

# Understanding T2 Standard Instance CPU Credits

*March 2019*



## Notices

Customers are responsible for making their own independent assessment of the information in this document. This document: (a) is for informational purposes only, (b) represents AWS's current product offerings and practices, which are subject to change without notice, and (c) does not create any commitments or assurances from AWS and its affiliates, suppliers or licensors. AWS's products or services are provided "as is" without warranties, representations, or conditions of any kind, whether express or implied. AWS's responsibilities and liabilities to its customers are controlled by AWS agreements, and this document is not part of, nor does it modify, any agreement between AWS and its customers.

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Contents

Introduction .....	1
Earned CPU Credits.....	1
Launch CPU Credits .....	1
CPU Utilization and CPU Credits.....	2
CPU Credit Earn Rates and CPU Utilization Rates .....	3
CPU Credit Earn Rates and Instance Sizes.....	4
Baseline Rates and Instance Sizes .....	5
CPU Credit Accrual Limits and the Discarding of Credits.....	6
The Five Phases in the CPU Credit System .....	7
Example: Tracking CPU Credit Usage .....	8
Period A — Balance at Maximum.....	9
Period B — Balance Stable .....	10
Period C — Balance Decreasing.....	11
Period D — Balance Decreasing.....	12
Period E — Balance Stable .....	13
Period F — Balance Decreases to Almost Zero .....	14
Period G — Balance at Minimum .....	15
Period H — Balance Increasing .....	16
Period I — Balance Increasing .....	17
Period J — Balance at Maximum .....	18
T2 Standard Instance Launch Credits .....	19
Launch Credit Allocation Limits .....	20
The Effects of Launch Credits on the CPU Credit Balance .....	21
Example: Tracking CPU Credit Accrual and Usage with Launch Credits .....	22
Period A — Launch Credits + 24 Hours of Earned Credits .....	23
Period B — Maximum Earned and Launch Credits .....	24

Period C — Spending Earned Credits.....	25
Period D — Balance Stable, 24 Hours of Earned Credits .....	26
Period E — Spending Earned Credits.....	27
Period F — Accruing Earned Credits .....	28
Period G — Balance Stable, 24 Hours of Earned Credits .....	30
Comparing T2 Instance Sizes With Identical Workloads .....	31
Scenario 1: Consuming CPU Credits at Different Rates .....	32
Scenario 2: Consuming 72 Credits Every 24 Hours .....	33
Scenario 3: Consuming 76 Credits Every 24 Hours .....	34
Scenario 4: Steady and Gradual Depletion of Credit Balance .....	35
Scenario 5: Variable CPU Utilization Rate .....	36
Scenario 6: Variable CPU Utilization Duration .....	36
Scenario 7: Consuming CPU Credits Immediately After Launch .....	37
Instances with Multiple vCPUs.....	38
Conclusion .....	39
Contributors .....	39
Further Reading.....	39
Document Revisions .....	39

## Abstract

Choosing the best Amazon EC2 instance type for your workload can be a challenge, especially if you are considering using a burstable instance type such as a T2 Standard instance. This document describes how a T2 Standard instance earns CPU credits, how launch credits are allocated, and how those launch and earned CPU credits are spent.

# Introduction

Most Amazon Elastic Compute Cloud (Amazon EC2) instance types provide a fixed level of CPU performance. However, the burstable performance instance types, T2 and T3, provide a baseline level of CPU performance with the ability to burst to a higher level (above that baseline) as required. The ability to use vCPUs at a rate higher than the baseline CPU utilization rate is governed by a CPU credit system.

Unlike the T2 Unlimited and T3 instance types, in addition to earned credits, T2 Standard instances can also be allocated launch credits. These two types of credits are treated differently and because the credit balance is presented as a single numeric value, it can be difficult to understand how the credits work.

## Earned CPU Credits

As a burstable instance type is running, it earns CPU credits. The rate at which an instance earns credits is based on the instance size—larger instance sizes earn CPU credits at a faster rate.

CPU credits are earned in fractions of credits and are allocated at 5-minute intervals. Up to 24hrs of earned credits can be accrued in the credit balance to be used later to burst above the baseline CPU utilization rate.

## Launch CPU Credits

A T2 Standard instance is allocated launch CPU credits during the instance launch, provided that the AWS account has not exceeded its launch credit limit. (See the [Launch Credit Allocation Limits](#) section for details.) These launch credits enable the instance to burst above the baseline CPU utilization rate immediately after launch—before any earned CPU credits have been accrued by the instance.

Launch credits are spent before earned CPU credits. Any unspent launch credits in the balance do not affect the accumulation of earned CPU credits.

**Note:** When a T2 instance is stopped (shut down), all CPU credits remaining within the CPU credit balance are forfeited.

## CPU Utilization and CPU Credits

During periods of CPU utilization (above 0%), CPU credits are redeemed for CPU time used. The utilization and corresponding CPU credit costs are calculated at millisecond granularity.

The following three vCPU utilization scenarios all result in the usage of 1 CPU credit:

- 1 vCPU @ 100% utilization for 60 seconds
- 1 vCPU @ 50% utilization for 120 seconds
- 2 vCPUs @ 25% utilization for 120 seconds

The following three vCPU utilization scenarios all result in the usage of 0.5 CPU credits:

- 1 vCPU @ 100% utilization for 30 seconds
- 1 vCPU @ 50% utilization for 60 seconds
- 2 vCPUs @ 25% utilization for 60 seconds

*Table 1: CPU Utilization Percentage vs Credit Utilization Rate Table*

vCPU Utilization Rate	Credits per Minute	Credits per Hour
100%	1	60
75%	0.75	45
50%	0.5	30
30%	0.3	18
25%	0.25	15
20%	0.2	12
15%	0.15	9
10%	0.1	6
5%	0.05	3
0%	0	0

## CPU Credit Earn Rates and CPU Utilization Rates

The CPU credit earn rate for an instance depends on the instance size and is directly related to the CPU utilization baseline. For example, a `t2.small` instance has a baseline CPU utilization rate of 20% and earns 12 CPU credits per hour.

In the next three examples, we see the effect of 3 different CPU utilization rates for a `t2.small` instance: below the baseline (10%), at the baseline (20%), and above the baseline (30%).

Example	CPU Utilization Rate	CPU Credits Spent	Description
T2 Small	10%	6 / hour	CPU credits are being spent at a slower rate than they are being earned.
T2 Small	20%	12 / hour	CPU credits are being spent at the same rate than they are being earned.
T2 Small	30%	18 / hour	CPU credits are being spent at a faster rate than they are being earned.

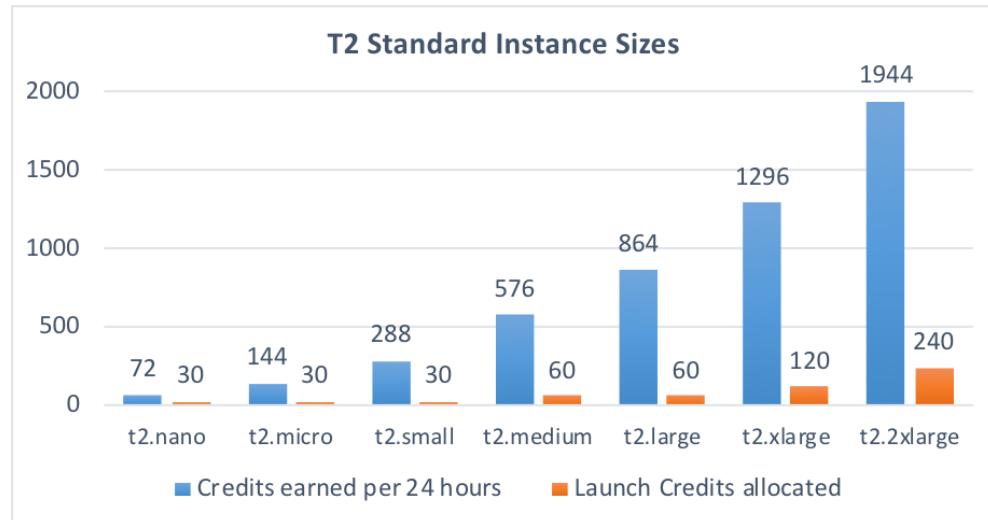
## CPU Credit Earn Rates and Instance Sizes

T2 Standard instances are available in multiple sizes to match different workloads. The number of vCPUs, the CPU credit earn rate, and the amount of memory varies by instance size as shown in the following table and graph.

Instance Size	CPU Credits Earned per 24 Hours	CPU Credits Earned per Hour	Maximum CPU Credit Balance *1	Baseline CPU Utilization *2	Launch Credits Granted	Number of vCPUs	Amount of Memory (GiB)
t2.nano	72	3	102	5%	30	1	0.5
t2.micro	144	6	174	10%	30	1	1
t2.small	288	12	318	20%	30	1	2
t2.medium	576	24	636	40%	60	2	4
t2.large	864	36	924	60%	60	2	8
t2.xlarge	1296	54	1416	90%	120	4	16
t2.2xlarge	1944	81	2184	135%	240	8	32

\*1 – The maximum CPU credit balance includes launch credits. Launch credits are allocated at launch and are not replenished after they are spent.

\*2 – Baseline CPU utilization is based on the equivalent utilization rate for a single vCPU. See the multiple vCPU section for details.

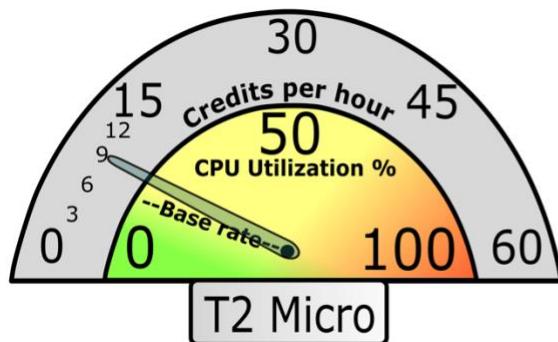


## Baseline Rates and Instance Sizes

The baseline CPU utilization rate for an instance is determined by the instance size—larger instance sizes have a higher baseline rate.

The per-vCPU utilization rate and the associated CPU credits spent do not vary with T2 Standard instance sizes. One minute of 100% vCPU utilization on a `t2.nano`, `t2.micro`, or `t2.small` equates to 1 CPU credit. (See “Instances with Multiple vCPUs” for information on CPU credit usage for `t2.medium` and larger instances.)

In the following example, the CPU utilization rate for both instances is 15% (9 CPU credits per hour). This utilization rate is above the baseline rate for a `t2.micro` instance but below the baseline rate for a `t2.small` instance:



### Instance Details

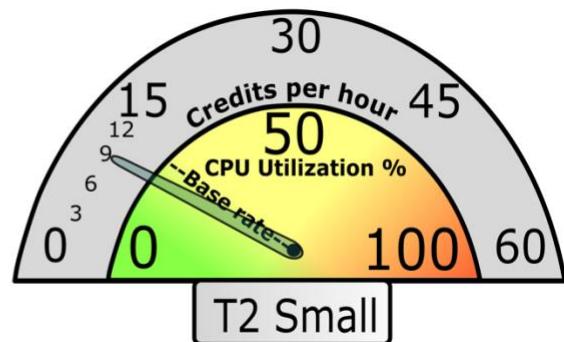
CPU base rate = 10%  
Credit earn rate = 6 / hour

### Current utilization details

CPU utilization rate – 15%  
Credit utilization rate – 9 / hour

### Result

Credits are being spent at a faster rate than they are being earned



### Instance Details

CPU base rate = 20%  
Credit earn rate = 12 / hour

### Current utilization details

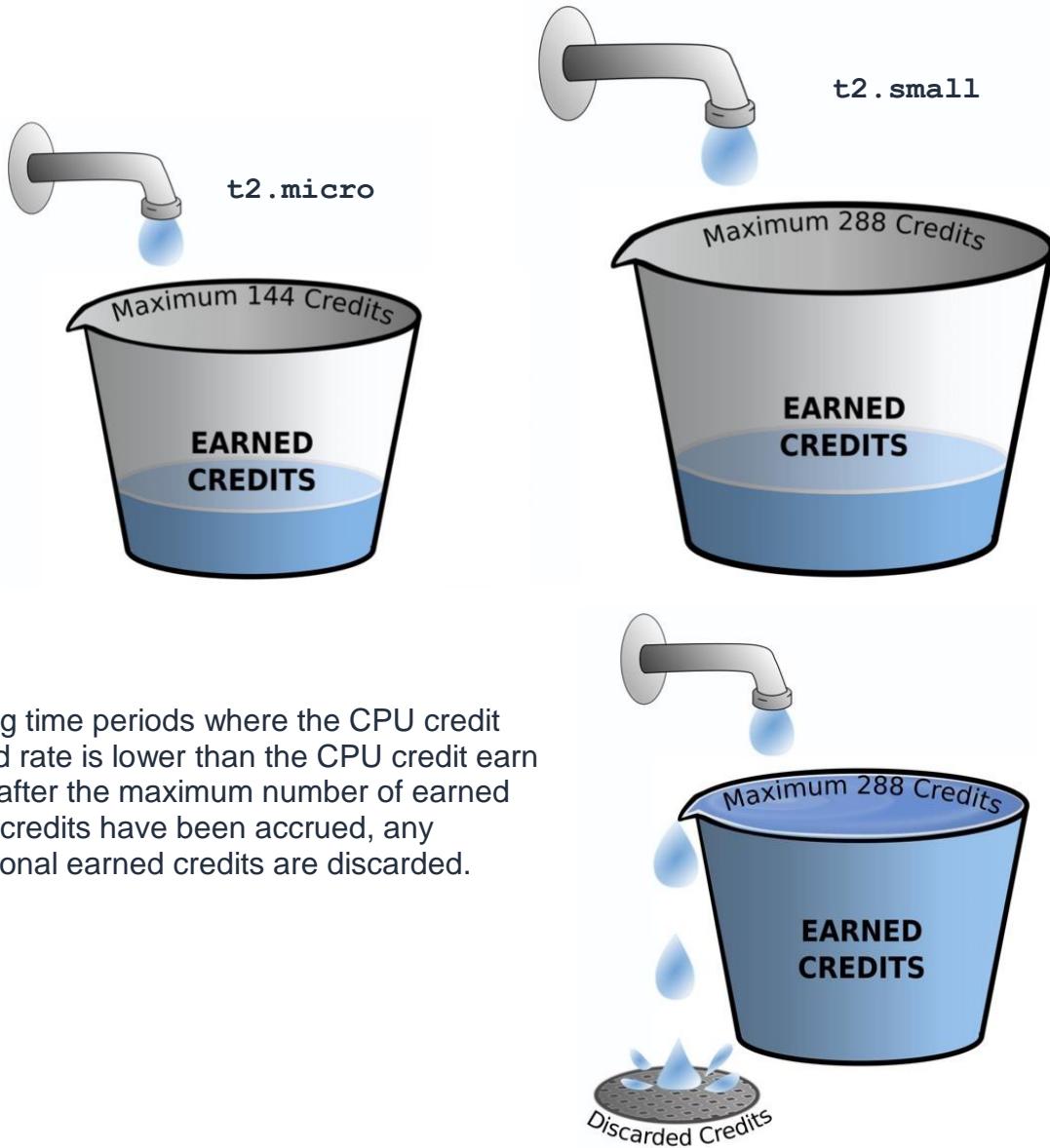
CPU utilization rate – 15%  
Credit utilization rate – 9 / hour

### Result

Credits are being spent at a slower rate than they are being earned.

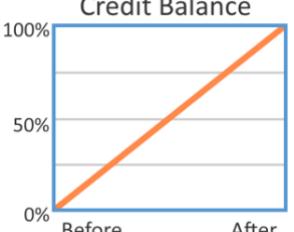
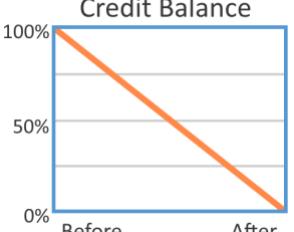
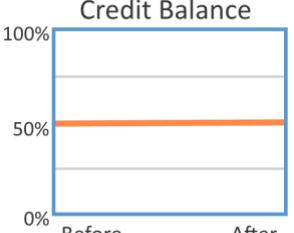
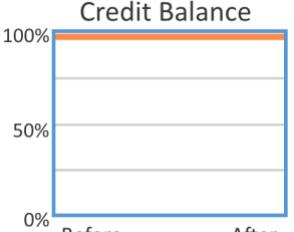
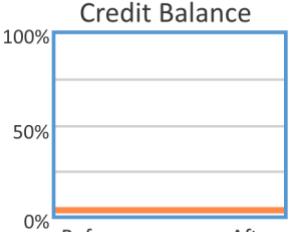
## CPU Credit Accrual Limits and the Discarding of Credits

The maximum number of earned CPU credits that can be accrued by a T2 Standard instance varies by instance size. As the following diagram shows, a larger instance size has a larger bucket for accruing CPU credits.



