

شبکه های محاسباتی Sun Grid Engine

گزارش فنی

پدرام حیاتی

pedram@iasbs.ac.ir

www.iasbs.ac.ir/students/pedram

مهندسی فن آوری اطلاعات 2

مقدمه

در گذشته یکی از بزرگترین مشکلات کسانی که نیاز به محاسبات طولانی و پیچیده داشتند نبود امکانات سخت افزار و نرم افزار مناسب جهت انجام محاسبات در یک زمان معقول و متناسب باعث شده بود گاهی بسیاری از محاسبات روزها، ماه ها و گاهی سالها به طول بیانجامد. در بسیاری از موارد دیده می شد که بعد از انجام یک دوره محاسبات طولانی، بعلت وجود داشتن ایرادی در قطعه کدی از برنامه کل محاسبات اشتباه از آب در می آمد و باید از سر گرفته می شد. بعد ها برای رفع این مشکلات از ابر کامپیوترها استفاده کردند که قدر پردازش بسیار بالایی داشت و متوانست محاسبات را سریعتر انجام دهد و از زمان بکاهد. استفاده از این ابر رایانه ها از دو حیث محدود می باشد:

1. قدرت پردازنده: هر قدر هم یک کامپیوتر قوی باشد باز به علت محدودیت های تکنولوژیکی سرعت

آن از یک حدی بالا نمی رود

2. محدودیت مالی. قیمت این ابر رایانه ها بسیار بالا است و قابل خریداری و استفاده توسط بسیاری از

ارگان ها نیست

انجا بود که پیشرو آن امر علوم کامپیوتر به دنبال چاره ای برای این مشکلات افتادند که نمود یکی از این مشکلات راه اندازی شبکه های محاسباتی گرید بود.

شبکه محاسباتی Grid چیست؟

به بیان ساده: مجموعه از چندین رایانه متصل به هم که از مجموع قدرت پردازنده های آنها برای انجام کارهای محاسباتی استفاده می کنند. در ساده ترین شکل Grid به شکل یک رایانه بسیار پر قدرت برای کاربر می باشد که از طریق آن می تواند به بسیاری از منابع گسترش یافته دسترسی پیدا کند. نرم افزاری که کارهایی که کاربران ارسال کرده اند را دریافت و روی این رایانه های متصل بهم پخش می کند نرم افزار مدیریت منابع است. یکی از این نرم

