LANCOM Tech paper

LANCOM fail-safe stacking

A stack is a group of switches that physically behave as a single device. A LANCOM stack can consist of up to 8 switches. All devices in the stack must be equipped with the same stacking interfaces (ports) and have an identical firmware version. In this specific case, this means that stacks can be formed with the following devices in each case:

- **LANCOM XS-6128QF (via the SFP-DD ports)**
- **LANCOM XS-5116QF and GS-4000 series (via the QSFP+ ports)**
- **LANCOM XS-5110F and GS-4000 series (via the SFP+ ports)**

LANCOM fail-safe stacking thus also supports mixed stacks of fiber-optic and copper switches. Members of the stack can all be interconnected at a single location, e.g. in a shared rack with a Direct Attach Cable (DAC), or they can be decentrally distributed across different locations. If the stack is distributed over two or more locations, the large distances between the stack member switches require the use of optical transceiver modules. For example, the LANCOM SFP-LR-LC25 module connects locations at distances of up to 10 km apart.

A stack can be constructed either as a ring or as a daisy chain. A chain is not considered further in this paper, since the network would be torn into two separate parts if one stack member failed. Two network segments that cannot see or communicate with one another is a worst-case scenario for any network administrator. For this reason, LANCOM recommends the use of the ring topology, as discussed in more detail in this tech paper.

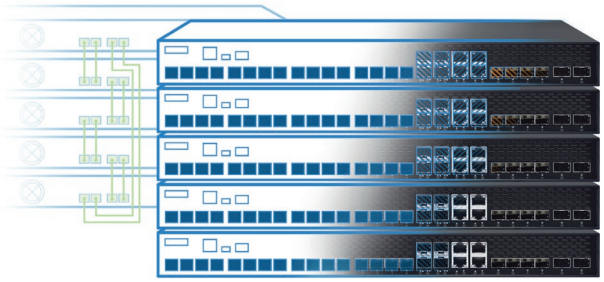

Fig. 1: LANCOM failsafe stacking

Stacking terminology and roles in a nutshell

- **Management switch (also manager, master):**
  Manages the stack and configures the stack's system and interface-related functions.
- **Operational standby (also backup, management standby):**
  The stack member with the highest priority after the management switch. This switch takes over the configuration if the management switch fails.
- **Stack member:**
  A switch in the stack that does not handle management-switch or standby-manager functions.
- **Stack unit:**
  Each device in the stack, including the management switch and operational standby.
- **Standalone unit:**
  A switch that is not part of a stack. The standalone unit can become a management switch if the stacking ports are set up accordingly and another switch is connected. In the same way, a management switch becomes a standalone unit if it loses all connections to other stack members.

Why stacking? – Advantages in switching and routing

Similar to a chassis or blade system, the stacking ports handle all of the data traffic in hardware and with the protocols that are precisely optimized for this.

In every multi-layer network, aggregation switches such as the XS-6128QF distribute the network data to lower-layer access switches. Instead of connecting these via conventional layer-2 spanning tree or layer-3 routing protocols, a better alternative is to connect them to a stack with its stack protocols for the following reasons:

- **Layer-2 simplification**
  All stack links are considered to be internal and are not treated as if they belong to the layer-2 network. This means that these connections are not considered to be closed, but rather to be a continuously available connection. This allows network traffic to be simultaneously transmitted over multiple connections, so maximizing throughput.
- **No layer-3 routing required**
  The intelligent distribution of the data stream within the stack does not require layer-3 routing, since the internal stacking protocols handle the connections as described above.
- **Fast failover and nonstop forwarding**
  Fast detection and link-recovery technologies mean that, in the event of an error, stack connections are transferred to other switches by "hitless failover", i.e. without data loss and without the user being aware of a service failure (also see the following chapter).
- **Non-blocking stacking**
  The total downlink capacity is fully matched by the sum of the stacking capacity, so avoiding any bottlenecks when switching between network ports.

Nonstop forwarding (NSF)

A switch can be understood to be three almost independent functional levels known as the forwarding plane, the control plane, and the management plane. The forwarding plane is implemented completely in hardware and simply forwards the data packets. The protocols on the control plane are intelligent and decide how the forwarding plane handles the packets. This plane decides which data packets can be forwarded and where they should go. The application software, i.e. the stacking protocol on the management switch, in this case can be understood to be the control plane. The management plane, i.e. the user interface, ultimately allows the network administrator to configure and monitor the device.

Nonstop forwarding means that, in case of a power failure, hardware error or software error impacting the management switch (master), the forwarding plane of the stack units continues to forward packets while the control and management planes are rebooted.

If the management switch fails, data streams entering and exiting the stack via network ports on a stack unit other than the management switch are continued with an interruption of no more than one second.

To ensure that the operational standby is prepared for a failover event, the management switch continuously feeds it with status information. For example, any changes to the current configuration are automatically copied to the backup. The MAC addresses stay the same during a nonstop forwarding failover so that the connected network neighbors (e.g. access switches) do not have to relearn them. In the event of a nonstop forwarding failover, the control plane on the operational standby boots from a partially initialized state and applies the status information from the last check-up with the manager. While this control plane is being initialized, the stack cannot react to external

changes such as changes in the network topology. As soon as the control plane is fully functional, it ensures that all internal hardware links are updated. The failover time of the control plane depends on the stack size, the complexity of the configuration, and the CPU speed. When a failover occurs, the management plane (i.e. the GUI) is rebooted, as the information about the new management connections has to be generated all over again.

In order for NSF to be as effective as possible, devices connected to the stack must not route traffic around the device that is rebooting. The LCOS SX 5.x operating system employs three techniques to prevent this:

1. Protocols can distribute the control plane across stack units so that it appears as if they are still functioning during the reboot. The spanning tree protocol and LACP/LAG use this technique.
2. In the event of an error, protocols can enlist the cooperation of stack neighbors using a technique known as "graceful restart". For example, the dynamic routing method OSPF (open shortest path first) uses this technique if it is enabled (also see IP Event Dampening Commands in the CLI manual).
3. After failover, protocols can simply restart if the connected network devices, like access switches, react so slowly that they usually do not detect the failure. The IP multicast routing protocols are a good example of this.

To take full advantage of nonstop forwarding, using an LAG group ensures that layer-2 connections from network devices like access switches are shared between two or more stack units. Similarly, layer-3 routes such as ECMP routes with next hops via physical ports should also be split between two or more stack units. The hardware quickly moves data streams from LAG members or ECMP paths on a failed unit to a stack unit that is still functioning.

Network management and stack configuration

If all devices in the stack are distributed decentrally, they can still be configured as a single unit at a single IP address. A decentralized stack thus behaves exactly like a stack at a single central location. One of these switches acts as a management switch and handles the configuration of the entire stack, thus controlling the configuration of all of the member switches. As an example, a firmware update is installed on the management switch first, and this then performs the updates on the other switches in the stack. The stack only works if all of the switches have the same firmware version. In the unlikely event that the management switch should fail, the operational standby defined previously takes over the management of the stack. If the stack is expanded to include additional devices, the management switch automatically configures them by means of zero-touch provisioning. Of course, this automated process can be deactivated and switched to manual mode. In this way, the network administrator has a free choice of how to perform a stack extension, a port reconfiguration (also see the tech paper on the LANCOM XS-6128QF Configuration Options), or other settings.

Stacking implementation XS-6128QF

As described in the tech paper "XS-6128QF Configuration Options", the aggregation switch LANCOM XS-6128QF has freely configurable stacking FleX ports. These ports can be operated either as stacking ports or as Ethernet ports. When used as stacking ports, connection options include the 50G DAC (LANCOM SFP-DD-DAC50), 25G DAC (LANCOM SFP-DAC25), or even optical transceiver modules (LANCOM SFP-SR-LC25 or SFP-LR-LC25). The latter enables decentralized stacking. See "Fig. 2: LANCOM XS-6128QF port layout".

For an overview of stacking accessories, see "Fig. 3: Stacking accessories LANCOM XS-6128QF".

