

**HLRN:** „Supercomputer werden brüderlich geteilt“

Aus den Erfahrungen mit dem NVV<sup>1</sup> entstand der „Norddeutsche Verbund für Hoch- und Höchstleistungsrechnen (HLRN-Verbund)“: Beteiligte Bundesländer sind Berlin, Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, Schleswig-Holstein und (seit dem 14.12.2012) Brandenburg.

Vorgesehen war ein einziger Supercomputer – doch in welchem Land sollte er installiert werden? An Berlin konnte man aus politischen Gründen nicht vorbeigehen, ebenso nicht an dem großen Bundesland Niedersachsen. Also entschied man sich für zwei identische Rechner beim ZIB in Berlin und im RRZN mit einer „Ein-System-Eigenschaft“: Das verteilte Gesamtsystem wurde einerseits für Anwender und Betreiber als eine Einheit angesehen, andererseits waren beide Systeme auch unabhängig voneinander betreibbar. Im Vordergrund stand die Betrachtung aus Nutzersicht: Die durch die Verteilung der Rechenressourcen vorhandene Komplexität sollte, soweit als möglich, verborgen bleiben und Leistungseinbußen, die zweifellos auftraten, sollten möglichst gering gehalten werden.

Im Juli 2001 trat das „Verwaltungsabkommen über den Norddeutschen Hoch- und Höchstleistungsrechner“ in Kraft. Wesentliches Ziel des Verwaltungsabkommens ist die *"gemeinsame(n) Förderung des Hoch- und Höchstleistungsrechnens in der Absicht, die bestehende regionale Infrastruktur in Wissenschaft und Wirtschaft durch den Aufbau und Betrieb eines Norddeutschen Verbundes für Hoch- und Höchstleistungsrechnen (HLRN-Verbund) als gemeinsame Verbundaufgabe zu verbessern."*



<sup>1</sup> Norddeutscher Vektorrechnerverbund (NVV) zur gegenseitigen Nutzung der Vektorrechner in den drei Bundesländern Berlin, Niedersachsen und Schleswig-Holstein, 1984 gegründet

## HLRN I: IBM

Das von der Fa. IBM im Jahr 2002 gelieferte HLRN-I-System bestand an beiden Standorten aus je 16 Rechenknoten IBM pSeries 690 mit jeweils 32 Prozessoren. Die peak performance lag bei zweimal 2,6 TFlop/s, der Hauptspeicher bei zweimal 1,280 GByte, der Plattenspeicher bei zweimal 26 TByte. Beide Systeme waren durch eine 1 Gbit/s Datenverbindung des DFN-Vereins direkt verbunden. Betriebssystem war AIX, die UNIX-Variante für IBM-Systeme.

Die Rechenleistung dieses komplexen IBM/HLRN-Systems entspricht etwa dem Faktor 100.000 zur CYBER 76 und ca. 1.000 zur S400/40

Mit der Installation des IBM-Systems konnte das RRZN (bzw. das HLRN-System zusammen mit dem ZIB) für kurze Zeit in die Spitzengruppe der „TOP 500-Liste“ der installierten Rechenleistung zurückkehren – eine Position, die das RRZN vorübergehend bei Installation der CYBER 76 (auch wenn es die Liste noch nicht gab) und später noch einmal bei der S400/40 innehatte.



HLRN auf Seite 1 der HAZ!!

Da verzeiht man auch den Fehler:

Nicht Hamburg, sondern Berlin „bekommt ein eigenes Modell“.