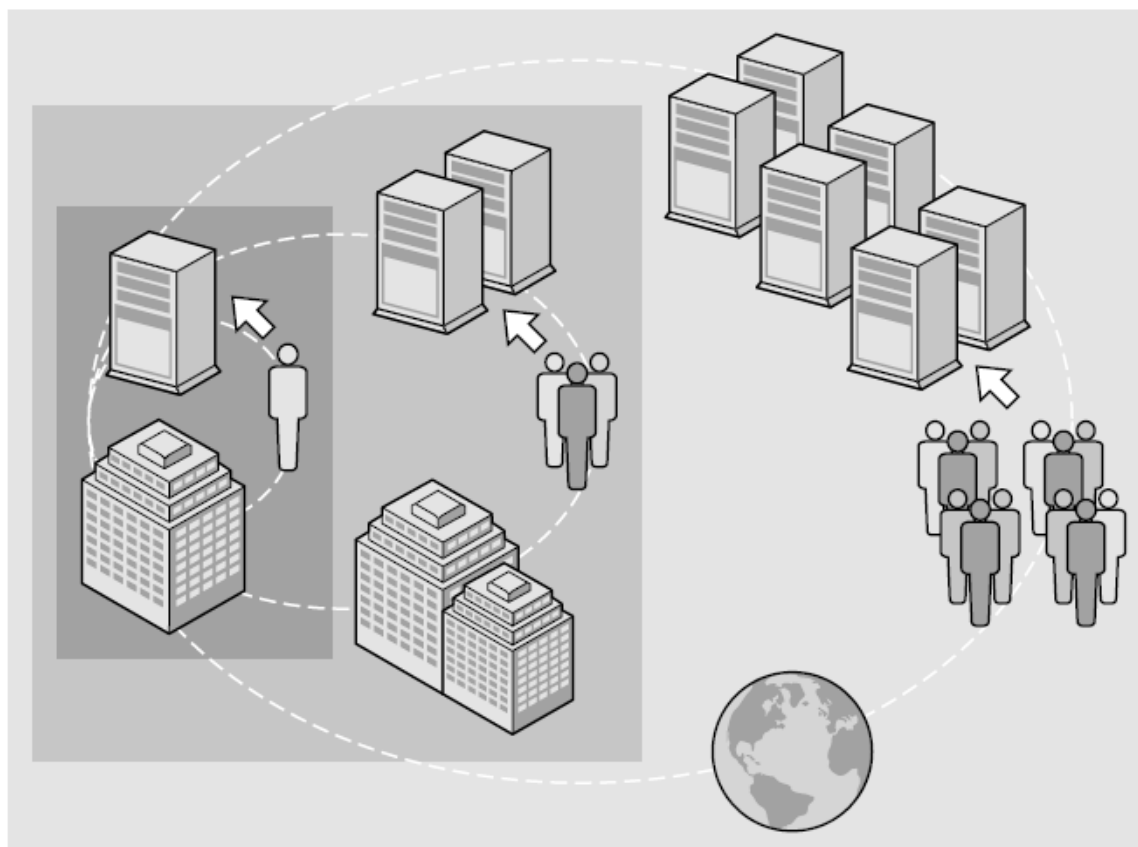
افزارها نرم افزار N1 Grid Engine 6 متعلق به شرکت Sun می باشد. کاربران در یک لحظه می توانند میلیون ها کار را وارد کنند بدون اینکه نگران نحوه و زمان اجرا شدن آنها باشند.

همانطور که در بالا اشاره شد Grid از مجموعه ای رایانه ها تشکیل شده است در نتیجه برای پیاده سازی این سیستم در سازمان های مختلف نیازی به تحمیل هزینه های مالی سنگین نیست همچنین قدرت شبکه محاسباتی Grid با افزودن رایانه های جدید به این شبکه افزوده می شود. در نتیجه یک راحل مناسب جهت انجام محاسبات پیچیده و طولانی در زمان کم می باشد.

هیچ Grid ای شبیه به هم نیستند و اندازه و شکل های متفاوتی دارند. در زیر 3 تقسیم بندی یا کلاس Grid های فعلی را که رایانه های فردی را تبدیل به یک مزرعه ابر رایانه با هزاران پردازنده را کرده است فهرست شده است:

- Grid خوشه ای ساده ترین نوع این سیستم هستند. از مجموعه ای از رایانه ها که با هم کار می کنند تشکیل شده اند و یک نقطه دسترسی به سیستم برای کاربران یک پروژه یا بخش سازمانی خاص ایجاد می کند.
- Grid پردیزه به پروژه ها و بخش های سازمانی متعدد اجازه می دهد که منابع رایانه ای خود را با هم به اشتراک گذارند. سازمانها می توانند از این نوع سیستم برای مدیریت کارهای زیاد، از فرآیند های تجاری سازمان تا داده کاوی استفاده کنند.
- Grid جهانی مجموعه ای از چندین Grid پردیزه می باشد که با بر فرآی مرزهای سازمانی گذاشته اند و یک سیستم مجازی بسیار بزرگ را ایجاد کرده اند. کاربران می توانند به منابع رایانه ی بیش از آنچه که در سازمان دارند دسترسی داشته باشند.

