CDFP MSA Delivers 400 Gb/s Today

The CDFP Form Factor

The CDFP module is the first 400 Gb/s form factor and will enable the highest port and bandwidth density of any pluggable form factor. With a 32mm pitch and 400 Gb/s of bandwidth per port, a 1U switch can deliver over 5 Tb/s of bandwidth. The CDFP module will support:

- 5 meter direct attach cables
- 100 meter multimode fiber
- 500 meter parallel single-mode fiber
- 2 kilometers of duplex single-mode fiber

With a wide range of supported distances for the data center, the CDFP port can be a very versatile solution for data center interconnects.

This paper discusses:

- Bandwidth trends driving 400GbE
- Limitations of current designs leading to the CDFP
- CDFP Retimer PHY
- CDFP 25Gb/s technologies

Industry Bandwidth Demands and Trends

The flow of information in the world today is truly amazing. Whether you are sitting at home surfing for details on your next purchase or connecting with a co-worker half way around the world, we now take it for granted that we have access to a global network that has tremendous capabilities. For business applications, we have full access to everything necessary to do our jobs whether we are in the office or traveling through some remote airport. Consumers are using the network in amazing ways whether it's to upload or download video or video chat with anyone anywhere.

The result of this network and an ever growing range of apps, is that traffic is growing "through the roof". By 2015, global IP traffic will exceed 1 zettabyte annually, and will reach 1.4 zettabytes per year by 2017, reflecting a compound annual growth rate from 2012 to 2017 of 23%. Perhaps, not surprisingly, the source of the traffic is shifting over this period and wireless traffic will surpass wired traffic by 2017. Only as recently as 2012, wired devices accounted for 59% of IP traffic[1]. In addition to the above IP network traffic growth, global data center IP traffic will reach 7.7 zettabytes by the end of 2017, representing a compound annual growth rate of 25% over the period of 2012 to

[1] Cisco, "Cisco Visual Networking Index: Forecast and methodology, 2012-2017" Cisco.com, 2013, <http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html>

2017.[2] All this growth requires significant increases in the capacity of network and data center equipment and connections.

In the past we have seen the networking and computing equipment market meet higher network capacity requirements by developing higher data rates. Currently, 25 Gb/s technologies are being developed to combine into 100 Gb/s channels that will replace previous 10 Gb/s technologies. It is anticipated these 25 Gb/s technologies can also be used to support next generation data rates beyond 100 Gb/s such as 400 Gb/s.

Over one billion smartphones shipped in 2013 for the first time and each of these smartphones drives a little traffic to the Internet and the data centers that support them. Estimates are that one server is needed for every 400 smartphones[3], so 2.5 million servers needed to be deployed in 2013 to support the phenomenal growth of smartphones - and deployments are increasing rapidly. The information technology industry is deploying about 10 million servers a year and server virtualization and massively parallel applications are driving higher bandwidths to and from each server. The need for bandwidth has never been greater and increasing the data rate of links to 400 Gb/s is helping meet that need.

Impact on Communication and Networking Equipment Architectures

Impact of the CDFP-MSA on Communications and Networking Equipment

One of the biggest decisions for the designer of communications and networking equipment is the one of which port to use. The equipment vendor would like their equipment to be used in as many applications as possible. To achieve that goal, they design the equipment around a common port, maximizing the number of options for cabling and transceivers. The form factor of the port often dictates the density and bandwidth of the switch, since only a limited number of ports can fit across the front face plate of the equipment.

Let's look at the density implications on two popular form factors, the SFP+ and the QSFP+. First, for 10 Gb/s SFP+ pluggable packages, switch vendors can fit as many as 60 ports in a 1 RU switch. This allows a switch bandwidth of 600 Gb/s. As evident from the photo in Fig. 1, it would be difficult to fit more than 60 ports in a 1 RU switch. Of course, other networking equipment consumes multiple rack units, allowing for more ports. But, often the best networking solution requires dense interconnect ports.

---

[2] Cisco, "Cisco Global Cloud Index: Forecast and Methodology, 2012-2017" Cisco.com, 2013, <http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud_Index_White_Paper.html>

[3] For every 400 smartphones, you need 1 server, Dean Takahashi, http://venturebeat.com/2013/11/21/for-every-400-smartphones-you-need-1-server/

Figure 1: An example of a networking switch with 60 SFP+ ports.

