The AES in a nutshell

G. Averkov
Brandenburg University of Technology

July 16, 2024

High-level idea

The AES is done over the field $F = \mathbb{F}_{256}$, realized as $F = \mathbb{Z}_2/g$ with $g = t^8 + t^4 + t^3 + t + 1$. The field F is an 8-dimensional vector space over $\mathbb{Z}_2$. So, F can be identified with $\mathbb{Z}_2^8$. The AES encrypts a list of 16 elements of F. So, we can view the encryption map as a map $E_k : F^{16} \to F^{16}$ but also, in view of the identification, as a map $E_k : \mathbb{Z}_2^{128} \to \mathbb{Z}_2^{128}$.

$F_{256} = \mathbb{Z}_2[t]/q$ with

$q = t^8 + t^4 + t^3 + t + 1$ (irreducible)

$x \in F_{256}$ can be written

as $x = \sum_{i=0}^{7} x_i \alpha^i$, where $\alpha = [t]_q$.

In other words, α is a formal element satisfying $\alpha^8 + \alpha^4 + \alpha^3 + \alpha + 1 = 0$

$x \in F_{256}$ (here) 8 bits of information

$x_0, \dots, x_7 \in \mathbb{Z}_2$.

$F := F_{256}$ ← just as a shorthand notation.

The AES encrypts 16 elements of F. If we interpret this in bits: the AES encrypts $16 \cdot 8 = 128$ bits (elements of $\mathbb{Z}_2$).

$$F^{16} \longrightarrow F^{16}$$

The incentive behind AES is to provide a map $E_k(x)$ that is highly non-linear in both k and x. Remember that linearity and non-linearity is introduced with respect to a field: so, in our case, we could mean it in the sense of the field $\mathbb{Z}_2$ and in the sense of the field F. The AES encoding map is built up of a sequence of operations that linear/affine and non-linear. Alternating between linear and non-linear operations, within the AES, one makes a complicated non-linear map.

- Difficult non-linear operations on small chunks of data
- Easier calculations on larger chunks of data.

The non-linear function on which AES relies

One has $x^{|F|} = x$ for every $x \in F$, which follows from Lagrange's theorem. So, raising to the power $|F|$ is a linear map. Thus raising to the power $x^{|F|-1}$ is a map that is equal to 1 on $x \in F \setminus \{0\}$ and equal to 0 on $x = 0$. This one is not invertible and not complicated (so, no good for the encryption). But raising to the power $|F| - 2$ is pretty much the same as the inversion of x (it gives x^{-1} for $x \neq 0$ and 0 for $x = 0$). Raising to the power $|F| - 2$ can be implemented efficiently (because the inversion in F can be implemented efficiently). The map is not only highly non-linear with respect to F, but also with respect to $\mathbb{Z}_2$.

Theorem 32 (A generalization of Fermat's Little Theorem)

Let F be a finite field. Then

$$a^{|F|-1} = 1 \quad \text{for every} \quad a \in F \setminus \{0\}.$$

Proof: Essentially, we repeat the proof of Fermat's Little Theorem. Let's fix $a \in F \setminus \{0\}$.

The function $f : F \setminus \{0\} \to F \setminus \{0\}$ given by $f(x) = a \cdot x$ is well-defined and bijective.

Well-defined, because one indeed has $a \cdot x \neq 0$ for every $x \in F \setminus \{0\}$ (since in a field there are no zero divisors). It's clear that the function f is bijective with the inverse function $f^{-1} : F \setminus \{0\} \to F \setminus \{0\}$ given by $f^{-1}(y) = \bar{a}^{-1} \cdot y$.

$$\Rightarrow \quad \prod_{x \in F \setminus \{0\}} x = \prod_{x \in F \setminus \{0\}} f(x) = \prod_{x \in F \setminus \{0\}} (a \cdot x)$$

$$= a^{|F|-1} \cdot \prod_{x \in F \setminus \{0\}} x$$

$$\Rightarrow \quad \underbrace{\prod_{x \in F \setminus \{0\}} x}_{\neq 0} = a^{|F|-1} \cdot \underbrace{\prod_{x \in F \setminus \{0\}} x}_{\neq 0}$$

$$\Rightarrow \quad 1 = a^{|F|-1}. \qquad \square$$

Remark How does the function $x \mapsto x^{|F|-2}$ act?

$$x^{|F|-2} = \begin{cases} x^{-1}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

We can calculate the inverse efficiently in a finite field using the EEA over a polynomial ring.