شکل 1 سه نوع کلاس Grid



- Cluster Grid**
 Single Owner
 Single Site
 Single Organization
- Campus Grid**
 Multiple Owners
 Single Site
 Single Organization
- Global Grid**
 Multiple Owners
 Multiple Sites
 Multiple Organizations

نرم افزار N1 Grid Engine جدید ترین محصول (در زمان نگارش این گزارش) مدیریت منابع شرکت Sun ، قوی ترین و منعطف ترین امکانات مورد نیاز برای پیاده سازی Grid پردیزه فراهم می کند. این محصول می تواند براحتی Grid خوشه ای را به Grid پردیزه مبدل کند. به بیان دیگر قابلیت گسترش بسیار بالایی دارد. در نتیجه انتخاب بسیار خوبی برای سازمان هایی که به برای اولین رو به سوی پیاده سازی Grid آورنده می باشد. نرم افزار Grid Engine با استفاده از "سیاست" های تعریف شده به دریافت کارها و گسترش آنها میان منابع می پردازد.

نکته مهم و حائز اهمیت در مورد Grid اینجاست که از ساعات مرده رایانه ها در سازمان و رایانه های اشتراکی استفاده می کند بدون اینکه نیاز به اختصاص رایانه هایی خاص برای راه اندازی Grid باشد. این بدان معناست که از کلیه رایانه های یک سازمان از رایانه مدیریت گرفته تا رایانه های داخل سایت عمومی سازمان و منشی ها می توان برای گسترش Grid استفاده نمود.

مدیریت منابع و سیاست ها

Grid Engine از قدرت سیستم مدیریت و تخصیص منابع محیط های UNIX استفاده می کند. که دارای قابلیت ها زیر است:

- تغییر زمانبندی پویا ئ مدیریت منابع به گونه این که سیاست های مختص محل را اجرا می کند.
- جمع آوری پویا داده ها تا زمانبند بتواند حداکثر استفاده از منابع ببرد.
- استفاده از قابلیت ها امنیتی و پروتکل های آن در نقل و انتقالات داده ها.
- مدیریت سطح بالا سیاست ها جهت پیاده سازی اهدافی چون: سودمندی، کاهش زمان و حداکثر استفاده از منابع.

در این سیستم کاربران می توانند کارهای دسته های (batch)، تعاملی و موازی را وارد کنند. این محصول با استفاده از قابلیت چک پوینت کارها را از یک رایانه به رایانه ای دیگر منتقل می کند بدون اینکه نیاز به مداخله کاربر باشد. در قسمت مدیریت ابزارهای توسعه یافته جهت مدیریت و کنترل کارها فراهم می آورد.

سیستم چگونه کار می کند؟

Grid Engine به نحوه زیر کار می کند:

- کارهای را از بیرون دریافت می کند. کارها درخواست های کاربران برای استفاده از منابع رایانه ای است.
- کارها را در محلی نگه داری می کنند تا زمان اجرا آنها فرا برسد.
- کارها را از محل نگه داری به جایی که باید محل اجرایی ارسال می کند.
- به مدیریت کارهای در حال اجرا می پردازد.
- به ثبت وقایع اجرای کار زمانی که کار تمام شد می پردازد.

نحوه تخصیص منابع

بزرگترین مشکلات در یک محیط مشترک تقسیم منابع است. همانطور که اشاره گردید Grid Engine از سیاست ها برای مدیریت منابع استفاده می کند. سیاست ها در Grid Engine به چهار دسته کلی تقسیم می شوند که بیشتر انعطاف پذیری را به کاربران اعطا می کند.

- اضطراری: یا استفاده از این سیاست اولویت کارها بر اساس میزان ضرورت آنهاست. میزان ضرورت کار از منابع مورد نیاز کار، زمان اتمام کار، و مدت زمان انتظار کار تعیین می شود.
- وظیفه ای: با استفاده از این سیاست، مسئول اصلی می تواند بر مبنای گروه کاربری، پروژه کاربر و ... اولویت هایی را به کارها تخصیص دهد.