## **HLRN II: SGI**

Das Compute-System setzt sich aus einem MPP-System und einem SMP-System zusammen, die u.a. aus wirtschaftlichen Gründen in mehreren Stufen geliefert wurden. Das MPP-System, die Hauptkomponente der Ausschreibung, war als ein massiv paralleles Rechnersystem für hochskalierende Anwendungen beschrieben. Das SMP-System war charakterisiert zur Nutzung von Anwendungen, die einen großen Hauptspeicherbedarf pro Knoten haben. Das SMP-System sollte etwa 10% der Leistung des Gesamtsystems besitzen. Die Inbetriebnahme der einzelnen Teilsysteme begann 2008 und endete 2010:

- MPP-1: SGI ICE-1, 5 Racks, 2.560 Cores, 30,7 Tflop/s peakperformance
- SMP-1: SGI XE, 3 Racks, 384 Cores, 4,6 Tflop/s peakperformance
- MPP-2: SGI ICE-2, 15 Racks, 7.680 Cores, 90 Tflop/s peakperformance
- SMP-2: SGI UltraViolet, 5 racks, 2.432 Cores, 22 Tflop/s peakperformance

Eine Besonderheit für speicherintensive Anwendungen war das SMP-2-System UltraViolet, da eine Anwendung in einer einzigen Linux-Instanz bis zu 496 Cores (1 Rack) und damit 1,9 TByte Hauptspeicher nutzen konnte.

## **HLRN III: Cray**

Die mit dem HLRN II begonnene Linie, neben den mächtigen MPP Systemen auch eine vergleichbar kleine SMP Komponente verfügbar zu haben, wurde beibehalten. Im Herbst 2013 wurden in der 1. Ausbaustufe jeweils in Berlin und Hannover je 4 Cray XC30 Racks installiert, ausgestattet mit insgesamt 744 Intel Xeon Rechenknoten je Standort.

Hannover erhielt zusätzlich eine auch auf Intel Xeon basierende SMP Komponente mit 32 MEGWARE Rechenknoten. Im Herbst 2014 wurde in der 2. Ausbaustufe das MPP System in Berlin um 6 und in Hannover um 5 Cray XC40 Racks erweitert, jeweils bestückt mit neueren Intel Xeon Prozessoren. Dadurch kamen in Berlin 1128 und in Hannover 936 Rechenknoten hinzu.

Die SMP Komponente Hannover wurde um weitere 32 MEGWARE Rechenknoten erweitert.

Im Endausbau besteht somit die MPP Gesamtkonfiguration im HLRN-III aus insgesamt 3552 Rechenknoten (1872 in Berlin, 1680 in Hannover). Diese Knoten beinhalten insgesamt 85248 Rechenkerne (Cores). Daraus ergibt sich eine theoretische Spitzenleistung (Peak Performance) von 2.7 PFlops/s. Das ist mehr als das 10-fache gegenüber dem HLRN II.

An Arbeitsspeicher stehen 222 TByte zur Verfügung (117 TByte in Berlin, 105 TByte in Hannover). Die SMP Komponente umfasst mit seinen 64 Knoten 2304 Rechenkerne und 24 TByte Arbeitsspeicher.

Das Betriebssystem basiert auf SuSE Linux Enterprise Server (SLES 11). Die Prozesskommunikation und Dateisystem I/O findet über das Cray Aries Netzwerk statt. Die beiden Standorte sind mittlerweile durch eine dedizierte optische 10Gbit/s-Verbindung gekoppelt, über die auch gegenseitig jeweils 0.5 PByte als NFS-Dateisysteme montiert sind.

Zudem stehen in jedem Standort 3.7 PByte unter einem Lustre-Paralleldateisystem zur Verfügung.

→ [www.hln.de](http://www.hln.de)

**2018 (?) HLRN IV**, der 7. niedersächsische Supercomputer

*wird nicht in Hannover installiert!*

„Aus Gründen, die ich als Außenstehender (allerdings habe ich in Hannover studiert) nicht erörtern oder kommentieren will, wird demnächst der dortige Höchstleistungsrechner aus dem norddeutschen Rechnerverbund nach Göttingen verlegt. Der Name RRZN wurde bereits durch den unsäglich schrecklichen Namen LUIS (Leibniz Universität IT Services) ersetzt. Dies hat viele, vor allem ältere Mitarbeiter hart getroffen.“

(Wilhelm Held, ehem. Leiter des RZ der Uni Münster, Chronist der deutschen Hochschul-RZ)