Remark For $x = \sum_{i=0}^{7} x_i \alpha^i \in F = \mathbb{F}_{256}$

the operation $x^{|F|-2}$ is not only non-linear in $x \in F$ but also in the bits $x_0, \dots, x_7 \in \mathbb{Z}_2$.

Non-linearity of inversion in a field - Example 1

If we had $F = \mathbb{Z}_2[t]/(t^3 + t + 1)$ with α being the coset of t modulo $t^3 + t + 1$, then

$$x = x_0 + x_1\alpha + x_2\alpha^2 \leftrightarrow (x_0, x_1, x_2)$$

raised to the power $|F| - 2 = 6$ would be

$$\begin{aligned} x^6 &= x_0 + x_1 + x_2 + x_1x_2 + (x_0x_1 + x_2)a + (x_0x_2 + x_1 + x_2)a^2 \\ &\leftrightarrow (x_0 + x_1 + x_2 + x_1x_2,\ x_0x_1 + x_2,\ x_0x_2 + x_1 + x_2). \end{aligned}$$

The dependence on $x_0, x_1, x_2 \in \mathbb{Z}_2$ is non-linear. The dependence of $x^{|F|-2}$ on $x_0, \ldots, x_{n-1}$ gets even more non-linear when F the dimension n of F over $\mathbb{Z}_2$ gets even higher.

Example Smaller field $F = \mathbb{F}_8 = \mathbb{Z}_2[t]/(t^3+t+1)$

$\alpha = [t]_{t^3+t+1}$, which means

$$\alpha^3 + \alpha + 1 = 0$$

$$x = x_0 + x_1\alpha + x_2\alpha^2$$

$$x^2 = (x_0 + x_1\alpha + x_2\alpha^2)^2$$

$$= (x_0 + x_1\alpha + x_2\alpha^2)(x_0 + x_1\alpha + x_2\alpha^2)$$

$$= x_0 + x_1\alpha^2 + x_2\alpha^4$$

$$= x_0 + x_1\alpha^2 + x_2\alpha^3 \cdot \alpha$$

$$= x_0 + x_1\alpha^2 + x_2(\alpha+1)\cdot\alpha$$

$$= x_0 + x_1\alpha^2 + x_2(\alpha^2+\alpha)$$

$$= x_0 + x_2\alpha + (x_1+x_2)\alpha^2$$

$$\underset{x}{(x_0, x_1, x_2)} \longrightarrow \underset{x^2}{(x_0, x_2, x_1+x_2)}$$

$$x^3 = x^2 \cdot x$$

$$= (x_0 + x_2\alpha + (x_1 + x_2)\alpha^2) \cdot (x_0 + x_1\alpha + x_2\alpha^2)$$

$$= x_0 + (x_0x_2 + x_0x_1)\alpha + (\cancel{x_0x_2} + x_1x_2 + x_0x_1 + \cancel{x_0x_2})\alpha^2$$

$$+ (x_2 + x_1 + x_1x_2)\underbrace{\alpha^3}_{\alpha+1}$$

$$+ (x_1x_2 + x_2)\underbrace{\alpha^4}_{\alpha^2+\alpha}$$

$$= (x_0 + x_1 + x_2 + x_1x_2) + (x_0x_2 + x_0x_1 + \cancel{x_1x_2} + \cancel{x_1x_2} + \cancel{x_1x_2} + \cancel{x_2})\alpha$$

$$(x_0x_1 + \cancel{x_1x_2} + \cancel{x_1x_2} + x_2)\alpha^2$$

$$(x_0, x_1, x_2) \xrightarrow{x^3} (x_0^3 + x_1^3 + x_2^3 + x_1 x_2,$$
$$x_0 x_2 + x_0 x_1 + x_1,$$
$$x_0 x_1 + x_2)$$

