

SCTE 2005 Conference on Emerging Technologies

The Statistics of Switched Broadcast

Nishith Sinha

iTV Systems Engineer, Cox Communications

1400 Lake Hearn Drive

Atlanta, GA 30319

+1 404 843 5958

Nishith.Sinha@cox.com

<http://www.cox.com>

Ran Oz

CTO, BigBand Networks

S.V. Vasudevan

Chief Architect, BigBand Networks

475 Broadway

Redwood City, CA 94063

+1 650 995 5000

Ran.Oz@bigbandnet.com

S.Vasudevan@bigbandnet.com

TABLE OF CONTENTS

INTRODUCTION	3
SWITCHED BROADCAST ARCHITECTURE AND OPERATIONS.....	4
ARCHITECTURAL COMPONENTS	4
SYSTEM OPERATION.....	5
TRIAL INFORMATION.....	6
TRIAL A.....	6
TRIAL B.....	7
STATISTICAL ANALYSIS	9
TRIAL A STATISTICS.....	9
INCREMENTAL STB IMPACT ON STATISTICS.....	9
TRIAL B STATISTICS.....	11
INCREMENTAL PROGRAMMING IMPACT ON STATISTICS	12
STATISTICAL MODELING OF SWITCHED BROADCAST	15
TRIAL A.....	16
TRIAL B.....	18
ANALYTICALLY ESTIMATING CHANNEL USAGE.....	19
SUMMARY.....	21

Introduction

Switched Broadcast is a method of delivering selected digital broadcast programming only to nodes where and when subscribers are actively requesting that programming. Using such a methodology enables the creation of “virtual” programming capacity without the directly correlated physical expense of creating and dedicating spectral resources. Recent expansions of broadcast programming lineups, particularly those in the bandwidth-hungry HD format, have elevated the consideration of a Switched Broadcast programming tier as a means of providing the capacity to deliver cost-effectively incremental broadcast content.

Most practical Switched Broadcast deployments will (for economic reasons) oversubscribe the number of offered channels to the number of physical stream resources. Thus, a 150-channel Switched Broadcast tier could be carried on as few as five QAM channels containing an aggregate total of 50 stream resources. This example configuration assumes that while subscribers may have access to all 150 programs, as a group they will never watch more than 50 different programs at any one time. Much of the value proposition of Switched Broadcast is tied to the statistical efficiencies that are obtained in specifying this oversubscription ratio. As these statistics are a function of the aggregate viewing behavior and patterns of digital television cable subscribers, it becomes clear that the support of empirical data would be essential in dimensioning a system.

Proof-of-concept trials of Switched Broadcast were successfully held in 2002 and 2003, demonstrating that Switched Broadcast services could be carried on existing set-top box platforms and requiring no modifications to the cable plant. The success of these trials provided the confidence to support deeper, statistics-focused follow-on trials in 2004, featuring expanded programming lineups and subscriber populations.

This technical paper reveals and analyzes the statistics from two trials held in two different systems, both offering a Switched Broadcast tier of dozens of channels, switched on-demand to hundreds of real-world subscribers. A transparent user interface was provided in both trials; most subscribers had no knowledge that the channels they were watching were being dynamically switched.

The data provides an insight into spatial and temporal subscriber viewing patterns, and statistical power-law models are applied as a means of characterizing the observed trial traffic as well as providing a means of predicting incremental resource demands and providing support for capacity planning in new or existing Switched Broadcast systems.

Switched Broadcast Architecture and Operations

A brief overview of a Switched Broadcast System is described. Other references¹ can provide a deeper understanding of the detailed system.

Architectural Components

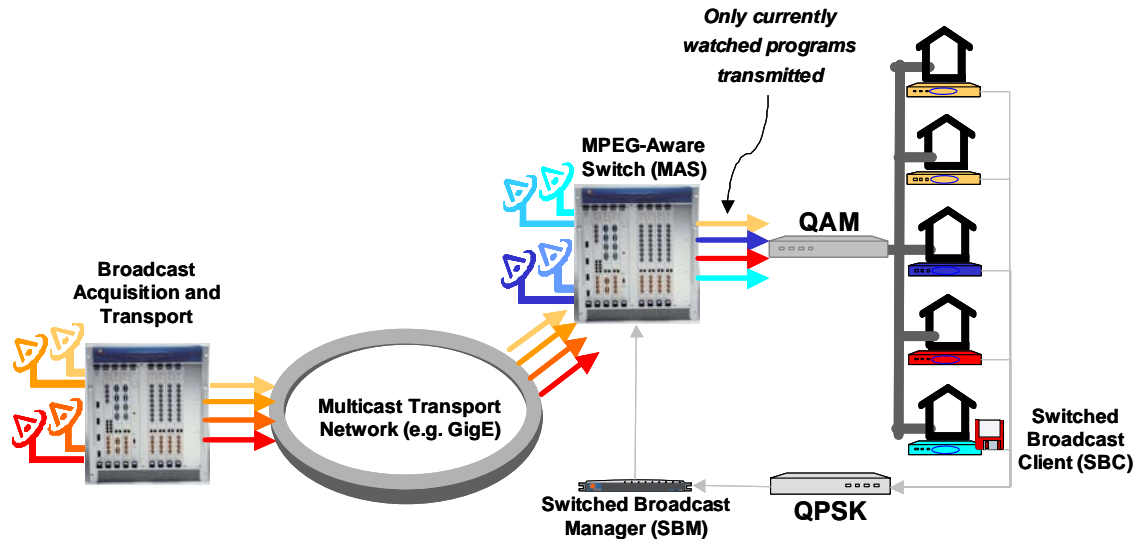


Figure 1 - Switched Broadcast Architecture

Figure 1 shows a generic Switched Broadcast architecture. Four subsystems provide the essential elements of the system.

A *Broadcast Acquisition and Transport* subsystem can be located in the headend. The Acquisition system performs a bit rate clamping function, where incoming broadcast programs are rate shaped to a predetermined constant bit rate (CBR) value. Though not absolutely required, the normalization of all programs to a CBR rate allows a faster and simpler switching mechanism, and paves the way for QAM resource sharing with similar services such as Video on Demand (VOD). The CBR rate chosen most often is one that resembles the program parameters for VOD streams. Using a Gigabit Ethernet transport structure, the clamped programs are multicast to all hubs.

A *Switched Broadcast Client (SBC)* is a small software application resident on the set-top box. This client functionality can just as easily be integrated into the tuning firmware of future set tops. When a Switched Broadcast program is selected, this software component conveys the channel request via the upstream cable plant, along with information that uniquely identifies the node-group location of the set-top box.

A *Switched Broadcast Manager (SBM)* application runs on a machine located in the hub or headend. The SBM uses the channel number received to identify the requested program, and consequently, the MPEG-Aware Switch (MAS) input port where the program is being received. Similarly, the SBM uses the set-top box ID and associates the node group information to determine the downstream connection (MAS output port) where the subscriber can be reached. In many cases, a subscriber can be reached by more than one downstream QAM channel. This collection of one or more QAM channels, and the dozen or so programs

that can be carried in each QAM channel, represent a pool of resources that the SBM has at its disposal to fulfill programming requests within each node group.

