemented very easily even at no additional cost if the AME license is already available.

To prove the strength of Db2 NX842 hardware compression for backups and log file archiving, we tested the available compression options for backup and log file archiving: No compression, Db2 software compression and Db2 NX842 hardware-accelerated compression.

Out of these three possibilities, NX842 backup compression was always the fastest backup option – even faster than uncompressed backups.

Despite the lowest backup runtime, the CPU consumption for NX842 backups was also lower than for uncompressed backups.

The bottom line is: You get fast compressed backups with the lowest CPU consumption if you use NX842 compression.

As the AIX kernel offers a software implementation of the NX842 decompression algorithm, compressed backups and archive log files can be used on POWER systems without an NX842 hardware compression unit.

Db2, POWER hardware and AIX represent an excellent combination that assures optimal resource usage at low cost and minimal administrative effort.

Why All That Fuss About Compression?

Databases are growing rapidly as more and more data is available in digitized form. To control that growth, all kinds of compression technologies are developed. Db2 offers several compression techniques to control the database size and make the database operation faster. But are backup image and log file compression really necessary? In the following, we share some thoughts that need to be considered to answer that question.

The most obvious reason for compressing backup images and log files is the reduced storage requirement. Both types of files are usually stored for a longer period of time and can cause a significant storage footprint. But it's not only the obvious storage footprint which should be considered, but also the question of speed. Smaller backups and smaller log files can be transferred much faster over the network – especially if the network is already heavily used – and therefore speed up the backup and log file archiving process. That way smaller files can also relieve the network and storage systems, and accelerate backup and log file archiving.

Now, are those benefits still present when all database compression methods are used excessively like data, index, or even BLU compression? Yes, they are. Of course, even well-compressed databases benefit from backup and log file compression. Experience has shown that additional reductions can be achieved from non-inlined LOBs, repeating patterns in database metadata, log record headers and even BLU-compressed pages – not to mention completely free pages or partially filled pages.

The bottom line is: Compression results in faster operations, resource optimization and hence cost optimization.

What NX842 Compression Is About

POWER7+™ and POWER8® processors contain special acceleration units for cryptography and compression (NX842).

By default, those accelerators are used to compress main memory, which is known as Active Memory Expansion (AME). AME is based on the 842 algorithm, which is related to Lempel-Ziv. The current implementation is optimized for memory compression and decompression and requires only a small footprint on the chip.

The accelerator unit receives incoming data in eight byte chunks per cycle, splitting the data in pieces of **8**, 2 x **4**, 4 x **2** bytes, and compares those pieces with already received data for data reduction – that's where the "842" in the name "NX842" comes from. The POWER7+™ accelerator unit is running at a clock speed of 2.3 GHz. A raw throughput of more than 18 GB/sec and unit is achieved at an extreme low power consumption of under 2 W. This is just a small example, the Power architecture offers a wide range of features and benefits like on-chip accelerators and algorithms. You can find more details in the excellent IEEE article listed in the References section at the end of this document.

The various Db2 compression techniques like deep row compression, adaptive compression, actionable compression for databases using BLU technology, and index compression make the additional memory compression offered by AME rather small as already compressed data cannot be compressed twice. So why not use the built-in AME accelerators for other important tasks within Db2?

Starting with Db2 11.1, the NX842 accelerators can be used to compress database archive log files and backup images. Before, only Db2 software-based compression was available to compress backups and archive log files. Compared to software compression, this hardware feature doesn't consume CPU resources. These savings (avoided CPU cycles) give the database additional headroom to handle the remaining database workload faster. Moreover, hardware compression is much faster than software compression.

The overall advantage of compression remains the same, that is, backup images are smaller, thus saving space in the backup infrastructure and network bandwidth. During the tests we found that the NX842 compressed backup runtime is lower than for uncompressed backups with a reduced CPU consumption.

On non-hardware-accelerated POWER-based hardware (prior to POWER 7+), the AIX kernel contains a software implementation of the decompression algorithm so that NX842 hardware-compressed backups can be restored and the rollforward operation process can be initiated.