The second, very popular, form factor example is the QSFP. The “Quad” SFP form factor offers many of the same advantages of the SFP+ but improves on the density dilemma by combining four ports into one. It is commonly used for 40 Gb/s Ethernet applications, but the four channel architecture is also used by InfiniBand. The big advantage of the QSFP over the SFP+ is that it has enabled equipment vendors to more than double the front panel density. Multiple equipment vendors offer 1 RU switches with 32 QSFP ports; some offer 36 QSFP ports. This is the equivalent of 128 or 144 SFP+ ports (1.28 Tb/s or 1.44 Tb/s) with 10 Gb/s lanes. With 25 Gb/s lanes, the QSFP28 yields up to 3.6 Tb/s of data. Below in Fig. 2 is one such example of a switch with 36 QSFP ports. Once again, as evident from the photo, the front panel space is totally consumed by QSFP ports. It would be difficult to design a 1 RU switch using more QSFP ports.

Figure 2: An example of a switch with 36 QSFP ports.

Multiple vendors implement the IEEE 802.3ba 40 Gb/s Ethernet standard using the QSFP form factor. This allows equipment vendors to treat the 40 Gb/s link as a single pipe, simplifying switching architectures and cable routing. The flexibility provided by the QSFP means that the equipment designer can design around an industry standard switch port; gain all of the advantages of the SFP+; and scale to 40 Gb/s pipes. Other vendors use a mix of SFP+ ports and QSFP ports. A switch with both 10 and 40 Gb/s ports allows a transition from 10 Gb/s equipment to 40 Gb/s equipment. The photo below in Fig. 3 shows the front panel of a switch fully utilized by SFP+ and QSFP ports. Once again, no room for more ports.

Figure 3: An example of a switch with a mix of 48 SFP+ ports and 4 QSFP ports.

The next generation in front panel density came with the development of the CXP connector which could support 10 lanes at 10 Gb/s for an aggregate bandwidth of 100 Gb/s per port. This solution allowed up to 16 ports per front panel for 1.6 Tb/s, while still providing the ability to choose either DAC copper cables or longer reach optical transceiver plug in solutions.

Although we are in the early stages for 4x25 Gb/s, 100 Gb/s applications, the QSFP form factor is expected to dominate again. The CFP4 form factor will also play a role in 4x25Gb/s for links greater than 100m and it can support 16 ports in the front panel. Equipment vendors are already designing switches with 100 Gb/s ports; some versatile switches will have a mix of 10, 40 and 100 Gb/s ports; others switches, to maximize switching bandwidth will squeeze as many 100 Gb/s ports as possible in the front panel.

So how do equipment vendors scale to 400 Gb/s pipes? What form factor could both support 400 Gb/s pipes and, at the same time, allow front panel density to double again? If the form factor doesn't solve the density problem, then it constrains the bandwidth of the switch. Is there a way to design a form factor which both supports 400 Gb/s pipes and doubles the front panel density?

Enter the CDFP

The CDFP-MSA offers the vision for a new form factor that supports both 400 Gb/s interconnects and increases front panel density. It is conceived around the same factors that made the SFP+ popular for 10 Gb/s networks and the QSFP popular for 40 & 100 Gb/s networks. This new form factor will more than double the front panel density when compared to QSFP. It is supported by a wide range of communications and networking equipment vendors; it is supported by a wide range of transceiver and cabling vendors; and, it is envisioned to support a range of technology and distance alternatives. Specifically, the CDFP pluggable package is designed to support:

1. Front panel density of 11 or 13 ports on a 365 (14.4") or 445 (17") faceplates. This could enable future 1 -RU switches and routers to support 4.4 or 5.2 Terabits per second in a single density configuration.
2. The 400G CDFP will support multiple power levels for a range of technologies.
3. The design intent of the CDFP is to be versatile enough to support technology choices:
    a. Direct attach copper cables (DACs)
    b. Active optical cables (AOCs)
    c. VCSEL transceivers (generally for reaches up to 100m)
    d. Silicon photonics and long wave transceivers (for reaches up to 2km or longer)
4. The design intent of CDFP transceivers is to support either single mode fiber (SMF) or multi-mode fiber (MMF).

