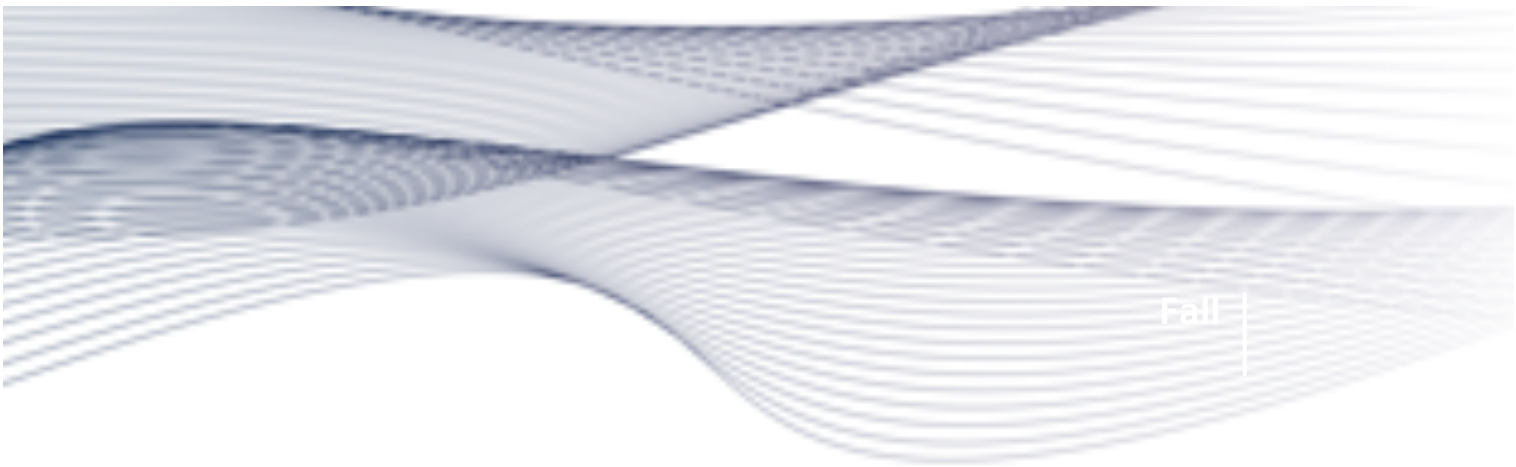


ARISTA WHITE PAPER

Arista EOS: Smart System Upgrade

KEEPING PACE WITH DATA CENTER INNOVATION

Deploying and taking advantage of new technology is top of mind for most organizations. Balancing the business benefits of adopting a rapid pace of innovation with the associated risks is a constant struggle, one most are losing. A major inhibitor to technology adoption is the ability to transparently insert new technologies into existing facilities without adversely impacting critical applications. Changes to the data center infrastructure can affect large numbers of applications and have the potential to cause significant service disruptions and extended outages. These concerns have had an almost paralytic effect on the adoption of beneficial new technologies. Unfortunately, operating with outmoded equipment can severely impact your bottom line causing further erosion in already shrinking budgets.

This paper will review the challenges with traditional software upgrade approaches and explain how Arista EOS® Smart System Upgrade provides a more reliable solution to the software upgrade process.

THE ISSUE WITH IN SERVICE SOFTWARE UPGRADES

When it comes to conducting data center network maintenance, the focus has been on upgrading software versions without impacting the flow of application traffic. One method of upgrading a particular network node is called in-service software upgrade (ISSU) and relies on highly complex system software that has proven unreliable and operationally difficult to implement. While the feature is sold with the idea that you can seamlessly move from one version of network OS to the next, the actual implementation is markedly different.

The traditional single-box ISSU approach has historically been burdened by the significant amount of complex software development required. ISSU-specific code had to be written to account for all possible checks and balances while the system attempts to maintain and convert all hardware and software state as it jumps between two different versions of software. This challenge is increased as the feature-set increases, features interact with each other, and the associated state information grows. The side effect of this complexity is that corner cases are common and it is not possible to address all scenarios. Even additional testing cycles, which further add an 'ISSU tax' to the release timeline, are not guaranteed to catch all cases.

