IMPROVEMENT OF THE ORACLE SETUP AND DATABASE DESIGN AT THE HEIDELBERG ION THERAPY CENTER

K. Höppner*, Th. Haberer, J. M. Mosthaf, A. Peters,
HIT, Heidelberg Ion Therapy Center, University Hospital, Heidelberg, Germany
M. Thomas, A. Welde, Eckelmann AG, Wiesbaden, Germany
G. Fröhlich, S. Jülicher, V. RW Schaa, W. Schiebel, S. Steinmetz,
GSI Helmholtzzentrum für Schwerionenforschung, Darmstadt, Germany

Abstract

The HIT (Heidelberg Ion Therapy) center is an accelerator facility for cancer therapy using both carbon ions and protons, located at the university hospital in Heidelberg. It provides three therapy treatment rooms: two with fixed beam exit (both in clinical use), and a unique gantry with a rotating beam head, currently under commissioning. The backbone of the proprietary accelerator control system consists of an Oracle database running on a Windows server, storing and delivering data of beam cycles, error logging, measured values, and the device parameters and beam settings for about 100,000 combinations of energy, beam size and particle rate used in treatment plans. Since going operational, we found some performance problems with the current database setup. Thus, we started an analysis in cooperation with the industrial supplier of the control system (Eckelmann AG) and the GSI Helmholtzzentrum für Schwerionenforschung. It focused on the following topics: hardware resources of the DB server, configuration of the Oracle instance, and a review of the database design that underwent several changes since its original design. The analysis revealed issues on all fields. The outdated server will be replaced by a state-of-the-art machine soon. We will present improvements of the Oracle configuration, the optimization of SQL statements, and the performance tuning of database design by adding new indexes which proved directly visible in accelerator operation, while data integrity was improved by additional foreign key constraints.

THE HIT MEDICAL ACCELERATOR

The heavy ion accelerator at HIT is used for raster-scanning radiation of cancer patients (cf. [1, 2] for an overview) with different types of ions from two sources (upgrade to three sources in progress [3,4]) in several treatment rooms, two with horizontal fixed beam exit and the heavy ion gantry with rotatable beam exit (under comissioning [5]), and a beam exit for experiments (see Fig. 1). Each combination of source and destination may be used for medical treatment, represented within the Accelerator Control System (ACS) by the so-callled *virtual accelerator* number. A radiation plan consists of a series of beam pulses chosen from a catalogue of 255 different energy values (88–430 MeV/u for carbon, 48–220 MeV/u for protons), 6 focus sizes, 15 intensity values (2×10^6–5×10^8 particles per second for carbon, 8×10^7–2×10^{10} pps for protons), and 36 exit angles in case of the gantry. These tuples of beam settings are named the MEFI combinations. Both the virtual accelerator as the MEFI combination may be changed from beam pulse to beam pulse (multi-plexed operation).

* klaus.hoeppner@med.uni-heidelberg.de

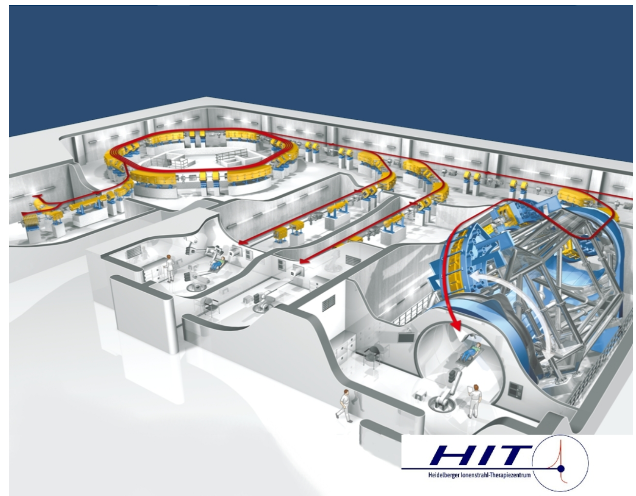

Figure 1: HIT accelerator facility with two ion sources, linear accelerator, synchroton, two horizontal beam exits and gantry for medical treatment. The experimental area is not shown.

DATABASE AS PART OF THE ACCELERATOR CONTROL SYSTEM

The database is a core part of the accelerator control system. It is used for

- a list of executed beam cycles for about 6 weeks together with measured values for e. g. currents of power supplies, data beam instrumentation (like ionization chambers), vacuum pressure, . . .
- list of devices that are active for a virtual accelerator,
- settings for all accelerator devices to be written into the device controllers,
- alarm and log messages,

- users and access rights,
- computers and applications.

The device settings are generated as binary large objects (BLOBs) by the DSM, the data supply module, from the physical beam properties optimized for a representative number of MEFI values and interpolation over the full MEFI range. Finally, the device blob per MEFI is stored within the Oracle database.

