A Utilization Study

Dealing with Fat Tailed Distributions and Long Term Correlation

Abstract

Last year I started looking into the nature of CPU utilization. This is an important subject because business value of infrastructure is closely tied to how busy or idle the infrastructure is. I presented the paper, "Heisenberg and Utilization" at CMG '14 about the underlying mechanism from which utilization measurements are taken. This asserted that utilization distributions are not "normal". Recently I looked at utilization data from 27 servers including the 6 that were examined in that paper. (The results show that utilization distributions are indeed "fat tailed". A fat tailed distribution has significantly more instances at values further from the mean than "normal". This implies that the "Normal" statistics that we use to understand server consolidations will be optimistic when we design for typical "2-3 sigma" service levels, expecting to cover 95 to 99.7% of the peaks. In many cases the appearance of high sigma results in measured data indicates an undefined underlying variance. When considering consolidations long term correlation of data due to the "workday" cycle is also common. This paper will examine these effects and suggest methods to deal with the results.

Description of the study

This is a study of load on 27 Intel L7555 servers. These are 1.87 GHz, 2 socket servers with 16 cores per socket. The study considered the consolidation of the loads on these servers into a single load to be running on a single large server. This was done using 4 days of 5 minute interval utilization data. Statistics for each load were extracted from the interval by interval machine utilization data. Statistics were used to generate a "peak of the sum" for the consolidated load used to drive the statistical utilization model of a sizer. The actual peak of the sum was then used to override the model with the data and the results were compared.

The Distributed Workloads

We will show a series of descriptive statistics for the set of workloads graphically.

They are:

1. Uavg – the average utilization
2. P/A – the peak to average ratio
3. k – the tolerance (number of standard deviations included at the peak)
4. c – the statistical variability (standard deviation / average)
5. HR – the normalized headroom HR = kc

First we plot: Umax v Uavg

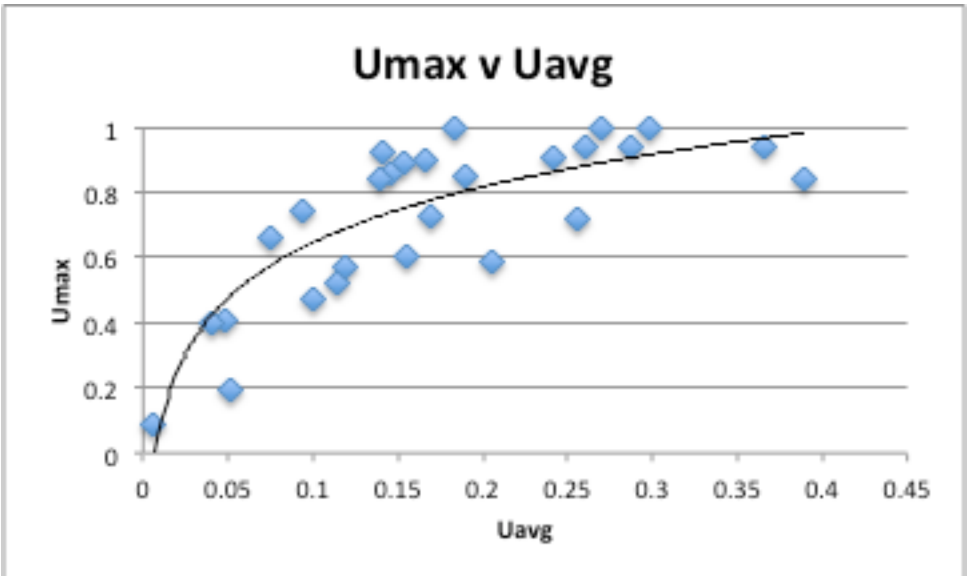


Figure 1: Umax v Uavg

We note that Umax is cannot be greater than 1. The highest peaks start occurring at about 15% average utilization. This means that low utilization implies a high P/A ratio.

