

سرور گرافیک

خلاصه

نیاز ما دسترسی به کامپیوتریه که بتونه حداقل پردازش گرافیکی رو انجام بده. از طرفی، ما برای شروع به حدودا 15 تا از این کاپیوترها نیاز داریم که ممکنه در آینده کمی بیشتر شه. روش بدیهی اینه که برای هر client یه سیستم مستقل (که پردازشگر گرافیکی هم داره) در نظر بگیریم. مشکل اصلی که با این روش رخ میده رو میشه از دو جنبه بررسی کرد:

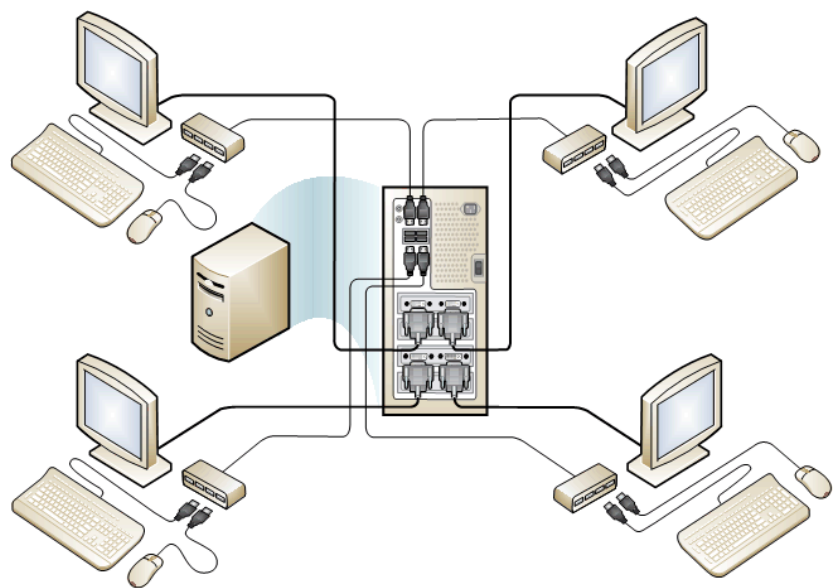
1. هزینه: هزینه این سیستم ها متناسب با پردازش گر گرافیکی که انتخاب میکنیم، میتونه خیلی زیاد باشه. این مشکل مخصوصا تشدید میشه وقتی client ها در اکثر مواقع از GPU کار نمیکشن (علا utilization پایین)
2. قدرت پردازش: اگر فرض کنیم یک پروژه با اندازه متوسط، از سقف پردازش (مخصوصا VRAM انحصاری GPU) منابع بیشتری بخواد، اونوقت دیگه مهم نیست که ما فرضا 15 تا GPU داشته باشیم، چون نمیشه GPU ها رو به هم وصل کرد (بیشتر از 2 تا)، عملا تمامی سخت افزار ما بی استفاده خواهد موند.
3. (اختیاری) scale up: بسیار بهتر خواهد بود اگه میشد کانفیک موجود رو بدون تغییر اساسی، بزرگتر کرد که دوباره سیستم های معمولی همچین توانی ندارن.

ما دوست داریم که بتونیم سه مشکل رو با هم حل کنیم که دقیقا شبیه به سیستم های cloud computing که داریم. به عبارت دیگه، هدف ما اینه که یه کامپیوتر خیلی قوی داشته باشیم (سرور گرافیکی) و تمامی client ها بدون هیچ سیستم مستقل (کیس جدا) بتونن وصل شن و حداکثر استفاده رو کنن. مشکل مخصوص ما اینه که ما کار گرافیکی میکنیم (نرم افزارهای adobe فرضا) و دوست داریم client ها بتونن تصویری که میبینن هم کیفیت خوبی داشته باشه.

مقدمه

ما دوست داریم چیزی شبیه به cloud computing که رایج هستن این روزا رو خودمون داشته باشیم. به نظر میرسه نزدیکترین تکنولوژی به این موضوع، [Windows Server MultiPoint System](#) هست. البته معادل لینوکسی این سیستم که به نظر خیلی بهتره هم ساپورت میشه هست، ولی مجددا، به خاطر نیاز به نرم افزارهایی که فقط روی ویندوز موجودن، مجبوریم که نهایتا client های ما به محیط ویندوز دسترسی داشته باشن.