When an available downstream QAM and the program resource are identified, the frequency and program information is returned via the downstream out-of-band channel to the set-top box, which decodes and displays the program using the normal tuning mechanisms. While the specific frequency and program number for a Switched Broadcast program may vary in time, the channel number as seen by the subscriber will always remain the same.

An *MPEG-Aware Switch (MAS)* takes direction from an SBM and connects programming sources with node-group destinations. The required switching capabilities can be integrated into an Edge QAM or the MAS can exist as a standalone entity.

System Operation

The actions that take place when tuning to a Switched Broadcast program differ from traditional broadcast tuning. In a traditional broadcast environment, *all* programming is sent via QAM/RF/HFC to set-top boxes, along with data streams that convey the Program Specific Information (PSI) and other channel map information. Using these tables, a set-top box can determine the specific frequency and program number for a television program, and command the set-top box to receive, decode and display the selected program.

In a Switched Broadcast environment, a slightly modified tuning methodology is used. The frequency and program number of any particular program varies in time, and so consequently the content of the tables that carry the frequency and program information for the broadcast channels now vary with time. Therefore, the set-top box uses the out-of-band data channel to acquire the now time-variant tuning information. When this information is retrieved, the tuning process proceeds as normal. To the end user selecting a program via an electronic program guide or remote control keypad, this can be designed to be a visually seamless process.

Trial Information

Two trials were conducted in 2004 to analyze statistical efficiencies and operational usage of a Switched Broadcast System*. The first system, known hereafter as “Trial A”, was deployed in Tyler, TX, a Cox system. This is a 550 MHz plant using Motorola set-top boxes and a Pioneer Electronic Program Guide (EPG). The second system, known hereafter as “Trial B”, was deployed in a 750 MHz system using Scientific-Atlanta set-top boxes and EPG. Trial B is included to illustrate the effect of increased programming on the switched broadcast tier.

	Trial A	Trial B
System	Cox; Tyler, TX. system	Top 5 MSO, United States
Plant	2-way 550 MHz	2-way 750 MHz
STB	Motorola	Scientific-Atlanta
Guide	Pioneer	SA
Digital Subscribers (trial)	603	915
Nodes (trial)	4	4
Number of Switched Channels	60	171
Number of Stream Resources	40	100
Trial Start Date	Spring 2004	Summer 2004
Trial End Date	Still active as of 10/27/04	Still active as of 10/27/04

Table 1 – Trial A and Trial B Parameters

Trial A

Trial A examined the viability of Switched Broadcast as an alternative to a plant upgrade. The trial was deployed on a 550 MHz system. In this trial, the digital channels already available as a broadcast tier were simply delivered in a switched capacity. Given the limited spectral capacity of a 550 MHz system, the number of offered programs (60) is much smaller than in Trial B.

CH NUMBER	SERVICE DESCRIPTION	CH NUMBER	SERVICE DESCRIPTION	CH NUMBER	SERVICE DESCRIPTION
45	TV LAND (analog/digital)	200	ENCORE EAST	600	ESPN NOW
69	ESPN2 (analog/digital)	201	ENCORE WESTERN EAST	601	ESPN PPV 1
76	SCI-FI CHANNEL (analog/digital)	202	ENCORE LOVE EAST	602	ESPN PPV 2
110	DISCOVERY-KIDS	203	ENCORE MYSTERY EAST	603	ESPN PPV 3
111	DISCOVERY-TIMES	204	ENCORE TRUE STORIES EAST	604	ESPN PPV 4
112	DISCOVERY-WINGS	205	ENCORE ACTION EAST	605	ESPN PPV 5
113	DISCOVERY-HOME & LEISURE	210	HBO EAST	801	PPV 1
114	DISCOVERY-SCIENCE	211	HBO PLUS EAST	802	PPV 2
116	DISCOVERY-EN ESPANOL	212	HBO SIGNATURE EAST	803	PPV 3
117	BBC AMERICA	213	HBO FAMILY EAST	804	PPV 4
118	GAME SHOW	214	HBO COMEDY	807	PPV 7
119	SPEED CHANNEL	215	HBO ZONE	808	PPV 8
121	FOX SPORTS WORLD	216	HBO LATINO	809	PPV 9
122	ESPNEWS	220	CINEMAX EAST	810	PPV 10
123	OUTDOOR LIFE	221	MORE MAX EAST	811	PPV 11
124	ESPN CLASSICS	222	THRILLER MAX	812	PPV 12
125	OUTDOOR CHANNEL	223	ACTION MAX	813	PPV 13
126	BIOGRAPHY CHANNEL	224	W MAX	814	PPV 14
127	HISTORY INTERNATIONAL	225	5 STAR MAX	815	PPV 15
130	DO IT YOURSELF DIY	240	SHOWTIME 1 EAST	816	PPV 16
132	FINE LIVING	241	SHOWTIME TOO EAST	817	PPV 17
138	TOON DISNEY	242	SHOWTIME SHWCASE EAST	818	PPV 18
140	ENCORE WAM	243	SHOWTIME EXTREME EAST	819	PPV 19
141	LIFETIME MOVIES	244	SHOWTIME BEYOND	820	PPV 20
142	TECH TV	250	THE MOVIE CH. EAST	821	PPV 21
143	BLOOMBERG	251	THE MOVIE CH. XTRA EAST	822	PPV 22
145	CNNFN	260	STARZ EAST	850	PLAYBOY 6 HR BLOCKS
183	MUCH MUSIC	261	STARZ 2 EAST	851	SPICE
198	FOX MOVIE CHANNEL	262	STARZ BET	852	HOT CHOICE
		263	STARZ FAMILY		
		264	STARZ CINEMA		

Table 2 – Switched Broadcast Channel Lineup, Trial A

* As of the time of this writing, the location and sponsoring MSO of trial B was not publicly disclosed.

Trial A was deployed to one node (each node = 1000HHP), and later expanded to four nodes. At its peak, 603 digital STBs were on the Switched Broadcast tier.

Trial B

Trial B aimed to conduct a comprehensive statistical qualification and analysis of Switched Broadcast. As such, the entire digital lineup, minus high-definition and music channels, was switched. Later in the trial, viewing statistics were also collected for the analog tier (though statistics were collected for the analog tier, the analog tier itself was not switched).

