

ECC: Error Correction Code or Error Correcting Code

The basis of all error detection and correction in hard disks is the inclusion of redundant information and special hardware or software to use it. Each sector of data on the hard disk contains 512 bytes, or 4,096 bits, of user data. In addition to these bits, an additional number of bits are added to each sector for the implementation of error correcting code or ECC (sometimes also called error correction code or error correcting circuits). These bits do not contain data; rather, they contain information about the data that can be used to correct any problems encountered trying to access the real data bits.

พื้นฐานของการตรวจจับข้อผิดพลาดและการแก้ไขข้อผิดพลาดในฮาร์ดดิสก์ ประกอบด้วยบิตข้อมูลที่เกินมาจะเป็นตัวบ่งบอก ใช้อุปกรณ์ฮาร์ดแวร์หรือซอฟต์แวร์พิเศษเข้ามาช่วยในการทำงาน ในแต่ละบล็อกข้อมูล (Sector data) ของฮาร์ดดิสก์บรรจุได้ 512 ไบต์ หรือ 4,096 บิต ไว้สำหรับให้ผู้ใช้ทำการบันทึกข้อมูล บิตที่เพิ่มเข้าเหล่านี้ เป็นการเพิ่มจำนวนบิตที่ได้ถูกเพิ่มเข้าไปในแต่ละบล็อกข้อมูล (Sector data) สำหรับช่วยสนับสนุนการทำงานของรหัสแก้ไขข้อผิดพลาด ECC (Error correction code) (บางครั้งก็อาจจะเรียกว่า Error correction code หรือ Error correcting circuits) โดยบิตรหัสแก้ไขข้อผิดพลาดเหล่านี้ไม่ได้อยู่ในข้อมูล แต่จะอยู่ข้างๆ นั่นเอง บิตรหัสแก้ไขข้อผิดพลาดจะเป็นตัวช่วยบอกว่าข้อมูลที่ได้อาจสามารถนำไปใช้ได้ หรือ ถูกต้องไหม และถ้าเกิดข้อผิดพลาดขึ้นก็จะแก้ไขเพื่อที่จะให้ได้ข้อมูลที่แท้จริง

There are several different types of error correcting codes that have been invented over the years, but the type commonly used on PCs is the Reed-Solomon algorithm, named for researchers Irving Reed and Gustave Solomon, who first discovered the general technique that the algorithm employs. Reed-Solomon codes are widely used for error detection and correction in various computing and communications media, including magnetic storage, optical storage, high-speed modems, and data transmission channels. They have been chosen because they are easier to decode than most other similar codes, can detect (and correct) large numbers of missing bits of data, and require the least number of extra ECC bits for a given number of data bits.

รหัสแก้ไขข้อผิดพลาดมีมากมายที่แตกต่างกันออกไปตามชนิดของรหัสแก้ไขข้อผิดพลาดที่ได้ถูกสร้างขึ้นในหลายปีที่ผ่านมา แต่ชนิดที่เป็นที่นิยมใช้ในคอมพิวเตอร์คือวิธีการของ Reed-Solomon สร้างขึ้นโดยสองนักวิจัยคือ Irving Reed และ Gustave Solomon เป็นคนแรกที่ทำให้คิดค้นวิธีการนี้และได้นำมาใช้ รหัส Reed-Solomon ได้ถูกนำมาใช้อย่างแพร่หลายในการตรวจสอบข้อผิดพลาดและการแก้ไขข้อผิดพลาดของ

ข้อมูล ในการคำนวณหลาย ๆ ด้าน และด้านเป็นตัวกลางการสื่อสาร รวมทั้งการบันทึกข้อมูลด้วยแม่เหล็ก การบันทึกข้อมูลทางแสง ในโมเดมความเร็วสูง และการส่งข้อมูลผ่านช่องสัญญาณ ก็ได้มีการเลือกที่จะใช้ การเข้ารหัสแบบ Reed-Solomon เพราะว่ามีคามง่ายในการเข้ารหัสกว่ารหัสแบบอื่นๆ มาก สามารถ ตรวจจับข้อผิดพลาด และทำการแก้ไขข้อผิดพลาด ที่จำนวนบิตของข้อมูลมากๆ ได้ ต้องการจำนวนบิต ECC น้อยๆ ก็พอสำหรับจำนวนบิตของข้อมูลได้

When a sector is written to the hard disk, the appropriate ECC codes are generated and stored in the bits reserved for them. When the sector is read back, the user data read, combined with the ECC bits, can tell the controller if any errors occurred during the read. Errors that can be corrected using the redundant information are corrected before passing the data to the rest of the system. The system can also tell when there is too much damage to the data to correct, and will issue an error notification in that event. The sophisticated firmware present in all modern drives uses ECC as part of its overall error management protocols. This is all done "on the fly" with no intervention from the user required, and no slowdown in performance even when errors are encountered and must be corrected.