**Note:** To avoid the complexity associated with launch credits, the next examples describing CPU credits exclude launch credits. See “T2 Standard Launch Credits” for a complete discussion.

## The Five Phases in the CPU Credit System

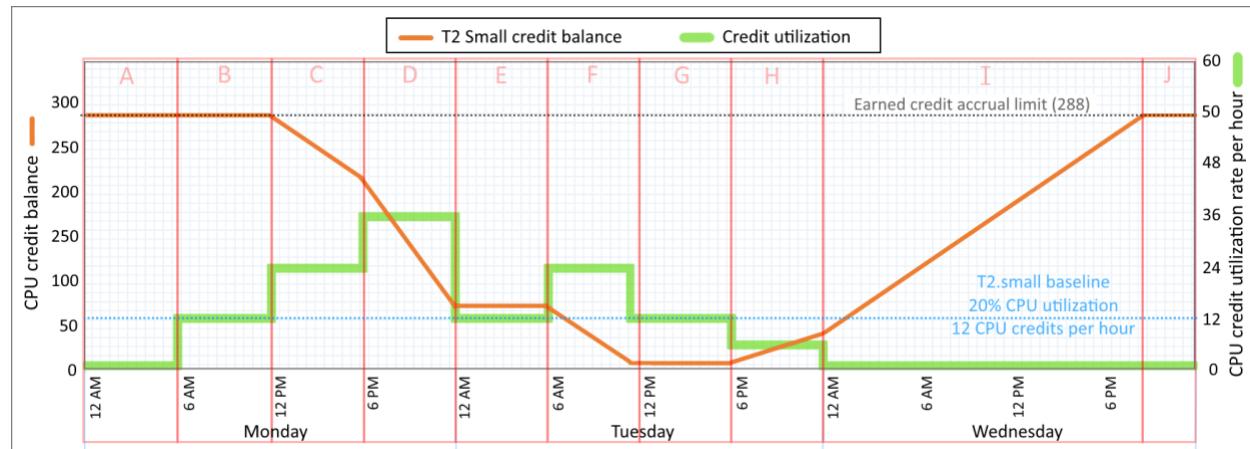
Phase Details	Credit Balance Chart	State Before	State After								
<b>Balance Increasing</b>  During periods where the CPU credit spend rate is less than the earn rate, you accumulate credits.	 <p>Credit Balance</p> <table border="1"> <tr> <td>100%</td> <td></td> </tr> <tr> <td>50%</td> <td></td> </tr> <tr> <td>0%</td> <td></td> </tr> <tr> <td>Before</td> <td>After</td> </tr> </table>	100%		50%		0%		Before	After	 <p>Discarded  Spent</p>	 <p>Discarded  Spent</p>
100%											
50%											
0%											
Before	After										
<b>Balance Decreasing</b>  During periods where the CPU credit spend rate is greater than the earn rate, your credit balance declines.	 <p>Credit Balance</p> <table border="1"> <tr> <td>100%</td> <td></td> </tr> <tr> <td>50%</td> <td></td> </tr> <tr> <td>0%</td> <td></td> </tr> <tr> <td>Before</td> <td>After</td> </tr> </table>	100%		50%		0%		Before	After	 <p>Discarded  Spent</p>	 <p>Discarded  Spent</p>
100%											
50%											
0%											
Before	After										
<b>Balance Stable</b>  During periods where the CPU credit spend rate is the same as the earn rate, the number of accumulated credits remains unchanged.	 <p>Credit Balance</p> <table border="1"> <tr> <td>100%</td> <td></td> </tr> <tr> <td>50%</td> <td></td> </tr> <tr> <td>0%</td> <td></td> </tr> <tr> <td>Before</td> <td>After</td> </tr> </table>	100%		50%		0%		Before	After	 <p>Discarded  Spent</p>	 <p>Discarded  Spent</p>
100%											
50%											
0%											
Before	After										
<b>Balance at Maximum</b>  During periods where the CPU spend rate is less than the earn rate and you have the maximum number of CPU credits accrued, additional earned credits are discarded.	 <p>Credit Balance</p> <table border="1"> <tr> <td>100%</td> <td></td> </tr> <tr> <td>50%</td> <td></td> </tr> <tr> <td>0%</td> <td></td> </tr> <tr> <td>Before</td> <td>After</td> </tr> </table>	100%		50%		0%		Before	After	 <p>Discarded  Spent</p>	 <p>Discarded  Spent</p>
100%											
50%											
0%											
Before	After										
<b>Balance at Minimum</b>  During periods where the credit balance is nearly depleted, the maximum utilization rate is restricted to the base rate. (The credit balance does not reach zero.)	 <p>Credit Balance</p> <table border="1"> <tr> <td>100%</td> <td></td> </tr> <tr> <td>50%</td> <td></td> </tr> <tr> <td>0%</td> <td></td> </tr> <tr> <td>Before</td> <td>After</td> </tr> </table>	100%		50%		0%		Before	After	 <p>Discarded  Spent</p>	 <p>Discarded  Spent</p>
100%											
50%											
0%											
Before	After										

## Example: Tracking CPU Credit Usage

In this section, we illustrate CPU credit usage over time and its effect on the CPU credit balance for a `t2.small` instance over 3 days. The 3 days are divided into 10 separate periods, identified by the letters A through J, and each period is described individually in the following sections.

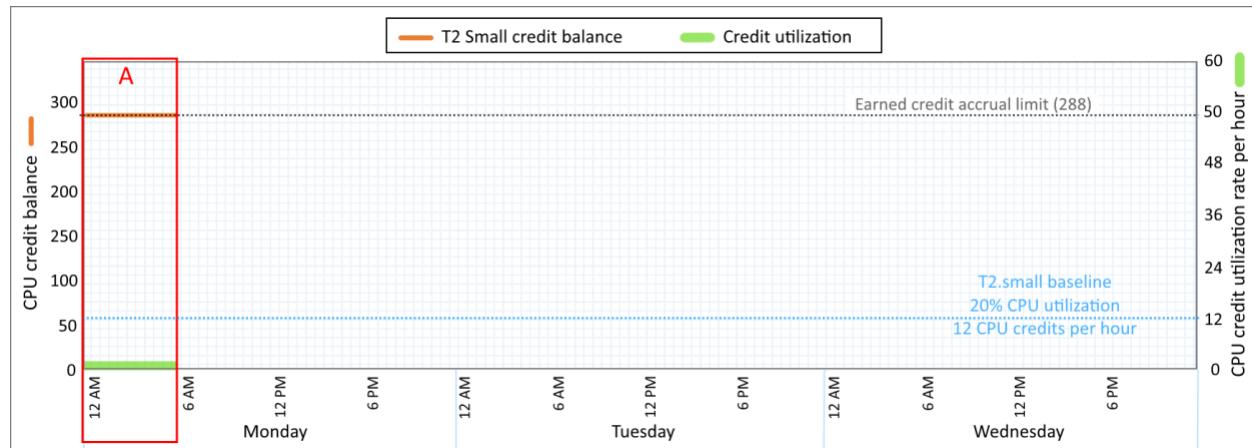
At the start of this example, assume the following:

- The credit balance contains the maximum number of earned CPU credits (288) that can be accrued by a `t2.small` instance.
- The credit balance consists only of earned CPU credits. There are no launch credits in the balance. (A later example includes launch credits.)



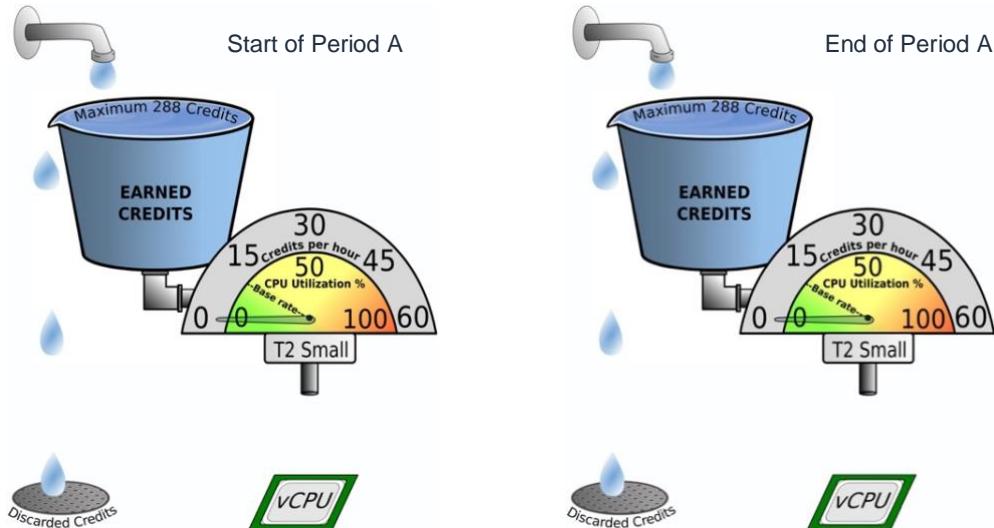
## Period A — Balance at Maximum

During this first period, the credit utilization rate is zero and the number of earned credits is at the maximum limit of 24 hours of earned credits (288). Any newly earned credits are discarded.



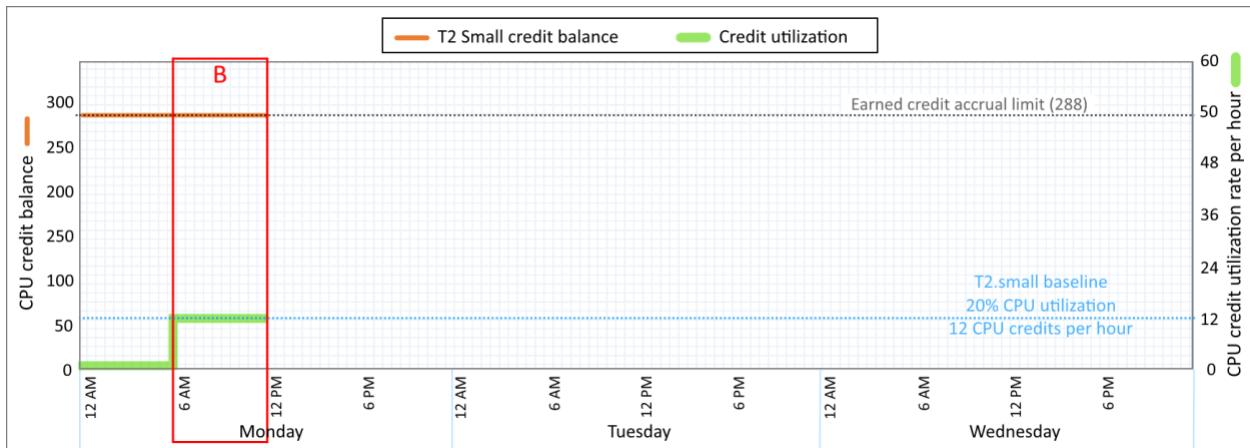
### Period A

<b>Credit Spend Rate</b>	0 credits per hour (0% of credit earn rate) 0% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	12 credits per hour (100% of credit earn rate)
<b>Credit Balance</b>	Balance is stable at 288 credits. (0 launch credits and 288 earned CPU credits)



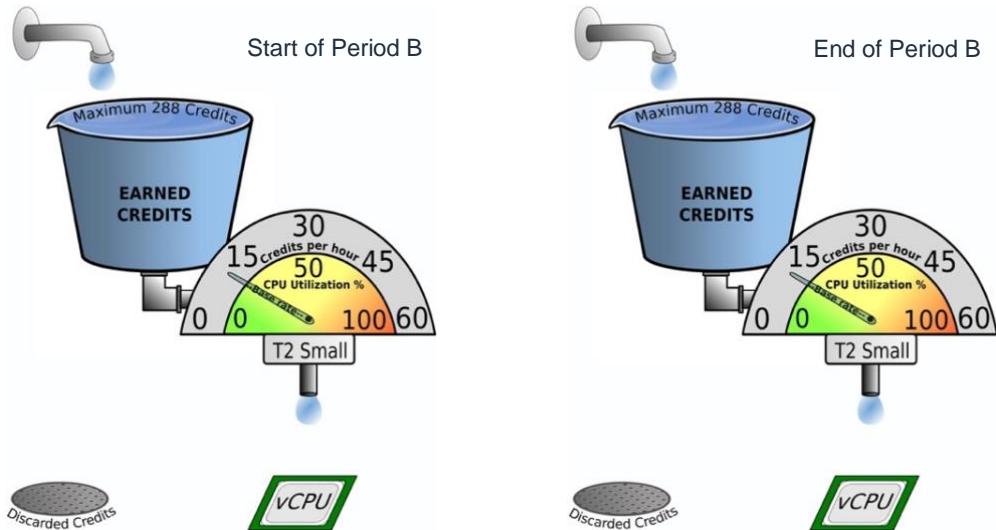
## Period B — Balance Stable

During this period, the credit utilization rate is equal to the credit earn rate, therefore credits are being replaced as they are spent. This results in the balance remaining unchanged at 288 credits.



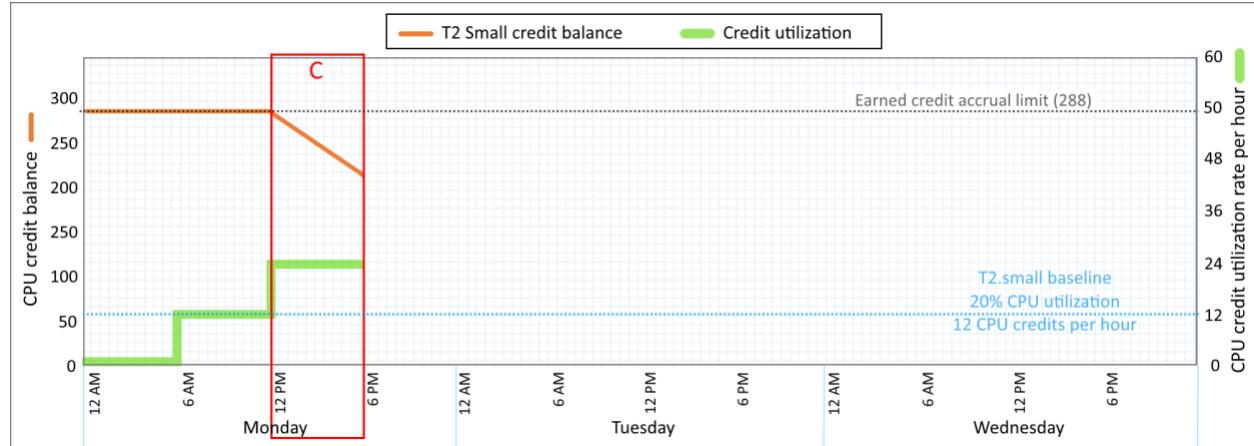
### Period B

<b>Credit Spend Rate</b>	12 credits per hour (100% of credit earn rate) 20% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour (0% of credit earn rate)
<b>Credit Balance</b>	Balance is stable at 288 credits. (0 launch credits and 288 earned CPU credits)



## Period C — Balance Decreasing

During this period, the credit utilization rate is two times the credit earn rate, therefore credits are being consumed from the credit balance faster than they can be replenished by earned credits.