CH	SERVICE DESCRIPTION	CH	SERVICE DESCRIPTION	CH	SERVICE DESCRIPTION
348	News 8 Traffic Now	207	DIY	906	ESPN GamePlan Full Court
359	New 8's Non-Stop Weather en Español	722	5 Star MAX	276	America's Store
490	Outdoor Channel	910	NBA League Pass/Direct Kick	979	NASCAR IN CAR
324	TXCN	912	NBA League Pass/Direct Kick	978	NASCAR IN CAR
470	Outdoor Life Network	913	NBA League Pass/Direct Kick	977	NASCAR IN CAR
1	Channel 1 RTE	914	NBA League Pass/Direct Kick	976	NASCAR IN CAR
283	Inspirational Life	915	NBA League Pass/Direct Kick	975	NASCAR IN CAR
613	Sorpresa	916	NBA League Pass/Direct Kick	950	Previews
102	Disney W	917	NBA League Pass/Direct Kick	434	Fuel
708	HBO W	918	NBA League Pass/Direct Kick	940	Playboy Pay-Per-View
709	HBO2 W	919	NBA League Pass/Direct Kick	941	TEN
710	HBO Signature W	290	EWTV	942	TEN*Blox
711	HBO Family W	930	NHL Center Ice/ MLB extra Innings	943	TEN*Clips
714	HBO Latino W	931	NHL Center Ice/ MLB extra Innings	140	Discovery Kids
723	CineMAX W	932	NHL Center Ice/ MLB extra Innings	237	Science Channel
724	MoreMAX W	933	NHL Center Ice/ MLB extra Innings	203	Discovery Home Channel
725	ActionMAX W	934	NHL Center Ice/ MLB extra Innings	250	Discovery Times
712	HBO Comedy W	935	NHL Center Ice/ MLB extra Innings	268	BBC America
713	HBO Zone W	936	NHL Center Ice/ MLB extra Innings	236	Discovery Wings
726	ThrillerMAX W	937	NHL Center Ice/ MLB extra Innings	640	Discovery en Español
382	C-SPAN3	938	NHL Center Ice/ MLB extra Innings	220	Health Network
405	ESPN News	939	NHL Center Ice/ MLB extra Innings	225	FitTV
525	GSN	921	NBA League Pass/Direct Kick	266	TRIO
368	Bloomberg	455	Tennis Channel	341	Newsworld Int'l
211	Lifetime Real Women	911	NBA League Pass/Direct Kick	738	Showtime W
103	Toon Disney	951	Movies and Events	739	Showtime Too W
420	ESPN Classic	952	Movies and Events	740	Showcase W
262	Ovation	953	Movies and Events	748	Movie Channel W
364	CNNfn	954	Movies and Events	749	Movie Channel Xtra W
750	STARZ!	955	Movies and Events	742	Showtime Beyond W
751	STARZ! Theater	956	Movies and Events	741	Showtime Extreme W
753	STARZ! Family	706	HBO Zone	743	Showtime Next W
752	Black STARZ!	701	HBO	745	Showtime Family Zone W
754	STARZ! Cinema	702	HBO2	744	Showtime Women W
760	Encore	703	HBO Signature	730	Showtime
761	Encore W	704	HBO Family	731	Showtime Too
768	WAM!	707	HBO Latino	732	Showcase
717	ActionMAX	716	MoreMAX	746	Movie Channel
762	Love Stories	718	ThrillerMAX	778	Sundance Channel
764	Mystery	705	HBO Comedy	747	Movie Channel Xtra
763	True Stories	715	CineMAX	734	Showtime Beyond
766	Westerns	280	TBN	733	Showtime Extreme
431	FOX Sports Digital Atlantic	381	C-SPAN2	735	Showtime Next
432	FOX Sports Digital Central	282	Inspiration Network	737	Showtime Family Zone
433	FOX Sports Digital Pacific	358	News 8's Non-Stop Weather	736	Showtime Women
770	FOX Movie Channel	611	Canal Sur	615	CNN Español
620	FOX Sports World en Español	632	Puma TV	571	MTV2
430	FOX Sports SW	575	Fuse	582	VH1 Classic
460	Speed Channel	361	CNBC World	150	Noggin
595	Great American Country	625	Cine Latino	122	GAS
719	WMAX	974	NASCAR IN CAR	121	Nick Too
720	@MAX	973	NASCAR IN CAR	609	mun2
721	OuterMAX	901	ESPN GamePlan Full Court	587	BET On Jazz
240	A&E	902	ESPN GamePlan Full Court	238	G4 Network
248	History International	903	ESPN GamePlan Full Court	435	NBA TV
774	Independent Film Channel	904	ESPN GamePlan Full Court	288	The Word Network
239	Tech TV	905	ESPN GamePlan Full Court		

Table 3 – Switched Broadcast Channel Lineup, Trial B

Trial B started with a single node containing 334 STBs, 149 channels of programming, and 160 stream resources. Since the number of stream resources initially exceeded the number of offered programs, there was no risk of denial of access for any service. As the trial progressed, the parameters were modified to exercise the system. Three nodes were added to the trial, yielding a final subscriber population of 915 digital STBs. More programming was subsequently added, growing the Switched Broadcast tier to 169 programs. Finally, the stream resource pool was reduced from 160 streams (16 256QAMs) to 100 streams (10 256QAMs).

Statistical Analysis

This section presents the observed statistics in each of the trials. Both trials produced an abundance of log files for analysis. A representative period of time was selected for analysis. In most cases, a one-week period was chosen to comprehensively capture subscriber viewing behavior (weekdays, weekends, etc.)

Trial A Statistics

The Figure below shows the maximum observed channel use and number of set-top boxes tuned to each channel across a one-week period in July 2004, in 5-minute increments.

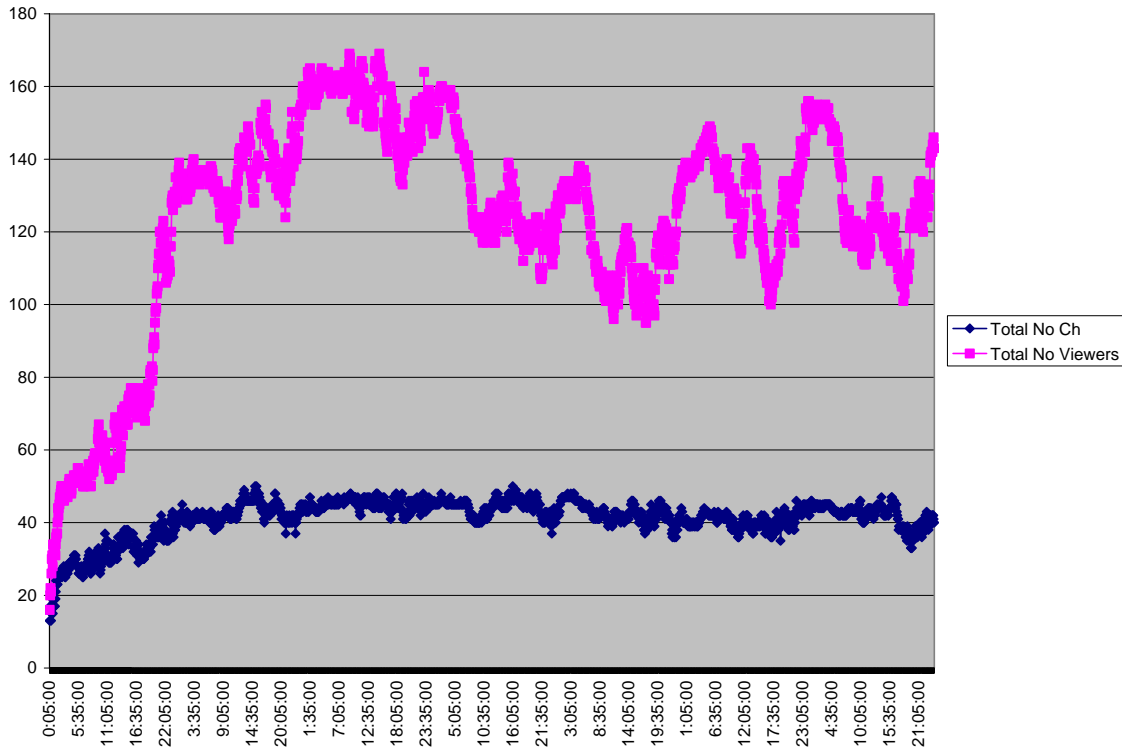


Figure 2 – Trial A, Subscribers and Watched Programs, All 4 Nodes, 7/9/04-7/15/04

Incremental STB Impact on Statistics

An important metric to study is the impact of incremental subscribers on the number of watched programs. Intuitively, we know that if more subscribers are added to the Switched Broadcast tier, the likelihood that they will have different viewing preferences and watch different programs increases. However, it is entirely possible that they will select the same programs as their peer subscribers, which would have no net effect on the number of watched programs.