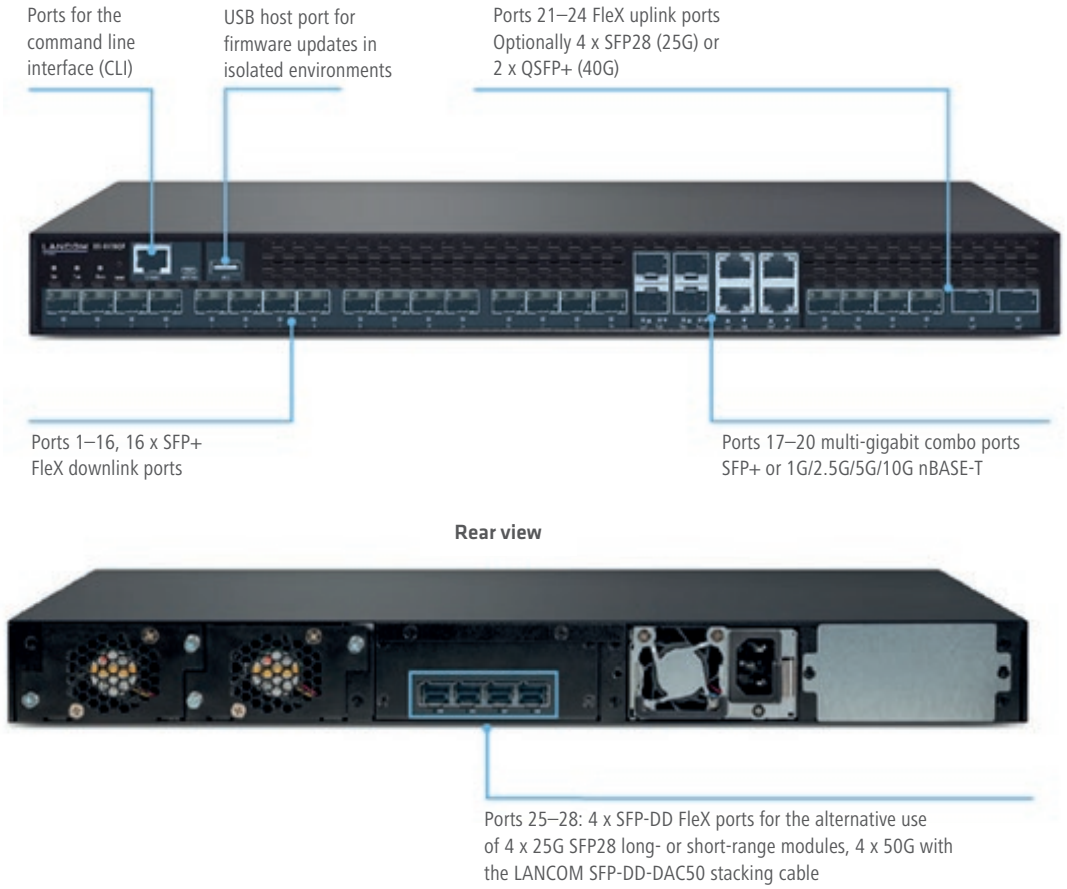

Fig. 2: LANCOM XS-6128QF port layout

The following describes how to optimize throughput with different numbers of switches in the stack. Note that the stacking protocol as described above works in the background without any special configuration.

Stack of 2 with the LANCOM XS-6128QF

When operating two LANCOM XS-6128QFs as a stack, and also with more than two switches as shown in the following scenarios, we recommend the ring topology. This means that the switches should have more than one physical connection by DAC or optical transceiver between them. The stacking protocol acts in the background to optimize network data flows without any further configuration of LACP or LAG groups.

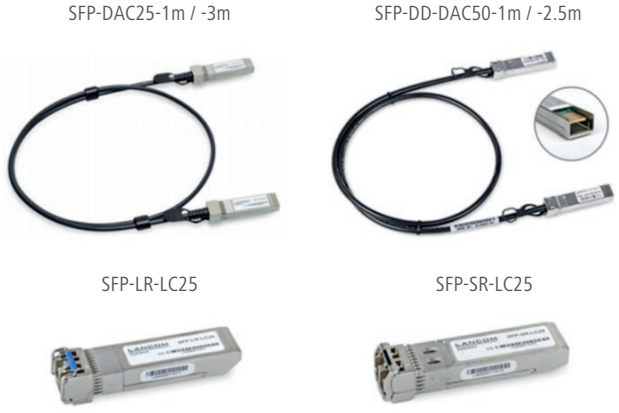

Fig. 3: Stacking accessories LANCOM XS-6128QF

For example, if two 50G LANCOM SFP-DD-DAC50 are used for stacking, these two cables are operated at 50G per direction to result in an overall bandwidth of 200G between the switches (full duplex).

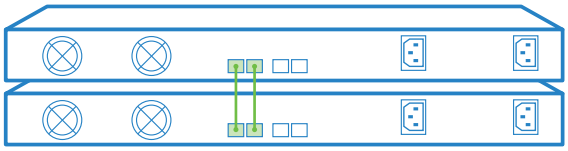

Fig. 4: Illustration of the cabling of a stack of 2 with the XS-6128QF with 2 connections

For the sake of completeness, we should mention that stacking can operate with just two switches connected, but this is not recommended due to the bandwidth limitations.

To further increase the bandwidth, four LANCOM SFP-DD-DAC50s can be used with two stacked switches. With 4x 50G per direction, this results in an overall bandwidth of 400G (full duplex):

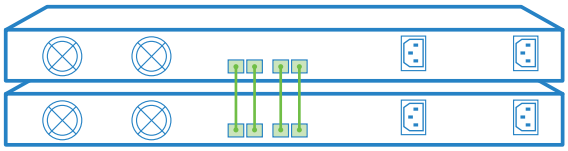

Fig. 5: Illustration of the cabling of a stack of 2 with the XS-6128QF with 4 connections

Stack of 3 to 8 with the LANCOM XS-6128QF

With a stack of 3 or more, the guideline for optimal cabling is as follows: To fully utilize all of the stacking ports, two LANCOM SFP-DD-DAC50s are used to connect each of the neighboring switches and, to complete the ring, two cables also connect the last switch to the first one:

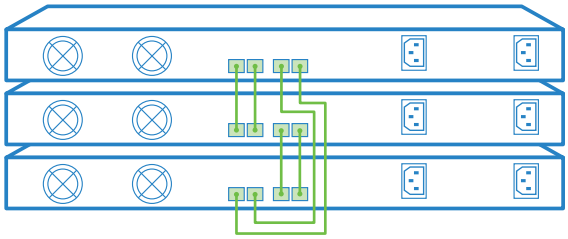

Fig. 6: Illustration of the cabling of a stack of 3 with the XS-6128QF

This illustration shows a stack of 8, but the cabling principle is the same for all other stacks:

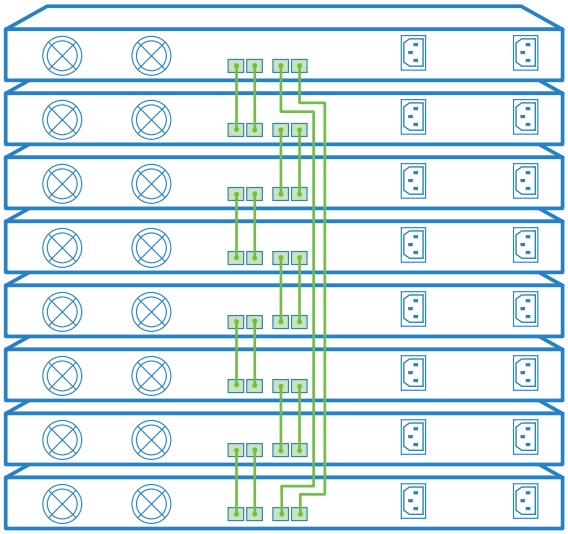

Fig. 7: Illustration of the cabling of a stack of 8 with the XS-6128QF

For the best possible distribution of the network load from the access switches, we further recommend that they should be connected redundantly to at least two different stacked switches via LAG or LACP. Example scenarios are outlined in the tech paper LANCOM XS-6128QF Configuration Options.