We now plot P/A v Uavg:

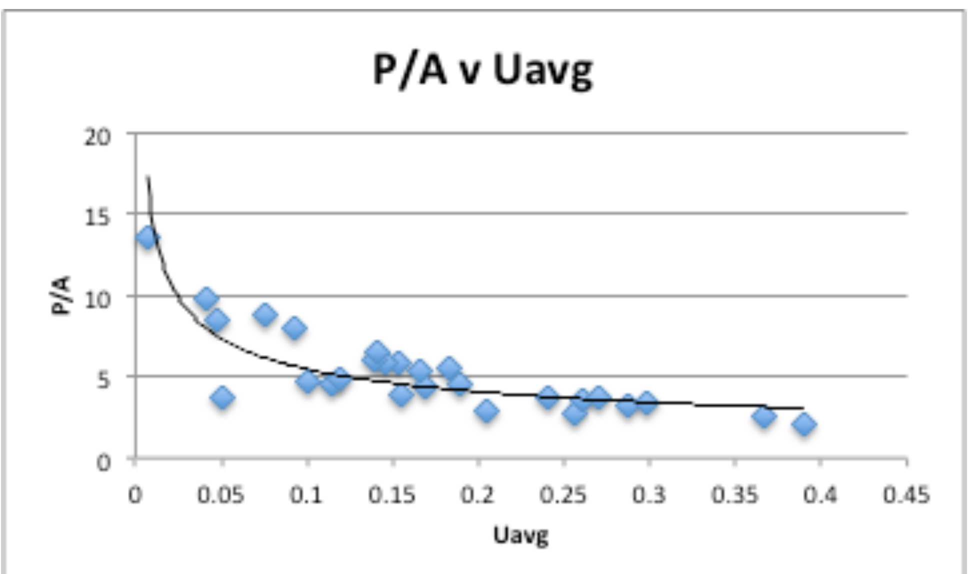


Figure 2: P/A Ratio v Uavg

As expected P/A has an inverse relationship with Uavg. Given this relationship we would expect the statistical “index of variability”, c would also have an inverse relationship to Uavg. Plotting this we find little or no correlation.

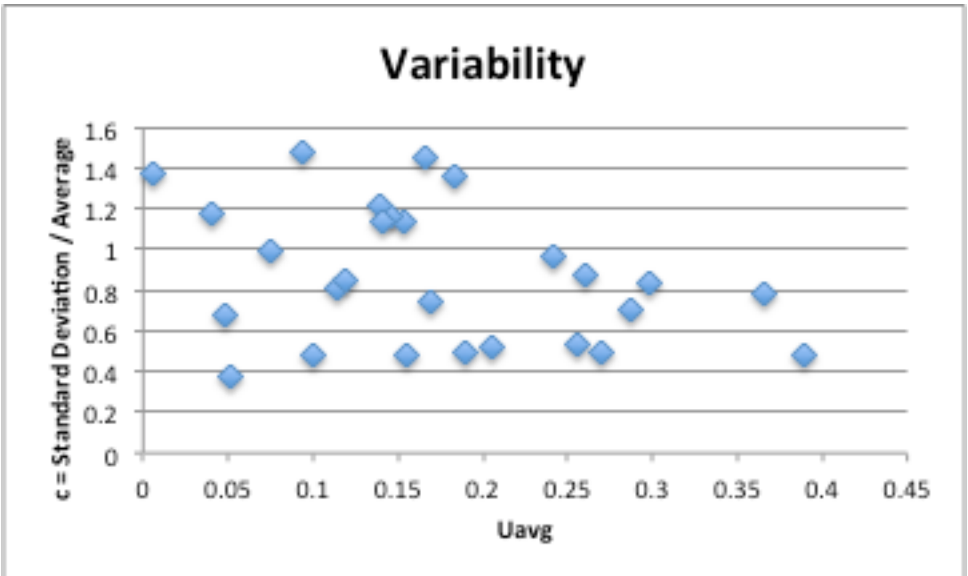


Figure 3: Index of Variability v Uavg

This illustrates the concept that fat tailed distributions have "undefined Variance". It also illustrates the notion of "burstiness" of distributed workloads. A bursty load has low overall variability, but exhibits extreme variability during relatively short intervals. The result is a high peak to average ratio and low index of variability. This is the situation that creates "fat tails".