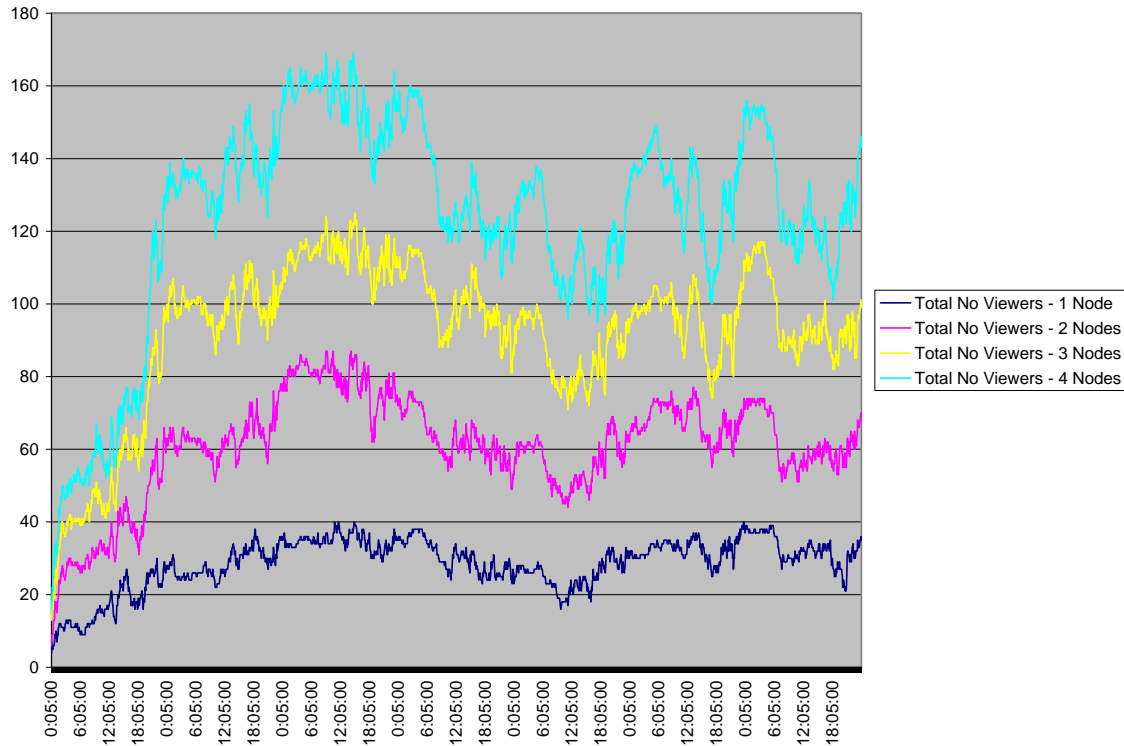


Figure 3 – Trial A, Number of Subscribers, 4 Nodes (cumulative), 7/9/04-7/15/04

Trial A was conducted in four nodes. The granularity of the log files allowed log data to be analyzed on a node-by-node basis. For the above graphs, data was plotted for node 1, node [1,2], node [1,2,3], and node [1,2,3,4]. This allows the examination of the effects of incremental subscribers on the overall statistics.

As is evident from the top plot, the number of set-top boxes simultaneously viewing programs tends to increase linearly as nodes (subscribers) are added to the switched broadcast tier. Note that the gap between the plots at any one point in time remains relatively constant.

The impact of additional subscribers to the Switched Digital tier also has an impact on the maximum number of channels used. The figure below demonstrates the increase in peak channel usage across the same period.

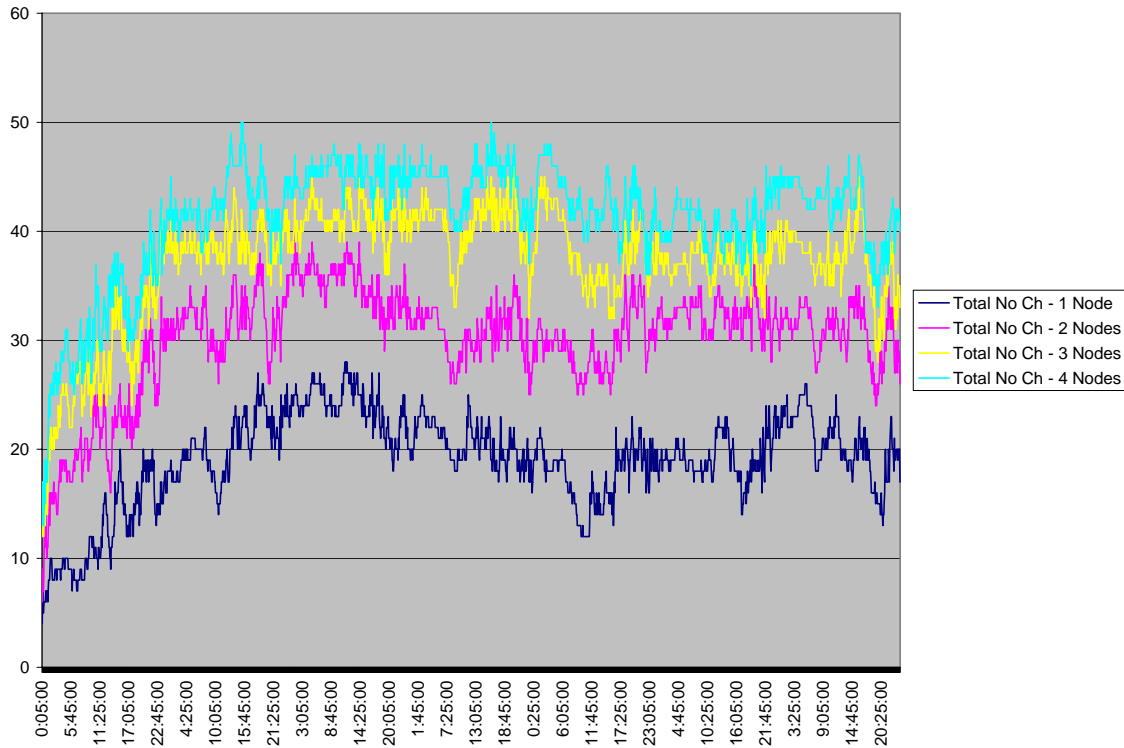


Figure 4 – Trial A, Number of Watched Programs, 4 Nodes (cumulative), 7/9/04-7/15/04

Unlike the previous graph, as nodes are introduced to the switched tier the number of channels used tends to increase, but in a logarithmic fashion. Note the gap between the graphs at any one point in time becomes smaller with the addition of each node. The data collected indicates that the channel usage peaked at 50 out of a possible 60 channels. This does not represent a large bandwidth savings, but there are two significant reasons why this occurred:

- When a set-top box in Switched Broadcast mode changes channels, it sends a “tune in” message to acquire the new channel information, and a “tune out” message to indicate that it is no longer watching the previous channel. In this trial, “tune outs” from set-top boxes that were turned off were not captured in the data. Hence, the total number of channels used (as indicated by the log file) is artificially higher than normally would be expected.
- In this trial, the top 60 channels were put on the switched tier. As will be shown in the modeling section below, substantial savings in simultaneous programs viewed is obtained when less popular channels are included in the mix. This is due to the fact that the likelihood of one or more of the less popular channels to not be viewed permits its use by another program.