Preparing the Server

For optimal performance on POWER8 servers, the following minimum firmware levels are needed for NX842 compression:

- FW820.50 - SC820_103
- FW830.30 - SC830_97
- FW840.40 - SC840_147

Using the NX842 accelerator unit requires the installation of the AME license for the server. Figure 1 shows the server settings for the Hardware Management Console (HMC) enhanced interface. The required licensed capabilities are selected:

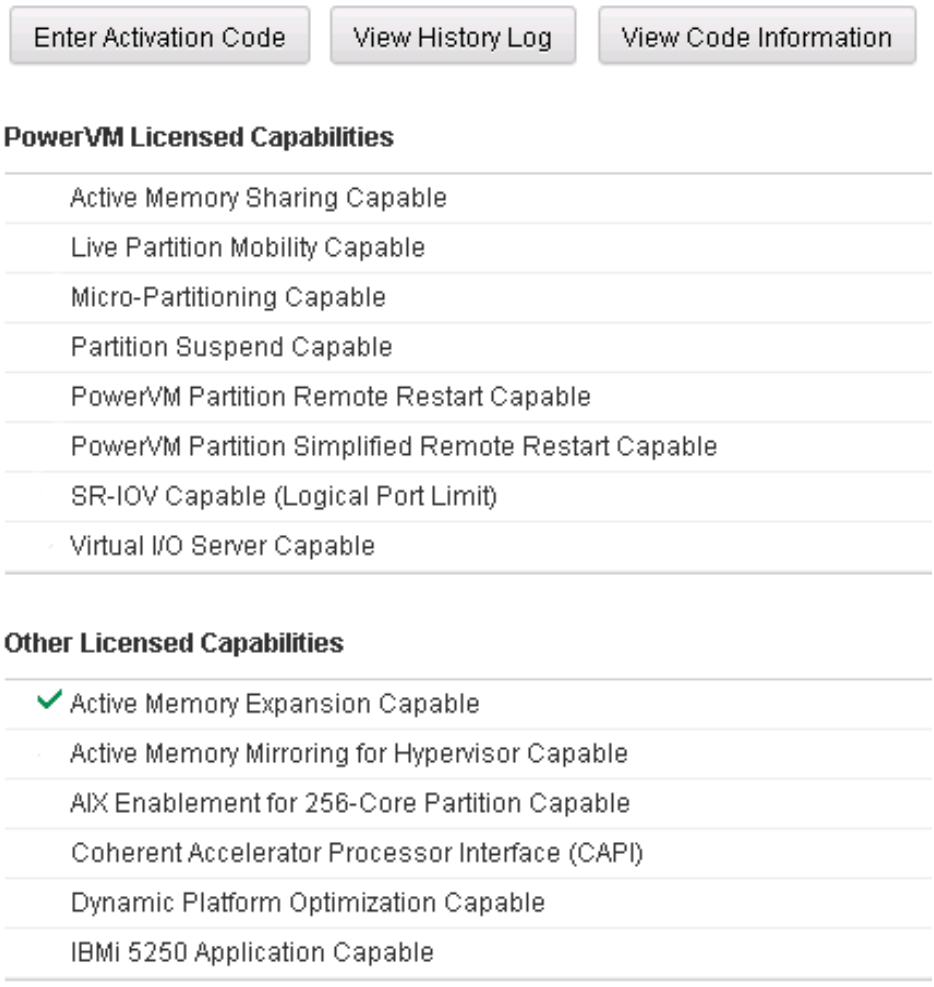

Figure 1: Required Licensed Capabilities POWER Server – HMC Enhanced Interface

Figure 2 shows the correct license settings on the HMC classic interface. The licensed capabilities are listed in more detail: Hardware-Accelerated Active Memory Expansion is listed in addition to Active Memory Expansion.