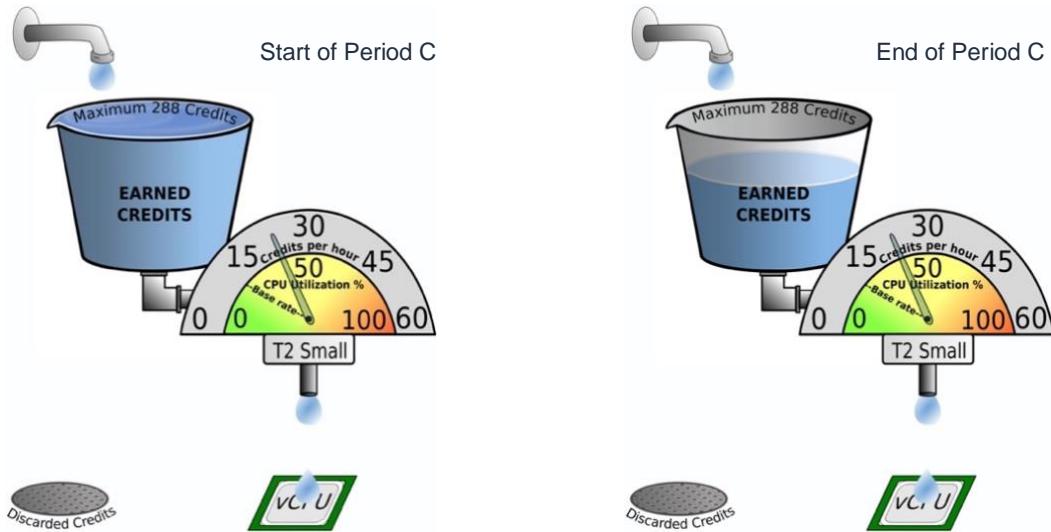
### Period C

**Credit Spend Rate** 24 credits per hour (200% of credit earn rate)  
40% CPU utilization

**Credit Earn Rate** 12 credits per hour

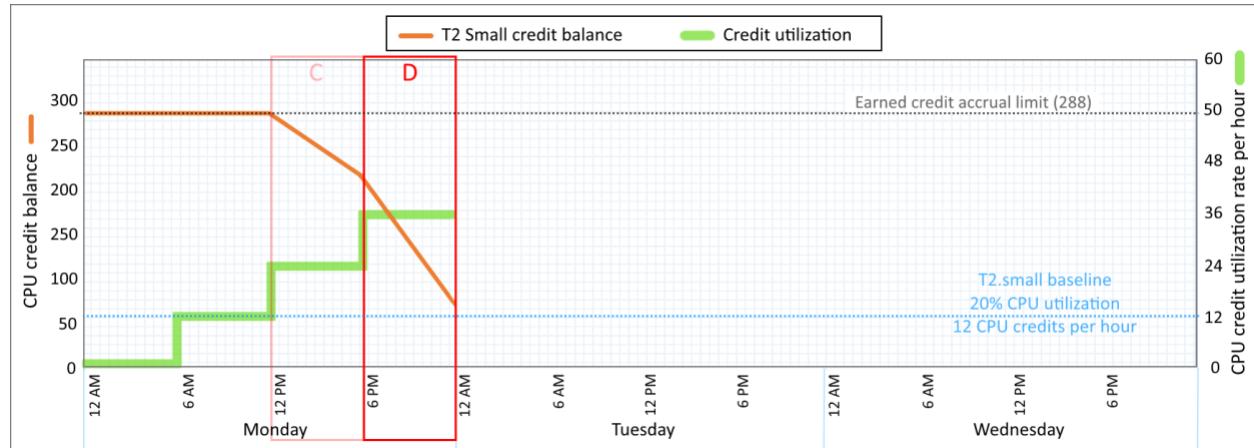
**Credit Discard Rate** 0 credits per hour (0% of credit earn rate)

**Credit Balance** Balance decreases at a rate of 12 credits per hour  
Change rate = earn rate (12) - spend rate (24)



## Period D — Balance Decreasing

During this period, the credit utilization rate is three times higher than the credit earn rate, therefore credits are being consumed from the credit balance at a faster rate than during period C.



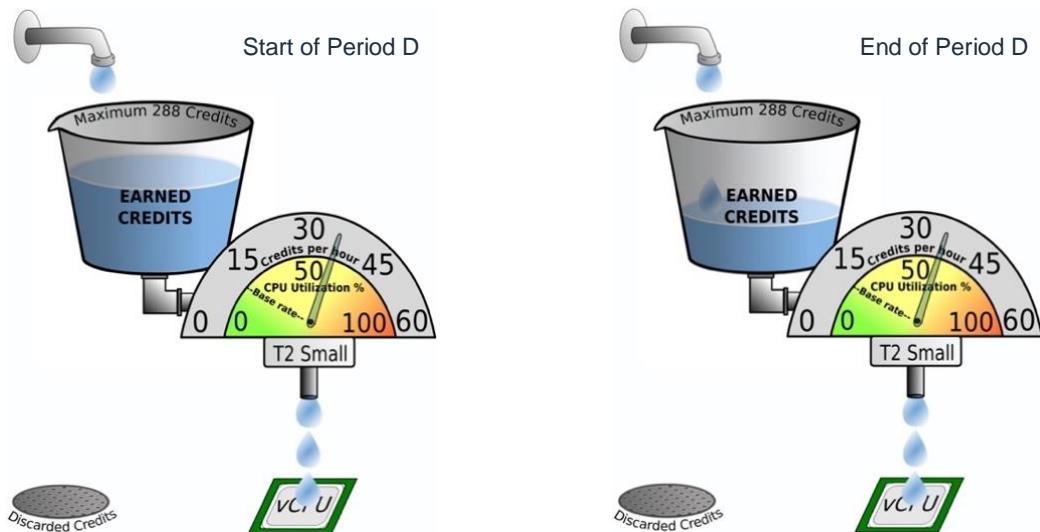
### Period D

**Credit Spend Rate** 36 credits per hour (300% of credit earn rate)  
60% CPU utilization

**Credit Earn Rate** 12 credits per hour

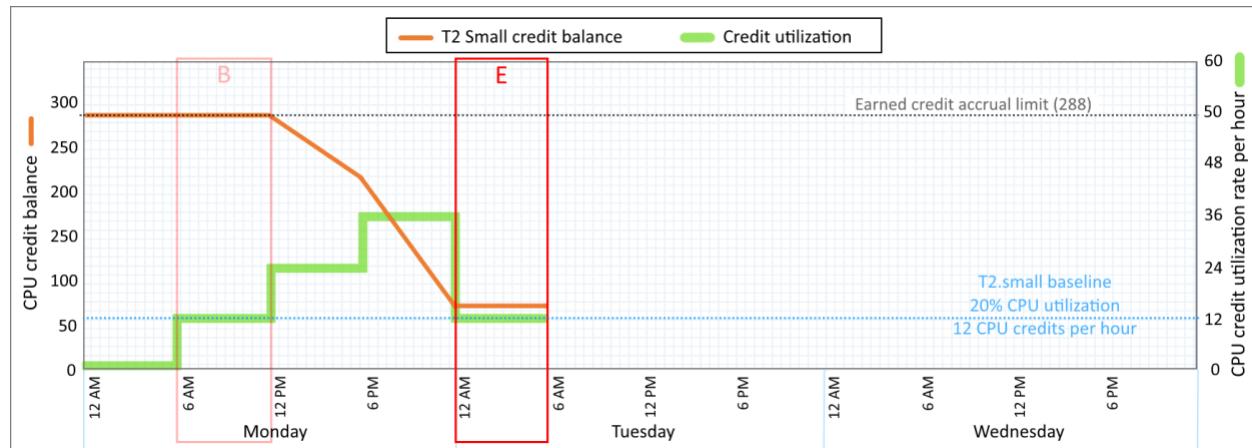
**Credit Discard Rate** 0 credits per hour (0% of credit earn rate)

**Credit Balance** Balance decreases at a rate of 24 credits per hour  
Change rate = earn rate (12) - spend rate (36)



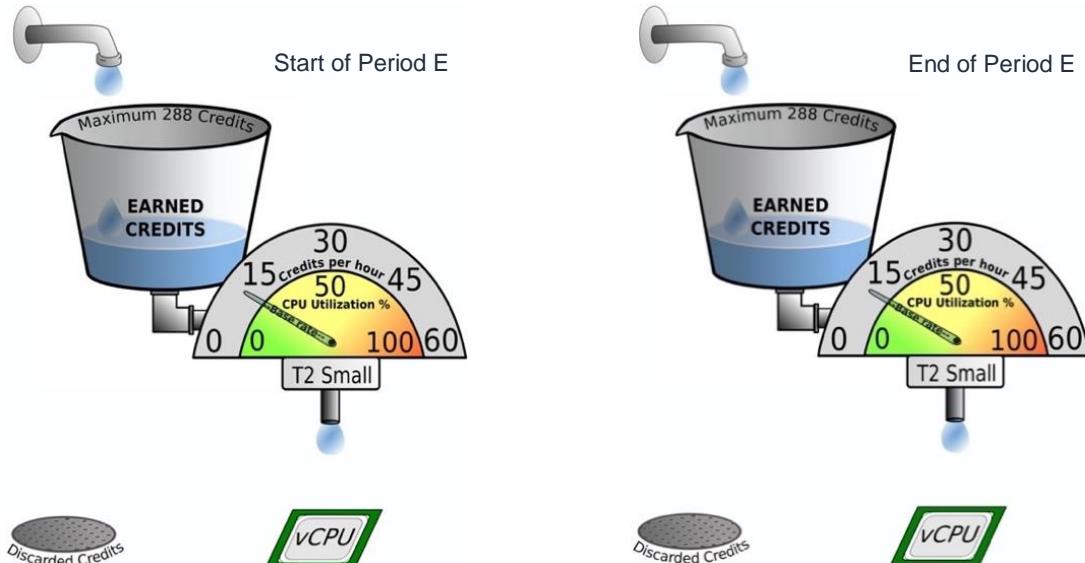
## Period E — Balance Stable

During this period, as in period B, the credit utilization rate is equal to the credit earn rate. Therefore, credits are being replaced as they are spent, resulting in the balance remaining stable.



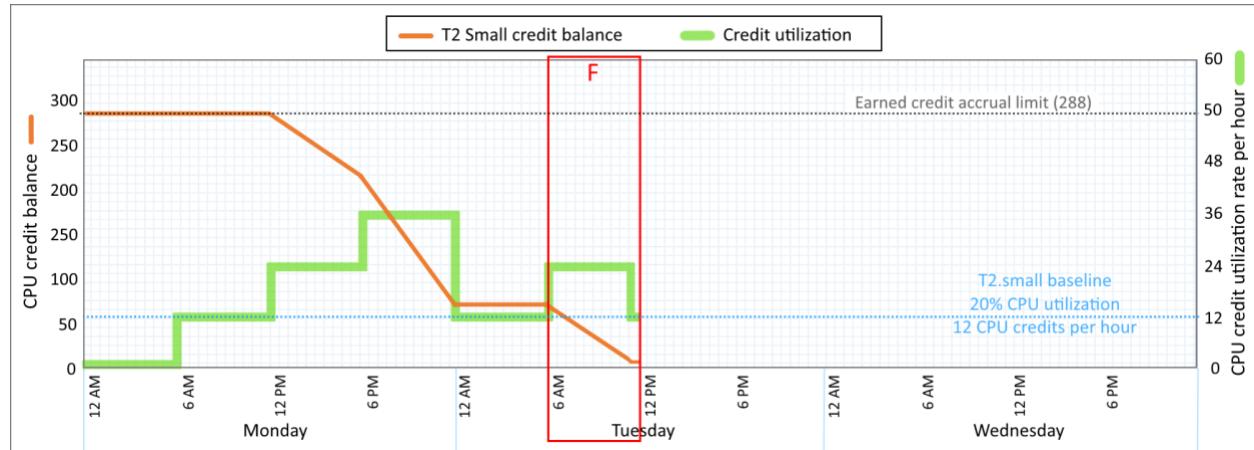
### Period E

<b>Credit Spend Rate</b>	12 credits per hour (100% of credit earn rate) 20% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour (0% of credit earn rate)
<b>Credit Balance</b>	Balance is stable at 72 credits (0 launch credits and 72 earned CPU credits)



## Period F — Balance Decreases to Almost Zero

During this period, the instance was consuming CPU credits two times faster than they are being earned. Because there were enough CPU credits in the credit balance, the workload was able to run unrestricted most of this period. However, near the end of the period, when the credit balance was nearly depleted, the CPU credit system restricted the maximum attainable CPU utilization to the base rate for a `t2.small` instance, 20%.



### Period F

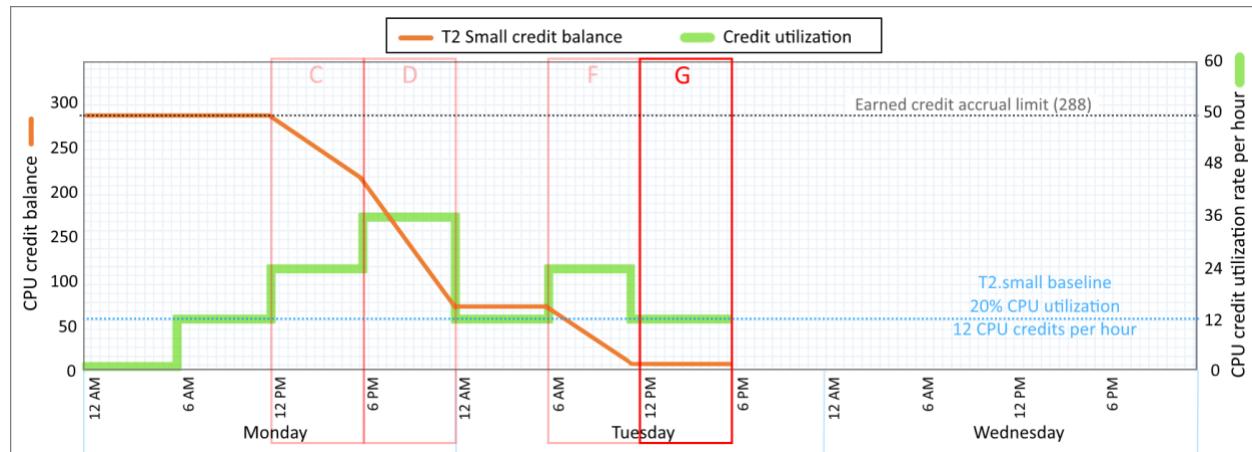
<b>Credit Spend Rate</b>	24 credits per hour (200% of credit earn rate) 40% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour (0% of credit earn rate)
<b>Credit Balance</b>	Balance decreases at a rate of 12 credits per hour Change rate = earn rate (12) - spend rate (24) At the end of period F, the credit balance is nearly depleted and the CPU utilization is limited to the base rate.



## Period G — Balance at Minimum

During this period, the credit balance remains stable near zero as the number of CPU credits are being spent as fast as they are earned.

When the credit balance is near zero, the maximum attainable CPU utilization is restricted to the baseline for the instance size, which is 20% in the case of a `t2.small`. Even if the workload required a similar vCPU utilization rate to what it had in periods C, D and F, the T2 Standard CPU credit system limits it to the base rate.