How can we detect this from the statistics? To explore this we examine "Rogers' Equation".

$$Uavg = 1/(1+kc)$$

We note that by definition

$$Uavg = Upeak / P/A$$

If we accept that in the short run Upeak = 1 unless the machine is highly over configured, by inspection:

$$P/A = (1+kc)$$

This establishes a link between the statistical parameters k and c and the intuitive P/A ratio as the metric for variability. We define:

$$HR = kc$$

And write $P/A = 1 + HR$

We note that P/A is high when HR is high. It is possible to have low variability and a high Peak to Average ratio only if k is large. When we use the sampled data to determine the variance and average and then calculate the standard deviation and c, bursty, fat tailed loads will have low c and very high k. High k in the sampled data means is the statistical indicator of a fat tailed distribution. It should be apparent from the definition of c as the ratio of the standard deviation to the average, c becomes undefined when the Variance is undefined. In this case the statistics k and c lose their independent meanings and we are left with HR as the statistic.

Here is a plot of c v kmax for the 27 workloads in this study:

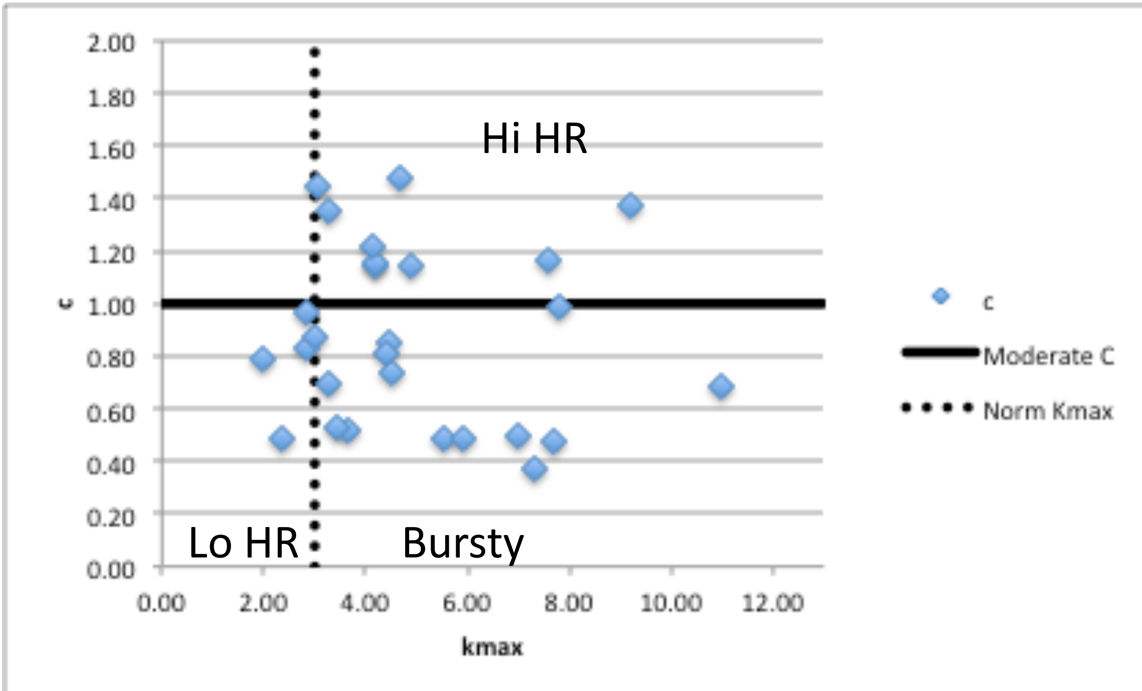


Figure 4: High HR but low variability indicates bursty loads with fat tailed distributions

