

## Gründe für C3800-Aufrüstung

- . Convex 3820 existiert
- . Systemsoftware und Anwendungspakete existieren
- . Anwendungspakete sind bereits installiert und laufen erfolgreich
- . Anwendungen verlangen z.T. einen HS > 512 MB
- . C3820 ist total überlastet.

Die Menge der Gesamtaufgaben ist z.Z. so groß, daß Jobs in den Queues der Convex-Rechner und des HP-Clusters tagelang, teilweise wochenlang, auf Abarbeitung warten. Es ergibt sich die Notwendigkeit Rechnungen in andere Rechenzentren auszulagern. Wir sind nicht in der Lage alle anstehenden Aufgaben im Haus bzw. auf den dezentralen Rechnern in begrenzter Zeit zu lösen.

Mit 2 CPU mehr können durchschnittlich 10 Programme zur gleichen Zeit **zusätzlich** an der C3800 abgearbeitet werden und die 1 GB HS stehen dann allen 20 an den 4 CPU durchschnittlich aktiven Anwendungen zur Verfügung.

Um mit Clusterausbau einen ähnlichen Effekt zu erzielen müßten

- . 3 Clusterknoten HP735 mit je 4 GB Disc und 400 MB HS  
angeschafft werden, Preis ca. 300.000,- DM.

- . die 400 MB pro Clusterknoten stehen nur Anwendungen auf den 3 neuen Knoten zur Verfügung
- . temporäre Files müssen weiterhin über NFS gehandelt werden --> lange Laufzeiten und große Netzbelastung
- . Aufgaben mit Einzelfiles > 2 GB können nicht am Cluster abgearbeitet werden --> Überlastung der C3820
- . die zusätzlichen Clusterknoten passen nicht mehr in den Schrank --> Kauf eines Schrankes

## **Wesentlicher Nachteil unseres WS-Clusters:**

### 1. fehlende Crossbar:

- jede CPU hat einen eigenen (kleinen) HS --> hoher Swap-Aufwand
- jede CPU hat eigenen Plattenspeicher
  - . begrenzt auf z.Z. intern 4 GB
  - . bei großen Anwendungen läuft I/O über das Netz --> Verweilzeiten vervielfachen sich

### 2. Begrenzung der Filegröße auf 2 GB

- einige Anwendungen laufen deshalb nicht auf dem Cluster

## **Aufgabengebiete und genutzte Programmpakete auf der C3800**

### Quantenchemie:

GAUSSIAN 90/92  
GAMES\_UK  
TURBOMOL  
HONDO7  
ADF  
GROMOS

### Agrarwissenschaften:

Zuchtwertschätzung  
(eigene Programmpakete PEST usw.)

### Physik:

hochvektorierte Programme  
(Eigenentwicklung)

### Meteorologie:

GESIMA

### Eigenschaften der Programme:

- hoher I/O-Anteil
- Einzelfiles werden oft größer als 2 GB
- der Plattenbedarf für Einzelanwendungen kommt in den Bereich 4-6 GB

1. Die hochvektorierten Programme der Hochenergiephysik und der Festkörpertheorie haben Programmlaufzeiten auf der Convex, die durchschnittlich 3 mal schneller sind als auf einer Workstation. Die CPU-Zeit auf der Convex beträgt trotzdem noch im Einzelfall mehrere Wochen.
2. Beispiel-Aufgabe für GAUSSIAN:  
Berechnung von 21 Atomen mit 60 Variablen benötigt 6 GB Plattenspeicher und 90 MB HS, dieses Aufgabenprofil ist auf einer Workstation nicht lösbar.
3. Aufgaben unter GAMESS\_UK:  
Bei großen Berechnungen werden Files über 2 GB groß, dadurch können diese Aufgaben nicht mehr auf einer Workstation gerechnet werden. Die Datenbinaries, als Ergebnis einer Rechnung sind auf den Rechnertypen Workstation und Convex unterschiedlich, daraus ergibt sich die Notwendigkeit der Festlegung auf einen Rechnertyp. Da für große Beispiele die Convex gewählt werden muß, müssen alle GAMESS-Jobs auf die Convex.
4. Agrarwissenschaftliche Programme für die Zuchtwertschätzung haben einen sehr hohen HS-Bedarf (zum Teil größer 512 MB) und die Dateninputfiles werden größer als 1 GB.

Allgemein ergeben sich bei Ein- und Ausgaben von Daten auf Workstation bei großen Datenfällen, die über NFS verarbeitet werden müssen, wesentlich höhere Laufzeiten.

Beispiel Aufgabe mit TURBOMOL:

Das Datenfile hat eine Größe von 1,8 GB, daraus ergibt sich eine verlängerte Laufzeit von 10h pro Iteration. Bei einem Programm, welches ca. 30-50 Iterationen durchführt, beträgt die Laufzeit ohne NFS-Handling 1,5 Tage pro Iteration und mit NFS-Handling verlängert sich die Laufzeit pro Iteration auf 2 Tage. Bei 50 Iterationen ergibt sich eine Laufzeit von 75 Tagen ohne NFS-Handling und von 100 Tagen mit NFS-Handling.

In Zeiten hoher Netzbelastung betragen die Laufzeiten von Aufgaben mit NFS-Datenanbindung das Drei- bis Vierfache.

### **Abarbeitung von Aufgaben im Batch-System**

Die Abarbeitung von Aufgaben im Batch-System an einer Convex ist komfortabler, da es nur hier die Möglichkeit des systemgestützten Checkpoints gibt, der einen verlustfreien Wiederanlauf ermöglicht. Da nur eine geringe Anzahl von Anwendungen eigene Restartfunktionen beinhalten, gibt es immer wieder Verluste bei Rechnungen auf dem Workstationcluster. Die Verluste sind besonders groß bei Anwendungen mit mehr als einem Monat Laufzeit.