

Mellanox Peer-to-Peer Einrichtung

Unraid <-> Windows

Vorwort:

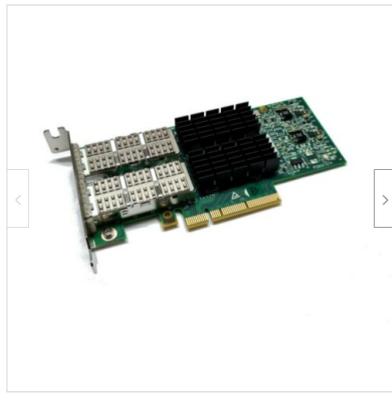
Dieser Guide richtet sich an die Leute die Daten sehr schnell im Heimnetz hin- und herschaufeln, oder wie ich einfach einen kleinen Server als NAS am laufen haben oder auch mit Parsec z.B. von einem leistungsfähigerem PC mit nativer Auflösung und High-Refresh-Rate sowie geringer Latenz streamen wollen.

Mein Szenario lag hierbei, dass ich zwei PCs habe mit jeweils einer Grafikkarte zum zocken (in meinem Fall in PC1 eine RX6900XT und in PC2 eine RTX3090).

In jedem PC habe ich natürlich mehrere SSDs verbaut worauf z.B. auch meine Games liegen welche ich zurzeit gerne zocke, und ab und an mal auch mit Parsec zum gleichzeitigen Arbeiten vom einem zum anderen PC streamen möchte oder mal schnelle Benchmarks zum Vergleich ausführe.

Daher dachte ich mir wenn ich die SSDs in den Server bekomme und dann evtl. auch noch eine NVMe SSD als Cache dazu packe, könnte ich annähernd mit nativer Zugriffsleistung auf die Laufwerke, meine PCs erstens von einer NVMe befreien (spielt ne Rolle wenn man zu wenige PCIe-Lanes zur Verfügung hat um im zweiten oder dritten PCIe-Slot noch Karten betreiben möchte, da bei vielen Consumer-Boards die Lane-Verteilung dann entweder einen Slot deaktiviert oder eben auf eine niedrigere PCIe-Gen schaltet und daher Performance verloren geht).

Soviel zum Einsatzzweck, nun zu meiner Einkaufsliste für dieses Vorhaben. Ich schaute mich in der Bucht um und fand Mellanox Connect-x3 VPI Karten:



Mellanox ConnectX-3 CX354A PCIe x8 10 40 GB QSFP+ Dual Port Server OEM 7046442
★★★★★ [Schreiben Sie die erste Rezension.](#)

Artikelzustand: **Gebraucht**

Multi-Rabatt:

Stückzahl: 4 oder mehr für EUR 77,50/Stk.
8 verfügbar
[369 verkauft](#) / [Bewertungen ansehen](#)

Preis: **EUR 79,90/Stk.**
(inkl. MwSt.)

Sie haben schon einmal bei diesem Verkäufer gekauft

369 verkauft	147 Beobachter
--------------	----------------

Der QFSP+ Zusatz lieferte mir auch gleich die passenden Kabel dazu:



Mellanox QSFP Kabel 40Gbps SFF-8436 - SFF-8436 1m - MC2206130-001
★★★★★ [Schreiben Sie die erste Rezension.](#)

Artikelzustand: **Vom Verkäufer generalüberholt**

Stückzahl: Mehr als 10 verfügbar / [8 verkauft](#)

Preis: **EUR 39,00**
(inkl. MwSt.)

Sie haben schon einmal bei diesem Verkäufer gekauft

30 Tage Rücknahme

Versand: **EUR 9,95** Standardversand | [Wieviele Details](#)

Achtung: wenn ihr Infiniband/Ethernet mit 56GB/s nutzen wollt dann solltet ihr darauf achten ein Kabel mit der Bezeichnung **MC2207130 (56Gbps)** nehmen anstatt dem hier gezeigten **MC2206130 (40Gbps)**.

Man findet oft Angebote mit dem Zusatz **MCX354A-„QCBT“** oder **MCX354A-„FCBT“** - gemeint ist damit, das im **Ethernet-Modus** die Karten mit **40GbE (FCBT-Variante)** bzw./oder nur einem Viertel der Übertragungsrate also **10GbE (QCBT-Variante)** arbeiten. Kaufen kann man beide Varianten da wir später noch durch „**crossflashen**“ der **Firmware** auch aus einer QCBT-Karte eine vollwertige FCBT Karte machen.