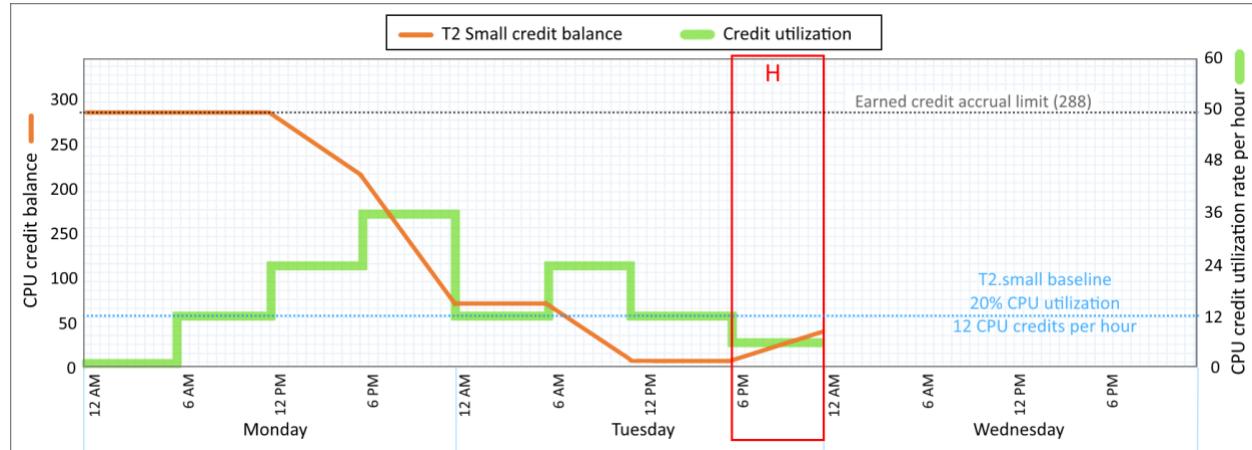
### Period G

<b>Credit Spend Rate</b>	12 credits per hour (100% of credit earn rate) 20% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour (0% of credit earn rate)
<b>Credit Balance</b>	Balance is stable at almost zero credits. (0 launch credits and almost zero earned CPU credits)



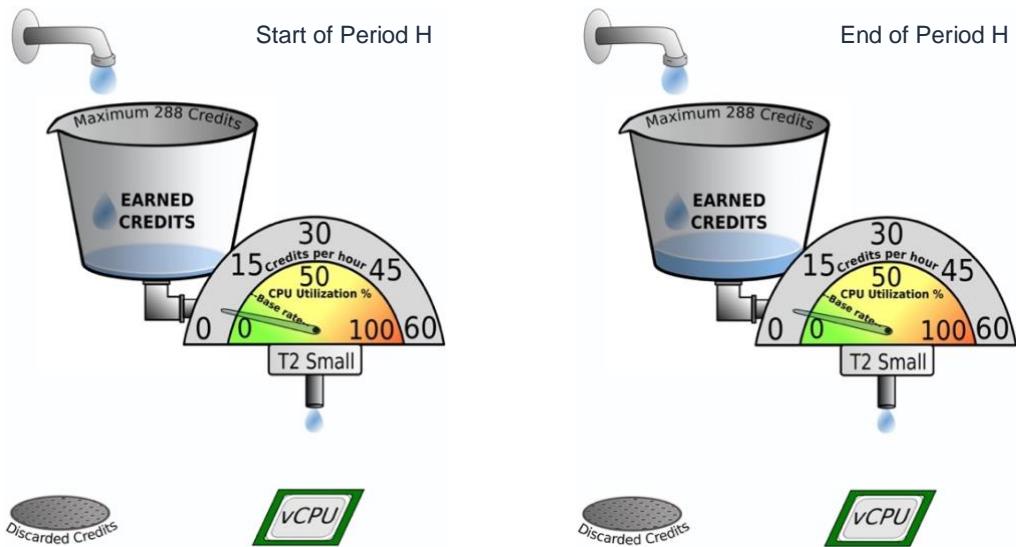
## Period H — Balance Increasing

During this period, the credit utilization rate is half of the credit earn rate and CPU credits are being added to the credit balance at a rate if 6 per hour.



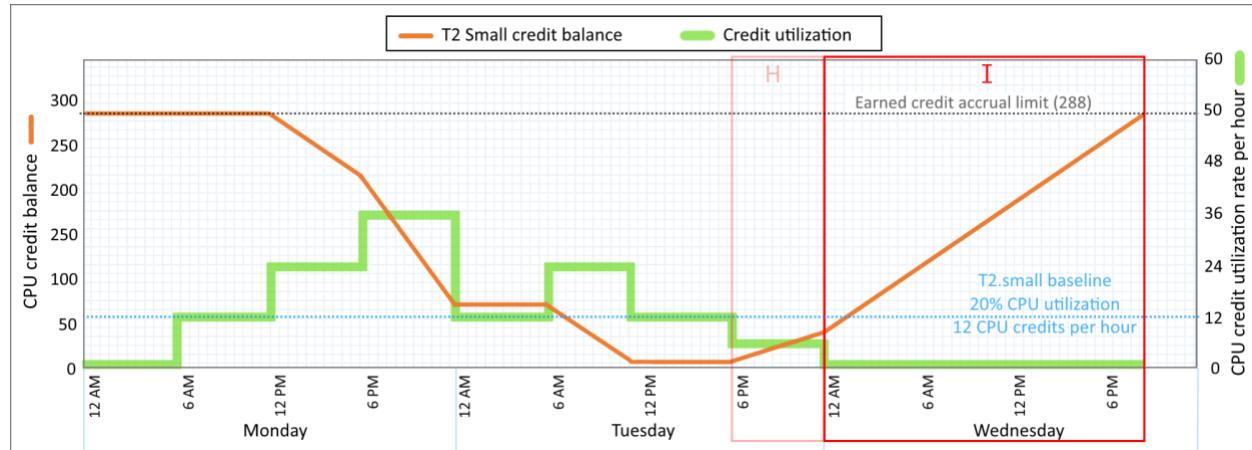
### Period H

<b>Credit Spend Rate</b>	6 credits per hour (50% of credit earn rate) 10% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour (0% of credit earn rate)
<b>Credit Balance</b>	Balance increases at a rate of 6 credits per hour. Change rate = earn rate (12) - spend rate (6)



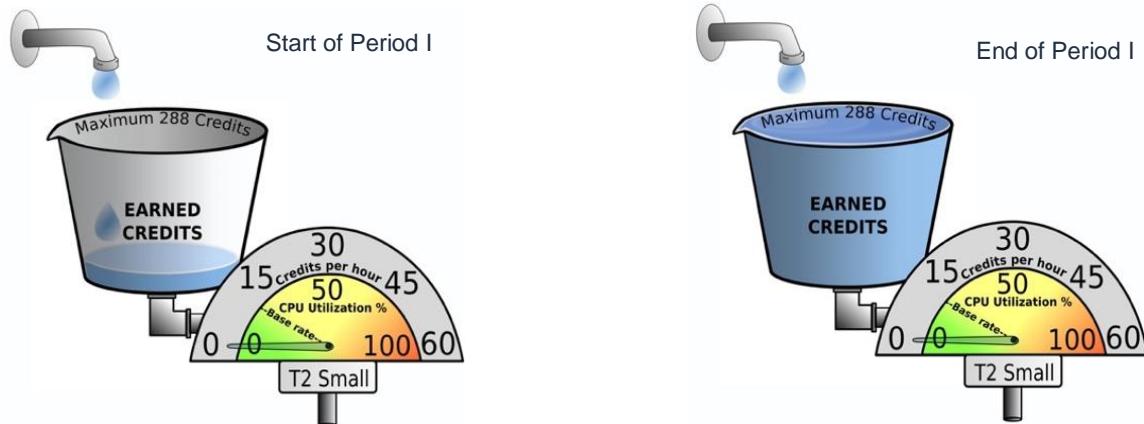
## Period I — Balance Increasing

During this period, the credit utilization rate is zero and all earned CPU credits are being added to the credit balance at a rate of 12 per hour, which is double that of period H. By the end of the period, the credit balance contains the maximum number of earned credits allowed.



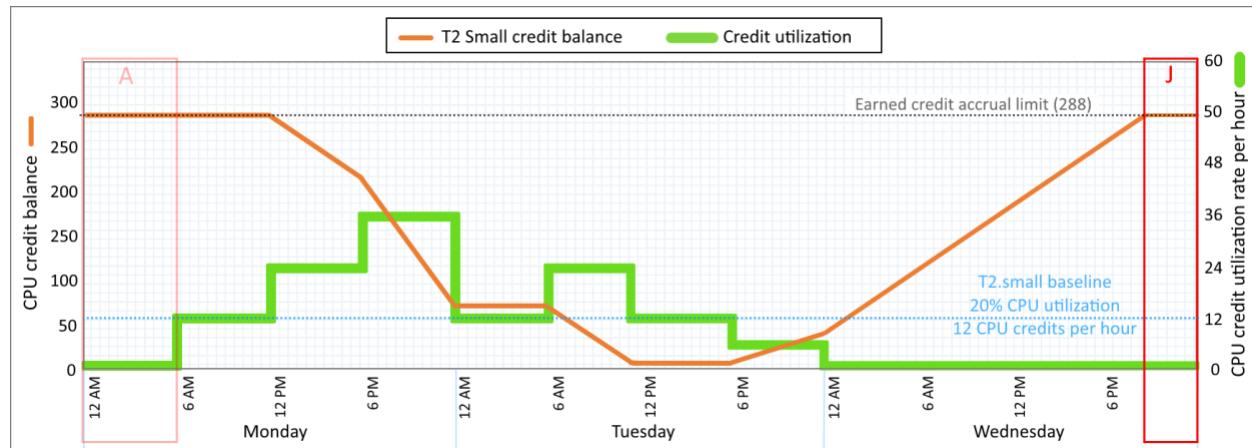
### Period I

<b>Credit Spend Rate</b>	0 credits per hour (0% of credit earn rate) 0% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour (0% of credit earn rate)
<b>Credit Balance</b>	Balance increases at a rate of 12 credits per hour. Change rate = earn rate (12) - spend rate (0)



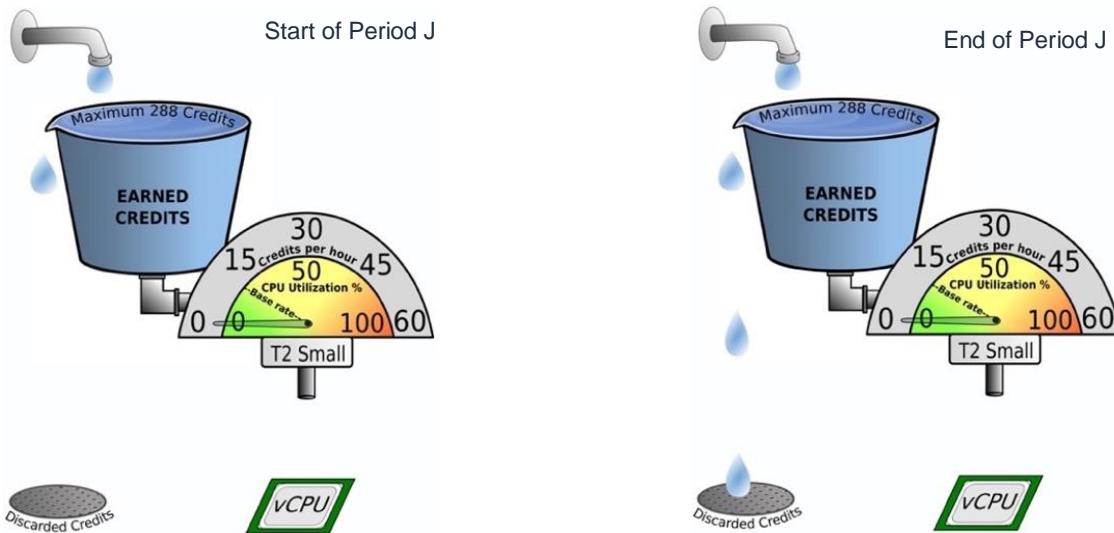
## Period J — Balance at Maximum

During this period, as in period A, the credit utilization rate is zero and the credit balance contains the maximum number of earned credits allowed (288). Any newly earned and unspent credits are discarded.



### Period J

<b>Credit Spend Rate</b>	0 credits per hour (0% of credit earn rate) 0% CPU utilization
<b>Credit Earn Rate</b>	12 credits per hour
<b>Credit Discard Rate</b>	12 credits per hour (100% of credit earn rate)
<b>Credit Balance</b>	Balance is stable at 288 credits. Change rate = earn rate (12) – spend rate (0) - discard rate (12)



## T2 Standard Instance Launch Credits

Launch credits enable a T2 Standard instance to burst above the baseline level of CPU utilization immediately after launch—prior to it having earned CPU credits and accruing them in the credit balance. Launch credits only apply to T2 Standard instances.

Launch credit features:

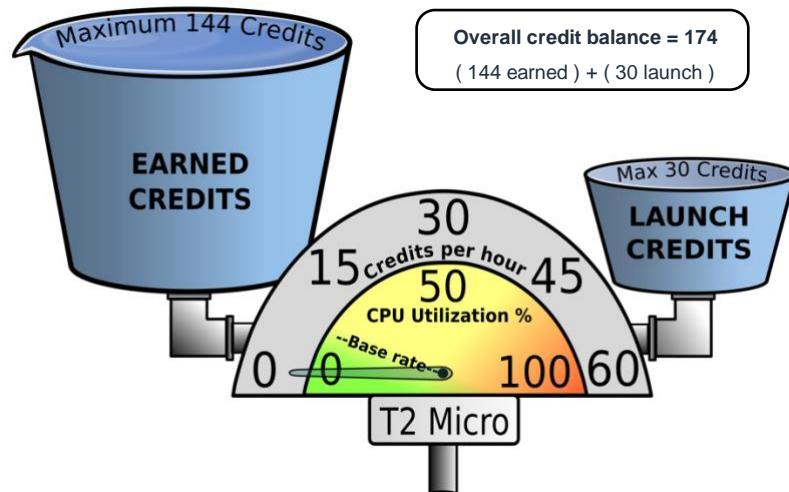
- Launch credits are added to the overall CPU credit balance.
- Launch credits are spent before earned CPU credits.
- Launch credits do not affect the accumulation of earned CPU credits.
- Launch credits do not get replenished while the instance is running.
- Launch credits are not allocated when the allocation limit is exceeded.

If you don't take these features into account, under certain circumstances the CPU credit balance can seem to behave in ways that you might not expect. For example:

- The CPU credit balance can plateau at different values.
- The CPU credit balance can experience different behavior over time, even if the workload CPU utilization rate is unchanged.

To better understand the effect of launch credits on the overall CPU credit balance, picture the credit balance as being comprised of two buckets of credits instead of one:

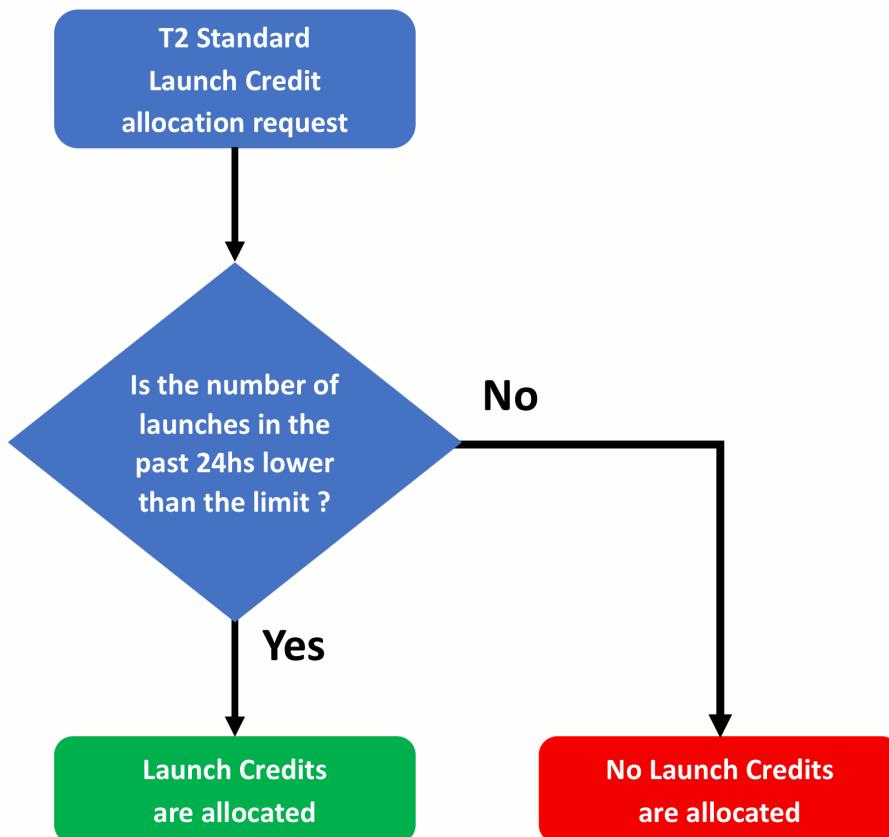
- A bucket for the accrued earned CPU credits, which is filled during times when the spend rate is lower than the earn rate.
- A second bucket for the launch credits that is filled at launch time but does not get replenished while the instance is running.