You can see that many of the loads are bursty (kmax>3 and c<1) leading to "fat tails". There are also very high HR values (k>3 and c>1). As expected the P/A ratios are high even though the loads have moderate c values:

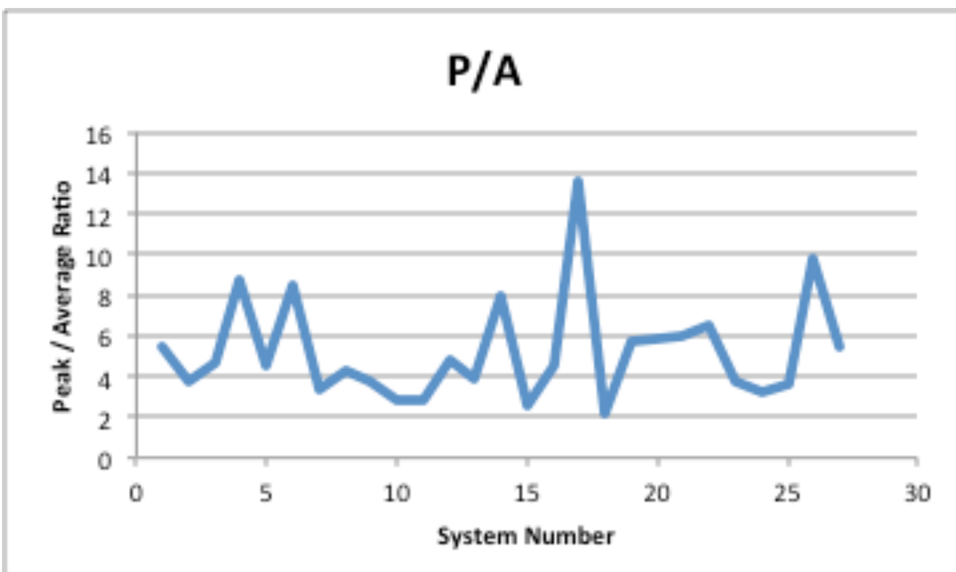


Figure 5: Most of the P/A ratios are greater than 4 indicating low average utilization

If we know the peak and the average utilization we can calculate HR

$$HR = Upeak/Uavg - 1$$

We now have two forms of the equation for utilization with HR as a parameter:

Uavg = Upeak/(1+HR)

Upeak = Uavg(1+HR)

Notice that we need any 2 of the three values to describe the usage pattern. If we measure the peak and the average we can determine HR. If we know only the average, we can assume the Upeak = 1 and determine HR.

We now turn to the statistical version of HR.

HR = kc

The statistic k is a design parameter. It is set by determining how many standard deviations from the mean will be "covered" by the solution. If we assume a "normal" distribution then k = 2 to 3 covers 95 to 97.7% of the possible peaks. We measure the sample variance to determine the standard deviation and from that calculate c. Multiplying by the design parameter k gives us HR. This allows us to determine the "covered peak".

When the Variance is undefined, c also becomes undefined. In that case HR rather than k is the design parameter. Given interval data we can still determine the sample variance and calculated a standard deviation and coefficient of variability, c. The resulting value of a c will understate the load's variability. Thus we cannot assume that the typical value of k = 2 to 3 will generate coverage of 95-97.7%. The data in this case suggests that if we use the sample variance to calculate c we need to approximately double k to get the result that we desire. Thus, we propose an initial rule of thumb that our procedure is to find the sample variance but set k = 4 to 6. To be well established this heuristic needs to be tested by more data and certainly more experience is needed to yield confidence in this practice.

The Consolidated Workload

The sum of the interval data looks like this: (What are the axis? And, you can remove the legend for a wider chart.)