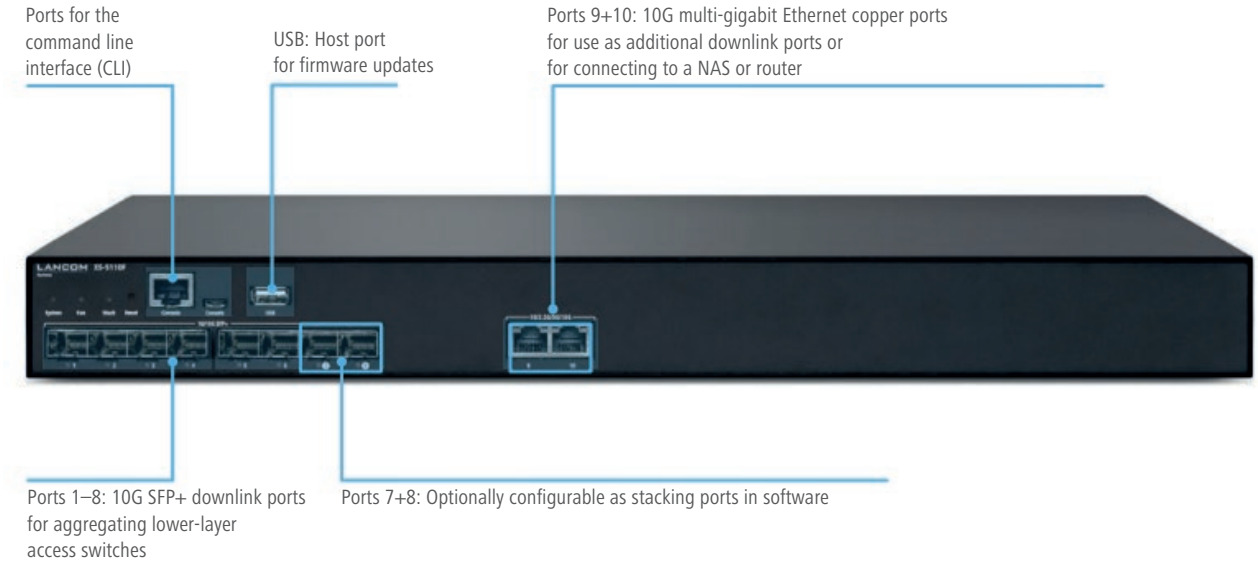

Fig. 8: LANCOM XS-5110F port layout

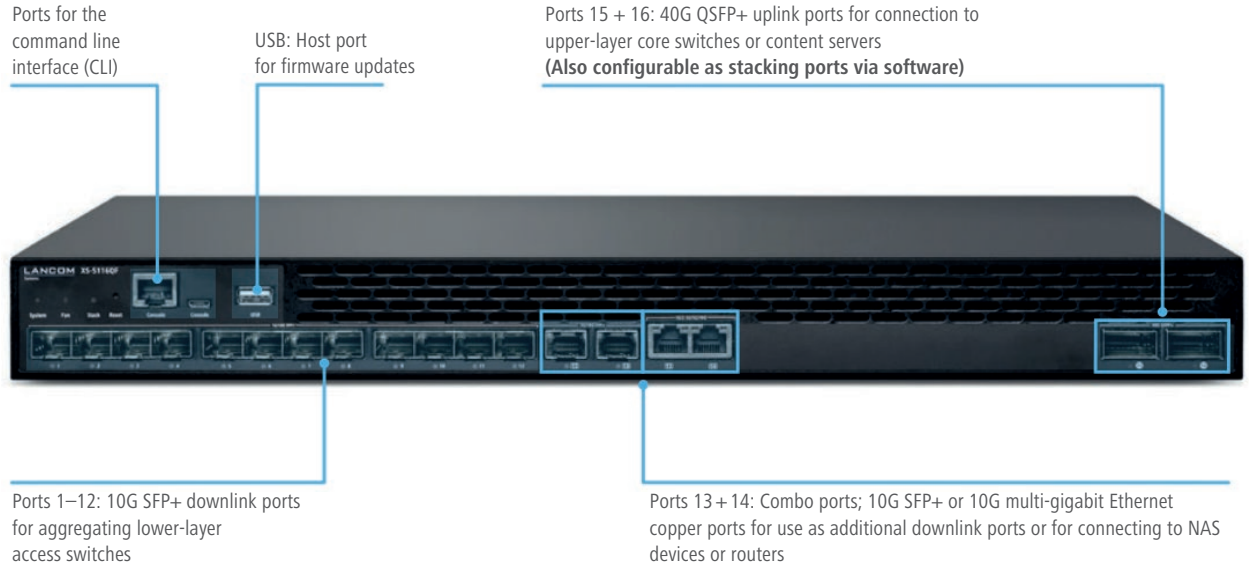

Fig. 9: LANCOM XS-5116QF port layout

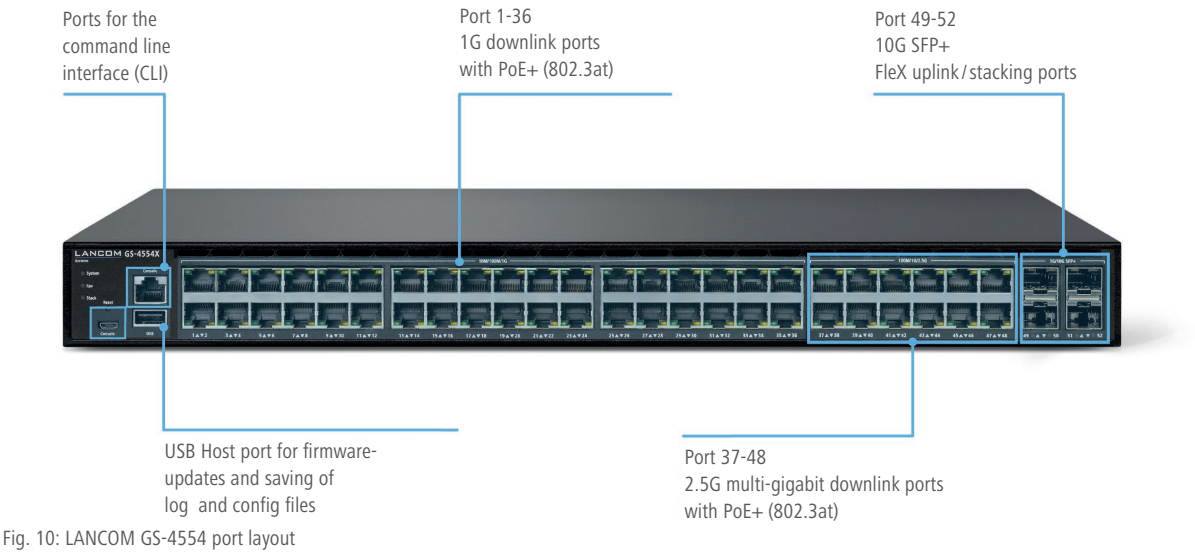

Fig. 10: LANCOM GS-4554 port layout

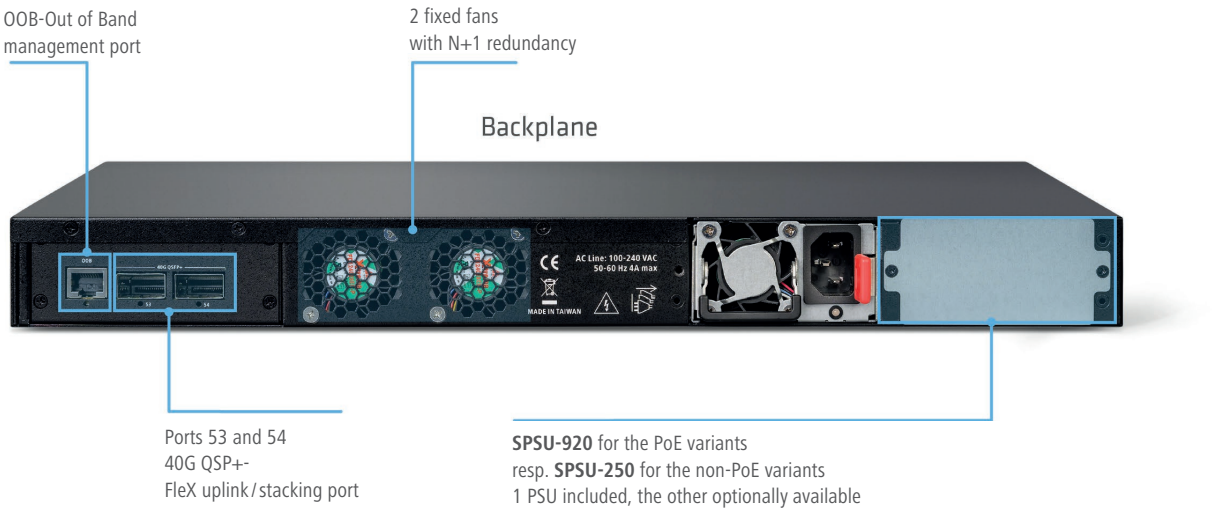

Fig. 11: LANCOM GS-4554 port layout (backplane)

Stacking implementation with LANCOM XS-5110F, XS-5116QF and GS-4000 series

In contrast to the LANCOM XS-6128QF, the aggregation switches LANCOM XS-5110F and LANCOM XS-5116QF as well as the stackable access switches of the LANCOM GS-4000 series have no rear stacking ports. To use the stacking function with the aggregation switches, you first have to use the CLI or WebGUI to configure the front SFP+ ports 7 and 8 (LANCOM XS-5110F) or QSFP+ ports 15 and 16 (LANCOM XS-5116QF) to be stacking ports. The default setting for these ports is "Ethernet".

The GS-4000 series can be stacked using two of the SFP+ ports 49 through 52 as well as the higher performance rear QSFP+ ports 53 and 54.

See "Fig. 8: LANCOM XS-5110F port layout", "Fig. 9: LANCOM XS-5116QF port layout", "Fig. 10: LANCOM GS-4554 port layout" and "Fig. 11: LANCOM GS-4554 port layout (backplane)".

The different stacking interfaces now also result in different accessories for the stacking function for these switches. See "Fig. 13: LANCOM SFP-DAC10-1m / 3m for LANCOM XS-5110F and GS-4000 series" and "Fig. 14: LANCOM SFP-DAC40-1m / 3m for LANCOM XS-5116QF and GS-4000 series".

Stacks of 2 with the LANCOM XS-5110F, XS-5116QF and GS-4000 series

As mentioned above these models each have two stacking ports, so the following connection guideline for a stack of two is quite self-explanatory:

Fig. 12: Illustration of the cabling of a stack of 2 with the XS-5110F

Here, too, the stacking protocol does all the work and the ring topology is recommended, so using the connections between two XS-5110Fs as shown above results in a throughput of 20G per TX or RX direction, and thus 40G overall bandwidth. This also applies to the GS-4000 series when using the SFP+ interfaces.