Trial B Statistics

In Trial B, a large number of programs, 171, were put on the switched digital tier. The graph below presents the number of channels used across a single day. (One of the busiest days of the week, a Sunday, was chosen).

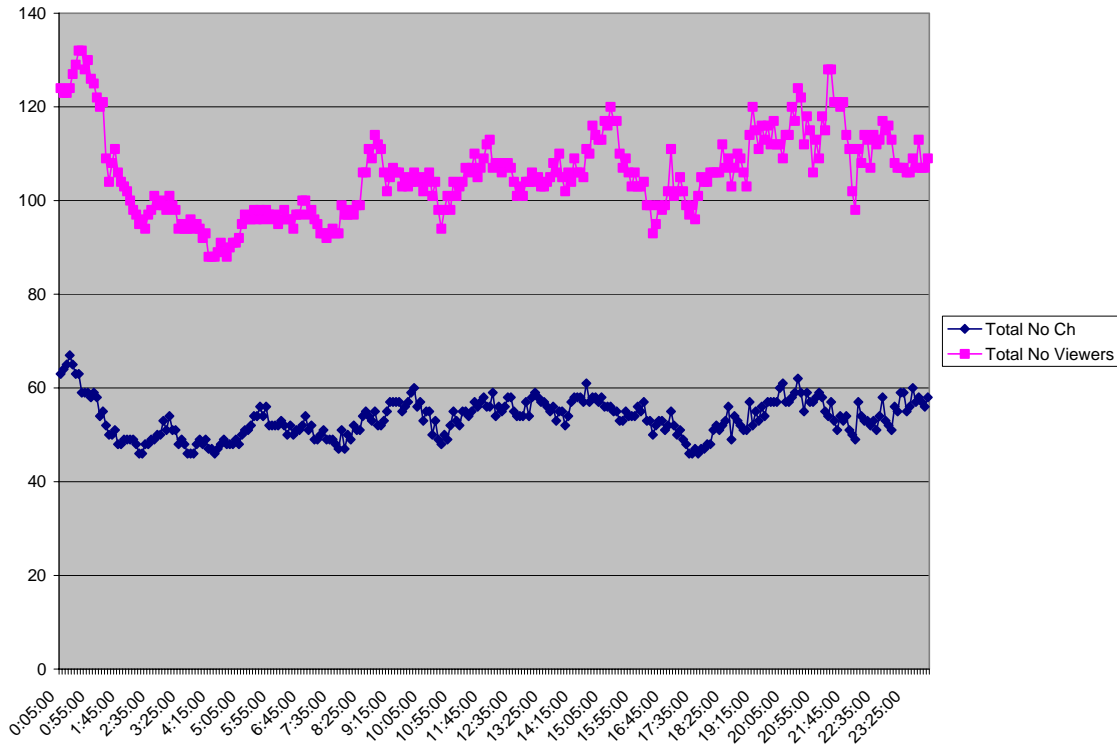


Figure 5 – Trial B, Subscribers and Watched Programs, All 4 Nodes, 9/26/04

On this day in Trial B, the maximum number of channels ever used never exceeds 67, a 60% savings in bandwidth use. In this case, both the number of programs and the number of subscribers that have access to the switched broadcast service are greater than Trial A. As expected, the resultant channel use is much smaller as the full suite of programming was included in the switched tier. Unlike Trial A, “tune outs” as a result of powered off set-top boxes were captured in the log file. In this trial, a significantly smaller percentage of the total number of channels, 40% versus 83% for Trial A, accounts for peak channel usage.

Incremental Programming Impact on Statistics

Another important consideration in Switched Broadcast is the effect of incremental programming on the overall system utilization. To study this effect, the data from Trial B was used. All 171 programs were ranked in popularity, with the number of users and the duration of viewing used as the factors in determining popularity. Then, in groups of 10, the most popular programs were removed from the analysis, and the utilization numbers were recalculated. This process was repeated again in groups of 10 until there were no more active programs. The plot of this exercise is shown below.

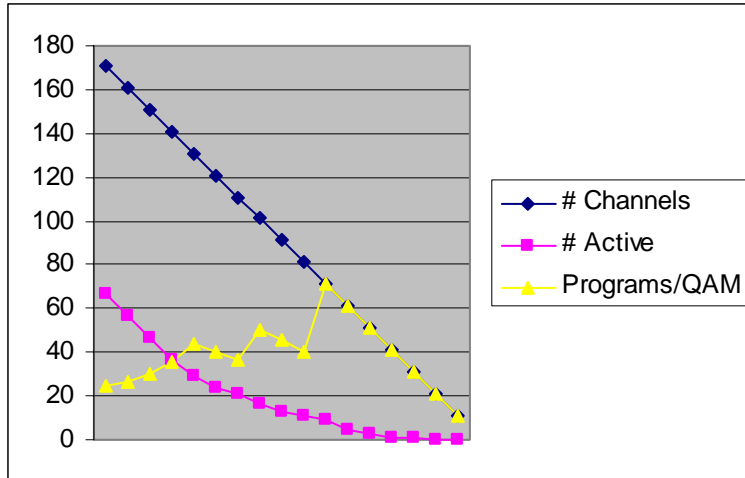


Figure 6 – Effect of Program Reduction on Switched Broadcast Utilization, Trial B, 9/26/04

The top plot shows the removal of programs from the analysis. The plot of the maximum number of active programs shows a corresponding decrease. At first, this decrease exhibits a linear pattern with an identical slope. This confirms that indeed the most popular programs are being removed from the Switched Broadcast tier. As more programs are removed, the decrease becomes more gradual.

The final plot of Programs/QAM is a rough calculation of the Switched Broadcast efficiency. Averaged over the entire tier, having more programs per QAM represents a higher efficiency. At the start point of the exercise, there are 171 programs available, and a maximum utilization of 67. The number 67 is rounded up to 70 (to represent seven QAMs of available capacity), and we state that for this day, 171 programs could have been carried in seven QAMs with no blocking. The number of programs/QAM is therefore $171/7 = 24.4$ programs/QAM. Note that this number already vastly exceeds the efficiencies that can be achieved with conventional broadcast techniques like RateShaping or closed-loop encoding. As programs are removed, the numerator and denominator values change, and the efficiency improves, up to a phenomenally high 71 programs/QAM (the least popular 71 programs occupied just nine active streams, or one QAM). Since just one QAM is now remaining, only the numerator is affected for the remainder of the plot, so the efficiency calculation tracks with the number of programs for the remainder of the plot.