Non-linearity of inversion in a field - Example 2

If we had $F = \mathbb{Z}_2[t]/(t^4 + t + 1)$ with α being the coset of t modulo $t^4 + t + 1$, then

$$
\begin{aligned}
& x^{14} = (x_0 + x_1\alpha + x_2\alpha^2 + x_3\alpha^3)^{14} \\
\leftrightarrow\ & (x_0x_1x_2 + x_1x_2x_3 + x_0x_2 + x_1x_2 + x_0 + x_1 + x_2 + x_3, \\
& x_0x_1x_3 + x_0x_1 + x_0x_2 + x_1x_2 + x_1x_3 + x_3, \\
& x_0x_2x_3 + x_0x_1 + x_0x_2 + x_0x_3 + x_2 + x_3, \\
& x_1x_2x_3 + x_0x_3 + x_1x_3 + x_2x_3 + x_1 + x_2 + x_3)
\end{aligned}
$$

The AES pipelines the initial data (the message), consisting of $8 \cdot 4 \cdot 4 = 128$ bits, through a sequence of transformations that jumble and reshuffle the data in each stage in a rather complicated way. The data is a 4×4 table (= a matrix), with each entry being an element of F, which corresponds to 8 bits of information. Imagine writing bits into the squares of the following table:

$$\begin{bmatrix} \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square \\ \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square \\ \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square \\ \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square & \square\square\square\square\square\square\square\square \end{bmatrix}$$

1. One type of the transformation is localized to the elements of F, and so it changes each of the 16 entries of the table independently of each other. So, this transformation is local, in the sense that it jumbles the 8-bit groups independently of each other. This transformation is highly non-linear.

1. One type of the transformation is localized to the elements of F, and so it changes each of the 16 entries of the table independently of each other. So, this transformation is local, in the sense that it jumbles the 8-bit groups independently of each other. This transformation is highly non-linear.
2. Another type of transformation reshuffles the bits within each of the 4 rows independently.

1. One type of the transformation is localized to the elements of F, and so it changes each of the 16 entries of the table independently of each other. So, this transformation is local, in the sense that it jumbles the 8-bit groups independently of each other. This transformation is highly non-linear.
2. Another type of transformation reshuffles the bits within each of the 4 rows independently.
3. Yet another transformation jumbles the bits within each of the 4 columns independently. This transformation is affine.

When all three types of transformations are applied successively, one makes sure that there is a lot of non-linearity in the process of encryption and that each of the 128 bits of the plain text affects each of the 128 bits of the cipher text and, vice versa, that each fo the 128 of the cipher text is affected in a complicated manner by each fo the 128 bits of the plain text.

It's like doing a fondue. Fondue is a liquid cheese in a pot made in Switzerland. So, the AES is a fondue done out of 128 bits. When you do the fondue, you stir the cheese in a pot. Our pot is consisting of 4×4 spots, each having 8 bits. There are several ways to stir. You can stir each of the 16 8-bit groups. You could stir row-wise, each of the 4 rows, and you can stir column-wise, each of the 4 columns. The AES encryption of a message is prepared by combining all these ways to stir.

The key of AES is $k \in F^{4\times 4}$ and, so it is also given by 128 bits. The AES has a procedure (to be described below) that associates to k a sequence of the so-called round keys $k^{(0)}, \ldots, k^{(10)} \in F^{4\times 4}$, calculated iteratively starting with $k^{(0)} = k$.

The plain text message $M = (m_{i,j})_{i,j=0,\ldots,3} \in F^{4\times 4}$ undergoes a sequence of transformations, carried out iteratively and producing a sequence $M^{(0)}, \ldots, M^{(10)} \in F^{4\times 4}$ with the final result $M^{(10)}$ being the cyphertext for M.