Fig. 13: LANCOM SFP-DAC10-1m / 3m for LANCOM XS-5110F and GS-4000 series

Fig. 14: LANCOM SFP-DAC40-1m / 3m for LANCOM XS-5116QF and GS-4000 series

With XS-5116QF and GS-4000 series, their 2x 40G QSFP+ ports result in 80G each in both directions, so a total of 160G data bandwidth in full-duplex operation.

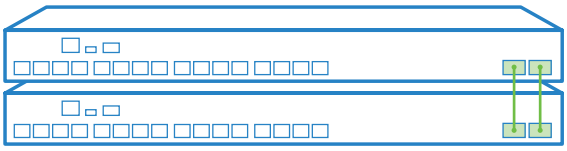

Fig. 15: Illustration of the cabling of a stack of 2 with the XS-5116QF

Stacks of 3 to 8 with the LANCOM XS-5110F and XS-5116QF and GS-4000 series

Starting with a stack of 3, the following connection schemes are used for the XS-5116QF as an example, which are also valid for the other switches via their stacking interfaces:

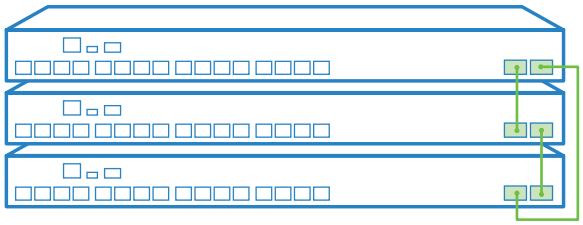

Fig. 16: Illustration of the cabling of a stack of 3 with the XS-5116QF

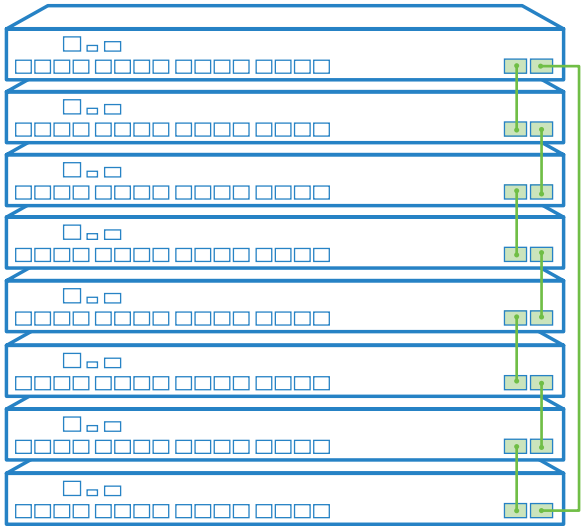

Fig. 17: Illustration of the cabling of a stack of 8 with the XS-5116QF

This connection guideline is quite self-explanatory as there is only one way to connect neighboring switches. This direct connection results in 40G (XS-5116QF, GS-4000 series) and 10G (XS-5110F, GS-4000 series) in each direction.

As with the XS-6128QF, the recommendation to connect subordinate access switches to at least two different stack units via LAG or LACP remains the same for these two switches. The stacking protocol running in the background ensures an optimal network load distribution.

As announced above, it should be emphasized once again that the LANCOM XS-5110F can also be stacked with the GS-4000 series via the SFP+ ports, and the LANCOM XS-5116QF can be stacked with the GS-4000 series via the QSFP+ ports. The resulting mixed stacks of copper (GS-4000 series) and fiber switches (XS-51xx) offer enormous flexibility when different types of cabling are used.

For further details please refer to the Design Guide.

Stack management by WebGUI

If all of the switches are interconnected as illustrated above, the stacking configuration in this example ("Fig. 18: Stacking GUI, XS-6128QF" and "Fig. 19: Rear view of "1", i.e. the management switch XS-6128QF") for an XS-6128QF is located under **1**. Clicking the Stacking tab in the dashboard takes you to the two submenus **Base** and **NSF**, which are discussed below.

Note that, with a stack in place, clicking on **2** toggles between the front and rear views of the selected stack unit. Without a stack configured, you will only ever see **1** displayed, since there are no other devices to view.

If you click on **1** and then **Base**, you will find all relevant stack configurations. The interface opens as shown in "Fig. 20: Stack Summary tab":

The **Summary** tab acts as a stack dashboard containing all information at a glance. It lists all connected stack units, in