FCBT = F(ull) C(hannel) B(andwith) T(ier)

QCBT = Q(arter) C(hannel) B(andwith) T(ier)

Ordering Part Number	VPI Ports	Dimensions w/o Brackets
MCX353A-QCBT	Single QDR 40Gb/s or 10GbE	14.2cm x 5.2cm
MCX354A-QCBT	Dual QDR 40Gb/s or 10GbE	14.2cm x 6.9cm
MCX353A-FCBT	Single FDR 56Gb/s or 40/56GbE	14.2cm x 5.2cm
MCX354A-FCBT	Dual FDR 56Gb/s or 40/56GbE	14.2cm x 6.9cm

*This product brief describes hardware features and capabilities. Please refer to the driver release notes on [mellanox.com](https://www.mellanox.com) for feature availability or contact your local sales representative.

Quelle: https://www.mellanox.com/related-docs/prod_adapter_cards/ConnectX3_VPI_Card.pdf

Das wars dann auch schon mit der Einkaufsliste, da diese Hardware aus dem (outdated) Server-Bereich kommt, gibt es davon zurzeit relativ viele Angebote.

Nach dem Einbau der Komponenten und der Verkabelung kommen wir nun zum Einrichtungsvorgang. Was wir dazu alles benötigen hier nachfolgend:

Firmware:

<https://www.mellanox.com/support/firmware/connectx3en>

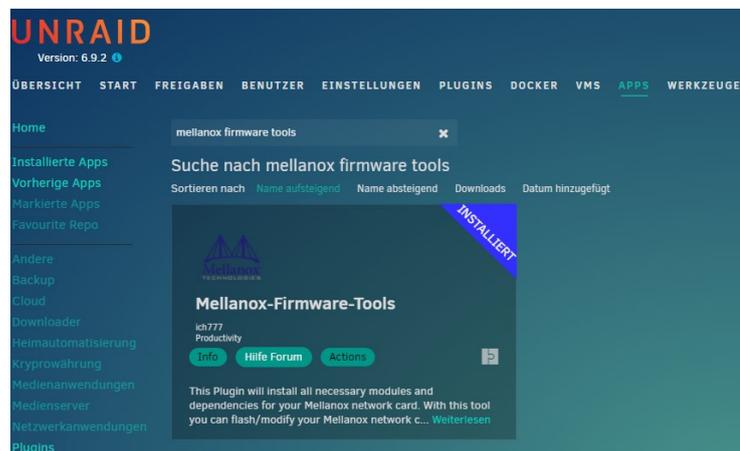
Flashsoftware für die Kommandozeile:

<https://www.mellanox.com/products/adapter-software/firmware-tools>

Windows Treiber und Tools:

<https://www.mellanox.com/products/adapter-software/ethernet/windows/winof-2>

Unraid Mellanox-Firmware-Tools: (in den Community Apps die Mellanox-Firmware-Tools herunterladen)



Firmware-flash:

(auch um 10GbE Connect-x3 Karten auf Full Bandwith 40GbE zu nutzen)

Flashvorgang unter Windows:

```
mst status
```

```
flint -d mt4099_pci_cro -i "firmwaredatei.bin" -allow_psid_change burn
```

Flashvorgang unter Linux:

```
lspci
```

```
mstflint -d 01:00:0 'firmwaredatei.bin' -allow_psid_change burn
```

Zuerst beide Karten in den Ethernet Modus setzen:

unter Linux:

```
mstconfig -d 01:00.0 s LINK_TYPE_P1=ETH  
mstconfig -d 01:00.0 s LINK_TYPE_P2=ETH
```

(nur bei den Dual-Port Karten)