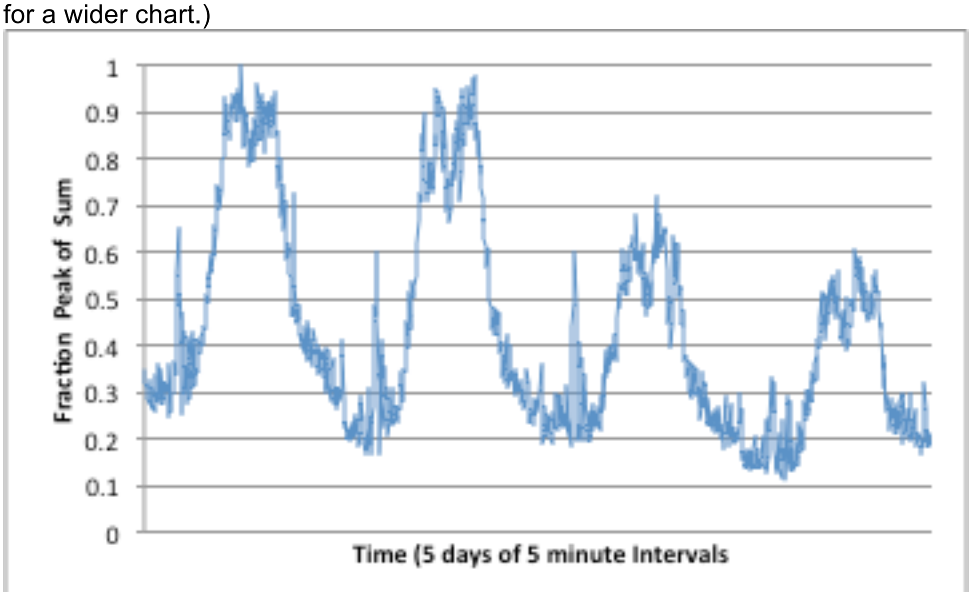


Figure 6: The Interval Data

The data on the "Y" axis is normalized to utilization with the peak interval is at 90% utilization on the target machine. The x axis is the 5 minute interval number.

The distribution of the consolidated load looks like this: (Again, label the two axes)

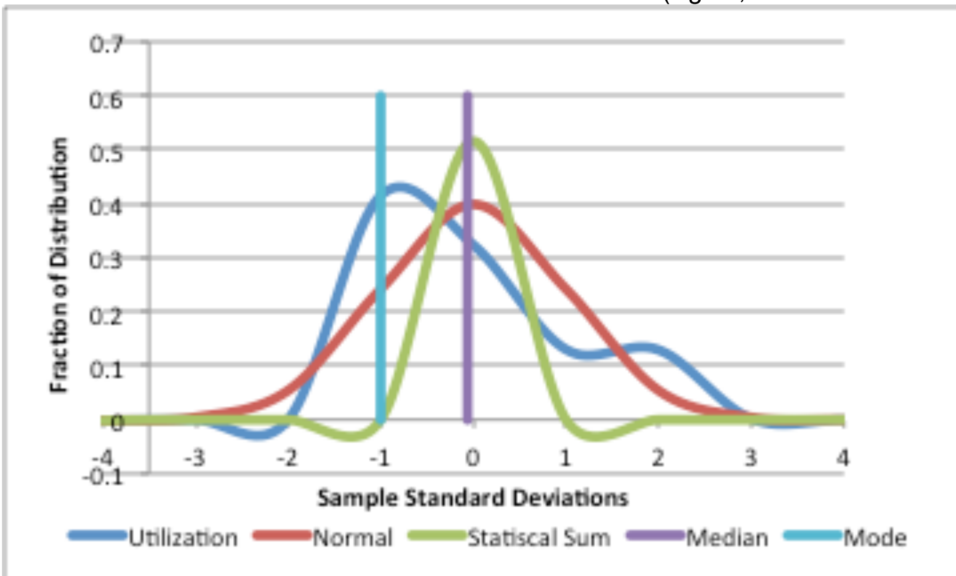


Figure 7: The distribution of consolidated load utilization

The blue line is smoothed curve of the actual interval data distribution. The red line is the "normal distribution" with the same mean and Variance as the interval data. The green line is the normal distribution but taking the "Statistical Sum" of the individual distributions. It is possible to adjust the model below to approach the normal distribution but not to approach the actual distribution. It is possible to use an Erlang, Gamma or other "skewed" distribution to model the actual interval data. Doing so moves changes the shape by moving the median and mode to the left, but it does not solve the problem of matching the "humps" in the right hand tail.