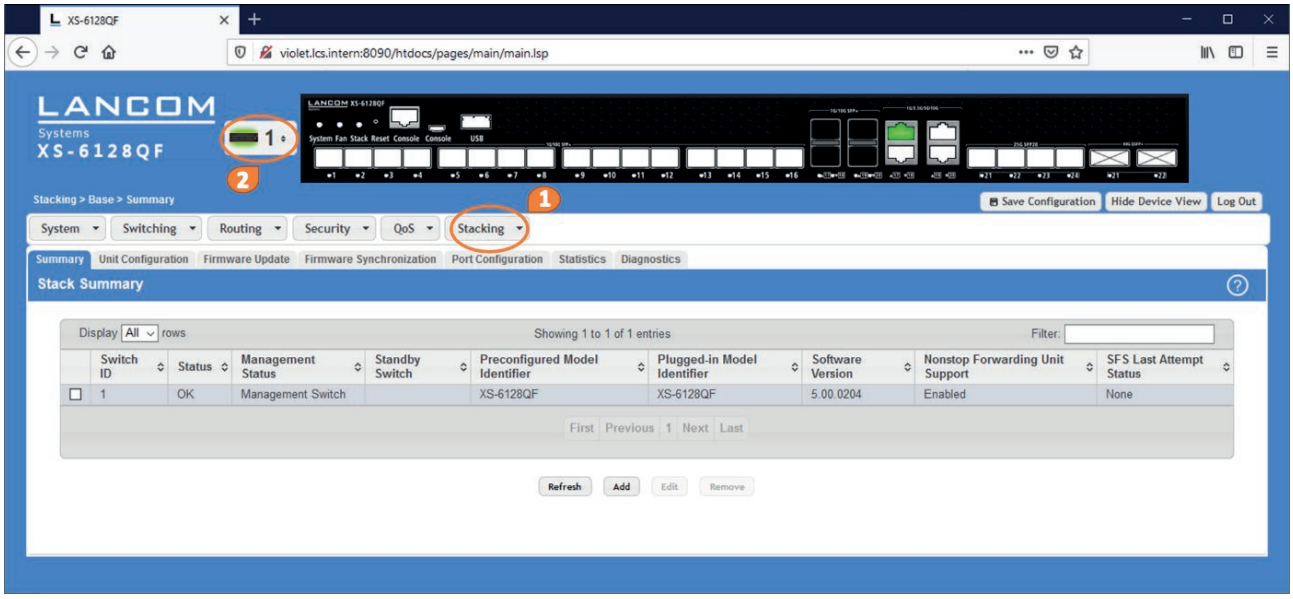

Fig. 18: Stacking GUI, XS-6128QF

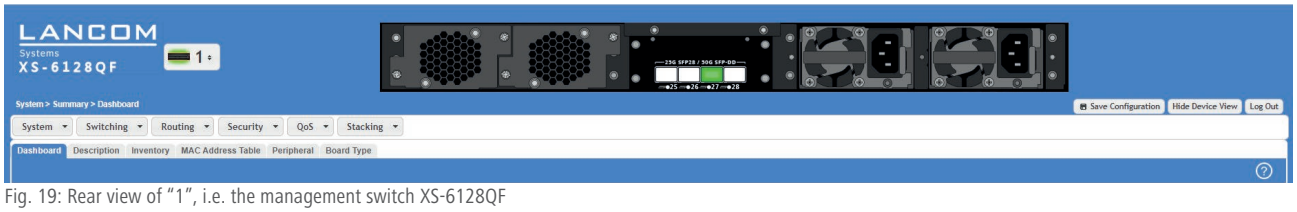

Fig. 19: Rear view of "1", i.e. the management switch XS-6128QF

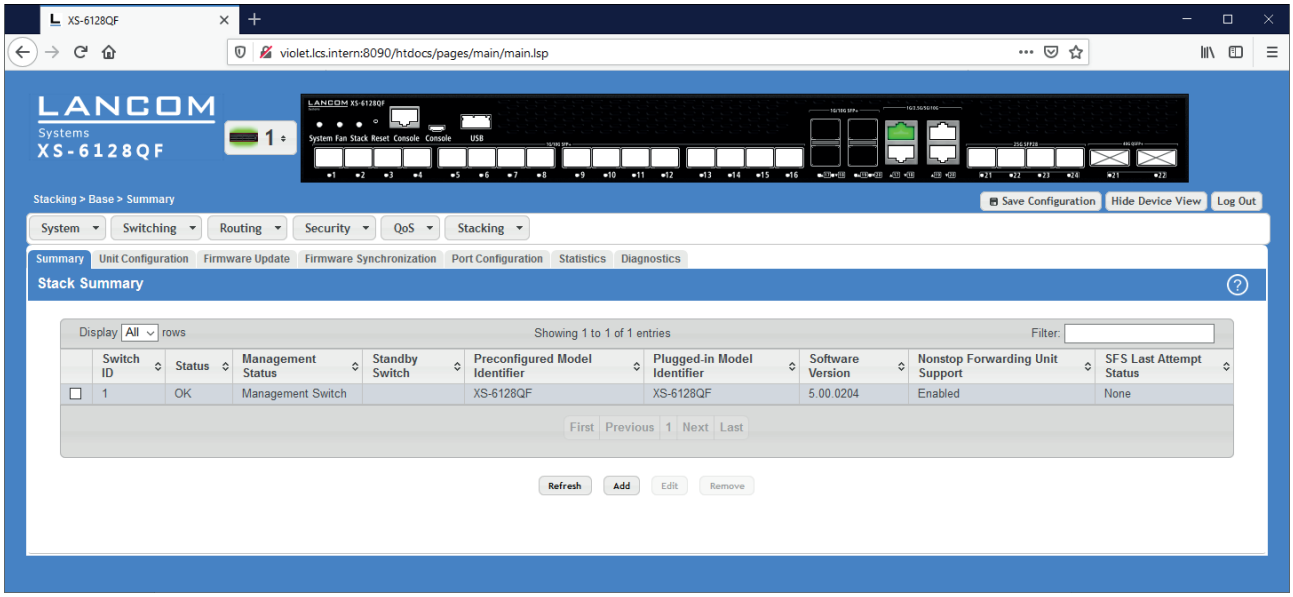

Fig. 20: Stack Summary tab

this case a single, standalone switch and thus a potential management switch. Additional information on the stack status is also displayed, such as the firmware versions of the members.

Under **Unit configuration** you see details such as serial number, status, etc. for all of the switches available under **1** and selected on this page. The item “Admin Management Preference” can be used to specify the role of this switch in a subsequent stack. The switch with the highest value becomes the manager when the stack is formed. See “Fig. 21: Unit Configuration tab”.

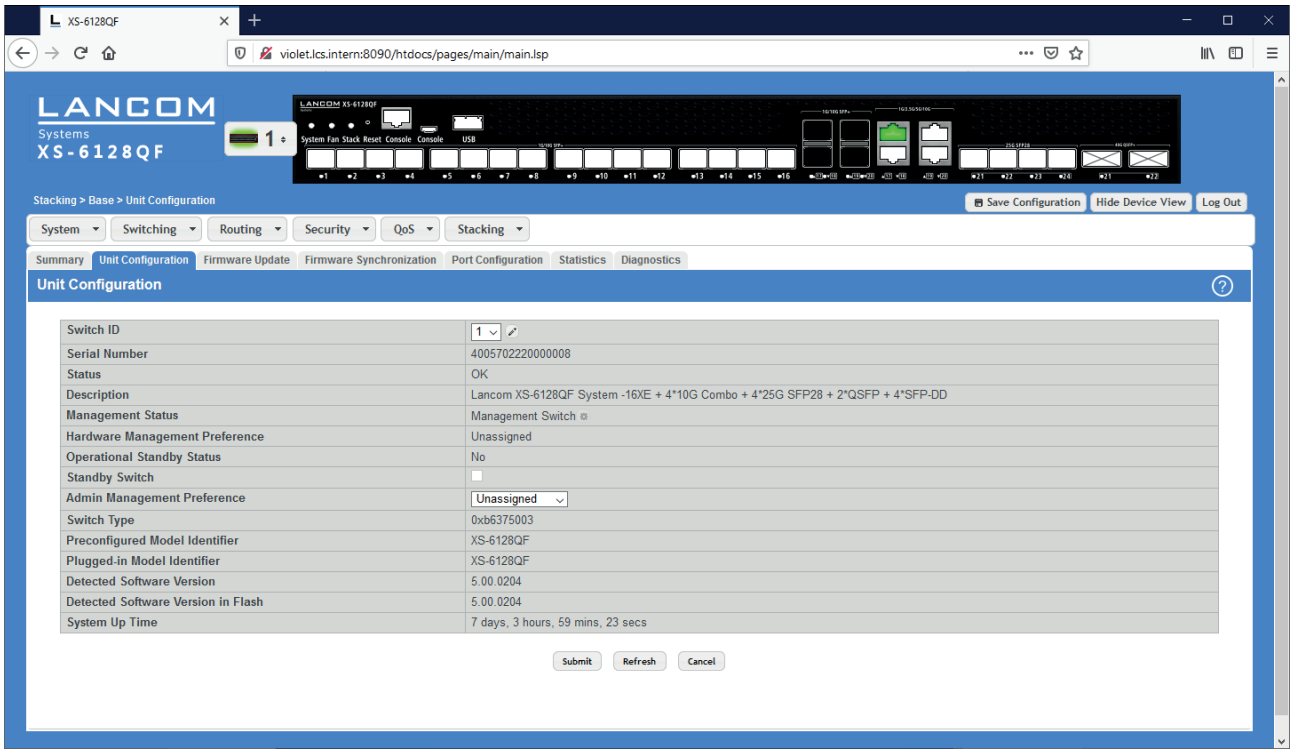

Fig. 21: Unit Configuration tab

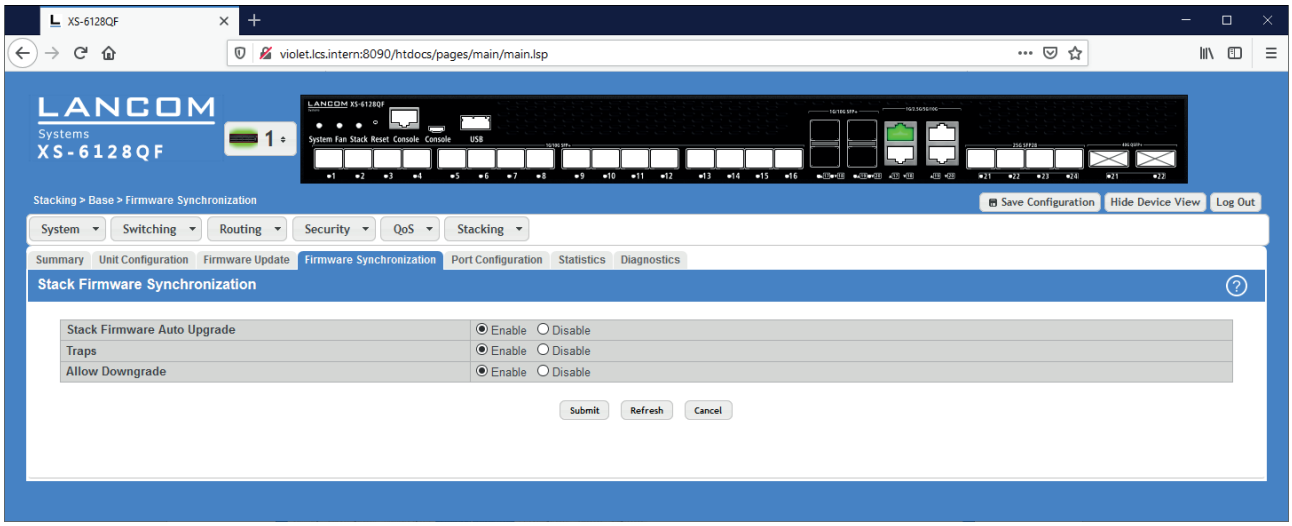

Fig. 22: Firmware Synchronization tab

The **Firmware Upgrade** tile allows the network administrator to manually control any firmware updates and to specify which is the next switch for the manager to provision with firmware. The status depends on the firmware version on the manager device.

The **Firmware Synchronization** tab controls whether or not the management switch manages and initiates automatic firmware updates. Also, SNMP traps can be output during the update process, and a setting determines whether a newly added switch with a potentially higher firmware version can be reset to an earlier version by the manager. See "Fig. 22: Firmware Synchronization tab".

The **Port Configuration** tab can be used to determine whether the rear ports are used for stacking or whether they should instead act as normal Ethernet ports. In the default state, these ports are configured as stacking ports. See "Fig. 23: Port Configuration tab".