Often times, the effects of an ISSU-related issue can take time to be exposed, many times lingering beyond the maintenance or change-control window associated with the particular upgrade. This adds additional concern, as it is difficult to know for certain if the upgrade was completed successfully or not.

A by-product of the fragile nature of this model is the ISSU version compatibility matrix. Since there are could be so many combinations of 'upgrade from' and 'upgrade to' software versions, this bounded matrix is needed to limit the test cases to a very few number of these upgrade combinations in an attempt to manage the complexity. This adds yet another restriction to the practical usefulness of ISSU as there is no guarantee that a customer's particular upgrade path will be supported in this matrix. The introduction of a multiple step upgrade procedure might be possible but only make the situation worse as the inherent device state is now subject to additional churn and an increased likelihood of instability.

For all of these reasons, the resulting ISSU feature is an incomplete solution that introduces unnecessary complexity and suffers from many caveats.

The intent behind ISSU is a good one, however. The need to provide a seamless upgrade of the network infrastructure without impact to the forwarding plane is still a problem in need of a better solution.

LESSONS LEARNED FROM THE CLOUD

The rise of cloud networking has driven many recent networking innovations. The increased network efficiency brought by focus on automation, orchestration, and programmability techniques is in large part due to the benefits seen in cloud network environments and then re-applied at varying scale to other network environments. So similarly, we can apply the lessons learned from the cloud network environments on their software upgrade approach.

Cloud networks are designed for simplicity, flexibility and repeatability. Built on basic design principles of spine/leaf designs with multiple active paths across the network, no single node in a cloud-based design should be so important that it couldn't be taken out of service for software maintenance. The perspective is raised from the availability of a 'box' to the availability of the overall 'fabric'. As a result, single-box software upgrade solutions, like ISSU, are not only complex and unreliable, but they explicitly are not required (and are correspondingly not used.) By moving the solution to the network level, the solution can leverage the existing set of rich networking tools and mechanisms to achieve the goal. This principle can be applied to datacenter networks of all types and sizes.

INTRODUCING ARISTA SMART SYSTEM UPGRADE

Smart System Upgrade (SSU) is an EOS network solution designed to allow maintenance to be performed on any infrastructure element, without adversely impacting application traffic flow. SSU is built on the same cloud networking principles of simplicity, flexibility, and repeatability. SSU uses a simplified stateless approach to ensure that application traffic flow is not impacted.

Combining native Arista Extensible Operating System (EOS) functionality and direct integration with other applications and infrastructure components, SSU takes a broader network perspective to software maintenance by allowing a network element to be transparently removed or added while traffic is either diverted or impact is altogether avoided. Designed to be a complete solution for data center infrastructure maintenance, Arista's SSU provides the following key benefits:

- Intelligent insertion and removal of network elements, customized to the spine role or the leaf role
- Programmatic upgrade to new software releases without causing systemic outages
- Open integration with all application and infrastructure elements
- Simplified solution: Intentionally avoids the complexity of heavy state maintenance and state conversion process needed with other approaches

Data center operations teams need more intelligent tools and extensible feature sets to manage today's 'always-on' data center infrastructures. Arista EOS provides the foundation for innovation, driving down operational cost while simultaneously increasing operating uptime.

CORE OPERATING SYSTEM

As data centers continue to grow and the services offered increasingly require 24x7 availability, the ability to transparently add and maintain these services is must-have functionality. Arista EOS significantly leapfrogs earlier network operating systems by providing the following key capabilities and benefits.

- In-service software patching - Ability to upgrade individual processes without system interruption
- Software Fault Containment (SFC) - Faults are contained to a single software process
- Stateful Fault Repair (SFR) – Restart of processes without the need to rebuild state information

Arista EOS delivers these benefits with a unique multi-process state sharing architecture that separates state information from the process. This reflects Arista's core software design philosophy and enables fault recovery and real-time software updates on a fine-grain process basis without affecting the running state of the system. In fact, EOS specifically avoids the complications of trying to maintain large amounts of software state.