The Statistical Model

The peak of the sum of the model is generated by following model:

Average(sum) = Sum(averages)

Variance(sum) = Sum(Variances)

Stdev(sum) = Sqrt(Variance(sum))

Peak(sum) = Average(Sum) + k x Stdev(sum)

The parameter k is design parameter usually taken as 2 - 3 in order to cover 95-99.7% of the peaks of normally distributed data. In this case k = 3 was used.

Sizing results

Here are the results from the sizer using the model above and the statistics from individual loads:

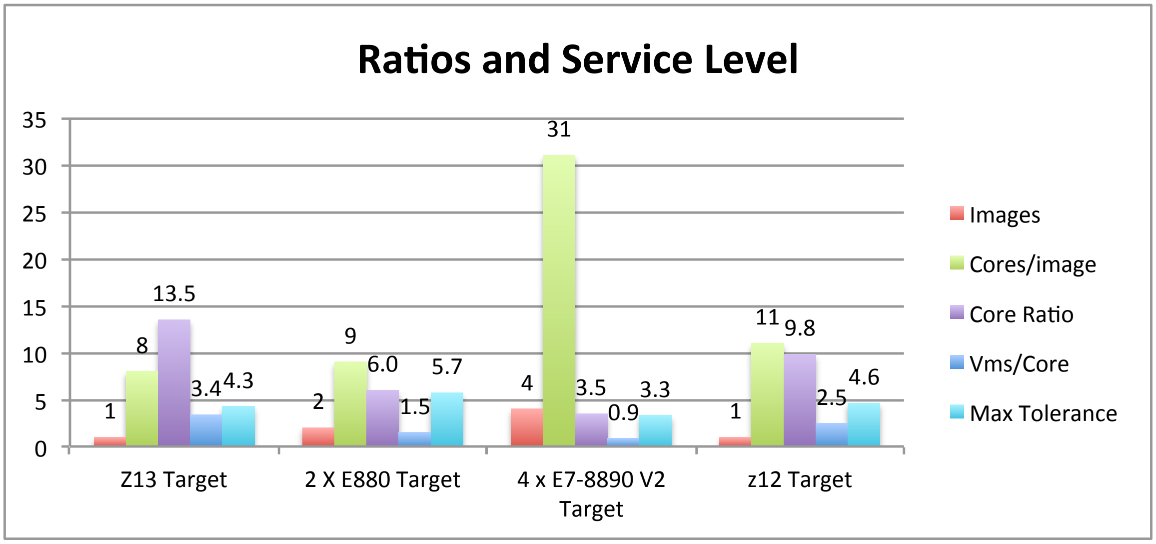


Figure 8: Modeled Sizer* (Will scrub this chart, labels are wrong and need to double check results)

Here are the results from using the actual interval data for the peak interval to drive the sizer.