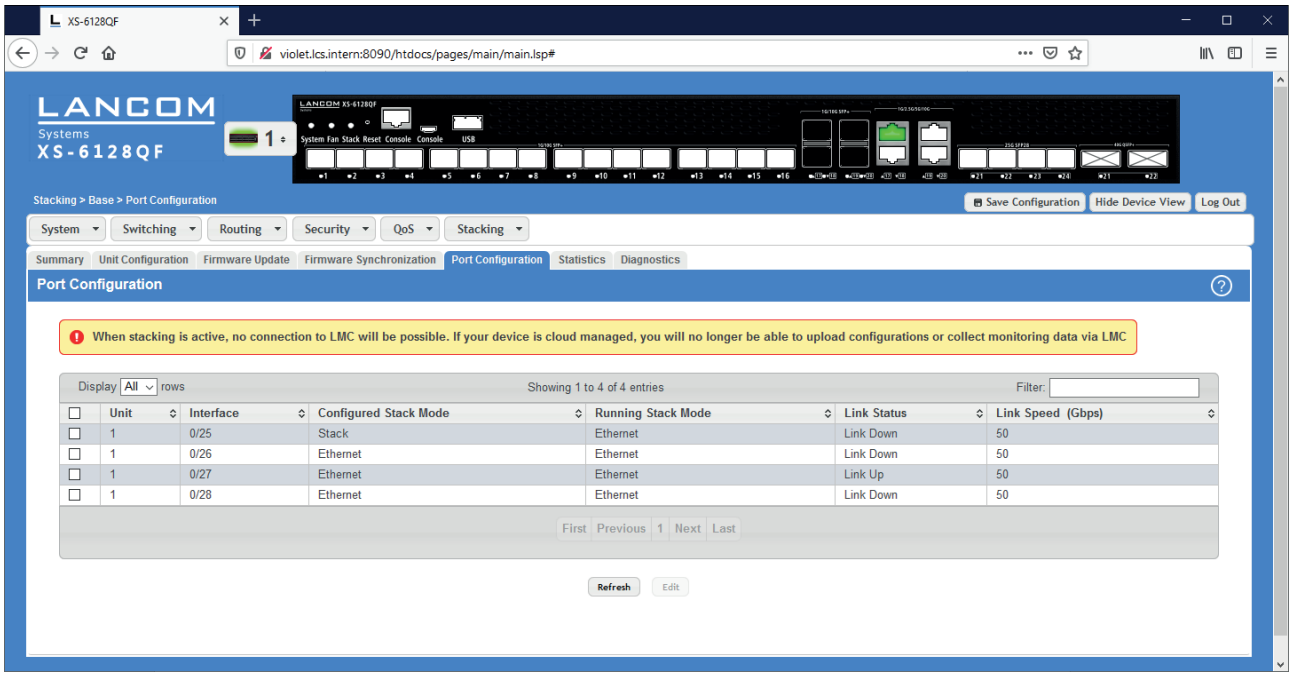

Fig. 23: Port Configuration tab

The **Statistics and Diagnostics** tab displays the current status of the stack and the data throughput (see "Fig. 24: Statistics and Diagnostics tabs").

Under **1**, the "Nonstop forwarding (NSF)" menu controls how the management switch behaves in the event of a failure. We recommend that you leave this function switched on and configure an operational backup accordingly (see "Fig. 25: Nonstop Forwarding tab").

Stack management by CLI

It would be impossible to outline all of the CLI commands for the stacking function here, so please refer to Chapter 2 of the highly detailed CLI manual for the XS switch family, available for example from the LANCOM product website under "Downloads & Links" (see "Fig. 26: Download & links on the product website for the LANCOM XS-6128QF").

Fig. 24: Statistics and Diagnostics tabs

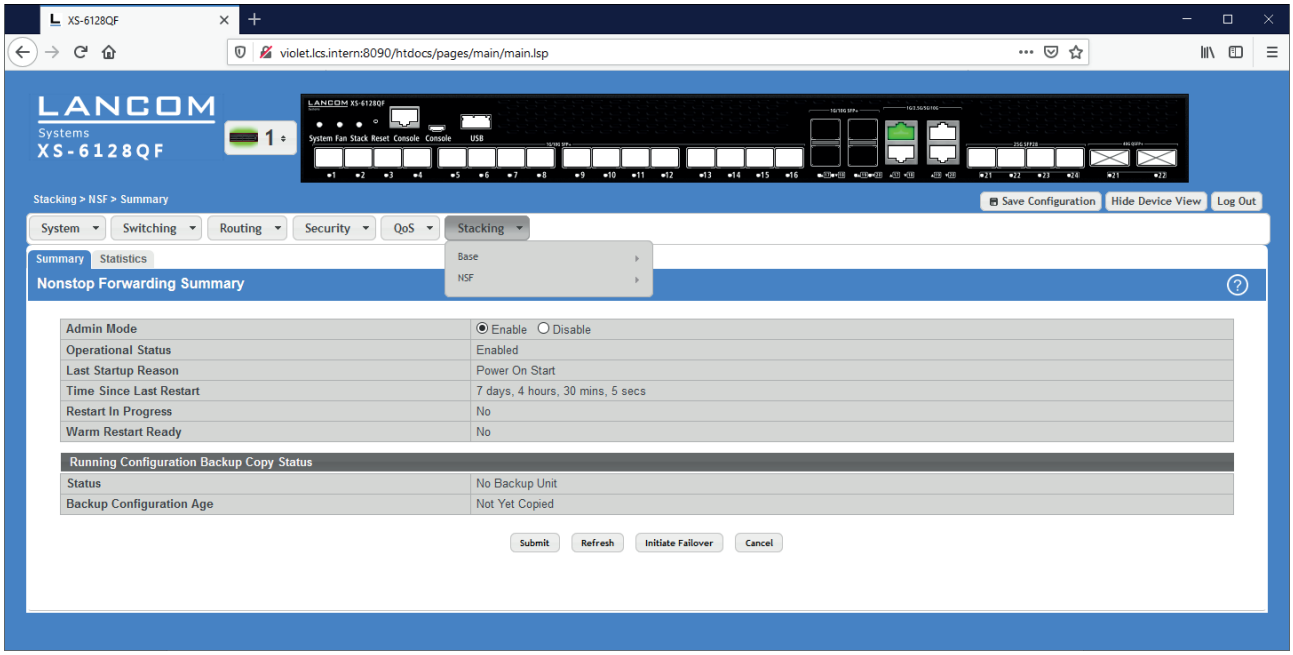

Fig. 25: Nonstop Forwarding tab

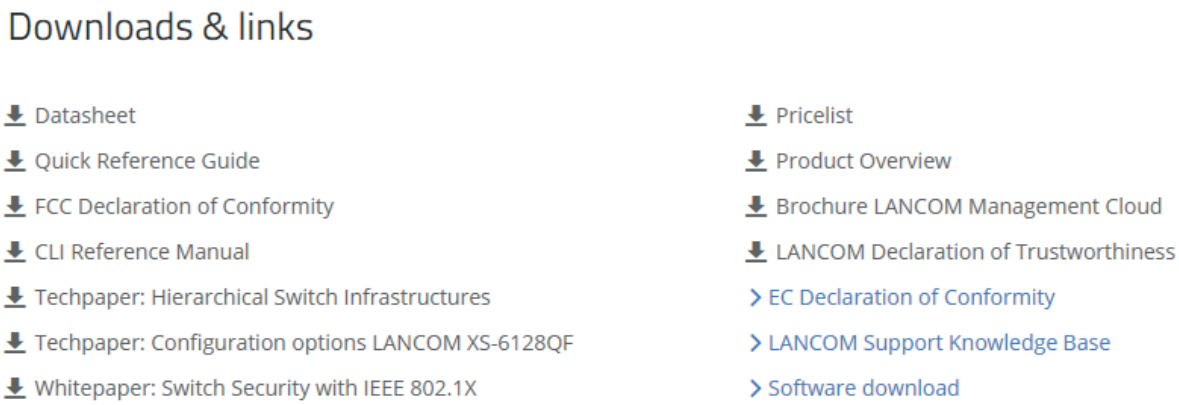

Fig. 26: Download & links on the product website for the LANCOM XS-6128QF

Port nomenclature for stacking

Once a stack has been configured and successfully booted, the next step is usually to configure the ports. Without going into every possible configuration now, for the sake of simplicity only the VLAN configuration page is shown here, which can be found under Switching and then **VLAN > Port Configuration** (see "Fig. 27: Port Configuration tab").

It is immediately apparent that the interface names consist of three digits:

The **first** digit indicates the **unit number** of the stacked switch. Strictly speaking this depends on the "Admin Management Preference" setting (see above), but for this example we assume that "1" is assigned to the management switch. The first digit "2" indicates the operational backup.