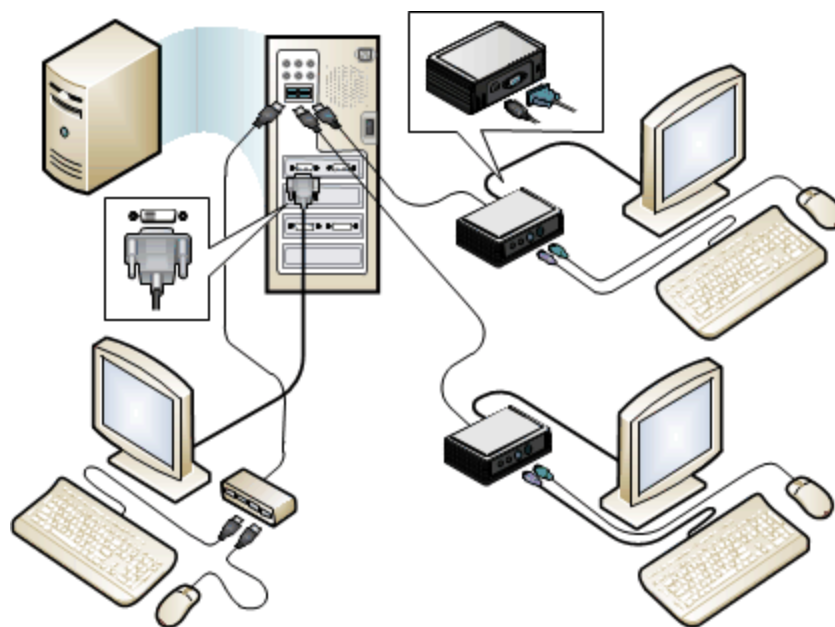
Windows MultiPoint System



ویژگی های این سیستم که برای ما جذابش میکنه:

1. مبتنی بر ویندوز و محصولات مخصوص ویندوز بهتر کار میکنن. نرم افزارهای گرافیکی (مثل adobe) نیاز اصلی ما هستن.
 2. کاربرها به صورت همزمان هر کدوم به محیط مستقلی ویندوزی در اختیار دارن
 3. هزینه کمتر راه اندازی نصب به خرید کامپیوتر مستقل برای هر کاربر. (در واقع هزینه GPU مخصوصا، چون share نمیشه – به نکته 4 نگاه کنید).
 4. هر کاربر میتونه از حداکثر توان سیستم بهره ببره (اگه بقیه کاربرها فشار نیارن). این مخصوصا خیلی مهمه چون برای مثال به پروژه متوسط motion graphic شاید حدودا 30 گیگ رم گرافیک بخواد. اینجا به جای اینکه 15 تا GPU معمولی داشته باشیم که هیچکدوم توان اجرای این رو ندارن، به GPU خیلی قوی مثل A16 داریم که میتونه اینکارو کنه و توان اشتراک گذاری هم داره.
- تصویر زیر حالتی رو نشون میده که 2 تا GPU داریم که هر GPU دو تا پورت تصویر داره. Windows Server وظیفه پخش حجم کار و غیره از طرف client ها روی GPU ها رو انجام میده. این قسمت مخصوصا به GPU انتخابی خیلی وابسته است که اتفاقا یکی از سوالات ما هم همینیه. برای نمونه، محصولات Nvidia مثل A16 به طور مخصوص از remote بودن سیستم های گرافیکی سنگین صحبت میکنه. GPU هم دوره با این محصول که مخصوص پردازش های هوش مصنوعی ساخته شده، A30 هم این قابلیت رو داره (MIG) با این تفاوت که برای نرم افزارهای گرافیکی ساخته نشده ولی توانایی استفاده وجود داره. اینجا Nvidia به طور مخصوص تکنولوژی هایی داره که به نظر دقیقا مناسب ولی هیچ ایده ای ندارم، مثل [Nvidia VirtualPC \(vPC\)](#).

برای توصیف بهتر، ما دوست داریم چیزی شبیه تصویر زیر رو ولی در تعداد بالاتر انجام بدیم:



تو این سیستم، اون کاربرهایی که بیشترین حجم مصرف گرافیکی دارن رو به صورت مستقیم به پورت تصویر GPU وصل میشن. فرض کنید ما دو تا GPU داشته باشیم که هر کدوم دو تا پورت داشته باشن. اینطوری 4 تا کلاینت (که کاربری سنگین تر دارن) وصل میشن. بقیه کاربرها، مثل تصویر بالا، با USB hub ها مانیتورها، کیبوردها و ... رو وصل میکنن و اینطوری به سیستم وصل میشن. برای نمونه فرض کنیم 10 تا کاربر (که کار گرافیکی ولی سبکتر میکنن) میتونن وصل شن.

توجه: متوجه شدیم که Windows MultiPoint دیگه ساپورت نمیشه و فقط تو Windows Server 2016 هست.

نتیجه گیری

در نتیجه به دنبال جواب این سوالیم که آیا:

1. همچنین سیستمی با نیاز نرم افزاری ما (محصولات adobe و ...) به صورت MultiPoint یا هر سیستم دیگه که امکانات متناظر رو بده، میشه ساخت؟
2. سخت افزار (و ایا نرم افزار) خاصی هست که باید تهیه کنیم؟