## Launch Credit Allocation Limits

Launch credits are only allocated to T2 Standard instances during their launch if the particular instance launch is within the account's Launch Credit Allocation Limit. The default limit is 100 launches or starts per account, per region, per rolling 24-hour period. The limit can be reached through any combination of launches (or stops and starts) within the same account and same region during the same rolling 24-hour period. For example:

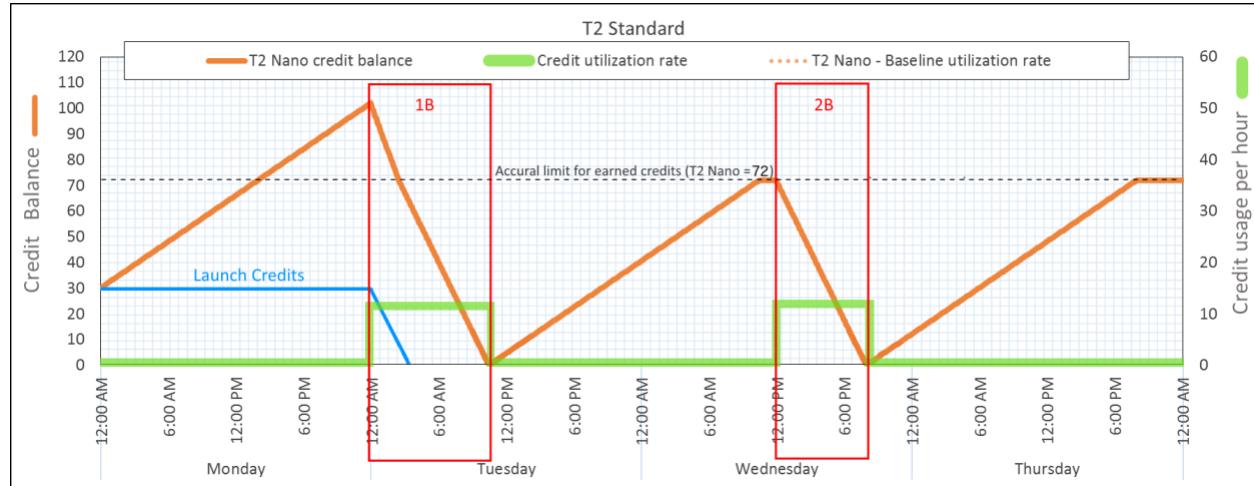
- 100 new T2 Standard instance launches, or
- 100 existing T2 Standard instance stops and starts, or
- 50 existing T2 Standard instance stops and starts and 50 new T2 Standard instance launches



**Note:** If you are regularly exceeding the launch credit allocation limit, you might want to switch to a T2 Unlimited or T3 instance instead.

## The Effects of Launch Credits on the CPU Credit Balance

If a T2 Standard instance is launched but does not consume all of the launch credits within the first 24 hours, then the credit balance will consist of the remaining launch credits plus 24 hours of earned credits. For example, a `t2.nano` instance could potentially accrue a total of 102 credits (72 earned credits plus 30 launch credits). The instance could then spend all 102 credits in a single continuous burst as illustrated in period 1B in the following graph.



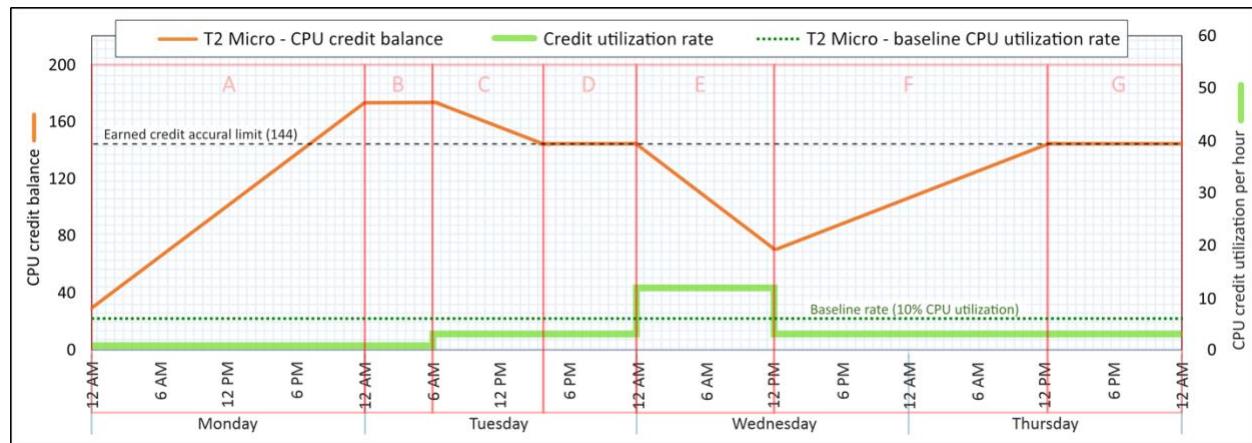
**Note:** Launch credits in the credit balance are illustrated by the blue line in the graph. Remember that Amazon CloudWatch only reports total credits—you cannot see the breakdown of launch credits and earned CPU credits.

Attaining a credit balance that is higher than the 24-hour earned CPU credit value can only be achieved one time per instance launch because after the launch credits are spent, they are not replenished. Any subsequent CPU credit accruals are limited to the value of 24 hours of earned credits, as illustrated at the start of period 2B in the graph.

## Example: Tracking CPU Credit Accrual and Usage with Launch Credits

In this section, we illustrate CPU credit accrual and usage for a `t2.micro` instance over a 4 day period, considering the effect that launch credits have on the credit balance. This example is specifically tailored to highlight some of the complexity that can be associated with launch credits.

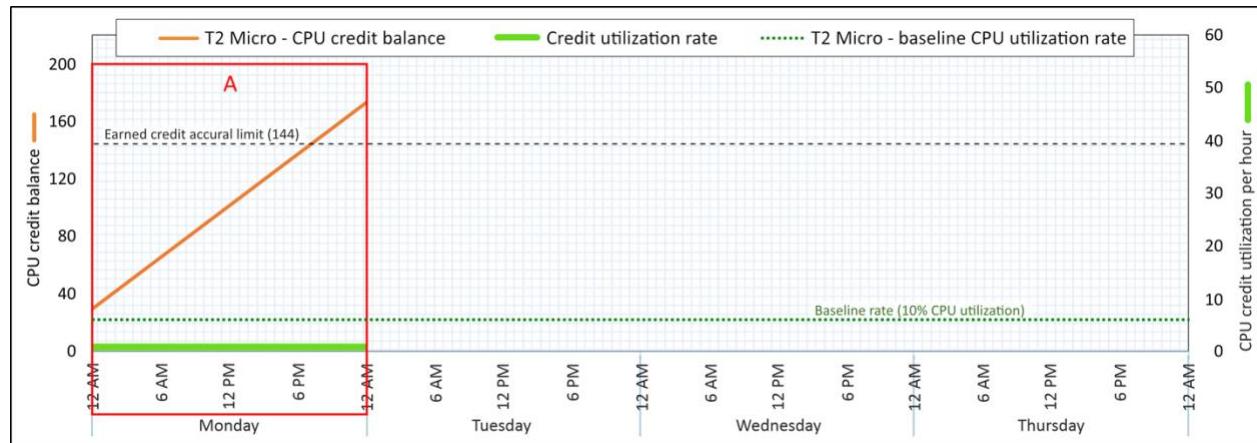
Because of launch credits, the credit balance for periods A, B, and C in this example is above the `t2.micro` 24-hour earned credit value of 144.



## Period A — Launch Credits + 24 Hours of Earned Credits

Immediately upon the launch of the `t2.micro` instance, 30 launch credits are added to the overall credit balance and the instance starts to earn credits.

Because no CPU credits are spent or discarded during this period, the credit balance increases at a rate of 6 credits per hour. In addition to the 30 launch credits, after 24 hours, the instance has accrued 144 earned CPU credits. The credit balance is able to increase above 144 credits because the unspent launch credits do not affect the accumulation of earned CPU credits.



### Period A (duration 24 hours) 12 AM Monday – 12 AM Tuesday

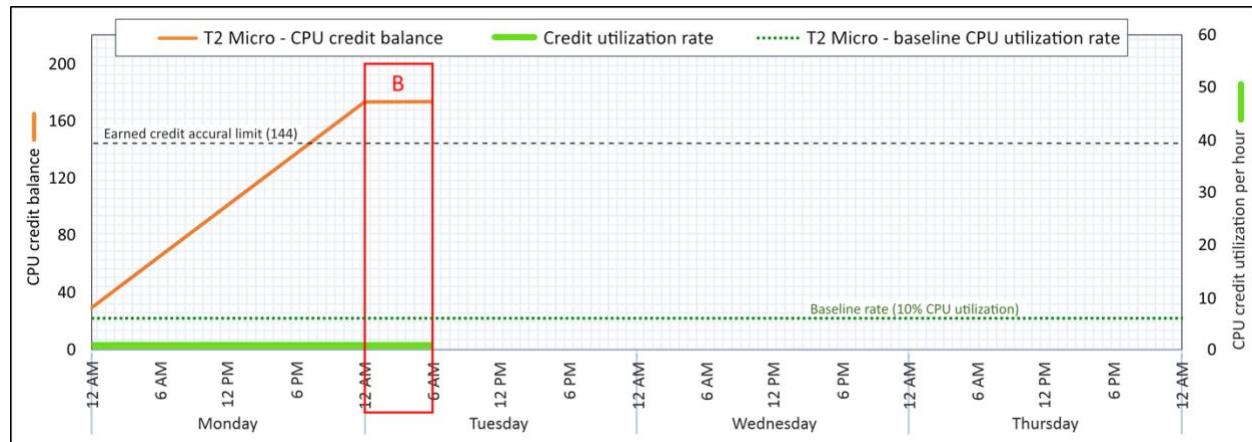
<b>Credit Spend Rate</b>	0 credits per hour (0% CPU utilization)
<b>Credit Earn Rate</b>	6 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour
<b>Credit Balance</b>	Balance increases from 30 at launch to 174 credits. (30 launch credits and 144 earned CPU credits)



## Period B — Maximum Earned and Launch Credits

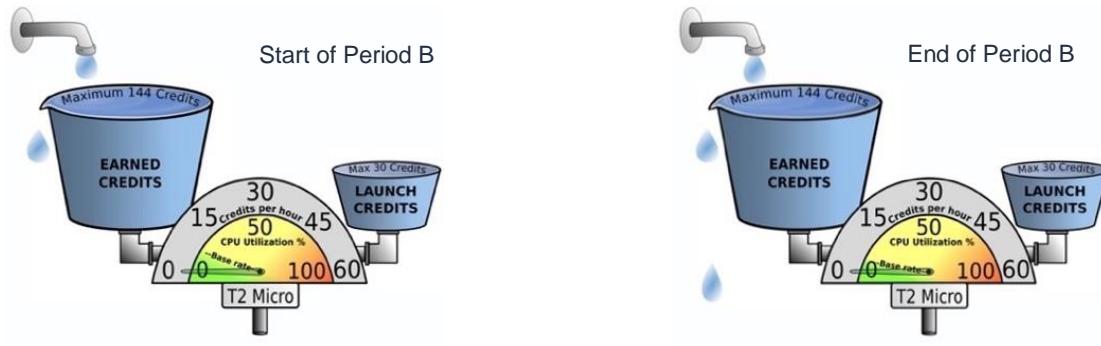
At the start of period B, the credit balance is 174 credits. The overall balance consists of 30 launch credits and 144 earned credits.

Because the credit balance contains the maximum number of earned CPU credits for a `t2.micro` instance (144 credits) any newly earned credits above this limit are discarded. This results in the credit balance plateauing at a value equal to 24 hours of earned credits (144) plus the unspent launch credits (30).



### Period B (duration 6 hours) 12 AM Tuesday – 6 AM Tuesday

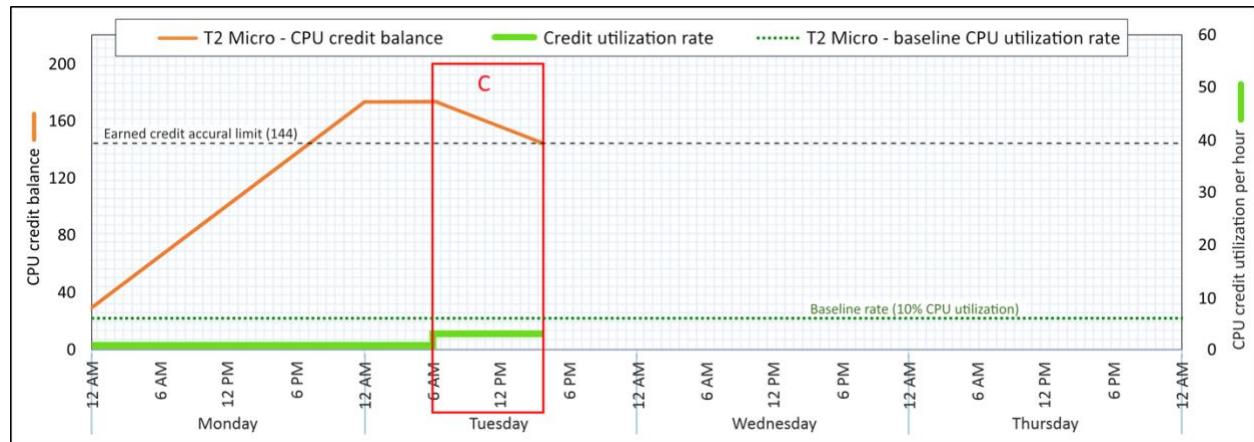
<b>Credit Spend Rate</b>	0 credits per hour (0% CPU utilization)
<b>Credit Earn Rate</b>	6 credits per hour
<b>Credit Discard Rate</b>	6 credits per hour
<b>Credit Balance</b>	Balance remains stable at 174 credits. (30 launch credits and 144 earned CPU credits)



## Period C — Spending Earned Credits

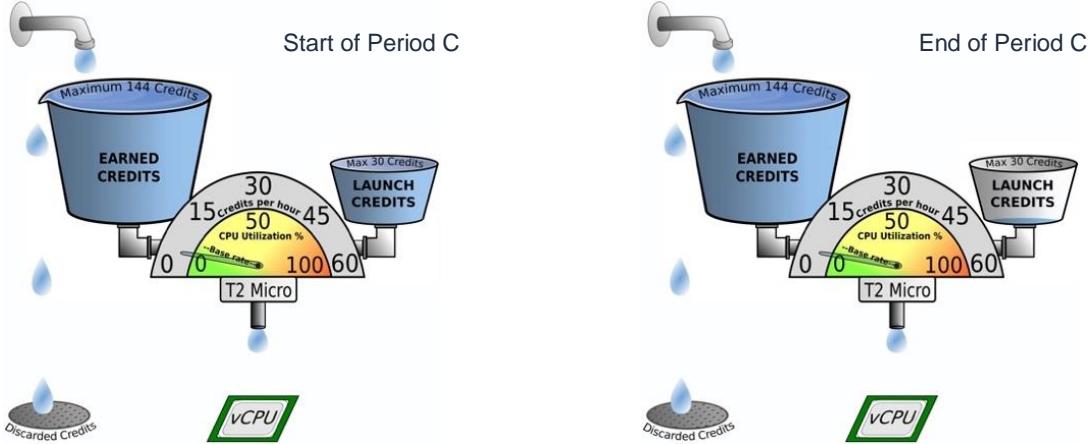
In period C, the instance consumes CPU credits at a rate of 3 credits per hour (50% of the credit earn rate). Despite the spend rate being less than the earn rate, the overall credit balance is decreasing at a rate equal to the credit spend rate (3 credits per hour).

This occurs because the non-replenishable launch credits are being spent first, and all freshly earned CPU credits are being discarded because the credit balance already has the maximum number of earned CPU credits (144).