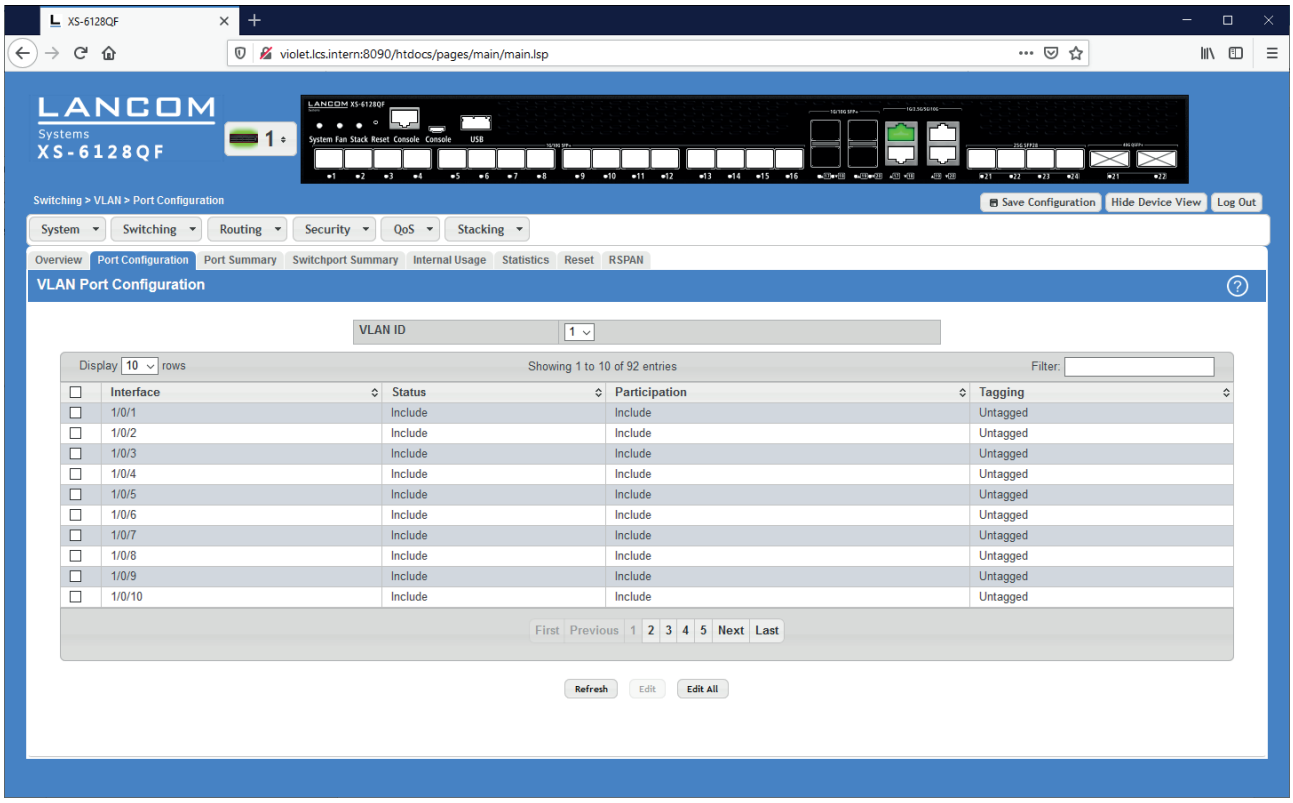

Fig. 27: Port Configuration tab

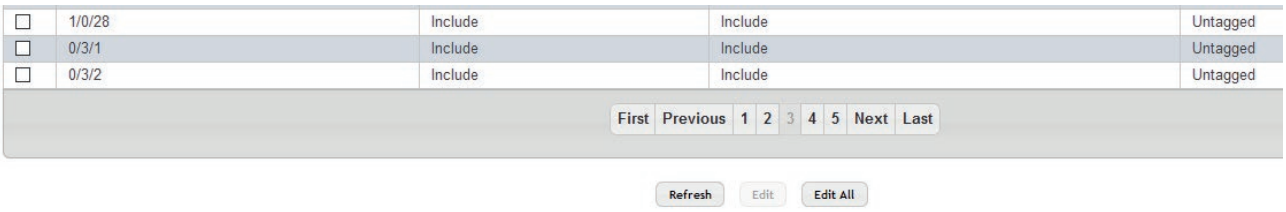

Fig. 28: Part of Port Configuration tab

The **second** digit indicates the blade or chassis slot and, in the case of the LANCOM XS-6128QF, this is **always "0"** because it does not have a modular structure.

The **third** digit indicates the **port number** of the selected stack unit.

Note that there are also interfaces that begin with a "0" and a subsequent "3". These are **reserved LAG groups** (see "Fig. 28: Part of Port Configuration tab").

As you can see, port 28 of the management switch, i.e. unit 1, is followed by the first reserved LAG group with 0/3/1.

Pairing of a switch stack with the LMC

Since setting up a stack is not remotely feasible for the network administrator, since the switches must be physically connected to each other, the local configuration of the stack is assumed as already seen above. Once the stack has been successfully formed, the entire stack with a device license can be paired with the LANCOM Management Cloud via the management switch ("management unit"). The cloud then automatically detects how many devices are in the stack and displays a corresponding stack icon and the switch type of the manager in the LMC device list (see Figure 29).

If you click on the stack name, the device detail view follows as for all device types, with the difference that up to 8 switches can then be hidden behind the stack (see Figure 30).

In addition to the device info, details about the state of the stack are also displayed via the green checkmark symbol for each stack unit, the up-time and other info such as the CPU load.

Configurations such as the assignment of individual VLANs or a LAG group can also be conveniently performed directly on this page by clicking on a port on the switch's stylized front panel. If a repetitive port configuration is required, the completed port configuration of a port can be conve-

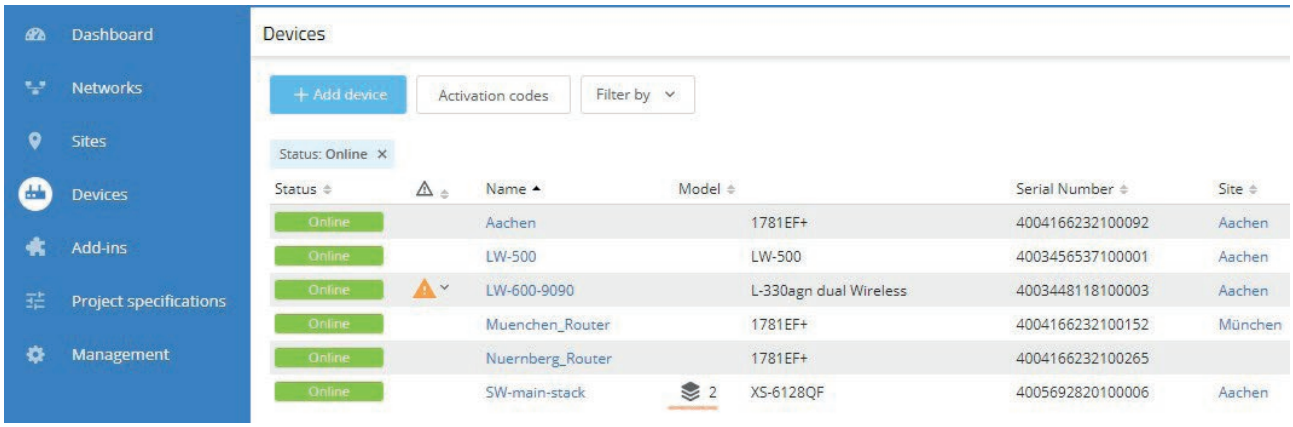

Fig. 29: LMC – Stack view in the device list

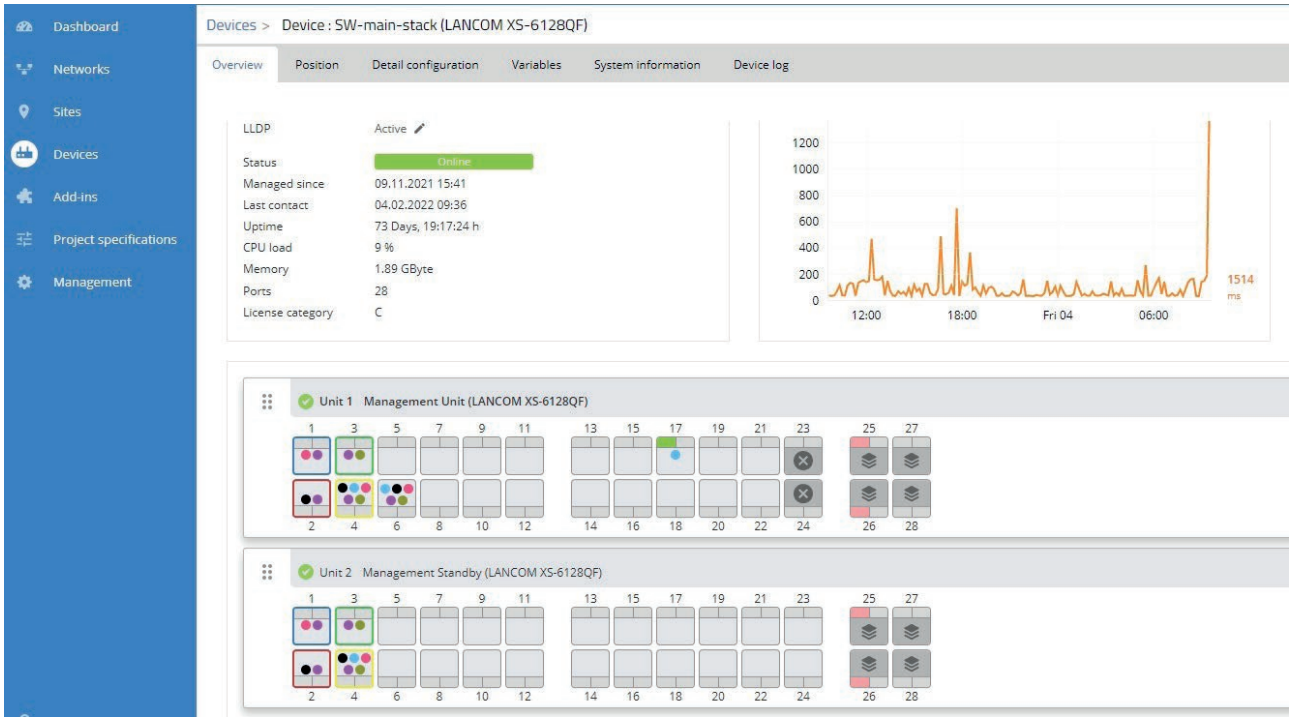

Fig. 30: LMC – Stack details view

niently applied to other ports to be selected via the copy icon (see Figure 31).