SubBytes

We introduce the transformation *SubBytes* SB : $F^{4\times 4} \to F^{4\times 4}$ that acts on $M = (m_{i,j})_{i,j=0,\dots,3} \in F^{4\times 4}$ by

$$\mathrm{SB}(M) = \left(T(m_{i,j}{}^{254}) + 1 + \alpha + \alpha^5 + \alpha^6 \right)_{i,j=0,\dots,3},$$

where $T : F \to F$ is the $\mathbb{Z}_2$-linear transformation

$$T(x_0\alpha + \cdots + x_7\alpha^7) = y_0\alpha + \cdots + y_7\alpha^7$$

given by... $x_0, \dots, x_7 \in \mathbb{Z}_2$ $y_0, \dots, y_7 \in \mathbb{Z}_2$

SubBytes matrix

$$y_2 = x_0 + x_1 + x_2 + x_6 + x_7$$
$$y_5 = x_1 + x_2 + x_3 + x_4 + x_5$$
$$y_7 = x_3 + x_4 + x_5 + x_6 + x_7$$

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}$$

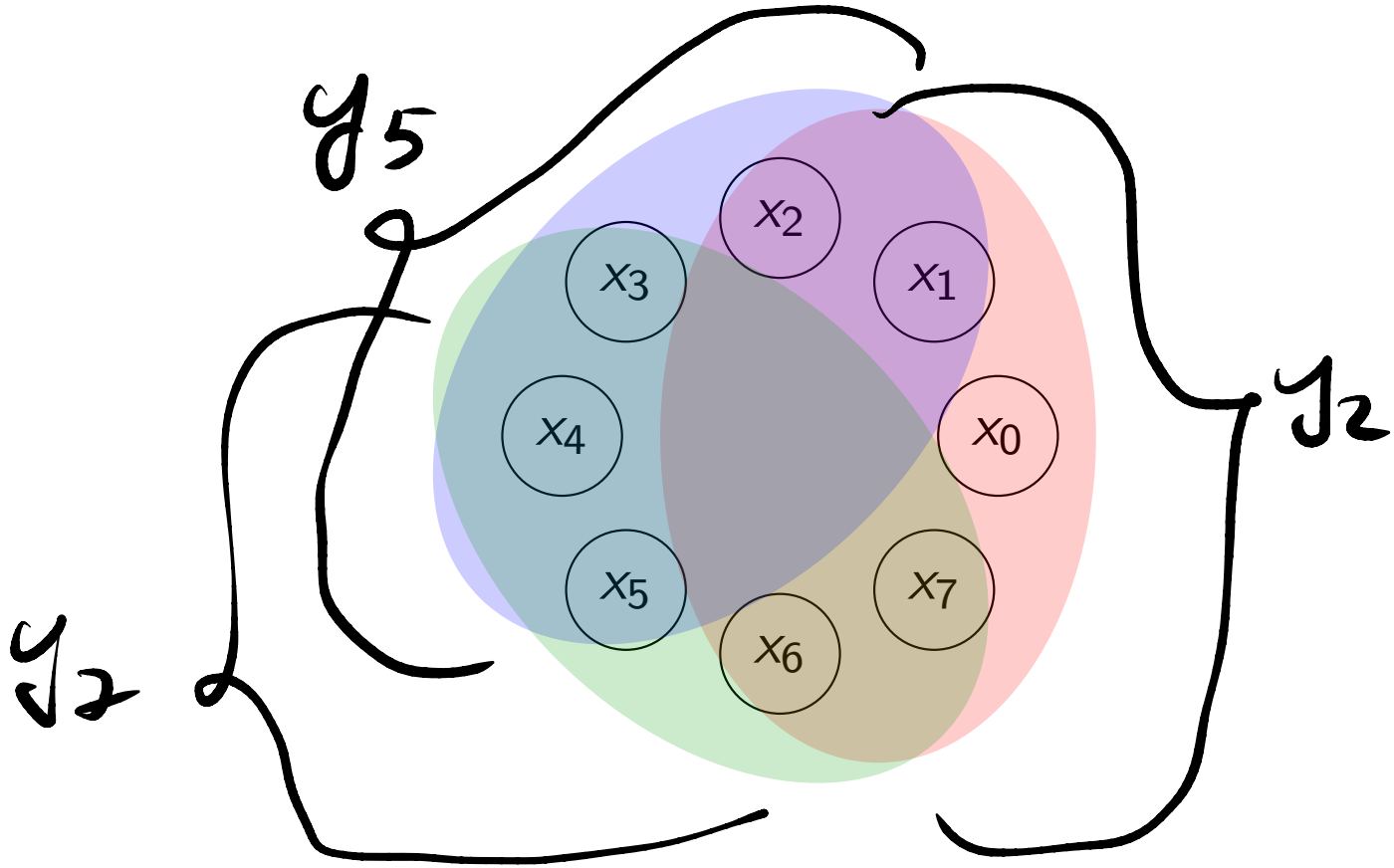
$$y_0 = x_0 + x_4 + x_5 + x_6 + x_7$$
$$y_1 = x_0 + x_1 + x_5 + x_6 + x_7$$
$$\vdots$$