There exist three kinds of device blobs to combine the need for adjusting the accelerator settings and test of different beam optics with the quality requirements for stable settings needed for medical treatment:

- flash blobs that are stored in non-volatile memory of the device controllers after verification for therapy,
- online blobs stored in the RAM of the device controllers that are changed during accelerator adjustment (without affecting the verified flash data), and
- offline blobs to save old device data, present in the database, only.

The checksums of the blobs stored in the device controllers are checked against the blob checksums stored in database on any beam request to ensure the integrity of the settings.

DATABASE PROBLEMS

Since about 2010 the existing Oracle database (Oracle 9i running on a server with a dual core Xeon CPU and 2 GB of RAM under Windows 2003) showed performance problems. This was both visible during therapy (when checking the blob checksums, writing measured values) as during accelerator adjustment (when writing generated online blobs into database, copying online blobs into flash blobs after verification).

Obviously, a major reason was the outdated hardware, that was upgraded to a modern blade center during summer 2011 (as presented in [6]), but also issues beyond hardware should be addressed. Thus, we established a working group consisting of HIT controls group, Eckelmann AG (as supplier of the proprietary control system) and the database experts from the accelerator controls department of the GSI Helmholzzentrum für Schwerionenforschung in Darmstadt, Germany.

The workgroup analyzed both the setup of the current Oracle installation (e. g. by statspack[1] reports of usual operating tasks) as well as the structural design of the database (tables, indexes, constraints).

RESULTS, RECOMMENDATIONS, AND REALIZATION

Hardware and Oracle Setup

As expected, both the CPU speed and the memory of the database server were considered insufficient. It was recommended to upgrade the Oracle installation to version 11g and use Oracle's Automatic Storage Manager (ASM) as file system which is optimized for the data storage of an Oracle instance. RMAN[2] should be used as backup tool instead of the previous daily tablespace exports.

With the installation of the new blade servers during the summer shutdown period, these proposals were put into practise.

Number of Commits

The statspack analysis showed a high number of about 20 commits per second. This is caused by several data gateway processes that insert measured values from the devices into the database. In the current realization of the control system, a commit command is sent per insert statement. It was proposed to decrease the number of commits per second.

Eckelmann AG was checking the impact of grouping several insert statements per commit on the overall performance of the ACS. Obviously, the longer the delay between insert and commit, the slower the GUIs will present state changes and measured values to the operator. Thus, a compromise had to be found. Additionally, since several applications are writing into the database, the effort to implement and test a better commit mechanism was rather high.

Finally, the changes will be installed at HIT in October.

Structure of Tables

Many requirements for accelerator operation changed until or even since the accelerator facility became operational for medical treatment, inspired by the lessons learnt during commissioning. Thus, there was an evolution of the ACS from its original design until its current state. The database design reflects this evolution.

It was the overall impression that the database design for rather static data (like definition of ion types, beamlines, list of devices and their assignment to beamlines) was good, with appropriate relations between tables defined by foreign key constraints. But in some cases, data columns with dynamic, volatile data were added into these "static" tables due to changing or adding new features for accelerator controls. Tables for dynamic data (beam cycle data, measured values, messages) were often lacking foreign key constraints.

The recommendations for an improvement of the database design were as follows:

- Separate tablespaces for static (configuration) data and dynamic (measured) data,
- avoid to mix columns with static and volatile data in tables,
- define proper relations between tables by adding foreign key constraints where missing,

[1] statspack is a performance monitoring and reporting tool provided by Oracle starting from version 8i, cf. `http://www.oracle.com` or `http://www.orafaq.com/wiki/Statspack`.

[2] RMAN is a tool provided by Oracle to backup, restore and recover Oracle databases, cf. `http://www.oracle.com`

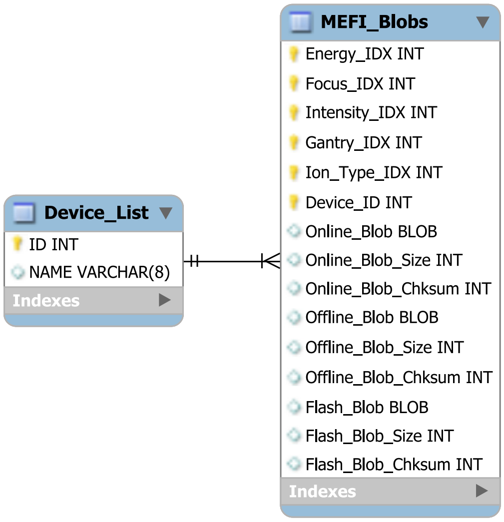

Figure 2: Current design of the table with MEFI device blobs, with three blobs per record.