5. The design intent of the CDFP interconnects is to be “hot pluggable.” Once again, this allows a single link to be replaced or upgraded without disturbing the rest of the network.
6. The CDFP-MSA is supported by a broad range of equipment vendors and interconnect suppliers

The CDFP-MSA is not envisioned to solve every 400 Gb/s application. Like any MSA, this MSA does have limits. The CDFP-MSA presently describes 16 lanes at 25Gb/s each. The requirements for 8 lanes at 50 Gb/s or 4 lanes at 100 Gb/s have so far not been included. Other 400 Gb/s networking applications may require longer reaches than the 2 km anticipated by this MSA.

The CDFP-MSA does solve many of the needs of the equipment vendor. It provides a next generation package architecture that greatly improves front panel density while at the same time supporting a variety of technology choices. The strong industry players backing this MSA provide confidence that an end-to-end, 400 Gb/s eco-system, can develop supporting the next generation of optical and electrical interconnects for data communications.

400 Gb/s discussions in the industry

The 40 Gigabit Ethernet (GbE) and 100 GbE standards were released in 2010 and the 400 GbE standard is expected to be released in 2017. To release 400 GbE products in 2017, the building blocks of 400 GbE, like the CDFP port, need to be in place years in advance. 400 GbE is going to be based on 16 x 25 Gb/s electrical signaling technology and the CDFP is set to be the first and most compact, pluggable form factor dedicated to 400 GbE. Eleven CDFP ports will enable over 4.4 Terabits/second of bandwidth in a 1 RU switch or a blade in a chassis switch. As the industry drives to smaller, faster and lower cost infrastructure, the CDFP is set to be the data center workhorse that delivers 400 Gb/s.

While the 400 GbE standard is still a few years away, the need for 400 Gb/s interfaces is here today. The CDFP form factor is already being used in proprietary interfaces to interconnect high performance servers and will soon be used to interconnect switch and router chassis. These proprietary chassis interconnects have always been massively parallel and will continue because they provide the massive bandwidth needed to interconnect equipment so that multiple chassis perform as one big chassis. While 16 lanes is a fairly wide interface, multiple applications need the maximum amount of bandwidth that can only be provided by many parallel lanes running at the fastest speed available.

Availability of enabling elements for the CDFP form factor

CDFP Re-Timer PHY Overview

The CDFP Re-Timer PHY(s) within the target system will be situated between the host IC and the module optics, or copper cable interface, depending upon the specific application. The Re-Timer PHYs are required to carry out several key functions in order to ensure overall system robustness and expected performance.

A functional block diagram based on four-lane PHYs is shown in Fig. 4. Each lane of the Re-Timer has an input equalization stage which typically includes a Continuous-Time Linear Equalizer (CTLE) amplification stage, a Clock and Data Recovery (CDR) stage and an output driver de-emphasis stage.