SubBytes matrix, continued

Equivalently, the dependence of y on x is

$$y_i = x_i + x_{i-1} + x_{i-2} + x_{i-3} + x_{i-4},$$

with the indexing modulo 8.
SB is an invertible transformation, because raising to the power 254 is invertible, and there is a one-to-one correspondence $x \leftrightarrow y$, as ...



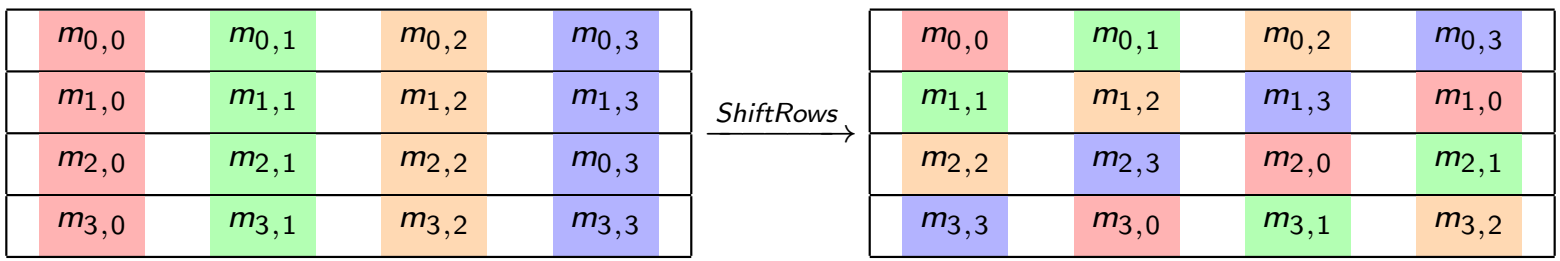
$$x_0 = y_2 + y_5 + y_7$$

Generally,

$$x_i = y_{i+2} + y_{i+5} + y_{i+7}.$$

ShiftRows

$$(m_{i,j})_{i,j=0,\ldots,3} \xrightarrow{ShiftRows} (m_{i,j-i})_{i,j=0,\ldots,3}.$$



MixColumns

$M = (m_{i,j})_{i,j=0,1,2,3}$

$M \longmapsto T \cdot M$

$$M \xrightarrow{MixColumns} TM$$

with

$$T = \begin{bmatrix} \alpha & 1+\alpha & 1 & 1 \\ 1 & \alpha & 1+\alpha & 1 \\ 1 & 1 & \alpha & 1+\alpha \\ 1+\alpha & 1 & 1 & \alpha \end{bmatrix}.$$

MixColumns transforms each column of M independently. Let's see how this transformation works.

Example: $F = \mathbb{F}_4 = \mathbb{Z}_2[t]/(t^2+t+1)$

$\alpha = [t]_{t^2+t+1}$

$\alpha^2 + \alpha + 1 = 0$

$f(x) = c \cdot x$

$c = 1 + \alpha$

$x \in F \qquad x = x_0 \cdot 1 + x_1 \cdot \alpha$

$1, \alpha$

$\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$

$\begin{bmatrix}x_0\\x_1\end{bmatrix} = x_0\begin{bmatrix}1\\0\end{bmatrix} + x_1\begin{bmatrix}0\\1\end{bmatrix}$

$$T = \begin{bmatrix} t_{00} & t_{01} \\ t_{10} & t_{11} \end{bmatrix} \qquad T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} t_{00} \\ t_{10} \end{bmatrix}$$

$$T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} t_{01} \\ t_{11} \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \longleftrightarrow 1 \longmapsto c \cdot 1 = 1 + \alpha \longleftrightarrow \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \longleftrightarrow \alpha \longmapsto c \cdot \alpha = (1+\alpha)\cdot\alpha = 1 \longleftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$cx \longleftrightarrow T \cdot \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

You can do this also for quotient rings.