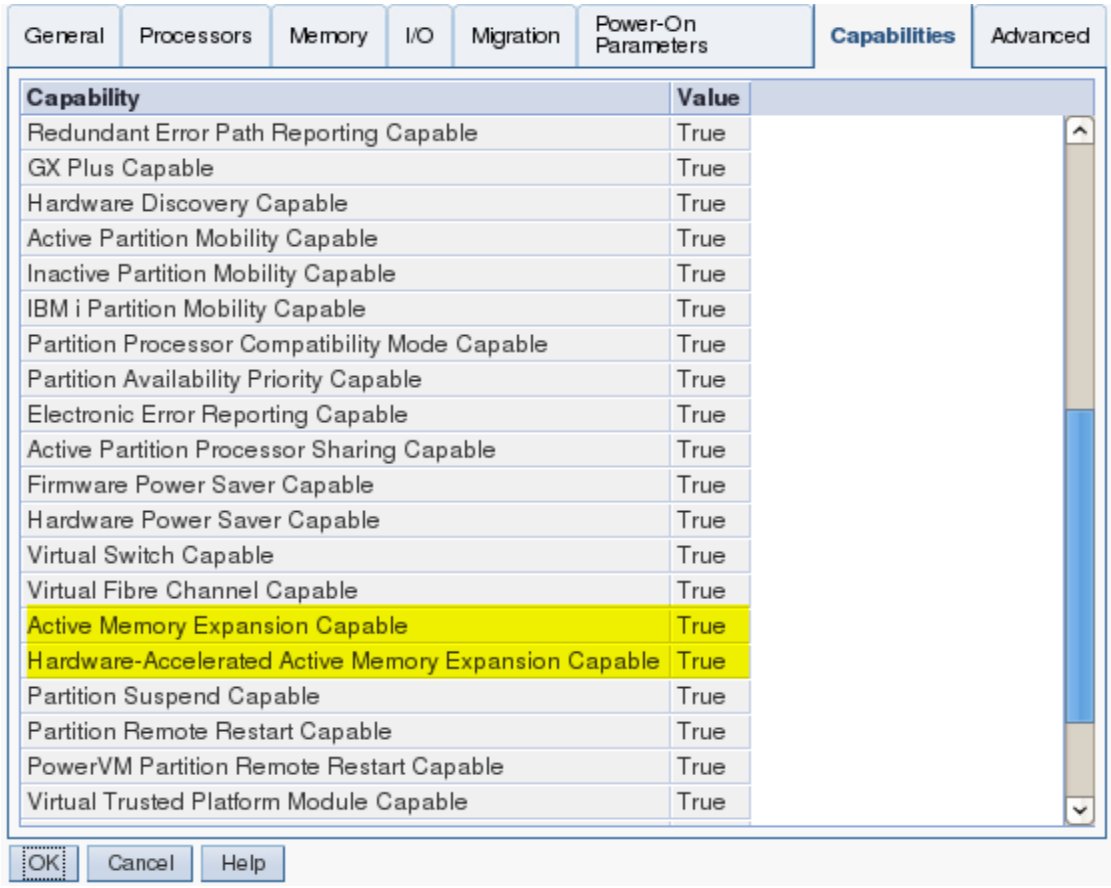

Figure 2: Required Licensed Capabilities POWER Server – HMC Classic Interface

Preparing the LPAR

NX842 compression with Db2 is currently only available on AIX®. The necessary minimum AIX levels are AIX V7 TL3 SP5 and AIX V6 TL9 SP3. If you meet the requirements for Db2 11.1, the OS requirements for NX842 are met as well.

After the AME license has been applied to the server, the LPAR needs to be rebooted.

As described in the IBM Knowledge Center for Db2 11.1, Active Memory Expansion (AME) and Active Memory Sharing (AMS) have to be switched off on the database LPAR. Otherwise, Db2 cannot take advantage of NX842 hardware compression.

You can use the AIX tool “amepat” to check if both AME and AMS are disabled. Figure 3 shows the required settings:

```
# amepat

Command Invoked                : amepat

Date/Time of invocation        : Thu Oct 13 15:24:33
Total Monitored time           : NA
Total Samples Collected        : NA

System Configuration:
---------------------
Partition Name                 : db6p041001
Processor Implementation Mode  : POWER8 Mode
Number Of Logical CPUs         : 12
Processor Entitled Capacity    : 3.00
Processor Max. Capacity        : 3.00
True Memory                    : 16.00 GB
SMT Threads                    : 8
Shared Processor Mode          : Disabled
Active Memory Sharing          : Disabled
Active Memory Expansion        : Disabled

System Resource Statistics:              Current
---------------------------          ----------------
CPU Util (Phys. Processors)                0.02 [  1%]
Virtual Memory Size (MB)                   8935 [ 55%]
True Memory In-Use (MB)                   16260 [ 99%]
Pinned Memory (MB)                         3277 [ 20%]
File Cache Size (MB)                       7050 [ 43%]
Available Memory (MB)                      6691 [ 41%]
```

Figure 3: Required AME/AMS settings for a database LPAR using NX842 compression

NX842 Compression with Db2 11.1

Power Servers w/o NX842 Support

Hardware-accelerated backups that were taken on POWER7+ and POWER8 servers can still be restored and rolled forward also on servers that do not offer hardware acceleration – like POWER6® or POWER7®. The AIX kernel contains a software implementation of the decompression algorithm so that decompression will work even on non-hardware-accelerated POWER-based hardware. Software decompression requires AIX and is not supported on other platforms.

In case hardware-accelerated log file compression was not disabled after a restore to POWER6, an error message is written in the db2diag.log. Due to the missing acceleration unit in POWER6, Db2 archives the log file uncompressed. To get rid of these error messages, the database configuration parameters LOGARCHCOMPR1 and LOGARCHCOMP2 need to be set to OFF (no compression) or ON (Db2 software compression).