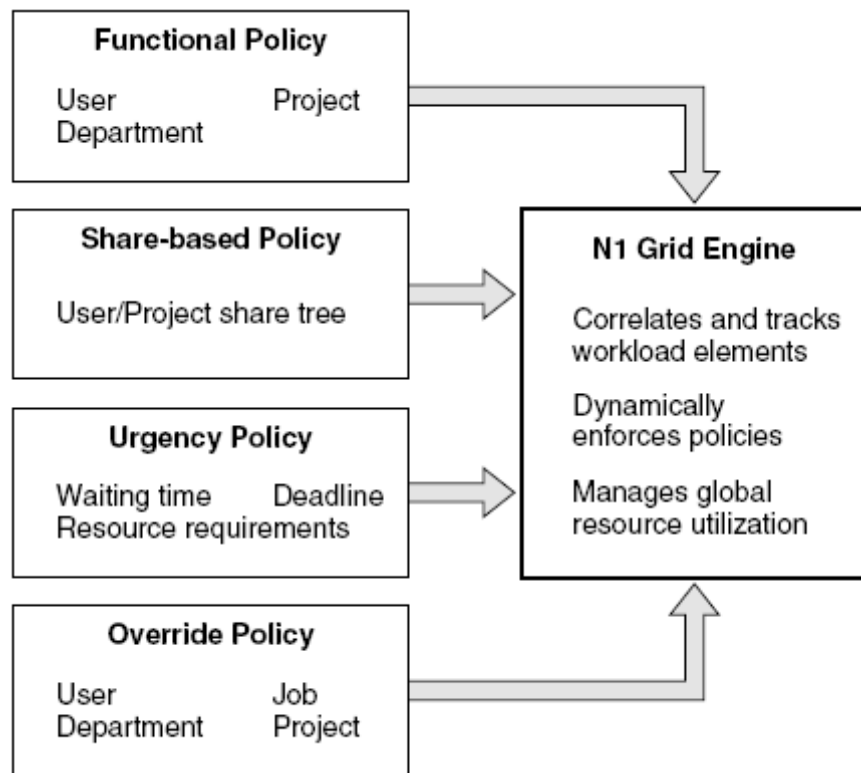
• اشتراکی: با استفاده از این سیاست، سطح سرویس دهی بستگی به کارهای گذشته کاربران، منابعی که قبلاً استفاده شده، کارهایی که در حال حاضر موجود است و میزان منابع به اشتراک گذاشته شده است.

• برتری: در این سیاست با دستکاری مستقیم مسئول اصلی بعضی کارهای اولویت بیشتری نسبت به بعضی دیگر پیدا می کنند.

استفاده از بلیط برای مدیریت سیاست ها

سیاست های وظیفه ای، اشتراکی و برتری توسط خود Grid Engine تعریف می شود. این سیاست ها از طریق دادن بلیط اجرا می شود. بلیط ها مانند سهام یک شرکت می باشد. هر چقدر شما سهام بیشتری داشته باشد شما برای آن شرکت اهمیت بیشتری داید. در نتیجه هر چقدر کاری بلیط بیشتری داشته باشد آن کار اهمیت و اولویت بیشتری دارد.

شکل 2 تصویری از سیاست های موجود در Grid Engine



اجزای فنی Grid Engine

یک Grid Engine از 4 نوع میزبان تشکیل شده است که عبارتند از:

- میزبان اصلی
- میزبان اجرایی

• میزبان مدیریتی

• میزبان ارائه

میزبان اصلی تمامی فعالیت های Grid Engine را مدیریت می کند. دو دامون (sge_qmaster daemon) و sqe_scheduler روی این میزبان در حال اجرا هستند. این دو دامون کنترل کل اجزای Grid Engine را مانند کارها، صف ها و ... دارند.

میزبان اجرایی سیستم هایی هستند که مجوز اجرا کارها را دارا می باشند. در نتیجه تنها دارای صف هستند. روی این میزبان ها دامون sqe_execd در حال اجراست.