Arista EOS architecture does not suffer from reliance on synchronous and in-order inter process communication infrastructure to trigger state machine changes. Instead the architecture is designed to be end state-driven with the System Database always carrying the final accurate state at any time. The advantage of this as opposed to other message-driven network operating systems is that it dramatically simplifies the upgrade process logic by eliminating the need to preserve and replay messages in order to achieve accurate process state. Furthermore, this allows EOS to provide consistent reliability even under heavy control plane and network loads without any degradation.

By providing features designed specifically for the 'always-on' infrastructure, Arista EOS focuses on intelligently maintaining your infrastructure.

NETWORK DESIGN

Arista's Universal Cloud Network, a non-blocking flexible architecture based on industry standard protocols and two-tier Spine/Leaf and single-tier Spline™ designs. In order to deploy and maintain this environment while providing maximum uptime, Arista has developed several differentiated features that are specifically designed for the particular place in the network.

Spine SSU with Maintenance mode

Maintenance mode is the Smart System Upgrade mechanism for the spine layer. Maintenance Mode focuses on operational procedures that take place to remove a device from the network by gracefully rerouting traffic, taking advantage of the inherent characteristics of redundancy through network design and distributed protocols. Intelligently removing devices from the forwarding topology and alternating system upgrades provides a rapid and pragmatic method to move directly between code versions without service impacting outages (see figure 2).

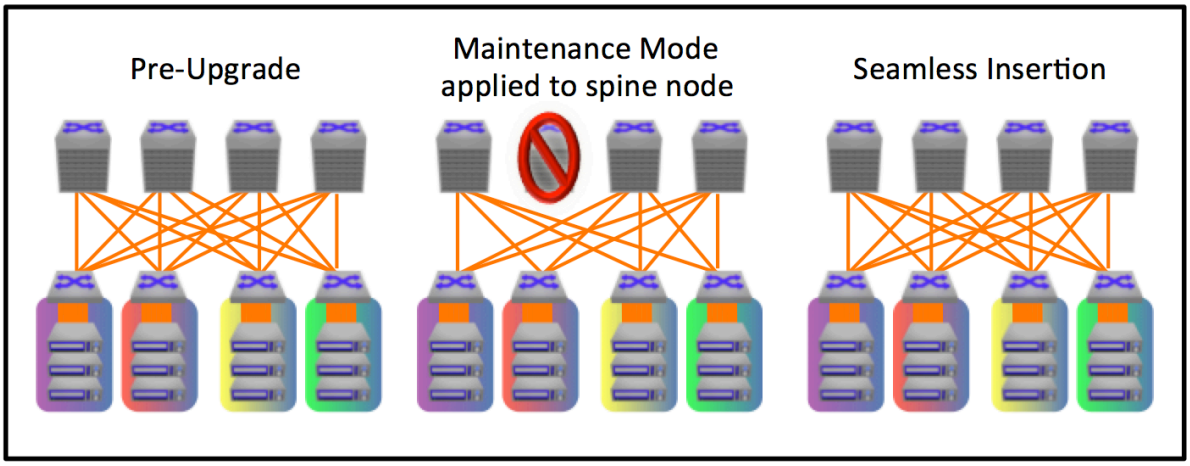


Figure 2: Arista Spine SSU with Maintenance Mode

Maintenance mode is applied to the device to-be-upgraded by invoking a single EOS command. This command will use standard route-weighting mechanisms to gracefully remove the node from attracting network traffic. The device will enter into maintenance mode only when all traffic has been drained on its configured interfaces. At this point, the operator can perform a clean software upgrade or any other system maintenance (hardware module swap, re-cabling activity, etc) on the particular device without impacting production traffic. Once the maintenance has

completed, the operator can then disable the maintenance mode to have the device gracefully re-inserted into the topology.

Leaf SSU

The leaf network layer is the most critical layer from an availability perspective. Often a single point of access for network-attached hosts, the challenge of updating the software is only increased for the network operator. The leaf layer cannot benefit from the same network redundancy approach that the spine layer leverages. As a result, the approach to solving leaf upgrades needs to be different from the spine upgrade.

