

Nützliche Abfälle: Vom WWW zum Grid

Erschienen am 18.10.2004 im Magazin für Computertechnik *c't*

<http://www.heise.de/ct/artikel/Nuetzliche-Abfaelle-289596.html>

Angela Meyer,

Das CERN, der Geburtsort des WWW, wird 50 - und spielt nach wie vor in der Forschung ganz vorne mit. Um den Large Hadron Collider, sein bisher größtes Projekt, nutzen zu können, treibt die Großforschungseinrichtung ein weiteres Mal auch die Entwicklung der Computerwelt mit voran.

Grundlagenforschung wie die vom Europäischen Labor für Teilchenphysik (CERN) bei Genf sucht eigentlich „nur“ nach Antworten auf Fragen wie die nach dem Ursprung des Universums oder den Prinzipien, die die Welt im Innersten zusammenhalten. Für die Suche nach dem Allerkleinsten sind die allergrößten Versuche notwendig - die Beschleunigerringe für die Experimente der Teilchenphysik haben mehrere Kilometer Umfang. Bei solchen baulich wie technisch herausfordernden Projekten fallen ähnlich wie in der Raumfahrt auch Antworten auf Detailfragen ab, die nicht nur bei Industrievorhaben, sondern auch im Alltag ganz unmittelbar von Nutzen sind. Das bekannteste Abfallprodukt der von 36 Staaten und internationalen Institutionen unterstützten CERN-Forschung ist das World Wide Web. Vor gut 15 Jahren machte der damalige CERN-Wissenschaftler Tim Berners Lee seinem Chef einen Vorschlag, wie man in der Großforschungseinrichtung mit tausenden Forschern, von denen ein Großteil jeweils nur wenige Monate oder Jahre dort arbeitet, Informationen dauerhaft für alle zugänglich halten kann, ohne eine zentrale Datenbank mit relativ starren Strukturen aufzubauen, die enormen Pflegeaufwand bedeutete. «CERN hat jetzt einige Probleme, denen sich der Rest der Welt bald gegenübersehen wird », heißt es in dem Vorschlag, aus dem Lee die Konzepte entwickelte, die sich hinter den Kürzeln HTML, http und URL verbergen. Wie recht er hatte - mit seiner Prognose wie mit seiner Lösung -, zeigte der rasche

Siegeszug des WWW in den Neunzigern.

Auch das Web war aber noch nicht die Antwort auf alle Fragen - mit Webserver und Browser alleine lassen sich nicht alle Probleme bei der Zusammenarbeit in großen Projekten lösen. Für das weltweit größte Projekt der Teilchenphysik, den Large Hadron Collider (LHC) und die in ihm ab 2007 geplanten Experimente, muss man jetzt wieder die Grenzen des technisch Machbaren an vielen Stellen ein gutes Stück weiter ausdehnen und dabei auch nochmal neue Wege im Umgang mit Informationen gehen.

Herausforderung Datenflut

In jedem der rund 10 000 Tonnen schweren Detektoren, die in mehreren großen Kavernen entlang des Rings aufgebaut werden, laufen hunderte von Millionen Messkanäle, um die sehr seltenen Zusammentreffen von Materieteilchen aufzuzeichnen. Auch nach der sofortigen Datenreduktion und -komprimierung werden im CERN für die LHC-Experimente pro Jahr immer noch etwa 109 Ereignisse und insgesamt mehr als 15 Petabytes (10^{15} Bytes) an Daten für die Analyse gespeichert und für Forscher weltweit zugänglich gemacht werden müssen.

Ähnliche Anforderungen gibt es auch in anderen Forschungsprojekten, vor allem in der Biologie, der Medizin, in der Klimaforschung und Erdbeobachtung oder der Astronomie. Weltweit sind daraus etliche Grid-Projekte entstanden Grid - damit ist hier, salopp gesagt, „Computer aus der Steckdose“ gemeint. Ähnlich wie bei der Information via Internet soll für den Benutzer völlig egal sein, ob die Daten, Programme oder Rechnerkapazitäten, die er nutzen will, lokal auf seinem Laptop zur Verfügung stehen oder in einem Rechenzentrum jenseits des Ozeans. [...] ■