### Period C (duration 10 hours) 6 AM Tuesday – 4 PM Tuesday

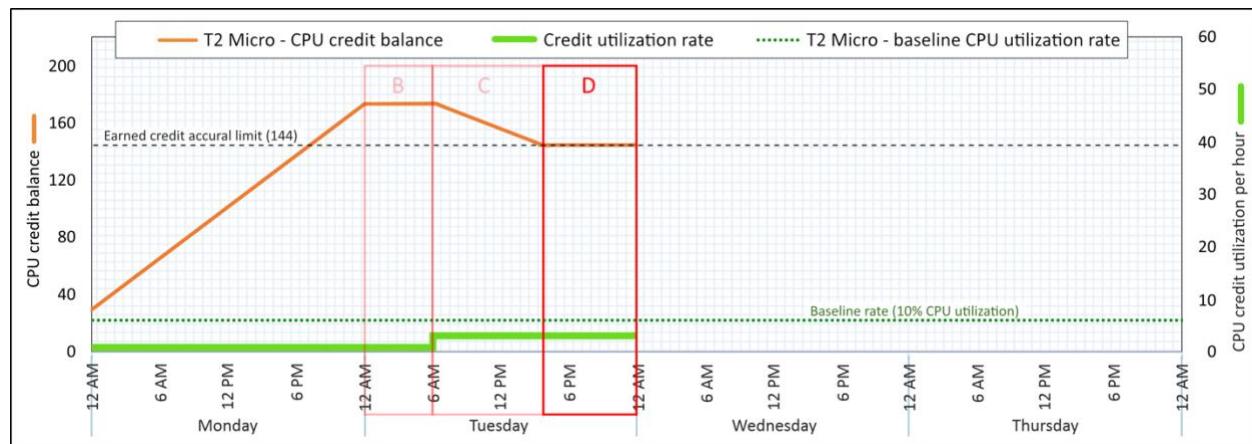
<b>Credit Spend Rate</b>	3 credits per hour (5% CPU utilization)
<b>Credit Earn Rate</b>	6 credits per hour
<b>Credit Discard Rate</b>	6 credits per hour
<b>Credit Balance</b>	Balance decreases from 174 to 144 credits. (0 launch credits and 144 earned CPU credits)



## Period D — Balance Stable, 24 Hours of Earned Credits

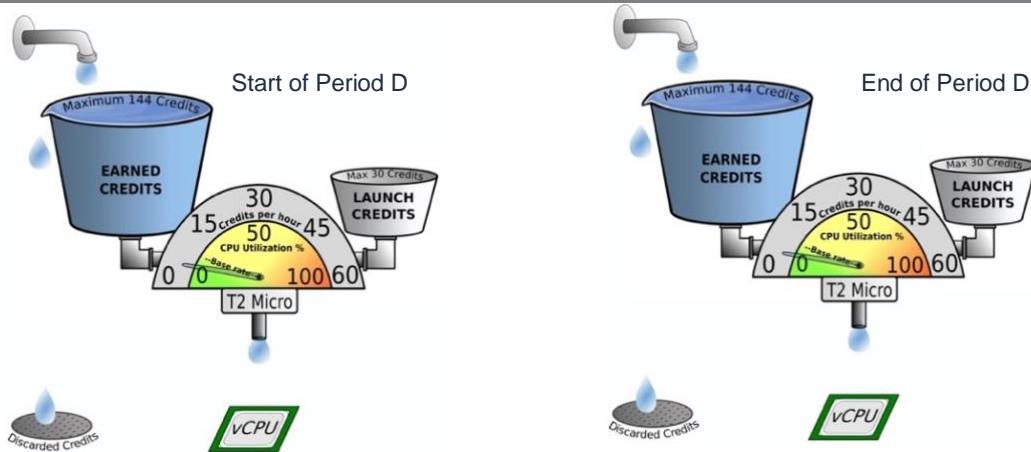
In period D, the instance continues to consume CPU credits at a rate of 3 credits per hour (50% of the credit earn rate) as it did in period C.

The credit balance contains the maximum number of earned credits (144). Half of the newly earned CPU credits are being spent while the other half are being discarded. Therefore, the balance now plateaus at 144 credits instead of at the 174 credit level seen in period B because there are no longer any launch credits in the credit balance.



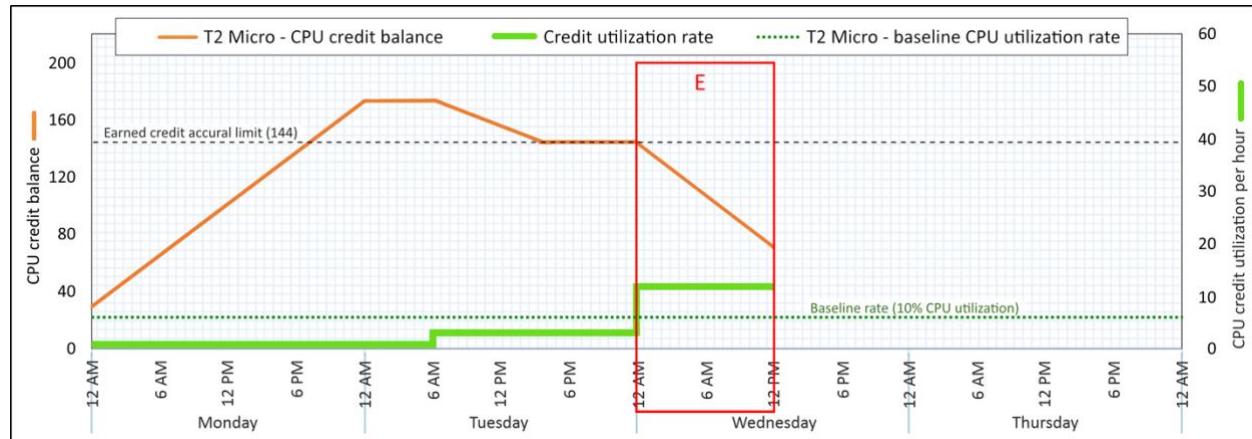
### Period D (duration 8 hours) 4 PM Tuesday – 12 AM Wednesday

<b>Credit Spend Rate</b>	3 credits per hour (5% CPU utilization)
<b>Credit Earn Rate</b>	6 credits per hour
<b>Credit Discard Rate</b>	3 credits per hour
<b>Credit Balance</b>	Balance is stable at 144 credits. (0 launch credits and 144 earned CPU credits)



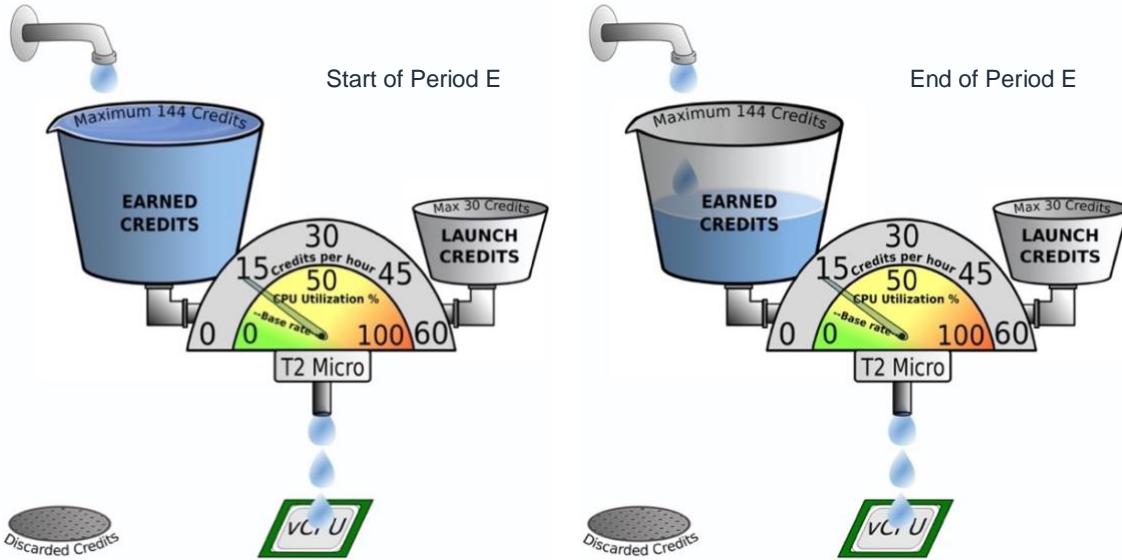
## Period E — Spending Earned Credits

In period E, the instance is consuming CPU credits at a rate of 12 credits per hour (200% of the credit earn rate). The credit balance decreases at a rate of 6 credits per hour from 144 to 72 credits.



### Period E (duration 12 hours) 12 AM Wednesday – 12 PM Wednesday

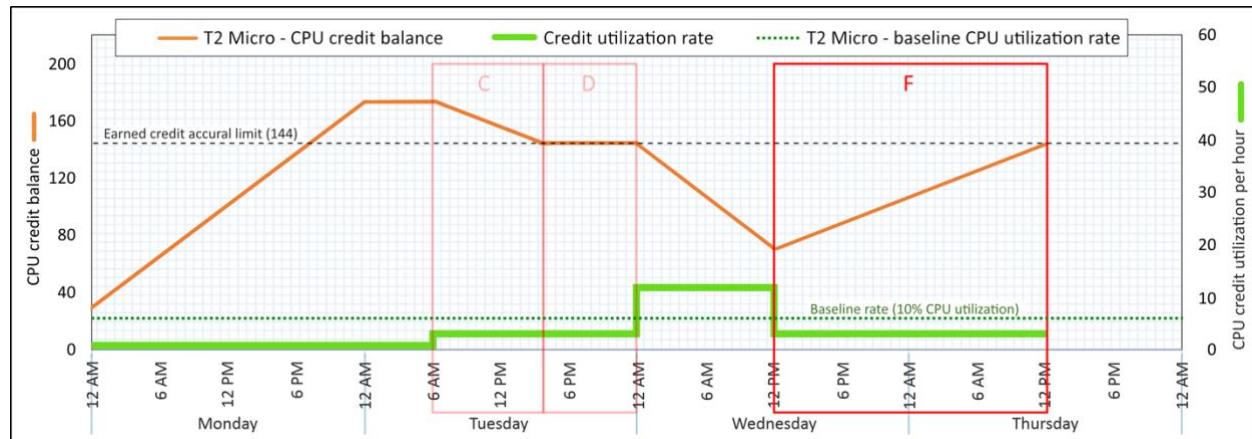
<b>Credit Spend Rate</b>	12 credits per hour (20% CPU utilization)
<b>Credit Earn Rate</b>	6 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour
<b>Credit Balance</b>	Balance decreases from 144 to 72 credits. (0 launch credits and 72 earned CPU credits)



## Period F — Accruing Earned Credits

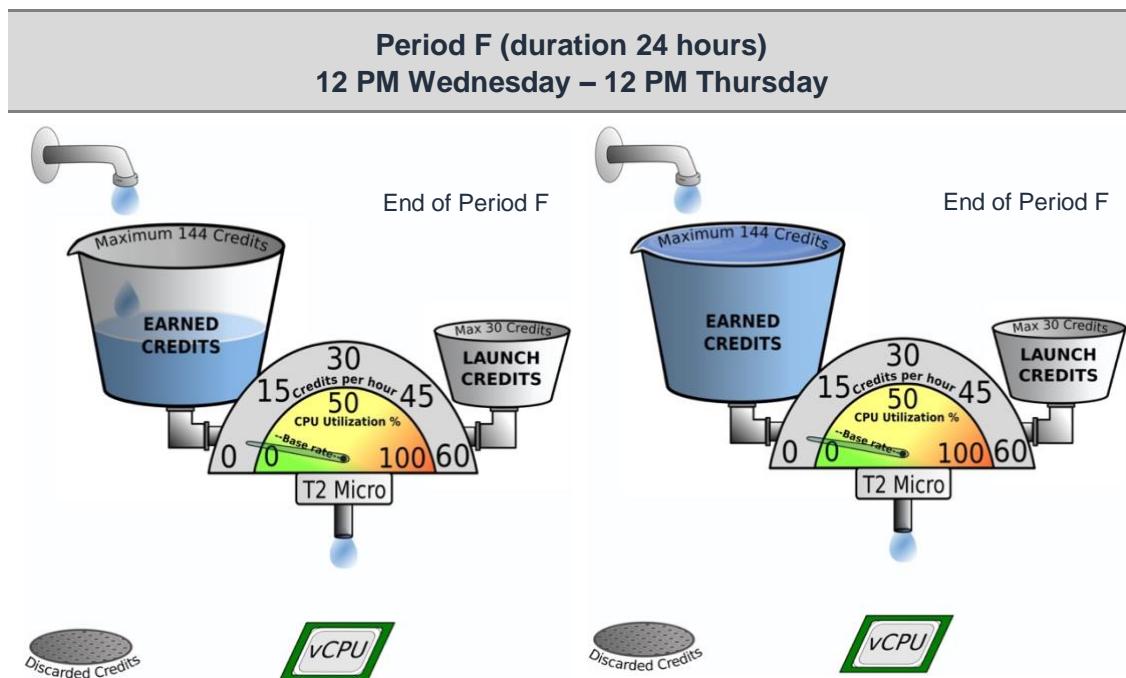
In period F, as in periods C and D, the instance is consuming CPU credits at a rate of 3 per hour (50% of the credit earn rate). The credit balance decreased during period C, was stable during period D, however it increases in period F. Why is that?

- In period C, the credit balance contained launch credits in addition to the maximum number of earned credits. The launch credits were being spent and all of the newly earned and unspent earned CPU credits were being discarded.
- In period D, the credit balance contained the maximum number of earned credits. Half of the newly earned CPU credits were being spent while the other half were being discarded.
- In period F, the number of earned CPU credits is under the 24 hour maximum (144). No credits are being discarded, half of the newly earned CPU credits are being spent while the other half are being accrued in the credit balance. This results in the overall credit balance increasing at half of the earn rate.



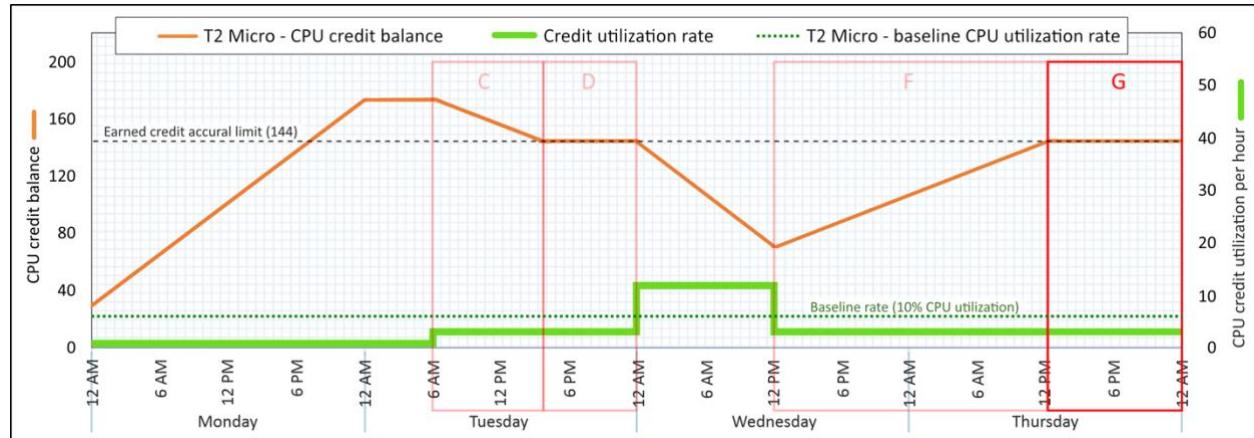
### Period F (duration 24 hours) 12 PM Wednesday – 12 PM Thursday

<b>Credit Spend Rate</b>	3 credits per hour (5% CPU utilization)
<b>Credit Earn Rate</b>	6 credits per hour
<b>Credit Discard Rate</b>	0 credits per hour
<b>Credit Balance</b>	Balance increases from 72 to 144 credits. (0 launch credits and 144 earned CPU credits)



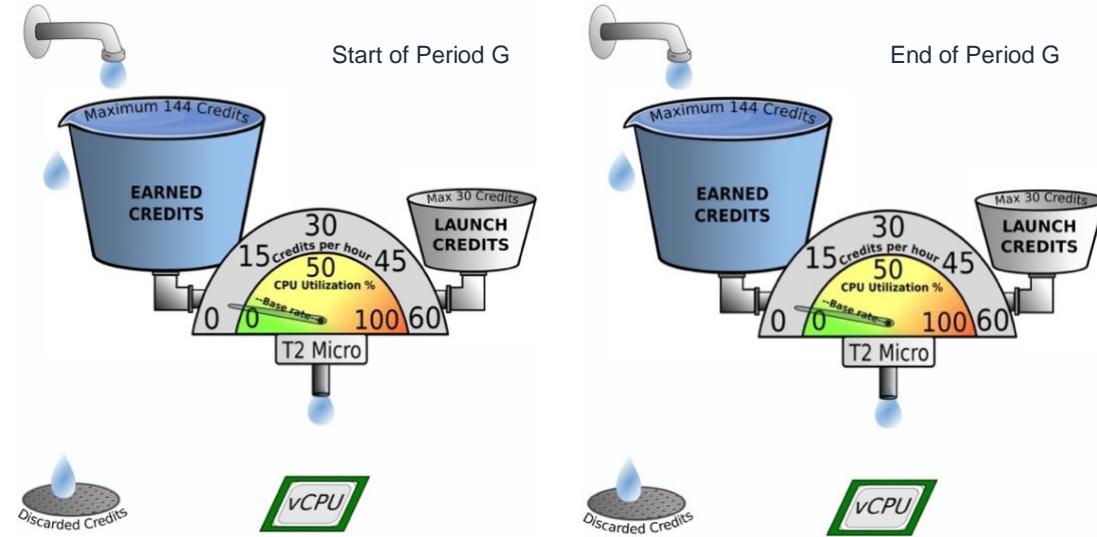
## Period G — Balance Stable, 24 Hours of Earned Credits

In period G, the instance continues to consume CPU credits at a rate of 3 per hour (50% of the credit earn rate), which is the same as periods C, D, and F. However, because the credit balance contains the maximum number of earned credits, any freshly earned but unspent CPU credits are discarded.