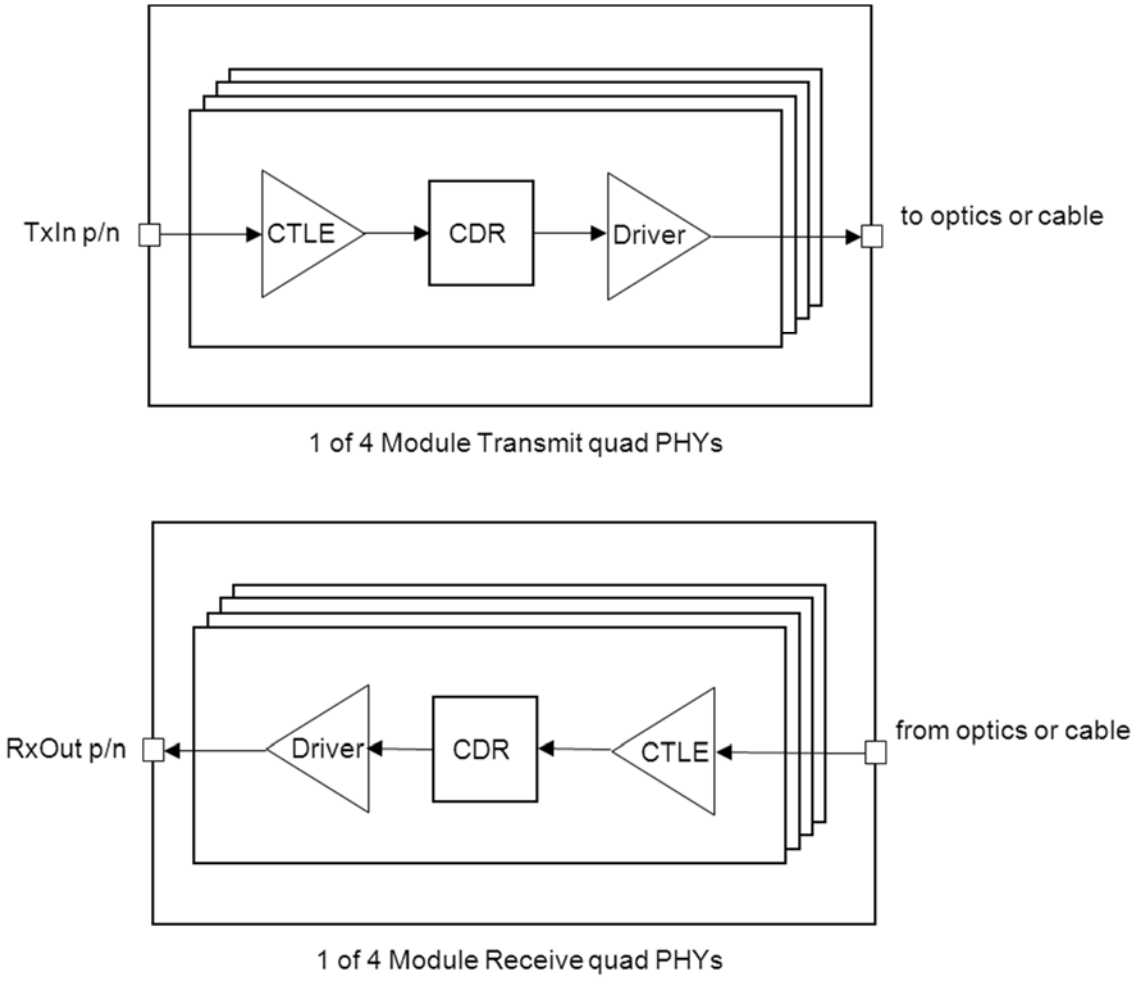

Figure 4: Quad Re-Timer PHY Functional Block Diagram

Within the Transmit PHY, the CTLE provides sufficient high frequency peaking response in order to be able to equalize a CEI-28G-VSR/CAUI-4 compliant channel which may have up to 10 dB loss from the host IC. In addition, the CTLE amplifier needs to be able to provide a Tx LOS alarm in order for the module to be able to enable a Tx squelch function. The CDR block recovers a clock from the incoming data and retimes the data using the recovered clock. This provides a data eye with sufficiently low jitter for the transmit optics. Ideally, the CDR should not require a reference clock as there is no provision for such a signal from the host within the CDFP connector and if required would need to be generated from a sufficiently stable clock source within the module itself. As well, the CDR must be able to provide a LOL alarm if unable to lock to the incoming data stream. The output driver provides sufficient voltage swing for the optical transmit input and usually has the ability to provide some output de-emphasis for the short channel between the output and optical driver input.

The required Receive PHY functionality is nearly identical to that of the Transmit PHY and needs to be able to perform the same data path functions. Differences may include a receiver CTLE/amplifier stage having different equalization capability than the transmit CTLE depending upon whether the application is for an optical or electrical interface.

All the above features and performance levels already exist for use in retimer PHY chips being developed for 4 x 25 Gb/s applications today that are being designed with the OIF CEI-28G-VSR and IEEE 802.3bm CAUI-4 specifications in mind. This is great news as re-use of existing elements means less risk, shorter lead times and improved economics.

25G Connector Technology

The challenge for I/O connectors is not just higher performance, but also minimizing the amount of face plate space required and maximizing port density, without increasing board layer count. The CDFP connector does all these things with its unique construction. To achieve this the CDFP connector offers straight back routing through the connector pin field footprint versus routing outside the connector pin field footprint.