- add indexes to optimize database queries,
- redesign "wide" tables to achieve a better in-memory-caching of data records within the Oracle database instance.

A good example for the last item is the table of device blobs per MEFI combination. As mentioned above, there exist online, offline and flash blobs. In the current table design, the table for MEFI device blobs has three blob columns (as shown in Fig. 2), one per blob type, i. e. when loading the data record for a MEFI combination and a device, three blobs are read and have to be cached in memory. Consider the workflow during accelerator adjustment, when online blobs are generated and written into database for all MEFI combinations and all devices. For downloading the new online blobs into the RAM of the device controllers, many records from the table with MEFI device blobs have to be read, and though only the online blobs are needed, the offline and flash blobs are read as well, consuming I/O ressources and memory of the database cache.

It was recommended to keep the blobs itself in an own table and replace the blob columns in the MEFI blob table by references. This would increase the number of tables itself, but keep the MEFI blob table compact. By appropriate joins between the MEFI blob table and the table(s) keeping the "real" blobs, the set of blobs that are read by Oracle would be reduced to those that are really needed. Figure 3 shows a possible redesign, where the blob columns are replaced by references to another table. Since changes to the table structure have impacts on the whole ACS, a redesign needs to be thorougly planned and tested. This process will be done until the end of this year.

Other recommendations listed above are already implemented or will be installed during the next ACS update in October 2011 (like adding foreign key constraints). The separation of static and volatile data in different tablespaces was realized during the migration from Oracle 9i to 11g.

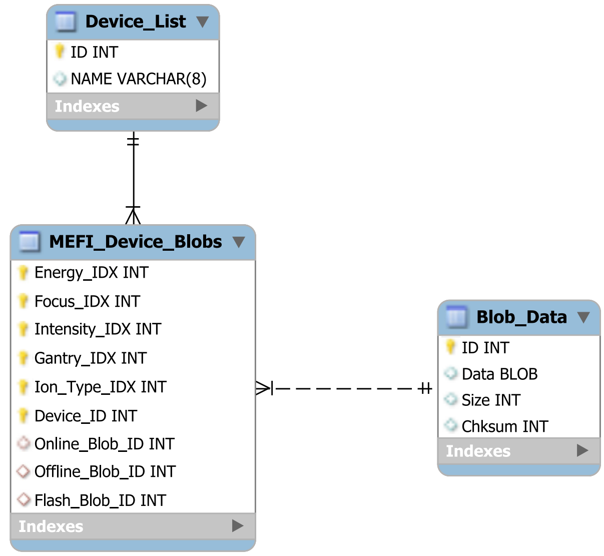

Figure 3: Possible change of table structure for MEFI device blobs by replacing the blob columns with a reference to another table keeping the blobs.

Several missing indexes were pointed out by the working group, that were added in the meantime, and some indexes were reorganized to achieve unique indexes as unique index scans are more efficient than normal index scans.

FIRST RESULTS

In combination with the new hardware platform, the changes to the Oracle database show already a positive performance effect.

Figure 4 compares the time between the stop and start broadcast signals (as sent by the central timing system), i. e. the dead time between beam cycles needed to prepare and setup the next cycle, for a week of therapy in September 2010 and September 2011, respectively. While the time

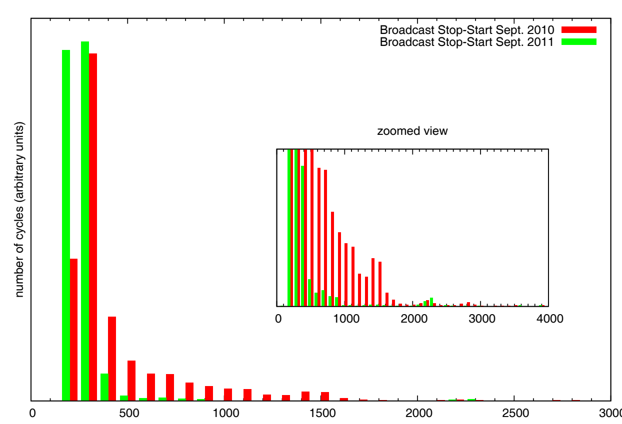

Figure 4: Comparison of dead time between beam cycles needed for preparation of the next beam pulse: September 2010 (red) and September 2011 (green).