میزبان مدیریتی وظیفه تخصیص مجوز به میزبان ها دیگر و چیزهای از این قبیل را دارد.

میزبان ارائه به کاربران اجازه ثبت کارهای خود در Grid Engine را می دهند.

نمونه آزمایشی Grid Engine در مرکز تحصیلات تکمیلی در علوم پایه زنجان

به صورت آزمایشی روی دو رایانه در مرکز تحصیلات تکمیلی در علوم پایه زنجان یک نمونه Grid Engine راه اندازی شده است. که دارای یک میزبانی اصلی، یک میزبان مدیریتی، یک میزبان اجرایی و ارائه می باشد. دو میزبان اول روی یک رایانه و دو میزبان بعدی روی یک رایانه دیگر پیاده سازی شده اند.

مشخصات فنی رایانه ها:

پردازنده: 2.4 و 3.2 گیگا هرتز

رم: 512 مگابایت

فضای هارد دیسک: 80 گیگابایت

سیستم عامل: Suse Linux 10.0

تصاویر زیر نحوه اجرا شده برنامه ساده را نشان می دهد. این برنامه تنها دو با تاریخ فعلی رایانه را با یک مکث 20 ثانیه ای در خروجی چاپ می کند. خروجی این برنامه در شاخه home رایانه ارائه دهنده بصورت یک فایل قابل دریافت است.

شکل 3- راه اندازی دامون های مربوط به سرور

```
linux:~ # /etc/init.d/sgemaster start
starting sge_qmaster
starting sge_schedd
error: commlib error: endpoint is not unique error (endpoint "linux.site/schedd/1" is already connected)
error: getting configuration: unable to contact qmaster using port 40151 on host "linux.site"
```

شکل 5- راه اندازی دامون میزان اجرای و ارائه دهنده

```
e11:/sge-root/default/common # ./sgeexecd start
starting sge_execd
```

شکل 6- سورس کد برنامه ای که برای اجرا ارسال خواهد شد. خروجی این برنامه تاریخ بروز کامپیوتر می باشد که ابتدا تاریخ اول را روی صفحه چاپ می کند سپس 20 ثانیه برنامه به حالت تعلیق در می آید و دوباره تاریخ را روی صفحه چاپ می کند.

```

Shell No. 2 - Konsole <@linux>
Session Edit View Bookmarks Settings Help

# obligations concerning the Software.
#
# The Initial Developer of the Original Code is: Sun Microsystems, Inc.
#
# Copyright: 2001 by Sun Microsystems, Inc.
#
# All Rights Reserved.
#
#####
# __INFO_MARK_END__
#
# This is a simple example of a SGE batch script
#
# request Bourne shell as shell for job
#$ -S /bin/sh
#
# print date and time
date
# Sleep for 20 seconds
sleep 20
# print date and time again
date
"/sge-root/examples/jobs/simple.sh" 47L, 1529C          47,1          Bot

```

شکل 7- ارسال برنامه برای اجرا با استفاده از دستور qsub. برنامه بعد از ارسال شماره 6 را به خود اختصاص می دهد. این شماره با توجه به تعداد برنامه هایی که در شبکه محاسباتی فعال هستند به کار اختصاص داده می شود.

```

e11:/sge-root/default/common # qsub /sge-root/examples/jobs/simple.sh
Your job 6 ("simple.sh") has been submitted