This routing allows the connector to be placed on the tightest port pitch without increasing board routing layer count. This style footprint is common in backplane connectors but not in I/O's due to the structure of the low cost I/O interface and pin out (edge card) that typically drives the connector structure and pin field footprint. The unique construction of the connector translates the interface pin-out to the footprint pin-out while maintaining the SI performance required to run at 25 Gb/s. To achieve the higher data rates in a system channel, lower loss material are recommended such as Megtron6. Also, since it is a press fit connector, back drilling is recommended and CDFP uses a small press fit pin that allows back drilling to within 1.0 mm from the top of the board to maintain high performance.

See Figure 5 for a density comparison of the 400 Gb/s CDFP connector to the 100 Gb/s CFP4 connector (a 4x25Gb/s solution). As discussed earlier, these comparisons can be made to any number of pluggable form factors such as SFP+, QSFP28 and CXP, all

with the same result; a significant increase in aggregate bandwidth and port density by using the CDFP form factor.

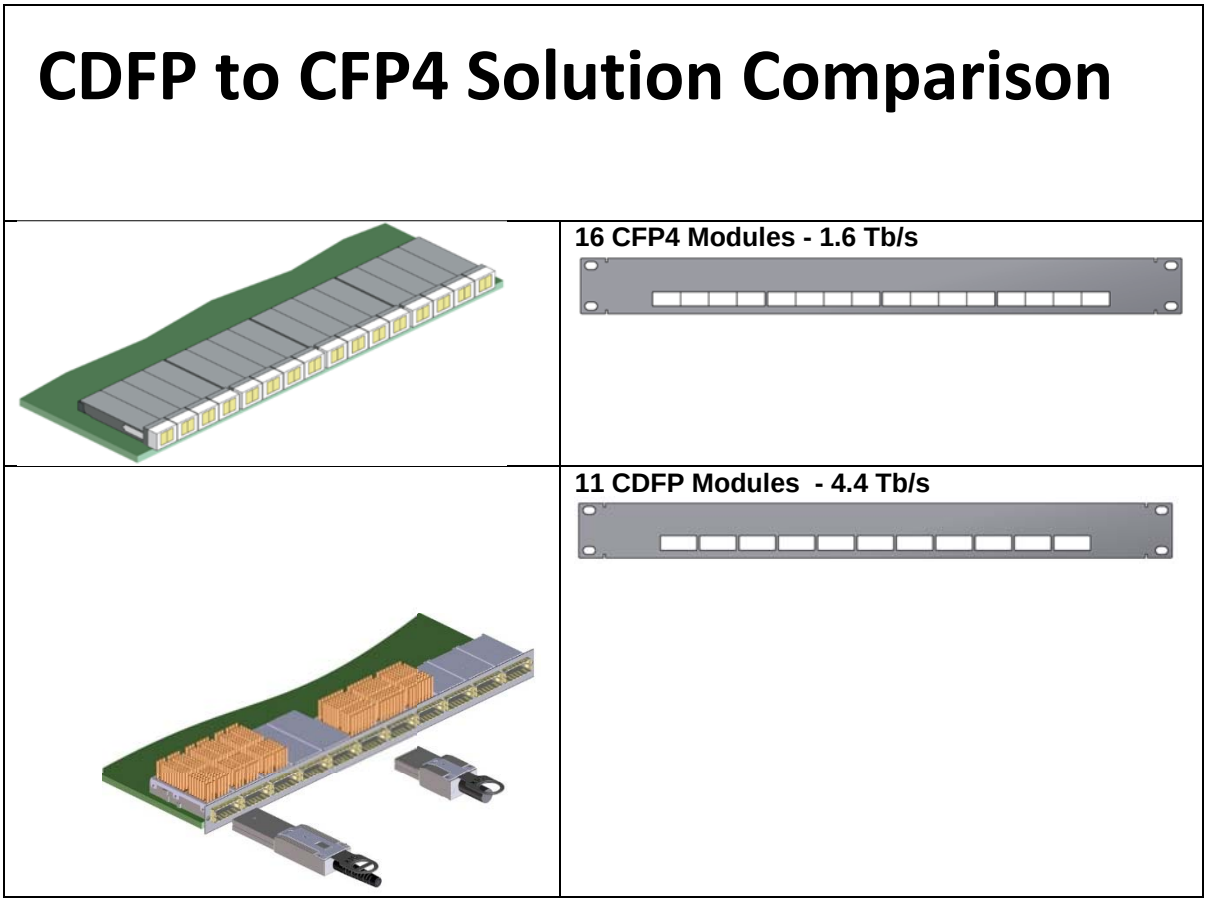

Figure 5: CDFP Connector to CFP4 Connector Density Comparison

25G Capable PCB Materials

Newer high performance printed circuit board (PCB) materials such as Panasonic's Megtron6 can provide the industry with the PCB "reach" necessary to realize conventional equipment architectures. The insertion loss of the PCB materials commonly used in 10 Gb/s era equipment would not be practical at 25 Gb/s. This new class of PCB materials offer a reduction in loss (or an increase in reach) of up to 42% compared to "improved" FR4 and enable practical implementation of OEM networking equipment based on 25 Gb/s lanes combined into 100 Gb/s and 400 Gb/s.

25G Copper Cables

The IEEE 802.3bj project is specifying 100 Gb/s copper cable assemblies based on 4 x 25 Gb/s with reach up to 5 meters with 24 gauge cable. For CDFP operation where 32

differential cable subassemblies would be required, 24 gauge cable may not be practical, but 30 gauge cables would result in a reach of 3 meters and only have a diameter of 12mm, while providing a low cost 400 Gb/s solution. With creative packaging, it may be possible to implement solutions with larger gauge cables allowing suppliers to get back to 5 meter passive solutions. Additionally, it is anticipated that active copper cable assemblies will also exist in this form factor as they have in other direct attach cable form factors (SFP+, QSFP+, etc.). An additional benefit of the CDFP cable assembly is that breakout solutions can be offered to support applications where a 400Gb/s CDFP port is connected to equipment that implements 400Gb/s based on four QSFP28 ports or the breakout cable can be used to enable high density 100Gb/s connectivity.

25G Optics