### Period G (duration 12 hours) 12 PM Thursday – 12 AM Friday

<b>Credit Spend Rate</b>	3 credits per hour (5% CPU utilization)
<b>Credit Earn Rate</b>	6 credits per hour
<b>Credit Discard Rate</b>	3 credits per hour
<b>Credit Balance</b>	Balance is stable at 144 credits. (0 launch credits and 144 earned CPU credits)



## Comparing T2 Instance Sizes with Identical Workloads

In this section, we will be repeating the same workload (green line) on different sizes of T2 Standard instances to illustrate the effect that different CPU credit earn rates have on the CPU credit balance.

All three of the instances, `t2.nano`, `t2.micro`, and `t2.small`, have a single vCPU and are allocated 30 launch credits. The instances have different CPU credit earn rates with the maximum earned CPU credit accrual limits being 72, 144, and 288 credits, respectively. Larger instance sizes have larger maximum credit balances.

If the same workload is repeated across the three instances, the credit balance changes will differ due to the different earn rates that are offsetting the same spending rate.

## Scenario 1: Consuming CPU Credits at Different Rates

In this scenario, the first utilization period had a vCPU utilization rate of 40% (24 credits per hour) which consumed a total of 100 CPU credits over the 250-minute duration of the period. The change in credit balance depends on the instance size:

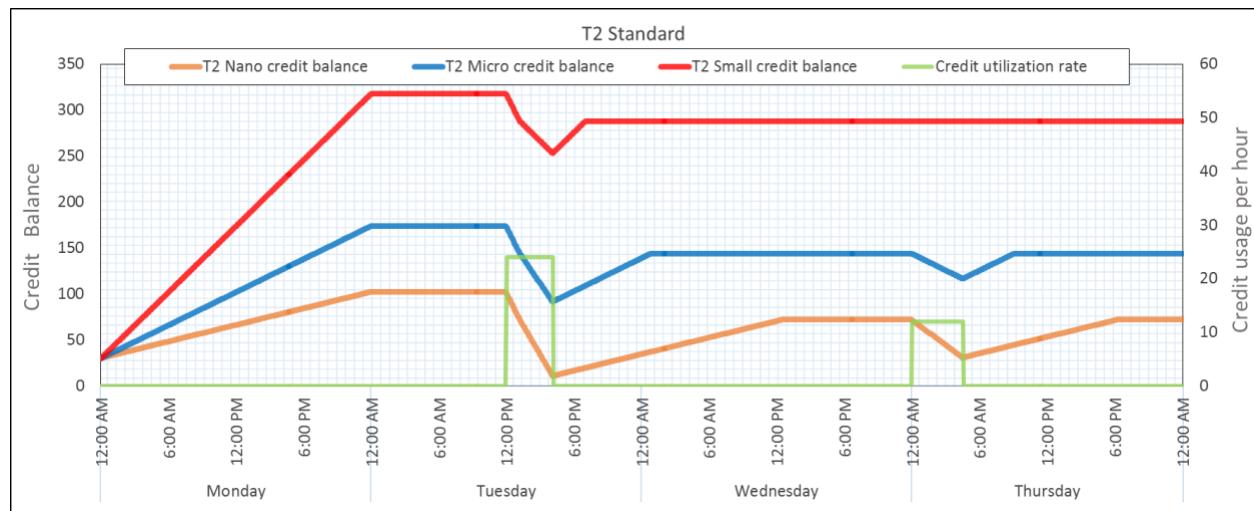
**t2.nano** — credit balance decreased by approximately 91 credits

**t2.micro** — credit balance decreased by approximately 82 credits

**t2.small** — credit balance decreased by approximately 65 credits

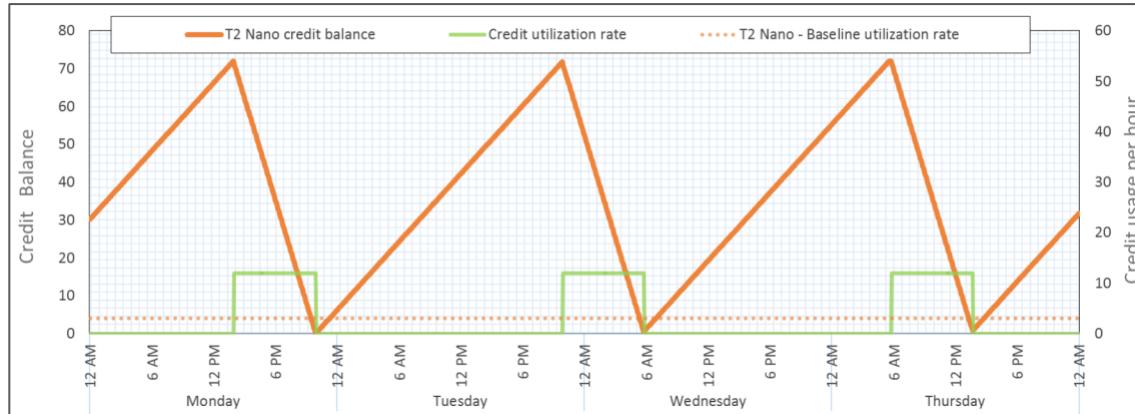
In the second utilization period the difference in the credit balance depletion rate is more apparent. The vCPU utilization rate of 20% (12 credits per hour) is equal to the CPU credit earn rate of a **t2.small** instance, so its credit balance does not decrease.

However, the credit balances for the smaller instances do decrease.

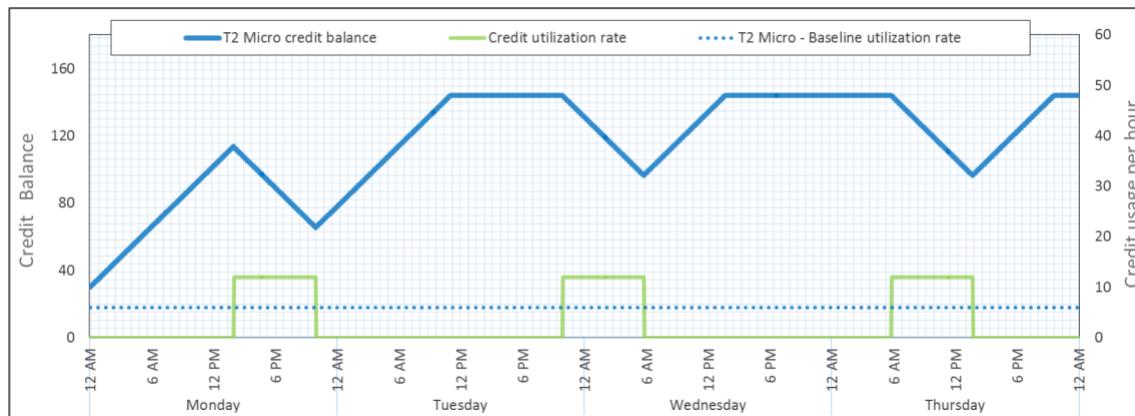


## Scenario 2: Consuming 72 Credits Every 24 Hours

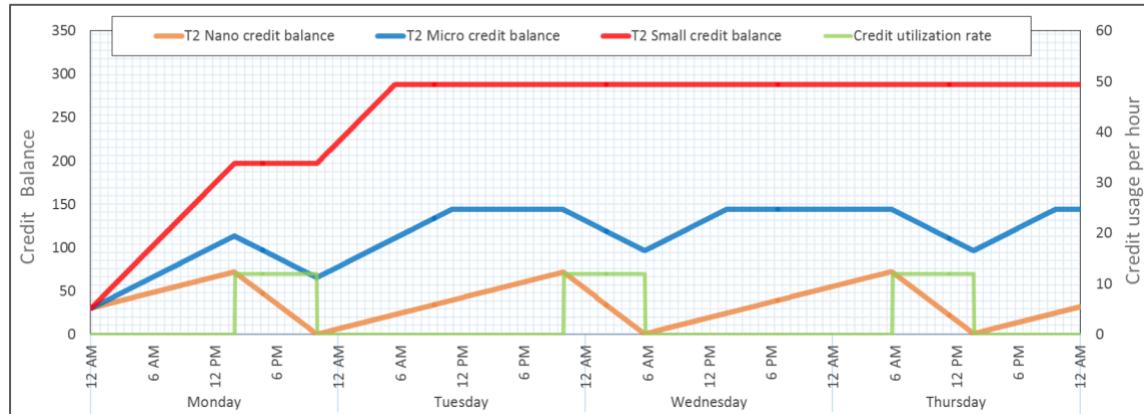
**t2.nano** — The daily credit utilization of 72 credits drains the entire credit balance of the instance during the 24-hour period.



**t2.micro** — The daily credit utilization of 72 credits partially depletes the credit balance during the 24-hour period.

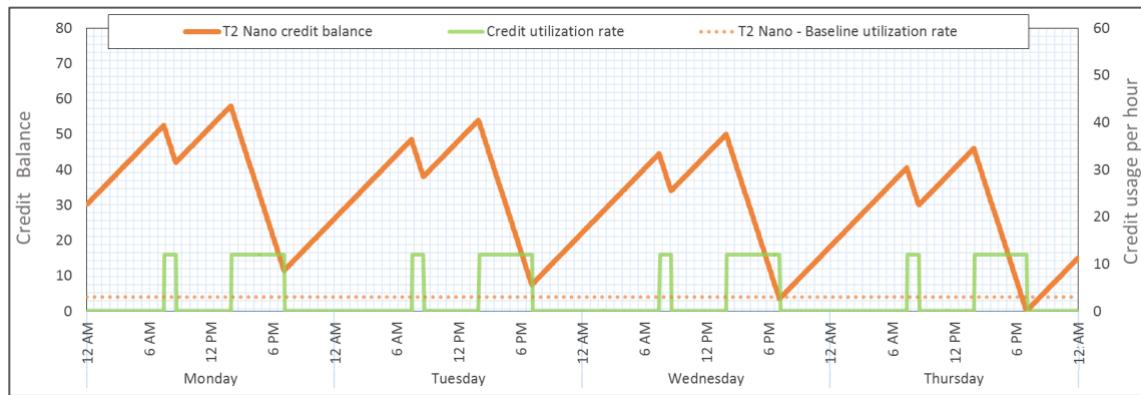


**t2.small** — The peak credit usage rate is lower than the credit earn rate of a t2.small instance, therefore the credit balance (red line) remains stable. This graph shows all three instance sizes for comparison.

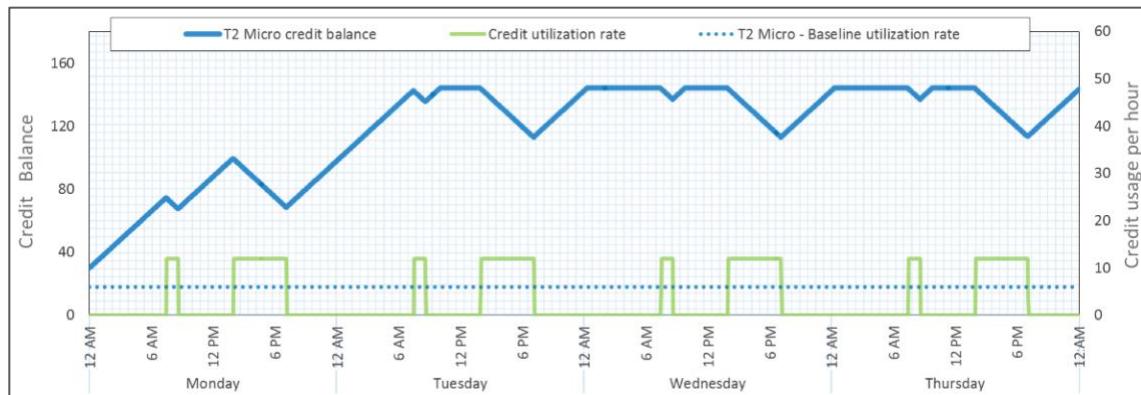


## Scenario 3: Consuming 76 Credits Every 24 Hours

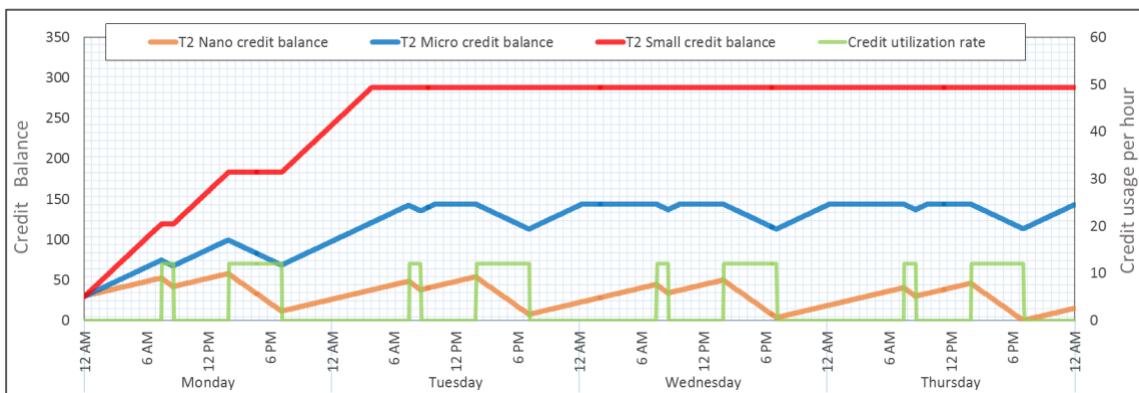
**t2.nano** — The daily credit usage rate exceeds the daily CPU credit earn rate. During periods of low CPU credit utilization, the balance is partially replenished. The credit balance will eventually be depleted over time.



**t2.micro** — The daily credit usage rate is lower than the daily CPU credit earn rate. During periods of low credit utilization, the credit balance is fully replenished.



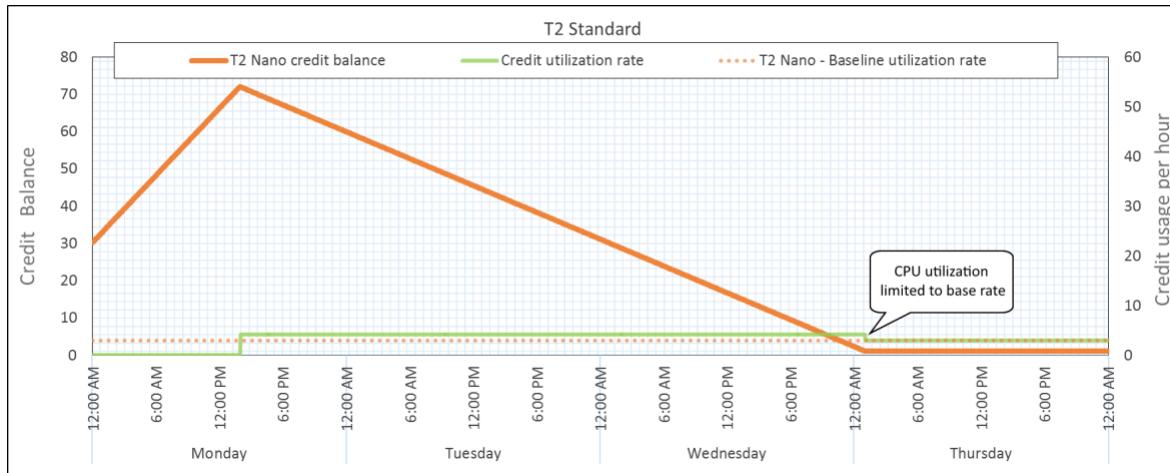
**t2.small** — The peak credit usage rate is lower than the CPU credit earn rate for a t2.small instance, therefore the credit balance (red line) remains stable. This graph shows all three instance sizes for comparison.



