# Dữ liệu

Trong máy tính, các dữ liệu nhị phân không được xử lý theo từng bit riêng lẻ, mà được xử lý thành từng khối 8 bit một, và đơn vị xử lý nhỏ nhất này gọi là byte.

Ví dụ, số nguyên 123456789 được biểu diễn dưới dạng nhị phân sẽ là:

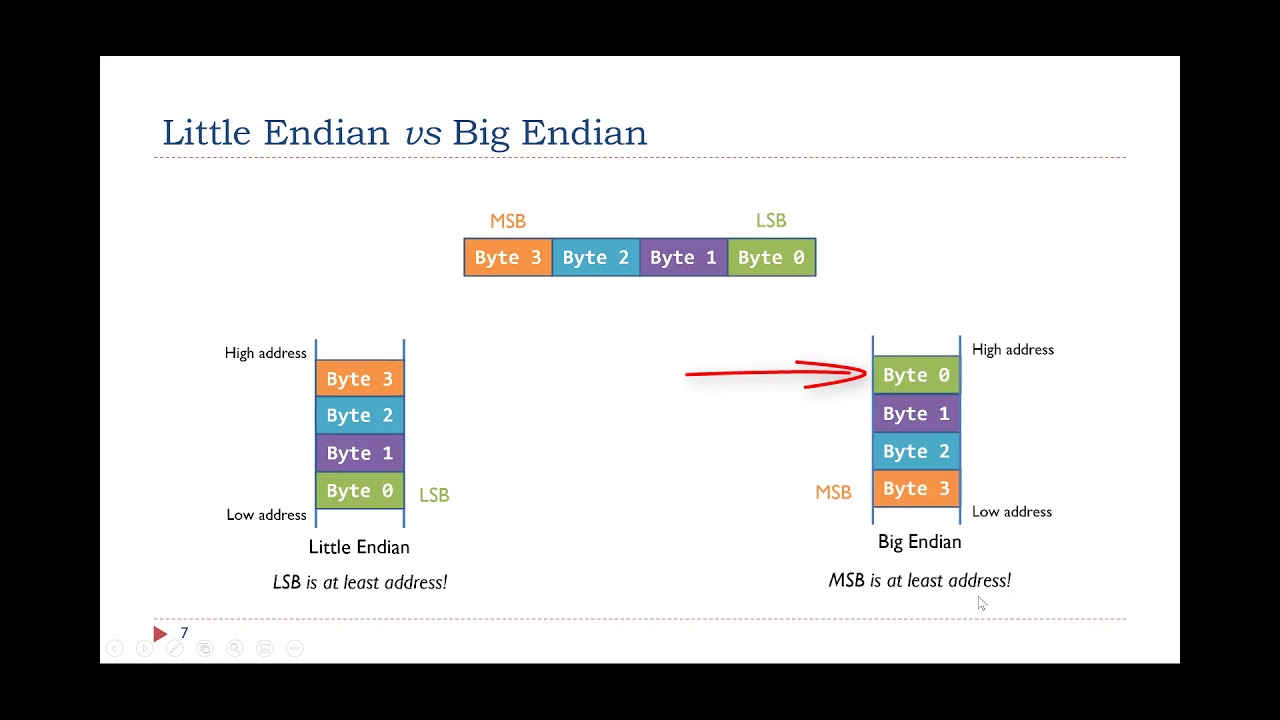
00000111 01011011 11001101 00010101

Để ngắn gọn, chúng ta có thể viết nó dưới dạng hexa như sau:

07 5b cd 15

**Để ghi dữ liệu này lên đĩa cứng, nó được ghi như thế nào?** Việc ghi dữ liệu như thế nào chính là lúc little endian và big endian được dùng đến.

# Little endian và big endian



Little endian và big endian là hai phương thức khác nhau để lưu trữ dữ liệu.

**Sự khác biệt của little endian và big endian khi lưu trữ** chính là ở **việc sắp xếp thứ tự các byte dữ liệu**.

* Little endian: byte cuối cùng trong biểu diễn nhị phân sẽ được ghi trước
* Big endian: cơ chế ghi dữ liệu theo thứ tự bình thường mà chúng ta vẫn dùng

# Dùng little endian và big endian trên các máy tính nào

Mọi hoạt động của máy tính đều sử dụng dữ liệu nhị phân, nên little endian/big endian hiện hữu trong mọi hoạt động của máy tính.

Sử dụng cơ chế lưu trữ phụ thuộc vào phần mềm (do dev quy định, hoặc do ngôn ngữ lập trình quy định trước) và bộ xử lý của máy tính đó.

* Intel: little endian
* ARM: bi-endian

# Cơ chế lưu trữ nào tốt hơn

Little endian hay big endian chỉ khác nhau ở việc lưu trữ thứ tự các byte dữ liệu. Cả hai phương thức đều không làm ảnh hưởng đến tốc độ xử lý của CPU.

Mỗi phương thức đều có những lợi thế nhất định:

* Little endian: vì byte nhỏ nhất luôn nằm bên trái nên nó cho phép chúng ta đọc dữ liệu với độ dài tùy ý. Nó thích hợp nếu chúng ta cần ép kiểu.
* Big endian: việc đọc dữ liệu byte lớn nhất trước, nó sẽ rất dễ dàng kiểm tra một số là âm hay dương, do byte chứa dấu được đọc đầu tiên.

# Các cơ chế lưu trữ này ảnh hưởng như thế nào đến việc lập trình

Về cơ bản thì little endian hay big endian không có ảnh hưởng lắm đến việc lập trình. Phần lớn các lập trình viên không cần quan tâm nhiều lắm, bởi mọi việc đã được các trình biên dịch/thông dich đảm nhiệm hết.

Tuy nhiên, một số trường hợp, chúng ta cần quan tâm, đặc biệt khi chuyển đổi dữ liệu giữa các máy tính khác nhau.

Ví dụ:

1. Ép kiểu cho các biến

A picture containing text, screenshot, font

Description automatically generated

Máy tính dùng cơ chế big endian sẽ cho ra kết quả là 256, còn với little endian kết quả sẽ là 1.

1. Trong trường hợp này, khi nhận file được tạo ra từ một máy tính khác, việc nó được ghi theo kiểu little endian hay big endian rõ ràng là ảnh hưởng rất nghiêm trọng, nếu sử dụng sai phương thức, chúng ta sẽ thu về dữ liệu sai.

# Vấn đề NUXI trong việc trao đổi dữ liệu giữa các máy có endian khác nhau