More complex configurations of multiple LAGs or VLANs can additionally be configured more effectively in a table view in the detailed configuration (see Figure 32).

Additional information about the stack, such as the serial numbers and MAC (partially grayed out here) of the paired devices, can be accessed via the "System Information" tile (see Figure 33).

Finally, it should be noted that so-called anomaly detection for enterprise switches, i.e. all stackable switches, can be activated via a switch in the project defaults (see Figure 34).

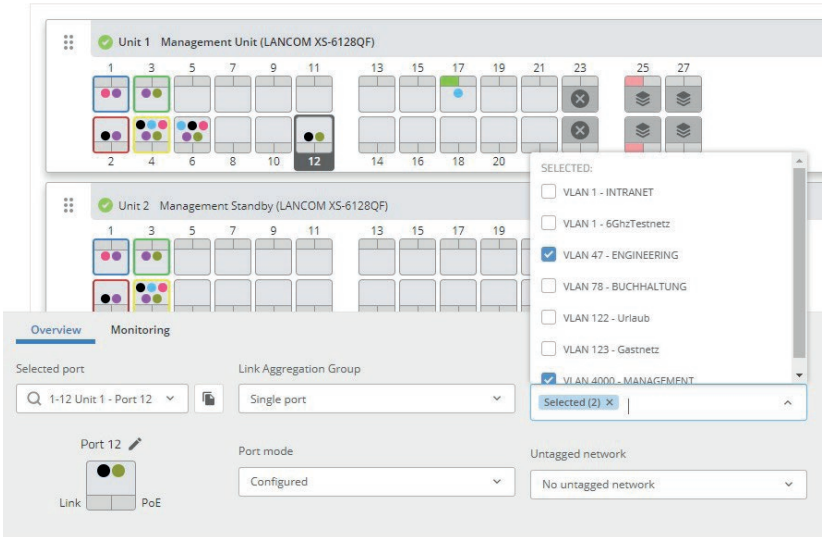

Fig. 31: LMC – Stack detail view with configuration options

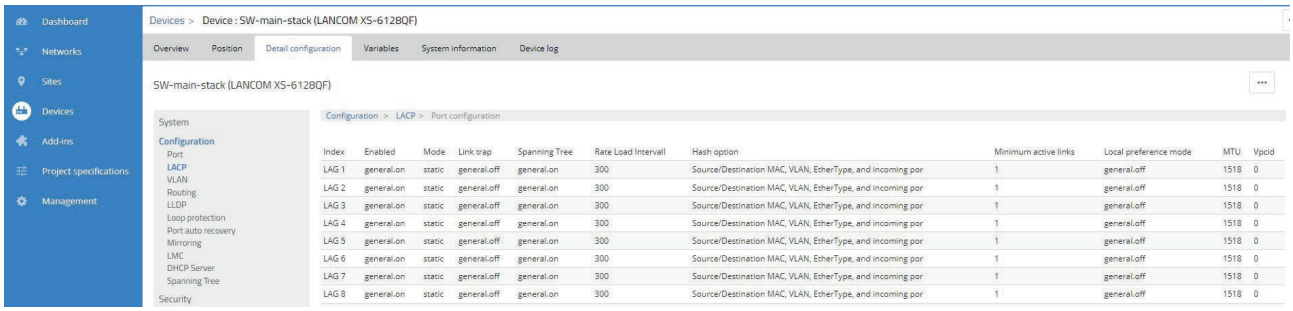

Fig. 32: LMC – Detailed configuration of LAGs as table view

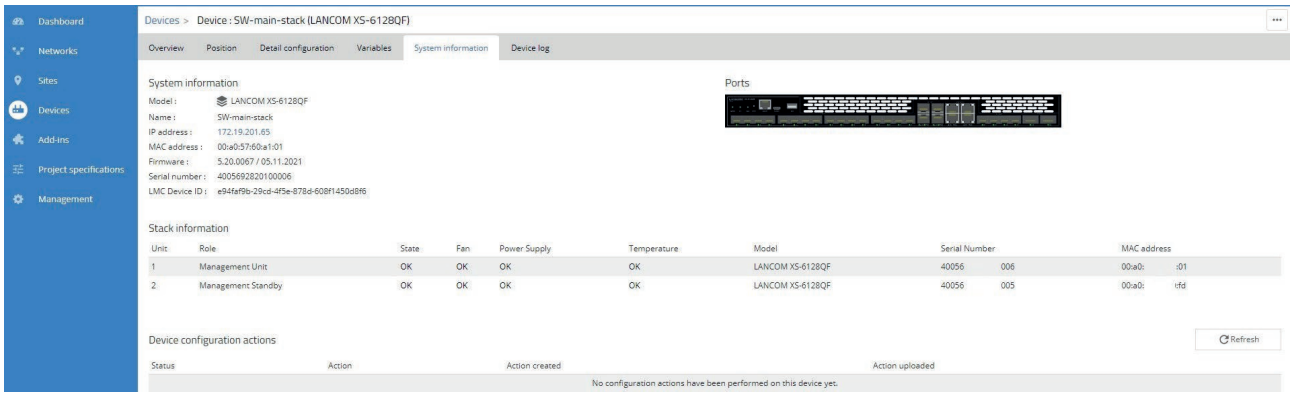

Fig. 33: LMC – System information of the stack

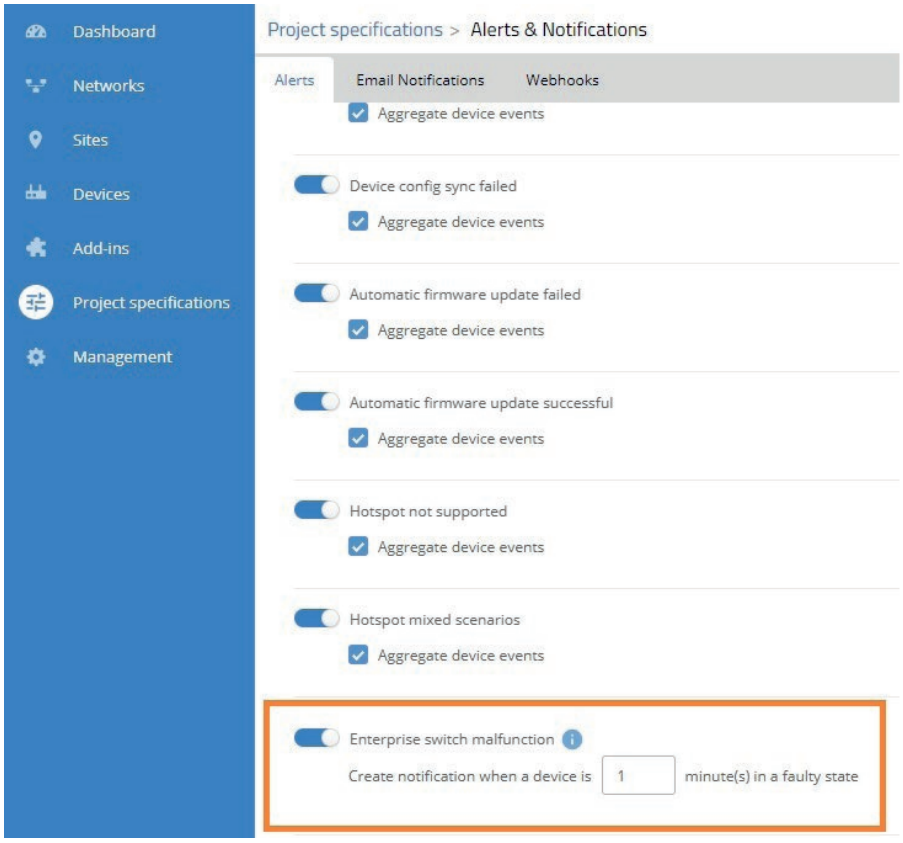

Fig. 34: LMC – Configuration of warnings in the project specifications

The monitoring function monitors power supply, fan, temperature and stack errors for all active stack units or individual devices of the LANCOM XS-51/61 and GS-4000 series.

Summary

Whether you require stacking at a single site or decentralized stacking over two or more locations, the LANCOM XS-6128QF is an ideal basis for a cost-efficient solution. The intelligent combination with the high-performance and cost-effective LANCOM access switches means that there are almost no limits on the possible applications—in particular for medium-sized companies.

If the requirements of the network are less complex and the stacking performance of the two smaller switches LANCOM XS-5110F and LANCOM XS-5116QF is sufficient, these aggregation switches form an excellent basis for small and medium-sized companies. Especially due to the LANCOM special feature of forming mixed stacks with the access switches of the GS-4000 series.

Whatever your situation, LANCOM Systems is the competent provider at your side. Experienced LANCOM technicians and the specialists from our system-vendor partners will help you with the planning and installation of a needs-based and future-proof LANCOM network.