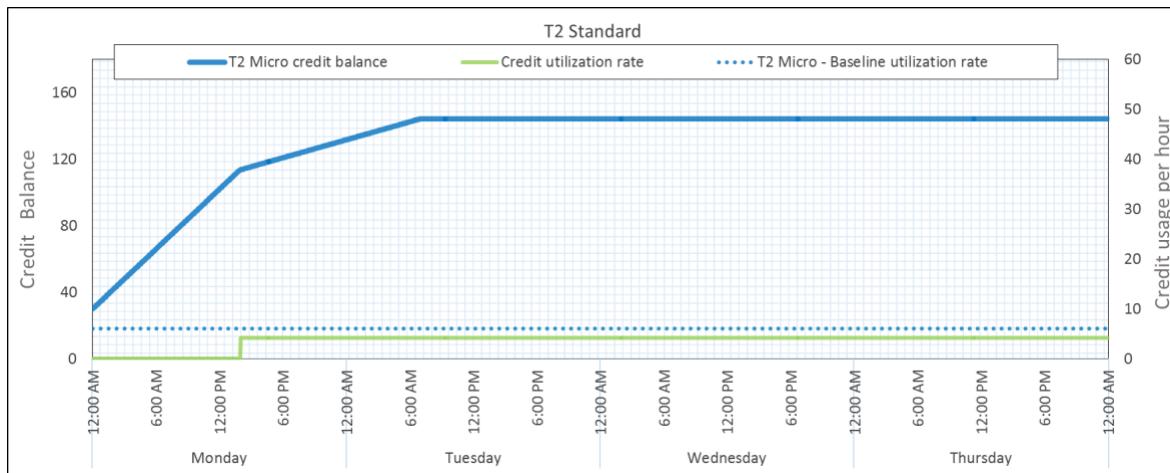
## Scenario 4: Steady and Gradual Depletion of Credit Balance

**t2.nano** — The 7% CPU utilization workload starts 14 hours after the instance is launched and consumes CPU credits at a rate of approximately 4 per hour. The spend rate is higher than the earn rate of 3 credits per hour for a t2.nano, therefore the credit balance gradually decreases.

The credit balance is depleted approximately 72 hours after launch at which point the maximum attainable CPU utilization is restricted to the base rate of 5% (3 credits per hour).

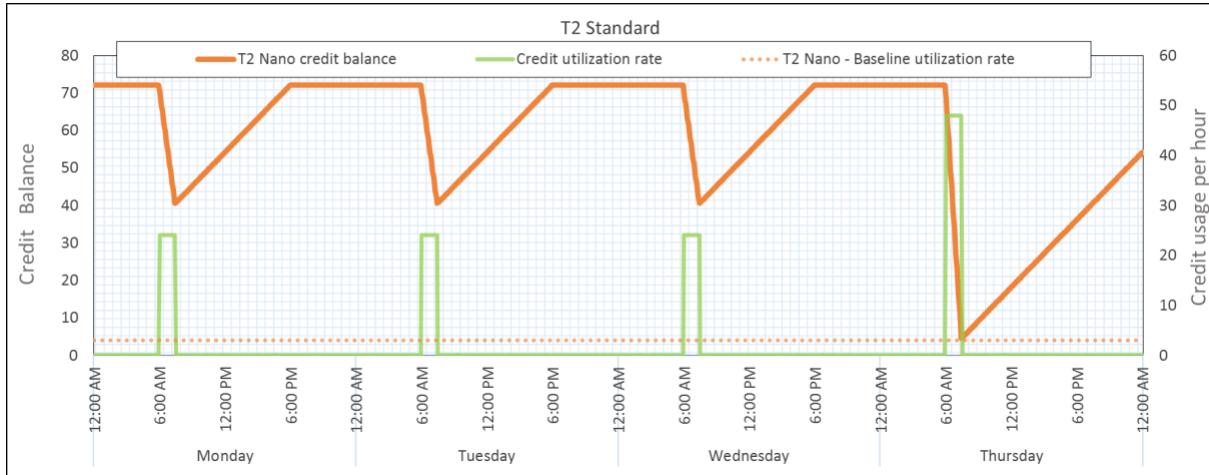


**t2.micro** — The 7% CPU utilization (approximately 4 credits per hour) workload is lower than the base earn rate of a t2.micro instance (10% or 6 credits per hour). Therefore, the credit balance does not decrease and the workload can continue at this utilization rate.



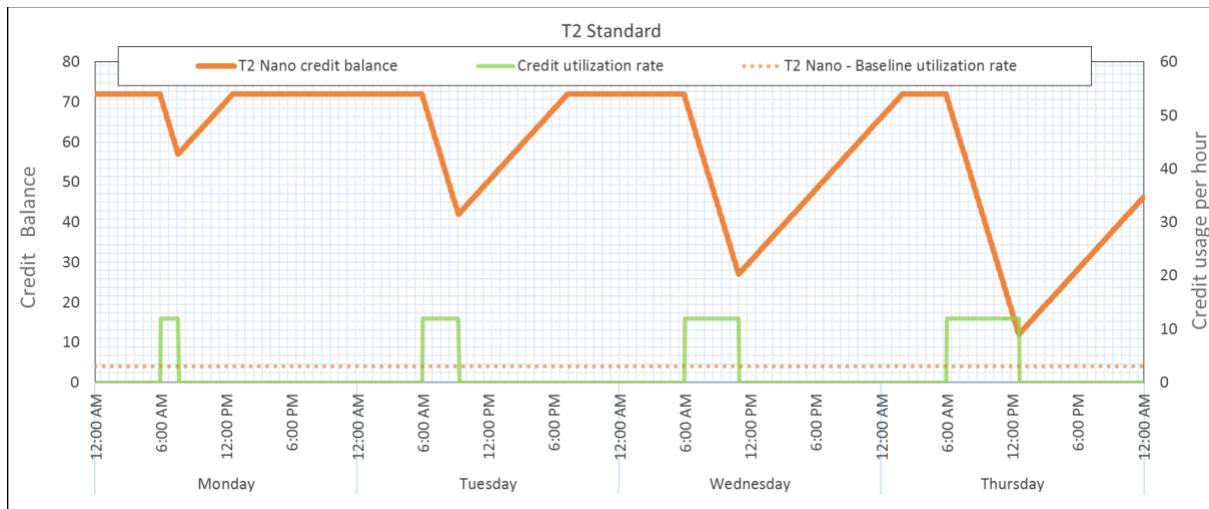
## Scenario 5: Variable CPU Utilization Rate

In this scenario, the duration of the daily workload varies. On Thursday, the workload increased to the point where it almost depleted the credit balance.



## Scenario 6: Variable CPU Utilization Duration

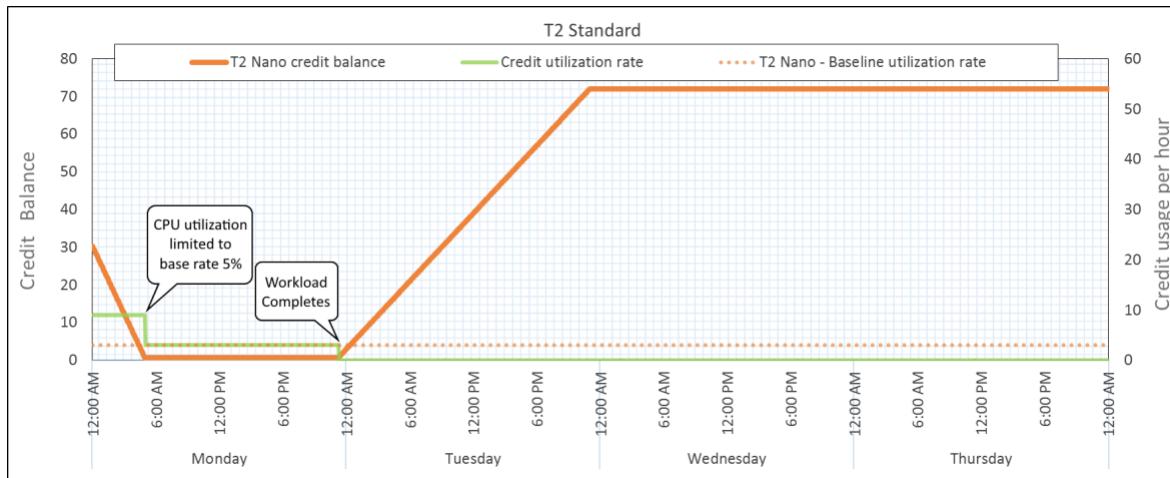
In this scenario, the duration of the daily workload is slowly and gradually increasing. If the workload continues to increase in this manner, it might result in a depletion of the credit balance.



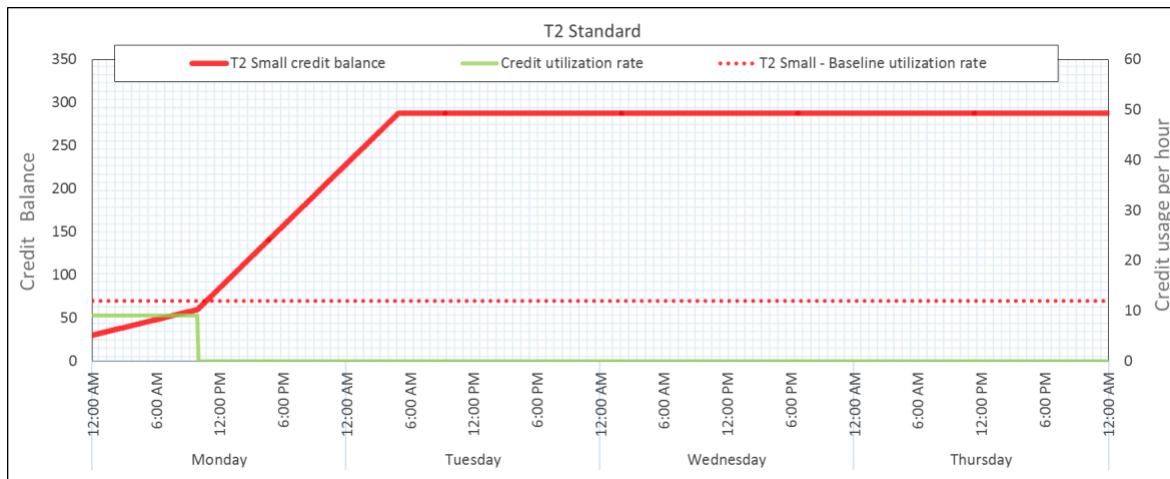
## Scenario 7: Consuming CPU Credits Immediately After Launch

A total of 99 CPU credits are required to complete this workload—ideally at a rate of 9 credits per hour.

**t2.nano** — The workload, running at a rate of 9 CPU credits per hour, consumes the 30 launch credits in approximately 3.3 hours and then begins to consume accrued earned CPU credits. Approximately 5 hours after launch, the credit balance is depleted and the maximum attainable CPU utilization is restricted to the base rate (5%). At this reduced utilization rate, the workload is restricted and requires approximately 23 hours to complete.



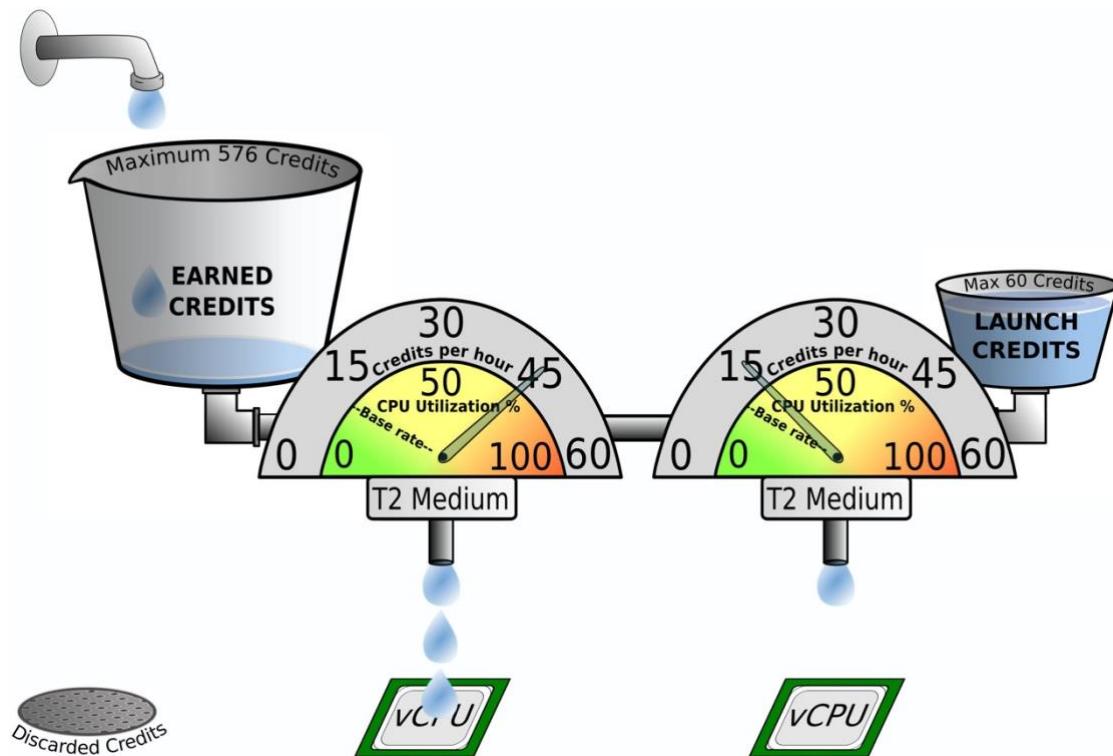
**t2.small** — The workload, running at a rate of 9 CPU credits per hour, has a lower spend rate than the base rate for a **t2.small** instance of 12 credits per hour. The workload can run unrestrained and requires approximately 11 hours to complete.



## Instances with Multiple vCPUs

T2 instance sizes larger than `t2.small` have more than 1 vCPU. The individual vCPUs consume credits from the single credit balance based on their individual CPU utilization rates. The CPU credit utilization rate for an instance is the aggregate of the credit utilization rate across all of the vCPUs.

In the following example, one vCPU is consuming 45 credits per hour while the other vCPU is consuming 15 credits per hours. Therefore, the total credit utilization for this instance is 60 credits per hour.



**Note:** A `t2.medium` or `t2.large` instance with 2 vCPUs can consume up to 2 CPU credits in 1 minute. A `t2.xlarge` instance with 4 vCPUs can consume up to 4 CPU credits in 1 minute. A `t2.2xlarge` instance with 8 vCPUs can consume up to 8 CPU credits in 1 minute.

An instance's specified baseline % rate is based on a single vCPU. For example a `t2.medium` baseline rate is specified as 40%. Which can equate to 1 x vCPU @ 40% utilization or 2 x vCPUs @ 20% utilization.

## Conclusion

Having an in-depth understanding of how the T2 Standard instance CPU credit system works will help you decide if this particular Amazon EC2 instance type is the best match for your workload. If so, this knowledge will assist you with optimizing your workload and obtaining the best cost and performance.

## Contributors

Contributors to this document include:

- Seamus Murray, Amazon Web Services

## Further Reading

For additional information, see:

- [AWS Documentation: Burstable Performance Instances<sup>1</sup>](#)

## Document Revisions

Date	Description
February 4, 2019	First publication
March 1, 2019	Second publication

## Notes

<sup>1</sup> <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/burstable-performance-instances.html>