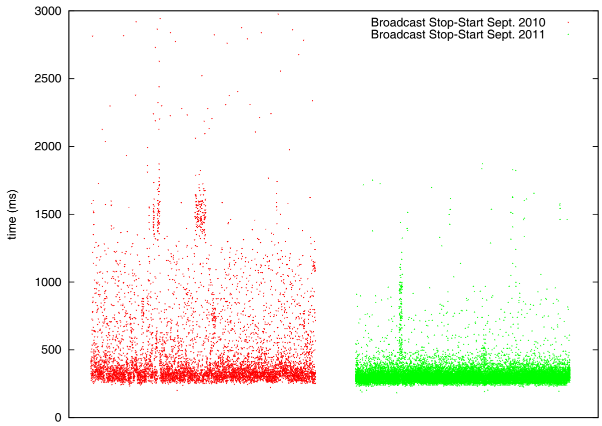

Figure 5: Scatter plot of dead time between beam cycles needed for preparation of the next beam pulse: September 2010 (red) and September 2011 (green).

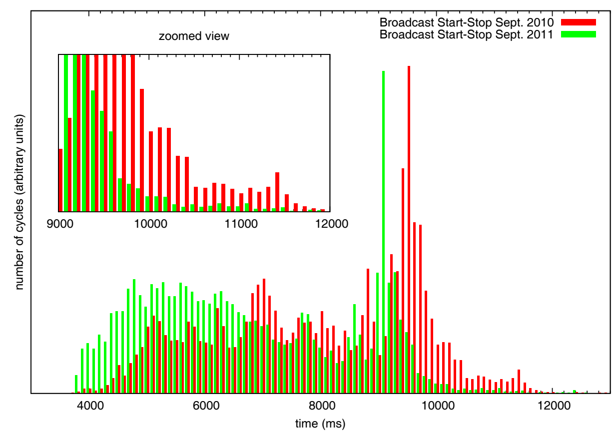

Figure 6: Comparison of beam cycle time (for up to 5 s tumor radiation): September 2010 (red) and September 2011 (green).

needed to prepare the next beam cycle after the end of the previous cycle is about 250 ms for most of the cycles, the average dead time was spoiled by many cycles with a much longer preparation time before the start broadcast signal was sent. The analysis of traces of servers and device controllers showed, that this effect was mainly caused by

- a delay in the Oracle database when creating the database record for the upcoming cycle and querying the list of devices that are active for the cycle,
- a read delay in the flash memory of the device controllers when checking the MEFI blob checksums.

After both problems were solved this year, it is immediately visible from the plot that most delayed cycles were eliminated, with the effect that the average preparation time was decreased from 490 ms in September 2010 to 390 ms in September 2011. The scatter plot in Fig. 5 also shows, that the number of cycles with preparation time above about 400 ms is decreased dramatically.

We compared also the times between the broadcast signals for cycle start and stop during therapy (based on medical treatment with up to 5 s radiation of the tumor per beam cycle) for September 2010 and 2011. As seen in Fig. 6 the cycle performance increased a lot, caused both by improvements in the therapy control system (TCS) and the communication between ACS and TCS, as well as the performance tuning of the database. The analysis of the traces showed that the time needed for updating the cycle status in the database was sometimes up to 200 ms.

CONCLUSION

As presented in the previous section, the accelerator performance was increased a lot during the last year. This is due to many efforts to optimize the ACS, both by hardware as well as by improvements to the database.

The improvements are also visible in GUIs for the operator that profit from better database querying performance, and time consuming tasks during accelerator adjustments, like interpolating and generating MEFI blobs or flashing are much faster than before. E. g., an interpolation of all devices for the gantry needs about 1.5 hours per ion type, instead of 6 hrs. before.

The recommendations didn't just address database performance, but data integrity, too. The realization of proper relations between the tables, as it is currently in progress or preparation, will help to avoid both redundant or even invalid data within the database.

REFERENCES

[1] Th. Haberer et al., "The Heidelberg Ion Therapy Center", Rad. & Onc., 73 (Suppl. 2), 186–190, (2004).

[2] S. Combs et al., "Particle therapy at the Heidelberg Ion Therapy Center (HIT)", Radiotherapy and Oncology, Vol. 95, 41–44, (2010).

[3] A. Peters et al., "Operational Status and Further Enhancements of the HIT Accelerator Facility", IPAC 2010, Kyoto, Japan, May 2010, MOPEA006, http://www.jacow.org.

[4] R. Cee et al., "Status of the Ion Source and RFQ Test Bench at the Heidelberg Ion Beam Therapy Centre", IPAC 2011, San Sebastián, Spain, September 2011, WEPS044, http://www.jacow.org.

[5] M. Galonska et al., "Commissioning of the Ion Beam Gantry at HIT", IPAC 2011, San Sebastián, Spain, September 2011, THOAB03, http://www.jacow.org.

[6] J. M. Mosthaf et al., "Upgrade of the Server Architecture for the Accelerator Control System at the Heidelberg Ion Therapy Center", ICALEPCS 2011, Grenoble, France, October 2011, MOMMU009, http://www.jacow.org.