Using NX842 Compression for Backup

Use the following command to run a backup with NX842 compression:

```
backup database <dbname> compress comprlib libdb2nx842.a
```

You can also use the registry variable DB2_BCKP_COMPRESSION to determine the standard compression method. The following command sets the default to NX842:

```
db2set DB2_BCKP_COMPRESSION=NX842
```

When setting the standard compression method to NX842, the backup command can be used as usual:

```
backup database <dbname> compress
```

Using NX842 Compression for Log Archiving

To use the new compression method also for log archiving, switch either LOGARCHCOMPR1 or LOGARCHCOMPR2, or both, to NX842:

```
UPDATE DATABASE CONFIGURATION FOR <dbname> USING LOGARCHCOMPR1 NX842
UPDATE DATABASE CONFIGURATION FOR <dbname> USING LOGARCHCOMPR2 NX842
```

Backup and Log File Archiving Performance

Main Steps of the Db2 Backup Process

Issuing the Db2 backup command starts or reuses a free db2agent as the controlling thread for the complete backup process. The selected parallelism value in the db2 backup command determines the amount of buffer manipulators started by the db2agent. Each buffer manipulator is responsible for reading the data from one tablespace. As the tablespaces usually consist of several container files, the buffer manipulators use prefetchers to read the container data from disk into the backup buffers in memory.

Depending on the number of backup targets used in the issued db2 backup command, a corresponding number of media controllers are started by the db2agent. The media controllers read the data from the backup buffers and write it to the backup target.

The time reading the data from disk into memory depends heavily on the installed hardware and is a major part of the backup runtime. Writing the data from the backup buffers into the backup target is the other major part influencing the backup runtime. The latter part of the runtime can be minimized by reducing the backup size through compression. This is where a fast backup compression option, like NX842 compression, can make the difference – especially for I/O bound systems.

You can find a more detailed description of the backup process in the IBM Redbook: Maximizing Performance of IBM Db2 Backups.

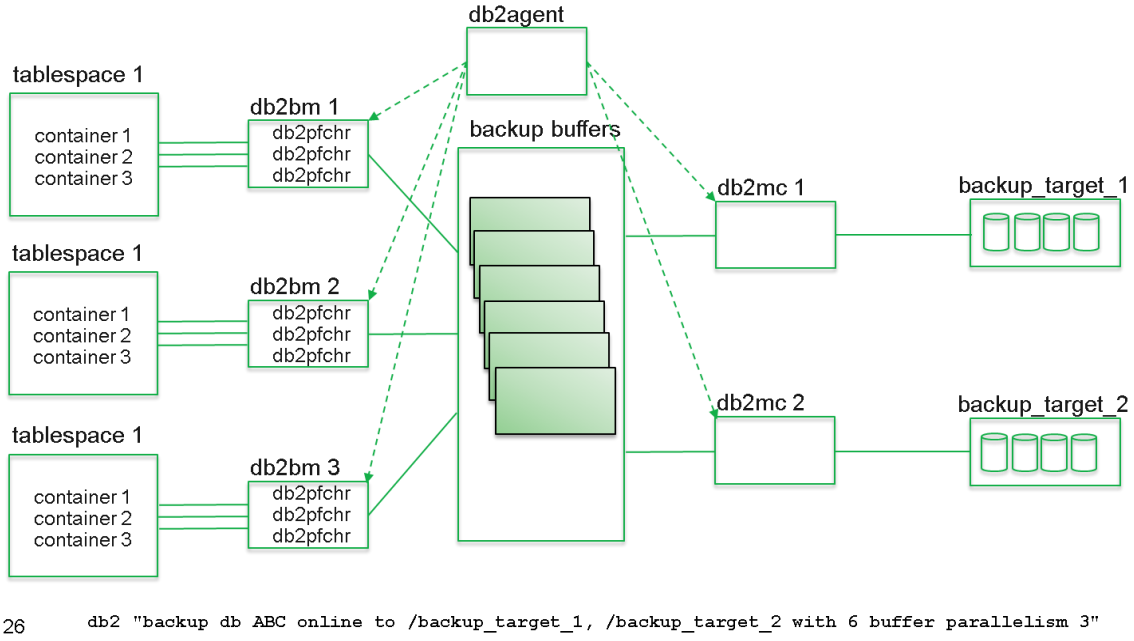

Figure 4: Overview of the Db2 backup process

Test Environment for Backup and Log File Archiving

The test LPAR was located on a POWER8 S824 server using AIX 7.1 TL3 SP 5 with four cores of an 8-core chip and 16GB of main memory. The data was distributed over four SAN disks, and the backup images were written to four additional disks.