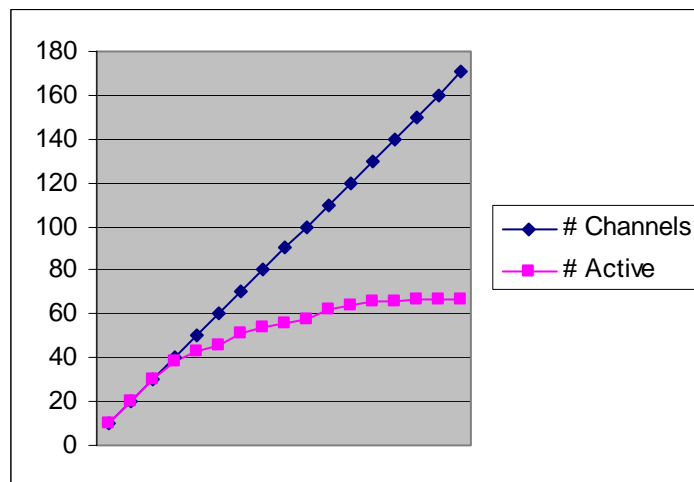


Figure 7 – Effect of Program Addition on Switched Broadcast Utilization, Trial B, 9/26/04

Figure 7 shows another representation of the same data. This time, we start from zero and add the most popular programs to the Switched Broadcast tier, 10 at a time. As expected, the number of active programs initially tracks the number of offered programs. Then, as less popular programs are added to the tier, the number of active programs tapers off. This is a strong visual demonstration of the power of Switched Broadcast.

Statistical Modeling of Switched Broadcast

Many man-made and naturally occurring phenomena, including city sizes, incomes, word frequencies, and earthquake magnitudes, are distributed according to a power-law distribution. A power-law implies that small or lower ranked occurrences are extremely common, whereas large, higher ranked instances are extremely rare. In essence, it is a mathematical description of the “80/20” rule. This regularity or “law” is sometimes also referred to as a Zipf² distribution.

It can be expressed in mathematical fashion as a power law, meaning that the probability of attaining a certain size x is proportional to x^{-r} , where r is generally less than or equal to 1. Unlike the more familiar Gaussian distribution, a power law distribution has no “typical” scale and is hence frequently called “scale-free”. A power law also gives a finite probability to very large elements, whereas the exponential tail in a Gaussian distribution makes elements much larger than the mean extremely unlikely. For example, city sizes, which are governed by a power law distribution, include a few mega cities that are orders of magnitude larger than the mean city size. On the other hand, a Gaussian distribution, which describes, for example, the distribution of heights in humans, does not allow for a person who is several times taller than the average. Interestingly, the data obtained in both Trial A and B exhibit a power law distribution that is essentially Zipf in nature in both peak simultaneous use and overall channel popularity

As mentioned above, mathematically a power law can be stated as follows:

$$p_k = Cr^{-\alpha} \quad (1)$$

p_k represents the popularity of the k -th ranked item in an ordered list of items ranked by popularity. Note that by multiplying the above equation by the total number of observations the expected number of each ranked item can be calculated. Hence,

$$N_k = Ar^{-\alpha} \quad (2)$$

N_k represents the expected number of the k -th ranked element. Note that the constant term, A will in most cases be different from the constant term C identified in equation 1.

An interesting consequence of either equation is that by taking the log of both sides:

$$\log(p_k) = \log(C) - \alpha \log(r) \quad (3)$$

$$\log(N_k) = \log(A) - \alpha \log(r) \quad (4)$$

Note that because of the equations are linear in $\log(\cdot)$, both the constant and the exponent term, α , can be obtained by linearly regressing $\log(p_k)$ on $\log(r)$ or $\log(N_k)$ on $\log(r)$.

When using the power law equation to express probabilities, the constant term can theoretically be calculated as follows:

$$C = \frac{1}{\sum_1^K \frac{1}{r^\alpha}} \quad (5)$$

K represents the total number of items in the rank order list. In a true Zipf distribution $\alpha=1$.

Trial A

In order to determine if peak simultaneous channel use follows a Zipf distribution, the log of channel popularity, defined as the number of unique STBs tuned to a given program, versus the log of the rank order of each of the programs, where the highest ranked program (rank=1) is defined by the channel to which the greatest number of STBs are tuned, is plotted. Equation 4 indicates that such a plot should be relatively linear. A representative example of one such plot is shown in the figure below. This plot shows the relationship between the two variables of interest for two nodes across the 5-minute period representing the largest simultaneous channel use (i.e., widest spread.) Note that the relationship is relatively linear. A summary of the regression output is provided in the Table below. The R^2 value, a measure of the degree to which the regression approximation captures the variability in the data, is 0.91 indicating a strong linear relationship.

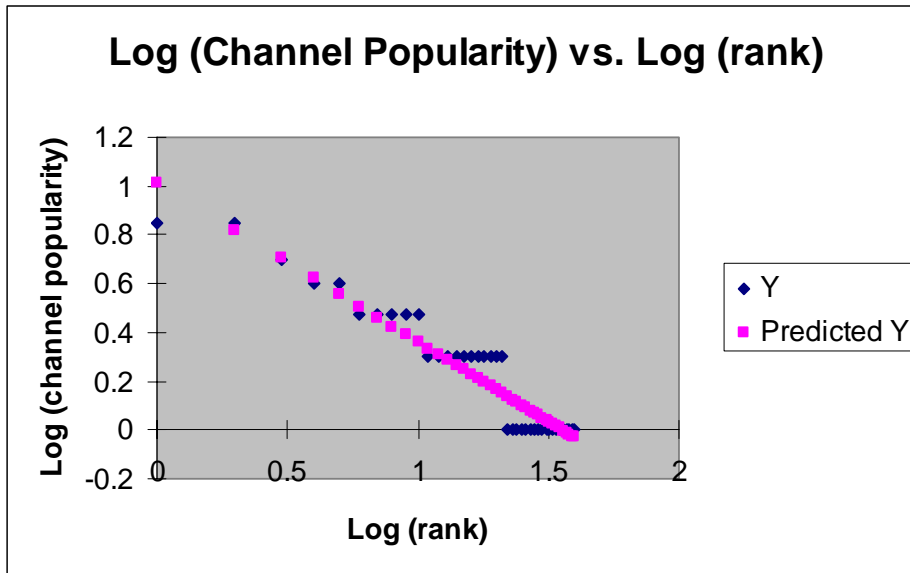


Figure 8 - Log (number of STBs tuned to r-th ranked program) vs. Log (rank) across 2 nodes, greatest channel usage across a 5 minute period

Also of interest are the Intercept and the slope of the regression line, denoted as x variable 1, which define the log of the constant term and exponent, α , respectively. Hence, the estimated equation based upon this data is:

$$N_k = 10.32 * r^{-0.65} \quad (6)$$

Regression Statistics	
Multiple R	0.95487601
R Square	0.911788195
Adjusted R Square	0.909466832
Standard Error	0.077975392
Observations	40
Coefficients	
Intercept	1.013682025
X Variable 1	-0.652389595

Table 4 - Summary of Regression analysis for Log (Channel Popularity) vs. Log(rank)

The same methodology was used to generate Zipf approximations for the 1, 2, 3, and 4 node cases for Trial A. A summary of the results is provided in the following table.