Interpretation of MixColumns

$$f := m_0 + m_1 u + m_2 u^2 + m_3 u^3 \in F[u]$$

$$\mapsto$$

$$(\alpha + u + u^2 + (1+\alpha)u^3) f \bmod (u^4 + 1).$$

This corresponds to a linear transformation on the ring $G = F[u]/(u^4 + 1)$. tt is invertible, because $(u^4 + 1) = (u + 1)^4$ in $F[u]$. This shows that $a(u) := \alpha + u + u^2 + (1 + \alpha)u^3$ and $u^4 + 1$ are relatively prime because $a(1) = \alpha + 1 + 1 + (1 + \alpha) = 1$.

$G = F[\beta]$, $\beta^4 = 1$

$$G = F[\beta]$$

$$1 \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\beta \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\beta^2 \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\beta^3 \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$y \longmapsto \underbrace{(\alpha + \beta + \beta^2 + (1+\alpha)\beta^3)}_{c} \cdot y$$

Round keys

$k = (k[0], \ldots, k[3]) \in F^{4\times 4}$. One generates

$$(k[0], \ldots, k[39]) \in F^{4\times 40}$$

by the following rule for $i \geq 4$:

$$k[i] = \begin{cases} k[i-4] + k[i-1], & i \text{ is not divisible by } 4, \\ k[i-4] + \mathrm{SB}(k[i-1]) + s & i \text{ divisible by } 4, \end{cases}$$

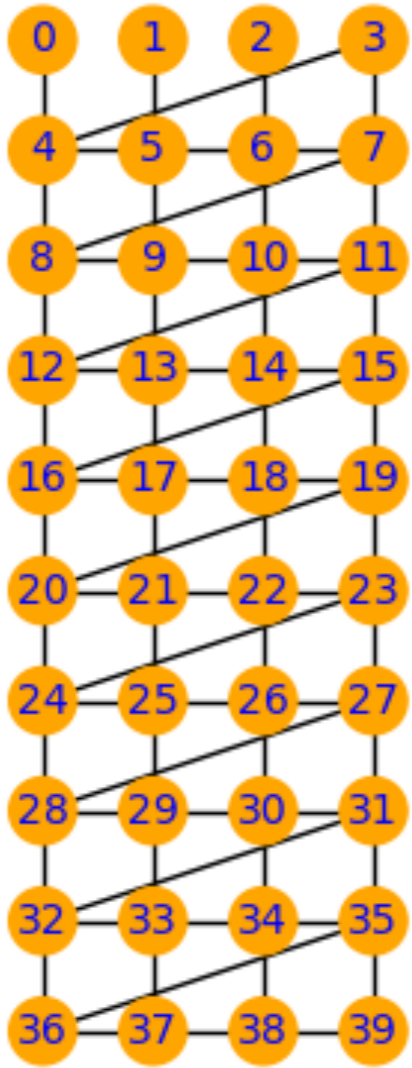
where

$$s = \begin{bmatrix} \alpha^{i/4-1} \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The round keys are

$$k^{(i)} = (k[4i], k[4i+1], k[4i+2], k[4i+3]) \in F^{4\times 4}.$$

AES keys


0 1 2 3
4 5 6 7
8 9 10 11
12 13 14 15
16 17 18 19
20 21 22 23
24 25 26 27
28 29 30 31
32 33 34 35
36 37 38 39

Encryption $(k, M) \mapsto M^{(10)}$

- $M^{(0)} = k^{(0)} + M$
- $M^{(i)} = k^{(i)} + \text{MixColumns}(\text{ShiftRows}(\text{SubBytes}(M^{(i-1)})))$ for $i = 1, \ldots, 9$
- $M^{(10)} = k^{(10)} + \text{ShiftRows}(\text{SubBytes}(M^{(9)}))$