Earlier in this paper it was mentioned that the CDFP form factor is intended to support multiple technologies, and optics is one of the areas where this is expected. The MSA does not specify reach, however it is anticipated that both VCSEL and Silicon Photonic technologies will play a role offering different reach solutions.

As the industry has participated in the development of the IEEE 802.3bm 40 Gb/s and 100 Gb/s next generation optics project, they have brought 100 Gb/s VCSEL (vertical cavity surface emitting laser) based product operating in 4x25Gb/s architectures. These developments are now being released to market as QSFP28 transceivers and active optical cable assemblies. The CDFP form factor will use these same VCSELs in a 16 x25 Gb/s architecture to realize 400 Gb/s optical links that will provide reaches of at least 100 meters. These 100m links have been demonstrated over temperature ranges of 0 to 85 degrees C. VCSEL's are particularly well suited to 400 Gb/s applications due to their demonstrated lower power consumption which results in reduced thermal management loads for application in the OEM equipment. Being able to re-use these 25 Gb/s VCSELs in CDFP applications allows product to be available in 2014.

Silicon Photonics based optical assemblies provide a low power solution for medium range interconnect solutions up to 2km. Silicon photonics AOC's use single mode fiber and as such can transmit the longer distances while still operating in the relatively modest power environment of the CDFP form factor. Yet they provide significantly lower cost and power consumption than traditional long range optics. They implement the same technology used in the 4 x 28G zQSFP+ AOC's, thus providing a proven path to the CDFP 400G requirement.

The CDFP optical solutions can also support breakout cable assemblies. There could be one end with a CDFP form factor and four QSFP28 form factors on the other end. There has been broad support for these types of breakout cable assemblies across the industry as they offer tremendous flexibility and density.

Bringing it all together

The market place is consuming more and more bandwidth annually as new powerful applications are developed and the masses expect to have access to any data anywhere all the time. To avoid network bottlenecks, higher data rate “pipes” are required to connect networking equipment. The CDFP MSA is leveraging all the good industry work that has been done to develop 25 Gb/s channel line rates in silicon devices, connectors, optical transceivers and copper cables to deliver a next generation 400 Gb/s form factor solution that offers significant line card density gains. Partially due to the technology re-use, this next generation bandwidth is available near term with a number of hardware displays and preproduction demonstrations taking place at OFC 2014. Visit www.cdfp-msa.com to stay current on the latest developments of the CDFP MSA.