```

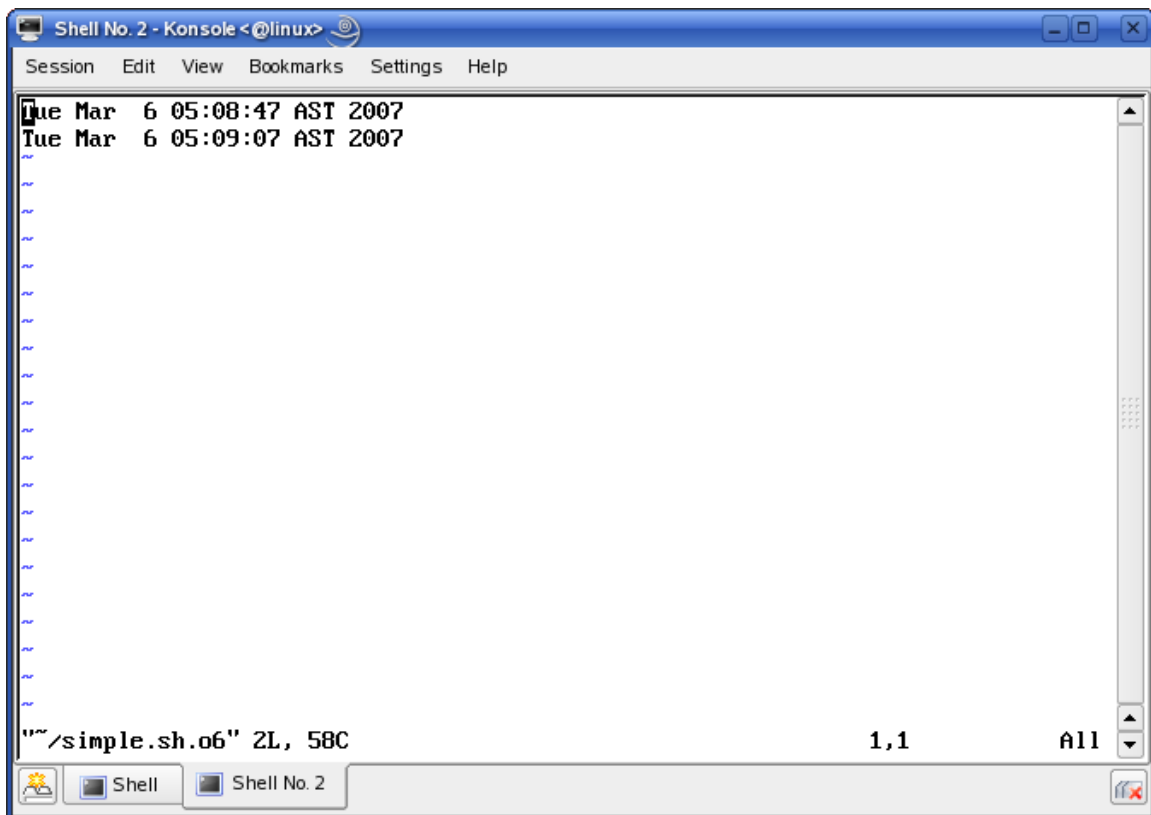
شکل 8- نمایش وضعیت صف و با وضعیت برنامه فرستاده شده یا استفاده از دستور qstat. این دستور دو بار اجرا شده است. در بار اول همانطور که در شکل مشخص است کار در وضعیت QW یعنی منتظر به سر می برد. بار بعدی که این دستور اجرا می شود کار در وضعیت I یعنی اجرا قرار گرفته است. در خروجی این دستور شماره شناسایی، کاربری که آن را وارد کرده، تاریخ ثبت و تاریخ شروع اجرا و نهایتاً صفی که کار به آن واگذار شده است مشخص است.


```

el1:/sge-root/default/common # qstat
job-ID prior   name         user          state submit/start at   queue
-----
        6 0.00000 simple.sh   root          qw    03/06/2007 16:15:35
        1
el1:/sge-root/default/common # qstat
job-ID prior   name         user          state submit/start at   queue
-----
        6 0.55500 simple.sh   root          r     03/06/2007 16:15:39 all.q@el1.site
        1

```

شکل 9- خروجی برنامه فوق بعد از اجرا در فایل به نام simple.sh.o6 به شکل زیر است



شکل 10- نمایی از برنامه qmon که یک محیط گرافیکی در اختیار کاربران برای ثبت کارها و نمایش وضعیت آن

می دهد.