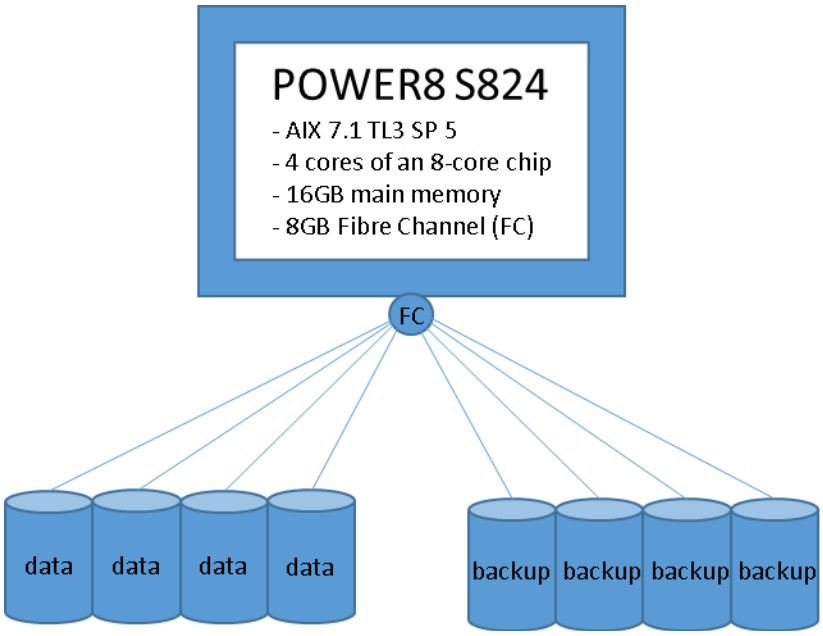

Figure 5: NX842 Test Environment

The backup tests were performed on a database with an overall compression rate of approx. 50% and a size of approx. 260GB. The overall compression rate was ascertained using the hundred largest tables, indexes and LOBs. Those three hundred database objects represented about 94 percent of the 260GB database size.

The following three backup types were tested:

- Uncompressed (u)
- Db2 software compressed (c)
- Db2 NX842 hardware compressed (n)

Three characteristics were used to judge the behavior of the three different backup types:

- Backup size
- Backup runtime
- CPU consumption during the backup

Test Results: Backup

Size view:

In the test environment, the size of an NX842 compressed backup was reduced to about 60% of the size of an uncompressed backup.

Software compression reduced the backup size to about 50% of an uncompressed backup.

Runtime view:

Figure 6 shows the backup runtime for the three different backup types.

NX842 compressed backups always had the fastest backup runtime and were even faster than uncompressed backups. As the hardware compression unit is extremely fast, there is no need to use high parallelism values to speed up the backup process for NX842 backups.

During the backup tests the NX842 compression backup runtime was 2.4 times faster than the software compression runtime.

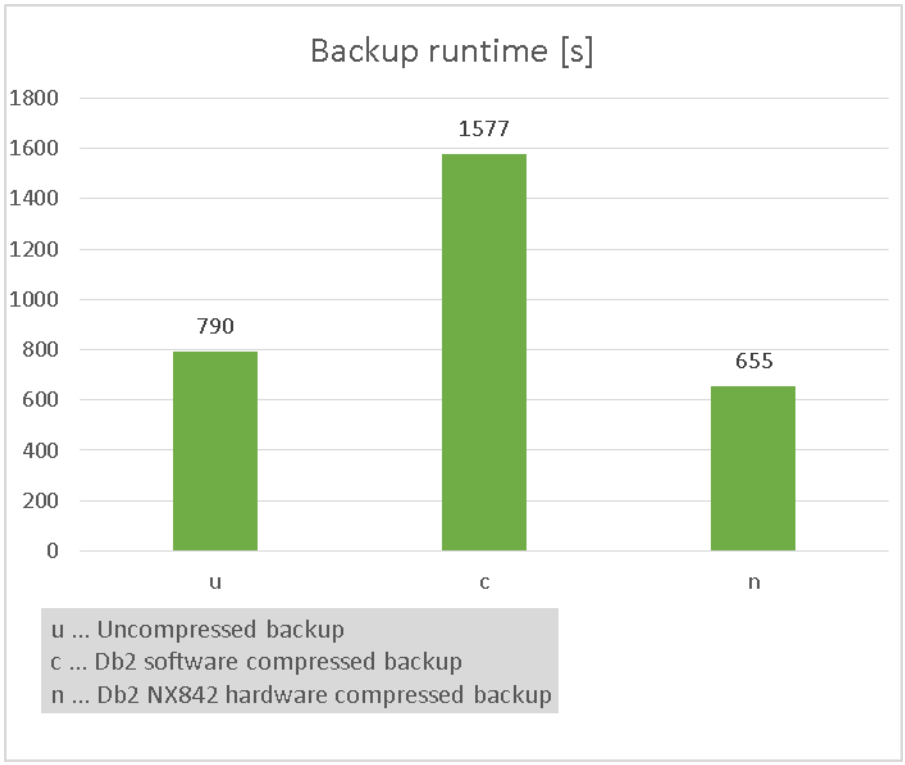

Figure 6: Backup runtime for the different backup types

CPU view:

Comparing the CPU usage during the different test cases led to impressive results for NX842 compression.