Arista EOS Leaf SSU provides a solution for upgrading these leaf devices in a hitless manner. The solution takes advantage of Arista EOS' ability to rapidly reload the operating system while providing the option of maintaining the forwarding state of all ASICs, ensuring that traffic is still forwarding while the control plane is updated. The reliable forwarding capabilities of Arista's EOS operating system allow the downstream network elements to continue to operate during the upgrade process without any knowledge that one of the primary upstream nodes is not available.

In addition to continuous forwarding within the forwarding plane, Leaf SSU enables Arista EOS to continue to participate with steady-state heartbeats for protocols like LACP and STP BPDUs to downstream devices while the control plane is being rebooted such that downstream devices aren't even aware that there is an software version change going on within the network device.

Combined, these enable the network to remain operationally available throughout the process, maximizing system uptime and network resource availability.

Zero Touch Provisioning (ZTP) / Zero Touch Replacement (ZTR)

When initially deploying or replacing a switch, ZTP can be used to configure a switch without user intervention; it is as simple as rack, connect and power-on. Built to fully leverage the power of Arista EOS, ZTP provides a flexible solution, provisioning the network infrastructure without requiring a network engineer to be present at the time of initial install or replacement. ZTR provides ZTP functionality for replacement of devices, with policy/configuration tied to the device location in the network and not to its MAC address.

The ZTP process runs by default at system boot, and based on administrator preferences can ensure the proper software image is installed and complete auto-provisioning of the switch is performed. Utilizing well-understood standards-based protocols, ZTP can leverage the same services that your servers already use, no retraining required. Used as a key component of a cloud data center, Arista ZTP/ZTR will help ensure that any switch is deployed running the proper version of software and with the proper configuration, and can be gracefully inserted into the network.

With tools like Spine SSU with Maintenance mode, Leaf SSU, and ZTP, the network operations team now has purpose build tools for each network role that enable system upgrades while minimizing the risk of service-impacting outages. Arista's SSU approach eliminates all of the complex message synchronization and backward compatible data structures that have plagued legacy operating system approaches to this issue. Further, the SSU implementation simplifies the upgrade process by allowing for complete changes between operating system versions without complex compatibility matrixes. The ability to transparently upgrade between any version minimizes certification-testing efforts, upgrade windows and potential adverse effects when compared to previous methods.

VIRTUALIZATION AND CLOUD MANAGEMENT

Advanced Event Management (AEM)

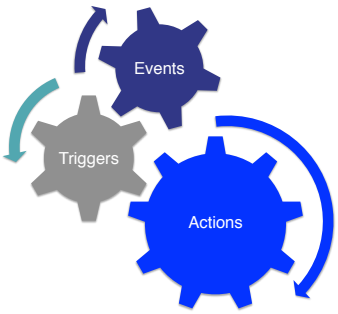
Advanced Event Management (AEM) provides a set of tools that can augment and enhance the capabilities of Smart System Upgrade, providing key indicators and actionable items for EOS resources in real time.

Leveraging Event Manager, SSU can react and respond to various activities happening in the network. For instance, you can be notified of the link status or protocol state of an interface or set of interfaces. Upon receiving the notification, a set of predetermined tasks can be performed. Tasks might include things such as changing the links IGP metric or informing the SSU workflow to move to the next step.

Event Monitor captures, catalogs and stores all moves, adds and deletes to ephemeral state tables in EOS. Using the information that is captured by Event Monitor, SSU can validate expected changes in these state tables (e.g., tables for routing, MAC, multicast routing and ARP). This validation helps ensure pre- and post-forwarding state is maintained and any discrepancies are highlighted to the operations team for validation or further troubleshooting.

Automation Integration with eAPI

The foundation that enables EOS to effortlessly carry out workflow automation is eAPI. Providing a pragmatic interface to EOS, eAPI allows the network engineering and operations team to build robust automation scripts. Unlike many operational interfaces, eAPI utilizes the same commands and structure as the command line interface, simplifying the building of tools and scripts that interact with the system. These automation scripts return structured data from the target EOS device. This structured data, in a JSON format, is easily parsable by any scripting language, providing the engineer or operator maximum flexibility in choice of tools.