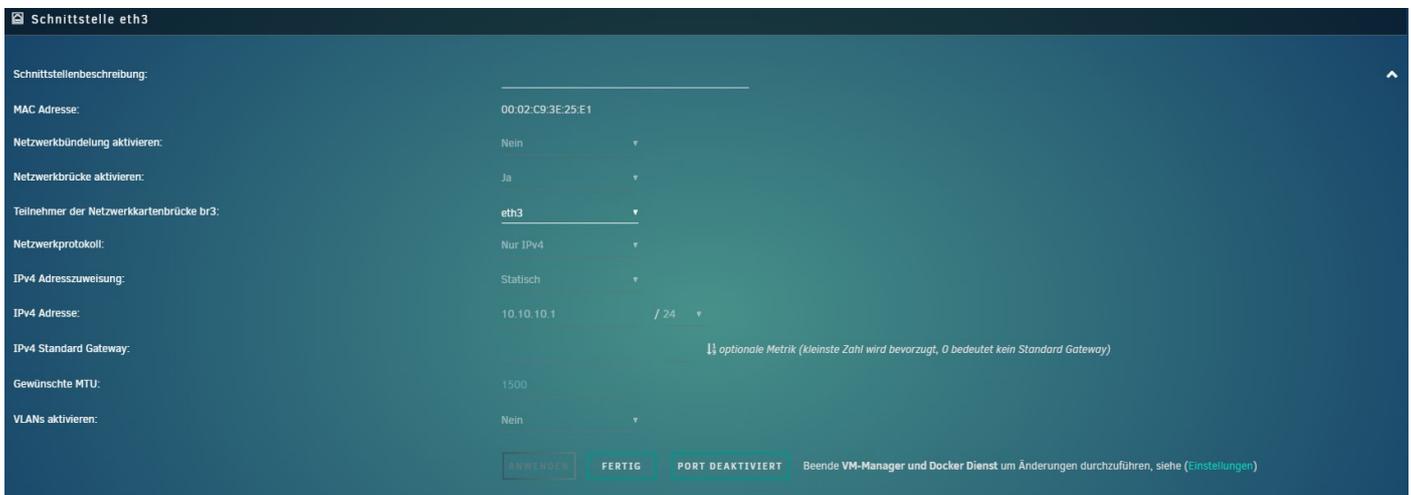
unter Windows:

entweder im Gerätemanager unter Port Config oder in der Kommandozeile wie folgt:

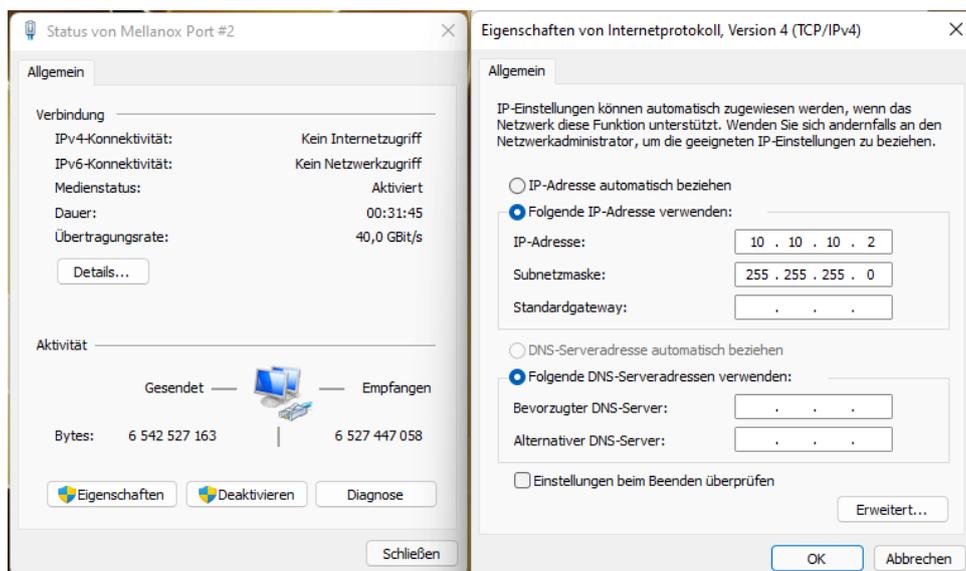
```
mlxconfig -d mt4099_pciconfo s LINK_TYPE_P1=ETH  
mlxconfig -d mt4099_pciconfo s LINK_TYPE_P2=ETH
```

(nur bei den Dual-Port Karten)

Server-Schnittstelle konfigurieren wie folgt: (ohne Gateway-Adresse)



Windows Schnittstellen wie folgt: (ohne Gateway- und DNS-Server Adressen)



Für ein *Peer-to-Peer* Netzwerk (ohne Switch dazwischen) in der Windows Hosts-Datei die „beiden“ IP-Adressen hinzufügen:

```
'C:\Windows\System32\drivers\etc\hosts'
```

```
10.10.10.1      server
10.10.10.2      client
```

Nun sollte die Verbindung beider Rechner stehen und das Netzwerk mit 40 Gb/s laufen.

Damit läuft mal der grundlegende Verkehr zwischen den Peers, möchte man nun noch die Connect-x3 Karte nutzen um damit **auch Internet über die Leitung des anderen Peers routen**, müssen wir im Linux Server noch die Firewallregeln dafür festlegen und IPv4-Forwarding aktivieren.

Dazu gehen wir wie folgt vor...

(Voraussetzung dafür ist das , iptables ' installiert ist – in Unraid ist dieses Paket schon standardmäßig für die Firewallregeln mit dabei, muss also nicht nachinstalliert werden.)