Software compression needed about twice as many CPU cycles than NX842 compression, whereas NX842 compression required fewer CPU cycles than backups without compression. Figure 7 shows the CPU usage for the test cases.

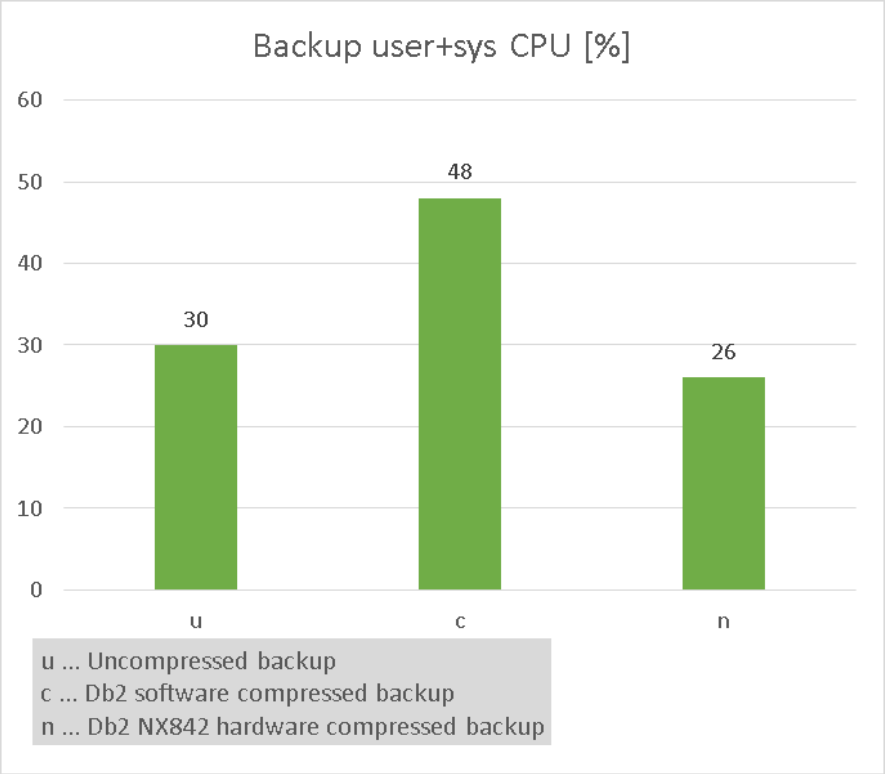

Figure 7: CPU usage of the different backup test cases

Test Results: Log File Archiving

Enabling log file archiving is an easy task as shown on page 11. The advantages are impressive especially for I/O-bound systems.
Figure 8 provides an overview of the benefits that compression offers for log file archiving.

The log file size, archiving time and CPU usage are used to compare the different compression options: no compression, Db2 software compression and NX842 hardware compression.

Runtime and size view:

Archiving a log file with Db2 NX842 compression is 25% faster than archiving a log file with Db2 software compression. At the same time, the log file size is reduced by 45% compared to an uncompressed log file.

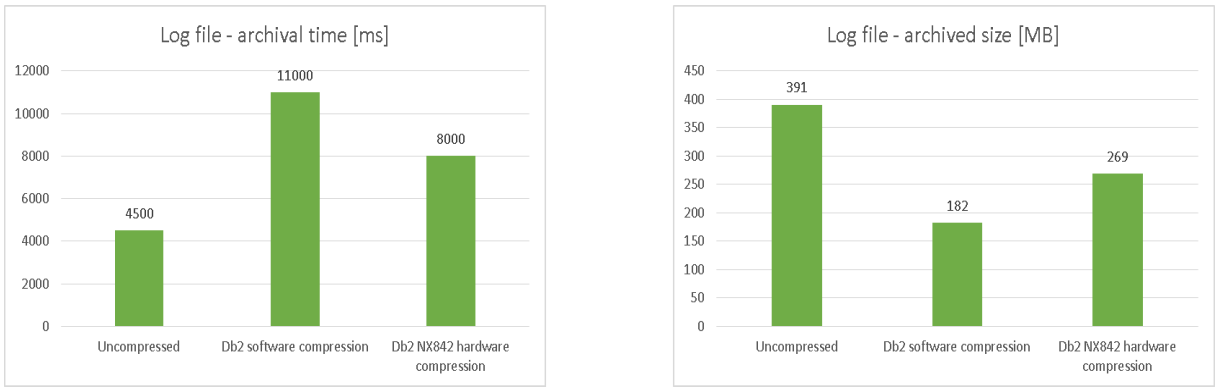

Figure 8: Benefits using archive log file compression

CPU view:

Due to the POWER coprocessor taking over the compression workload for log file archiving when NX842 compression is used, software compression consumes more than twice as many CPU cycles (+116%).