เมื่อทำการเขียนข้อมูลลงบนฮาร์ดดิสก์ ก็ถือได้ว่ารหัส ECC ได้ถูกสร้างขึ้นด้วยเช่นกันซึ่งจะถูก บันทึกอยู่ในส่วนพื้นที่ที่เป็นบิตที่สงวนไว้สำหรับรหัส ECC หลังจากนั้นเมื่อทำการอ่านข้อมูลจากฮาร์ดดิสก์ ก็จะได้ข้อมูลร่วมกับส่วนที่เป็น ECC บิตด้วย สามารถบอกตัวควบคุมได้ว่าถ้าเกิดการผิดพลาดในระหว่าง การอ่านข้อมูล ข้อผิดพลาดที่เกิดขึ้นนั้นสามารถแก้ไขให้ถูกต้องได้โดยใช้บิตข้อมูลที่เกินมาช่วยในการแก้ไข ก่อนที่จะทำการส่งข้อมูลไปยังส่วนอื่นๆ ของระบบ ระบบก็เช่นเดียวกันสามารถบอกได้ว่าเมื่อเกิดบิต ผิดพลาดมากๆ ก็ให้ทำการแก้ไขให้ถูกต้อง โดยจะบอกให้ทราบก่อนที่จะเกิดการผิดพลาดขึ้นได้ ปัจจุบันมี ตัวฮาร์ดแวร์แบบใหม่ที่ใช้กับฮาร์ดดิสก์รุ่นใหม่ๆ จะทำการจัดการกับข้อผิดพลาด ได้ทั้งหมด และกำลัง พัฒนาไปเป็นแบบไม่ให้อยู่ระหว่างที่ผู้ใช้ต้องการ สมรรถนะการทำงานก็ไม่ได้ลดลงเมื่อเกิดข้อผิดพลาดขึ้น และถ้ามีก็แก้ไขให้ถูกต้องได้

The capability of a Reed Solomon ECC implementation is based on the number of additional ECC bits it includes. The more bits that are included for a given amount of data, the more errors that can be tolerated. There are multiple tradeoffs involved in deciding how many bits of ECC information to use. Including more bits per sector of data allows for more robust error detection and correction, but means

fewer sectors can be put on each track, since more of the linear distance of the track is used up with non-data bits. On the other hand, if you make the system more capable of detecting and correcting errors, you make it possible to increase areal density or make other performance improvements, which could pay back the "investment" of extra ECC bits, and then some. Another complicating factor is that the more ECC bits included, the more processing power the controller must possess to process the Reed Solomon algorithm. The engineers who design hard disks take these various factors into account in deciding how many ECC bits to include for each sector.

ความสามารถในการแก้ไขข้อผิดพลาดของข้อมูลด้วยรหัส Reed-Solomon ECC ขึ้นอยู่กับจำนวนของบิต ECC ที่ได้เพิ่มหรือเขียนลงไป ซึ่งจำนวนบิต ECC ที่เพิ่มหรือเขียนลงไปมีหลายบิตก็เท่ากับมีจำนวนข้อมูลมากขึ้นตามด้วย และขึ้นอยู่กับว่าจะสามารถยอมรับข้อผิดพลาดที่เกิดขึ้นได้มากเท่าไร นั่นก็คือทำการแบ่งออกเป็นหลายๆ ส่วนเพื่อเป็นส่วนช่วยในการตัดสินใจเลือกใช้ ECC ขนาดก็บิต จำนวนบิต ECC ต่อหนึ่งบล็อกข้อมูล (Sector data) มีขนาดมากก็ทำให้ข้อมูลที่ได้มีความสมบูรณ์ดี และถ้าเกิดข้อผิดพลาดขึ้นก็สามารถที่จะตรวจจับได้แล้วทำการแก้ไขข้อผิดพลาดนั้นได้ แต่ก็หมายความว่าจะสามารถใส่จำนวนบิตบล็อกข้อมูลได้สองถึงสามบล็อกในแต่ละแทร็ก (Track) พบว่าแทร็กจะเพิ่มขึ้นเป็นเชิงเส้นกับระยะทางเมื่อไม่มีบิตของข้อมูล ถ้าจะทำให้ระบบมีความสามารถเพิ่มขึ้นโดยสามารถตรวจจับและแก้ไขข้อผิดพลาดได้ดี ก็ทำได้โดยการเพิ่มความจุต่อหนึ่งหน่วยพื้นที่ หรือทำการปรับปรุงสมรรถนะการทำงานตัวอื่นๆ ซึ่งก็จะทำให้ได้การลงทุนที่คุ้มค่ากับการเพิ่มบิต ECC เป็นต้น ตัวพารามิเตอร์ตัวอื่นๆ ที่จะสามารถทำให้การเพิ่มบิต ECC ได้ก็อย่างเช่น กำลังงานที่ให้ในระบบตัวควบคุมของวิธีการ Reed-Solomon วิศวกรคนไหนที่ทำหน้าที่ในการออกแบบฮาร์ดดิสก์ต้องมีข้อมูลตัวแปรหรือพารามิเตอร์ในการช่วยตัดสินใจว่าจะใช้จำนวน ECC ขนาดกี่บิตที่จะเขียนหรือบันทึกลงไปในแต่ละบล็อกข้อมูล (Sector data)

แปลและเรียบเรียงโดย
อดิศร.

Reference

<http://www.storagereview.com>