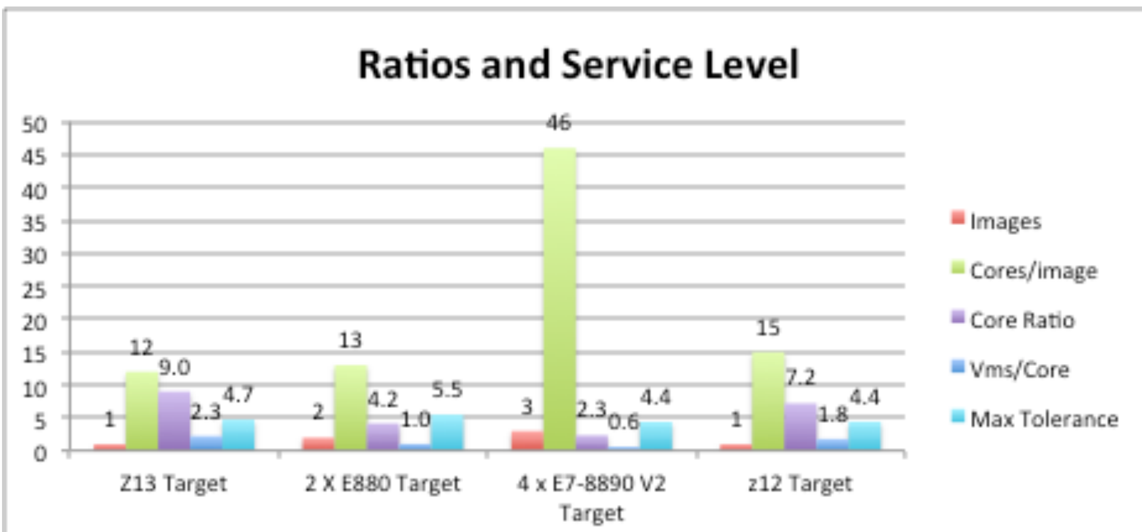


Figure 9: Sizer Results using the peak interval data (Also need scrubbing)

These results show that the while the ratios between the target machines remain similar, all the results require more cores than the statistical model predicts.

Here is a table of the resulting ratios.

Fat Tailed/Norm Ratios

	Images	Cores/Image	Core Ratio	VMs/Core
z13	1	1.5	0.7	0.7
E880	1	1.4	0.7	0.7
8890 V2	1.3	1.5	0.7	0.7
zEC12	1	1.4	0.7	0.7

The fat tailed sizing shows 40-50% more cores for all target machines

Is there an Analytic Approach?

Assume that for each component workload we have data, estimates or assumptions for the average (Uavg) and peak (Upeak) utilization. Further assume that there is a specified fraction of the possible peaks covered by the machines before overloading occurs. Our task becomes estimating the consolidated peak load from these machine parameters.

We know that:

Uavg = 1/(1+HR)
Upeak = Uavg(1+HR)

The simplest normal case

The simplest consolidation case is for workloads independent workloads of similar size and variability, that are independent and not correlated over time and are "normally" distributed.

In this case we have the following:

c_{cons} = c/sqrt(n) where n is the number of workloads consolidated

This is a direct result summing statistical addition by summing averages and summing variances for the case where averages and variances are the same.

Given c_{cons} we have:

Uavg = $1/(1+kc_{cons})$

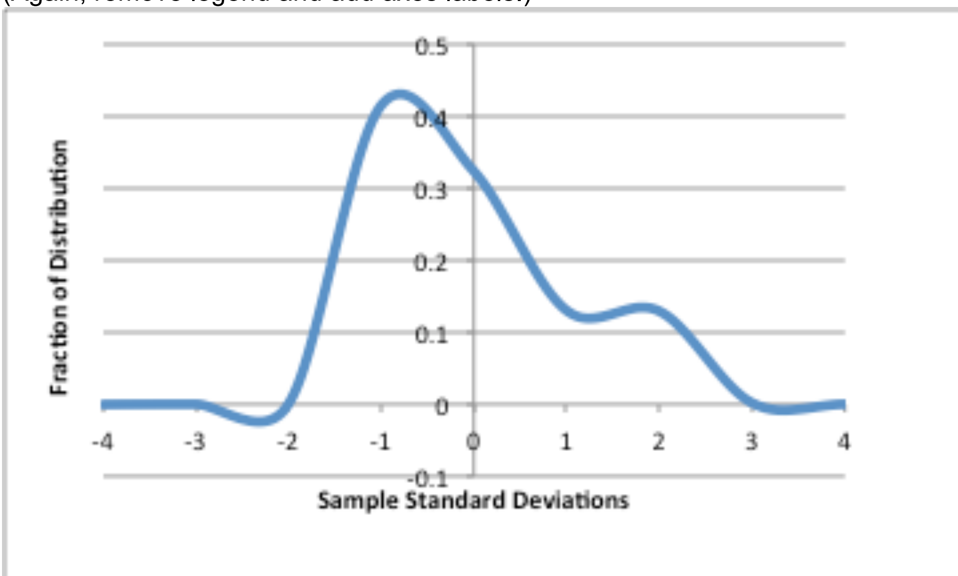
and

$U_{peak} = U_{avg} (1+ kc_{cons})$

This is in effect what the sizer does with its statistical model.