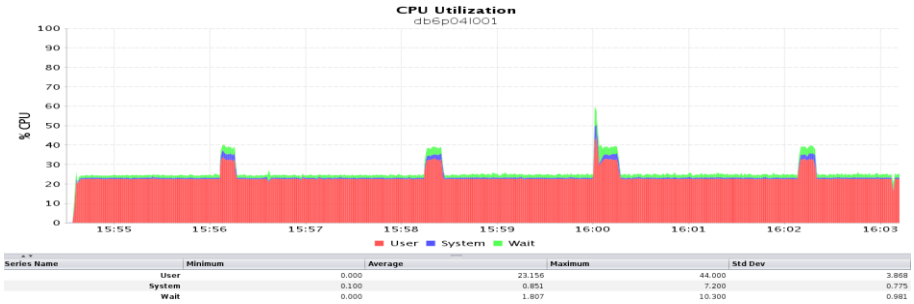

Figure 9: Db2 archive compression ~ +13% CPU usage during archival

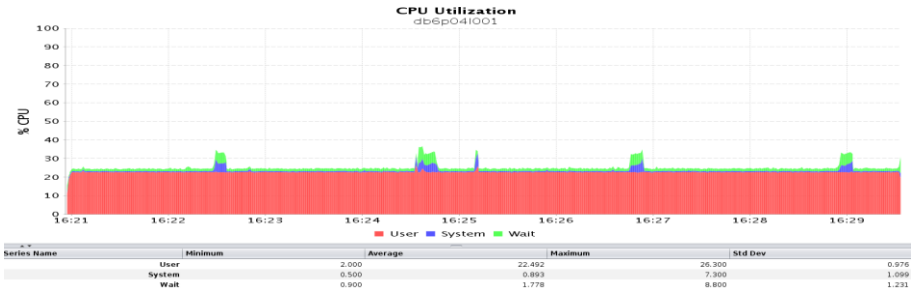

Figure 10:NX842 archive compression ~ +6% CPU usage during archival

Monitoring and Diagnostics of NX842 Compression

Backup Statistics

Figure 11 shows the performance statistics for the NX842 backup (from figure 6 and 7) and can be found in the db2diag.log after each finished backup.

Each line (000 to 003) under *BM#* provides detailed information about the time a buffer manipulator (BM) is spending with reading data. Under MC# (000 to 003), the information for the media controllers, which are writing data to disk, is listed in the same way. The number of buffer manipulators listed is defined via the parallelism option in the backup command, while the number of media controllers is defined by the number of backup targets.

```
2016-08-25-12.01.57.841616-300 E15769840A2034       LEVEL: Info
PID     : 10092648             TID : 39895          PROC : db2sysc 0
INSTANCE: db2vfb               NODE : 000           DB   : VFB
APPHDL  : 0-2350               APPID: *LOCAL.db2vfb.160825165101
AUTHID  : DB2VFB               HOSTNAME: rubix-lp02.aus.stglabs.ibm.com
EDUID   : 39895                EDUNAME: db2agent (VFB) 0
FUNCTION: DB2 UDB, database utilities, sqluxLogDataStats, probe:395
MESSAGE : Performance statistics
DATA #1 : String, 1514 bytes

Parallelism       = 4
Number of buffers = 8
Buffer size       = 16781312 (4097 4kB pages)
```

1 BM#	2 Total	3 I/O	4 Compr	5 MsgQ	6 WaitQ	Buffers	7 MBytes	8 Compr MBytes	
000	655.56	588.94	62.05	0.06	0.30	3406	74941	100569	25614
001	655.56	599.01	37.13	0.02	17.29	1665	54944	67462	12509
002	655.34	586.01	48.81	0.03	17.41	2523	62758	81687	18928
003	655.34	588.57	46.71	0.03	17.57	2212	66360	82893	16533
TOT	2621.81	2362.55	194.71	0.15	52.58	9806	259004	332613	

MC#	Total	I/O	MsgQ	WaitQ	Buffers	MBytes
000	655.56	122.44	533.07	0.00	2474	39577
001	655.62	121.34	533.96	0.24	2441	39033
002	655.56	121.34	533.90	0.24	2458	39321
003	655.56	120.33	534.93	0.24	2437	38985
TOT	2622.31	485.46	2135.88	0.72	9810	156918

```
2016-08-25-12.01.57.841914-300 E15771875A477       LEVEL: Info
PID     : 10092648             TID : 39895          PROC : db2sysc 0
INSTANCE: db2vfb               NODE : 000           DB   : VFB
APPHDL  : 0-2350               APPID: *LOCAL.db2vfb.160825165101
AUTHID  : DB2VFB               HOSTNAME: rubix-lp02.aus.stglabs.ibm.com
EDUID   : 39895                EDUNAME: db2agent (VFB) 0
FUNCTION: DB2 UDB, database utilities, sqlubcka, probe:1070
MESSAGE : Backup complete.
```