Number of nodes	Simultaneous Users	R-square value	Peak simultaneous Channel use	Zipf Approximation Equation
1	40	0.85	28	$4 * r^{-0.5}$
2	83	0.91	40	$10.32 * r^{-0.65}$
3	98	0.90	47	$12.1 * r^{-0.68}$
4	136	0.86	50	$18.9 * r^{-0.74}$

Table 5 - Summary of the Zipf approximation for 1 – 4 nodes in Trial A, 60 total channels in the switched tier

Note that the R^2 value never falls below 0.85 indicating a relatively strong log linear relationship. Also of note are the Zipf approximation equations whose exponent terms increase in magnitude from 0.5 to 0.74. This results in a greater spread of channel use as dictated by the observed data. The key point to note here is that an increase in the subscriber base while holding the number of offered programs constants tends to increase the spread of channel usage. This makes intuitive sense, as more subscribers will, in general, create a wider distribution of viewed channels. Both actual and Zipf approximation plots are shown in the figure below. Note that the approximations do a fairly good job of tracking the overall shape of the observed data. In the samples analyzed there were several channels to which zero set-top boxes were tuned, which could not be incorporated in the log – log plots. Thus, the regression equation tends to overestimate the number of set-top boxes tuned to those channels.

Observed Channel Popularity and Estimated Zipf Approximation for 1 - 4 nodes, Trial A

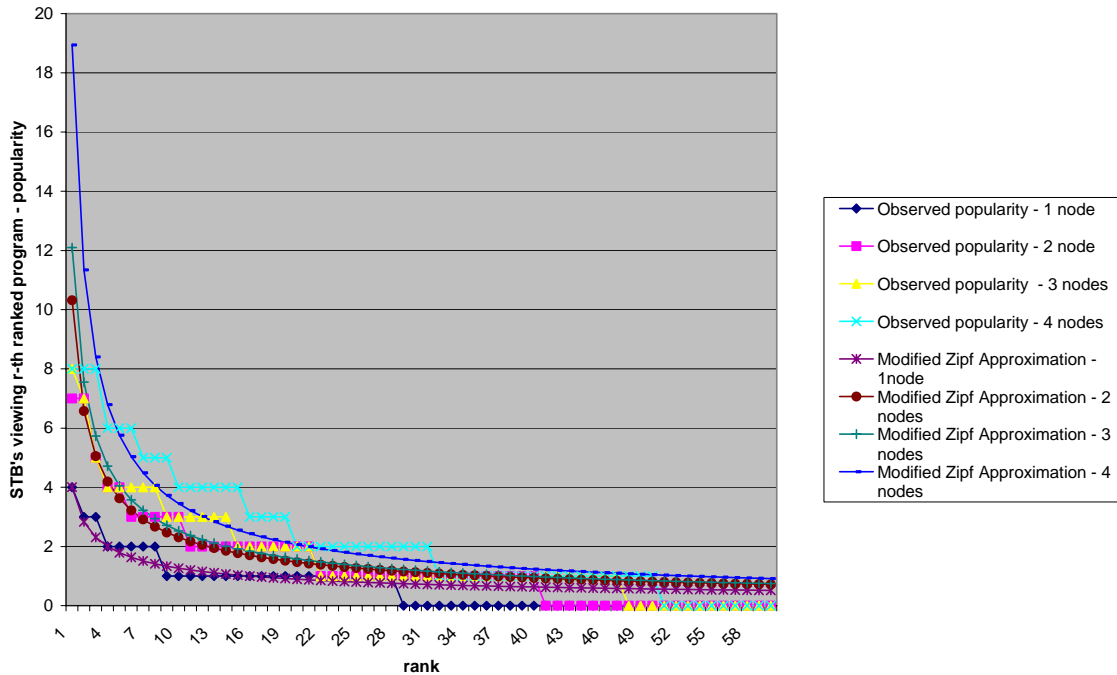


Figure 9 - Observed Channel Popularity versus rank and the associated Zipf approximation

Trial B

Trial B includes the addition of a much greater number of channels, 171, in the Switched Broadcast tier. The same methodology was applied to this data to obtain the Zipf approximation. Of interest is the peak channel usage as the number of offered programs is increased. We were able to find a 5-minute period in Trial A's empirical data which contained the highest channel usage and approximately the same number of simultaneous users as a similar observation in Trial B's data sample. The resulting actual data and associated Zipf approximations are presented in the figure below. Note that as the number of programs is increased the percentage of the number of simultaneous channels used to the total number of channels offered decreases dramatically. In our sample, Trial A shows that $50/60 = 83.3\%$ of the total number of channels is required whereas in Trial B, only $72 / 171 = 40\%$ is required. Note this gain is obtained by adding lower popularity channels to the switched digital tier.

This phenomenon is also observed from the Zipf approximation. The table below compares the approximation for the relevant data sets of both Trials.

Trial	Nodes	Channels occupied / Total channels offered	Total number of STB's	R ²	Zipf approximation
A	4	50 / 60	136	0.86	$18.9 * r^{-0.74}$
B	4	72 / 180	147	0.92	$13.5 * r^{-0.64}$

Table 6 - Summary of Trial A and Trial B data set and associated Zipf approximation

Note that the exponent, α , decreases significantly in magnitude in Trial B thereby tightening the spread of channel use. The approximation tracks the non-zero number of STBs viewing the r-th ranked program fairly well. However, it tends to overestimate the popularity of the lower ranked programs. Nonetheless, in both cases by using the approximation a fairly accurate estimate of the total channels occupied can be obtained.