The simplest case with fat tails

Here we keep the same assumptions as above except for the assumption of normalcy. In this case the variances become “undefined”. This means that they indeterminably large. This means that we lack analytic means to determine c. Recognizing that c_{cons} underestimates the real variability we and that k is a design variable we can define k as larger than would be indicated by “normal” statistics. Here is the consolidated distribution versus k given a standard deviation calculated from the sample variance. Note the large hump at k = 2. At k=3 there is residual distribution greater than 0.3%. We choose k = 5 as being well beyond the sample distribution. (Again, remove legend and add axes labels.)



Thus we have:

Normal Case

$U_{peak} = nU_{avg} (1+ 3c_{cons})$

Fat Tailed Case

$U_{peak} = nU_{avg} (1+ 5c_{cons})$

The Fat Tail / Normal Ratio

$(1+ 5c_{cons}) / (1+ 3c_{cons})$

Here, $c_{cons} \sim 0.5$ so

Ratio = 2.5/1.5 ~ 1.4 which is what we measured in the sizing results.

Note that picking k = 5 required “judgment” and confidence will only come with experience and more data.

A Dose of Reality

We are still not out of the woods. The sum of the variances only works if the workloads are uncorrelated over time. In many cases the “workday” will create workload correlation even for independent workloads. As a result, the summation of individual variances will underestimate the variability, even if the distribution is not fat tailed. We find that taking the geometric mean of the statistical “peak of the sum” with the sum of the peaks compensates for the correlation. Correlation models can do this more precisely but the math gets complex and cumbersome for any but small consolidations. It is far easier to gather and directly use interval data. However, the geometric mean of the peak of the sums and the modeled sum of the peak is usually a better estimate than the modeled results alone.

Fat Statistics

Utilization distributions are not only fat tailed, they are “skewed”. This means that the median utilization is significantly lower than the average utilization. Another way to compensate for the underestimation of variability by using the sampled data is to calculate the “fat variance”, by using the deviations from the median rather than the deviations from the average. From this we can calculate Fat Deviation, Fat Index of Variability, and Fat HR. When the average utilization is well below 50%, the use of these statistics improves our estimate of the sum of the peaks. For the case in hand we have: (Is the “Sum Peaks” supposed to be in the Interval Data column?)

	Interval Data	Norm Stats	Fat Stats
Max	1101	566	688
	Sum Peaks	1952	1952
Geomean	*1103*	1051	1159

Notice that the correlation compensated “normal statistics” overestimates the peak of the sum and the correlation compensated “fat statistics” over estimate it. The Geometric mean of the 2 estimates closely approximates the measured max load in this case.

Conclusions:

We come to the following conclusions:

1. Specifying “2 or 3 sigma” service levels or expecting “normal” statistical modeling at k=2 or 3 to deliver 95 to 97.7% service levels is risky and will likely lead to “black swan” overload events.
2. The Fat Tailed nature of workloads even when consolidated means that design points should be k greater than 2 to 3. An initial rule of thumb is to double this to k = 4 to 6.
3. When dealing with estimated statistics and assumptions for component workloads it is best to be conservative and use higher values of k than “normal statistics” would indicate or to estimate the median and use “Fat Statistics”

4. Because variance becomes undefined, it is far better to collect actual interval data and determine the peak of the interval by interval sum than to use the statistical sum of random variables. While there is labor and expense involved in handling the data, the results are better and the math is less arcane.
5. The analytic approaches suggested above needs to be vetted by applying them to more cases allow us to develop better intuition and judgment for sizing workload consolidations when there is little usage pattern data or limited usage statistics in hand.