Specific to Smart System Upgrades, eAPI allows customers to define customized workflows to perform specific actions to control network resource state in EOS, as well as connect to external systems for validation and reporting. Even more so, Arista EOS is integrated with all of the major automation frameworks available today. Integration with tools like Puppet, Chef and Ansible enhance the native capabilities provided by EOS and eAPI to deliver customized workflows that focus on adding additional workflow intelligence to perform system upgrades. By leveraging the programmatic interfaces available in EOS, network engineering and operations teams can focus on developing the right operational plan as opposed to spending their time adapting to a specific toolset.

System Snapshot

One of the key requirements to reliably returning a network device back into operation is validating that the system state has normalized. In order to perform such an operation, operations teams typically compare state tables, along with the output from a variety of “show” commands to validate that the system has returned to its expected operational state. The individual or team executing this process is typically aggregating the data, in some cases on the fly, as output from various commands are provided. This process can be cumbersome and is prone to human error.

```
z#snapshot compare base routes

     Snapshot Summary

Metric                           Snapshot A        Snapshot Z       Changed
-------------------------------- ----------------- ---------------- -------
Baseline Summary
   Num of Interfaces               8                 8
   Num of Vlans                    12                113              *
   Num of LLDP Neighbors           29                29
   Num of MAC Entries              4                 0                *

Interfaces
   Num of Eth Interfaces           7                 7
   Num of Eth Interfaces Up        7                 7
   Num of Eth Interfaces Dwn       0                 0
   Num of Eth Interfaces Other     0                 0

   Num of Vlan Interfaces          0                 0
   Num of Vlan Interfaces Up       0                 0
   Num of Vlan Interfaces Dwn      0                 0
   Num of Vlan Interfaces Other    0                 0

   Num of Lag Interfaces           0                 0
   Num of Lag Interfaces Up        0                 0
   Num of Lag Interfaces Dwn       0                 0
   Num of Lag Interfaces Other     0                 0
```

System Snapshot, implemented using Arista eAPI, provides the capability to automatically collect the information to validate system state both before and after a change management window. Through the use of value difference checking, Snapshot can quickly scan through hundreds or even thousands of data points to ensure validation and congruency from one state transition to the next. The result of this scan can then be reported to the operations team in the form of a summary report that can be used to quickly pinpoint areas of concern or provide validation that the system state has normalized.

Partner Integration

Many of the features available in EOS provide an opportunity to bring more intelligence into handling system upgrades, as the extensible nature of EOS is not only limited to Arista devices. By taking advantage of the programmability across Arista’s partner ecosystem, SSU can integrate more broadly to extend the intelligent nature of the upgrade process up the stack to include systems and applications.

For instance, Arista has demonstrated intelligent upgrades that work in concert with F5 iControl rules and VMware SDK to signal network resource state changes. This allows devices such as firewalls, load balancers and compute infrastructures to take precautionary measures prior to removing a resource from the forwarding plane. Coordinating system upgrades allows the entire data center infrastructure to be aware of system state throughout the process, allowing systems to take appropriate actions each step along the way.

SUMMARY

Arista’s SSU is designed to address the most challenging IT workflows. Combining protocol-based graceful shutdown, insertion and unique redirection techniques, traffic can be diverted around the device under maintenance without the applications being affected. Building on Arista’s open philosophy, all of these tools are customizable, allowing them to be tailored to your specific operating environment. Now operations teams, network, server and virtualization administrators have the ability to perform seamless maintenance to any infrastructure element faster, with a summary report detailing the changes.

Arista Networks strives to bring industry leading hardware and software capabilities through a robust product portfolio, which is augmented by focused network designs that drive today’s largest data center infrastructures. By leveraging the features that form the foundation of Arista’s EOS operating system, customers are now able to drive

new intelligence into managing and upgrading systems throughout the data center. By combining operational innovation with well-architected, software-defined procedures, Arista Networks is delivering on the promise of smarter systems upgrades, today.

For more information, visit: http://www.arista.com/en/products/eos

ARISTA

Santa Clara—Corporate Headquarters
5453 Great America Parkway
Santa Clara, CA 95054
Tel: 408-547-5500
www.arista.com

Ireland—International Headquarters
Hartnett Enterprise Acceleration Centre
Moylish Park
Limerick, Ireland

Singapore—APAC Administrative Office
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989