Figure 11: db2diag.log performance statistics for each completed backup

1 *BM* (buffer manipulator): Number of buffer manipulators launched

MC (media controller): Number of media controllers launched

2 *Total*: Time a BM or MC exists which equals the backup runtime

3 *I/O*: Time spent for read and write operations

4 *Compr*: Time spent for the compression

5 *MsgQ*: Time waiting for an I/O buffer

6 *WaitQ*: Time waiting for next instruction from the db2agent

7 *Mbytes*: Amount of uncompressed and compressed data read or written by the different BM or MC. *Mbytes TOT* shows the size of the uncompressed and compressed backup.

8 *Compr MBytes*: quantity of uncompressed data that was compressed

Typical Error Conditions

Short error messages will be given on the command line as a direct response to a failing command. You can always find more details in the db2diag.log.

The following message typically means that the backup is either not running on POWER7+ or POWER8 systems, or that the AME license is not installed. Similar messages might occur if AME and/or AMS are enabled.

```
db2 backup db sample compress comprlib libdb2nx842.a

SQL2062N  An error occurred while accessing media "libdb2nx842.a".
Reason
code: "1".
```

The db2diag.log contains more information about the current error condition and explains how to overcome the issue. The following two db2diag.log entries provide the exact reason why the above command failed with an error message.

The first db2diag.log entry shows that the error is caused by the enabled AME functionality on the LPAR:

```
2016-03-16-09.27.16.004937-240 E5606375A501         LEVEL: Error
PID     : 31720066             TID : 34954          PROC : db2sysc
INSTANCE: jklauke              NODE : 000           DB   : SAMPLE
APPHDL  : 0-11                 APPID: *LOCAL.jklauke.160316132714
AUTHID  : JKLAUKE              HOSTNAME: hal9000
EDUID   : 34954                EDUNAME: db2bm.34697.0 (SAMPLE)
FUNCTION: DB2 UDB, database utilities, InitCompression, probe:414
MESSAGE : AME is enabled.
```

The second db2diag.log entry shows that the error is caused this time by the missing hardware acceleration unit in POWER6:

```
2016-06-29-18.04.30.411779+120 E270031A402          LEVEL: Error
PID     : 10747954             TID : 3857           PROC : db2sysc 0
INSTANCE: db2vfb               NODE : 000           DB   : VFB
HOSTNAME: db6p1lp06
EDUID   : 3857                 EDUNAME: db2logmgr (VFB) 0
FUNCTION: DB2 UDB, database utilities, InitCompression, probe:376
MESSAGE : The compression accelerator is not available.
```

References

- IBM POWER7+ processor on-chip accelerators for cryptography and active memory expansion: http://ieeexplore.ieee.org/document/6665020/

- AIX Knowledge Center
  - Active Memory Expansion (AME) Introduction: https://www.ibm.com/support/knowledgecenter/ssw_aix_71/com.ibm.aix.performance/intro_ame_process.htm
  - Active Memory Expansion Planning and Advisory Tool *amepat:* https://www.ibm.com/support/knowledgecenter/ssw_aix_71/com.ibm.aix.cmds1/amepat.htm
  - accel_compress() subroutine using 842 algorithm for compressing data in POWER8 accelerator units: https://www.ibm.com/support/knowledgecenter/ssw_aix_71/com.ibm.aix.basetrf1/accel_compress.htm

- Db2 Knowledge Center
  - Hardware accelerated backup and log file compression: https://www.ibm.com/support/knowledgecenter/SSEPGG_11.1.0/com.ibm.db2.luw.admin.ha.doc/doc/c0062003.html
  - Backup and restore statistics: https://www.ibm.com/support/knowledgecenter/SSEPGG_11.1.0/com.ibm.db2.luw.admin.ha.doc/doc/c0060744.html

- IBM Redbook *Maximizing Performance of IBM Db2 Backups:* http://www.redbooks.ibm.com/technotes/tips1344.pdf

Disclaimers

The information provided by this media supports the products and services described with consideration for the conditions described herein.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:
INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON - INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at:
http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:
AIX®, DB2®, Db2®, Power Architecture®, POWER®, POWER7®, POWER7+™, POWER8®

The following terms are trademarks of other companies:
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Other company, product, or service names may be trademarks or service marks of others.