Also zuerst sehen wir mal nach wie die Schnittstelle heißt mit dem Befehl:

- (wir suchen hier nach der IP-Adresse auf einer der Schnittstellen die wir vorher vergeben haben)

ifconfig

```
br3: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 10.10.10.1 netmask 255.255.255.0 broadcast 0.0.0.0
    ether 00:02:c9:3e:25:e1 txqueuelen 1000 (Ethernet)
    RX packets 1215913 bytes 17986892427 (16.7 GiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 2022336 bytes 2557297703 (2.3 GiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
```

Nun wissen wir das die Schnittstelle *br3* (falls Bridging aktiviert ist) heißt bzw. *eth3* (wenn kein Bridging aktiviert ist) Dies muss nachfolgend beachtet, und die farbig markierten Stellen mit dem Namen eurer Schnittstelle ausgetauscht werden.

Firewallregeln per iptables ergänzen: (192.168.0.0/24 steht hier für das **Subnetz eures bestehenden LAN-Netzwerks**)

```
iptables -w -t nat -A POSTROUTING -s 192.168.0.0/24 -o br3 -j MASQUERADE &&
echo "iptables -w -t nat -A POSTROUTING -s 192.168.0.0/24 -o br3 -j MASQUERADE" >> /boot/config/go
```

```
iptables -w -t filter -A FORWARD -m conntrack --ctstate ESTABLISHED,RELATED -j ACCEPT &&
echo "iptables -w -t filter -A FORWARD -m conntrack --ctstate ESTABLISHED,RELATED -j ACCEPT" >> /boot/config/go
```

```
iptables -w -t filter -A FORWARD -i br3 -j ACCEPT &&
echo "iptables -w -t filter -A FORWARD -i br3 -j ACCEPT" >> /boot/config/go
```

wir öffnen nun mit *vim* oder *nano* die Datei */etc/sysctl.conf* und entfernen das #-Zeichen vor der Zeile mit dem Inhalt:

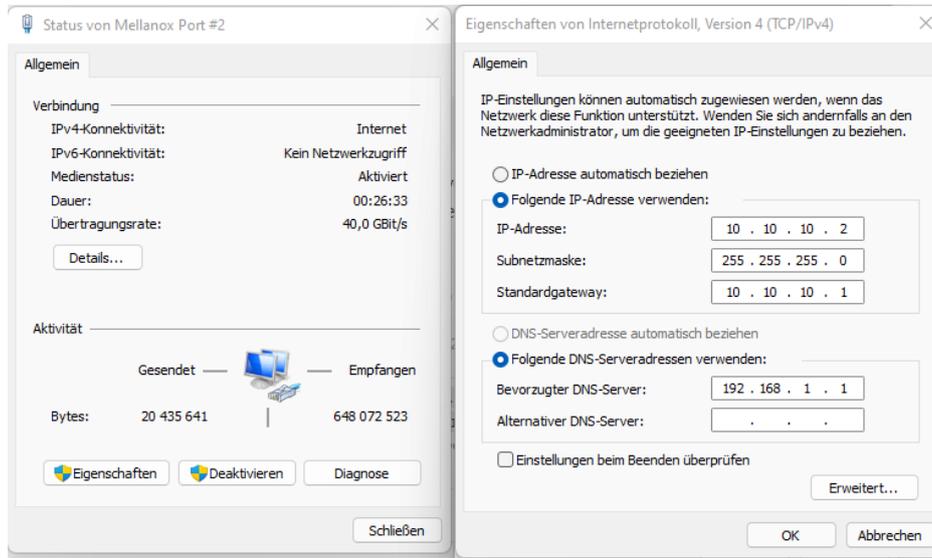
```
net.ipv4.ip_forward=1
```

abspeichern und danach mit

```
sysctl -p
```

aktivieren.

Am (Windows/Linux/whatever)-Client vergeben wir nun zusätzlich noch die IP-Adresse des Servers(Unraid) als Gateway und die Adresse des Routers als DNS-Server:



Nach einem Reboot des Windows-Clients sollte nun auch das Internet laufen und ihr könnt die Schnittstelle am Mainboard anderwärtig verwenden.