Impact of additional programming on Observed popularity and Zipf Approximation

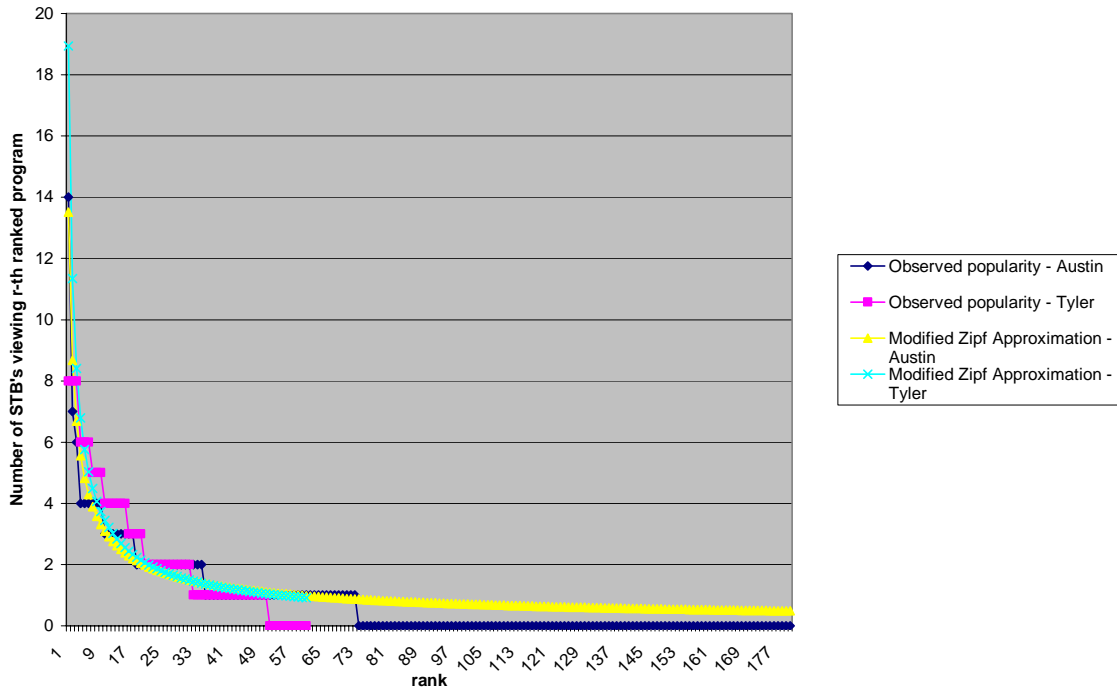


Figure 10 - Observed and Zipf estimated simultaneous channel popularity for Trial 1 (60 offered channels) and Trial 2 (180 offered channels). The number of simultaneous users was relatively constant

Analytically Estimating Channel Usage

The previous analysis affords us the opportunity to use a Zipf approximation to calculate the expected channel use if 500, 1000, or even 1500 programs were to be offered on the Switched Broadcast tier. Based on the analysis, the following has been empirically shown:

1. The addition of subscribers (nodes) tends to linearly increase the number of simultaneous STBs tuned to the view programs.
2. The addition of subscribers (nodes) tends to increase the number of channels used for switched broadcast. However, the rate of this increase diminishes with the addition of subscribers (nodes).
3. Adding programming to the switched service while holding the viewing population relatively constant narrows the channel popularity distribution. This results in a smaller proportion of the total channels offered accounting for a larger percentage of the total views.

Given the above, a Zipf approximation can be used to estimate peak channel use, making the following assumptions:

1. The digital subscriber population and node size is approximately consistent with that observed in Trial B.
2. We chose a very conservative 0.74 value for α as the starting point for 500 channels. This value is considerably higher than that observed in the Trial B data sample with only 180 services.
3. The constant term in equation (1) represents the popularity of the highest ranked channel (i.e., with $r=1$.) We estimated this by using ~ 0.5 of the value obtained from equation (5), with $\alpha = 1$. Note that this approximate relationship was observed in the data sets we analyzed.

4. We decreased α by ~ 0.2 with each increase of 500 channels. Note that this again is a very conservative decrease when compared with the observed data. In our sample, when the channel pool in the switched tier was increased from 60 to 180, α decreased from 0.74 to 0.64.
5. We used the Zipf distribution to estimate the number of channels, K , required to accommodate at least 99.5% of total channel use.

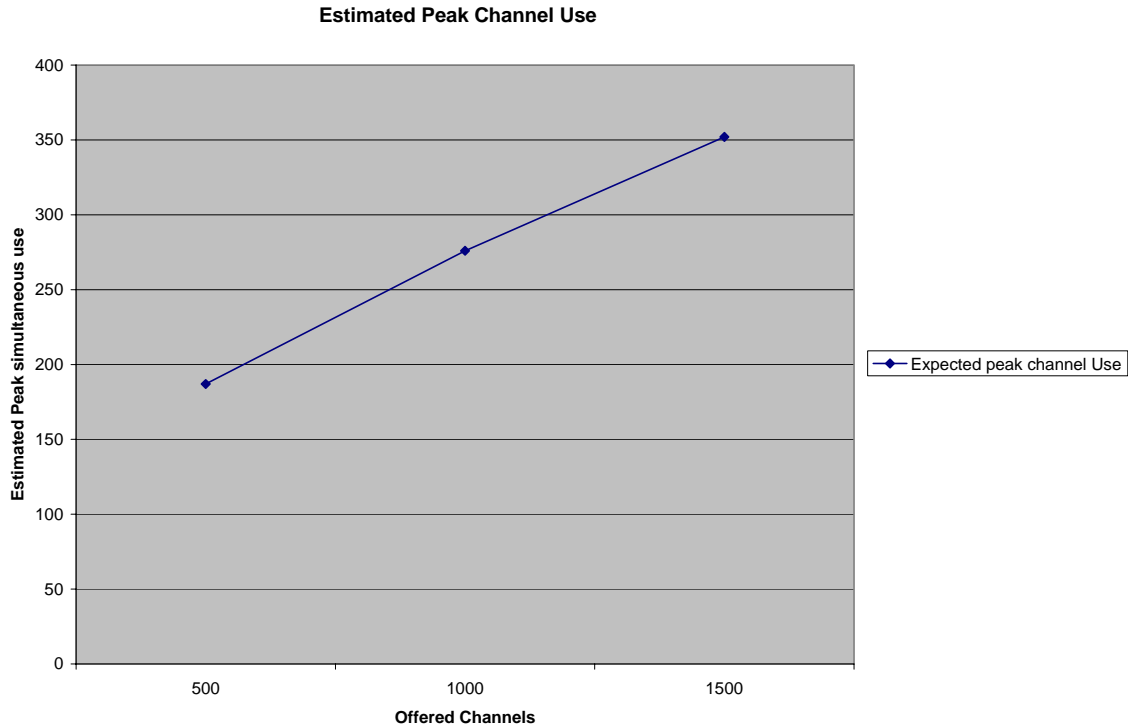


Figure 11 – Expected utilization of 500-channel, 1000-channel and 1500-channel Switched Broadcast tiers based on Zipf modeling of trial data

Note that despite fairly conservative parameter estimates, considerable channel savings can be obtained by putting a large number of services on the switched broadcast tier. A 500-channel system is calculated to require 187 active streams, or 19 256QAMs. A 1000-channel system is calculated to require 276 active streams, or 28 256QAMs. And a 1500-channel system is calculated to require 352 streams, or 36 256QAMs. This yields a remarkable ratio of $1500/36 = 41$ programs/256QAM, easily three times the efficiency achievable by even the best closed-loop encoders on the market today.

Summary

The statistics and subsequent analysis of data from real-world Switched Broadcast trials have provided a before-unseen insight into the viewing patterns of digital subscribers. Furthermore, a mathematical framework can be fitted to this viewer behavior. This allows the development of tools that can reliably assist in the dimensioning of Switched Broadcast system designs, as well as assist the development of accurate forecasting of capital needs as Switched Broadcast systems are expanded.

Although first coined by cable industry executives, the proverbial “500-channel system” was first realized by the satellite industry. Switched Broadcast allows the leveraging of cable’s two-way interactive network to offer and deliver catalogs of programming that are not practically achievable using a pure broadcast paradigm.

¹ S.V. Vasudevan and Paul Brooks, “Planning for and Managing the Rollout of Switched Broadcast Services”, SCTE Conference on Emerging Technologies, Miami, January 2003.

² G.K. Zipf, *Selective Studies and the Principle of Relative Frequency in Language* 1932.