Wollt ihr auch aus eurem bestehenden Subnetz heraus auf den Adressbereich der Mellanox Karten zugreifen, müsst ihr hierzu eine statische Route im Router anlegen wie folgt:

Adressbereich	Subnetzmaske	Gateway	Verbindung
10.10.10.0	255.255.255.0	192.168.1.100	LAN

Falls ihr aus einer VM am Unraid-Server eure Karte im PC nicht pinggen könnt dann ist es auch hier notwendig mit Routen in erwähnter Form zu arbeiten und die **Firewall Regel unter erweiterte Einstellungen -> Eingehende Regeln** wie folgend zu aktivieren:

Überwachung für virtuelle Computer (Echoanforderung - ICMPv4 eingehend) Überwachung für virtuelle Computer Alle Ja Zulassen

Falls erforderlich (wenn kein ping möglich ist zum Client-PC) dann noch hier wie folgt für -

- Linux Clients statische Route:

```
ip route add 10.10.10.0/24 via 192.168.1.100
```

(Diesen Befehl könnt ihr danach in die Datei `/etc/rc.local` eingeben und abspeichern um bei einem Neustart die Route wieder aktiv zu halten.)

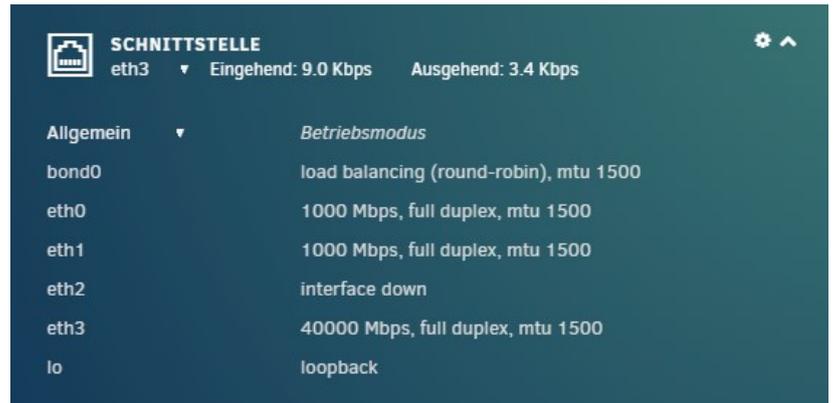
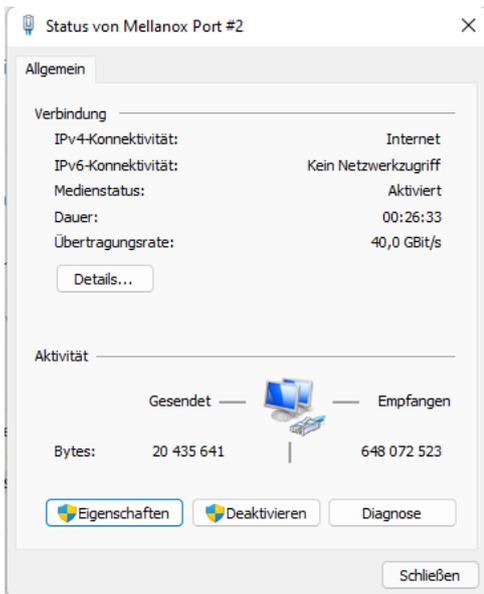
- Windows Clients statische Route:

```
route ADD -p 10.10.10.0 MASK 255.255.255.0 10.10.10.1
```

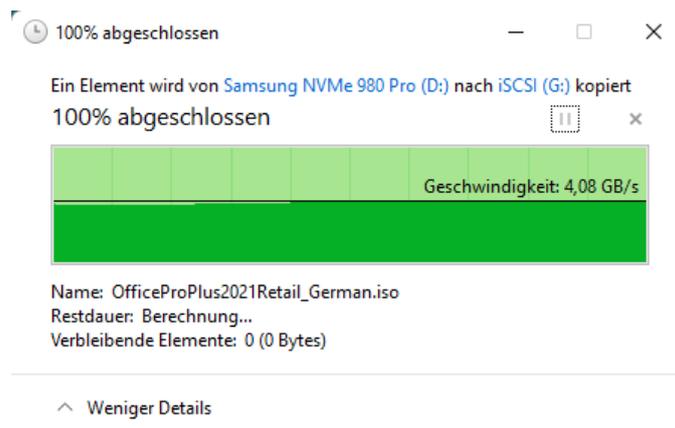
und damit sind wir fertig! 👍

Wenn euch diese Anleitung gefallen hat, oder ihr auch über andere Projekte in diesem Bereich erfahren wollt, lasst es mich wissen 😊

Nachfolgend noch ein paar Eindrücke zur erfolgreichen Umsetzung:



iSCSI Laufwerk (Samsung NVMe 980 Pro per Passthrough)



SMB-Share als Netzlaufwerk eingebunden: (Samsung 840 Evo per SATA mit aktiviertem Cache)

